Universidad de Oviedo

TESIS DOCTORAL

UN MODELO DE SISTEMA DE INFORMACIÓN PARA LA EFICIENCIA EN LA GESTIÓN DE LAS ORGANIZACIONES

Programa de Doctorado en Ingeniería de Producción, Minero-Ambiental y de Proyectos

Autora: D. Bárbara Egea Oliver

Directora: Dra. Nieves Roqueñí Gutiérrez

Codirector: Dr. Enric Brull Alabart

Oviedo, Junio 2016

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma: **UN MODELO DE SISTEMA DE INFORMACIÓN PARA LA EFICIENCIA EN LA GESTIÓN DE LAS ORGANIZACIONES**	Inglés: **AN INFORMATION SYSTEM MODEL FOR THE EFFICIENCY IN THE MANAGEMENT OF THE ORGANIZATIONS**

2.- Autor	
Nombre: BÁRBARA EGEA OLIVER	DNI/Pasaporte/NIE:
Programa de Doctorado: INGENIERÍA DE PRODUCCIÓN, MINERO AMBIENTAL Y DE PROYECTOS	
Órgano responsable: CENTRO INTERNACIONAL DE POSTGRADO	

RESUMEN (en español)

El trabajo realizado trata de responder a la pregunta básica de cómo deberían ser los sistemas de información para ser útiles a las organizaciones y permitirles aumentar la eficiencia en su gestión. Una organización eficiente es aquella que comprende cuales son las necesidades de sus grupos de interés y es capaz de satisfacer sus requerimientos presentes y futuros, para obtener resultados positivos de forma sostenible en el tiempo. Disponer de información relevante y contextualizada de todos los ámbitos de su actividad así como del entorno, y de sus procesos o servicios es una variable de gestión clave para la organización a la hora de conseguir este objetivo.

En un entorno de crisis global e incertidumbre, las organizaciones buscan la eficiencia para ser más competitivas y generar más valor en sus actividades. Implantar un sistema de información facilita que la organización aprenda, que haya una retroalimentación permanente y que se mejoren los sistemas de toma de decisiones fundamentados en datos. Se trata de una herramienta esencial para los procesos de planificación, ejecución y revisión orientados a la mejora continua de la organización, como camino hacia la excelencia.

Los sistemas de información obtienen grandes cantidades de datos de cada vez más fuentes de información. Este trabajo de investigación presenta un modelo para recoger estos datos y colocarlos de una forma estructurada, preparándolos para aplicar técnicas de análisis estadístico apropiadas, presentando los resultados mediante componentes de visualización de información, y elaborando con todo ello un sistema de informes que se publiquen en un portal de informes único y accesible. Estas partes del modelo se presentan para ser implantadas de forma desacoplada permitiendo así su actualización y mejora asíncrona sin alterar el resultado final.

La validación del modelo, que se presenta como modelo SIEGO, se ha realizado para cada

FOR-MAT-VOA-010-BIS

una de sus dimensiones –medir, analizar, visualizar, informar y publicar- comprobando hasta qué punto la organización que lo haya implantado, en todo o en parte, percibe una mejora en sus resultados. Para ello se ha llevado a cabo un estudio empírico basado en los datos de una encuesta dirigida a empresas y organizaciones españolas de distintos sectores, tamaños y ámbitos geográficos.

Los resultados de la encuesta muestran un grado de implantación desigual en el desarrollo de los sistemas de información en las organizaciones evaluadas. Destaca la aplicación casi generalizada de la dimensión de medición y recopilación de datos y de la dimensión de elaboración de informes, aunque no ocurre lo mismo con las dimensiones de visualización y de análisis estadístico, cuya utilización se ve muy influenciada por la existencia de herramientas informáticas más sofisticadas y de recursos humanos con mayor nivel de cualificación técnica. Por último, la dimensión de publicación de informes, con percibirse como importante, no se desarrolla.

Se ha buscado medir el impacto que la implantación de un sistema como el modelo SIEGO tiene en el proceso de toma de decisiones de la organización, y hay unanimidad por parte de las organizaciones encuestadas en que se produce una reducción de los tiempos de respuesta ante cualquier solicitud de información. Este resultado lleva a la organización a un nivel mayor de eficiencia.

La orientación del sistema de información hacia la planificación estratégica mejora la percepción de resultados de las organizaciones, y esto se hace más plausible en organizaciones que han adoptado modelos de gestión de excelencia como el modelo EFQM.

El modelo presentado, posibilita el análisis del pasado, el conocimiento del presente y aporta su máximo valor en la predicción del futuro de la organización.

RESUMEN (en Inglés)

The work done tries to answer the fundamental question on how should the information systems be useful to the organizations, hence allow them to increase the efficiency in their management. An efficient organization is one which understands what are the needs of their stakeholders, and is able to satisfy its present and future requirement, in order to get positive results in a sustainable way in time. To get relevant and contextualized information in all the fields of activity as in their environment, and of their process or services is a concordant management key to the organization itself to obtain this goal.

In a global crisis environment and uncertainty, the organizations search for efficiency in order to be more competitive and generate more value in their activities. To introduce an

informational system contributes to the learning of the organization, to a continuous feedback and improve the decision system based on information.

It is an essential tool for the planning, execution and revision process focused on the continuous improvement of the organization, as a way to excellence.

The informational systems get great quantities of data from more and more sources of information. This research introduces a model to collect this information and set them in a structured way, preparing them to implement appropriate statistical analysis techniques, showing the results through visual information components, and elaborate on the whole a reporting system that can be published in an accessible and unique website. These parts of the model are presented to be implemented in a disengaged way, hence allowing their update and improvement asynchronous without modifying the final result.

The validation of the model, presented as the SIEGO model, has been carried out for each one of their aspects -measure, analyze, visualize, inform and publish- verifying to what extent, the organization which has implemented it, wholly or partly, senses an improvement in its results. An empirical study has been accomplished based on an informational survey addressed to Spanish companies and organizations of various sectors, sizes and geographical fields.

The results of the survey show an uneven scale of implementation in the development of the information systems of the assessed organizations. An almost general application is highlighted in the data dimension and data collection and the reporting dimension, although not the same thing happens with the visual and statistical analysis aspect, whose use is very influenced by the existence of more sophisticated information tools and more qualified technical human resources. Finally, the information publication, sensed as important, is not developed.

A measurement of the impact of the implementation of a system such as SIEGO has been reviewed, in the decision making process of the organization, and there is unanimity on behalf of the surveyed organizations that a reduction in the reply is produced in any information request.

To focus the information system on strategic planning increases the perception of organizations results, and this is done more plausible in organizations that have adopted Excellency management models, such as EFQM.

The given model, enables the analysis of the past, the present knowledge and contributes to its highest value in the prediction of the future of the organization.

A l'Agustí, el meu company de viatge, i als dos estels que guien el nostre camí en el procés d'aprenentatge de la vida, la Sílvia i la Núria, us estimo molt.

Vull agrair, quiero agradecer:

Primerament al Dr. Enric Brull la inspiració, l'acompanyament i la codirecció d'aquesta tesi. Si tu no m'ho haguessis dit, no hagués pensat que podia fer-ho.

Muy especialmente a la Dra. Nieves Roqueñí Directora y Tutora de esta tesis, por el apoyo prestado, por el tiempo que me has dedicado y por encontrar siempre la solución y guiarme en el camino. Me lo has puesto siempre fácil exigiéndome el máximo. Sin tu ayuda no lo habría conseguido.

A Josep Sánchez Pagès, Cap del Gabinet de Presidència de la Diputació de Tarragona, per creure i confiar en mi, en el meu treball, però sobretot per tant que he après al teu costat.

A la Diputació de Tarragona, al seu equip directiu, als representants polítics i als companys: som una organització excel·lent sobretot per les persones que la formem. En especial als companys de l'Àrea de Coneixement i Qualitat que em pateixen de prop cada dia.

A l'Eloy Hernández, al Ferran Rodríguez, al Carles Llorach, al Toni Sunyer, a l'Agustí Porta, a l'Emma Gorgori i a la Lourdes López, pel vostre suport i paciència en aquest darrers temps. Sou els millors companys que he tingut mai. Gràcies per les estones de discussió, pel sentit de l'humor i per la feina tant gran que feu.

A la Unidad de Consultoría Estadística de la Universidad de Oviedo, y en especial a Tania Iglesias, por su colaboración en la parte empírica de este trabajo de investigación, a los servicios académicos, así como a toda la Universidad de Oviedo.

A l'Abel González de TSD Consulting, pels anys de treball plegats, per permetre'm descobrir i evolucionar en el mon del Business Intelligence.

A Miquel Romero Director de Socios y Conocimiento del Club d'Excelencia de Gestión, pel seu suport en la difusió del qüestionari de la tesi.

Als meus pares, per seguir estimant-vos i estimant-nos, incondicionalment. Passi el que passi a la meva vida, sé que sempre estareu al meu costat. Us estimo molt. Soc el que soc gràcies a vosaltres.

Al meu germà Saxa i la meva tieta Lourdes, tot i que de vegades no us veig, sé que esteu aquí al meu costat.

A la Rosa Soriano pels ànims, el suport i per ser incondicional, i a la Mireia Sánchez per les estones de ball i running per aquestes muntanyes, que tant m'han ajudat a aclarir-me la ment.

Escribir una tesis es recorrer un largo y difícil camino, que se hace a solas pero con la compañía de muchos.

A todos los que os habéis cruzado en mi camino…

A tots vosaltres moltes gràcies!

Índice de contenidos

Capítulo I. Presentación y motivación 9
1. Introducción y motivación 11
2. Estructura de la tesis 17

Capítulo II. Marco y fundamento teórico 19
1 La gestión de las organizaciones 21
1.1 Círculo de Deming o PDCA (Plan, Do, Check, Act) 21
1.2 Modelos de Excelencia 22
1.2.1 Premio Deming 27
1.2.2 Malcom Baldridge National Quality Award 28
1.2.3 El modelo EFQM de Excelencia 30
1.2.4 Danish Business Excellence Index (DBEI) 31
1.2.5 Singapore Quality Award Framework (SQAC) 32
2 De la Excelencia a los Sistemas de Información 34
3 Los Sistemas de Información 35
3.1 Concepto de Sistema de Información para la Gestión 37
4 El Business Intelligence como sistema de información 38
4.1 Data Warehouse 42
4.2 Herramientas de análisis e Informes 42
4.3 La creación del Data Warehouse y las ETL 43
4.4 El diseño del Data Warehouse, punto clave 44
4.4.1 Tablas de dimensión 45
4.4.2 Tablas de hechos 46
4.4.3 Data Marts 46
4.4.4 Esquemas 47
4.5 Las ETL (Extract, Transform and Load) 48
4.5.1 Extracción (Extraction) 49
4.5.2 Transformación (Transformation) 50
4.5.3 Carga (Loading) 50
5 Las soluciones actuales para el Business Intelligence 51

Capítulo III. Del modelo a la realidad 57
1. SIEGO, el modelo propuesto 59
2. Dimensión Datos 62
1 Marco general de la dimensión datos 64
1.1 La gestión del conocimiento 65
1.2 Proceso de gestión del conocimiento 69
1.3 Fuentes del conocimiento 71
2 La dimensión Datos en el modelo SIEGO 73
2.1 Medición 76
2.2 Tipos de escalas 78

2.2.1 Nominal 79
2.2.2 Ordinal 79
2.2.3 Intervalo o Discretas 79
2.2.4 De razón o Continuas 80
2.2.5 Alternativas 80
2.3 Requerimientos de la medición 81
2.3.1 Validez 81
2.3.2 Detalle suficiente 81
2.3.3 Constructos 81
2.3.4 Calidad de datos 82
2.4 La matriz de datos 83
2.4.1 Matrices de Datos de Elementos 84
2.4.2 Matrices de Datos de Hechos 86
2.4.3 Tipos de información 86
2.4.4 Medidas de indicadores 87
3 Propuesta de estructura del sistema de datos 90
3. Dimensión Análisis 94
1 Técnicas analíticas 96
1.1 Análisis univariable 103
1.1.1 Serie temporal 104
1.1.2 Medidas de centralidad 105
1.1.3 Medidas de dispersión 106
1.1.4 Medidas de distribución 107
1.2 Análisis bivariable 108
1.2.1 Correlación 108
1.2.2 Regresión 109
1.2.3 Tablas de contingencia 109
1.3 Análisis multivariable 110
1.3.1 Análisis factorial 110
1.3.2 Análisis de correspondencias 111
1.3.3 Análisis de clúster 112
1.3.4 Análisis discriminante 112
1.3.5 Análisis de regresión múltiple 113
1.3.6 Análisis estructural 113
1.4 Data Mining 114
2 Selección del tipo de análisis 118
Dimensión Visualización 121
1 Antecedentes históricos de la representación de datos 123
1.1 Medida y teoría a lo largo de la historia 125
1.2 Evolución de los gráficos relacionados con datos estadísticos 127
1.3 La influencia de las nuevas tecnologías y la informática en la representación gráfica 137
2 La visualización de datos 139
2.1 La visualización como parte intrínseca de la estadística 139

2.2 La percepción humana de los gráficos 140
2.3 Leyes de la visión aplicadas a la visualización de gráficos 143
2.3.1 Ley de la proximidad 144
2.3.2 Ley de la semejanza o equivalencia 144
2.3.3 Ley de Prägnanz o de la buena forma 145
2.3.4 Ley de cierre 145
2.3.5 Ley de la simetría 146
2.3.6 Ley de la continuidad 147
2.3.7 Ley de la experiencia 148
2.3.8 Ley de la figura-fondo 148
2.4 El pensamiento visual 150
2.5 Conclusiones para una correcta representación de datos 153
3 Definición de un catálogo de componentes para visualización 154
4 Propuesta de un Catálogo de componentes para visualización de datos 156
4.1 Tablas de datos 157
4.2 Gráfico de líneas 158
4.3 Gráfico de coordenadas, gráfico de dispersión. 159
4.4 Gráfico de barras o columnas 160
4.5 Gráfico de secciones o circular 161
4.6 Mapa 162
4.7 Diagrama de caja 166
4.8 Treemap, mosaic plot 167
4.9 Sankey 168
4.10 Heat map 171
4.11 Sparklines 171
4.12 Nubes de palabras 172

Dimensión Informes 175
1 Introducción al sistema de informes 178
2 El diseño de la información 181
2.1 Arquitectura de la información 183
3 Tipos de informes 184
3.1 Reporting 185
3.2 Dashboard y cuadro de mando 185
3.2.1 Concepto de cuadro de mando 188
3.2.2 Cuadro de mando: Tableau du Board 188
3.2.3 Cuadro de mando: Balance Score Card 188
3.3 OLAP 190
3.4 Infografía 194
4 Diseño de informes 202
4.1 Formato físico del informe 204
4.2 Componentes 205
4.3 Ejemplos de informes 206
5 El Sistema de Informes propuesto en el modelo SIEGO 209

5.1 Estructura 209
5.2 La Gestión de los Informes 213
5.2.1 Identificar el mensaje 213
5.2.2 Planificar el informe 213
5.2.3 Elaborar el informe 214
6 El proceso de forma gráfica 217

Dimensión Publicar 219
1 Publicar los informes 221

Capítulo IV. Estudio Empírico 227
1. Preguntas de tesis y metodología 229
1 Objetivos del estudio empírico 231
2 Hipótesis 232
3 Posicionamiento epistemológico 233
4 Elección del método 235
5 Unidad de análisis y población 236
6 Muestra 238
6.1 Cálculo del tamaño muestral 238
7 Cuestionario y preguntas 239
7.1 Preguntas sobre resultados 243
7.2 Herramienta utilizada 245
7.3 Tipo de Escala y propiedades socio métricas asociadas 246
7.4 Fiabilidad de las escalas seleccionadas en la encuesta 247
7.5 Validez 250
2. Análisis de los resultados 253
1 Datos obtenidos 255
2 Metodología 256
2.1 Tabla descriptiva de variables, preguntas y dimensiones 257
2.2 Fiabilidad 259
2.3 Segmentación 260
2.4 Organizaciones con sellos de excelencia y/o certificados de calidad 262
3 Resultados de las dimensiones del modelo propuesto SIEGO: 263
3.1 Dimensión Datos 263
3.1.1 Utilización de bases de datos (Data Warehouse) con los datos organizados por tipos y materias (Data Marts). 263
3.1.2 Datos del Sistema de Gestión 264
3.2 Datos sobre Estrategia 265
3.2.1 Objetivos estratégicos (o líneas, o áreas, etc.) y estado de ejecución. 265
3.2.2 Datos sobre Clientes y Ventas 267
3.2.3 Nivel de satisfacción de los clientes. 268
3.2.4 Datos sobre Procesos / Productos / Servicios 269
3.3 Datos sobre Personas 270

3.3.1 Datos sobre la composición del equipo (número de personas, género, edad, nivel de estudios...) 270
3.3.2 Datos sobre el potencial de las personas, plan de carrera y plan de formación. 271
3.4 Dimensión Análisis 272
3.4.1 Técnicas estadísticas univariables 273
3.4.2 Técnicas estadísticas bivariables 274
3.4.3 Técnicas estadísticas multivariables 274
3.5 Dimensión Visualización 276
3.5.1 Utilización de tablas de datos 276
3.5.2 Gráficos de líneas 277
3.5.3 Uso de gráficos de columnas 278
3.5.4 Uso de gráficos treemap 278
3.5.5 Gráficos bubble o de burbuja 279
3.6 Dimensión Informes 281
3.6.1 Utilización de un sistema de informes para la presentación periódica de datos en papel o bien en formato digital (cuadros de mando, reporting, mapas, etc.). 281
3.6.2 Herramientas de análisis de datos como por ejemplo tablas dinámicas de Excel o cubos multidimensionales (OLAP). 282
3.7 Dimensión Publicación 283
3.7.1 Utilización de un espacio común (portal web, intranet, etc.) en el que se publican los informes. 283
3.8 Dimensión Resultados 284
3.8.1 Percepción sobre el clima laboral de la organización 284
3.8.2 Percepción sobre los resultados de la organización 285
3.8.3 Un sistema de información como el propuesto, SIEGO, mejoraría los resultados de la organización. 286
3.9 Dimensión Tiempo de respuesta 287
3.10 Comentarios de los encuestados 288
4 Análisis de las relaciones entre las dimensiones 289
4.1 Descriptivo por dimensiones del modelo propuesto y para resultados 290
4.2 Análisis de correlaciones entre las dimensiones 290
5 Comprobación de las hipótesis 292
5.1 H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados 292
5.2 H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas. 295
5.3 H3 Las organizaciones que disponen de datos perciben mejores resultados 296
5.4 H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados 298
6 Análisis multivariante: modelo de regresión 299
Capítulo V. Conclusiones 301
1 Objetivos 303
2 Resultados contrastados 304

3 Conclusiones y Contribuciones 306
4 Líneas de futuro 308

Bibliografia 311

Tabla de Ilustraciones

Ilustración 1. Clasificación de los criterios en funcion de su importancia relativa según Talwar. 12
Ilustración 2. Criterios del estudio de Talwar. Fuente: Elaboración propia 13
Ilustración 3 Edward Tufte Graphical Excellence 14
Ilustración 4. EIS = Executive information systems - Uso de sistemas de información ejecutiva. Fuente: José Luís Cano Giner (2014) 14
Ilustración 5. PDCA. Fuente: Elaboración propia (2016) 21
Ilustración 6. Business Excellence Models / National Quality Awards. Fuente: Talwar (2009) 23
Ilustración 7. Modelos de la Global Excellence Model Council. Fuente: www.efqm.org (2011) 24
Ilustración 8. Modelos relacionados con el GEM. 24
Ilustración 9. Tabla comparativa modelos de Excelencia. 25
Ilustración 10. Esquema Premio Deming. 28
Ilustración 11. Criterios Malcolm Baldrige National Quality Award (MBNQA), Fuente: http://www.nist.gov/baldrige/publications/business_nonprofit_criteria.cfm (junio, 2011) 29
Ilustración 12. Esquema del modelo EFQM. Fuente: www.efqm.org 31
Ilustración 13. Danish Business Excellence Index (DBEI) 32
Ilustración 14. Singapore Quality Award Framework. Fuente: SQA (2016) 32
Ilustración 15. Singapore Quality Award Framework. Fuente: SQA (2016) 33
Ilustración 16. Circulo PDCA. Fuente: Elaboración propia. 34
Ilustración 17. Componentes de un sistema de BI. Fuente: Forrester Research Inc. 41
Ilustración 18. Esquema de datos en estrella. Fuente: De Jesuja - Trabajo propio, GFDL, https://commons.wikimedia.org/w/index.php?curid=3886973 (2016) 46
Ilustración 19. Esquema Data Mart ventas. Fuente: Brull Alabart (2009) 47
Ilustración 20. Esquema en estrella o en copo de nieve. Fuente: Advanced Data WareHouse Design, Elzbieta Malinowski 48
Ilustración 21. Esquemas de un Data Mart. Fuente: Elaboración propia (2016) 48
Ilustración 22. Prioridades 2016. Fuente: Gartner (Feb 2016) 52
Ilustración 23. Búsqueda en Google Trends Business Intelligence 53
Ilustración 24. Búsqueda Google Trends Business Analytics 53
Ilustración 25. Distribución de usuarios en función de la sofisticación de la solución de Business Intelligence 53
Ilustración 26. Cuadrante de Gartner para Business Intelligence y Analytics. Fuente: Gartner (Feb 2016) 54
Ilustración 27. Business Analytics vs. Business Intelligence. Fuente: https://wiki.smu.edu.sg/is101_2012/img_auth.php/e/ec/Business_Analytics.jpg 56
Ilustración 28. Equivalencia entre un sistema de BI y el modelo SIEGO. Fuente: Elaboración propia (2016) 60
Ilustración 29. El modelo SIEGO. Fuente: Elaboración propia (2016) 61
Ilustración 30 Convertir datos en sabiduría. Fuente: Internet (2015) 66
Ilustración 31. De datos a sabiduría 66
Ilustración 32. Datos convertidos en información. Fuente: Internet (2016) 67
Ilustración 33. Arbol de configuración de valor. Fuente: Roos et al. (1997) 67
Ilustración 35. El proceso de creación de conocimiento y capital intelectual. Fuente: E. Bono (1999) 69
Ilustración 38. Equivalencia actividad con elementos del modelo SIEGO. 73
Ilustración 39. Datos, información, conocimiento 74
Ilustración 42. Escalas, transformaciones y estadísticas significativas 78
Ilustración 44. Modelo de Constructo. 82

Ilustración 45. La matriz de datos. Fuente elaboración propia. 84
Ilustración 46. Matriz de datos de elementos. Fuente: Elaboración propia. 84
Ilustración 47. Matriz de datos de hechos. Fuente elaboración propia. 86
Ilustración 48. Sistema de Indicadores propuesto. Fuente: Brull Alabart (2014) 89
Ilustración 53. Catálogo de Técnicas de Análisis. Fuente: Elaboración propia. 103
Ilustración 54. Serie temporal. Fuente: Elaboración propia 104
Ilustración 55. Media móvil. Fuente: Elaboración propia. 104
Ilustración 56. Tabla ejemplo evolución de una serie. Fuente: Elaboración propia. 105
Ecuación 1. Teorema de Chebyshev, 106
Ecuación 2. Cálculo del Teorema de Chebyshev con k=3 107
Ilustración 57. Simetrias y asimetrias. 108
Ilustración 58. Curtosis 108
Ilustración 59. Ejemplo tabla de contingencia bivariable. Fuente: Diputación de Tarragona (2016) 110
Ilustración 60. Análisis de correspondencias. Fuente: https://israelcruzcorona.wordpress.com/2014/03/05/analisis-factorial-de-correspondencias/ 111
Ilustración 61. Análisis de clúster. 112
Ilustración 62. Análisis discriminante. 113
Ilustración 63. Ejemplo de análisis estructural de un servicio. Fuente: Wikipedia (2015) 114
Ilustración 64. Selección de la técnica de tratamiento de datos más adecuada (Fuente: http://1.bp.blogspot.com/-ME24ePzpzIM/UQLWTwurfXI/AAAAAAAAANw/W3EETIroA80/s1600/drop_shadows_background.png 119
Ilustración 65. Distancia de Toledo a Roma. Fuente: Michael van Langren. 126
Ilustración 66. Años de vida en función de la edad. Fuente: Christian Huygens. 126
Ilustración 67. Mapa de previsiones de viento. Fuente: Edmond Halley. 126
Ilustración 68. Datos históricos en una línea de tiempo. Fuente: Priestley. 127
Ilustración 69. Biografia gráfica de personajes históricos. Fuente: Priestley. 128
Ilustración 70. Gráfico de coordenadas. Fuente: Playfair. 128
Ilustración 71. Diagrama de barras. Fuente: Playfair. 129
Ilustración 72. Evolución deuda Inglaterra. 129
Ilustración 73. Gráfico circular de la población de distintas capitales europeas. El tamaño del círculo es proporcional al dato (Statistical Breviary). 129
Ilustración 74. Población, impuestos y extensión superficial de distintos países europeos. Aparece por primera vez el gráfico de sectores (Statistical Breviary). 130
Ilustración 75. Gráfico de dispersión, Francis Galton 130
Ilustración 76. Gráfico de dispersión que recoge los datos de consumo de fertilizantes y producción de cereales en varios países europeos. 131
Ilustración 77. Destino del ejército de Napoleón en Rusia, E. Tufte. 132
Ilustración 78. Rutas de emigración en el mundo en 1858. El ancho de las bandas es proporcional al número de emigrantes. 132
Ilustración 79. Representación de las bajas en el ejército británico en la guerra de Crimea. 133
Ilustración 80. Diagrama de Gantt. 135
Ilustración 81. Diagrama de Hertzsprung-Rusell. 135
Ilustración 82. Otto Neurath, ISOTYPE. 136
Ilustración 83. Plano del metro de Londres, Henry C. Beck. 136
Ilustración 84. Diagrama de caja. 138
Ilustración 85. Prinicipio emergencia, figura dálmata. 142
Ilustración 86. Imágenes efecto reitificación. 142
Ilustración 87. Cubo de Necker y Vaso de Rubin. 143
Ilustración 88. Ley de proximidad. 144
Ilustración 89. Ley de semejanza o equivalencia. 144
Ilustración 90. Ley de Prägnaz o de la buena forma. 145
Ilustración 91. Ley de cierre. 146
Ilustración 92. Ley de simetría. 146
Ilustración 94. Ley de la experiencia. 148
Ilustración 95. Jarrón de Rubin. 149
Ilustración 96. Mariposa a la Flor, de Octavio Ocampo. 149

Ilustración 97. Preferencias de la mirada. 150
Ilustración 98. Perversiones de la mirada. 150
Ilustración 99. Teoría de la percepción gráfica, Cleveland & McGill. 151
Ilustración 100. Comparar gráfico de barras, burbujas y tonalidades. 152
Ilustración 101. Tabla de elementos de visualización de Ralph Lengler & Martin J. Eppler. 155
Ilustración 102. Gráfico de líneas. Fuente: Elaboración propia. 158
Ilustración 104. Variantes del gráfico de columnas. 160
Ilustración 105. Gráfico de barras o columnas. Fuente: Elaboración propia. 161
Ilustración 106. Gráfico de sección o circular. 162
Ilustración 107. Mapa de coropletas. 163
Ilustración 110. Gráfico de Voronoi. 165
Ilustración 111. Diagrama de caja. 166
Ilustración 112. Diagrama de caja enriquecido. 166
Ilustración 114. Treemap municipios 2012. Fuente: Diputació de Tarragona (2012). 168
Ilustración 117. Gráfico de coordenadas paralelas. 170
Ilustración 119. Muestra de sparklines. Fuente: Elaboración propia. 172
Ilustración 120. Muestra sparklines en tabla. 172
Ilustración 124. Tiempo de respuesta a una pregunta. Fuente: Elaboración propia (2016) 181
Ilustración 125. Arquitectura de la información. 184
Ilustración 132. Cubo OLAP de la Diputación de Tarragona. Muestra de drill-down. 194
Ilustración 152. Formato de un informe. 215
Ilustración 153. Tipografía. 215
Ilustración 154. Escala de colores. 215
Ilustración 155. Portal de informes de la Diputació de Tarragona (2016) 222
Ilustración 156. Tiempo de respuesta a una pregunta. Fuente: Elaboración propia (2016) 223
Ilustración 158. Sistema de carpetas en el portal de informes de la Diputació de Tarragona (2016) 225
Ilustración 159. Portal de informes Diputació de Tarragona. Fuente: Diputació de Tarragona (2016). 225
Ilustración 160. Elección del método. Fuente: Domínguez Amorós et al. (2000:93) 235
Ecuación 3. Tamaño muestral 239
Ecuación 4. Tamaño muestral, con valores 239
Ecuación 5. Proporción de pérdidas. 239
Ilustración 161. Imagen cuestionario, respuesta múltiple. 247
Ilustración 162. Tabla descriptiva variables, bloques, dimensiones y preguntas. 259
Ilustración 163. Segmentación por Sector de Actividad. Fuente: Elaboración propia (2016) 261
Ilustración 164. Segmentación por tamaño del equipo. Elaboración propia (2016) 261
Ilustración 165. Segmentación por sello de calidad o excelencia. Elaboración propia (2016) 262
Ilustración 166. Opinión encuestados. Fuente: Elaboración propia (2016) 289
Ilustración 167. Correlación entre dimensiones 292
Ilustración 168. Correlación entre dimensiones y resultados. 293
Ilustración 169. Modelo SIEGO. Fuente: Elaboración propia (2016) 294
Ilustración 170. Gráficos de caja relación datos con resultados. 297
Ilustración 171. Gráfico Resultados - Sello de calidad 298
Ilustración 172. Resultados - Sello EFQM. 299
Ilustración 173. Gráfico de la función. Fuente: Elaboración propia (2016) 300

Tabla de Ecuaciones

Ecuación 1. Teorema de Chebyshev, 106
Ecuación 2. Cálculo del Teorema de Chebyshev con k=3 107
Ecuación 3. Tamaño muestral 239
Ecuación 4. Tamaño muestral, con valores 239
Ecuación 5. Proporción de pérdidas. 239

Capítulo I. Presentación y motivación

"La perfección no se alcanza cuando no hay nada más que añadir, sino cuando no hay nada más que quitar"

Antoine de Saing-Exupéry

1. Introducción y motivación

El trabajo realizado trata de responder a la pregunta básica de cómo deberían ser los sistemas de información para ser útiles a las organizaciones y permitirles aumentar la eficiencia en su gestión. Una organización eficiente es aquella que comprende cuales son las necesidades de sus grupos de interés y es capaz de satisfacer sus requerimientos presentes y futuros, para obtener resultados positivos de forma sostenible en el tiempo. Disponer de información relevante y contextualizada de todos los ámbitos de su actividad así como del entorno, y de sus procesos o servicios es una variable de gestión clave para la organización a la hora de conseguir este objetivo.

En un entorno de crisis global e incertidumbre, las organizaciones buscan la eficiencia para ser más competitivas y generar más valor en sus actividades. Implantar un sistema de información facilita que la organización aprenda, que haya una retroalimentación permanente y que se mejoren los sistemas de toma de decisiones fundamentados en datos. Se trata de una herramienta esencial para los procesos de planificación, ejecución y revisión orientados a la mejora continua de la organización, como camino hacia la excelencia.

Las organizaciones disponen cada vez de más datos y sus rutinas se automatizan en mayor medida, dejando rastro digital de cada operación, proceso, actividad, etc.

Asimismo, cada ámbito de la organización dispone de su sistema de datos independiente (operaciones, proyectos, personas, gestión económica, etc.) de forma que los datos no pueden relacionarse de forma fácil, periódica y significativa.

Un incremento de información o incluso de la comunicación, no genera conocimiento por sí mismo, sino que, en general, cuanta más información, más dificultad para saber. "La hiper información y la hiper comunicación no inyectan ninguna "luz" en la oscuridad." Byung-Chul Han, "la sociedad de la transparencia", Herder, 2013.

Los datos creados deberían ser utilizados para generar conocimiento, y medir el comportamiento de la organización de forma habitual, para de este modo entender la situación actual y sus causas, así como para establecer la planificación de futuro.

Por ello es necesario un sistema de información que permita recoger los datos de las diferentes fuentes de información de la organización para convertirlos en conocimiento explícito que aporte valor a la organización y permita llevarla a un escenario de mayor eficiencia, sobretodo en la toma de decisiones.

Existen en el mercado infinidad de soluciones comerciales cerradas para implantar sistemas de información, pero el objetivo planteado en este trabajo es desarrollar un modelo de sistema de información desacoplado en el que cada uno de los elementos pueden ser seleccionados, instalados, actualizados o mejorados de forma asíncrona. Además la organización puede

implementar el modelo en todo o en parte, en función de sus necesidades, recursos, disponibilidad, conocimiento, etc.

Por otra parte, no todos los datos tienen el mismo valor. Es necesario determinar qué datos son los más relevantes, para ello en la tesis se investiga en la búsqueda de modelos de referencia contrastados en la literatura internacional que determinen qué conceptos/criterios deben medirse en las organizaciones que buscan ser las mejores y que quieren mantener sus resultados de forma sostenible en el tiempo.

En este sentido se analizan los requisitos de los modelos de excelencia empresarial y de los sistemas de calidad que permiten a las empresas autoevaluarse periódicamente con respecto a un modelo creado por una organización o institución con la intención de llevar las empresas que lo siguen a la calidad total o excelencia.

Para la tesis se toma como referencia el estudio de Talwar en el que se analizan más de 347 trabajos de investigación relacionados con los modelos de calidad total o excelencia, se identifican cuáles son los elementos o criterios más utilizados para evaluar las organizaciones, y estos son los criterios que se han tratado de validar en el marco del sistema de información propuesto. Mayoritariamente los modelos coinciden en la evaluación de 9 criterios, que son los siguientes:

- Liderazgo
- Planificación estratégica
- Personas (empleados)
- Proveedores
- Clientes
- Gestión del conocimiento y TI
- Procesos
- Sociedad
- Resultados

Una vez identificados los criterios, Talwar realizó una valoración de la importancia relativa de

Table IV Classification of criteria of BEMs based on their relative importance

Framework criteria	*Total value*[a]	*Average relative importance*[b]	*Rank range*	*Percentage score*	*EM/NQAs giving highest rank*
Core criteria – a must for survival					
Customer	43	2.15	1-4	9-30	Sweden, Japan, Singapore, India, Brazil, Mongolia, Thailand
People	49	2.45	1-5	9-20	EFQM, Canada, UK, GPNQA (India), Hungary, Belarus
Business results	51	2.55	1-8	12-46	MBNQA, Brazil, Mongolia, Israel, Malaysia
Internal environment criteria – the differentiators					
Processes	70	3.50	1-7	9-20	Deming (Japan), France, Israel
Leadership	75	3.75	1-7	9-33	Fiji, Japan, Deming (Japan), Australia
Strategic planning	118	5.90	3-9	6-10	Deming (Japan), EFQM, UK, RGNQA (India)
Knowledge and information management	126	6.30	4-9	6-10	RGNQA (India)
Goodwill criteria – ensuring sustenance					
Society	148	7.40	4-9	0-10	RGNQA (India), Brazil, EFQM, UK
Suppliers/partners	160	8.00	5-9	0-11	Deming (Japan), Belarus

Note: [a]TV; [b](TV)÷/20

Ilustración 1. Clasificación de los criterios en funcion de su importancia relativa según Talwar.

cada uno de ellos en los diferentes modelos de excelencia, estableciendo así una clasificación jerárquica en función de la comparación de los criterios:

De este modo estableció que algunos de los criterios formaban los criterios principales o CORE CRITERIA, que serían los relativos a **Clientes, Personas y Resultados.**

En segundo lugar estableció que los criterios de **Procesos, Liderazgo, Planificación estratégica y Gestión del conocimiento** formarían los criterios internos o INTERNAL ENVIROMENT CRITERIA.

Finalmente, el criterio de **Sociedad** y el de **Proveedores**, son los criterios deseables para asegurar la sostenibilidad, a los que Talwar llama GOODWILL CRITERIA.

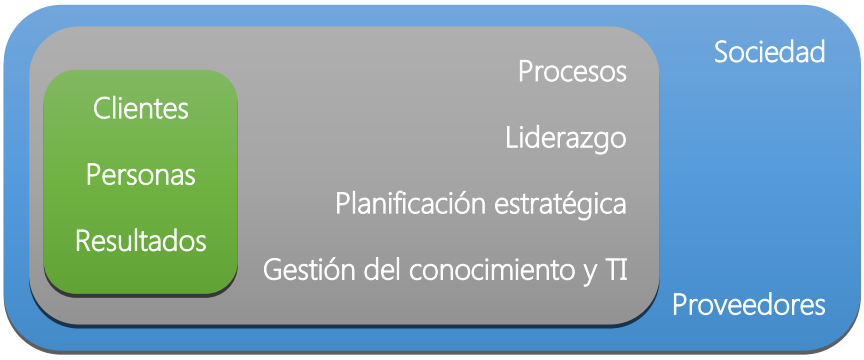


Ilustración 2. Criterios del estudio de Talwar. Fuente: Elaboración propia

Se establecen entonces por el autor cuales son los criterios que una organización excelente debe asegurar.

Por ello, se ha tomado esta referencia para validar si las organizaciones disponen de este tipo de datos en sus sistemas de información cuando se hace el análisis empírico del modelo.

Otra aportación que nos ayuda a desarrollar el sistema de información que se propone en este trabajo, son las publicaciones de Edward **Tufte**, de las cuales se extrae la teoría de la **visualización** y arquitectura de datos, entendida como la habilidad para ordenar, seleccionar y analizar los datos, aspecto que se considera básico en el modelo planteado.

La arquitectura de datos de una organización debe permitir al usuario encontrar la vía de navegación hacia el conocimiento y la **compresión**, cuanto más rápida posible, de la información (**Wurman**, 1996). El informe excelente es el que da al espectador el mayor número de ideas, en el menor tiempo y con la menor cantidad de tinta, el espacio más pequeño… y que dice la **verdad** acerca de los datos (**Tufte**, 1999).

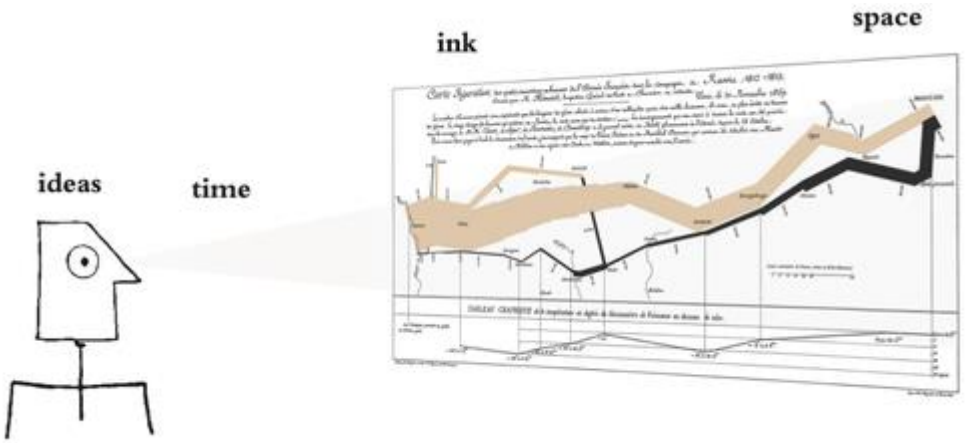


Ilustración 3 Edward Tufte Graphical Excellence

Tufte, incide desde sus investigaciones en que en la interpretación de la información gráfica existe un cierto periodo de aprendizaje, y que ello debe utilizarse en las organizaciones para asegurar la familiarización con los datos, presentándolos siempre con los mismos patrones o esquemas.

Asimismo, José Luis Cano, establece en su tesis doctoral "¿Qué grupos de factores creen los ejecutivos sénior que afectan más al uso de los sistemas de información ejecutiva?", que los tres factores principales son:

- El contenido y la accesibilidad a la información
- La contribución al rendimiento y la vinculación con los objetivos de negocio
- El diseño

Average position	Cluster name	Average rating
1	EIS content and access to information	4.05
2	EIS focus	3.99
3	EIS design	3.97
4	Executives' involvement in EIS	3.76
5	EIS accessibility	3.73
6	EIS confidence	3.63
7	Importance of EIS use in the organization	3.59
8	Executives' proximity to EIS	3.54
9	EIS ease of use	3.46
10	EIS organizational behavior	3.40
11	Executives' attitudes	3.19
12	Executives' ability to use EIS	2.99
	Total	3.59

Ilustración 4. EIS = Executive information systems - Uso de sistemas de información ejecutiva. Fuente: José Luís Cano Giner (2014)

De todo lo comentado hasta ahora, se considera apropiado plantear en este trabajo de investigación un modelo que dé cumplimiento a las necesidades de información de las empresas o instituciones, mediante un sistema de información para la eficiencia, que se ha determinado llamar SIEGO y que consta de los elementos siguientes:

1. Un **modelo estructurado de datos** que presenten la información y las relaciones entre los criterios fundamentales de una organización.
2. Se propone también un catálogo de **técnicas analíticas,** para aplicarlas en la revisión y estructuración de los datos.
3. Asimismo, se presenta un inventario de **componentes de visualización** de datos.
4. Con todo ello, se propone un **sistema de informes** corporativo para presentar el sistema de información a los grupos de interés de la organización. Presentar los datos de la forma más eficiente e intuitiva para generar el mayor conocimiento, en el menor tiempo posible.
5. Un espacio común y único para **publicar los informes**, en el que las personas de la organización encuentren toda la información que precisan para su gestión, de forma estructura y organizada por temas, con una política de permisos y contenidos adecuados a cada perfil y cada grupo de interés.

Una vez planteado el modelo de sistema de información se ha considerado imprescindible validar la adecuación de su estructura y de los contenidos que incorpora para las organizaciones que decidan implantarlo. Para ello se ha desarrollado una investigación empírica en la que se intenta contrastar si las organizaciones que disponen de un modelo parecido obtienen **mejores resultados**.

Igualmente, se propone relacionar la **rapidez en la respuesta** a las solicitudes de información, con la existencia o no de un sistema de información como el que se propone, ya que esta cuestión se considera básica para que los procesos de toma de decisiones de la organización se ejecuten de forma eficiente.

El método escogido para contrastar las hipótesis, es el del cuestionario cuantitativo enviado a una muestra de organizaciones.

La propuesta de modelo no sólo nace de la reflexión de la profunda revisión bibliográfica realizada, sino que es en gran parte fruto de la experiencia de la doctoranda en el desarrollo de la aplicación de este sistema en una organización de más de 1.200 trabajadores.

La organización, que en el año 90 se planteó iniciar el camino de la calidad y más tarde el de la Excelencia, es en este momento una de las únicas en su sector que se encuentra en un nivel de excelencia de más de 400 puntos EFQM, y que muy probablemente este año conseguirá el sello de oro 500+ EFQM.

Ello llevó a la organización a buscar un sistema de gestión de su conocimiento que en un principio se consideró para la información económica y presupuestaria, y que más tarde se quiso ampliar a todas las demás actividades de la organización. Este sistema debía permitir tomar decisiones con mayor eficiencia y participar en la obligación de rendir cuentas a los grupos de interés.

Se empezó con la implantación de una solución propietaria de licenciamiento, y pronto se vio que ello no permitía la evolución a otros ámbitos y la integración con otros sistemas.

Por ello se desacopló el sistema y se empezó a trabajar con herramientas de plataformas open source (de código abierto, gratuitas) que hasta la fecha están resolviendo enteramente las necesidades del sistema de información de esta organización.

La experiencia de la doctoranda como protagonista en este proceso ha sido el motivo y la fuente de inspiración para plantear este trabajo. Se puede decir que la motivación personal que ha llevado a realizar esta tesis ha sido doble: por un lado tratar de recoger en un modelo toda la experiencia ganada con el desarrollo, implantación y mejora de un sistema de información en una organización y que esto pueda servir para quien se proponga realizar un proceso similar; y por otro validar, contrastando con la literatura y con la experiencia de otras organizaciones, que el modelo planteado resulta realmente útil y responde de forma eficaz a los requerimientos de una organización excelente.

2. Estructura de la tesis

Este trabajo de investigación se estructura en cinco capítulos.

En el primer capítulo se hace introducción al tema, se da cuenta de la motivación y se explica la estructura de la tesis.

El segundo capítulo es el correspondiente a la explicación del marco y fundamento teórico de este trabajo, que finaliza con la justificación de la necesidad de un modelo como el que se presenta en este trabajo. El marco teórico parte de los diferentes referentes de sistemas de gestión para llegar a los modelos de excelencia existentes en el mundo. Estos modelos de excelencia cubren diferentes criterios de la organización que deben ser medidos y mejorados, lo cual justifica la necesidad de un sistema de información. Se revisa cómo debe ser este sistema y cuáles son los sistemas de información que actualmente se utilizan. El Business Intelligence como sistema de información más utilizado actualmente, sus componentes, y la descripción de las herramientas que proporciona. El Data Warehouse, las ETL (herramientas de carga de datos) y los sistemas de reporting existentes. También se menciona hacia donde evolucionará el mercado de soluciones Business Intelligence según Gartner, y el concepto de Business Analytics. Finalmente se mencionan las ventajas de implantar un sistema de información

En el tercer capítulo se presenta y explica detalladamente el modelo que se propone, el modelo SIEGO. Se trata cada una de las dimensiones separadamente: la preparación y estructuración del sistema de datos en la dimensión datos, el catálogo de técnicas de análisis en la dimensión análisis, la dimensión visualización que contiene el catálogo de componentes de visualización de datos, la elaboración y estructuración del sistema de informes así como su publicación, en la dimensión publicación.

Para la preparación del sistema de datos, se tiene en cuenta la aproximación de la gestión del conocimiento, el proceso en sí y sus fuentes. A continuación se trata el dato como medida, los tipos de escalas y los requerimientos de la medición. Finalmente, se explica la matriz de datos, que es la que fundamentalmente da respuesta a las peticiones de información. La dimensión datos termina con una propuesta de estructura de datos, dimensiones y tablas de hechos.

En lo referente a la dimensión análisis, se propone un diccionario de técnicas de análisis y se describe cada una de ellas, sean de tipo univariable, bivariable o multivariable.

La dimensión visualización se inicia con un repaso histórico de la representación gráfica de datos, y de cómo influye la percepción humana en la comprensión de los datos. Se explican las diferentes leyes de la visión que la Gestalt formuló. Finalmente se resumen unas conclusiones sobre la visualización de datos y se define el catálogo de componentes propuesto, que se clasifican en básicos o avanzados. Para cada componente se indica hace una descripción, se indica para qué tipo de datos es adecuado, qué variantes existen y cuáles son las aplicaciones

más indicadas. Al final de esta dimensión se muestra un gráfico para escoger el tipo de componente adecuado en función de los datos que se quieren mostrar.

El sistema de informes que se propone se determina en cuanto a cómo diseñarlo y estructurarlo y en cuanto a qué metodología seguir para su elaboración. Se indican que tipos de informes existen, y para qué tipo de fines son adecuados.

Para finalizar este capítulo se indica cómo debe de ser el portal de informes de la organización, pieza final del modelo.

En el cuarto capítulo se hace el análisis empírico de la validación del modelo propuesto y de las preguntas e hipótesis de tesis, presentando las preguntas y la metodología de una parte, y los resultados obtenidos de otra.

Finalmente, el capítulo cinco presenta las conclusiones a las que la autora del trabajo ha llegado con el análisis efectuado en esta investigación.

Capítulo II. Marco y fundamento teórico

"Cuando se les habla de un nuevo amigo, jamás preguntan sobre lo esencial del mismo. Nuncas se les ocurre preguntar: ¿Qué tono tiene su voz? ¿Qué juegos prefiere? ¿Le gusta coleccionar mariposas?

Pero en cambio preguntan: ¿qué edad tiene? ¿cuántos hermanos? ¿cuánto pesa? ¿cuánto gana su padre? Solamente con estos detalles creen conocerle.

Pero nosotros, que sabemos comprender la vida, nos burlamos tranquilamente de los números"

Antoine de Saing-Exupéry

1 La gestión de las organizaciones

Un sistema de gestión es un esquema general de procesos y procedimientos que se emplea para garantizar que la organización realiza todas las actividades necesarias para alcanzar sus objetivos" (Glosario de términos de la EFQM, Modelo EFQM de Excelencia de 2003).

Existen en el mundo una gran cantidad de referencias en el ámbito de la gestión de las organizaciones. Hay referenciales que son fruto de aportaciones de autores y los hay que son originales de determinadas organizaciones estatales o supra estatales.

A continuación se explican aquellos referenciales que se consideran más fundamentados, integrados y sencillos para construir un Sistema de gestión.

1.1 Círculo de Deming o PDCA (Plan, Do, Check, Act)

La primera referencia, realmente simple pero a la vez muy potente, es la conocida como círculo de Deming o PDCA (del inglés Plan –Planificar-, Do –Ejecutar-, Check –Revisar-, Act –Actuar), que especifica que un sistema está bien gestionado cuando se cumplen las cuatro fases del círculo continuo: se planifica, se ejecuta, se revisa y se actúa en la mejora continua, es decir se mejora.

Ilustración 5. PDCA. Fuente: Elaboración propia (2016)

a) Planificar

Integra aquellas actividades destinadas a pensar en qué/como/cuando… haremos las cosas. Significa situarse, a nivel de pensamiento, en el futuro para decidir la planificación estratégica y táctica. Requiere habilidades estratégicas y de imaginación de escenarios.

b) Ejecutar

Integra aquellas actividades que se llevan a cabo para conseguir la planificación prevista en el Plan. Requiere habilidades interrelacionales y de coordinación.

c) Revisar

Integra aquellas actividades destinadas a gestionar el conocimiento de la organización. Implica ver la organización como un sistema, comprendiendo sus recursos y capacidades. Requiere habilidades de análisis complejo, desde todas las perspectivas de los grupos de interés.

d) Actuar

Integra aquellas actividades que mejoran la organización y parte del conocimiento extraído en la fase de Revisar. Requiere habilidades de creatividad y de innovación.

1.2 Modelos de Excelencia

Los enfoques de gestión han evolucionado desde una posición reactiva hacia una posición preventiva, desde el control de calidad estadístico, el aseguramiento de la calidad, la gestión de la calidad total (TQM) hacia los modelos de excelencia. Una característica destacable de los modelos de excelencia es la adición de aspectos éticos, ambientales y sociales; es decir, van tendiendo hacia la integración de diferentes áreas gestoras.

Actualmente, el modelo de gestión más evolucionado son los denominados Modelos de Excelencia. Los modelos de excelencia tienen la característica de tener una visión global de la organización.

Excelencia en el negocio, según lo descrito por la European Foundation for Quality Management (EFQM), se refiere a "las prácticas sobresalientes en la gestión de la organización y el logro de resultados, todo ello basado en un conjunto de ocho conceptos fundamentales."

En general, los modelos de excelencia empresarial han sido desarrollados por organismos nacionales como base para programas de premios. Para la mayoría de estos organismos, los premios son un vehículo para facilitar la adopción generalizada de los conceptos de excelencia empresarial, que en última instancia llevan a un mejor desempeño económico nacional. La mayoría de las organizaciones que utilizan estos modelos ponen en práctica la autoevaluación, a través de la cual se pueden identificar puntos fuertes y áreas de mejora.

Diferentes autores han opinado sobre las ventajas del uso de los modelos de excelencia (Grigg y Mann, 2008; Oakland y Tanner, 2008).

Aun considerando que el uso de modelos de excelencia es de buena ayuda, existen algunas opiniones que defienden que la relación entre la aplicación de los criterios y los resultados de la organización no son tan evidentes (Williams, Bertsch, Van der Wiele, Van Iwaarden y Dale, 2006).

Oakland (2008) afirma que en el escenario europeo, la Excelencia Empresarial se considera que va más allá de Gestión de la Calidad Total (TQM), pero en los EE.UU. los términos de Calidad Total y Excelencia Empresarial son intercambiables. Bauer (2002) hizo una revisión exhaustiva de la literatura y argumentó que la Excelencia Empresarial es una extensión de la TQM. En este trabajo utilizaremos el término Excelencia como una evolución de la TQM.

Talwar (2011:24) revisa 100 BEMs/NQAs (Business Excellence Models / National Quality Awards) que están siendo patrocinados en 82 países del planeta.

Region	No. of awards	Countries
Sub-Saharan Africa	3	Mauritius, South Africa, Southern African Development Community (SADC)
East Asia and the Pacific	17	Australia, Brunei, Fiji, Hong Kong, Indonesia, Japan (two), Republic of Korea, Malaysia, Mongolia, New Zealand, The Philippines, Singapore, Republic of China, Thailand, Vietnam, International Asia Pacific Quality Award
South Asia	7	India (four), Nepal, Pakistan, Sri Lanka
Europe and Central Asia	50	Austria, Belarus, Belgium, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany (two), Greece (two), Hungary, Iceland, Ireland (two), Italy, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Luxembourg, Malta, The Netherlands, Norway, Poland, Portugal, Romania, Russia (two), Scotland, Slovakia, Slovenia, Spain (two), Serbia and Montenegro, Sweden, Switzerland (two), Turkey (two), Ukraine, UK (two), Wales, European Union (EQA), Ibero-American Quality Award
Middle East and North Africa	9	Egypt, Israel (two), Iran, Jordan, Morocco, Oman, Qatar, United Arab Emirates
Americas	14	Argentina, Brazil (two), Canada, Chile, Colombia, Costa Rica, Ecuador, Mexico, Paraguay, Peru, USA (two), Uruguay

Notes: Total: 100 awards in 82 countries
Sources: Talwar (2009), Mavroidis *et al.* (2007), UNECE (2004), Calingo (2002)

Ilustración 6. Business Excellence Models / National Quality Awards. Fuente: Talwar (2009)

Algunos de ellos están agrupados en una organización denominada Global Excellence Model Council (GEM) cuyos miembros son los principales impulsores de los modelos de excelencia de todo el planeta.

Ilustración 7. Modelos de la Global Excellence Model Council. Fuente: www.efqm.org (2011)

Los modelos relacionados con el GEM son:

Europa	EFQM
América	EEUU, Baldrige National Quality Program Iberoamerica, Redibex (representado por FNO Brazil & IFCT Mexico)
Asia	Japon, Japanesse Productivity Development for Socio-Economic Development Singapur, Spring Singapore India, Confederation of Indian Industry
Oceanía	Australia, SAI-Global

Ilustración 8. Modelos relacionados con el GEM.

En la actualidad, a nivel internacional, los tres modelos más conocidos son los basados en el premio nacional de Estados Unidos, el Malcolm Baldrige (MBNQA), el basado en el premio europeo a la calidad (EFQM) y el premio nacional a la calidad en Japón (Deming). Su alto impacto mediático a nivel mundial está soportado porque a priori deberían contribuir a que las industrias en sus respectivos países incrementasen sus resultados y su competitividad.

Según Sharma y Talwar (2007), esta circunstancia ha hecho que estos modelos se hayan consolidado como una referencia internacional para alcanzar la excelencia, siendo la base de la gran mayoría de modelos desarrollados a final de los años ochenta y en la década de los

noventa, (Bohoris, 1995; Cauchick, 2001; Hughesy Halsall, 2002; Koura, 2009; Kouray Talwar, 2008; McDonald, et al., 2002; Powell, 1995; Sharma y Talwar, 2007).

En esta línea, Tan et al. (2003) indican que la mayoría de modelos de excelencia nacionales están basados bien en el modelo MBNQA o en el Modelo EFQM. Esta evidencia se constata igualmente en otros estudios (Cauchick, 2005; NIST National Institute of Standards and Technology, 2010).

Talwar (2011:29) afirma "Tres son los modelos de excelencia más implantados en el mundo, el MBNQA (Premio Malcom Baldrige Nacional de Calidad, EFQM y el Premio Deming, que son reconocidos en todo el mundo y se utilizan como base de la mayoría de los modelos de excelencia en otros países".

A continuación se muestra una tabla comparativa de cuáles son los criterios en cada uno de los modelos principales, Deming, Malcom Baldridge y EFQM:

Modelo Malcolm Baldrige	Modelo de excelencia de la EFQM	Modelo Premio Deming
1. Liderazgo	1. Liderazgo	1. Políticas de dirección y su despliegue
2. Planificación estratégica	2. Política y estrategia	2. Desarrollo de nuevos productos, innovación de procesos
3. Enfoque en el cliente y el mercado	3. Personas	3. Mantenimiento y mejora
4. Medida, análisis y gestión del conocimiento	4. Alianzas y recursos	4. Recopilación, transmisión y utilización de la información sobre calidad
5. Enfoque en los recursos humanos	5. Procesos	5. Sistemas de dirección
6. Gestión por procesos	6. Resultados en los clientes.	6. Análisis de información y utilización de las TIC
7. Resultados	7. Resultados en las personas	
	8. Resultados en la sociedad	
	9. Resultados clave	

Ilustración 9. Tabla comparativa modelos de Excelencia.

Los modelos de excelencia (George, 1992) propugnan modelos multifactoriales con la finalidad de obtener mayor eficacia de las organizaciones.

El número de modelos y premios asociados se ha incrementado considerablemente en las últimas dos décadas. Varios artículos académicos han realizado una recopilación de los mismos.

- Koura (2009) identifica 94 premios en 77 países.
- Un informe de UNECE (2004), citado en Koura y Talwar (2008), resalta que hay al menos 90 premios en como mínimo 75 países.
- Sharma y Talwar (2007) incluyen dos referencias, la primera a un informe de Martin Searle (2005) que muestra que al menos 53 países han basado sus premios nacionales en el Modelo Malcolm Baldrige y la segunda, Kozo y Yoshizawa (2005) en la que se indica que 26 países han basado su premio en el Modelo EFQM.
- Mavroidis *et al.* (2007) analizan 25 premios que han sido desarrollados en las últimas dos décadas en Europa.
- Cauchick (2005) también hace referencia a que al menos 76 países disponen de premios de excelencia, resaltando que existen tres premios regionales («Asia-Pacific Quality Award», «Iberoamerican Quality Award» y «European Quality Award»).
- Tan (2002) comparó 16 premios nacionales de calidad estableciendo recomendaciones para aquellos países que todavía no disponían de premio.
- En un trabajo del National Institute of Standards Technology (NIST, 2010) se muestran 95 premios.

Koura (2009) resalta el carácter dinámico de los modelos así como las revisiones periódicas que llevan a cabo sus entidades promotoras para actualizar sus contenidos, lo que justifica que las diferencias entre los modelos se estén reduciendo, al incorporar los distintos modelos las tendencias en gestión. Koura también lleva a cabo una comparación de los valores fundamentales de 16 modelos de gestión de la calidad.

Sila y Ebrahimpour (2003) analizaron 76 estudios y concluyeron que los factores más comunes de la TQM son:

- alta dirección: compromiso y liderazgo
- empoderamiento de los empleados
- gestión de proveedores
- planificación estratégica
- gestión de recursos humanos
- formación
- participación de los empleados
- diseñar el trabajo en equipo
- satisfacción de los empleados
- productos y servicios
- gestión de procesos
- control de procesos
- orientación al cliente

- información y análisis
- benchmarking
- la mejora continua
- la garantía de calidad
- la responsabilidad social

A continuación mostramos un resumen de algunos de estos modelos de excelencia así como otros derivados de estos, por su relevancia o nivel de implantación.

1.2.1 Premio Deming

El Premio Deming, que se otorgó por primera vez en el año 1951, es un premio anual al que se presenta una organización que ha implementado prácticas de excelencia adecuadas a su filosofía de gestión, alcance / tipo / escala de negocio. Sin importar el tipo de operaciones, cualquier organización puede solicitar el Premio bajo ciertas condiciones, ya sea pública o privada, grande o pequeña, nacional o extranjera, o parte o toda la organización.

El Premio Deming se da a las organizaciones que han ejercido una inmensa influencia directa e indirectamente en el desarrollo del control y la gestión en Japón.

Las organizaciones solicitantes desarrollan métodos eficaces de gestión de calidad, establecen las estructuras para la implementación y ponen en práctica los métodos.

Comúnmente, los que han desafiado al Premio comparten la sensación de que han tenido una experiencia valiosa y que el principio de gestión de lograr un éxito empresarial a través de la mejora de la gestión ha funcionado realmente. A través de presenciar el éxito de estas organizaciones, muchas otras organizaciones se han inspirado para comenzar su propia búsqueda.

Los criterios básicos del premio Deming son:

1. Políticas de dirección y su despliegue
2. Desarrollo de nuevos productos, innovación de procesos
3. Mantenimiento y mejora
4. Sistema de gestión
5. Análisis de información y utilización de las tecnologías de información
6. Desarrollo de recursos humanos

El peso en puntos de los diferentes criterios es el que se indica a continuación:

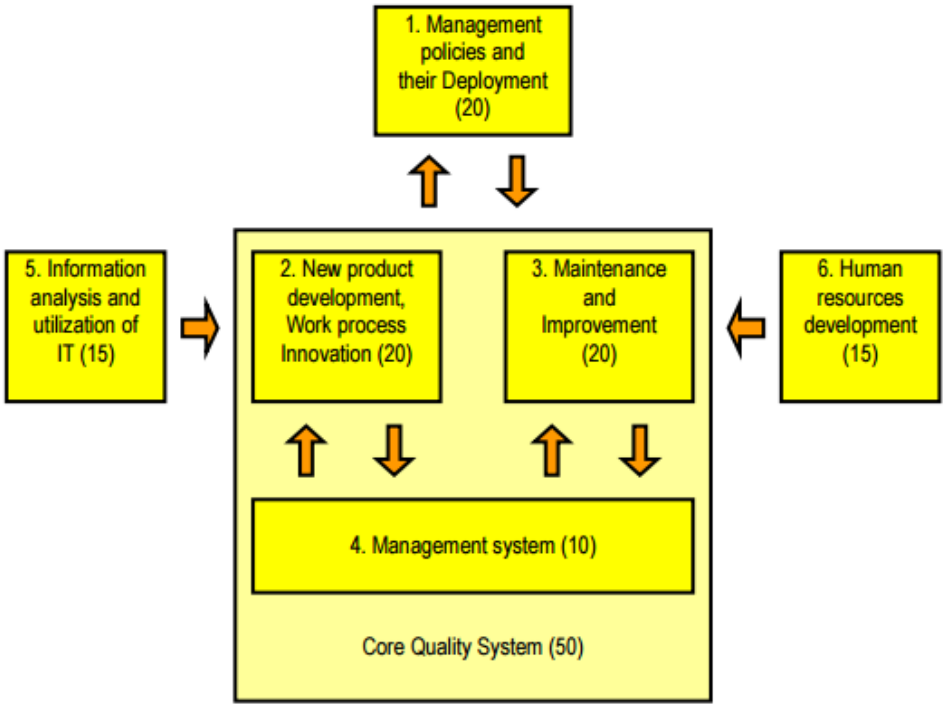


Ilustración 10. Esquema Premio Deming.

1.2.2 Malcom Baldridge National Quality Award

Es el modelo de los EE.UU. Malcolm Baldrige National Quality Award (MBNQA), incorpora un proceso de adjudicación administrado por la American Society for Quality (ASQ) y gestionado por el Instituto Nacional de Ciencia y Tecnología (NIST), una agencia del departamento de Comercio de EE.UU.

Los conceptos básicos de los Criterios Baldrige de Excelencia son:

1. Liderazgo
2. Planificación estratégica
3. Enfoque al cliente y al mercado
4. Medida, análisis y gestión del conocimiento
5. Enfoque en los recursos humanos
6. Gestión por procesos
7. Resultados organizativos

Ilustración 11. Criterios Malcolm Baldrige National Quality Award (MBNQA), Fuente: http://www.nist.gov/baldrige/publications/business_nonprofit_criteria.cfm (junio, 2011)

Además de los criterios básicos, el modelo se construyó sobre unos valores o conceptos fundamentales. Las organizaciones excelentes deben difundir estos valores y conceptos para que la se incorporen a la cultura de la organización:

- Liderazgo visionario
- Excelencia orientada al consumidor
- Aprendizaje organizativo y personal
- Valorar a empleados y socios
- Agilidad
- Enfoque en el futuro
- Gestión para la innovación
- Gestión por hechos
- Responsabilidad social
- Enfoque en resultados y en creación de valor
- Perspectiva del sistema

Jayamaha, Grigg y Mann (2009) analizó los resultados de las autoevaluaciones de diversas organizaciones.

Tiene un premio asociado, el Malcolm Baldrige Award, que se otorgó por primera vez el año 1987.

1.2.3 El modelo EFQM de Excelencia

Es un modelo creado a finales de los años 80 y que presenta tres partes básicas: los Conceptos Fundamentales, la Estructura y la Lógica Radar.

El Modelo EFQM se basa en ocho conceptos fundamentales que resultan aplicables a organizaciones de todo tipo, independientemente de su sector o tamaño, y que constituyen la base del modelo. La versión actual està soportada por los siguientes conceptos fundamentales (EFQM, 2013):

1. Añadir valor para los clientes.
2. Crear un futuro sostenible.
3. Desarrollar la capacidad de la organización.
4. Aprovechar la creatividad y la innovación.
5. Liderar con visión, inspiración e integridad.
6. Gestionar con agilidad.
7. Alcanzar el éxito mediante el talento de las personas.
8. Mantener en el tiempo resultados sobresalientes.

El denominado Modelo o estructura del Modelo que tiene un primer nivel de 9 criterios y un segundo nivel de 32 subcriterios. Además, cada subcriterio tiene diferentes elementos. Los 9 criterios son:

1. Liderazgo
2. Estrategia
3. Personas
4. Alianzas y recursos
5. Procesos, productos y servicios
6. Resultados cliente
7. Resultados personas
8. Resultados sociedad
9. Resultados clave

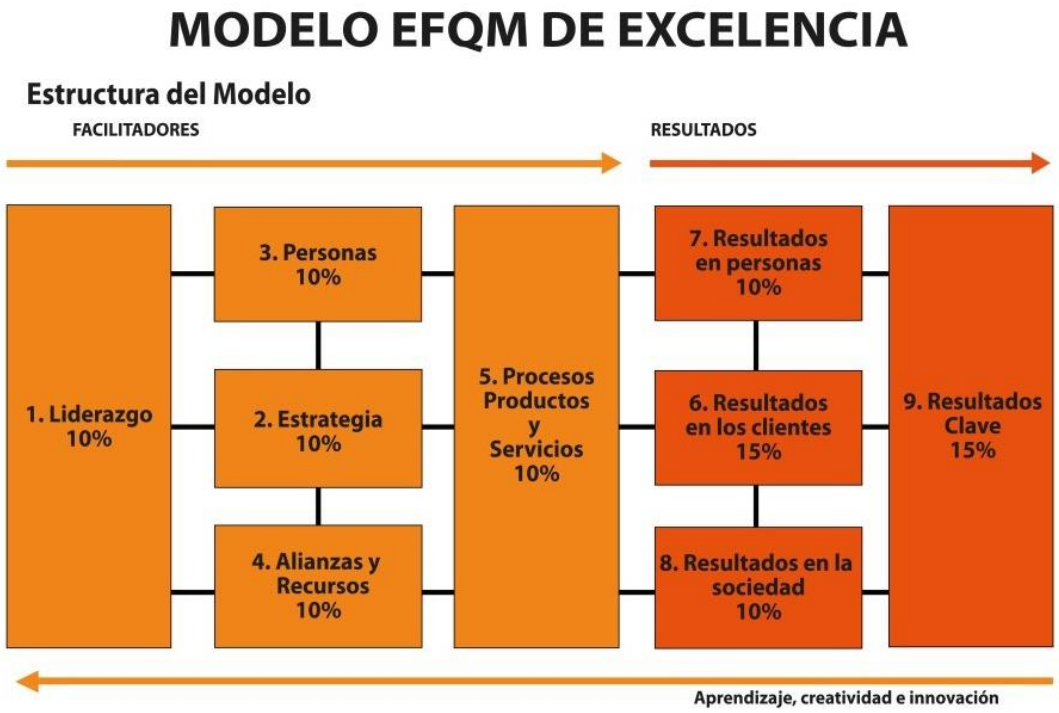


Ilustración 12. Esquema del modelo EFQM. Fuente: www.efqm.org

Derivados del modelo EFQM nos encontramos, entre otros, el Modelo Danés, el Iberoamericano, el de la India y el de Singapur.

Se puede profundizar el comportamiento de este modelo en Haffer y Kristensen (2008).

Tiene un premio asociado que se otorgó por primera vez el año 1991.

1.2.4 Danish Business Excellence Index (DBEI)

Está basado en el Modelo EFQM de Excelencia.

Está publicado por la Danish Association for Quality Management en asociación con Aarhus School of Business y la empresa Markeds/Consult A/S.

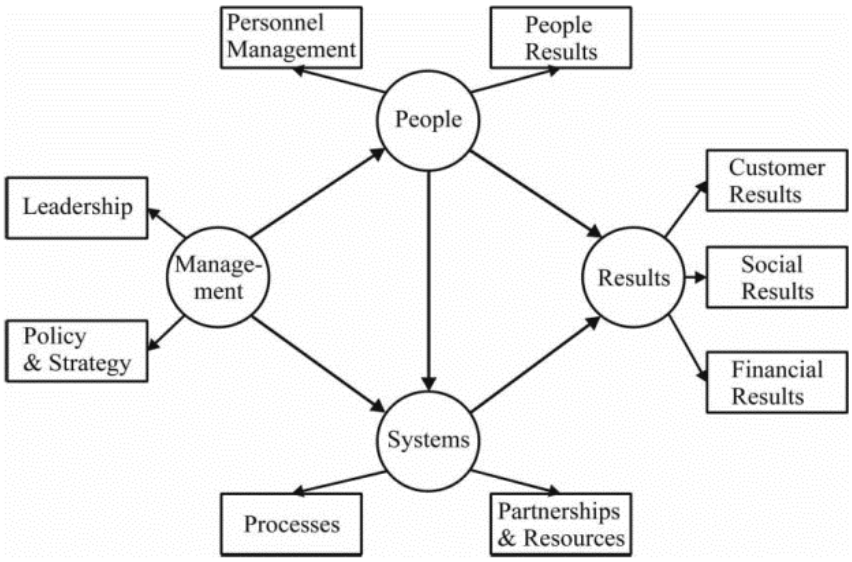


Ilustración 13. Danish Business Excellence Index (DBEI)

El modelo consta de cuatro componentes fundamentales:

1. Gestión
2. Personas
3. Sistemas
4. Resultados

Se puede profundizar el comportamiento de este modelo en Haffer y Kristensen (2008).

1.2.5 Singapore Quality Award Framework (SQAC)

El Singapore Quality Award es modelo derivado del Modelo EFQM de Excelencia (europeo).

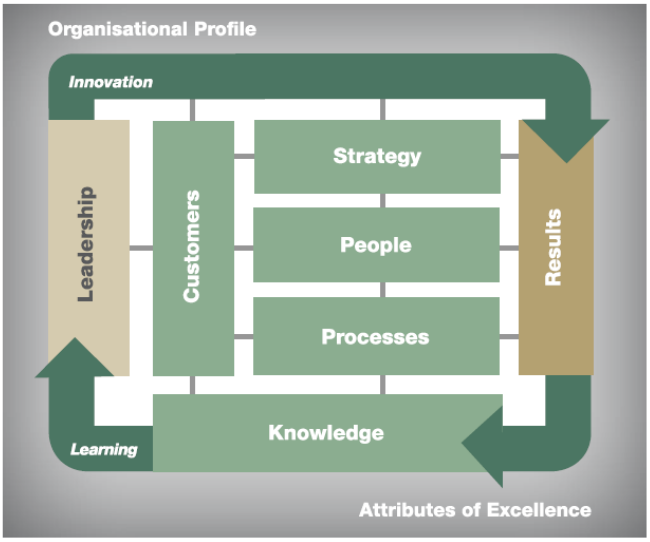


Ilustración 14. Singapore Quality Award Framework. Fuente: SQA (2016)

El Marco del Premio de Singapur a la Calidad (SQA) se utiliza como base para la evaluación de las organizaciones de Singapur para los más altos estándares de calidad y excelencia empresarial. El premio tiene por objeto establecer a Singapur como un país comprometido con la excelencia empresarial de clase mundial.

El marco se basa en los siguientes valores fundamentales y conceptos:

1. Liderazgo visionario
2. La calidad basada en el cliente
3. Centrado en la innovación
4. Valoración de las personas y socios
5. Agilidad
6. El conocimiento impulsado por el sistema
7. Responsabilidad social
8. Orientación hacia los resultados
9. Sistemas de perspectiva

El marco consta de siete categorías:

1. Liderazgo
2. Planificación
3. Información
4. Personas
5. Procesos
6. Clientes
7. Resultados

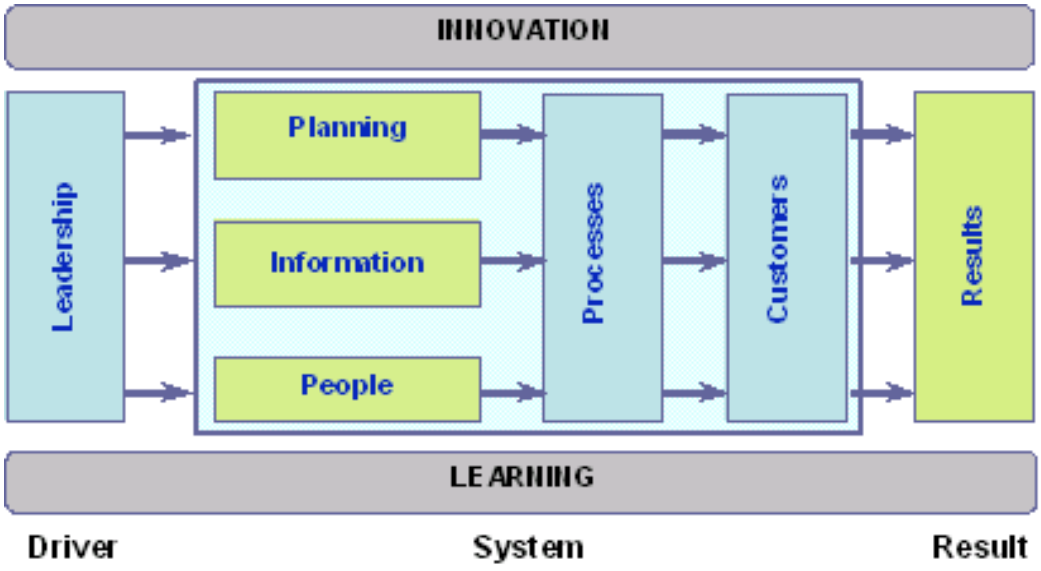


Ilustración 15. Singapore Quality Award Framework. Fuente: SQA (2016)

2 De la Excelencia a los Sistemas de Información

Del análisis de los sistemas de gestión y premios a la excelencia se desprende que los principales valores que comparten los modelos son los siguientes:

- Liderazgo y constancia de propósito.
- Orientación al cliente.
- Mejora continua, innovación y aprendizaje.
- Gestión de procesos basados en hechos.
- Desarrollo e implicación del personal.
- Responsabilidad social.
- Visión de futuro.
- Orientación a resultados y creación de valor.
- Desarrollo de parternariado

Estos valores constituyen los pilares sobre los que los modelos de excelencia se han desarrollado y en los que las empresas han encontrado una filosofía de gestión orientada al éxito a largo plazo y a la mejora de sus resultados. Un factor común a todos los modelos es su utilización por parte de las organizaciones como referencia para llevar a cabo sus autoevaluaciones. Las autoevaluaciones, permiten identificar áreas de mejora y puntos fuertes sobre los que desarrollar planes de mejora que pueden ser integrados en el proceso de planificación estratégico de la organización, constituyendo una herramienta imprescindible para la mejora continua.

Según Membrado (2002), se entiende por «Modelo de Excelencia» a un conjunto de criterios agrupados en áreas o capítulos y que sirven como referencia para estructurar un plan que lleve a una empresa u organización o a una parte de la misma hacia la mejora continua de su gestión y sus resultados.

Ilustración 16. Circulo PDCA. Fuente: Elaboración propia.

Las organizaciones excelentes gestionan sus actividades y recursos mediante un conjunto de sistemas, procesos y datos que les permiten la toma eficaz de decisiones.

En este contexto, ser excelente implica, entre otros aspectos, llevar a cabo mediciones periódicas tanto de los enfoques aplicados como de su despliegue, con el fin de disponer

de datos objetivos que sustenten el avance de la organización hacia unos buenos resultados en relación con sus grupos de interés.

El desarrollo de sistemas de indicadores para estructurar la medición en las organizaciones es sin duda necesario para asegurar la calidad de la información empleada en dicha toma de decisiones.

El sistema de indicadores de una organización debe estar sustentado en un sistema de información robusto y fiable, actualizado y de calidad. El sistema de información debe permitir gestionar la información de los procesos y operaciones en tiempo real. La información es la materia prima principal para la toma de decisiones, y su adecuado tratamiento es clave para identificar las necesidades, cumplir los objetivos y analizar las relaciones causa-efecto.

Las organizaciones requieren de sistemas de medición de los resultados que permitan la autoevaluación continua de los criterios de los modelos de gestión. Las organizaciones implantan un sistema de información, con un sistema de indicadores integrado que a su vez, es la base de su sistema de gestión.

3 Los Sistemas de Información

La toma de contacto con los sistemas de información para la mayoría de las organizaciones se justificaba con el objetivo de mejorar su eficiencia interna: se implementaban sistemas de medición, se automatizaban las operaciones y, en esencia, la eficiencia significaba un intento de reducir puestos de trabajo y niveles jerárquicos. La evolución natural de los sistemas de negocios, del gobierno y personales, ha impulsado el uso generalizado de las redes de comunicación de datos para interconectar estos diversos sistemas (Fitz, 1992, p.45).

La rápida evolución de la tecnología de la información disponible en el mercado, y la complejidad del entorno en el que se mueven las empresas, han motivado la evolución de los Sistemas de información en las organizaciones, desde un planteamiento operativo hasta un planteamiento estratégico.

A medida que van creciendo las necesidades de información para la dirección, se comienza a realizar un esfuerzo por conocer las relaciones tanto verticales como horizontales de la información; empieza a plantearse la posibilidad de su integración: es la etapa en que se produce el desarrollo de los primeros sistemas integrados para el tratamiento de la información.

El desarrollo de los sistemas de información en las organizaciones, empezó en la década de los setenta como un instrumento al servicio de la función meramente operativa (sistemas de procesamiento de transacciones). Las nuevas tendencias, en virtud de los rápidos adelantos tecnológicos, tratan de encauzarlos hacia las funciones tácticas y estratégicas, pero sin

perturbar el rendimiento de su primera aplicación: sistemas de información de gestión o administrativos, sistemas de soporte a la decisión, sistemas de información para los ejecutivos.

En el periodo 1950-1960 se consideraba la información como algo malo, pero necesario, un requisito burocrático, y los sistemas de información se orientaban a una aceleración del proceso de datos. En 1950 existían sistemas de lote con archivos discretos y los usuarios transferían sus documentos en papel a la computadora para su procesamiento. La comunicación de datos en esta época implicaba que el usuario llevara físicamente documentos en papel. Durante los años sesenta se agregaron los circuitos de comunicación (líneas telefónicas) y se dotó a los usuarios con terminales en línea. En los últimos años de dicha década y principios de los setenta se empezaron a desarrollar los sistemas en línea de tiempo real, con lo cual los usuarios cambiaron los procesamientos de lote a un procesamiento por transacción individual. En este período se hizo necesaria la comunicación de datos.

En la década siguiente, 1960-1970, la información pasa a ser considerada como un soporte a necesidades de control de gestión y nacen los MIS (*Management Information Systems* o Sistema de Información para la Gestión). Se agregaron los sistemas de manejo de base de datos, los cuales reemplazaron a los antiguos archivos discretos.

Más adelante, en 1970-1980, se busca un soporte adecuado a las necesidades de control de los directivos y aparecen los DSS (*Decision Support Systems* o Sistema de Soporte a la Toma de Decisión) y los ESS (*Expert Support Systems* o Sistemas Expertos de Soporte a la Toma de Decisión). Por último, en la etapa que va de 1985 a 2000 la información es vista como un recurso estratégico y se plantean los Sistemas de Información como armas básicas para mejorar la posición competitiva de las organizaciones.

Uno de los aspectos más importantes del trabajo diario de la empresa es la toma de decisiones; por esta razón es necesario que los dispositivos de proceso de datos sean más valorados y que les sea encomendada la tarea de seleccionar, filtrar y presentar la información al directivo, para que éste cuente con el apoyo necesario. Los sistemas de información han variado de manera decisiva la forma de hacer negocios, precisamente por su capacidad para evitar la sobreabundancia de datos, lo cual puede bloquear o retardar la toma de decisiones. El entorno empresarial exige reacciones rápidas y oportunas. El entorno actual de las organizaciones obliga a un estilo de gestión fundamentalmente basado en la información, debido a cambios sustanciales que llevan a un nuevo contexto.

La creciente complejidad de las tareas de gestión, originada por un entorno cambiante que evoluciona muy rápidamente, ha incrementado el protagonismo y el interés por los sistemas de información de las empresas. Estos sistemas deben ser mecanismos rápidos, con una gran capacidad de adaptación frente a los cambios del entorno y de la propia organización, exhaustivos en su análisis, que aseguren la eficacia del resultado a la hora de alcanzar los objetivos marcados y eficientes respecto al uso de recursos, sobre todo de la información.

El proceso de toma de decisiones es una de las actividades que se realizan con mayor frecuencia en el mundo de las organizaciones. Se presenta en todos los niveles de la organización, desde asistentes o auxiliares, hasta los directores generales de las empresas. Según el nivel en cual se tome una decisión, así será el impacto de ésta (Cohen, 1999, p.73).

La información circula por toda la organización como si fuera un fluido, por causas formales e informales y en sentido horizontal y vertical. El Sistema de Información es la estructura que debe administrar dichos flujos de información con la máxima eficacia en su capacitación, proceso, almacenamiento y distribución. El Sistema de información resulta básico para apoyar el desarrollo de cualquier actividad operativa de la empresa, así como para la coordinación de los procesos entre organizaciones. La importancia del mismo en la actualidad obliga a una implicación en el tema por parte de los directivos, existiendo progresivamente una mayor interdependencia de estrategias, procesos y políticas con el software y las tecnologías de la información.

El sistema de información es un elemento capaz de integrar no sólo actividades dentro de la organización, sino entre organizaciones, considerando cada vez más un entorno competitivo en el ámbito global, en el cual las tecnologías de la información contribuyen a romper las barreras geográficas, idiomáticas y culturales. Lo importante es reunir las informaciones de la organización para poder así proporcionar suficiente información global para dar soporte a los objetivos.

La implantación de sistemas de información adecuados requiere una sintonía de los aspectos tecnológicos con los aspectos organizativos y humanos, ya que existe una interacción entre todos ellos. Los sistemas de información modifican sustancialmente la naturaleza de los trabajos, permiten combinar funciones anteriormente agregadas y, sobre todo, suponen una nueva forma de hacer las cosas. Integran y hacen compatibles los conocimientos en la organización y entre organizaciones, facilitan las operaciones, el control y la planificación, modificando las relaciones entre una organización y su entorno, permitiendo mejorar el control, aportando información para la toma de decisiones, reducciones de coste, plazo y mejora de la calidad.

3.1 Concepto de Sistema de Información para la Gestión

Los Sistemas de información para la gestión o *Management Information Systems* (MIS) son el conjunto de medios humanos y materiales encargados del tratamiento de la información empresarial, mediante la entrada de datos, su procesamiento y almacenamiento y su posterior salida, presentación y, a veces, su explotación, Manfred (1995, p.23).

Los MIS proporcionan información a los directivos mediante informes o acceso directo a consultas informáticas. Suelen estar orientados a la información interna más que a la externa y, generalmente, están alimentados por los sistemas operacionales; muchos sistemas informáticos

cubren ambos aspectos en el mismo software. Se trata normalmente de salidas de información periódica que aportan visión de conjunto para el seguimiento y control de un aspecto específico. Efectivamente, son sistemas estructurados con salidas predeterminadas y poca capacidad analítica o de simulación; los requisitos de información son conocidos y estables y suelen requerir largos periodos de diseño y ajuste hasta que son operativos y de aplicación a los usuarios Morete y Pérez-Alcobas (1998, p.12).

El objetivo de los Sistemas de información tiene una triple dimensión. En primer lugar, deben facilitar, simplificar o realizar procesos de medición que tradicionalmente se realizaban de forma manual, evitando errores y mejorando la velocidad de medida. En segundo lugar, deben proporcionar informaciones y datos que faciliten la toma de decisiones, dejando libertad plena a los usuarios. En último lugar, deben permitir interacciones con el usuario en un plano más profundo que el mero apoyo al proceso de toma de decisiones. Para alcanzar sus objetivos, el sistema de información debe ser capaz de: recibir datos con el menor coste posible y sin errores, evaluar la calidad e importancia relativa de los datos de entrada, procesar la información sin alterarla y transformarla para que sea útil al usuario (errores de redondeo inapreciables conducen muchas veces a resultados absurdos e inútiles), almacenar los datos de forma que estén accesibles cuando se requiera y ofrecer la información de acuerdo con las necesidades del usuario Monforte (1995, p.19-20).

El sistema de información tiene como objetivo proveer a una organización de la información necesaria para apoyar las operaciones, la administración y la toma de decisiones.

4 El Business Intelligence como sistema de información

Para gestionar el conocimiento explícito en la organización se han desarrollado diferentes tecnologías de soporte. La tecnología que hoy en día mejor contribuye a la eficiente gestión del conocimiento la conocemos como Business Intelligence (o inteligencia de negocio).

El cambiante entorno económico y la problemática de sistemas descrita anteriormente impulsaron el surgimiento del llamado Business Intelligence, el cual es un concepto que trata de englobar todos los Sistemas de Información de una organización para obtener de ellos no solo información o conocimiento, sino una verdadera inteligencia que le confiera a la organización una ventaja competitiva por encima de sus competidores.

Hackney (2001) nos dice que el Business Intelligence se compone de todas las actividades relacionadas con la organización y entrega de información, así como el análisis del negocio. Esto incluye Minería de Datos, Administración del Conocimiento, Aplicaciones Analíticas, Sistemas de Reportes y principalmente Data Warehousing. Michel (2000) afirma que el conjunto de tecnologías que usan Data Warehousing y OLAP (On-line Analitic Procesing, procesamiento analítico en línea), combinado con herramientas de reporte, son referidas como

Business Intelligence, porque ayudan a las compañías a ganar inteligencia en operaciones y desempeño.

Buksard, Mollot y Richards (2000) comentan en su artículo que la necesidad de nuevas herramientas de acceso y reporte de información, para diversos tipos de usuarios, ha impulsado la creación de nuevas herramientas, colectivamente conocidas como Business Intelligence. Business Intelligence no es una sola tecnología o aplicación. No es una "cosa", sino que se trata de un "suite" de productos que trabajan de manera conjunta para proveer datos, información y reportes analíticos (Informes) que satisfagan las necesidades de una gran variedad de usuarios finales.

Cano (1999) propone que Business Intelligence es la habilidad de consolidar información y analizarla con la suficiente velocidad y precisión para descubrir ventajas y tomar mejores decisiones de negocios. Definición compatible con la necesidad actual de los negocios que ante la presión de ser cada día más competitivos, para mantenerse tienen la doble tarea no sólo de permanecer sino de ser lucrativos.

Howard Dresner, citado en Hilson (2001), vicepresidente y director de investigaciones del Grupo Gartner, agrega: "Business Intelligence es simplemente la habilidad de los usuarios finales para acceder y analizar tipos cuantitativos de información y ser capaz de actuar en consecuencia".

Business Intelligence ha tomado la delantera en los últimos años, los proveedores de soluciones cuentan con tecnología más amigable y presentan datos más fáciles de analizar, apostando hacia la adopción masiva de dichos sistemas por parte de las organizaciones (Sullivan, 2001).

Las compañías actualmente usan una amplia gama de tecnologías y productos para saber qué es lo que está pasando en la organización. Las herramientas más comunes (simple consulta y reporte de datos, procesamiento analítico en línea, análisis estadístico, predicciones y minería de datos) pueden ser usadas de una gran variedad de formas. El objetivo de todo esto es transformar las montañas de datos en información útil para la empresa (McGeever, 2000).

El radical crecimiento de nuevas formas de "inteligencia" generada por computadora es una de las dos revoluciones en materia de tecnologías de información que se están dando en la actualidad. La segunda es el Internet, del cual ya todo mundo conoce (Martín, 2001). Las dos revoluciones van a soportar y amplificar una a la otra. El conjunto de ambas resultará en una radical reinvención de los negocios.

En los inicios de la década de los 90, Giner de la Fuente, (2004:129) tuvo lugar un progresivo proceso de convergencia entre el conjunto de empresas que configuraban el sector de informática y de tecnologías de la información. Este proceso de convergencia propuso una nueva filosofía de organizar y de explotar los datos con el objeto de obtener información orientada al análisis y la gestión de las organizaciones.

La nueva propuesta filosófica se vio impulsada y soportada por los fabricantes de:

- bases de datos, que construyeron potentes bases de datos relacionales con capacidad de almacenar y mantener un elevado número de datos, en conectividad constante con los sistemas transaccionales.
- productos de software que crearon herramientas que permitían transformar los datos en información con un uso poco intensivo de programación.

Esta nueva filosofía se ha llamado Business Intelligence. Podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten:

- reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la organización) en información estructuradapara su análisis y conversión en conocimiento, dando así apoyo a la toma de decisiones sobre el negocio.

Podemos definir dos partes fundamentales en un sistema de Business Intelligence:

Data Warehouse: contiene los datos almacenados en la forma adecuada para ser consultados y obtener respuestas ágiles. Generalmente, las aplicaciones transaccionales de la organización (programa de nóminas, de almacén, de ventas, de márqueting, CRM – consumer relationship management, ERP -Enterprise Resource Planning; Planificación de Recursos Empresariales-, etc.) están pensadas para almacenar datos, pero no para que estos datos estén disponibles para su análisis.

Herramientas de análisis de información: permiten interrogar, confeccionar y presentar la información elaborada en base a los datos que están en el Data Warehouse. En la actualidad se pueden clasificar en herramientas de:

a. Reporting, que permiten obtener los clásicos 'reports' o Informes, en forma de listados con distintas agrupaciones y componentes visuales.
b. Dashboard y Scorecard, que permiten diseñar vistas interactivas con diferentes componentes visuales. Son frecuentemente utilizadas para visualizar los Cuadros de Mando.
c. OLAP, que partiendo de unas dimensiones prefijadas, permiten combinarlas de muy diferentes maneras para obtener vistas flexibles. Es parecido a las conocidas tablas dinámicas de Excel y cumplen fundamentalmente con la función de análisis interactivo.
d. Data mining, que ayuda a las organizaciones a evidenciar información que no es perceptible de forma directa, como por ejemplo patrones de comportamiento, relaciones, asociaciones, etc… que nos permitan tomar mejores decisiones. También, a través del análisis del pasado, y aplicando algoritmos, se construyen predicciones que nos permiten mejorar nuestra eficiencia y conseguir así la estrategia de la organización.

Los sistemas BI están formados por un conjunto de tecnologías y disciplinas, que se pueden agrupar en el siguiente gráfico. Este diagrama fue creado por Forrester Research sobre las diferentes tecnologías y aspectos que forman parte de un sistema BI:

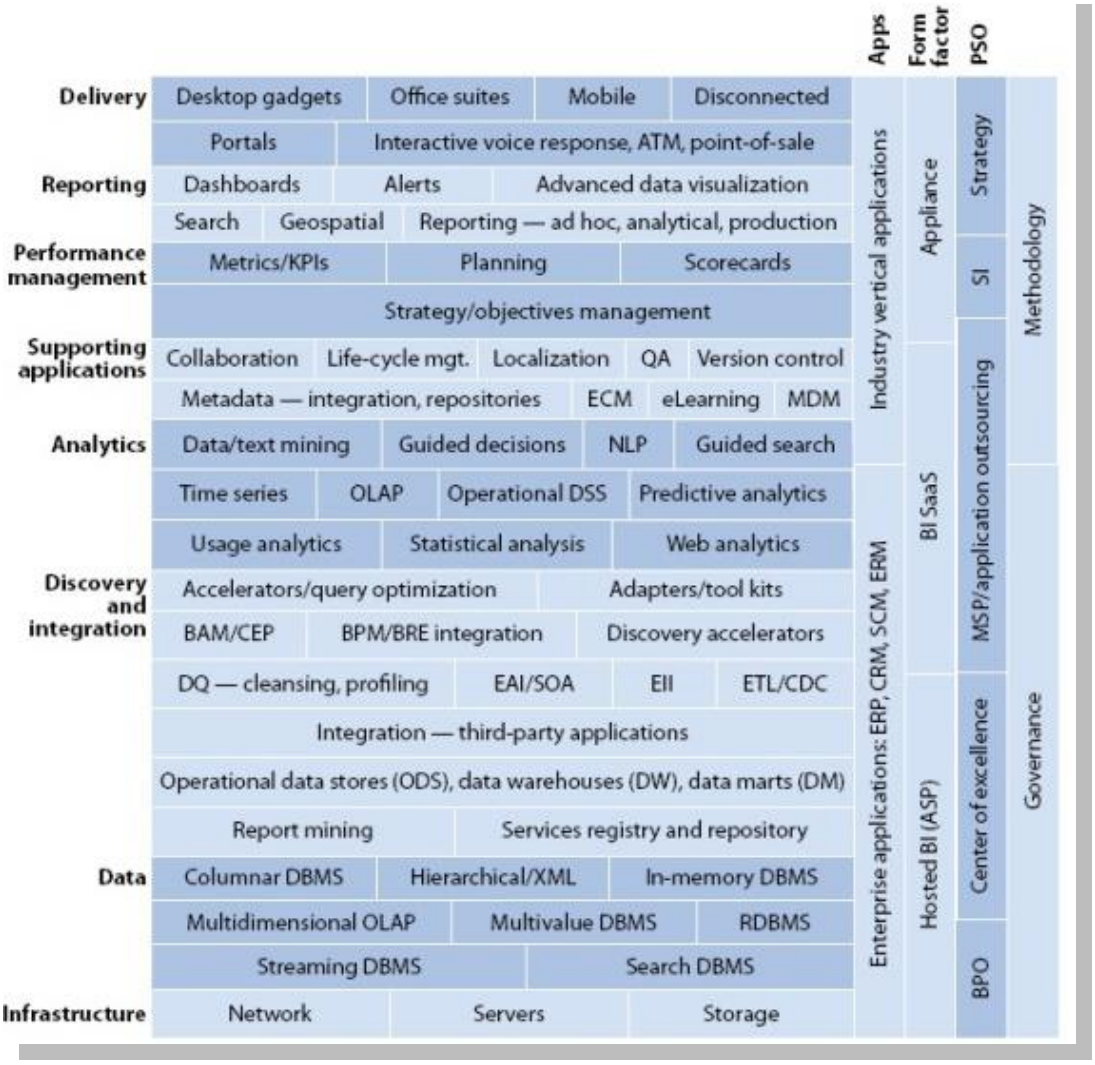


Ilustración 17. Componentes de un sistema de BI. Fuente: Forrester Research Inc.

La gestión del conocimiento eficiente requiere una adecuada estructura de gestión de la información ágil y precisa. Ágil significa que debe responder en tiempo adecuado a las necesidades de información. Un buen sistema de información responde en menos de 5 minutos a las solicitudes de información.

El BI está situado entre dos capas:

- Las bases de datos dispersas entre las diferentes aplicaciones operacionales de la organización y la información en el exterior.
- Los usuarios que desean obtener información de la operación de la organización.

A continuación se detalla la información referente a los dos componentes principales de un sistema de Business Intelligence, el Data Warehouse y las herramientas de análisis de información.

4.1 Data Warehouse

El Data Warehouse contiene aquellos datos extraídos de las aplicaciones operacionales y del exterior que, una vez tratados, se almacenan en una base de datos independiente con la estructura y el formato adecuado para ser analizados. La herramienta que recoge, trata y almacena los datos en el Data Warehouse se denomina ETL (del inglés Extract, Transform and Load).

Es conveniente que el Data Warehouse sea 'altamente' independiente de las herramientas utilizadas para el análisis de información. Por este motivo, resulta de utilidad crear una base de datos especializada que residirá en una máquina dedicada. De esta manera el Data Warehouse se circunscribe a dos elementos: la herramienta ETL y el Data Warehouse propiamente (base de datos).

La base de datos del Data Warehouse se estructura en dos tipos de tablas básicas: las tablas de hechos (tablas f: facttables) y las tablas de dimensiones (tablas d: dimensión tables). Las tablas de hechos están formadas por registros de hechos y se recomienda que contengan datos de movimientos y de los estados de los movimientos, así como las métricas o medidas. Un registro de movimiento se caracteriza por tener campos de tiempo (fecha/hora), campos de dimensiones y campos de indicadores o métricas. Los campos de dimensiones estarán relacionados con las tablas de dimensiones que contienen datos que caracterizan la dimensión.

Así la tabla de la dimensión producto, contiene los datos de un producto (el precio, el tamaño, el color, el código, el nombre comercial, el tipo de embalaje, etc.). La tabla de hechos ventas, contiene una fila para cada venta de productos, en la que se registra la fecha, el código del producto vendido, y las métricas asociadas, como podría ser el precio de venta.

4.2 Herramientas de análisis e Informes

Contiene una serie de herramientas cuya finalidad es realizar análisis y construir y presentar informes.

Es conveniente disponer de una variada gama de herramientas de análisis. Además de las herramientas univariables, recomendamos la inclusión de herramientas que permitan análisis e informes complejos, con varias variables (bivariable y multivariable). Las herramientas de Data mining suelen incluir herramientas de análisis multivariable especializadas, y facilitar el análisis visual de los datos. No debemos olvidar que hay mucha información que no es evidente; hay muchas relaciones entre las variables que no son obvias a simple vista, y que este tipo de análisis puede explicitar.

También es conveniente disponer de una potente gama de herramientas de generación de informes que cubran todas las necesidades de información de la organización. Cada una de

estas herramientas debe ofrecer formatos de salida normalizados. Es frecuente encontrar formatos tipo pdf, html, Word, Excel, PowerPoint, etc.

4.3 La creación del Data Warehouse y las ETL

Una buena parte de los sistemas de información existentes están diseñados desde la óptica de almacenar los datos en bases de datos operacionales (datos para ser guardados), en lugar de hacerlo desde la óptica de extraerlos para ser analizados.

El Data Warehouse es "Es una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis…". Ralph Kimball

Para disponer de un sistema de medición ágil es conveniente crear un 'Data Warehouse' que contendrá toda aquella información que necesitemos para su posterior análisis y que deberá extraerse de las bases de datos operacionales y almacenarse en forma y precisión adecuada con fines de análisis de gestión. Se trata de crear un espacio (una base de datos) 'nuevo' donde almacenaremos la información estructurada de tal manera que permita una lectura óptima.

Un Data Warehouse se caracteriza (Bill Inmom):

- Integrado: los datos almacenados en el *Data Warehouse* deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- Temático: sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del *Data Warehouse*.
- Histórico: el tiempo es parte implícita de la información contenida en un *Data Warehouse*. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el *Data Warehouse* sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el *Data Warehouse* se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- No volátil: el almacén de información de un *Data Warehouse* existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del *Data Warehouse* la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

4.4 El diseño del Data Warehouse, punto clave

Existen dos figuras muy importantes en el desarrollo conceptual del datawarehousing: Ralph Kimball y Bill Inmon. A pesar de que las teorías de ambos persiguen los mismos objetivos, la forma de conseguirlos difiere bastante, lo que ha generado dos filosofías a la hora de diseñar la estrategia de datos.

La visión de Kimball se basa en que son los procesos de negocio los que deben de marcar la forma en la que diseñamos el Data Warehouse. Admite un punto de partida en la que ya existen datos más o menos organizados en Datamarts, y que estos son la base del futuro Data Warehouse.

Para Kimball lo más importante es que el cálculo de los datos que servirá para la toma de decisiones sea rápido, por lo que estructura los datos del Data Warehouse siguiendo patrones dimensionales. Esto suele mejorar mucho su rendimiento a la hora de realizar consultas y además organiza los datos de una forma más intuitiva y natural para los usuarios.

Por todo esto se considera a la arquitectura de Kimball como una aproximación "bottom-up" del problema, ya que se parte de los datos y procesos existentes y se modela el Data Warehouse para que se adapte a ellos, tomando como premisas la eficiencia en tiempo y la representación natural de datos a costa de la normalización.

Por otra parte, su colega Inmon (quien es considerado el padre del concepto de Data Warehouse) nos plantea una estrategia "top-down" del problema. Lo primero que haremos a la hora de desarrollar del Data Warehouse será establecer la estructura de datos perfectamente normalizada y limpia. Los datos que se insertarán en esta estructura generalmente procederán de un "área de carga" en la que los datos son depurados antes de pasar a la estructura normalizada del Data Warehouse.

A partir de esa estructura, se pueden establecer una serie de Datamarts que agrupen de una forma más lógica (y si se quiere multidimensional) la información del Data Warehouse principal. Como se puede observar, esta forma de trabajar es mucho más organizada pero menos flexible que la anterior, ya que aquí es la estructura y la normalización de los datos lo que marcará la pauta a la hora de trabajar en vez de que sean los procesos existentes en la organización.

¿Cuál es mejor? No se considera una mejor que otra… simplemente cada una tiene un punto diferente sobre lo que debe primar a la hora de diseñar un Data Warehouse. Kimball representa la relación de los datos con el usuario final, la flexibilidad y la rápida explotación de la información. Por otra parte Inmon representa la pulcritud en el diseño y el respeto por una serie de normas que garanticen la exactitud de los datos, su integración y su coherencia.

Con el tiempo, los profesionales del datawarehousing han aprendido a fusionar ambas tendencias en modelos híbridos que recogen lo mejor de ambas soluciones. Estructuras que

cierran bien el ciclo de vida de la información mediante procesos ETL y a su vez modelos multidimensionales para representar la información de una forma que facilite su explotación.

Para comprender los datos multidimensionales, primero piense en una hoja de cálculo que muestra cómo se venden distintos productos en mercados diferentes. Esta hoja de cálculo muestra las ventas de cada producto por zona. Los datos tienen dos dimensiones: Producto (en las filas) y Zona (en las columnas).

Una hoja de cálculo como ésta tiene un uso limitado ya que sólo muestra el rendimiento de las ventas en un momento dado. Para realizar el seguimiento del rendimiento de los distintos períodos de tiempo, los analistas necesitarían apilar varias hojas de cálculo, una por cada período de tiempo.

Juntas, estas hojas de cálculo mostrarían una tercera dimensión, Tiempo, que agrega a las dos anteriores, Producto y Zona.

Otra forma de representar estas celdas de datos apilados es mediante un cubo de datos.

El cubo de datos permite que los analistas dividan los datos de distintas formas para obtener respuestas a diferentes preguntas, como:

- ¿Nuestros productos se venden por regiones (vista *Producto* por *Zona*, igual que antes)?
- ¿Cómo se venden nuestros productos en las distintas épocas del año en cada mercado (*Producto* por *Zona* por *Tiempo*)?
- ¿Cómo se venden nuestros productos en las distintas épocas del año (vista *Producto* por *Tiempo*)?

Este ejemplo sólo incluye tres dimensiones.

Una cuarta dimensión podría permitir que los analistas dividieran los datos por *Cliente*, por ejemplo.

Este planteamiento es más difícil de representar que tres dimensiones, pero funciona de la misma forma; con otra dimensión, existen más formas posibles de dividir el cubo, para proporcionar respuestas a preguntas del tipo: ¿Cuál es el cliente de determinado país que compra más, o menos, un producto determinado en Navidades?

4.4.1 Tablas de dimensión

Las tablas de dimensión contienen los atributos (normalmente textuales) que nos permiten filtrar y agrupar los datos.

Una dimensión es un conjunto datos que nos permiten filtrar, agrupar o seccionar la información. También se denomina "atributo", "característica", "propiedad", "campo", etc.

Es frecuente utilizar el concepto de dimensión, también, como una jerarquía. Es conveniente que cada tabla contenga los datos de cada jerarquía. Una jerarquía podría ser Año, Trimestre, Mes, Semana, Día, por ejemplo.

4.4.2 Tablas de hechos

Las tablas de hecho contienen las medidas numéricas provenientes de los orígenes transaccionales.

Denominamos "hechos" a las medidas de la organización. Por ejemplo, son "hechos" las ventas, los pedidos, los envíos, las reclamaciones, las compras, etc. Es decir, son todas aquellas medidas numéricas que incluiremos en nuestro sistema de datos. Técnicamente, una tabla de hecho es la tabla central de un modelo en estrella.

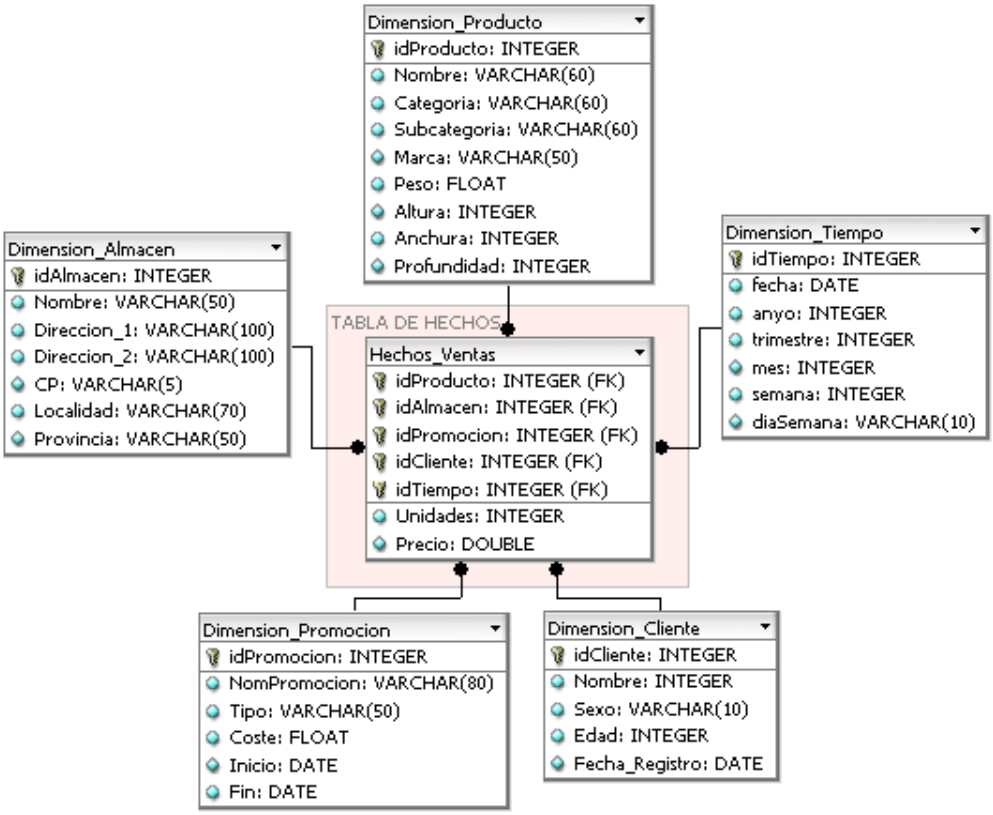


Ilustración 18. Esquema de datos en estrella. Fuente: De Jesuja - Trabajo propio, GFDL, https://commons.wikimedia.org/w/index.php?curid=3886973 (2016)

4.4.3 Data Marts

Un Data Warehouse está compuesto por distintos Datamarts que contienen información homogénea y relacionada con aspectos concretos de la organización (por ejemplo: Data Mart de personas, de clientes, de reclamaciones, de productos o servicios, etc.).

Un Data Mart es un subconjunto de datos estructurados de un tipo específico que tiene el propósito de facilitar la extracción de información para ayudar a la toma de decisiones en las organizaciones.

Una forma habitual de definir los Data Marts es creando uno para cada área funcional; por ejemplo, el Data Mart de Gerencia, Marqueting. Operaciones, Recursos Humanos y Finanzas.

4.4.4 Esquemas

A su vez, cada Data Mart está compuesto por uno o varios 'Esquemas'. Existen fundamentalmente dos tipos de Esquemas: en estrella y en copo de nieve. Mostramos un Esquema de 'Ventas' tipo 'estrella'.

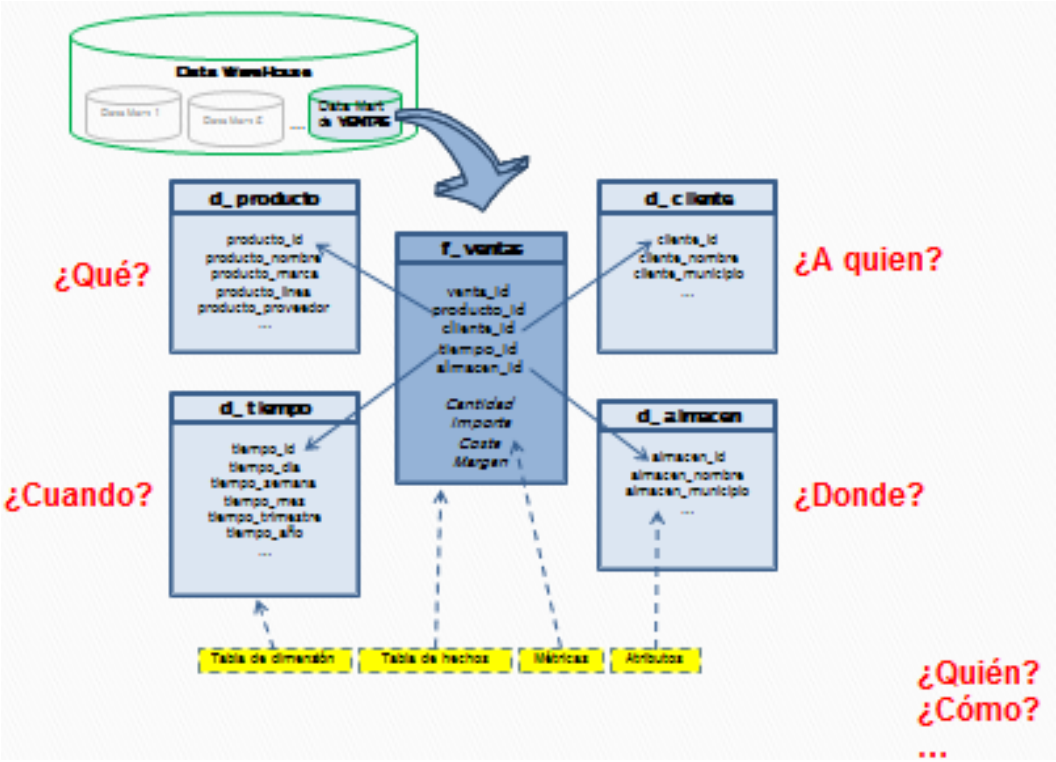


Ilustración 19. Esquema Data Mart ventas. Fuente: Brull Alabart (2009)

El esquema en forma de estrella solo tiene un nivel de relación, el esquema en copo de nieve puede tener más de un nivel, como se muestra en el siguiente gráfico:

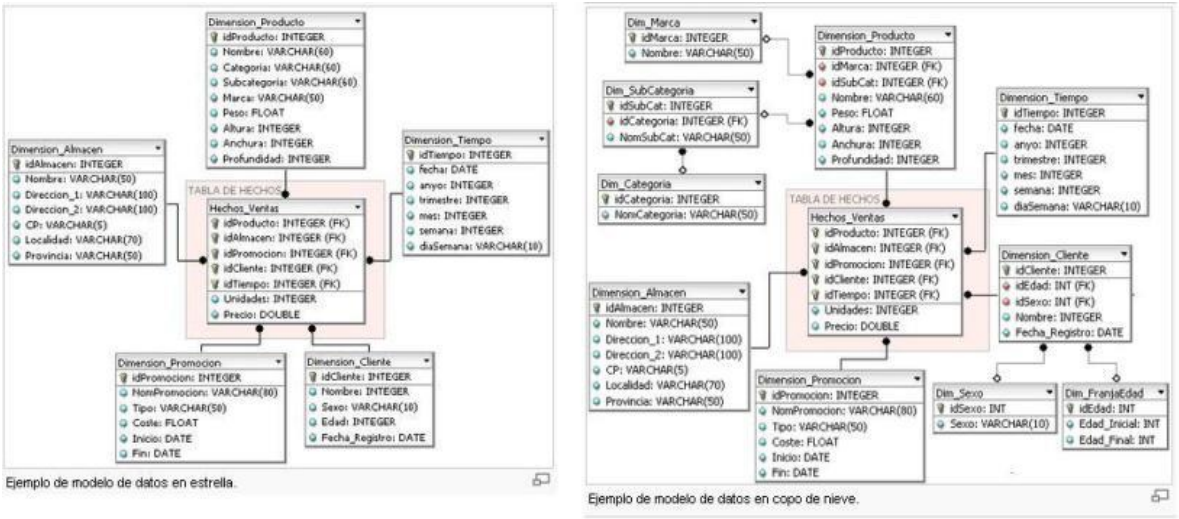


Ilustración 20. Esquema en estrella o en copo de nieve. Fuente: Advanced Data WareHouse Design, Elzbieta Malinowski

4.5 Las ETL (Extract, Transform and Load)

El Data Warehouse no deja de ser una estructura de campos, de una tipología determinada, ordenadas en tablas de hechos y tablas de dimensión. Para llenar de datos el Data Warehouse–además de poder rellenarlos manualmente-, se suelen utilizar herramientas de Business Intelligence específicas, concretamente las utilidades de extracción, transformación y carga (ETL), proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, Data Mart, o Data Warehouse para analizar. Las utilidades ETL facilitan la extracción de datos ubicados en diferentes orígenes y su almacenamiento en el espacio Data Warehouse.

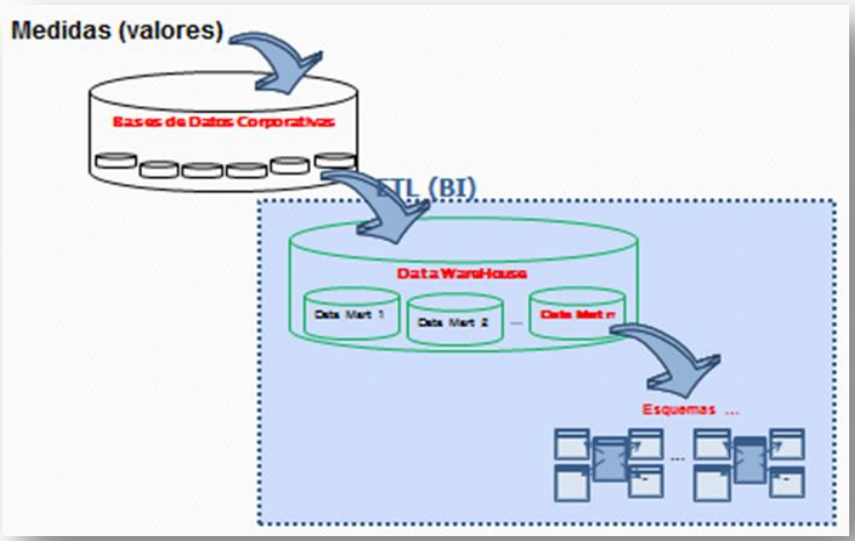


Ilustración 21. Esquemas de un Data Mart. Fuente: Elaboración propia (2016)

Esta etapa es típicamente la más subestimada de las tareas en un proyecto de Data Warehouse.

Para poblar el Data Warehouse se deben mover bloques de datos, muchas veces desde diferentes sistemas operativos, estructuras de archivos y bases de datos diferentes, mediante procesos programados en el tiempo, que se ejecutan frecuentemente fuera del horario de trabajo para no consumir tiempo de procesamiento de la organización, entorpeciendo la operatoria.

Los subsistemas para poblar el Data Warehouse se pueden construir utilizando herramientas y productos disponibles en el mercado, programas y procesos codificados desde cero, o combinaciones de estos elementos.

Al construir los sistemas para poblar el Data Warehouse, se debe considerar la posibilidad de que estos permitan regular el crecimiento evolutivo del Data Warehouse, brindando escalabilidad y soporte para grandes cantidades de datos y consultas complejas. Se pueden encontrar dificultades adicionales dependiendo de las fuentes de datos que se tengan disponibles, que implican el uso de diferentes herramientas y tecnologías para acceder a cada uno de ellos.

4.5.1 Extracción (Extraction)

Se definen como procesos de extracción a aquellos requeridos para obtener los datos que permitirán efectuar la carga del Data Mart definido. El propósito principal de la fase de extracción es capturar y copiar los datos requeridos de uno o más sistemas operacionales o fuentes de datos. Los datos que se extraen son colocados en un archivo intermedio con un formato definido, que luego será utilizado por la siguiente fase del proceso.

Los registros que sean rechazados en el proceso deben ser registrados en un archivo o log de rechazos para que puedan ser analizados posteriormente y así tener la posibilidad de cargarlos en el Data Warehouse correctamente. Además, esto permite descubrir los errores que han ocurrido en los procesos de creación de los datos operacionales. Ejemplos de estos errores son violaciones de integridad, claves duplicadas, formatos de datos incorrectos y datos inválidos como campos vacíos, fechas futuras e importes negativos cuando estos no correspondan. Otros ejemplos: números telefónicos inválidos, estándares no respetados, llamadas duplicadas y campos nulos.

Hay que tener en cuenta que después de la fase inicial de carga del Data Warehouse sólo es necesario cargar los datos nuevos o los que hayan sido modificados.

4.5.2 Transformación (Transformation)

Se definen como procesos de transformación los procesos para convertir o recodificar los datos de la fuente a fin de poder efectuar la carga efectiva del data mart.

Las funciones básicas a ser realizadas en esta fase consisten en leer los archivos intermedios generados por la fase de extracción, realizar las transformaciones necesarias, construir los registros en el formato del Data Warehouse y crear un archivo de salida conteniendo todos los registros nuevos a ser cargados en el Data Warehouse. La mayor parte del trabajo en esta fase involucra el efectuar las transformaciones necesarias. Estas transformaciones incluyen:

- Combinar campos múltiples de nombres y apellidos en un solo campo.
- Fusionar campos o datos homónimos.
- Separar un campo de fecha en campos de año, mes y día.
- Cambiar la representación de los datos, como TRUE (verdadero) a 1, y FALSE (falso) a 0, o códigos postales numéricos a alfanuméricos, respetando los estándares definidos.
- Cambiar un dato que tiene múltiples representaciones a una sola representación, como por ejemplo definir un formato común para números telefónicos, o establecer un término común para los nombres de los campos o los valores de los datos que sean sinónimos.

4.5.3 Carga (Loading)

Por otra parte, los procesos de carga de datos son los procesos requeridos para llenar el Data Warehouse.

El objetivo de esta fase consiste en tomar los registros formateados por la fase de transformación y cargarlos en el Data Warehouse, que es el contenedor para todos los datos informativos (actuales e históricos) requeridos por las operaciones del Data Warehouse. Generalmente los datos son insertados en el Data Warehouse, rara vez son actualizados o eliminados.

Todas estas tareas son altamente críticas pues tienen que ver con la materia prima del Data Warehouse: los datos. La desconfianza y pérdida de credibilidad del Data Warehouse serán resultados inmediatos e inevitables si el usuario se encuentra con información inconsistente. Es por ello que la calidad de los datos es un factor determinante en el éxito de un proyecto de data warehousing. Es en esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuente.

Como advierte Kimball (1998), el proceso de Data Staging es el iceberg de un proyecto de data warehousing. Son muchos los desafíos que deben enfrentarse para lograr datos de alta calidad de los sistemas fuentes. Como ya se ha comentado, en general es una de las etapas más subestimadas, que siempre termina tomando más tiempo del previsto. Ralph Kimball propone entonces un plan de 10 ítems que ayudarán a guiar esta etapa:

Planificar:

1. Crear un diagrama de flujo fuente-destino esquemático, de una página y de muy alto nivel.
2. Probar, elegir e implementar una herramienta de data staging.
3. Profundizar en detalle por tabla destino, gráficamente describir las reestructuraciones o transformaciones complejas. Gráficamente ilustrar la generación de las claves subrogadas. Desarrollo preliminar de la secuencialidad de los trabajos.

Carga de dimensiones:

4. Construir y probar la carga de una tabla dimensional estática. La principal meta de este paso es resolver los problemas de infraestructura que pudieran surgir (conectividad, transferencia, seguridad, etc.)
5. Construir y probar los procesos de actualización de una dimensión.
6. Construir y probar las cargas de las restantes dimensiones.

Tablas de hechos y automatizaciones:

7. Construir y probar la carga histórica de las tablas de hechos (carga masiva de datos). Incluyendo búsqueda y sustitución de claves.
8. Construir y probar los procesos de cargas incrementales.
9. Construir y probar la generación de agregaciones.
10. Diseñar, construir y probar la automatización de los procesos

5 Las soluciones actuales para el Business Intelligence

Gartner Inc. es una empresa consultora y de investigación de las tecnologías de la información con sede en Stamford, Connecticut, Estados Unidos. Gartner incluye entre sus clientes a algunas de las más grandes empresas, agencias de gobierno, empresas tecnológicas y agencias de inversión como BT, CV, The Wall Street Journal, etc. La empresa se concentra en la investigación, programas ejecutivos, consultas y eventos. Fue fundada en 1979; y en 2014 contaba con 5300 empleados, incluyendo a 1250 analistas y clientes en 85 países por todo el mundo. Gartner proporciona el análisis de investigación y el consejo para profesionales de las TIC (tecnologías de la información y la comunicación), empresas de tecnología y la comunidad de la inversión en varios formatos: reuniones informativas, servicios de pares en red (peer networking service) y programas de socios diseñados explícitamente para CEOs y otros directores ejecutivos. Gartner utiliza para presentar sus análisis los conocidos como Cuadrantes Mágicos.

Según el Informe Gartner Magic Quadrant for Business Intelligence and Analytics Platforms 2016 la principal prioridad de los responsables de tecnología (CIO, Chief of Information Office) este año será precisamente ésta, encontrar los mejores sistemas de Business Intelligence y Business Analytics.

2016 CIO TECHNOLOGY PRIORITIES		
◀▶	1	BI/Analytics
▲	2	Cloud
▲	3	Mobile
▲	4	Digitalization/Digital Marketing
▼	5	Infrastructure & Data Centre
▼	6	ERP
◀▶	7	Security
▲	8	Industry Specific Applications
◀▶	9	CRM
▼	10	Networking/Voice/Data Communications

Ilustración 22. Prioridades 2016. Fuente: Gartner (Feb 2016)

El Business Intelligence (y con él la Analítica Digital) continúa siendo una de las áreas de más rápido crecimiento en las compañías de hoy. El incremento exponencial del volumen de datos que se generan a diario hace imprescindible el uso de sistemas que sean capaces de analizarlos, lo cual implica necesariamente la incorporación de herramientas de Business Analytics o Smart Data.

Se entiende que la diferencia entre Business Intelligence y Business Analytics, es que este último hace un uso extensivo de los datos, análisis estadístico, modelos explicativos y predictivos para impulsar la toma de decisiones. En cambio, Business Intelligence está más ligada a la generación de los datos e información para apoyar el mismo proceso.

Una definición de Business Analytics podría ser "la práctica iterativa, la exploración metódica de datos en una empresa u organización con énfasis en el análisis estadístico…para la toma de decisiones basada en los datos". Y luego se podría definir el Business Intelligence como "las aplicaciones y tecnologías para la recolección, almacenamiento, análisis y acceso a los datos para ayudar a una empresa u organización a tomar mejores decisiones de negocios".

La integración de datos se asume ahora como una prioridad y todo apunta a que en los próximos dos años muchos perfiles profesionales dentro de la organización tendrán acceso a diferentes herramientas donde podrán "autosatisfacer" sus distintas necesidades de datos.

Si buscamos la evolución en Google Trends del término Business Intelligence, vemos que las búsquedas han ido disminuyendo desde el año 2004:

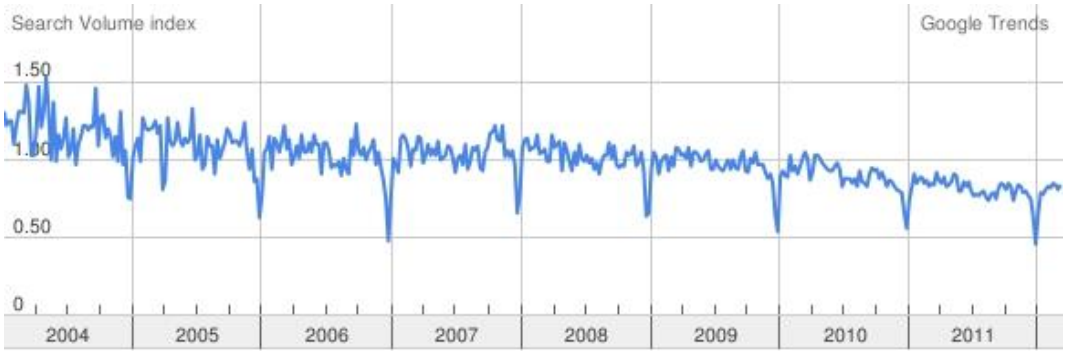


Ilustración 23. Búsqueda en Google Trends Business Intelligence

Por contra, la misma gráfica para la búsqueda Business Analytics, muestra una tendencia incremental desde el año 2009:

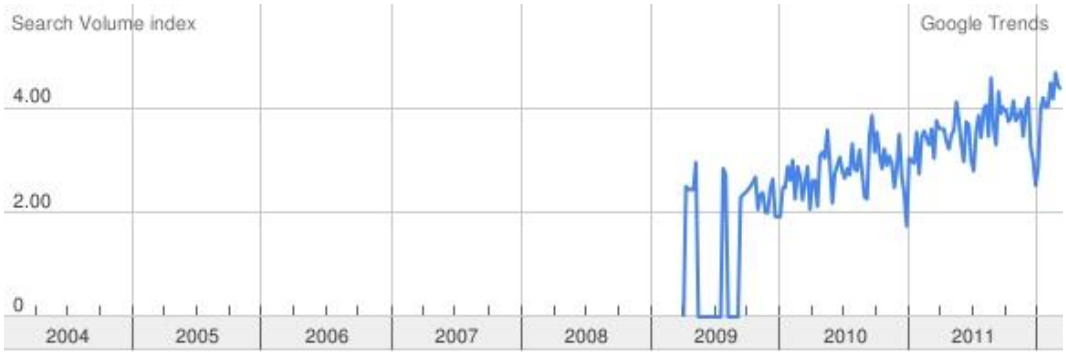


Ilustración 24. Búsqueda Google Trends Business Analytics

La evolución y sofisticación de las capacidades de preparación de datos de autoservicio y de descubrimiento de datos disponibles en el mercado ha desplazado el foco de los compradores en el mercado de plataformas de BI y análisis, hacia herramientas fáciles de usar que no requieren una participación destacada de los responsables de informática para predefinir los modelos de datos iniciales como requisito previo para el análisis.

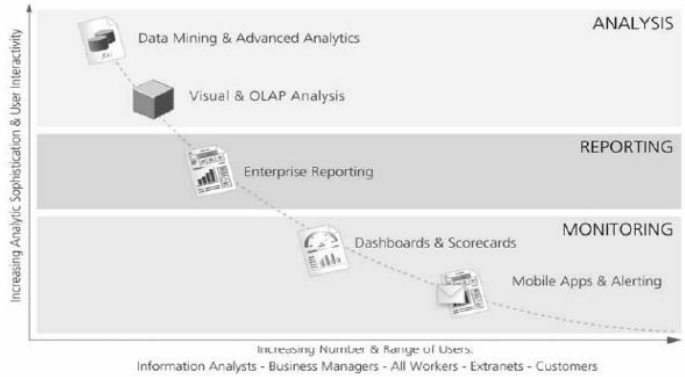


Ilustración 25. Distribución de usuarios en función de la sofisticación de la solución de Business Intelligence

En este gráfico se muestra la distribución de usuarios en función de la sofisticación de la solución de Business Intelligence o Business Analytics.

A continuación se presenta el Cuadrante mágico para Business Intelligence publicado en Febrero de este mismo año, en el que se puede ver la posición de cada proveedor de soluciones de Business Intelligence y Analytics en el mercado actual:

Ilustración 26. Cuadrante de Gartner para Business Intelligence y Analytics. Fuente: Gartner (Feb 2016)

El cuadrante de Gartner muestra todas las soluciones disponibles para uns sistema de Business Intelligence o para un parte de él (ninche players), y contemapla tanto las soluciones de licenciamiento (de pago) como las soluciones open source (de código abierto, generlmente gratuitas).

Según la visión de Gartner, el Business Intelligence y Business Analytics deberían conformar una plataforma única, la cual debería ofrecer hasta 15 principales funcionalidades, agrupadas en tres grandes categorías:

INTEGRACIÓN

- Infraestructura BI. Debe ser única y común para gestionar todos los recursos y funcionalidades de la plataforma.
- Gestión de metadatos. Deben ser reaprovechables y base para definir nuevos objetos. La plataforma de brindar las herramientas para gestionarlos.
- Herramientas de desarrollo. Debe brindar los recursos necesarios para diseñar aplicaciones analíticas que puedan ser integradas a los procesos de negocio u otras aplicaciones.
- Colaboración. Debe facilitar el intercambio de idea o diálogos entre usuarios que analizan la información a través del uso de anotaciones, chats o foros.

DISTRIBUCIÓN DE LA INFORMACIÓN (INFORMATION DELIVERY)

- Informes (Reporting). Facilidades para crear informes con formato e interactivos, con o sin parámetros y posibilidades de programar su distribución
- Tableros de datos (Dashboards). Informes para Web o dispositivos móviles con pantallas táctiles que indiquen el estado de una métrica de rendimiento en comparación con un objetivo o valor objetivo.
- Consultas personalizadas (Ad hoc query). Permitir a los usuarios elaborar sus propias consultas, para lo que resulta muy necesario contar con una capa semántica.
- Integración con Microsoft Office. Especialmente con MS Excel, ofreciendo las posibilidades de análisis integradas con todas las características que ofrece este programa de hojas de cálculo.
- Búsquedas (Search-based BI). La interfaz de usuario deberá ofrecer funcionalidades de búsqueda de datos estructurados y no estructurados.
- Movilidad (Mobile BI). Facilitar la distribución de contenidos BI o analíticos a través de los dispositivos móviles ya ea a través de publicaciones y/o modo interactivo.

ANÁLISIS (ANALYSIS)

- Procesamiento analítico en línea (Online analytical processing – OLAP). Permite a los usuarios acceder a los datos a través de consultas ad hoc con cálculos y resultados rápidos, posibilitando los análisis desde diferentes niveles de agregación de la información utilizando técnicas como "slicing and dicing" o "drilldown / drillup". Del mismo modo, esta característica debería facilitar la construcción de escenarios "What if". La posibilidad de esta característica podría implicar arquitecturas da datos multidimensionales, relacionales o inclusive técnicas de procesamiento en memoria.
- Visualización interactiva (Interactive visualization). Posibilidad de mostrar los importes a través de imágenes y gráficos, en vez de tablas de filas y columnas.
- Minería de datos y modelado predictivo (Predictive modeling and data mining). Aplicación de algoritmos matemáticos y técnicas estadísticas para facilitar la identificación de características en un conjunto de datos a través del tratamiento de variables (Variable = atributo o propiedad que admite diversos datos. Las Variables Categóricas clasifican a los sujetos facilitando su agrupación. Las Variables Continuas toman un número infinito de valores).
- Cuadros de mando (Scorecards). Alineación de los indicadores claves de rendimiento (KPIs) con los objetivos estratégicos de los negocios a través del diseño de paneles.
- Modelado Prescriptivo, simulación y optimización (Prescriptive modeling, simulation and optimization). El análisis prescriptivo es el siguiente paso o un complemento del análisis predictivo, el cual, a partir de un modelo predictivo sugiere la decisión o decisiones que podrían tomarse.

A continuación se muestra un gráfico que relaciona la ventaja competitiva con el grado de inteligencia implantado en una organización. La relación entre estas dos variables determina el paso del análisis del pasado, pasando por la generación de alertas en el presente, evolucionando hacia la predicción del futuro, fundamentalmente dividido en dos fases: toma de decisiones reactivas y toma de decisiones proactivas.

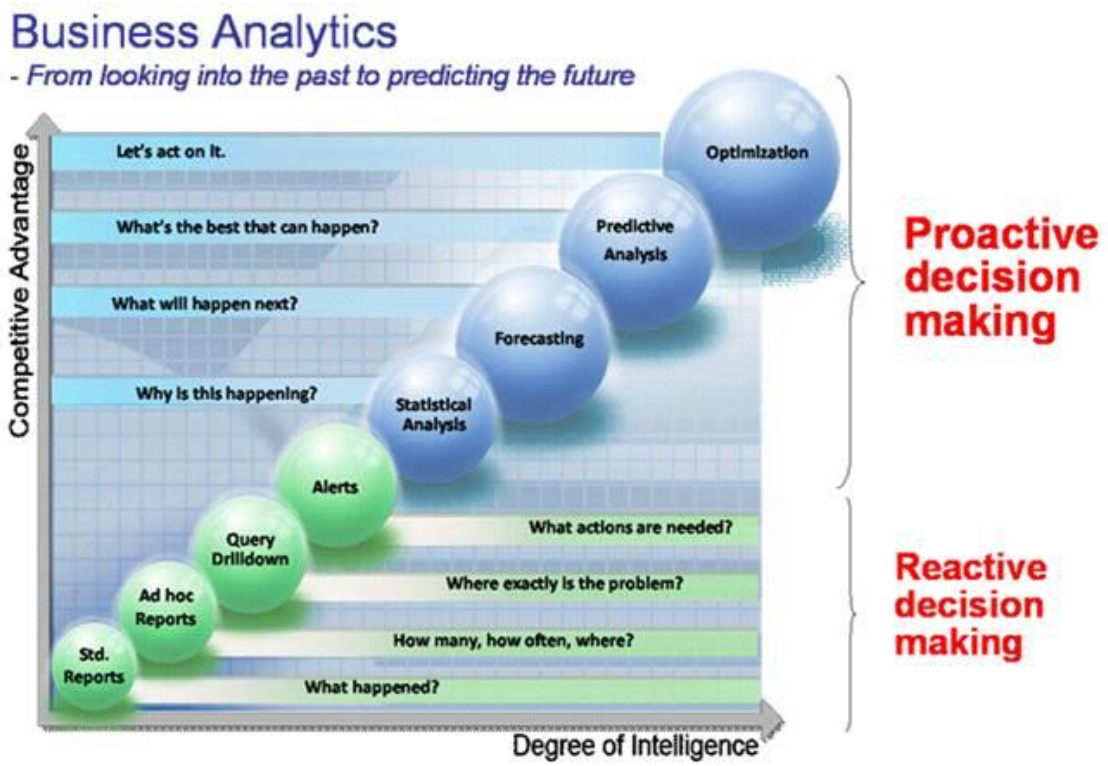


Ilustración 27. Business Analytics vs. Business Intelligence. Fuente: https://wiki.smu.edu.sg/is101_2012/img_auth.php/e/ec/Business_Analytics.jpg

Es decir, que evolucionamos del Business Intelligence tradicional al Business Analytics gracias al nuevo paradigma que trae el Big Data y los métodos de procesamiento de datos más avanzados. Con estos servicios de Business Analytics, básicamente, una organización puede efectuar dos tipos de explotaciones de datos:

- Informar: ver lo que ha ocurrido en el pasado, y tomar decisiones reactivas (Business Intelligence).
- Predecir: inferir lo que puede ocurrir en el futuro, y tomar decisiones proactivas (Business Analytics)

El Business Analytics es una inteligencia de negocios enriquecida a través de modelos estadísticos que permiten descubrir nuevas estructuras, patrones, relaciones entre variables, etc.

Capítulo III. Del modelo a la realidad

1. SIEGO, el modelo propuesto

Una vez revisados la estructura y requerimientos de un sistema de información basado en Business Intelligence, la presente tesis se plantea un modelo de sistema de información al que llamamos SIEGO, que estará constituido por cinco elementos independientes entre sí pero integrados a los que llamaremos dimensiones del modelo.

Se considera deseable que las herramientas utilizadas en cualquiera de los ámbitos del modelo SIEGO sean independientes entre sí; es decir, que cada nivel pueda funcionar de forma totalmente autónoma y desacoplada del resto de los elementos. Asimismo, cada elemento debe de estar integrado en el sistema convenientemente. El motivo de esta propuesta es que el ritmo de avance y desarrollo de las herramientas que conforman un sistema de Business Intelligence es desigual y el grado de substitución o agregación de otras herramientas debería poder hacerse con facilidad.

Para diseñar el Sistema de Información de una organización se pueden tomar dos caminos:

a. Diseñar un sistema de información como el que se propone en este trabajo, desacoplado pero integrado, en el que cada componente o dimensión del sistema se escoge entre las herramientas más adecuadas para el propósito de la organización. Generalmente se utilizan entonces herramientas especializadas en cada una de las dimensiones del modelo (Dimensión Datos, Dimensión Análisis, Dimensión Visualización, Dimensión Informes y Dimensión Publicar) o de los ámbitos de un sistema clásico de Business Intelligence.
b. Diseñar un sistema con un producto cerrado comercial. La mayoría de soluciones comerciales ofrecen productos con todos los componentes integrados. A cambio, los costes de implantación suelen ser altos, y además, tienen otros costes adicionales de mantenimientos elevados, justificados por consultorías y soporte muy especializado.

Esta es la eterna discusión en el ámbito de la tecnología: construir o comprar. Se considera que la mejor solución a medio plazo es tomar la primera opción que permitirá disponer de un sistema de información altamente flexible, desacoplado pero integrable con los sistemas de la organización. El inconveniente que presenta tomar esta opción es que precisa disponer de conocimientos técnicos más avanzados. Aun así, se entiende que un proyecto de Business Intelligence, por la importancia estratégica que tiene, debe diseñarse de manera que el control total sobre el sistema lo tenga la propia organización y no esté gobernado por un solo producto o proveedor, ni tampoco por una única empresa consultora especialista.

Si se escoge un solo producto, lo que se suele llamar una solución vertical, no seremos partícipes de las decisiones de evolución del producto, y puede darse el caso que nuestras necesidades como organización evolucionen de forma diferente a las del mercado y los clientes de la solución escogida, y que en determinado momento nos encontremos con una aplicación que no cumple con nuestras necesidades ni expectativas.

En el grafico siguiente se establece una comparación entre los elementos de un modelo clásico de Business Intelligence y el modelo propuesto SIEGO:

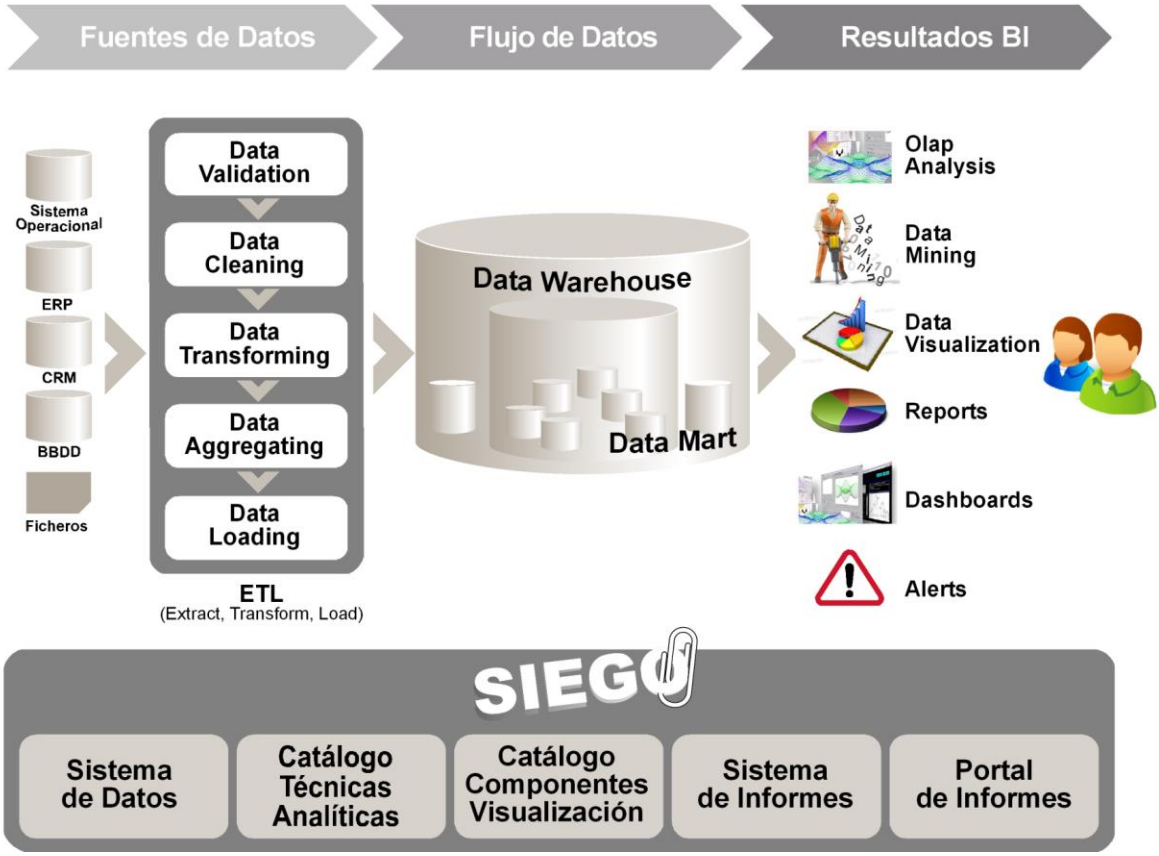


Ilustración 28. Equivalencia entre un sistema de BI y el modelo SIEGO. Fuente: Elaboración propia (2016)

El modelo SIEGO es un sistema de información que pretende aportar la fase de revisión al sistema de gestión de la organización. La organización planifica sus actividades, mediante la fase de Planificar. Una vez ha realizado el plan, empieza a ejecutarlo. Para poder mejorar la planificación y por tanto la ejecución de las actividades de la organización, sean proyectos, procesos, servicios, etc., la fase de revisión es clave puesto que recoge el conocimiento explícito de la organización y lo pone a disposición de los responsables.

La fase de revisión consiste en cinco actividades básicas: Medir, analizar, visualizar, informar i publicar. El modelo define una dimensión equivalente a cada una de estas actividades, como se describe en el gráfico siguiente:

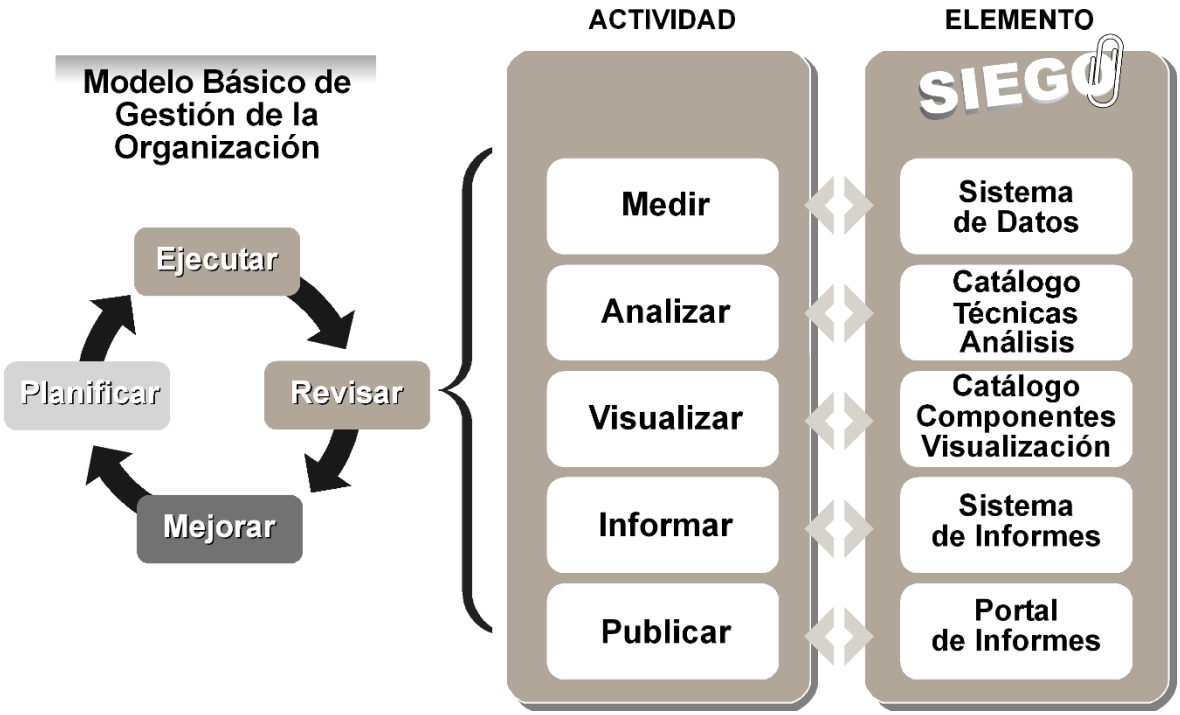


Ilustración 29. El modelo SIEGO. Fuente: Elaboración propia (2016)

En los siguientes apartados se describe en detalle cada uno de los elementos o dimensiones del modelo, que a continuación se analiza mediante los resultados del estudio empírico realizado.

2. Dimensión Datos

1 Marco general de la dimensión datos 64
1.1 La gestión del conocimiento 65
1.2 Proceso de gestión del conocimiento 69
1.3 Fuentes del conocimiento 71
2 La dimensión Datos en el modelo SIEGO 73
2.1 Medición 76
2.2 Tipos de escalas 78
2.2.1 Nominal 79
2.2.2 Ordinal 79
2.2.3 Intervalo o Discretas 79
2.2.4 De razón o Continuas 80
2.2.5 Alternativas 80
2.3 Requerimientos de la medición 81
2.3.1 Validez 81
2.3.2 Detalle suficiente 81
2.3.3 Constructos 81
2.3.4 Calidad de datos 82
2.4 La matriz de datos 83
2.4.1 Matrices de Datos de Elementos 84
2.4.2 Matrices de Datos de Hechos 86
2.4.3 Tipos de información 86
2.4.4 Medidas de indicadores 87
3 Propuesta de estructura del sistema de datos 90

1 Marco general de la dimensión datos

En este apartado se presenta la primera de las dimensiones del sistema propuesto, que recoge los datos de la organización y los estructura para poder ser consultados.

Las organizaciones actuales generan una gran cantidad de información; una parte de esta información está estructurada y almacenada en los sistemas informáticos y otra parte está fuera de ellos.

La información es el activo más importante en las organizaciones, ello es en parte debido a que su éxito depende de cuánto conozcan a sus clientes, cómo entiendan sus procesos internos y cómo de efectivo sean para realizar todas sus operaciones. La información adecuada es el único medio por el cual una organización puede conocer tales cuestiones. Además, el entorno socioeconómico actual exige a las organizaciones dar cuenta de sus procesos y resultados, lo cual requiere el tratamiento de los datos que se generan en las operaciones.

Las organizaciones son juzgadas no únicamente por la calidad de sus productos o servicios, sino también por el grado de conocimiento de sus clientes, empleados y socios. Cuánto más ampliamente disponible tenga la información una empresa, ésta se vuelve más valiosa. Cuando un departamento de marketing tiene información precisa de sus productos y servicios, está más capacitado para desarrollar promociones mejor enfocadas. Cuando los clientes pueden fácilmente verificar que un producto está en el inventario, es mucho más probable que estos realicen la compra. Cuando el directivo tiene acceso instantáneo a datos de tendencias, pueden introducir un cambio en una dirección que desemboque en unas ganancias millonarias. Esto es sólo por citar algunos ejemplos de la importancia de la información actualizada, organizada, enfocada y disponible en todo momento.

La finalidad de la información almacenada es servir de soporte para la toma de decisiones. Para ello se analizará aquella información necesaria –tanto la originada en la organización como la recogida de fuentes externas– para convertirla en conclusiones. Para llevar a cabo este proceso es recomendable el uso de métodos científicos.

En consecuencia, se propone presentar qué métodos tenemos a nuestro alcance para analizar la gran cantidad de datos que hay en las organizaciones. Las razones que hacen conveniente realizar el estudio de algunos de los métodos de análisis más relevantes son porque podremos:

- Fundamentar decisiones.

- Llegar a conclusiones que no se pueden obtener mediante la simple observación de los datos.
- Dar cuenta de forma clara y objetiva de la actividad y los resultados de la organización.
- …

Una buena gestión de la información, permite llevar a la práctica el concepto de la Gestión del Conocimiento. Según Hibbard (1996), la Gestión del Conocimiento es el "proceso de captura de la experiencia o know-how colectivo de una organización, donde sea que éste resida -en bases de datos, papel, o en las cabezas de las personas- y su distribución a donde sea que pueda ayudar a producir la mayor rentabilidad. Asimismo, Davenport (2003) modela la gestión del conocimiento como la búsqueda de ideas, el dejarlas en un formato adecuado, la promoción de estas ideas y su implementación.

1.1 La gestión del conocimiento

La Gestión del conocimiento, se reconoce como un nuevo campo de investigación ligado a las organizaciones y se da en la confluencia entre sistemas de información, teoría de las organizaciones y estrategias gerenciales (Rodriguez G. 1998). La Gestión del conocimiento se define como:

Gopal y Gagnon (1995): la identificación de categorías de conocimiento necesario para apoyar la estrategia empresarial global, evaluación del estado actual del conocimiento de la empresa, y transformación de la base de conocimiento actual en una nueva y poderosa base de conocimiento, rellenando las lagunas de conocimiento.

Bueno (1998): el conjunto de procesos que permiten utilizar el conocimiento como factor clave para añadir y generar valor.

Garvin (1994), incluye no solo los procesos de creación, adquisición y transferencia del conocimiento, sino el reflejo de ese nuevo conocimiento en el comportamiento de la organización.

Existe un acuerdo en que el origen del término Gestión del Conocimiento fue de Nonaka (1995):

- la socialización: que facilita compartir experiencias por medio de exposiciones orales, documentos, manuales y tradiciones orales y que añade el conocimiento novedoso a la base colectiva que posee la organización.
- la exteriorización: que supone hacer tangible mediante el uso de metáforas, conocimiento de por sí
- difícil de consumir, integrándolo en la cultura de la organización.
- la combinación: que crea nuevo conocimiento por integración de bases cognitivas.
- la interiorización: que analiza las experiencias adquiridas en la puesta en práctica de los nuevos conocimientos.
- la última fase de asimilación: que sintetizan dichas experiencias en la base cognitiva del trabajador o grupo.

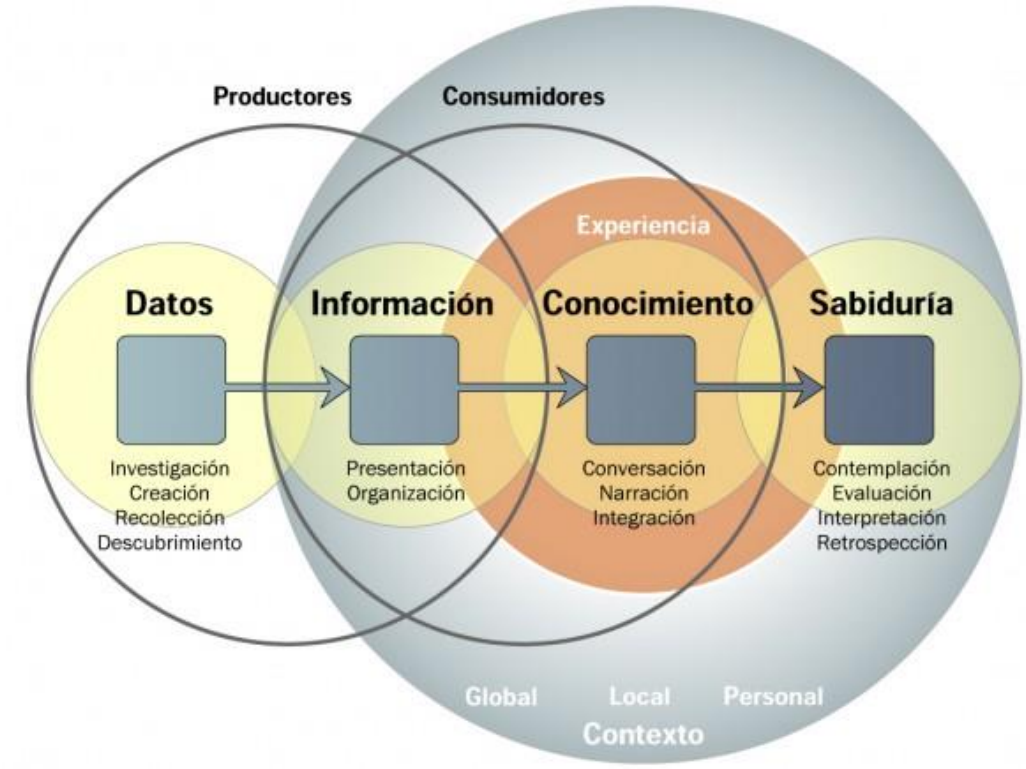


Ilustración 30 Convertir datos en sabiduría. Fuente: Internet (2015)

En la escala evolutiva, el conocimiento es la consecuencia de gestionar la información, y la información es la consecuencia de gestionar los datos. Esta posición explica que para gestionar el conocimiento es necesario que las organizaciones hayan recorrido los anteriores estadios.

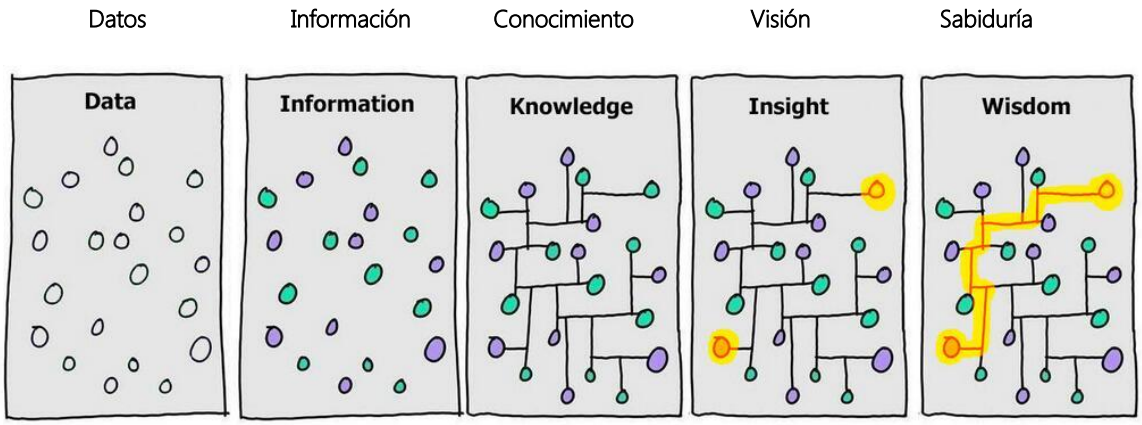


Ilustración 31. De datos a sabiduría

Datos: Cifras o valores fuera de contexto. Sin relaciones entre ellos.

Información: Cuando organizamos los datos de forma estructurada, clasificada, convertimos los datos en información:

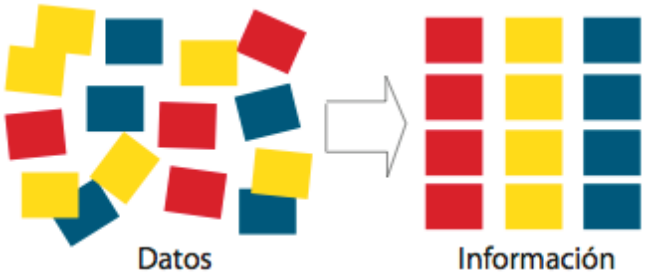


Ilustración 32. Datos convertidos en información. Fuente: Internet (2016)

Conocimiento: Se produce al combinar diferentes fuentes de datos.

Visión: Se establece cuando pueden conectarse los datos.

Sabiduría: Cuando se puede establecer relaciones claras entre datos e información, asociaciones, patrones, predicciones, etc.

El conocimiento se puede calificar como uno de los capitales de las organizaciones. Roos (1997) aportó que el valor total de una organización está compuesto per diferentes capitales:

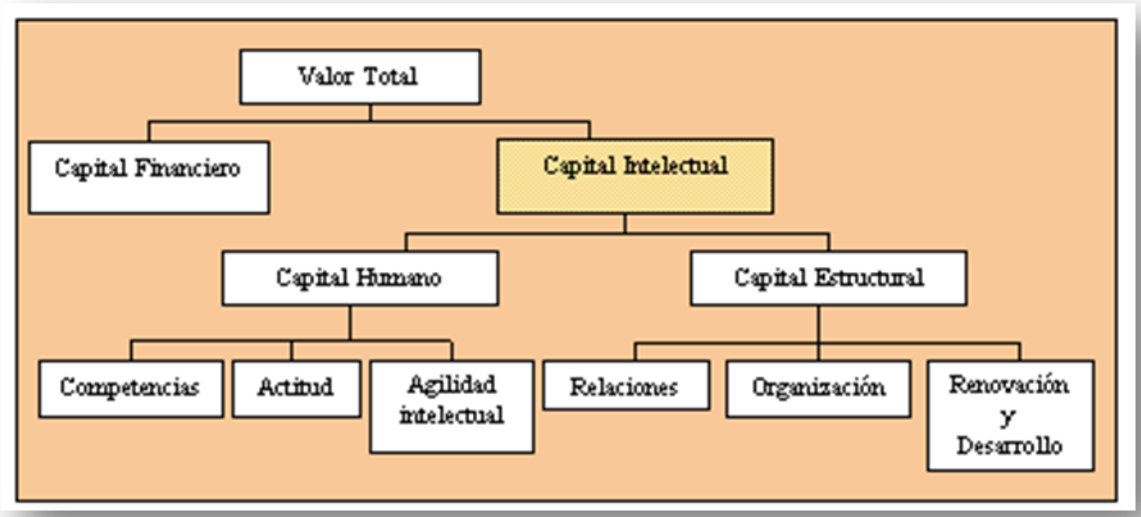


Ilustración 33. Arbol de configuración de valor. Fuente: Roos et al. (1997)

El capital humano se refiere al conocimiento (explícito o tácito) útil para la empresa que poseen las personas y equipos de la misma, así como su capacidad para regenerarlo; es decir, su capacidad de aprender. El Capital Humano es la base de la generación de los otros dos tipos de Capital Intelectual. Una forma sencilla de distinguir el Capital Humano es que la organización no lo posee, no lo puede comprar, sólo alquilarlo durante un periodo de tiempo.

AUTORES	DEFINICIÓN
List (1841)	El resultado de la acumulación de los descubrimientos, invenciones, mejoras, perfeccionamientos y esfuerzos de todas las generaciones que nos han precedido" (Capital Intelectual de la raza humana).
Kendrick (1961)	El resultado de las inversiones en el descubrimiento y difusión del conocimiento productivo.
Bontis (1996)	La relación de causalidad entre el capital humano, relacional y organizativo.
Brooking (1996)	La combinación de activos intangibles que permiten a la empresa funcionar.
Bradley (1997)	La habilidad de transformar el conocimiento y los activos intangibles en recursos creadores de riqueza para las empresas y para los países.
Edvinsson y Malone (1997)	La posesión de conocimientos, experiencia aplicada, tecnología organizativa, relaciones con los clientes y destrezas profesionales que proporcionan una ventaja competitiva en el mercado.
Stewart (1997)	El conocimiento, la información, la propiedad intelectual y la experiencia que pueden ser utilizados para crear nueva riqueza.
Sveiby (1997)	La combinación de activos intangibles que generan crecimiento, renovación, eficiencia y estabilidad en la organización.
Lev (2001)	Representa las relaciones principales, generadoras de activos intangibles, entre innovación, prácticas organizativas y recursos humanos.
Bueno (2002)	Representa la perspectiva estratégica de la "cuenta y razón" de los intangibles de la organización.
Mølbjerg-Jørgensen (2006)	Desde una perspectiva filosófica, entendido como conocimiento sobre el conocimiento, creación de conocimiento y apalancamiento del mismo en valor social o económico.
Kristandl y Bontis (2007)	Recursos estratégicos organizativos que permiten a la misma crear valor sostenible, pero que no están disponibles a un gran número de empresa (escasez). Generan beneficios potenciales futuros que no pueden ser tomados por otros (apropiabilidad) y que no son imitables por los competidores o sustituibles por otros recursos. No son transferibles debido a su carácter organizativo.

Ilustración 34. Defininiciones del Capital Intelectual. Fuente: E. Bueno (2008)

El capital estructural es el conocimiento que la organización consigue explicitar, sistematizar e internalizar y que en un principio puede estar latente en las personas y equipos de la organización. Quedan incluidos todos aquellos conocimientos estructurados de los que depende la eficacia y eficiencia interna de la empresa: los sistemas de información y comunicación, la tecnología disponible, los procesos de trabajo, las patentes, los sistemas de gestión… El Capital Estructural es propiedad de la organización, queda en ella cuando sus personas la abandonan. Un sólido Capital Estructural facilita una mejora en el flujo de conocimiento e implica una mejora en la eficacia de la organización.

El capital intelectual para Bueno (1998), está integrado por cuatro componentes básicos: el capital humano, el capital organizativo, capital tecnológico y capital relacional.

1.2 Proceso de gestión del conocimiento

Eduardo Bueno Campos (1999) aportó una interesante visión del proceso de gestión del conocimiento:

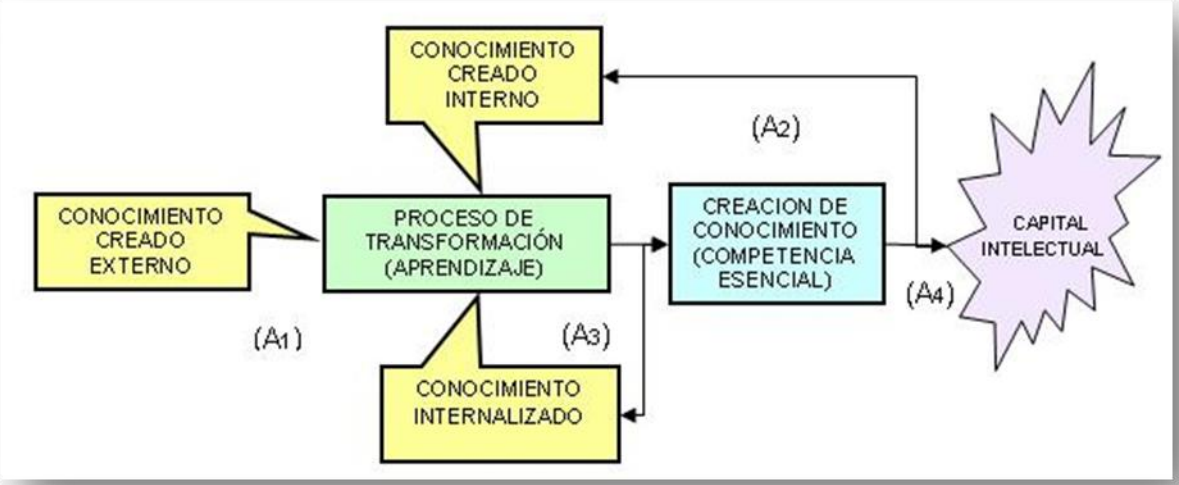


Ilustración 35. El proceso de creación de conocimiento y capital intelectual. Fuente: E. Bono (1999)

El primer flujo (A1) representa la adquisición de conocimientos del exterior, necesarios para llevar a cabo la actividad económica de la organización, según cual sea su actividad de la organización, lo que Eduardo Bueno denomina el "algo" necesario para justificar la gestión; conjunto de conocimientos que suelen ser de carácter explícito.

El segundo flujo (A2) indica como la organización es capaz de crear su propio conocimiento, el cual se utilizará como recurso para llevar a cabo su transformación en determinada competencia esencial; este conocimiento suele ser tanto explícito como tácito.

El tercer flujo (A3) representa la internalización del conocimiento o de las capacidades desarrolladas que se incorporan como "rutinas organizativas" o "procesos de acción" que hacen a la organización inteligente, es decir, dotada de capacidad para aprender; conocimiento que representa normalmente la calidad de conocimiento tácito.

Los flujos (A2) y (A3) existen gracias a la aplicación de un proceso de aprendizaje, el cual expresa la capacidad de la entidad a nivel individual, particular y colectivo de insertar nuevas ideas, conceptos, competencias en los colaboradores y en los sistemas de algoritmos que desarrollan.

El capital intelectual expresado en la figura con la forma de una explosión, representa el valor creado por el sistema que representa la gestión del conocimiento. Bueno (1999) agrega:

"En conclusión, se puede indicar que la *gestión del conocimiento* es una función dinámica o un concepto dinámico relacionado con la dirección o administración de un conjunto de flujos de conocimientos (externos e internos, captados o creados, explícitos o tácitos). De otra parte

el *aprendizaje* es el proceso de transformación y de incorporación del conocimiento tanto a nivel de persona, como de grupo o de organización en su conjunto. El *aprendizaje* en equipo, se basa en procesos de observación interpersonal y en compartir el grupo sus conocimientos individuales, para ello es fundamental trabajar en red y son, por lo tanto, técnicas facilitadoras las de la información y las comunicaciones. Finalmente, el *capital intelectual* es la medida del valor creado, es una "variable fondo" que permite explicar la eficacia del *aprendizaje organizativo* y evaluar, en suma, la eficiencia de la *gestión del conocimiento."*

No obstante, es mucho más explícito Plaz Landaeta (2005, a) a la hora de detallar con precisión los contenidos del ciclo de gestión del conocimiento.

Plaz hace uso de un concepto denominado por él "Ciclo de Gestión del Conocimiento" [CGC], asegura Plaz (2005, a) que "no existe linealidad en la gestión del conocimiento, sino espirales positivas [positive loops] que retroalimentan el sistema y permite a su vez su auto sustentación":

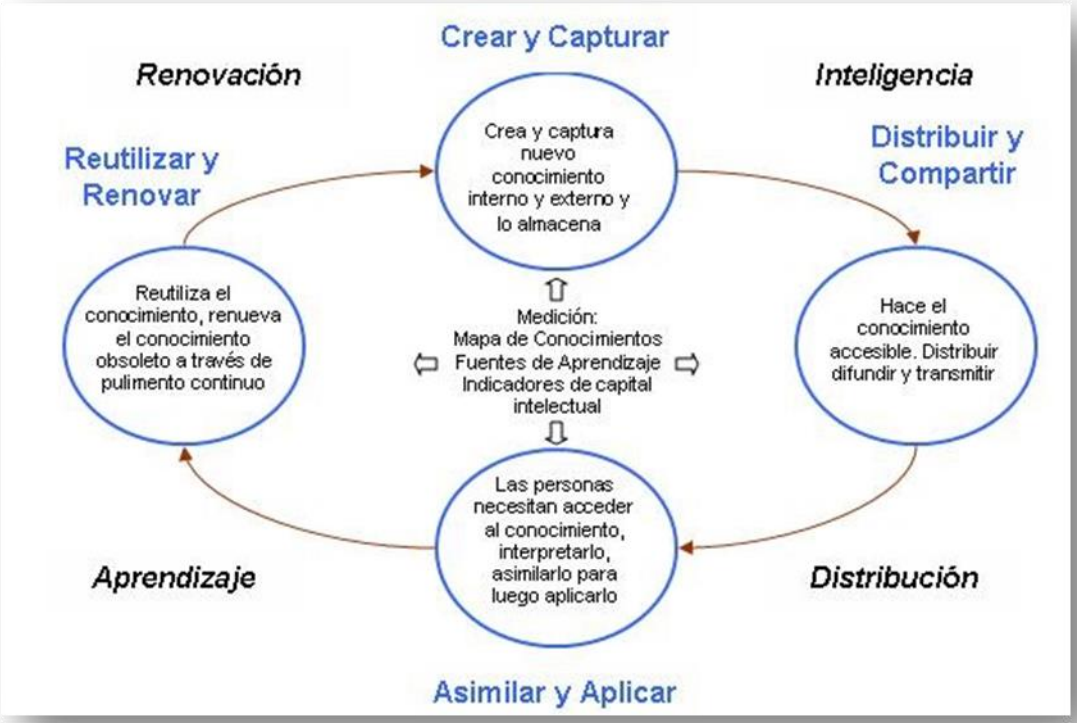


Ilustración 36. El Ciclo del Conocimiento de Plaz Landaeta. Fuente: Plaz Landaeta (2005, a)

En la dimensión inteligencia, se crea y captura nuevo conocimiento, sea este tácito o explícito; en la dimensión distribución se codifica, transmite y difunde el conocimiento a través de modelos de asociación, en la magnitud aprendizaje se asimila e internaliza la información, por último, la dimensión renovación pretende reorganizar y crear nuevo conocimiento, utilizando el juicio existente, las prácticas vividas y las disciplinas aprendidas.

Es decir, la información se convierte en conocimiento cuando se ha producido el aprendizaje, tanto el aprendizaje individual como el organizacional, aquí es donde hace aparición el

modelo mental requerido tanto en el individuo como en la organización, para que esa información creada en la primera dimensión y transmitida en la segunda se transforme en conocimiento en la tercera dimensión, se requiere el rol del modelo mental apropiado, que facilite la percepción de la anteriormente citada información y el traslado aplicativo de ese conocimiento en el contexto.

1.3 Fuentes del conocimiento

Tiessen, Andriessen y Deprez (1998) proponen 4 fuentes de conocimiento:

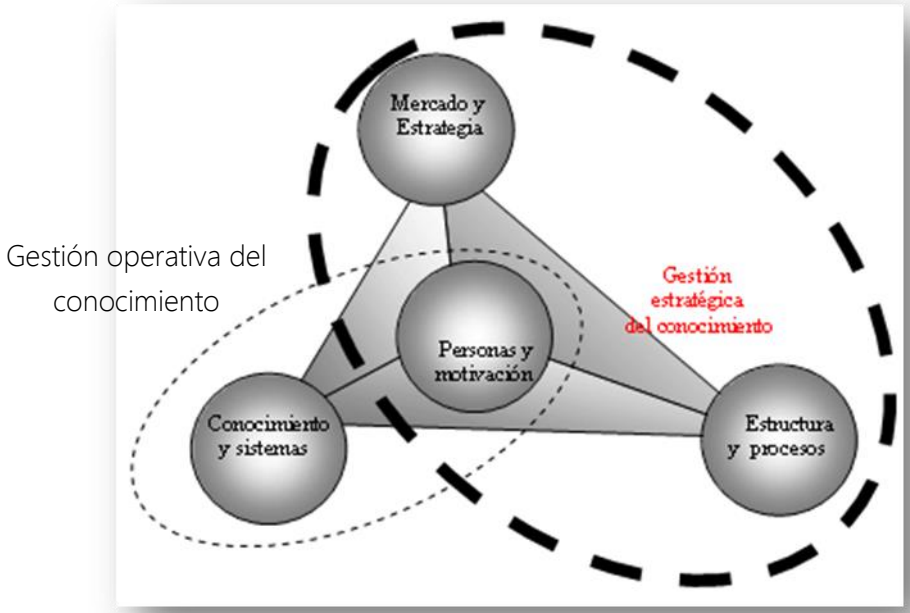


Ilustración 37. Fuentes del conocimiento. Fuente: Adaptado de Tiessen, Andriessen y Deprez (1998: 26, 34)

Siguiendo a Tiessen, Andriessen y Deprez (1998) podemos diferenciar entre gestión estratégica y gestión operativa del conocimiento. La gestión operativa utiliza las tecnologías de la información para organizar y distribuir la información hacia y procedente de los empleados. La gestión estratégica es un proceso que relaciona el conocimiento de la empresa con: 1) el diseño de estructuras organizativas que fomentan el conocimiento, 2) la estrategia empresarial, y 3) el desarrollo de profesionales del conocimiento.

Brull Alabart (2009), propone determinar que el conocimiento en una organización proviene de diferentes fuentes que se pueden situar alrededor de los grupos de Interés y de la propia organización:

Grupos de interés

- Clientes
- Personal

- Propiedad
- Partners
- Sociedad
- …

Organización

- Estrategia
- Procesos
- Estructura
- Operaciones
- Mejora
- …

De las operaciones de la organización nacen oportunidades de conocimiento nuevas, que pueden proceder de las siguientes fuentes:

Incidencias:

- De las causas de las Incidencias/No Conformidades detectadas o potenciales
- De las quejas y sugerencias, de clientes, trabajadores, proveedores, etc.

…

Informes:

- Informes (del Sistema de Informes Corporativo)
- Encuestas (clientes, clima, sociedad, …)
- Informes especializados (de auditoria, de cuentas, etc.)
- Evaluaciones EFQM
- Auditorias (ISO 9000, ISO 14000, de madurez, …)

…

Exterior

- Congresos, Jornadas, ferias, formación, …
- Internet (Blocs, webs, PLE's –Entorno personal de aprendizaje- , …)
- Revistas especializadas, catálogos, …
- Universidades, patentes, leyes y normas, …
- Benchmarking: proveedores, clientes, competencia,…
- Networking (asociaciones, etc.)

…

Personas

- Concursos o campañas de ideas (guiados o abiertos)
- Espacios de creatividad (individual y de grupos), redes

…

"El cruce de las fuentes de conocimiento con los elementos da un conjunto de formas concretas de gestionar el conocimiento. Cada organización debería identificar aquellas formas que mejor responden a sus necesidades", Brull Alabart (2009).

En este apartado del trabajo, proponemos un modelo de datos que recoja todo el conocimiento que constituye el capital estructural de la organización, que es el conocimiento que se consigue explicitar, que facilita un avance en el flujo de conocimiento e implica una mejora en la eficacia de la organización. Este proceso constituye la Dimensión Datos del modelo propuesto SIEGO.

2 La dimensión Datos en el modelo SIEGO

Medir es la primera fase de la 'Revisión' en un sistema de gestión básico y consiste en recoger información (datos o valores) de la actividad de la organización, de sus procesos y operaciones. Es frecuente designar como 'indicadores' o 'métricas' las magnitudes que se miden para poder evaluar un proceso. Es necesario que las medidas sean tomadas de acuerdo a su diseño. No olvidemos que las medidas obtenidas se utilizarán como soporte para la toma de decisiones.

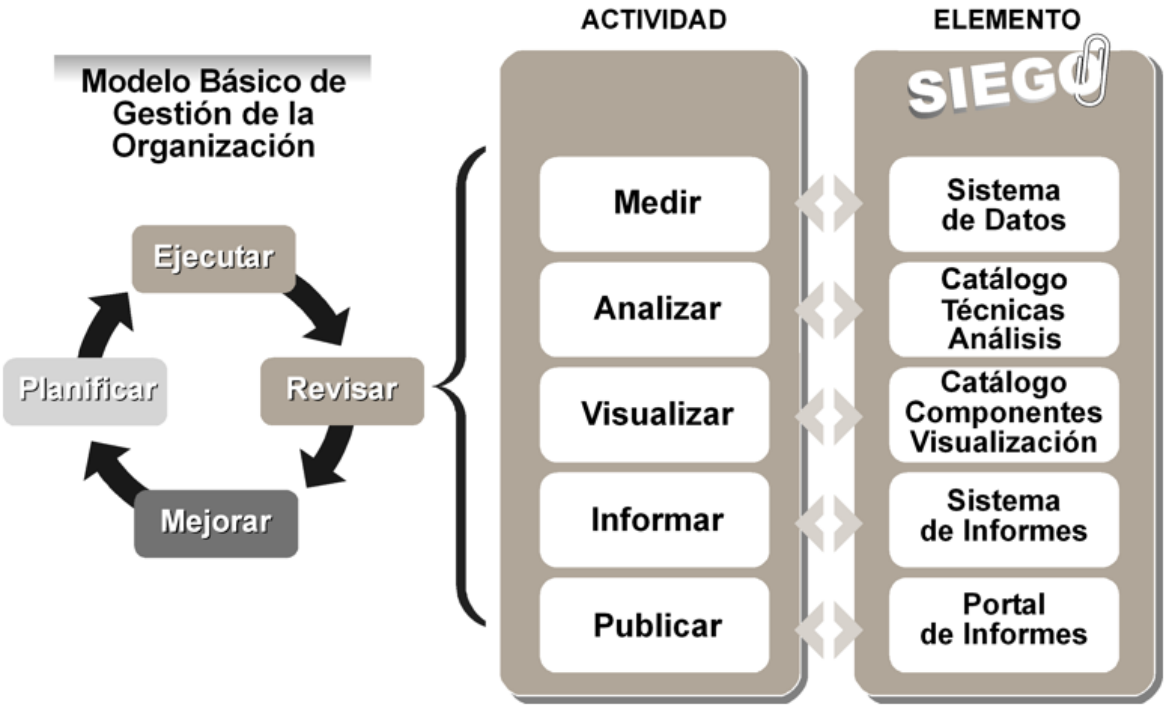


Ilustración 38. Equivalencia actividad con elementos del modelo SIEGO.

Dato:

> "Es el antecedente necesario para llegar al conocimiento exacto de una cosa o para deducir las consecuencias legítimas de un hecho."
>
> "Es el resultado de un proceso de elaboración, hay que construirlo." Gil (1994).

Un dato Galtung (1966) tiene una estructura compuesta por tres elementos:

- Es una Unidad de Análisis
- que es una Variable
- y que asumirá un determinado Valor.

La Unidad de Análisis son los elementos menores y no divisibles que componen el universo de estudio de una investigación. Sobre dichos elementos se estudia el comportamiento de las variables.

Las Variables son como 'aspectos o dimensiones' de un fenómeno capaces de asumir diferentes cualidades y/o valores. Expresan un concepto cuantificable en la unidad de análisis. En sociología una variable es un atributo o cualidad que presentan los individuos o los hechos sociales susceptibles de ser observados y medidos de alguna forma.

Los Valores o categorías son las diferentes posiciones o alternativas que presenta la Variable y adopta alguna Unidad de Análisis. Se pueden expresar cualitativamente a través de una clasificación por ausencia o presencia, por jerarquía u orden o, sino, cuantitativamente a través de magnitudes.

Los datos constituyen la materia prima de la fase de revisión y análisis, pudiéndose establecer distintas clasificaciones en función de la forma en que éstos vengan dados (Arriaza, 2008:6). Se obtienen datos al realizar cualquier tipo de prueba, experimento, valoración, medición, observación, etc. Una vez obtenidos los datos pueden generarse nuevos datos mediante transformación y/o combinación de los datos originales. Al conjunto de datos convenientemente organizados se le llamará *modelo de datos*.

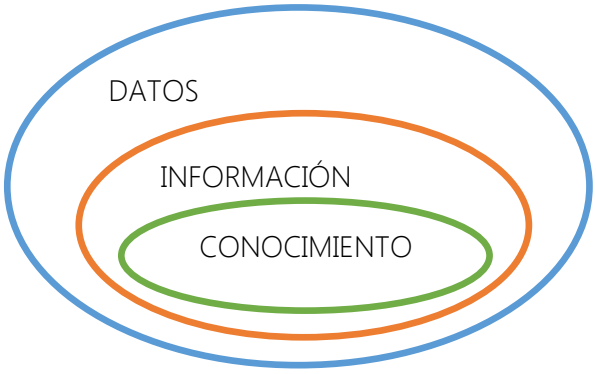


Ilustración 39. Datos, información, conocimiento

Otra aproximación relevante a la definición de datos, es la de la norma UNE-ISO/IEC 15939:2009 Ingeniería del software y sistemas. Procesos de medición, que define cada concepto involucrado en el sistema de datos de una organización.

Entidad: Objeto que va a ser caracterizado, por la medición de sus atributos. Por ejemplo un proceso, un producto, un servicio.

Atributo: Propiedad o característica de una entidad que puede ser identificada cuantitativamente o cualitativamente por medios humanos o automáticos. El atributo se puede medir (cuantificar) por medio de una métrica directa o indirecta.

Medida (substantivo): Número asignado a un atributo de una entidad como resultado de una medición.

Medida básica: Medida definida en términos de un atributo y del método para cuantificarlo.

Medida derivada: Medida que es definida como una función de dos o más medidas básicas

Indicador: Medida que contiene una estimación o evaluación de atributos específicos derivados de un modelo con respecto a necesidades de información definidas. Es el método de cálculo y la escala definidos, además del modelo y criterios de decisión con el fin de proveer una evaluación o estimación de un concepto medible con respecto a una necesidad de información.

Criterio de decisión o meta: Umbral, objetivo o patrón, utilizado para determinar la necesidad de acción o más investigación, o para describir el nivel de confianza en un determinado resultado.

Valor: Resultado numérico o categoría, asignado a una medida básica, una medida derivada o un indicador.

Proceso de medición: El proceso para establecer, planear, realizar y evaluar mediciones de software dentro de un proyecto o de una estructura de mediciones en una organización

Método de cálculo: Secuencia lógica de operaciones, descriptas genéricamente, utilizado para cuantificar un atributo con respecto a una escala especificada.

Procedimiento de medición: Conjunto de operaciones, descriptas específicamente, utilizadas en la realización de una medición específica siguiendo un método dado.

Unidad organizacional: La parte de una organización elegida para realizar mediciones

Dueño del proceso de medición: Persona o individuo responsable del proceso de medición.

Método (de medición, cálculo) Secuencia lógica de operaciones, expresadas de forma genérica, que permite la realización de una descripción de actividad. El tipo de método de medición va a depender de la naturaleza de las operaciones utilizadas para cuantificar el atributo. Pueden distinguirse dos tipos:

- Subjetivo: Cuando la cuantificación supone un juicio realizado por un ser humano.
- Objetivo: Cuando la cuantificación está basada en métodos numéricos.

Escala: Conjunto ordenado de valores continuos o discretos o conjunto de categorías con las que el atributo es comparado. Un conjunto de valores con propiedades definidas:

- Escala Numérica (Continua o Discreta)
- Escala Categórica

- Tipos de Escala (Nominal, Ordinal, Intervalo…)

Métrica: El método de medición definido y la escala de medición. Es la correspondencia de un dominio empírico (mundo real) a un mundo formal, matemático. La medida incluye al valor numérico o nominal asignado al atributo de un ente por medio de dicha correspondencia. [Fenton]. Las métricas NO pueden interpretar por sí solas un concepto medible, precisan de indicadores para ello.

Métrica directa: Una métrica de un atributo que no depende de ninguna métrica de otro atributo.

Métrica indirecta: Una métrica de un atributo que se deriva de una o más métricas de otros atributos. Se formaliza por medio de una Función de Medición (Fórmula, Ecuación).

Función de Medición: Un algoritmo o cálculo que permite combinar dos o más métricas.

KPI (Key Performance Indicators o Indicadores clave de rendimiento): son indicadores, acordados de antemano, que reflejan los factores críticos de éxito de una organización. Cualesquiera que sean los KPI escogidos, deben reflejar los objetivos de la organización, deben ser la clave para su éxito, y deben ser cuantificables (medibles). Los indicadores clave de rendimiento suelen ser consideraciones a largo plazo. La definición de lo que son y cómo se miden no cambian a menudo. No así las metas, que suelen cambiar a medida que se acerca la meta.

Un ejemplo de KPI

Una empresa de transporte de viajeros por carretera, tiene como uno de sus KPI fundamentales la puntualidad de sus vehículos en las líneas interurbanas de las que es titular.

Este indicador se controla desde la Dirección General y desde la Dirección de Operaciones todos los días (cada cambio de turno en concreto) y se comunica a los inspectores de las líneas y a sus conductores.

Este indicador es un KPI que tiene un impacto en:

- La reducción de costes (menos combustible, vehículos de refuerzo, etc.).
- La satisfacción de los clientes, que tendrán en cuenta el autobús frente a medios de transporte alternativos y a la competencia.
- El ambiente laboral de los conductores, ya que no tendrán que soportar las quejas de los viajeros, alargar sus turnos y, además, cobrarán incentivos.
- La capacidad de planificación de las operaciones de la empresa, evitando inversiones innecesarias.

2.1 Medición

La medida no es una actividad natural. La naturaleza no mide, sólo trata con 'patrones que conectan', no con cuantificaciones.

Con Aristóteles, la medida era comparativa: más alto, más rápido, etc.; Galileo (XVI) midió el movimiento; Descartes representó el movimiento en un diagrama (coordenadas cartesianas) y Newton estableció las leyes matemáticas que vinculan el movimiento con la gravitación.

El nuevo paradigma enfoca medir para informar, complementario a medir para aprender.

Es relevante destacar que a menudo no podemos obtener información o conocimiento directamente de los datos, que debemos aplicar ciertas técnicas de análisis y ciertos componentes de visualización para entender mejor los datos. Por ejemplo, en la siguiente tabla de datos no podemos realizar ningún análisis a simple vista, excepto el valor de x4 que es siempre 8 excepto para uno de los datos.

x1	y1	x2	y2	x3	y3	x4	y4
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Ilustración 40. Sample data sets reacrated from Asncombe, Francis J. (1973) Graphs in statistical analysis. American Statistician, 27, 17-21.

Cuando representamos estos datos en un gráfico de coordenadas:

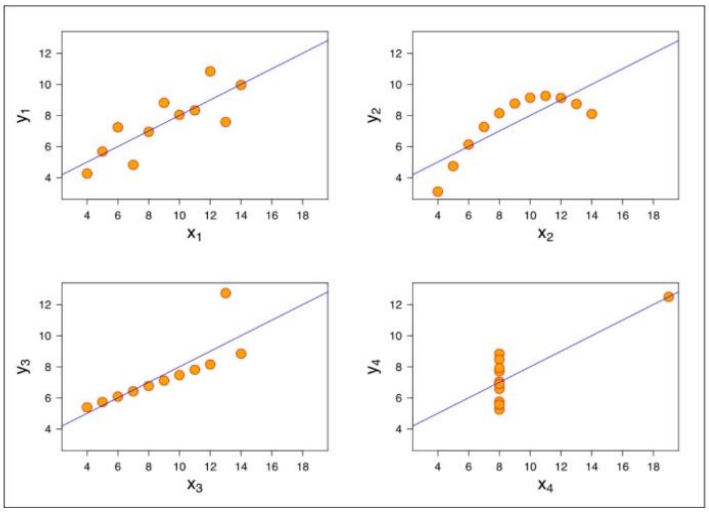


Ilustración 41. www.commons.wikimedia.org/wiki/File-Ascombe%27s_quartet_3.svg

Inmediatamente somos capaces de detectar ciertos patrones simplemente creados por las relaciones entre el cruce de los valores x y y de los cuatro juegos de datos.

Este gráfico representa los cuatro conjuntos de datos definidos por Francis Anscombe en 1973, para los que algunas de las propiedades estadísticas habituales (media, varianza, correlación y regresión) son los mismos, a pesar de que los conjuntos de datos son diferentes.

Fueron construidos para demostrar tanto la importancia de la representación gráfica de los datos antes de analizarla como el efecto de los valores extremos en propiedades estadísticas, como un intento de atacar la impresión entre los estadísticos que " los cálculos numéricos son exactos, pero los gráficos son aproximados".

2.2 Tipos de escalas

En este trabajo se clasifican las escalas en cuatro tipos fundamentales:

a. Cualitativas o de atributos:
    - Nominales
    - Ordinales
b. Cuantitativas o de variables:
    - Discretas
    - Continuas

Tipo de Escala	Transformación Admisible	Estadísticas Significativas	Ejemplos
Nominal	Equivalencia uno-a-uno	Moda, Frecuencia Test no paramétricos	Clasificación, etiquetado
Ordinal	Funciones monotónicas incrementales	Mediana, Spearman y Kendall Tau, W Test no paramétricos	Grado de disponibilidad, funcionalidad, complejidad, etc.
Intérvalo	m1 = a m + b	Media Aritmética, Desvío estándar Coef. de correlación de Pearson Test no paramétricos (Todo lo previo)	Temperatura (Celsius, Fahrenheit) Tiempos relativos, etc.
Proporción	m1 = a m	Porcentajes, Media Geométrica Test paramétricos y no paramétricos (Todo lo previo)	Preferencia de calidad, tamaño (LOC), longitud, etc.
Absoluta	m1 = m (Identidad)	(Todo lo previo)	Conteo

Ilustración 42. Escalas, transformaciones y estadísticas significativas

2.2.1 Nominal

Se realiza una medida nominal cuando la propiedad estudiada en los objetos solo puede agruparse en categorías lógicamente exhaustivas y mutuamente exclusivas, de tal modo que puedan establecerse claramente equivalencias o diferencias.

Sus valores se han obtenido con una escala de medida nominal.

Ejemplos: El género es una variable cualitativa nominal con dos categorías (o atributos): Hombre y Mujer; Municipio de residencia; las variables 'sí o no' 'proceder o no proceder', 'aceptar o rechazar', etc.

Algunos ejemplos de preguntas que utilizan medidas Nominales (Si/No):

- ¿Llegó el empleado a tiempo a su trabajo?
- ¿Se escribió la carta sin errores?
- ¿Está el departamento por debajo del presupuesto?
- ¿Se inició la reunión de conformidad con el programa?
- ¿Se devolvió el informe de acuerdo con lo programado?
- ¿Se contestó el teléfono a más tardar al segundo timbrazo?

2.2.2 Ordinal

Sus valores se han obtenido con una escala de medida ordinal. Tienen un carácter nominal, pero admiten una ordenación.

Ejemplos: El Diagnóstico de un Paciente es una variable ordinal que puede tomar como valores las siguientes categorías: mucho mejor, mejor, igual, peor. La Formación Académica={primaria, secundaria, media, superior, doctorado]; opinión={muy insatisfecho, insatisfecho, indiferente, satisfecho, muy satisfecho}

2.2.3 Intervalo o Discretas

Las medidas discretas, suministran una historia más detallada de los procesos. Esto abarca la recolección de valores numéricos que cuantifican la medición.

Sólo pueden tomar ciertos valores dentro de un campo de variación. No tienen sentido los valores intermedios.

Algunos ejemplos de indicadores que utilizan medidas de intervalo:

- Número de granos de una espiga de trigo
- Número de clientes diarios
- Número de veces que suena el teléfono antes de ser contestado
- Número de trabajadores
- Número de días al año que los empleados se toman por concepto de enfermedades

2.2.4 De razón o Continuas

Puede tomar cualquier valor y tienen un cero absoluto.

Algunos ejemplos de indicadores que utilizan medidas de razón:

- Costo del correo de entrega inmediata
- Índice de satisfacción
- Número de horas para procesar una solicitud de cambio en ingeniería.
- Valor del inventario en dinero
- Temperatura ambiente de la sala de clientes
- Distancia entre tiendas

2.2.5 Alternativas

Hemos dicho que una variable puede ser medida de diferentes maneras. Su elección depende del tipo de investigación.

Así, por ejemplo, si nos interesamos por el consumo de alcohol, el investigador puede trabajar con cualquiera de las siguientes variables:

- Nominal (2 categorías): Si/no
- Ordinal (3 categorías): Abstemio/Bebedor moderado/Gran bebedor
- Intervalo: número de copas y combinados bebidos durante una semana
- Razón: gramos de alcohol ingeridos durante una semana

Los "gramos de alcohol ingeridos durante una semana" es la variable de mejor calidad y la que contiene el máximo de información.

Una misma variable puede diseñarse en forma de atributo o en forma de variable:

Atributos	Variables
¿Llegó el empleado a tiempo a su trabajo?	¿Con cuántos minutos de retraso llego el empleado?
¿Se escribió la carta sin errores?	¿Cuántos errores había por cada 100 palabras registradas?
¿Está el departamento por debajo del presupuesto?	¿De qué tamaño son las variaciones entre gastos reales y el presupuesto?
¿Se inició la reunión de conformidad con el programa?	¿Con cuántos minutos de retraso se dio comienzo a la reunión?
¿Se devolvió el informe de acuerdo con lo programado?	¿Con cuántas horas de anticipación o retraso se entregó el informe?
¿Se contestó el teléfono a más tardar al segundo timbrazo?	¿Cuántas veces timbró el teléfono antes de ser contestado?

Ilustración 43. Atributos y Variables. Fuente: Elaboración propia (2015)

2.3 Requerimientos de la medición

Kaydos (1991) Medir todas las variables (o indicadores) correctas en un sistema de producción es la primera condición para un sistema de mediciones de resultados efectivos, pero no es suficiente. Para ser efectiva, la medición debe satisfacer los siguientes requerimientos.

2.3.1 Validez

Las mediciones de resultados deben ser válidas. Esto es, deben medir lo importante y ser aceptadas y entendidas por los usuarios.

2.3.2 Detalle suficiente

En la teoría de la medición hay un concepto conocido como variabilidad. Puede entenderse como el número de diferentes modos de respuesta de un sistema. Conforme se incrementa la variabilidad de un sistema, más información se necesita para describir su conducta. Para poder controlar un sistema, el sistema de control debe tener, al menos, tanta variabilidad como el sistema a controlar. En otras palabras, los sistemas complejos no pueden controlarse mediantes mecanismos simples. La frecuencia de las mediciones debe ser consistente con el tiempo de respuesta del sistema. Son suficientes de 6 a 10 muestras durante el periodo que se estime en el que una variable pueda cambiar significativamente. Por ejemplo: una variable que pueda variar considerablemente en una semana, debe medirse diariamente; si puede variar considerablemente en un día, debe medirse cada tres horas como mínimo.

El sistema de medición debe identificar, al menos, el 80% de la desviación entre los resultados actuales y los deseados. Si no es así, no podrá identificar todas las causas significativas de las variaciones del output del sistema de gestión y, por lo tanto, no conocerá realmente lo que está sucediendo. Así, pues, el 20% de las causas será responsable del 80% de las variaciones del output. Las otras serán tan pequeñas, que no necesitaran ser medidas.

Si las mediciones no se hacen con la frecuencia suficiente, el resultado puede ser un cuadro distorsionado. (Kaydos, 1991)

2.3.3 Constructos

Conocer todos los aspectos, facetas o dimensiones que determinan un concepto complejo -también llamado constructo- como la orientación al mercado, la innovación o la satisfacción de los clientes es necesario de cara a poder ofrecer una definición completa del mismo, y posteriormente desarrollar estrategias de mejora basadas en su explotación.

Para dar respuesta a este tipo de situaciones, las investigaciones utilizan las llamadas escalas múltiples o sistemas de medida. A través de estas escalas es posible establecer un sistema de medida donde se recogen y se cuantifican, todas y cada una de las dimensiones que tiene un

concepto complejo o constructo, de tal manera que al final es posible establecer un valor para este concepto basado en una combinación lineal de todas y cada una de estas dimensiones. La medida agrupada se conoce con el nombre de constructo, mientras que cada una de estas dimensiones estarán representadas por uno o varios indicadores (Muñiz *et al.*, 2005) que presentan un peso específico determinado en el resultado final.

Ilustración 44. Modelo de Constructo.

Es habitual encontrar los constructos:

- Satisfacción de clientes
- Satisfacción de personas (clima laboral)
- Índice de calidad de vida

2.3.4 Calidad de datos

Levis et al. (2007) mencionan ejemplos de escenarios en los que datos con niveles de calidad inadecuados originan problemas que afectan negativamente a los Sistemas de Información y, por tanto, al rendimiento organizacional. Las causas más comunes de esos niveles inadecuados son una serie de obstáculos a lo largo del ciclo de vida de los datos en el sistema de información como los descritos por Strong et al. (1997). Por su parte, los problemas producidos por esta falta de calidad pueden ser clasificados a diferentes niveles según su naturaleza: técnicos (como los referidos a la implementación de almacenes de datos), organizacional (tales como pérdida de clientes, grandes pérdidas financieras o incluso insatisfacción de los trabajadores) y legales (como la violación de la Ley de protección de datos).

Sin embargo, y a pesar de haber obtenido un conjunto de conceptos con bastantes elementos comunes, se da la circunstancia de que normalmente se han usado diferentes términos para el mismo concepto, lo que dificulta la puesta en práctica de los resultados que se han obtenido en distintos escenarios. Para tratar de paliar esta inconsistencia en la nomenclatura, Caballero et al. (2007) han analizado los términos de los autores más referenciados y han propuesto una terminología unificada de acuerdo con el Modelo de Información de Medición de Software

propuesto en ISO/IEC 15939 y la ha ampliado con los aspectos específicos de calidad de datos. El resultado es un Modelo de Información de Medición de Calidad de Datos (Data Quality Measurement Information Model, DQMIM). Un valor añadido de ISO/IEC 15939 es que proporciona una metodología para la definición de planes de medición de software observando aspectos organizacionales. Esta metodología podría ser adaptada y extendida con los aspectos específicos de la calidad de los datos a los que antes se hacía referencia para cubrir el hueco existente en el campo de calidad de los datos. La principal contribución de este artículo al campo de la calidad de los datos es MEPLAMECAL, una metodología basada en la propuesta por ISO/IEC 15939 para la elaboración de planes de medición de calidad de datos.

La definición de la metodología implica no sólo la identificación de las actividades y tareas correspondientes, sino también la identificación de los productos de entrada y salida para cada una de ellas. Además, se identifican, a partir de las utilizadas en el campo de Ingeniería del Software, algunas técnicas y herramientas útiles para transformar las entradas en salidas.

De acuerdo al modelo propuesto, las medidas de calidad de datos pueden ser de uno de los siguientes tipos: medidas base, medidas derivadas o indicadores.

Las entidades a medir se pueden encontrar en almacenes de datos de distinta naturaleza como, por ejemplo, bases de datos relacionales o semi-estructuradas, ficheros de acceso secuencial, documentos XML u hojas de cálculo.

Para planificar la medición, es preciso localizar la ubicación de las entidades cuyo nivel de calidad de datos va a ser medido. Estas entidades tienen atributos medibles que deben ser inspeccionados a la hora de trazar el plan de medición.

Algunos ejemplos de entidades pueden ser esquemas de datos, valores de datos, dominios de datos, reglas de negocio, o interfaces de usuario.

Una de las principales técnicas para localizar estas entidades son las sesiones de trabajo con los responsables de los recursos de datos de la organización.

2.4 La matriz de datos

Entendemos que un punto clave es disponer de Matrices de Datos bien diseñadas. De las matrices de datos se obtienen los posteriores indicadores y son la base de los distintos análisis.

Es una forma de sistematizar la información recogida de la realidad para investigar un problema y tratar de obtener conocimiento científico que intente explicar dicho problema a través del método de investigación científica.

Los datos obtenidos se organizan en matrices de i x j, donde cada fila (o registro) representa una Unidad de Análisis y las columnas (o variables) las características observadas.

Las casillas de cruce x_{ij} representa el comportamiento de la Unidad de Análisis I_i respecto a la Característica C_j, es decir, el valor.

	C1	C2	...	Cj
UA1	x11	x12	...	x1j
UA2	x21	x22	...	x2j
...	...	...	...	...
UAi	xi1	xi2	...	xij

Ilustración 45. La matriz de datos. Fuente elaboración propia.

Las características pueden ser o no susceptibles de medida cuantitativa. En el primer caso nos encontramos ante una variable y en el segundo ante un atributo.

La variable tomará distintos valores observados cuantitativamente mediante una medida y una escala determinadas. Para designar variables utilizaremos las últimas letras del alfabeto en mayúsculas (..., X, Y, Z) y para designar los valores que puede tomar una variable determinada, la letra correspondiente en minúsculas y afectada por un subíndice de orden (x1, x2, xn; y1, y1, ... yn; z1, z2,... zn)

El atributo tomará distintas modalidades observadas cualitativamente. Para designar los atributos utilizaremos la primeras letras del alfabeto en mayúsculas (A, B, C, ...) utilizando las minúsculas para designar las distintas modalidades que puede tomar un atributo determinado.

La Matriz de Datos es la manera más básica de crear estructuras de datos

Podemos encontrar dos tipos de Matrices de Datos básicas, según que la estructura de sus filas sea de elementos o de hechos

2.4.1 Matrices de Datos de Elementos

Las Matrices de Datos de Elementos constan de datos tomados en el mismo momento del tiempo, no repetidos y que corresponden a un período determinado.

	C1: Preg01	C2: Preg02	C3: Preg03	...	Cn: Pregnn
Individuo 1	7	5	8		7
Individuo 2	7	6	7		6
Individuo 3	6	6	7		6
...					
Individuo i	Di1	Di2	Di3		Dij

Ilustración 46. Matriz de datos de elementos. Fuente: Elaboración propia.

Corresponden a las tablas 'd' -de dimensiones- de un *Data Warehouse* y se presentan en la Segunda Forma Normal (2FN).

Algunos ejemplos de Matrices de Datos de Elementos son:

- Clientes
- Artículos (catálogo)
- Individuos (personas)
- Puestos de trabajo
- Municipios (Geografía)
- Inventario
- Acciones formativas
- Proyectos
- Partidas
- Terceros
- Proveedores
- …

Una dimensión especial es la de 'Tiempo' que incluye todas las fechas de un período con datos detallados.

Existe la posibilidad de incluir dimensiones anidadas jerárquicas. La información se podrá presentar en mayor o menor nivel de agrupación (granularidad).

Por ejemplo:

- Jerarquía geográfica (Continente / País /Región/Provincia/Municipio)
- Jerarquía de producto (Marca/Familia/Producto/Tipo de Presentación)
- Jerarquía comercial (País / Zona/ Punto de venta)
- Jerarquía temporal (Año/ Trimestre/Mes/Día)

Representadas por un ordenamiento lógico dentro de la dimensión, se encuentran formadas por los diferentes tipos de relaciones entre los atributos de una misma dimensión. Pueden existir múltiples jerarquías dentro de una dimensión pero siempre es posible identificar una jerarquía principal o columna vertebral de la dimensión y jerarquías secundarias o descriptivas compuestas por atributos característicos definidos desde la jerarquía principal.

Dentro del contexto de navegación del modelo dimensional, se puede decir que las diferentes jerarquías definen el mapa de caminos para el "drilling" o la "navegación" de los datos. Se hace drill-up o roll-up cuando se mueve los datos hacia un atributo superior dentro de la jerarquía (navegación ascendente), drill-down o roll-down cuando se analiza información a mayor nivel de detalle (navegación descendente), drill-within cuando se mueven los datos entre la jerarquía principal y la característica desde un atributo hacia cualquier otro que no sea ni descendiente ni ancestro dentro de la misma dimensión (navegación intradimensional) y drill-across para analizar información sobre diferentes dimensiones (navegación interdimensional).

2.4.2 Matrices de Datos de Hechos

Las Matrices de Datos de Hechos constan de datos tomados en distintos momentos de tiempo; siempre tienen una característica de tiempo. Corresponden a las tablas 'f' –de *facts*- de un *Data Warehouse* y se representan en la Tercera Forma Normal (3FN).

	C1: Fecha	C2: Cliente	C3: Producto	...	Cn-1: Cantidad	Cn: PVP
Venta 1	1-3-10	ABC, SA	CX-3		50	145
Venta 2	4-3-10	Paint, SL	CZ-2		210	356
Venta 3	6-3-10	ABC, SL	CX-3		100	290
...	...	...	...		...	...
Venta i	Di1	Di2	Di3		...	Dij

Ilustración 47. Matriz de datos de hechos. Fuente elaboración propia.

Algunos ejemplos de Matrices de Datos de Hechos son:

- Ventas
- Asignación (PersonaPuesto)
- Compras
- Estrategia
- Contabilidad
- Presencia
- Mantenimiento
- Comunicación
- Operaciones
- Reclamaciones
- Proyectos Mejora

...

Denominamos "hechos" a los indicadores de negocio. Por ejemplo, son "hechos" las ventas, los pedidos, los envíos, las reclamaciones, las compras, etc. Es decir, son todas aquellas medidas numéricas que incluiremos en nuestro sistema de datos. Técnicamente, una tabla de hecho es la tabla central de un modelo en estrella.

2.4.3 Tipos de información

Los resultados alcanzados dependen mucho de la calidad de la información, de ahí la importancia que la misma tiene. Los tipos de información con los que se deben trabajar serán:

- Series de tiempo
- Series de corte transversal
- Combinación de las dos anteriores
- Datos Panel

a) Series de Tiempo

Son aquellas que se almacenan durante un período de tiempo. Ejemplo el PNB, cifras de empleo, salarios etc. Se recogen en intervalos regulares de tiempo ya sea anuales, mensuales, trimestrales, diarios, etc. Pueden ser cuantitativas o cualitativas.

b) Corte Transversal

Se recogen de acuerdo a una o más variables en un momento dado en el tiempo. Ejemplo: los indicadores anuales de una empresa; el Censo de Población, que se hace cada 10 años; la evaluación anual del rendimiento del personal o del clima laboral. También denominados como datos en sección cruzada que miden una variable particular en un período de tiempo, para diferentes entidades.

c) Datos Combinados

Se tiene la información tanto de serie de tiempo como de corte transversal. Ejemplo: Tasa de Inflación de 5 países en un año, que se tomaría de las Series de Tiempo de la tasa de inflación de 5 países, durante 20 años.

Se debe señalar que cualquier columna del cuadro es una sección cruzada mientras que cualquier fila es una serie de tiempo.

Generalmente cuando se trabaja con series de tiempo el subíndice que se utiliza es "t", mientras que cuándo se trabaja con series de corte transversal el subíndice que se utiliza es "i".

d) Datos Panel

También se pueden presentar los llamados Datos Panel (ó datos longitudinales) que son un tipo especial de datos combinados de sección cruzada y series de tiempo, por ejemplo, el gasto familiar donde ciertas familias elegidas mantienen registros de sus gastos a lo largo de varios años.

Los Datos Panel por lo general son micro datos (pertenecen a agentes económicos individuales como las familias o las empresas). Estos son preferibles a los macro datos porque se evita la agregación.

2.4.4 Medidas de indicadores

Con los datos estructurados en Matrices de Datos, se pueden obtener una gran cantidad de medidas de indicadores. La herramienta ortodoxa para gestionar las matrices de datos es el Data Warehouse.

Uno de los factores determinantes para que todo proceso se lleve a cabo con éxito, es implementar un sistema adecuado de indicadores para medir la gestión de los mismos, con el fin de que se puedan implementar indicadores en posiciones estratégicas que reflejen un resultado óptimo en el mediano y largo plazo, mediante un buen sistema de información que permita medir las diferentes etapas del proceso.

Es común, que los datos se midan a nivel de alguna agrupación. Por ejemplo, la cifra de ventas no se mide a nivel de cada acto de venta sino, por ejemplo, a nivel de 'Cifra de ventas mensual'. Para ello, partiendo de la Matriz de Datos de hechos de Venta, mediante una consulta (SQL) se extrae la información agregada necesaria.

La granularidad expresa el nivel de detalle de los datos. Una variable que sus valores se tomen diariamente tendrá más granularidad que si sus valores se toman mensualmente.

Los indicadores básicos, primitivos o crudos existen físicamente en el Warehouse junto a los atributos que los caracterizan, pueden venir de diferentes sistemas fuentes y tener distintos niveles de granularidad o agregación. Por ejemplo, la variable Venta (€) es llevada diariamente mientras que el indicador Unidades en Stock es seguido semanalmente.

Por otro lado, los indicadores derivados o métricas calculadas se construyen a partir de los indicadores base y pueden o no estar almacenados físicamente en el Data Warehouse (es típicamente una decisión de tunning, de ajuste fino). Un ejemplo clásico de métrica derivada es Margen de Ganancia, la cual se define como la resta entre Precio y Costo (ambos indicadores básicos).

Los indicadores son, por consiguiente, magnitudes asociadas a los procesos que permiten evaluarlos. Un proceso, por definición, tiene que generar un valor añadido alguno de los grupos de interés de la organización y, este valor, se tiene que poder medir.

Los indicadores son un conjunto de magnitudes que se tendrán que medir en el momento de ejecución y que permitirán determinar la eficiencia del proceso.

Los atributos (metadatos) básicos de un indicador son, de acuerdo con la norma UNE 66175:2003 Sistemas de gestión de la calidad. Guía para la implantación de sistemas de indicadores:

- La denominación
- La forma de cálculo
- La frecuencia de medida
- La definición de responsabilidades
- La definición de una meta o límites (opcional)

Un indicador numérico, como regla general, puede ser:

- Un valor (número de, cualquier magnitud, etc.)
- Un índice (tanto por ciento, por uno, etc.)
- Un ratio (cociente entre dos valores o una operación "compleja" entre varios indicadores).

Los indicadores deben tener una meta asignada. Una meta es un valor que un indicador debe tener.

Una manera útil de facilitar la gestión de los indicadores puede ser agrupándolos en tipos de indicadores. La propuesta en este trabajo es definir indicadores de:

a) cantidad
b) calidad
c) entrega
d) importe

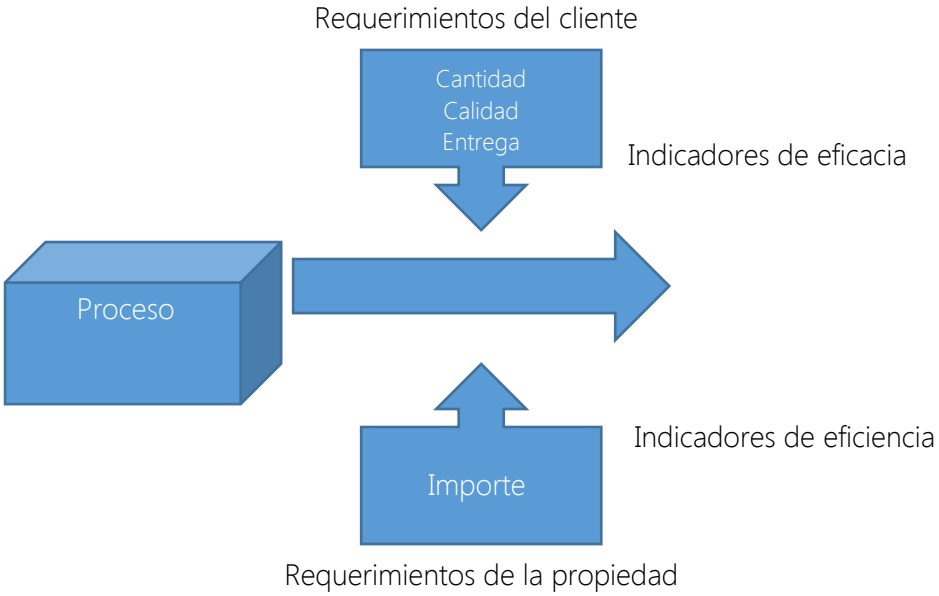


Ilustración 48. Sistema de Indicadores propuesto. Fuente: Brull Alabart (2014)

Otra manera como propone Kaplan (1997) es agrupándolos en indicadores de:

a. finanzas
b. clientes
c. procesos
d. formación y crecimiento

También podríamos aplicar una agrupación de acuerdo con criterio del modelo EFQM en indicadores de:

a. clientes
b. personas
c. sociedad
d. rendimiento clave

A modo de ejemplo mostramos los indicadores asociados al proceso "Definir el Plan Anual"

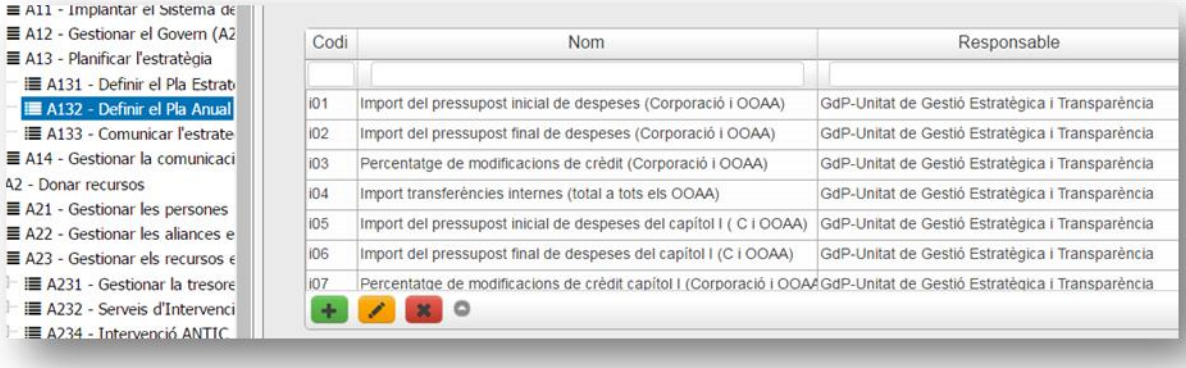


Ilustración 49. Ejemplo de indicadores asociados a un proceso. Fuente: Diputació de Tarragona (2016)

Las organizaciones excelentes se destacan por gestionar sus actividades y recursos mediante un conjunto de sistemas, procesos y datos que les permiten la toma eficaz de decisiones.

En este contexto, ser excelente implica, entre otros aspectos, llevar a cabo mediciones periódicas tanto de los enfoques aplicados como de su despliegue, con el fin de disponer de datos objetivos que sustenten el avance de la organización hacia unos buenos resultados en relación con sus grupos de interés.

El desarrollo de sistemas de indicadores para estructurar la medición en las organizaciones es sin duda necesario para asegurar la calidad de la información empleada en dicha toma de decisiones.

La norma UNE 66175:2003 ofrece una guía para la implantación de este tipo de sistemas, facilitando la identificación y el establecimiento de los indicadores relevantes y cuadros de mando de la organización.

3 Propuesta de estructura del sistema de datos

A continuación se muestra una visión general de algunos Data Marts y tablas de un Data Warehouse que se propone como modelo, así como de sus relaciones.

Data mart		f_venta	f_compra	f_contrato	f_plantilla	f_formación	f_planificación	f_costes	f_presencia	f_rendimiento	f_macroentorno
Clientes											
	d_cliente	x					x	x			
Personas											
	d_persona	x			x	x	x	x	x	x	
	d_puesto				x				x		
	d_actividadformativa					x					
Procesos / Operaciones											
	d_servicio_output	x		x			x	x			
	d_producto	x	x				x				
	d_servicio_input			x			x				
Estrategia											
	d_objetivos						x			x	
	d_proyecto						x			x	
	d_tarea						x				
Gestión del conocimiento y TI											
	d_fecha	x	x	x	x	x	x	x	x	x	x
	d_geografia	x					x				
	d_inventario		x				x				
	d_contabilidad	x	x	x			x	x			
Sociedad											
	d_variable_sociedad										x
Proveedores											
	d_proveedor		x	x			x				

Ilustración 50. Relaciones entre tablas f, tablas d y data marts. Fuente: Elaboración propia (2015)

La finalidad de disponer de esquemas de datos es permitir el acceso de herramientas para aplicar aquellas técnicas de análisis de datos que se requieran en cada caso.

Se puede obtener la Matriz de Datos de un Esquema con la agrupación –de manera adecuada- de todas las tablas que configuran un Esquema.

Se muestra un ejemplo de una Matriz de Datos del Data Mart de ventas. Cada registro de la Matriz de Datos contiene un conjunto de actos de venta relacionados y que la caracterizan: Datos de Tiempo, de Producto, de Cliente,… y métricas de importe de la venta, cantidad,….

Producte		Mercat		Venedor		Geografia				Temps				Vendes
Familia	**Item**	**Sector**	**Client**	**Venedor**	**Edad**	**CA**	**Provincia**	**Comarca**	**Municipi**	**Any**	**Trimestre**	**Mes**	**Dia**	**Vendes**
Mecanica	mpx-435	Quimic	BCN, S.A.	Joan Ramió	45	Catalunya	Barcelona	Barcelones		2005	4	11	03/11/2005	23.456
Electrica	egq-001	Energetic	Pujol, S.L.	Pere Estiles	38	Catalunya	Tarragona	Tarragones		2005	4	11	10/11/2005	45.789
Electrònica	tyu-385	Mobles	Rubin, S.A.	Vicent Puig	56	Valencia	Castelló	Maestrat		2005	4	12	07/12/2005	12.678
Construcció	cgq-001	Construcció	Calzada	Joan Ramió	45	Catalunya	Tarragona	Baix Camp		2005	4	12	10/12/2005	9.876
Electrònica	tyu-385	Alimentació	Rubin, S.A.	Vicent Puig	56	Valencia	Castelló	Maestrat		2005	4	12	10/12/2005	45.900
Mecanica	mpx-435	Construcció	AIMA, S.L.	Enric Gatell	23	Aragon	Huesca	Cariñena		2006	1	1	08/01/2006	5.678
Electrònica	tyu-386	Sabates	Ristra, S.A.	Josep Torruella	30	Catalunya	Tarragona	Tarragones		2006	1	1	15/01/2006	3.167
Construcció	cgq-001	Sabates	Calzada	Joan Ramió	45	Catalunya	Tarragona	Alt Camp		2006	1	1	20/01/2006	9.876
Mecanica	mpx-435	Consultoria	Padró, S.L.	Enric Gatell	23	Catalunya	Barcelona	Barcelones		2006	2	4	20/04/2006	6.789
Construcció	cgq-001	Mobles	Rubin, S.A.	Vicent Puig	56	Valencia	Castelló	El Puig		2006	2	4	22/04/2006	45.908
Electrònica	tyu-386	Construcció	AIMA, S.L.	Pilar Cachirulo	29	Aragon	Huesca	Cariñena		2006	2	4	24/04/2006	23.589
Mecanica	mpx-435	Quimic	Nurel	Pilar Cachirulo	29	Aragon	Zaragoza	Moncayo		2006	2	4	30/04/2006	42.378
Construcció	cgq-001	Turisme	Hotels ACE	Vicent Puig	56	Valencia	Valencia	La Plana		2006	2	5	02/05/2006	26.789
Electrica	egq-002	Oci	Eclectic	Pilar Cachirulo	29	Aragon	Teruel	Calanda		2006	2	5	20/05/2006	42.378
Construcció	cgq-002	Turisme	Benidorm	Vicent Puig	56	Valencia	Alacant	Ravalet		2006	2	5	28/05/2006	45.908
Construcció	cgq-002	Alimentació	Fruites	Pere Estiles	41	Catalunya	Lleida	Segria		2006	2	6	02/06/2006	33.222
Electrònica	tyu-386	Turisme	Turistics	Narcis Olot	42	Catalunya	Girona	L'Emporda		2006	2	6	05/06/2006	67.900
Electrica	egq-001	Quimic	BCN, S.A.	Enric Gatell	23	Catalunya	Barcelona	Baix Llobregat		2006	3	7	07/07/2006	36.700
Electrica	egq-001	Construcció	Calzada	Joan Ramió	45	Catalunya	Tarragona	Tarragones		2006	4	10	30/10/2006	9.876

Ilustración 51. Tabla matriz de datos de ejemplo. Fuente: Elaboración propia (2014)

Un adecuado diseño del Data Warehouse es fundamental para que el modelo aporte la información necesaria en cada momento. Debe dedicarse gran parte del tiempo de un proyecto de implantación de un sistema como el que se propone al diseño del Data Warehouse. De no ser así, al llegar al final de la implantación, en el proceso de publicación de los informes, podemos echar en falta información. Si el sistema de datos está bien estructurado, ello no comportará más tarea que la de añadir el dato en la tabla de hechos o dimensión adecuada, del Data Mart correspondiente.

3. Dimensión Análisis

1 Técnicas analíticas 96
1.1 Análisis univariable 103
1.1.1 Serie temporal 104
1.1.2 Medidas de centralidad 105
1.1.3 Medidas de dispersión 106
1.1.4 Medidas de distribución 107
1.2 Análisis bivariable 108
1.2.1 Correlación 108
1.2.2 Regresión 109
1.2.3 Tablas de contingencia 109
1.3 Análisis multivariable 110
1.3.1 Análisis factorial 110
1.3.2 Análisis de correspondencias 111
1.3.3 Análisis de clúster 112
1.3.4 Análisis discriminante 112
1.3.5 Análisis de regresión múltiple 113
1.3.6 Análisis estructural 113
1.4 Data Mining 114
2 Selección del tipo de análisis 118

1 Técnicas analíticas

A continuación se expone la Dimensión Análisis del modelo, que se conforma mediante un catálogo de Técnicas de análisis.

Los datos obtenidos de la fase de Medida (ver capítulo anterior Dimensión Datos), deben ser analizados mediante las técnicas estadísticas y las herramientas adecuadas.

Según la finalidad del análisis, habrá unas herramientas de análisis más adecuadas que otras. Hay una gran cantidad de ellas; desde muy elementales hasta muy sofisticadas.

No es el objetivo de este trabajo hacer una descripción exhaustiva de los métodos estadísticos existentes, sino la de dar una información elemental sobre las técnicas de análisis básicas que toda organización puede utilizar.

Dado que la finalidad de los indicadores es conocer la eficacia y la capacidad asociada a los procesos, es importante en este punto introducir estos dos conceptos:

- La eficacia del proceso se refiere a si los resultados son adecuados o suficientes para alcanzar los resultados planificados.
- La capacidad de un proceso está referida a la aptitud para cumplir con unos determinados requisitos.

Cuando el número de datos es alto, se hace necesario la implantación de herramientas de estadísticas informatizadas que permitan los distintos análisis univariables, bivariables y multivariables.

Las técnicas estadísticas son herramientas de apoyo para medir, describir, analizar, interpretar y modelar la variación y para resolver y prevenir respecto a los orígenes de dicha variación. Estas herramientas nos permiten obtener conclusiones a partir de la investigación empírica usando modelos de estructura matemática.

Para la interpretación del modelo SIEGO, es sustancial destacar que todas la técnicas analíticas tienen una representación gráfica recomendada o relativa, pero que la técnica analítica en sí no es una representación gráfica. Podemos ver los diferentes componentes de representación visual en el capítulo siguiente, Componentes de visualización. Asimismo, siempre será mucho más entendible una técnica analítica si se muestra gráficamente representada.

La norma ISO/TR 10017:2003, "Guía de técnicas estadísticas para ISO9001:2000", se desarrolló para apoyar a los responsables en sus procesos de toma de decisiones, identificando las técnicas estadísticas que pueden ayudar a mejorar la calidad de su sistema de gestión, productos y procesos.

TÉCNICA O FAMILIA DE TÉCNICAS ESTADÍSTICAS	DESCRIPCIÓN
Estadística Descriptiva	Método para presentar la distribución y asociación de la data cuantificable. Ejemplo: Histogramas, Cuadros de Tendencia, Cuadros de Dispersión.
Diseño de Experimentos	Método de investigación a través de la experimentación y presentación de las caracterización de entrada y los resultados, mediante modelos matemáticos. Ejemplo: Cuadro de Probabilidades, Cuadro de Análisis de Varianza.
Testeo de Hipótesis o Pruebas de Hipótesis	Método para determinar la relación entre un conjunto de data y las hipótesis asociadas.
Análisis de Incerteza de Mediciones	Método para establecer los parámetros de incerteza de los sistemas de medición.
Análisis de Capacidad de Proceso	Método para examinar la variabilidad y su distribución, de un proceso, con la finalidad de establecer su nivel de capacidad para producir los resultados esperados.
Análisis de Regresión	Método para establecer la causa de la variación y los factores determinantes.
Análisis de Confiabilidad	Método para estudiar la probabilidad de fallas y establecer el grado de confianza para lograr los resultados esperados.
Muestreo	Método para establecer el comportamiento de una muestra representativa, que nos brinde información sobre la caracterización de aquello de donde se obtiene la muestra.
Simulación	Método para establecer matemáticamente el comportamiento de un ente o elemento y poderlo estudiar.
Cuadros de Control Estadístico de Procesos	Método para ordenar data y establecer los límites del proceso de manera gráfica. Se usa generalmente para evaluar el grado de estabilidad de los procesos.
Tolerancia Estadística	Método para establecer la tolerancia mediante modelos matemáticos.
Análisis de Serie Tiempo	Familia de métodos para establecer la relación data-tiempo. Se usa generalmente para predecir valores y tendencias en el tiempo.

Ilustración 52. Técnica o familia de técnicas estadísticas. Fuente: Norma ISO 10017:2003.

La mejora es el foco de los sistemas de gestión, para lograr la satisfacción de las partes interesadas y generar el desarrollo sostenible de la organización. Aproximadamente los costos de no calidad representan el equivalente a un 15 a 40% de los ingresos de una organización. Contentarse con operar con una tasa de aciertos (conformidad) de sólo 94% equivale a tener costos de no calidad que pueden representar el 25% de los ingresos.

Para ser más eficientes y eliminar los costos de no calidad, las organizaciones deben controlar y mejorar sus sistemas y procesos, mediante el uso de técnicas estadísticas que les permita obtener rentabilidad, social y económica.

Para entender mejor el uso de las técnicas estadísticas a continuación se presenta una tabla basada en la guía ISO/TR 10017. Se puede utilizar el cuadro para determinar qué técnicas

estadísticas se usarán en la organización para obtener y evaluar información que permita dar cumplimiento a los requisitos exigidos en las diferentes cláusulas de la Norma ISO 9001:2000.

Tabla basada en el documento ISO/TR 10017:2003

"Necesidades relacionadas con los datos cuantitativos y las técnicas estadísticas de apoyo, por cláusulas de la Norma ISO 9001:2000"

Capítulo/apartado ISO 9001:2000	Necesidades	Técnicas estadísticas
4 Sistema de Gestión de la Calidad 4.1 Requisitos generales	Identificadas en diversos apartados de la norma.	
4.2 Requisitos de la documentación 4.2.1 Generalidades	Ninguna identificada	
4.2.2 Manual de la calidad	Ninguna identificada	
4.2.3 Control de los documentos	Ninguna identificada	
4.2.4 Control de los registros	Ninguna identificada	

Capítulo/apartado ISO 9001:2000	Necesidades	Técnicas estadísticas
5 Responsabilidad de la dirección 5.1 Compromiso de la dirección	Ninguna identificada	
5.2 Enfoque al cliente	Determinar los requisitos del cliente Evaluar la satisfacción del cliente	Véase 7.2.2. en esta tabla Véase 8.2.1 en esta tabla
5.3 Política de la calidad	Ninguna identificada	
5.4 Planificación 5.4.1 Objetivos de la calidad	Ninguna identificada	
5.4.2 Planificación del sgc	Ninguna identificada	
5.5 Responsabilidad, autoridad y comunicación 5.5.1 Responsabilidad y autoridad	Ninguna identificada	
5.5.2 Representante de la dirección	Ninguna identificada	
5.5.3 Comunicación interna	Ninguna identificada	
5.6 Revisión por la dirección 5.6.1 Generalidades	Ninguna identificada	
5.6.2 Información para la revisión a) resultados de auditorías	Evaluar los datos de la auditoria	Estadística descriptiva; muestreo
b) retroalimentación del cliente	Obtener y evaluar la retroalimentación del cliente	Estadística descriptiva; muestreo
c) desempeño de los procesos y conformidad del producto	Evaluar el desempeño del proceso y conformidad del producto	Estadística descriptiva; análisis de la capacidad del proceso; muestreo; gráficos de control estadístico de procesos (CEP)
d) estado de las acciones correctivas y preventivas	Obtener y evaluar datos para las acciones preventivas y correctivas	Estadística descriptiva

5.6.3 Resultados de la revisión	Ninguna identificada	

Capítulo/apartado ISO 9001:2000	Necesidades	Técnicas estadísticas
6 Gestión de los recursos 6.1 Provisión de recursos	Ninguna identificada	
6.2 Recursos humanos 6.2.1 Generalidades	Ninguna identificada	
6.2.2 Competencia, toma de conciencia y formación 6.2.2 a)	Ninguna identificada	
6.2.2 b)	Ninguna identificada	
6.2.2 c) evaluar la eficacia de las acciones tomadas	Evaluar la competencia y eficacia de la formación	Estadística descriptiva; muestreo
6.2.2 d)	Ninguna identificada	
6.2.2 e)	Ninguna identificada	
6.3 Infraestructura	Ninguna identificada	
6.4 Ambiente de trabajo	Hacer seguimiento del ambiente de trabajo	Estadística descriptiva; gráficos de CEP

Capítulo/apartado ISO 9001:2000	Necesidades	Técnicas estadísticas
7 Realización del producto 7.1 Planificación de la realización del producto	Ninguna identificada	
7.2 Procesos relacionados con el cliente 7.2.1 Determinación de los requisitos relacionados con el producto	Ninguna identificada	
7.2.2 Revisión de los requisitos relacionados con el producto	Evaluar la capacidad de la organización para cumplir los requisitos definidos	Estadística descriptiva; análisis de la medición; análisis de la capacidad del proceso; muestreo; fijación de tolerancia estadística
7.2.3 Comunicación con el cliente	Ninguna identificada	
7.3 Diseño y desarrollo 7.3.1 Planificación del diseño y desarrollo	Ninguna identificada	
7.3.2 Elementos de entrada para el diseño y desarrollo	Ninguna identificada	
7.3.3 Resultados del diseño y desarrollo	Verificar que el resultado del diseño satisface los requisitos de los elementos de entrada	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de regresión; análisis de confiabilidad; muestreo; simulación; análisis de series de tiempo
7.3.4 Revisión del diseño y desarrollo	Ninguna identificada	
7.3.5 Verificación del diseño y desarrollo	Verificar que los resultados del diseño satisfacen los requisitos de los elementos de entrada	Estadística descriptiva; diseño de Experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del

		proceso; análisis de regresión; análisis de confiabilidad; muestreo; simulación; análisis de series de tiempo
7.3.6 Validación del diseño y desarrollo	Validar que el producto cumple el uso y las necesidades establecidas	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo; simulación
7.3.7 Control de los cambios del diseño y desarrollo	Evaluar, verificar y validar el efecto de los cambios del diseño	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo; simulación
7.4 Compras 7.4.1 Proceso de compras	Asegurarse de que el producto adquirido cumple los requisitos de compra especificados Evaluar la capacidad de los proveedores para suministrar productos que cumplan con los requisitos de las organizaciones	Estadística descriptiva; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo Estadística descriptiva; diseño de experimentos; análisis de la capacidad del proceso; análisis de regresión; muestreo
7.4.2 Información de las compras	Ninguna identificada	
7.4.3 Verificación de los productos comprados	Establecer e implementar la inspección y otras actividades necesarias para asegurarse de que el producto comprado cumple los requisitos especificados	Estadística descriptiva; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; muestreo
7.5 Producción y prestación del servicio 7.5.1 Control de la producción y de la prestación del servicio	Hacer seguimiento y control de la actividad de producción y servicio	Estadística descriptiva; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo; gráficos de CEP; análisis de series de tiempo
7.5.2 Validación de los procesos de la producción y de la prestación del servicio	Validar, hacer seguimiento y controlar los procesos cuyos resultados no pueden ser medidos fácilmente	Estadística descriptiva; análisis de la capacidad del proceso; análisis de regresión; muestreo; gráficos de CEP; análisis de series de tiempo
7.5.3 Identificación y trazabilidad	Ninguna identificada	
7.5.4 Propiedad del cliente	Verificar las características de la propiedad del cliente	Estadística descriptiva; muestreo
7.5.5 Preservación del producto	Hacer seguimiento del efecto de la manipulación, el embalaje y el almacenamiento sobre la calidad del producto	Estadística descriptiva; análisis de regresión; análisis de confiabilidad; muestreo; gráficos de CEP; análisis de series de tiempo

7.6 Control de los dispositivos de seguimiento y de medición	Asegurar que el seguimiento y la medición del proceso y el equipo es coherente con los requisitos Evaluar la validez de las mediciones previas , cuando se requiera	Estadística descriptiva; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; muestreo; gráficos de CEP; fijación de tolerancias estadísticas; análisis de series de tiempo Estadística descriptiva; prueba de hipótesis; análisis de la medición; análisis de regresión; muestreo; fijación de tolerancias estadísticas; análisis de series de tiempo
Capítulo/apartado ISO 9001:2000	Necesidades	Técnicas estadísticas
8 Medición, análisis y mejora 8.1 Generalidades	Ninguna identificada	
8.2 Seguimiento y medición 8.2.1 Satisfacción del cliente	Hacer seguimiento y analizar la información relacionada con la percepción del cliente	Estadística descriptiva; muestreo
8.2.2 Auditoría interna	Planificar el programa de auditoría interna y el informe de los datos de auditoría	Estadística descriptiva; muestreo
8.2.3 Seguimiento y medición de los procesos	Hacer seguimiento y medición de los procesos de gestión de la calidad, para demostrar la capacidad del proceso de alcanzar los resultados planificados	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; muestreo; gráficos de CEP; análisis de series de tiempo
8.2.4 Seguimiento y medición del producto	Hacer seguimiento y medir las características del producto en las etapas apropiadas de realización, para verificar que los requisitos se cumplen	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo; gráficos de CEP; análisis de series de tiempo
8.3 Control del producto no conforme	Determinar la cantidad de productos no conformes entregados Una nueva verificación para asegurar su conformidad	Estadística descriptiva; muestreo Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la medición; análisis de la capacidad del proceso; análisis de regresión; análisis de confiabilidad; muestreo; gráficos de CEP; análisis de series de tiempo
8.4 Análisis de datos	Obtener y analizar los datos para evaluar la eficacia del sistema de gestión de la calidad, y para evaluar las posibilidades de	Véase 8.2.1 en esta tabla

	mejora, con relación a La satisfacción del cliente La conformidad con los requisitos del producto Las características del proceso y tendencias Los proveedores	Véase 8.2.4 en esta tabla Véase 8.2.3 en esta tabla Véase 7.4.1 en esta tabla
8.5 Mejora 8.5.1 Mejora continua	Mejorar los procesos del sistema de gestión de la calidad a través del uso de los datos cuantitativos, en las áreas de: - diseño y desarrollo - compras - producción y prestación del servicio - control del seguimiento y medición de los procesos	 Véanse 7.3.3, 7.3.5 y 7.3.6 en esta tabla Véanse 7.4.1 y 7.4.3 en esta tabla Véanse 7.5.1, 7.5.2 y 7.5.5 en esta tabla Véase 7.6 en esta tabla
8.5.2 Acción correctiva	Analizar los datos relacionados con las no conformidades para ayudar a entender sus causas	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la capacidad del proceso; análisis de regresión; muestreo; gráficos de CEP; análisis de series de tiempo
8.5.3 Acción preventiva	Analizar los datos relacionados con las no conformidades y no conformidades potenciales para ayudar a entender sus causas.	Estadística descriptiva; diseño de experimentos; prueba de hipótesis; análisis de la capacidad del proceso; análisis de regresión; muestreo; gráficos de CEP; análisis de series de tiempo

De todas las técnicas de análisis que se presentan en la norma anterior, en este trabajo nos centraremos en el análisis de datos mediante técnicas descriptivas.

El análisis estadístico es el más utilizado para el análisis numérico de datos y en este trabajo se presenta mediante el planteamiento del siguiente Catálogo de Técnicas de Análisis. El objeto de este diccionario es el de explicar básicamente cada una de las técnicas analíticas, así como indicar para qué tipo de análisis es adecuado para cada tipo de dato.

Catálogo de Técnicas de Análisis

Análisis univariable

- Serie temporal
  - Contar
- Medidas de centralidad
  - Media
  - Mediana

- Moda
- Medidas de dispersión
  - Varianza
  - Desviación estándar
  - Máximo
  - Mínimo
  - Rango
  - Rango intercuartílico
- Medidas de distribución
  - Frecuencia
  - Asimetría o sesgo
  - Curtosis o apuntamiento

Análisis bivariable

- Correlación
- Regresión
- Tablas de contingencia (2 variables)

Análisis multivariable

- Tabla de contingencia (> de 2 variables)
- Regresión lineal múltiple
- Análisis factorial
- Análisis de componentes principales
- Análisis de correspondencias
- Análisis discriminante (o clasificación DM)
- Clasificación jerárquica (o clúster DM)
- Análisis estructural

Ilustración 53. Catálogo de Técnicas de Análisis. Fuente: Elaboración propia.

Además, en el caso de querer relacionar los resultados obtenidos de una muestra sobre una población, será necesario considerar los contrastes de cada una de las técnicas relacionadas.

Una organización que quiera implantar un sistema como el que se propone en este trabajo, debería dominar como mínimo las técnicas estadísticas básicas (univariables) y dar formación y acompañamiento a las personas de su organización para comprenderlas apropiadamente y poder generar el conocimiento adecuado.

1.1 Análisis univariable

Permite conocer el comportamiento de una variable aislada. Estas técnicas buscan encontrar y representar características propias de las mismas, aunque suelen ser la primera fase del análisis que luego se realiza con técnicas bivariables o multivariables. Para el análisis univariable se utilizan 4 tipos básicos de técnicas.

1.1.1 Serie temporal

Permiten visualizar los datos sobre un eje x temporal. Con las herramientas adecuadas se pueden calcular si existe periodicidad y estacionalidad en la serie.

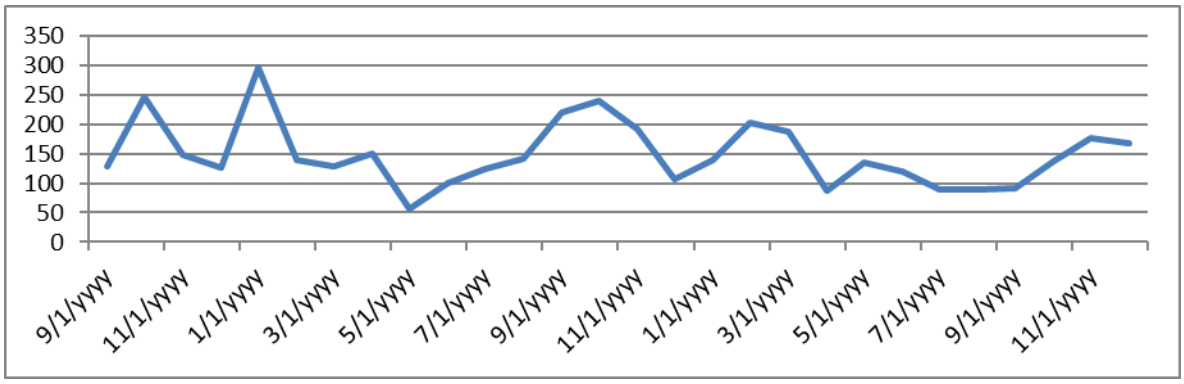


Ilustración 54. Serie temporal. Fuente: Elaboración propia

La TAM (Tendencia o Total Anual Móvil, o más genéricamente, medias móviles) permite ver la serie temporal mediante los valores acumulados del periodo. Mostramos un ejemplo con la media móvil de doce meses.

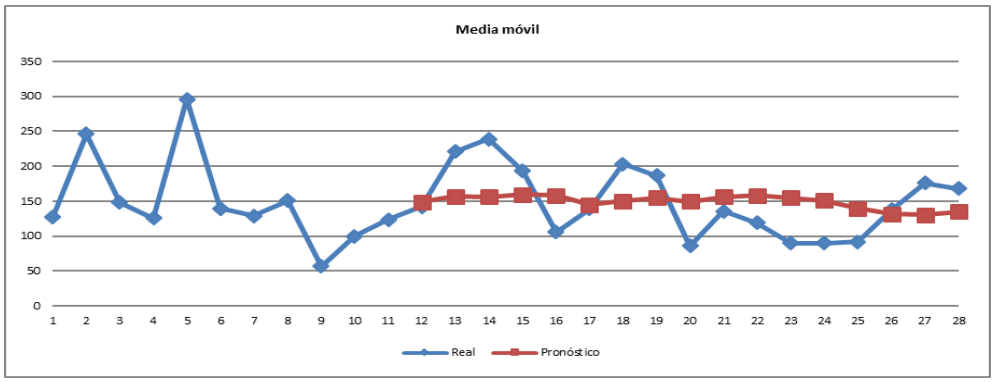


Ilustración 55. Media móvil. Fuente: Elaboración propia.

Contar: Es el valor resultante de sumar el número de elementos. [Excel: contar()]

Contar	30

Aunque la función contar sea muy simple, en general un primer análisis de los datos puede ser difícil de analizar a simple vista. Un simple cuenteo de los datos nos puede dar una primera información de la evolución de una serie, que no parece evidente en la tabla de datos.

Fecha	nClientes	Satisfacción	Temps de cua
30/09/2004	128	56,7	3,84
31/10/2004	246	90,3	6
30/11/2004	148	76,9	4,44
31/12/2004	126	34,1	3,78
31/01/2005	296	55	8,88
28/02/2005	140	54,5	4,2
31/03/2005	129	88,9	3,87
30/04/2005	151	78,5	5
31/05/2005	57	76,9	1,71
30/06/2005	100	76,2	3
31/07/2005	124	56,9	3,72
31/08/2005	142	73,6	4,26
30/09/2005	221	45,6	6,63
31/10/2005	239	56,8	7,17
30/11/2005	193	78,5	5,79
31/12/2005	106	75,9	3,5
31/01/2006	139	74,8	4,17
28/02/2006	203	76,3	6,09
31/03/2006	187	56	5,61
30/04/2006	86	44,6	3
31/05/2006	136	79	4,08
30/06/2006	119	86,7	3,57
31/07/2006	90	45,8	3
31/08/2006	90	33,6	2,7
30/09/2006	92	65,9	2,76
31/10/2006	138	69,2	4,5
30/11/2006	176	67,1	5,28
31/12/2006	168	75,9	5,04

Indicadores agrupados 2005:
Número de clientes = 1.898
Media mensual= 158,16

Indicadores agrupados 2006:
Número de clientes = 1.624
Media mensual= 135,33

Ilustración 56. Tabla ejemplo evolución de una serie. Fuente: Elaboración propia.

1.1.2 Medidas de centralidad

Dan una idea del centro de la distribución de los datos. Las principales medidas de este tipo son la media o promedio aritmético, la mediana y la moda.

Media aritmética: Es el valor obtenido de la suma de los valores dividida por el número de valores. [Excel: promedio()].

Las medias tienen un problema fundamental: Si una sección tiene unas ventas de 50 unidades, y otra sección no ha vendido, el promedio o media es de 25. Podría darse el caso que el objetivo de ventas fuera de 22, por ejemplo, y no ser visible que la sección primera está teniendo unas ventas excepcionales y que la segunda está muy lejos de cumplir los objetivos. Por ello, la media no es una técnica estadística recomendable si no se acompaña de más información.

Media	148,9286

Mediana: Es el valor del elemento central cuando los datos están ordenados. [Excel: mediana()]

Mediana	138,5

Moda: Es el valor que se repite un mayor número de veces. [Excel: moda()]

Moda	90

1.1.3 Medidas de dispersión

Expresan el grado de concentración o dispersión de los datos con respecto al centro de la distribución. Entre las principales medidas de este tipo están la varianza, la desviación estándar, el rango intercuartílico. Aparte también hay medidas de posición, como son los cuartiles, deciles y percentiles.

Varianza: Es la suma de los cuadrados de las desviaciones de los valores respecto a su media, dividida por el número de valores. [Excel: var()] Indica en qué medida están dispersos los datos en una distribución. A mayor medida, mayor dispersión. Habitualmente es un número muy alto, y por tanto se vuelve difícil de interpretar.

Varianza	3037,106

Desviación estándar. Se calcula con la raíz cuadrada de la varianza. [Excel: desvest()].Para solucionar las complicaciones de comprensión de la varianza, se halla la raíz cuadrada de la misma, es decir, se calcula la desviación estándar, la cual es un número pequeño expresado en las mismas unidades de los datos originales y que tiene un significado lógico respeto a los mismos. A pesar de lo anterior, es difícil describir exactamente qué es lo que mide la desviación estándar. Sin embargo, hay un resultado útil, que lleva el nombre del matemático ruso Pafnuty Lvovich Chebyshev, y se aplica a todos los conjuntos de datos. Este teorema de Chebyshev establece que para todo conjunto de datos, por lo menos 1- 1/k2 de las observaciones están dentro de k desviaciones estándar de la media, en donde k es cualquier número mayor que 1. Este teorema se expresa de la siguiente manera:

$$1 - \frac{1}{k^2}$$

Ecuación 1. Teorema de Chebyshev,

Así por ejemplo, si se forma una distribución de datos con k=3 desviaciones estándar por debajo de la media hasta 3 desviaciones estándar por encima de la media, entonces por lo menos

$$1 - \frac{1}{3^2} = 1 - \frac{1}{9} = \frac{9-1}{9} = \frac{8}{9} = 0{,}8889 = 88{,}89\%$$

Ecuación 2. Cálculo del Teorema de Chebyshev con k=3

Interpretación: El 88,89% de todas las observaciones estarán dentro ± 3 desviaciones de la media.

DesvEst	55,10994

Rango o recorrido: Es la diferencia entre el valor mayor o máximo y el valor menor o mínimo. [Excel: max(), min()].

Max	296
Min	57
Rango	239

Rango intercuartílico: la diferencia entre el tercer y el primer cuartil, nos permite conocer cuántas unidades de los valores que toma la variable se concentra el cincuenta por ciento central de los casos. Para ello, determinamos los valores que dividen el recorrido de datos en k partes de la misma frecuencia (percentiles: 100 partes, cuartiles: 4 partes, etc.).

1.1.4 Medidas de distribución

La tabla de frecuencias representa la frecuencia de cada intervalo.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Menos de 5 minutos	0	21	20,79%
Entre 5 minutos y 60 minutos (1 hora)	1	44	43,56%
Entre 1 hora y 8 horas (1 día)	2	19	18,81%
Entre 1 día y 5 días (1 semana)	3	11	10,89%
Entre 1 semana y 4 semanas (1 mes)	4	4	3,96%
No puedo responder al 95% de las peticiones de información	5	2	1,98%
		101	100,00%

La tabla de frecuencias la podemos representar con el histograma, que permite visualizar la frecuencia de cada valor o categoría de valores. Es útil para poder aplicar algunos cálculos; por ejemplo, cuando la distribución es normal.

Cuando se analiza la distribución de una variable, suele realizarse el análisis de la asimetría, el cual puede dar información al gestor en función de la naturaleza de los datos:

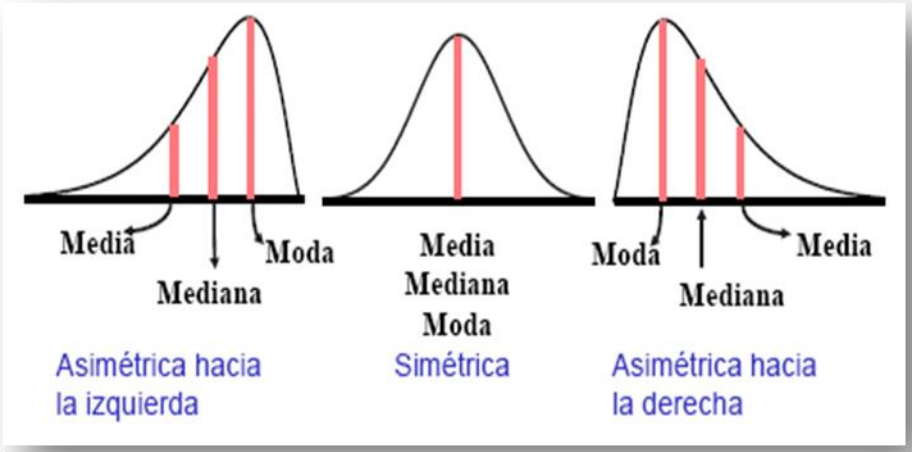


Ilustración 57. Simetrias y asimetrias.

También suele analizarse con la distribución el índice de curtosis, que también puede dar información al gestor en función del tipo de datos analizados.

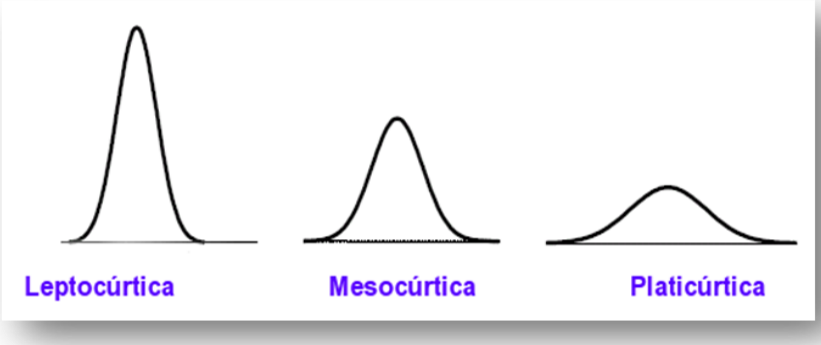


Ilustración 58. Curtosis

1.2 Análisis bivariable

Permite conocer la relación entre dos variables. Para el análisis bivariable se utilizan, básicamente, tres tipos básicos de técnicas.

1.2.1 Correlación

Se utiliza un coeficiente, por ejemplo, el de correlación de Pearson [Excel: coef.de.correl()], que da un valor entre -1 y +1, y que caracteriza la magnitud de la relación entre las dos variables.

	Columna 1	*Columna 2*
Columna 1	1	
Columna 2	0,08014761	1

La representación gráfica de las dos variables y de sus relaciones, muestra el conocido diagrama de dispersión. (Ver apartado xxx de Visualización)

1.2.2 Regresión

Se trata de obtener una fórmula que relacione una o varias variables (independientes) con una variable (dependiente).

Ver apartado Estudio empírico.

1.2.3 Tablas de contingencia

Se trata de obtener una tabla donde hay una variable en filas y una variable en columnas. En las celdas de cruce se calcula el número de ocurrencias entre las dos variables. Las herramientas OLAP (dentro de la tecnología de Business Intelligence) así como las Tablas dinámicas de Excel, permiten construir tablas de contingencia de forma ágil. Ver apartado Dimensión Informes.

			#null	BAS	CIQ	GDP	INT	PDT	RHP	SAC
Servei_Area	Servei_Categoria	Servei_Nom	Quantitat (Σ)	Quantitat (Σ)	Quantitat (Σ)	Quantitat (Σ)	Quantitat (Σ)	Quantitat (Σ)	Quantitat (Σ)	Quanti (Σ)
CIQ	Altres serveis	Suport Informàtic			50					
	Aplicacions de Gestió	Administració de Bases de Dades	4		127	1			1	
	CAU	Assistent creació tíquets	9		3	2	2	2	3	
		Emissió d'informes	4		29	5		5	7	
		Identitats usuaris	33	59	24	9	9	6	645	
		Integritat/Seguretat de la informació	55	5	41	28	47	14	98	
		Petició aprovisionament maquinari del lloc de treball	78		99	44	13	12	112	1
		Petició confecció formularis PDF	4		2	7	3	23	10	
		Petició d'accés a aplicacions	95	13	111	64	74	20	194	1
		Petició de Comunicacions	148	11	322	80	39	81	137	1
		Petició de Sistemes	22	1	83	13	10	9	19	
		Petició de canvi en aplicacions	86	1	46	36	111	18	96	
		Petició de consumible	66		3	13	14	5	37	
		Petició de programari de treball	110	2	74	59	50	26	181	1
		Petició usuaris ALTANET	34	1	1.093				1	
		Resolució Incidències Serveis	3		469					
		Resolució incidències amb aplicacions de gestió	151	27	101	52	180	30	267	1
		Resolució incidències de Comunicacions	24	6	76	20	5	10	25	
		Resolució incidències de Sistemes	7	2	5	3	16	8	18	
		Resolució incidències maquinari de lloc de treball	190		75	71	57	46	163	3
		Resolució incidències programari de treball	161	2	102	78	114	69	137	2
		Resolució incidències usuaris Altanet	11		95				1	
	ISO 20.000	Disseny i Transició de Serveis nous o modificats								
		Gestió de Problemes	2		2				3	
		Gestió de la Entrega i el Desplegament	1		42					
		Relacions amb el Negoci	2		5					
	Infraestructures	Petició General de Sistemes	4		30	1	2	3	3	
	Sistema de Gestió	Canvis en condicions d'entrega d'un servei	1							
		Millora-Projectes			39	12			3	
		Millora-Reptes			26	4			6	
		No Conformitats			28					
		Petició d'informe			15	7			5	
		Vigilància Tecnológica			21					
		Grand Total	1,305	130	3,238	609	746	387	2,172	1,6

Ilustración 59. Ejemplo tabla de contingencia bivariable. Fuente: Diputación de Tarragona (2016)

1.3 Análisis multivariable

Delante de una cantidad apreciable de datos estructurados en forma de tabla (por ejemplo, variables en columnas e individuos en filas), además de los análisis univariables o bivariables comentados, puede ayudarnos conocer el nivel de agrupación que podamos hacer de filas o de columnas. En líneas generales la agrupación de filas la podemos hacer con técnicas de análisis de clúster y la agrupación de columnas la podemos hacer con técnicas de análisis factorial.

1.3.1 Análisis factorial

Es el tipo de análisis que pone énfasis en la interrelación entre las variables. Podemos destacar:

El *análisis factorial exploratorio*, es el análisis que trata de evidenciar la estructura de un número relativamente grande de variables. La hipótesis *a priori* del investigador es que pueden existir una serie de factores asociados a grupos de variables. Las *cargas* de los distintos factores se utilizan para intuir la relación de éstos con las distintas variables. Es el tipo de análisis factorial más común.

El *análisis factorial confirmatorio*, intenta determinar si el número de factores obtenidos y sus *cargas* se corresponden con los que cabría esperar a la luz de una teoría previa acerca de los datos. La hipótesis *a priori* es que existen unos determinados factores preestablecidos y que cada uno de ellos está asociado con un determinado subconjunto de las variables. El análisis factorial confirmatorio entonces arroja un nivel de confianza para poder aceptar o rechazar dicha hipótesis.

1.3.2 Análisis de correspondencias

En estadística multivariante, el análisis de correspondencias es una técnica descriptiva. Suele aplicarse al estudio de tablas de contingencia y es conceptualmente similar al análisis de componentes principales con la diferencia de que en el análisis de correspondencias los datos se escalan de modo que filas y columnas se tratan de modo equivalente.

El análisis de correspondencias descompone el estadístico del test de la ji-cuadrado asociado a una tabla de contingencia en componentes ortogonales. Dado que se trata de una técnica descriptiva, puede aplicarse incluso en circunstancias en las que la prueba anterior no es apropiada.

Ilustración 60. Análisis de correspondencias. Fuente: https://israelcruzcorona.wordpress.com/2014/03/05/analisis-factorial-de-correspondencias/

1.3.3 Análisis de clúster

Es el análisis de la interrelación entre los sujetos analizados.

Análisis tipológico. Dado un conjunto de individuos se busca una partición de este de tal manera que se consiga que:

- Los grupos sean tan homogéneos como sea posible en su interior, similares
- Los grupos sean tan diferentes como sea posible con los otros

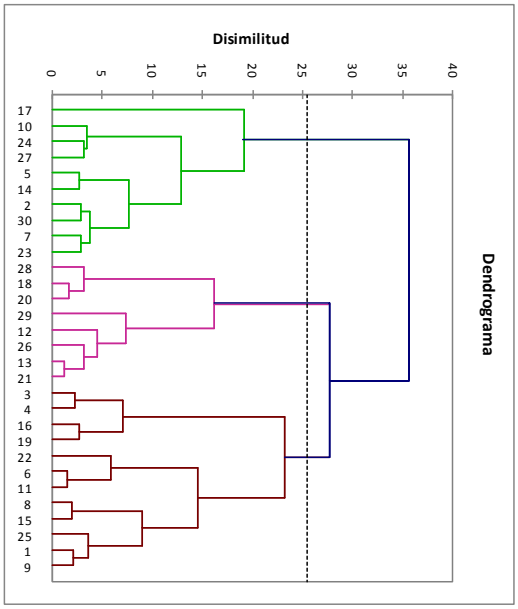


Ilustración 61. Análisis de clúster.

Otro tipo de análisis que podemos ubicar en este apartado son las tablas OLAP con más de 2 dimensiones. Se trata de obtener una tabla donde hay varias variables en filas y varias variables en columnas y, en las celdas de cruce se calcula el número de ocurrencias entre las diversas variables.

1.3.4 Análisis discriminante

En el análisis discriminante, se clasifican las variables en grupos sistemáticamente distintos, utilizando también técnicas factoriales. El objetivo de esta técnica es explicar la pertenencia de cada individuo a un grupo según la variable aleatoria y predecir a qué grupo pertenecería un individuo nuevo, del que conocemos el valor de la variable clasificadora o explicativa.

Puede aplicarse para explicar la diferencia entre los distintos tipos de objetos, contrastar diferencias significativas entre poblaciones o decidir dónde clasificar un individuo.

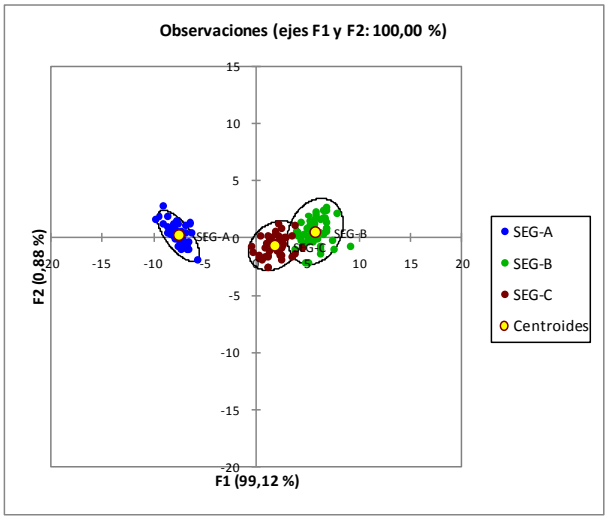


Ilustración 62. Análisis discriminante.

1.3.5 Análisis de regresión múltiple

Consiste en explicar una variable (dependiente) mediante un conjunto de otras variables (explicativas o independientes)

Lineal	$Y = a_0 + a_1X$
Parábola o cuadrática	$Y = a_0 + a_1X + a_2X^2$
Curva cúbica	$Y = a_0 + a_1X + a_2X^2 + a_3X^3$
Curva cuártica	$Y = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4$
Curva de grado n:	$Y = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$
Hipérbola	$Y = 1 / (a_0 + a1\ X)$
Exponencial	$Y = a_0\ (e^{b1})$
Geométrica	$Y = a_0\ X^b$
Gompertz	$Y = pq^{b\wedge x}$
Logística	$Y = 1 / (a\ b^x + g)$

1.3.6 Análisis estructural

Los Modelos de Ecuaciones Estructurales están pasando a convertirse en una de las opciones metodológicas más utilizadas en las ciencias sociales. El más conocido y empleado de los métodos de ecuaciones estructurales es el enfoque basado en la covarianza (LISREL es uno de los programas que mejor ejemplifican este enfoque). Existe además otro enfoque alternativo, Partial Least Squares (PLS), que en determinadas circunstancias puede llegar a ser más apropiado.

Los modelos de ecuaciones estructurales son una técnica estadística multivariante para probar y estimar relaciones causales a partir de datos estadísticos y asunciones cualitativas sobre la causalidad. Esta definición ha sido articulada por el genetista Sewall Wright (1921), el

economista Trygve **Haavelmo** (1943) y el científico cognitivo Herbert **Simon** (1953), y formalmente definido por Judea **Pearl** (2000) utilizando el cálculo contrafactual.

Esta técnica combina el análisis factorial con la regresión lineal para probar el grado de ajuste de unos datos observados a un modelo hipotetizado y expresado mediante un diagrama de senderos. Como resultado, los modelos de ecuaciones estructurales proporcionan los valores pertenecientes a cada relación, y más importante, un estadístico que expresa el grado en el que los datos se ajustan al modelo propuesto, confirmando su validez.

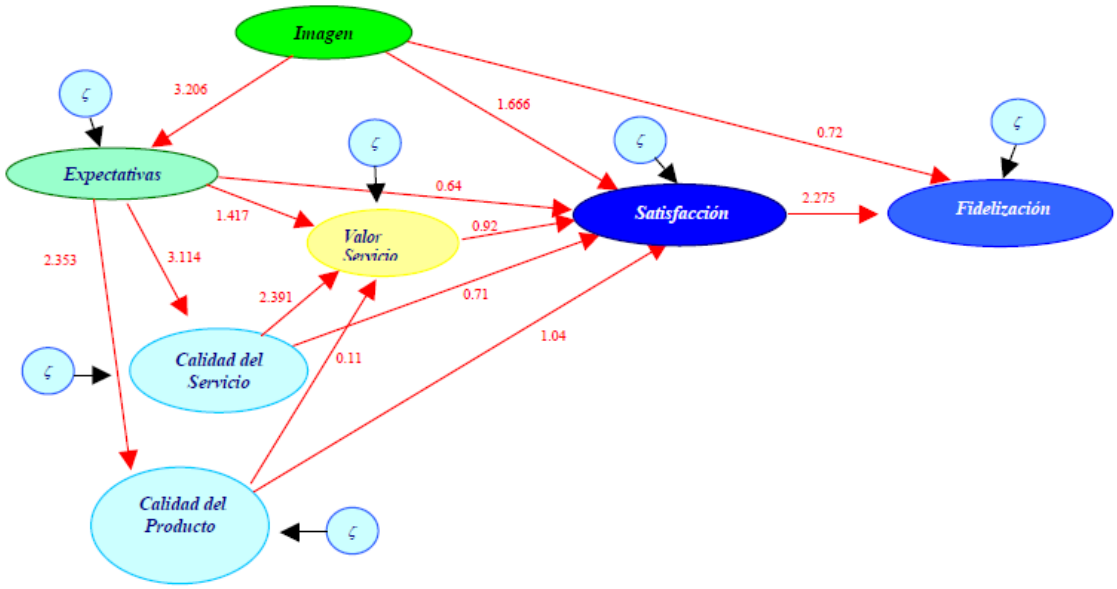


Ilustración 63. Ejemplo de análisis estructural de un servicio. Fuente: Wikipedia (2015)

1.4 Data Mining

A la hora de acotar este término se puede recurrir a diversas definiciones que han dado autores sobre el mismo, así el Data Mining o minería de datos se concibe como el "Proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos." **Witten and Frank**, 2000.

"La minería de datos (Data Mining) ayuda a las organizaciones a encontrar información que no es perceptible de forma directa, como por ejemplo patrones de comportamiento, relaciones, asociaciones, etc… que nos permitan tomar mejores decisiones. A través del análisis del pasado, y aplicando algoritmos, se construyen predicciones que nos permiten mejorar nuestra eficiencia y conseguir así una mayor rentabilidad de la actividad de negocio". (Oracle, 2010)

La **minería de datos** consiste en la extracción no trivial de información que reside de manera implícita en los datos. (La extracción de información oculta y predecible de grandes bases de datos).

"Data Mining es el análisis de (a menudo grandes) conjuntos de datos observacionales para encontrar relaciones insospechadas y resumir los datos de nuevas formas que son tanto comprensibles como útiles para el propietario de los datos". Hand, Mannila y Smyth, (2001).

"En la literatura existen muchas definiciones para caracterizar expresiones como "descubrimiento de conocimiento en base de datos" (KDD: Knowledge Discovery in Databases") y "data mining": véase por ejemplo Frawley et al. 1991), donde se hacen las siguientes aseveraciones"

- "El descubrimiento de conocimiento en bases de datos es el proceso no trivial de identificar patrones en datos que sean válidos, novedosos, potencialmente útiles y, por último, comprensibles" (Fayyad 1996)
- "… data mining se refiere al acto de extraer patrones o modelos a partir de los datos …" (Fayyad 1996)

Data mining dispone de diversas técnicas para encontrar patrones en grandes conjuntos de datos. Este enfoque multidisciplinar combina los resultados e intuiciones provenientes de varias ramas científicas tales como:

- la estadística
- el aprendizaje de máquina (machine learning)
- tecnologías difusas (fuzzy technologies)
- y redes neuronales (véase por ejemplo, Adriaans y Zantinge 1996).

"El proceso (SEMMA) de Seleccionar (Selecting), Explorar (Exploring), Modificar (Modifying), Modelizar (Modeling) y Valorar (Assessment) grandes cantidades de datos con el objetivo de descubrir patrones desconocidos que puedan ser utilizados como ventaja comparativa respecto a los competidores." Instituto SAS.

De una forma general, el objetivo principal del data mining es analizar los datos para extraer conocimiento. Este conocimiento puede ser en forma de relaciones, patrones o reglas inferidos de los datos y desconocidos o bien en forma de descripción más concisa, es decir, un resumen de ellos. Las relaciones o resúmenes constituyen el modelo de los datos analizados y se pueden representar de formas diferentes. Los modelos tradicionales están basados en la existencia de hipótesis o modelos previos, el Data Mining, no.

El software de minería de datos analiza las relaciones y patrones en los datos de transacción almacenados sobre la base de consultas de los usuarios. Existen varios tipos de software de análisis que están disponibles: estadísticos, de aprendizaje automático, y redes neuronales. En general, se buscan cualquiera de estos cuatro tipos de relaciones:

- Clases: Los datos almacenados se utilizan para localizar datos en grupos predeterminados. Por ejemplo, una cadena de restaurantes podría minar datos de compra del cliente para

determinar cuando los clientes los visitan y lo que normalmente ordenan. Esta información podría utilizarse para aumentar el tráfico al tener especiales del día.

- Grupos: Los elementos de datos se agrupan de acuerdo a las relaciones lógicas o preferencias de los consumidores. Por ejemplo, los datos pueden ser extraídos para identificar segmentos de mercado o afinidades de los consumidores.
- Asociaciones: Los datos pueden ser minados para identificar asociaciones. El ejemplo de cerveza-pañal (las familias que compran pañales también compran cerveza, debido seguramente a que el hecho de tener bebés en casa no permite salir a tomar la cerveza fuera, y condiciona un incremento de la compra de cerveza) es un ejemplo clásico de minería asociativa.
- Patrones secuenciales: Los datos se minan para anticipar patrones de comportamiento y tendencias. Por ejemplo, un distribuidor especializado en sistemas al aire libre puede predecir la probabilidad de compra de una mochila basado en la compra de un consumidor de sacos de dormir y calzado de senderismo.

Un estudio de Kdnuggets obtiene conclusiones sobre cuáles son las metodologías de análisis más utilizadas por las organizaciones, siendo por orden las siguientes en función del porcentaje de organizaciones que las usan:

- Arboles de decisión (186) 59.8 %
- Regresión (180) 57.9 %
- Agrupación/Clúster (163) 52.4 %
- Estadísticos descriptivos (149) 47.9 %
- Técnicas de visualización (119) 38.3 %

Los modelos de análisis aplicados a minería de datos a su vez pueden ser descriptivos o bien predictivos.

Los modelos descriptivos identifican patrones que explican o resumen los datos. (Pérez, 2006:16): Su meta es encontrar una descripción de los datos del estudio. Por ejemplo, cuales son los clientes de una organización (características) o encontrar productos que, frecuentemente se compran juntos.

Los modelos descriptivos describen las relaciones en los datos para poder clasificar a los clientes en grupos. A diferencia de modelos de predicción que se centran en predecir el comportamiento de un único cliente (como el riesgo de crédito), los modelos descriptivos identifican diferentes relaciones entre clientes o productos. Pero los modelos descriptivos no clasifican a los clientes según su probabilidad de tomar una acción en particular. Los modelos descriptivos se utilizan a menudo "offline"; por ejemplo, agrupar a los clientes por las preferencias de los productos según la etapa de la vida. Las herramientas de modelado

descriptivo pueden ser utilizadas para desarrollar modelos basados en agentes simulando una gran cantidad de agentes individuales pudiendo predecir acciones futuras.

Podemos considerar modelos descriptivos:

a. Clustering (análisis de segmentación). No jerárquico, se compara cada registro con todos los segmentos o semillas creados por la función de creación de semillas y se mide la distancia asignando el registro de entrada al segmento correspondiente. Es el método de la K-medias.

El Jerárquico se utiliza cuando no conocemos los grupos en los que se clasifican los clústeres. También se utilizan las redes neuronales basadas en aprendizaje no supervisado como, por ejemplo, los mapas de Kohonen.

a. Reglas de asociación. El objetivo es encontrar elementos que implican la presencia de otros elementos dentro de la misma transacción. El resultado son reglas del tipo If x Then y. Uno de los algoritmos más utilizados es Apriori (Agrawal 94)

b. Patrones secuenciales. Tratan de descubrir patrones entre transacciones en las que un conjunto de elementos va seguido de otro conjunto de elementos distanciados en un período de tiempo determinado.

c. Patrones en series temporales. Se pretende descubrir ocurrencias o secuencias similares a un dato en una base de datos que almacene información que represente una serie temporal, como puede ser la evolución de los precios de mercado.

d. Correlaciones.

Los modelos predictivos pretenden estimar valores futuros o desconocidos de variables de interés, que denominamos variables objetivo o dependientes, usando otras variables a los que nos referiremos como variables independientes o predictivas.

Se basan en el análisis de los datos actuales e históricos para hacer predicciones sobre futuros eventos. Dichas predicciones raramente suelen ser afirmaciones absolutas, pareciéndose más a eventos y su probabilidad de que sucedan en el futuro.

El objetivo es explotar los patrones de comportamiento encontrados en el pasado para poder identificar riesgos y oportunidades. Los modelos capturan las relaciones entre muchos factores permitiendo capturar riesgos potenciales asociados a un conjunto de condiciones, guiando así a la toma de decisiones.

Los modelos predictivos analizan los resultados anteriores para evaluar qué probabilidad tiene un cliente para mostrar un comportamiento específico en el futuro. Esta categoría también incluye modelos que buscan patrones discriminatorios de datos para responder a las preguntas sobre el comportamiento del cliente, tales como la detección de tipos de fraudes. Los modelos de predicción a menudo realizan cálculos en tiempo real, durante las operaciones, por ejemplo,

para evaluar el riesgo o la oportunidad de un determinado cliente o transacción, a fin de orientar una decisión.

Podemos considerar modelos predictivos:

a. Clasificación. Se utiliza el aprendizaje supervisado. Se suelen utilizar arboles de decisión, regresiones logísticas y redes neuronales.

b. Regresión o Predicción de valores. Se utilizan, junto con los anteriores, la regresión lineal y la no lineal.

2 Selección del tipo de análisis

Son muchas las situaciones en las que una organización deberá aplicar técnicas estadísticas a sus datos. Cuando la organización quiere analizar los datos debe cuestionarse con que objeto realizará el análisis:

- Quiero conocer el comportamiento pasado de una variable, entonces el análisis monovariable me dará la media, la varianza, la distribución, etc.
- ¿Quiero conocer las relaciones entre dos datos? Los análisis de correlación o regresión me mostraran cuanto de significantes son estas relaciones.
- ¿Agrupar las mediciones por un comportamiento común? Las técnicas de clusterización permiten agrupar observaciones por patrones similares.
- ¿Observar la estructura de un conjunto de datos? Las técnicas de reducción de la dimensionalidad son las que pueden servir para este objetivo.
- ¿Predecir una cantidad numérica? Aquí los métodos de regresión serán los indicados.
- ¿Predecir una categoría? Los métodos clasificadores pueden servir para alcanzar estos objetivos.

Además de saber con qué objeto se analizan los datos, también es relevante quien va a realizar el análisis. Los análisis univariables y la mayoría de los bivariables pueden realizarlos personas con cierto conocimiento estadístico pero no es necesario que sean personas especializadas. Cuando hablamos del análisis multivariable, ya precisamos de un perfil profesional muy especializado y con un conocimiento estadístico amplio. La organización que se plantee evolucionar el conocimiento de las personas de su organización deberá vincular el uso de determinadas técnicas estadísticas con un plan de formación acorde.

Lo mismo sucede si hablamos de los recursos técnicos necesarios para ejecutar los análisis. En el caso de los análisis univariables y bivariables, estos análisis pueden realizarse con una herramienta al alcance de todas las organizaciones, como es MS Excel, pero para el análisis

multivarible, en general es precisa una herramienta específica de análisis estadístico como R o SPSS.

A continuación se muestra un diagrama de flujo para la elección del método estadístico multivariable en función de las necesidades de análisis:

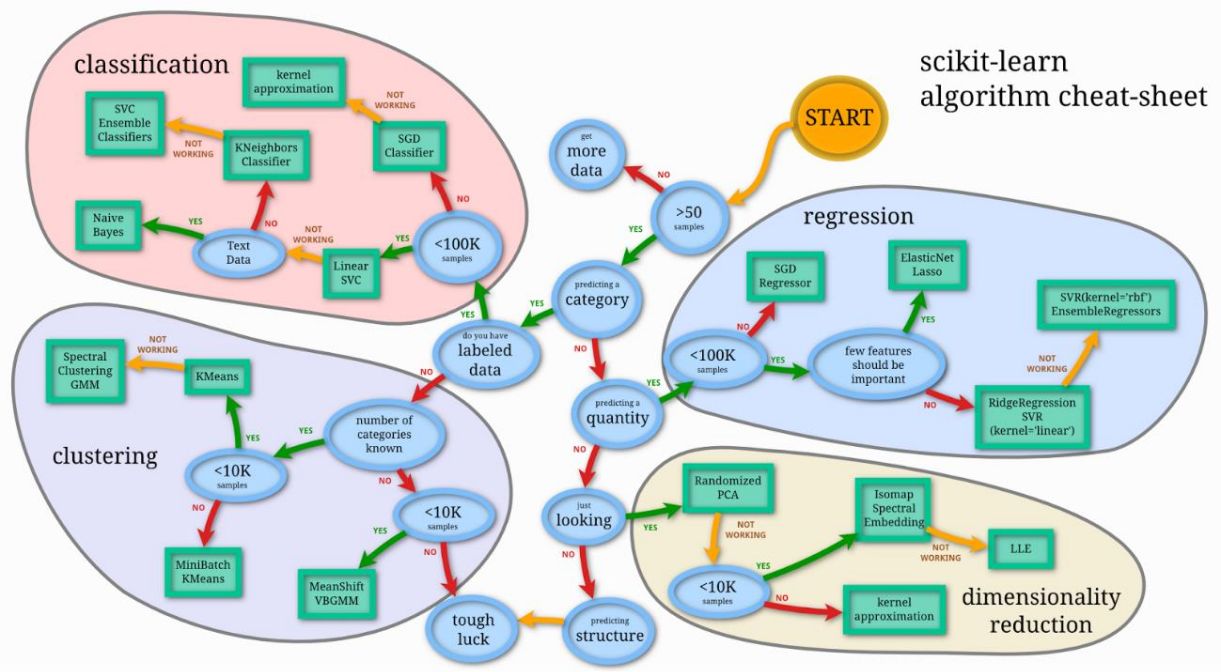


Ilustración 64. Selección de la técnica de tratamiento de datos más adecuada (Fuente: http://1.bp.blogspot.com/-ME24ePzpzIM/UQLWTwurfXI/AAAAAAAAANw/W3EETIroA80/s1600/drop_shadows_background.png

Dimensión Visualización

1 Antecedentes históricos de la representación de datos 123
1.1 Medida y teoría a lo largo de la historia 125
1.2 Evolución de los gráficos relacionados con datos estadísticos 127
1.3 La influencia de las nuevas tecnologías y la informática en la representación gráfica 137

2 La visualización de datos 139
2.1 La visualización como parte intrínseca de la estadística 139
2.2 La percepción humana de los gráficos 140
2.3 Leyes de la visión aplicadas a la visualización de gráficos 143
2.3.1 Ley de la proximidad 144
2.3.2 Ley de la semejanza o equivalencia 144
2.3.3 Ley de Prägnanz o de la buena forma 145
2.3.4 Ley de cierre 145
2.3.5 Ley de la simetría 146
2.3.6 Ley de la continuidad 147
2.3.7 Ley de la experiencia 148
2.3.8 Ley de la figura-fondo 148
2.4 El pensamiento visual 150
2.5 Conclusiones para una correcta representación de datos 153

3 Definición de un catálogo de componentes para visualización 154

4 Propuesta de un Catálogo de componentes para visualización de datos 156
4.1 Tablas de datos 157
4.2 Gráfico de líneas 158
4.3 Gráfico de coordenadas, gráfico de dispersión. 159
4.4 Gráfico de barras o columnas 160
4.5 Gráfico de secciones o circular 161
4.6 Mapa 162
4.7 Diagrama de caja 166
4.8 Treemap, mosaic plot 167
4.9 Sankey 168
4.10 Heat map 171
4.11 Sparklines 171
4.12 Nubes de palabras 172

1 Antecedentes históricos de la representación de datos

En este apartado se hace una análisis detallado de como debería de ser la dimensión visualización del modelo, iniciando la definición desde los antecedentes históricos, pasando por la leyes de la visión hasta llegar a la propuesta de Catálogo de Componentes del modelo.

Para iniciar el análisis de la representación gráfica, hay que remontarse a la época de la cultura egipcia. El arte egipcio se caracterizaba por el canon de perfil, que consiste en representar la cabeza y las extremidades de perfil, pero los hombros y los ojos de frente. Además, las figuras se representan con criterio jerárquico, el faraón tiene un tamaño más grande que los súbditos o los enemigos que están a su lado, que generalmente se representan en yuxtaposición y en planos superpuestos.

Se considera que la cultura egipcia constituye el primer ensayo de representación de la realidad. La siguiente gran etapa está protagonizada por la cultura de la Grecia clásica que supuso una nueva manera de representar la realidad.

La historia de los gráficos está ligada a la de los mapas en su inicio. Los mapas de pastilla de Babilonia, creados hace aproximadamente 4500 años, siempre se consideraron los mapas más antiguos, hasta que, en el año 1963, el arqueólogo James Mellaart (1925-) descubrió en Ankara (Turquía) un mapa que pertenece al año 6200 antes de Cristo (AC). Al parecer, el primer mapa mundial también fue creado en Turquía, hacia el año 550 AC, por Anaximander (610 AC 546 AC). Este mapa es una ilustración sencilla en la que el mundo está sintetizado en un círculo, dentro del cual se encuentran tres continentes: Asia, Europa y África.

Hacia el año 366 AC, se ubica el que puede denominarse como el primer mapa de ruta (carteroutière), representando el sistema de caminos del imperio Romano. Fue descubierto por el alemán Konrad Peutinger (1465-1547) en 1494. También, se conoce el mapa del geógrafo Claudius Ptolemy (90 DC-160 DC), que data del año 150 después de Cristo (DC), que fue uno de los primeros en utilizar el principio de coordenadas para la construcción de mapas.

En la antigüedad, los egipcios desarrollaron e incorporaron el papiro para escribir sus manuscritos, marcando un importante adelanto dentro de la comunicación visual. En el año 170 AC se inventó el pergamino, material que facilitó la creación de manifestaciones visuales

gracias a que permitía ser impreso en ambos lados (frente-dorso), principal diferencia con su antecesor el papiro.

En el año 105 DC, el chino Tsai Lun (50 DC 121 DC) produjo el nacimiento del papel, reemplazando las inscripciones en madera, paño o piedras. El proceso para la fabricación del papel se mantuvo sin cambios hasta que se mecanizó su producción (siglo XIX) en Inglaterra.

La posibilidad de reproducir las representaciones visuales de una forma no manual fue un paso muy importante dentro de la historia del lenguaje visual. En este sentido, el primer gran invento en la evolución de las técnicas de reproducción llegó desde China con la creación de la impresión en relieve (Xilografía). Para Meggs una posible fecha de su creación podría haber sido alrededor del año 165 DC. La impresión en relieve se complementó con la invención del tipo móvil, alrededor del año 1045, gracias al alquimista chino Pi Sheng (1022-1063). Los tipos estaban elaborados con una mezcla de arcilla y pegamento. En 1453 se produjo la posibilidad de imprimir mediante el uso de piezas de metales independientes, movibles y reutilizables, cada una de las cuales tenía la forma de una letra realzada en la parte superior. Esta técnica (impresión tipográfica) creada por el alemán Johann Gutenberg (1398-1468) posibilitó la rápida propagación del conocimiento debido a la producción masiva de ejemplares, que a su vez incrementó la capacidad de leer y escribir, y permitió fusionar en una misma página imágenes y textos.

En el siglo XVI, se crearon técnicas e instrumentos para precisar la observación y la medición de cantidades físicas. Estos pasos comprenden los inicios de lo que podría denominarse la disciplina de la representación visual. Los primeros ejemplos que se conocen de visualización de datos corresponden a diagramas geométricos, a tablas de las posiciones de las estrellas y otros cuerpos celestes; y al armado de mapas para ayudar en la navegación y la exploración.

En el tema de representación de datos, en su libro *Historical Development of the Graphical Representation of Statistical Data*, ubica alrededor del año 950 DC la que considera la primera construcción gráfica. Esta imagen descubierta en 1877, por el alemán Sigmund Günther, formaba parte de un manuscrito que pertenece a la *BayerischeStaats-Bibliothek* de Munich. Estaba estructurada en una cuadrícula y acompañaba una descripción de los movimientos planetarios a través del zodíaco en función del tiempo.

En cambio, otros autores, como Paul Mijksenaar (1944-), sitúan el trabajo de Nicole Oresme (1323-1382) como la primera representación gráfica de datos.

El 1375 se realizó un Atlas mundial conocido como el Atlas de conocimientos geográficos visuales más completo de los últimos años de la edad media. Este ejemplar era una completa obra visual de cosmografía, junto con un calendario perpetuo y una representación temática del mundo conocido. El Atlas fue un encargo de Carlos V de Francia (1338-1380) al catalán Abraham Cresques (¿-1387).

Podría decirse que las representaciones gráficas mencionadas hasta este punto del recorrido histórico, mediados del siglo XIV, responden a conocimientos que el ser humano tenía del mundo, sin basarse en explicaciones científicas o teóricas. Esta información podría resumirse con el concepto de mimesis, que consiste en la expresión del conocimiento que se tiene de la naturaleza y del universo en el que vivimos. Este tipo de reproducciones gráficas podría pensarse que todavía no respondían a códigos preestablecidos, ya que un código puede definirse como un conjunto de normas sistemáticas que regulan unitariamente una materia determinada. El concepto de mimesis está cercano a la reproducción estética, en cambio, el concepto de código está relacionado con reglas y parámetros preestablecidos.

1.1 Medida y teoría a lo largo de la historia

Los problemas que caracterizaron al siglo XVII se centraron en las mediciones físicas de tiempo, distancia y espacio; en la realización de mapas de navegación y expansión territorial. Se crearon objetos para lograr mayor exactitud en las ilustraciones técnicas y facilitar la reproducción de figuras en diferentes escalas.

En 1626, Christopher Scheiner (1575-1650) produce una secuencia visual que representaba los cambios de los puntos solares con el paso del tiempo desarrollando una de las principales herramientas contemporáneas para la visualización de información (*Small Multiple*).

En 1637 René Descartes en su libro "Geometrie" presentó las coordenadas cartesianas como el modelo básico de representación de datos, que tanto ha aportado a la ciencia y a la tecnología.

La geometría analítica, iniciada por Descartes (1596-1650) en 1637, ayudó a determinar la forma de representar gráficamente una curva de una ecuación, determinando la relación existente entre las coordenadas de cada punto de dicha curva, y cómo cada ecuación puede ser representada en una curva, lo que aportó una nueva metodología en la representación gráfica de datos.

En 1644 de Michael van Langren, un astrónomo holandés que trabajó en la corte de Felipe IV, que recogía las distintas estimaciones que se habían hecho de la distancia que separa Toledo de Roma (12 en total).

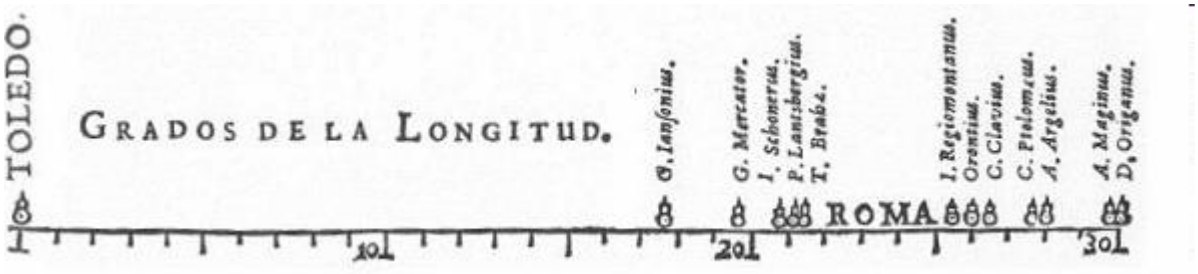
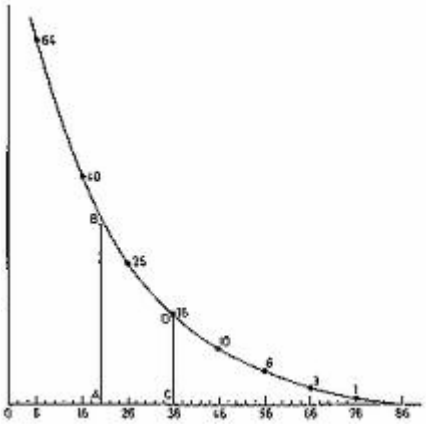


Ilustración 65. Distancia de Toledo a Roma. Fuente: Michael van Langren.



Durante este siglo también se produjo un gran avance en el corpus del conocimiento teórico (teoría de la probabilidad y estadísticas demográficas y políticas) y la puesta en práctica de teorías científicas que dio origen a las primeras construcciones de tablas y gráficos, y a la recopilación sistemática de datos estadísticos.

En 1669 de Christian Huygens elabora un gráfico que muestra los años restantes de vida en función de la edad.

Ilustración 66. Años de vida en función de la edad. Fuente: Christian Huygens.

La combinación de la cartografía y las habilidades estadísticas produjo mapas de datos. El primer mapa de datos, creado por el inglés Edmond Halley (1656-1742) en 1686, estaba orientado hacia la representación del estado del tiempo, mostrando previsiones del viento en un mapa geográfico.

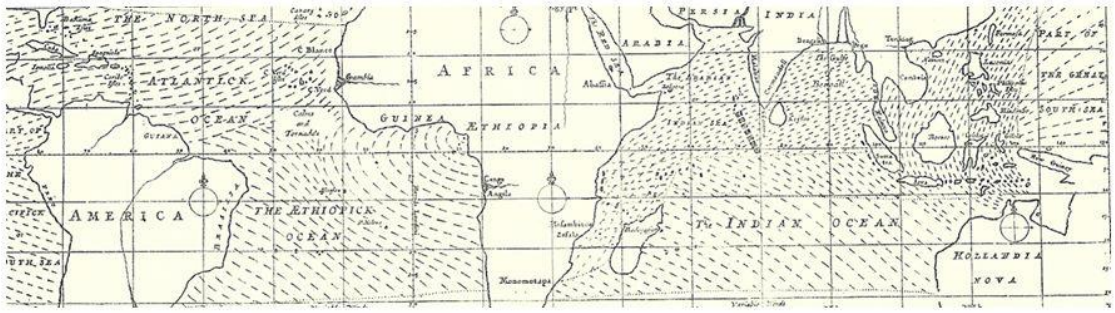


Ilustración 67. Mapa de previsiones de viento. Fuente: Edmond Halley.

Hacia finales de siglo, los elementos necesarios: datos reales de interés significante, teorías que les den sentido, e ideas para su representación visual para el desarrollo de construcciones gráficas estaban planteados. Podría afirmarse que, durante el siglo XVII, se obtuvieron los ingredientes necesarios para la confección disciplinar del pensamiento visual.

El siglo XVIII potenció y diversificó las primeras creaciones sobre visualización de datos que fueron planteadas durante los años anteriores. Realizadores de mapas comenzaron a tratar de

mostrar más que sólo posiciones geográficas, surgiendo como resultados nuevas tipologías gráficas. Junto con los comienzos de la teoría estadística y la colección sistemática de datos empíricos, fueron introducidos gráficos abstractos y gráficos de funciones. Las innovaciones tecnológicas (impresiones color, litografía) generaron la materia prima que facilitó la reproducción y creación de imágenes de datos. En una primera instancia, muchas de estas formas gráficas aparecían en publicaciones de circulación limitada, de manera que no eran conocidas popularmente.

Una de las construcciones gráficas que más evolucionó fue el mapa, creándose un mapa diferente para representar cada necesidad temática como mapas de contorno, mapas cronológicos (1753) y mapas geológicos (1775). El desarrollo de la geometría descriptiva, en 1776, conocida como el sistema Monge (1746-1818) aportó una gran ayuda a la representación de la realidad de tres dimensiones en la bidimensionalidad del papel.

1.2 Evolución de los gráficos relacionados con datos estadísticos

Para **Funkhouser**, Joseph **Priestley** (1733-1804) ha sido el creador de los gráficos estadísticos, mientras que para Edward **Tufte** este lugar pertenece a Johann **Heinrich Lambert** y William **Playfair** (1759-1823). Históricamente, el primer caso registrado en considerar el tiempo como una línea gráfica (*time-series*) se encuentra en una publicación de **Priestley**.

En 1765 J. **Priestley** fue el primero en representar los datos históricos en una línea de tiempo.

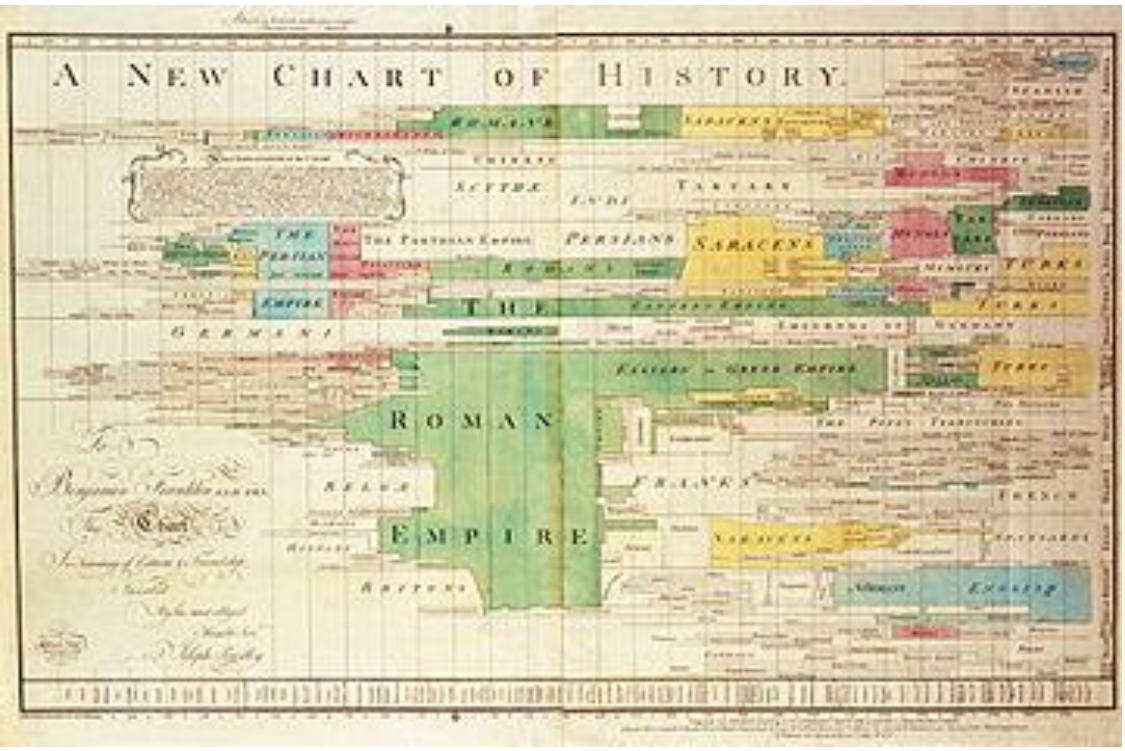


Ilustración 68. Datos históricos en una línea de tiempo. Fuente: Priestley.

En este gráfico se representan las fechas de nacimiento y muerte de 2000 personajes ilustres de la Historia.

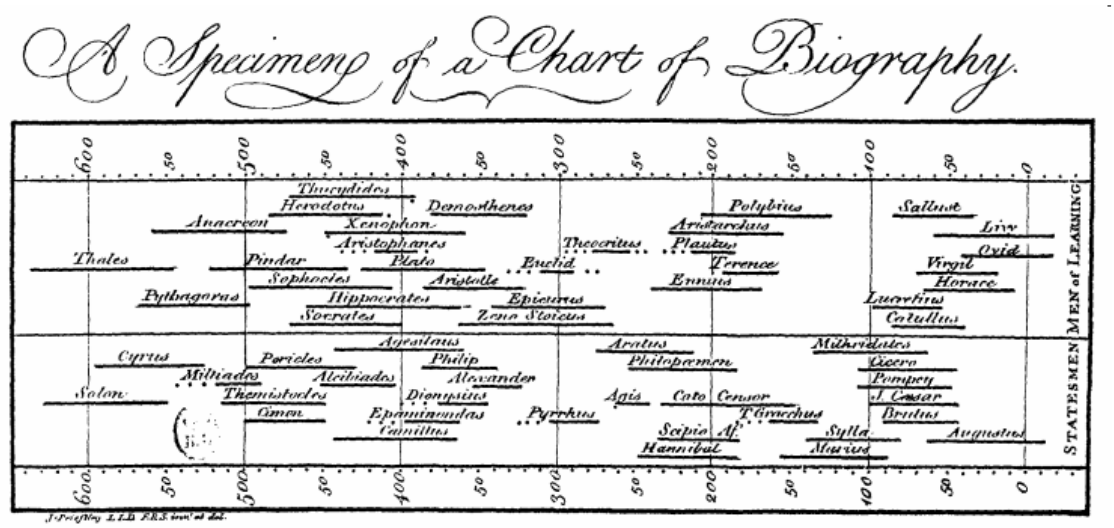


Ilustración 69. Biografia gráfica de personajes históricos. Fuente: Priestley.

Abarca un intervalo de tiempo muy amplio, desde 1200 a.C. hasta 1800 d.C. y cuenta con dos mil nombres.

Priestley clasificó a los personajes en seis categorías. En el ejemplo se muestra un fragmento del gráfico referido a la biografía de personajes del año 600 a.C. al año 0 relativo a las categorías *eruditos* y *estadistas*. Los puntos que aparecen al lado de algunos de los segmentos indican que las fechas de nacimiento o muerte son inciertas.

El gráfico de coordenadas (*time-series*), publicado en 1786 en el Atlas de Playfair *The Comercial and Political Atlas* (Londres), así como gran parte de sus gráficos, representaba información sobre economía. La importancia del estudio realizado por Playfair fue la creación de un nuevo método gráfico que contrastaba con la presentación tabular de la información conocida hasta ese momento. En palabras del propio Playfair: "Los gráficos son preferibles a las tablas porque los gráficos muestran la forma de los datos desde una perspectiva comparativa".

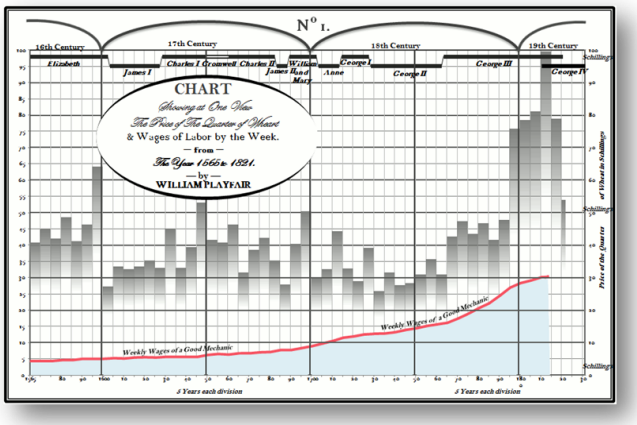


Ilustración 70. Gráfico de coordenadas. Fuente: Playfair.

William **Playfair** (1759-1823) ideó los gráficos de barras tal y como hoy los conocemos. Los gráficos de serie de tiempo dieron fama al médico Joseph **Priestley** en 1765. En ellos se inspiró William Playfair para inventar el diagrama de barras, que aparece por primera vez en su *Atlas Comercial y Político*, de 1786.

Playfair, que opinaba que los gráficos eran más explicativos que las tablas, se considera inventor del polígono de frecuencias, el diagrama de barras y el gráfico de tarta.

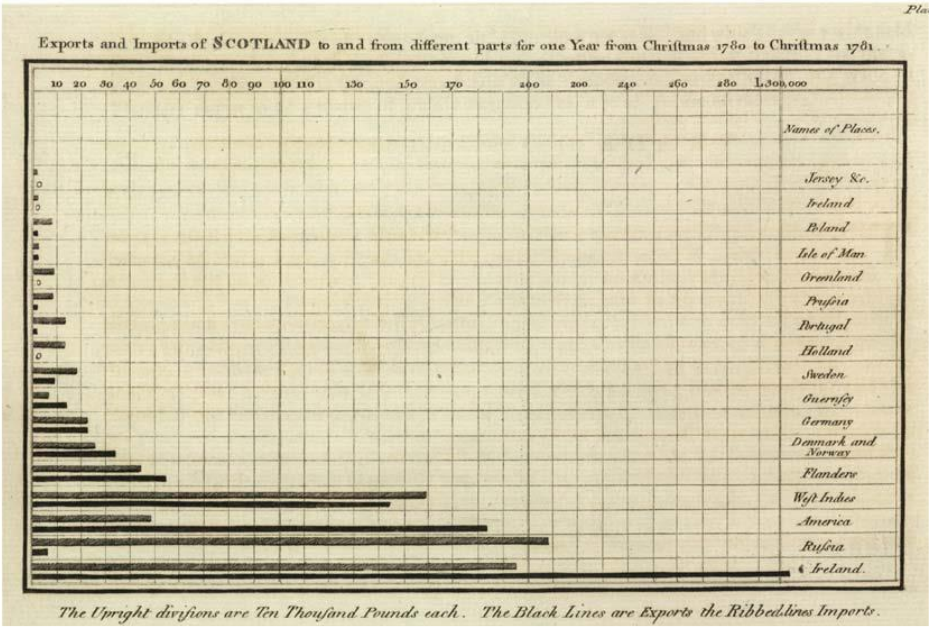


Ilustración 71. Diagrama de barras. Fuente: Playfair.

Gráfico de la deuda de Inglaterra (The Commercial and Political Atlas).

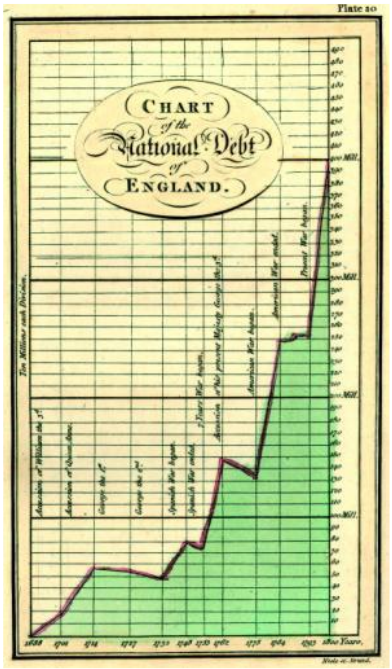


Ilustración 72. Evolución deuda Inglaterra.

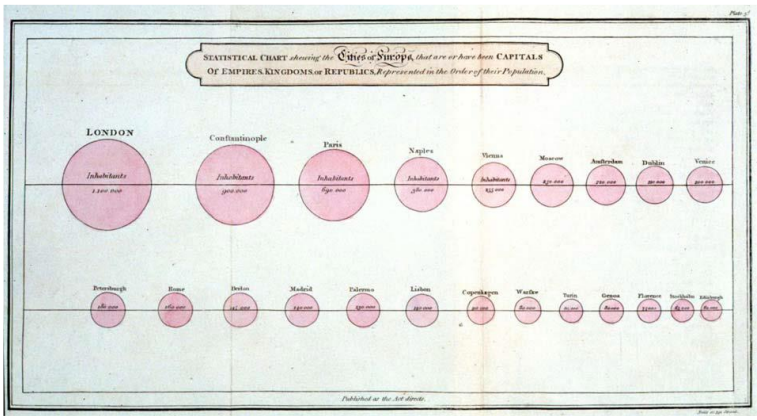


Ilustración 73. Gráfico circular de la población de distintas capitales europeas. El tamaño del círculo es proporcional al dato (Statistical Breviary).

Gráfico circular de la población de distintas capitales europeas. El tamaño del círculo es proporcional al dato (Statistical Breviary).

El Breviario Estadístico, publicado en Londres en 1801, contiene el considerado como primer diagrama circular.

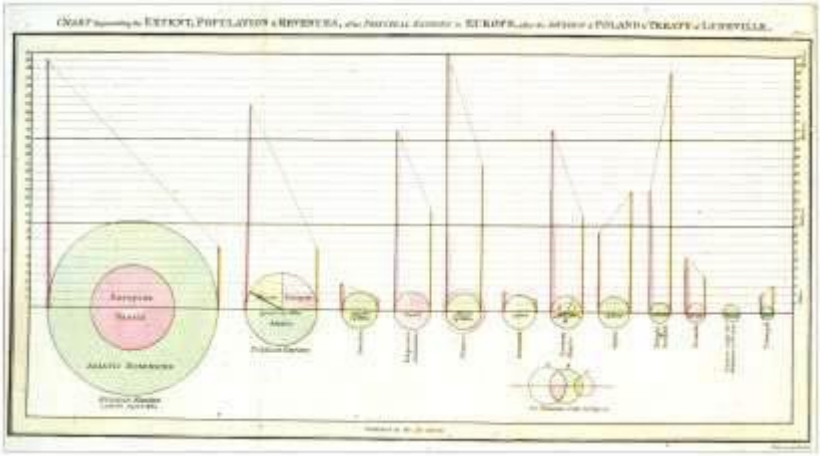


Ilustración 74. Población, impuestos y extensión superficial de distintos países europeos. Aparece por primera vez el gráfico de sectores (Statistical Breviary).

Población, impuestos y extensión superficial de distintos países europeos. Aparece por primera vez el gráfico de sectores (Statistical Breviary).

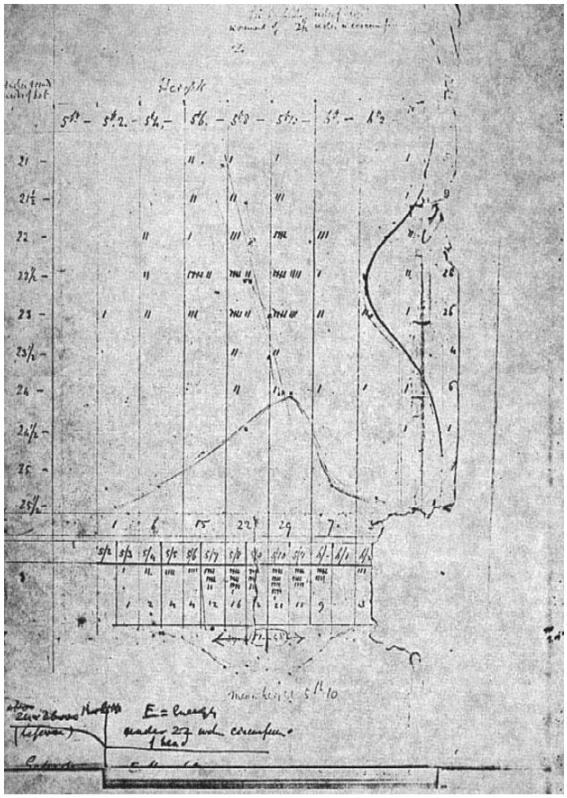


Ilustración 75. Gráfico de dispersión, Francis Galton

Francis Galton (1822 –1911), explorador, antropólogo y estadístico británico, es conocido por introducir la idea de las curvas de regresión, que pueden utilizarse para predecir el valor de una variable aleatoria en función de otra, y la noción de correlación. Hizo varios estudios sobre herencia genética y publicó los resultados de sus trabajos en su obra Natural inheritance.

Gracias a sus estudios y novedosas técnicas de análisis de datos y representaciones gráficas incluidas, y a los trabajos de astronomía de John F. W. Herschel (científico británico nacido en 1792, fallecido en 1871) se ideó y extendió el uso del gráfico de dispersión, utilizado para estudiar las relaciones que puede haber entre

dos variables referidas al mismo conjunto de individuos.

Gráfico de correlación (no de dispersión) de Galton que muestra la relación entre la circunferencia de la cabeza de un individuo y su estatura. Galton describió el proceso de construcción de un diagrama de dispersión en un artículo de 1890.

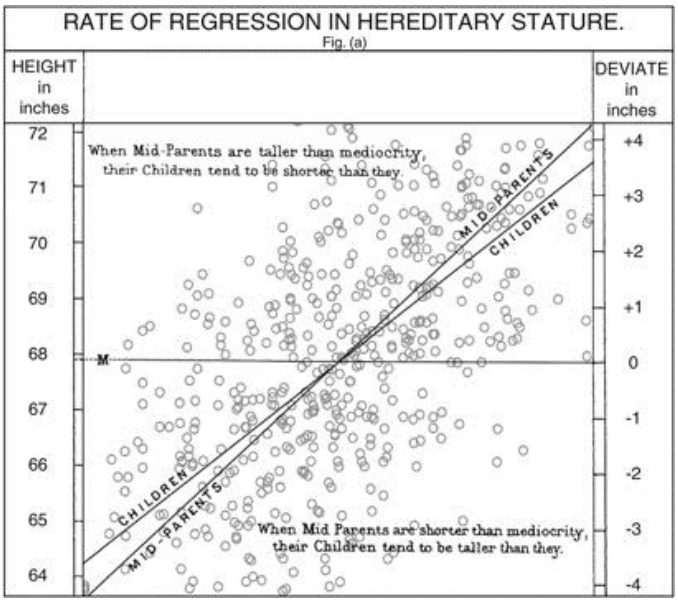


Ilustración 76. Gráfico de dispersión que recoge los datos de consumo de fertilizantes y producción de cereales en varios países europeos.

Gracias a las innovaciones en diseño y avances técnicos, la primera mitad del siglo XIX, experimentó un crecimiento explosivo de gráficos estadísticos y mapas temáticos. Fueron inventados todo tipo de formas gráficas visuales, como gráficos de barras y circulares (*pie chart*), histogramas, gráficos de línea, líneas de tiempo, mapas de contornos (1843), gráficos isotérmicos (1817), gráficos en forma de ojiva, cartogramas (1826), mapas de puntos (1830, *dot-maps*), mapas de transporte (1837), gráficos compuestos.

Los análisis demográficos y estadísticos (denominados estadísticas morales: crímenes, suicidios, literatura) comenzaron a ser representados gráficamente en cuadros, diagramas y mapas temáticos. André Michel Guerry (1802-1866) es considerado el fundador del análisis de la estadística moral, en temáticas como la criminología, la sociología y las ciencias sociales.

Alrededor del 1835, el matemático, estadístico y sociólogo Lambert Adolphe Jacques Quételet (1796-1874) fue el primero en aplicar los métodos estadísticos en las ciencias sociales, utilizando las leyes estadísticas para explicar patrones de comportamiento en crímenes, casamientos y suicidios.

Otra persona importante de mencionar es el ingeniero francés Charles Joseph Minard (1781-1870) quien incorporó diagramas estadísticos y gráficos de barras a mapas cartográficos, y es

considerado por **Tufte** como el creador del mejor gráfico estadístico de todos los tiempos. Este gráfico, realizado en 1861, forma parte de una publicación recopilatoria de sus trabajos y representa el terrible destino del ejército de Napoleón en Rusia.

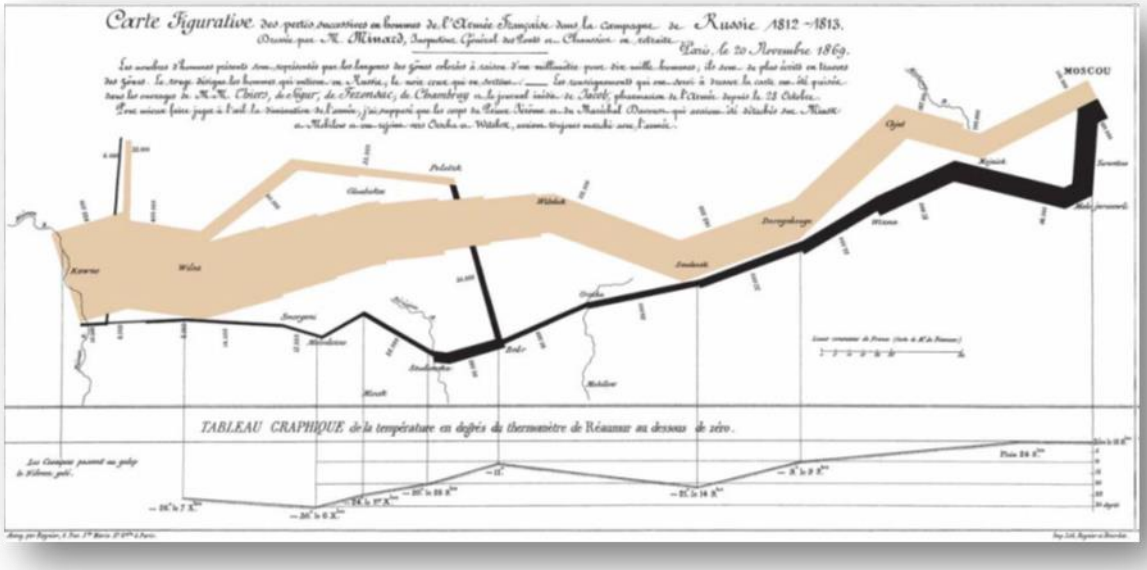


Ilustración 77. Destino del ejército de Napoleón en Rusia, E. Tufte.

La campaña rusa del ejército de Napoleón: En beige las tropas que partieron hacia Moscú, y en negro las que regresaron. Se representa también la temperatura del suelo. El grosor de las bandas es proporcional al número de soldados.

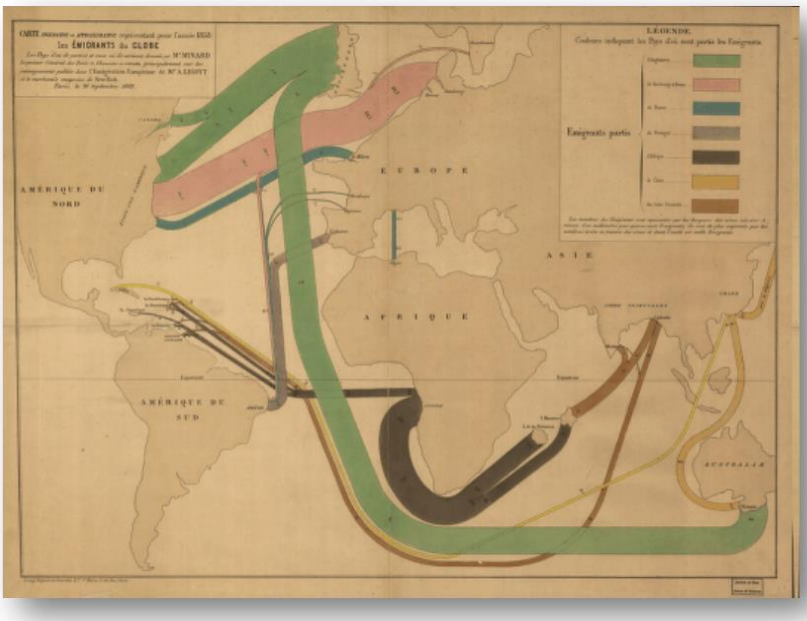


Ilustración 78. Rutas de emigración en el mundo en 1858. El ancho de las bandas es proporcional al número de emigrantes.

En el siglo XIX, pueden distinguirse dos metodologías para la visualización de datos. Por un lado, se encontraban los que apoyaban las representaciones o los gráficos tabulares, y por otro, los que comenzaban a encontrar ventajas en los métodos gráficos de representación visual. El primer grupo estaba integrado por economistas y políticos que, en un principio, criticaron al método gráfico argumentando falta de rigor.

El primer paso para el crecimiento de la representación visual de datos fue el establecimiento, en Europa, de oficinas oficiales de estadísticas. Esto evidenció un reconocimiento hacia la creciente importancia de la información estadística para el planeamiento social, la industrialización, el comercio y el transporte. Entre 1830 y 1850 creció el interés por realizar estudios sobre datos estadísticos. Se formaron numerosas sociedades estadísticas y agencias y se crearon estadísticas municipales en varios países. También aumentó el público interesado en los resultados de las investigaciones sobre la evolución del método gráfico estadístico. Estos análisis conjuntos llevaron a la creación del Congreso Internacional de Estadística. El primer congreso se realizó en Bruselas en 1853, celebrándose ocho congresos más: París 1855, Viena 1857, Londres 1860, Berlín 1863, Florencia 1867, La Haya 1869, San Petersburgo 1872, y Budapest, 1876. En 1858, Florence Nightngale elaboró gráficos que comparaban las causas de mortalidad en los hospitales de campaña durante la guerra de Crimea.

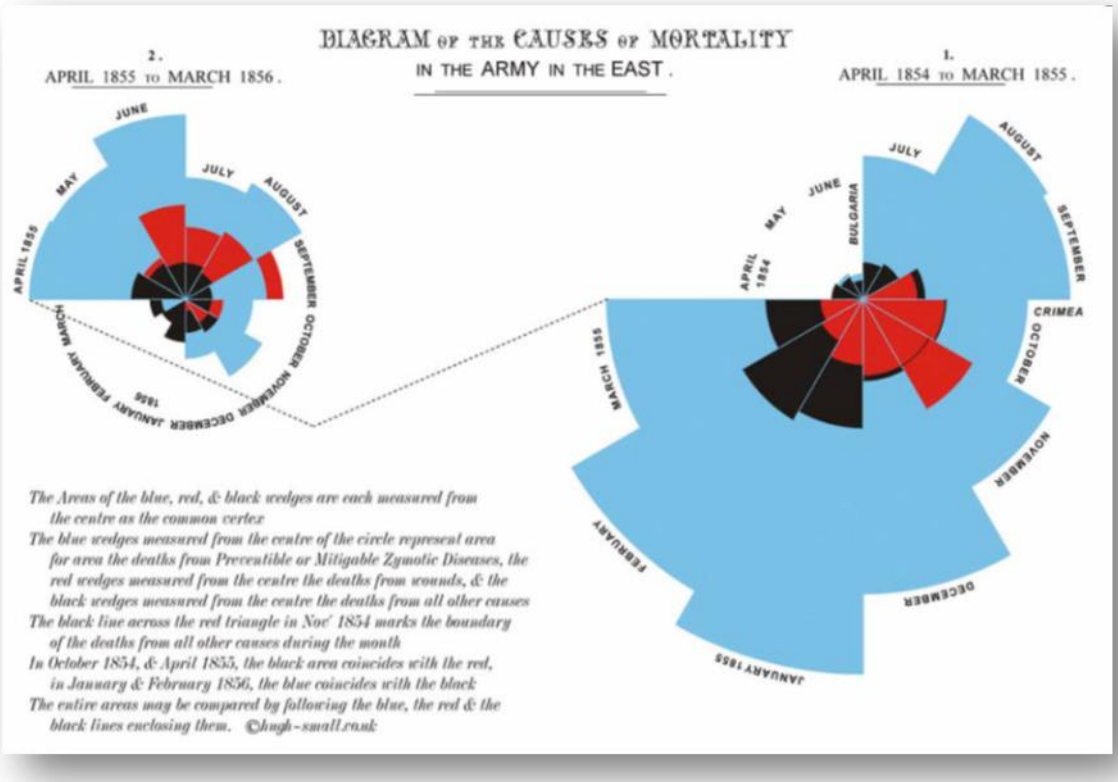


Ilustración 79. Representación de las bajas en el ejército británico en la guerra de Crimea.

Representación de las bajas en el ejército británico en la guerra de Crimea, por causa y mes, entre 1854 y 1856. El azul representa las defunciones por enfermedades infecciosas, el rojo, por heridas de guerra, y el negro por otras causas. El tamaño de las áreas es proporcional al número que representan.

El período entre los años 1860 y 1890 es muy importante en la consolidación del método gráfico, y podría denominarse los años dorados del método gráfico. Todos aquellos con imaginación capaces de realizar representaciones visuales a partir de información comenzaron a experimentar con el nuevo lenguaje. El método gráfico fue oficialmente reconocido por las agencias gubernamentales y aceptado en las publicaciones como un elemento más de información. En las publicaciones comenzaron a ser incorporadas diferentes tipologías de representaciones visuales.

Las primeras décadas del siglo XX fueron una época de cambios en la vida social, política, cultural y económica que alteraron de manera radical los diferentes aspectos de la condición humana. El desencadenamiento de la primera guerra mundial (1914-1919) movilizó a la civilización occidental generando un cambio en la concepción de la vida. En medio de este contexto, las formas gráficas de comunicación experimentaron una serie de revoluciones creativas que cuestionaron sus valores, su aproximación a la organización del espacio y su función en la sociedad.

Por estos motivos, los comienzos del siglo XX pueden ser llamados los oscuros años modernos de la representación visual de datos. El entusiasmo por la representación visual había sido reemplazado por una producción masiva de gráficos que sobrevaloraban la estética pero descuidaban el contenido. Este proceso se agudizó al punto que las construcciones gráficas pasaron a convertirse sólo en imágenes, atractivas y evocadoras, pero incapaces de representar un hecho o comunicar datos con rigor.

Otra reacción adoptada por la sociedad fue la aplicación y popularización del método gráfico en todo lo que tuviera a su alcance, estandarizándose su uso en el gobierno y el comercio, además de la ciencia. Las características de esta etapa pueden definirse como un retroceso, un estancamiento dentro de su evolución y una pérdida de pulcritud técnica.

En 1911, Henry L. Gantt planificó los calendarios de producción industrial usando por primera vez el diagrama de Gantt que aún hoy es la representación fundamental para la gestión de proyectos, mediante la cual se planifica la duración de las tareas y sus relaciones.

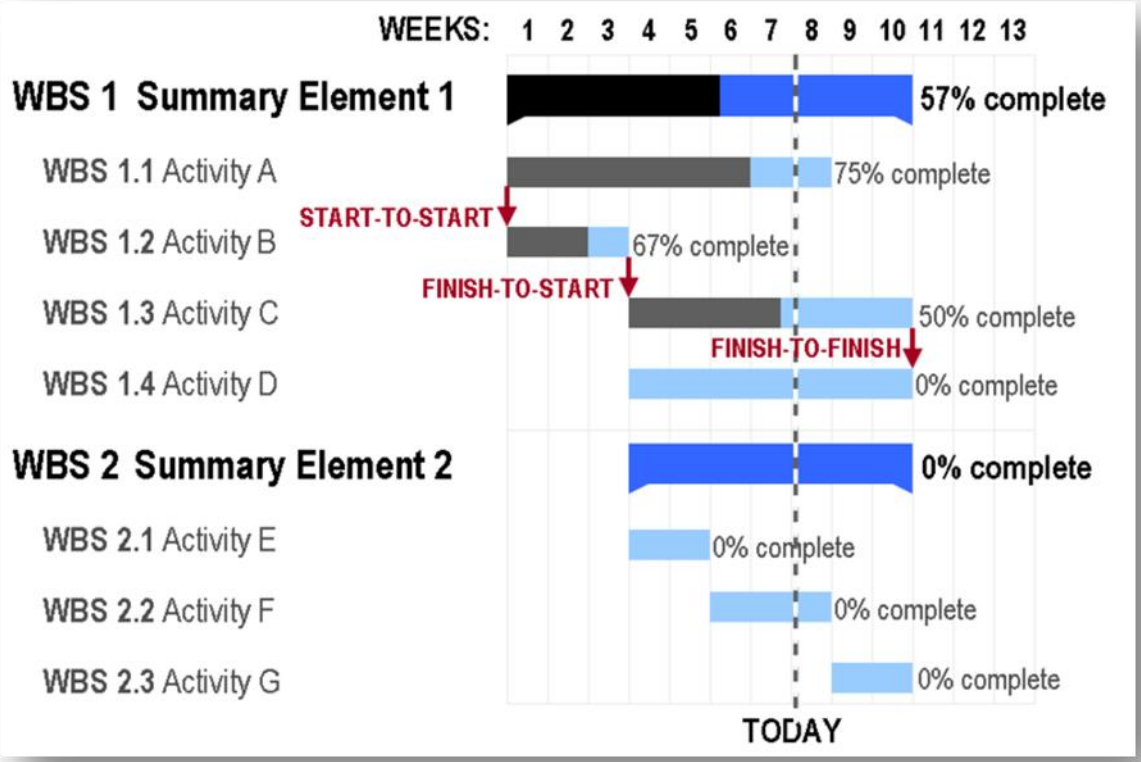


Ilustración 80. Diagrama de Gantt.

En 1913, el diagrama de Hertzsprung-Rusell, se utilizó por primera vez para el estudio de la luminosidad y temperatura de color de las estrellas.

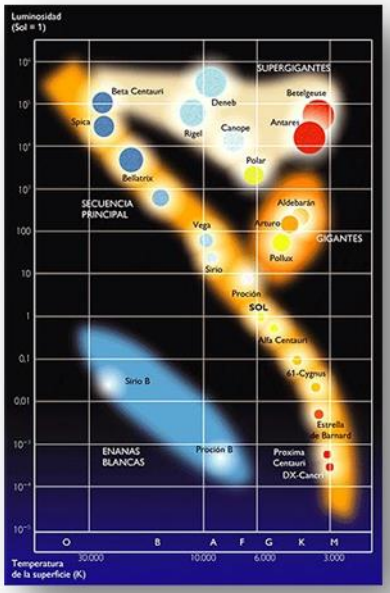


Ilustración 81. Diagrama de Hertzsprung-Rusell.

En relación al método gráfico puede mencionarse, por un lado, la creación del lenguaje ISOTYPE (Internacional System Of Typographic Picture Education) en la década de los 30, en Viena. Este lenguaje de imágenes combinaba las reglas de la ciencia y las del diseño en la construcción de sus signos visuales y fue desarrollado por Otto Neurath (1882-1945). Este sistema permitió la producción de un gran número de imágenes que representaban palabras y conceptos de uso cotidiano, logrando un lenguaje icónico universal.

Ilustración 82. Otto Neurath, ISOTYPE.

Por otro lado, el nuevo diseño introducido por Henry C. Beck (1903-1974) en el sistema de metro de Londres creó una nueva forma de representación de la realidad: la esquematización de un mapa.

Ilustración 83. Plano del metro de Londres, Henry C. Beck.

A pesar de la diversificación de tipologías gráficas los mapas aún conservaban las mismas características de los últimos 130 años. Por ejemplo, los mapas de transporte continuaban tratando de ser una copia fiel de la realidad sobre la que se estructuraban los recorridos de los medios de transporte. Gracias a Beck la representación gráfica de transportes comenzó a favorecer la comprensión de cómo llegar a un determinado lugar, dejando de lado las características geográficas reales. Un mapa en el que los rasgos geográficos desaparecen casi en su totalidad, y son reemplazados por un nivel mayor de abstracción de la realidad se denomina Plano. A partir de este momento, lo importante en una construcción gráfica de transportes fue representar un recorrido, sin especificar si en medio de dicho recorrido se encontraba una colina, un río o un precipicio.

Con el transcurso de los años el concepto de abstracción y geometrización empezó a aplicarse a distintas tipologías esquemáticas.

1.3 La influencia de las nuevas tecnologías y la informática en la representación gráfica

Durante el siglo XX, la gestación de nuevas ideas y el desarrollo de las tecnologías permitieron explorar las posibilidades representativas de las dos dimensiones del plano y generar nuevos métodos que permitieran representar información de tres dimensiones, favoreciendo el desarrollo de gráficos multivariables. Grandes corporaciones estadounidenses cumplieron una función importante en el desarrollo de productos y servicios, gracias a los cuales fue posible que el diseño de información comenzara a salir de la inactividad. La invención del primer ordenador, en 1944, hizo posible realizar pruebas y experimentar distintas técnicas gráficas para la elaboración de diagramas y mapas. También con la informática se crearon nuevas incertidumbres, motivando a retomar los estudios relacionados con el método gráfico.

En Francia, Jacques Bertin (1918-) publicaba la obra *Semiologie Graphique* (1963), convirtiéndose en el libro esencial para organizar los elementos visuales y perceptuales de acuerdo a características y relaciones de la información. La obra de Bertin posibilitó los cimientos para los principios de la semiología gráfica, desarrollando una teoría sobre los modos de representación y símbolos gráficos.

En los EEUU, John W. Tukey (1915-2000), en su publicación *El futuro del análisis de información*, realiza una llamada al reconocimiento del análisis de información como una rama legítima de la estadística diferente a las estadísticas matemáticas. Estos avances gráficos, publicados en 1977, lograron hacer que la gráfica de datos comenzara a ser respetada y tomada en cuenta nuevamente.

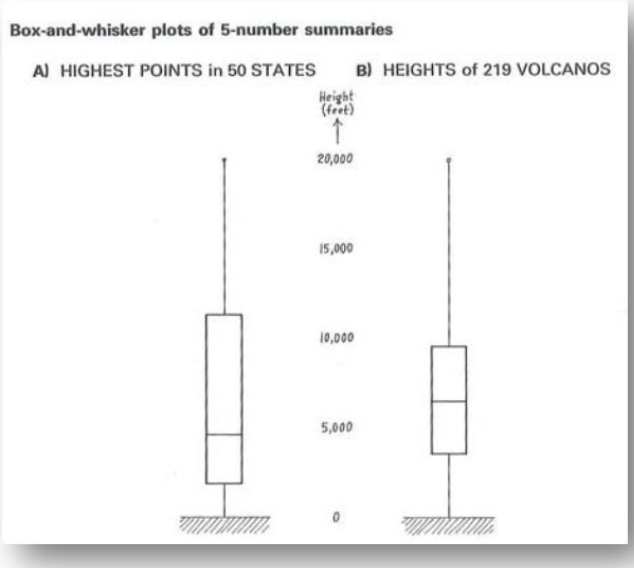


Ilustración 84. Diagrama de caja.

Diagramas de cajas y bigotes en los que se representan los datos de los puntos más altos en los 50 estados de EE. UU y la altura de 219 volcanes, del libro *Exploratory data analysis*.

Los ordenadores ofrecieron la posibilidad de construir viejas y nuevas formas gráficas con programas especializados (software). Se desarrollaron gráficos de alta resolución, aunque se tardaron algunos años más en que se popularizara el uso de la informática.

En 1974, Jerome H. Friedman y Tukey dieron comienzo a los gráficos estadísticos interactivos (sistema PRIM-9) y en tres dimensiones.

Hacia el final de este período la tecnología desplegó todo su potencial: se desarrolló la investigación científica por ordenador, se produjeron herramientas de software, lenguajes específicos (C, UNIX) y avances en el análisis de datos, exhibidores y técnicas de reproducción (plotters, terminales gráficas, tabletas digitales, el mouse, etc.). Estos avances técnicos y metodológicos de trabajo enriquecieron las construcciones gráficas y mecanizaron su desarrollo, generando a su vez nuevos paradigmas. En otras palabras, comenzaron a surgir nuevas problemáticas: representaciones visuales de información multivariable, animaciones del proceso estadístico, y teorías basadas en la percepción, relacionadas en averiguar la forma para mejorar la comunicación y el entendimiento entre la información y el lector.

En conclusión, los avances en las técnicas y métodos de representación visual de datos están íntimamente relacionados con la evolución tecnológica y del corpus del conocimiento teórico. Gracias a la combinación de estos avances y teorías, en la actualidad se están produciendo soluciones para algunos de los problemas mencionados a lo largo del recorrido histórico analizado: pobreza de recursos gráficos, universalidad de lenguaje visual, tiempos de creación y construcciones gráficas pobres comunicacionalmente.

2 La visualización de datos

Todos hemos oído el dicho: una imagen vale más que mil palabras, o que mil datos! Se considera que una de las mejores técnicas para hacer comprensibles los datos es la representación de los números mediante imágenes. Con ello se hacen más evidentes los patrones o se ponen de relieve cuestiones, que de otra manera pasarían desapercibidas.

Existe una amplia variedad de gráficos, desde sencillas líneas de puntos a complejos componentes gráficos enriquecidos. También existen los gráficos planos, estáticos, y aquellos por los que se puede navegar, que son interactivos.

2.1 La visualización como parte intrínseca de la estadística

La producción gráfica es una parte esencial del proceso estadístico, así como la visualización de datos es una parte importante de la comunicación de las tendencias y relaciones.

Es mucho más fácil comprender los datos presentados como gráficos o mapas que en largas listas de números, o tablas de datos, asumiendo por supuesto que los gráficos están correctamente realizados.

Las representaciones gráficas deben presentar las tendencias y las relaciones de forma rápida y clara. Son una manera eficiente de transmitir la información desde el sistema de información hacia el sistema de conocimiento de la organización.

Una mala representación de la información estadística puede resultar engañosa. Hay muchas maneras de proporcionar información engañosa, bien sea intencionadamente, o bien como ocurre muy frecuentemente, por error.

Las representaciones complejas, a menudo confunden al usuario, de forma que éste debe de invertir más tiempo y esfuerzo en la comprensión del sistema que en el de los datos propiamente.

Palyfair publica dos libros sobre el desarrollo del comercio mundial a las puertas del siglo XIX (1786 y 1801), en el que crea el "método gráfico". Este nuevo enfoque, promocionaba las ventajas de las presentaciones visuales con los siguientes argumentos:

a. Las presentaciones gráficas ayudan a simplificar complicadas relaciones que pueden ser difíciles de percibir a simple vista.
b. Los responsables de la toma de decisiones, necesitan un atajo visual a la esencia de la información estadística.
c. Las tablas, especialmente las voluminosas, pueden ser tediosas y difíciles de comprender. Las presentaciones gráficas ayudan al usuario.
d. Los gráficos atraen la vista. Playfair enfatizó esto al subrayar la utilidad de este método para la comprensión de las relaciones y las cantidades.

e. El método gráfico atrae y desafía a la mente, ayudando no sólo a la percepción, sino también en la comprensión de estructuras y relaciones.

2.2 La percepción humana de los gráficos

Los modelos actuales de la visión humana se basan en un estudio interdisciplinario de la **percepción visual** que tiene en cuenta como mínimo los siguientes aspectos:

- fisiológico
- neurológico
- psicológico

En ellos el ojo actúa como receptor y el proceso perceptivo es principalmente neurológico y cognitivo.

En el caso de la percepción visual el proceso empieza en los ojos. La luz que llega a ellos estimula los órganos receptores de la retina que convierten el estímulo lumínico en impulso eléctrico y lo transmiten, a través de los axones del nervio óptico, hacia el cerebro.

La información se encuentra en la conexión cerebral llamada *quiasma* y continúa hasta llegar a una zona específica del tálamo (los núcleos geniculados laterales) desde donde es enviada al córtex visual situado en el lóbulo occipital. La información procedente del ojo derecho es enviada al córtex visual del hemisferio izquierdo y al revés.

Podemos dividir este proceso en tres etapas principales:

a. Fotorrecepción: la luz que llega al ojo estimula en el fondo de la retina las células fotorreceptoras, que transmiten la señal al nervio óptico.
b. Transmisión y procesamiento: en la retina empieza un primer nivel de procesamiento que se irá haciendo complejo hasta llegar al tálamo y de éste al córtex cerebral.
c. Percepción: en el lóbulo occipital (en la zona visual primaria y en la zona de asociación visual) se completa el proceso de percepción y podemos hablar de consciencia de la imagen vista.

Dentro del ámbito de la psicología, el análisis de la percepción visual humana ocupa un lugar relevante. Las teorías y leyes expuestas por la Escuela de la Gestalt, aunque no explican el porqué del proceso perceptivo, siguen siendo un referente importante en tanto que describen y predicen algunos principios de su funcionamiento.

La psicología de la Gestalt es una corriente de pensamiento dentro de la psicología moderna, surgida en Alemania a principios del siglo pasado. Sus pensadores más reconocidos han sido los teóricos Max Wertheimer, Wolfang Köler, Kurt Koffka i Kurt Lewin. Es importante distinguirla de la terapia Gestalt, nacida en estados unidos en la década de los 60.

El término Gestalt proviene del alemán y fue introducido por primera vez por Christina Von Ehrenfels. No tiene una traducción única, se entiende generalmente como "forma". Aunque también se podría traducir como figura, configuración, estructura o creación.

La mente configura a través de ciertas leyes los elementos que llegan a ella a través de los canales sensoriales (percepción) o de la memoria (pensamiento, inteligencia y resolución de problemas). En la experiencia del medio ambiente, esta configuración tiene un carácter primario por encima de los elementos que la conforman, y la suma de estos por si solos no podría por tanto llevarnos a la comprensión del funcionamiento mental. Este planteamiento se ilustra con el enunciado: "el todo es más que la suma de las partes", con el cual se ha identificado con mayor frecuencia a esta escuela psicológica.

La psicología de la Gestalt es un movimiento experimental que empieza justo antes de la Primera Guerra Mundial. Su psicología está centrada en el estudio del conocimiento, la percepción y los procesos mentales. Es decir, en cómo nuestra percepción determina y condiciona nuestro pensamiento y, fundamentalmente, en el estudio del grado de conocimiento ya presente en el simple acto de percibir.

Para los psicólogos de la Gestalt, la visión consiste en un proceso de percepción activa en el que, más allá de la recepción pasiva de los rayos lumínicos, aparece implicada una serie básica de procesos activos de organización y estructuración de los estímulos.

La psicología de la Gestalt parte de una concepción sintética de la visión, expresada en una premisa básica:

Las formas exteriores son percibidas como totalidades o conjuntos superior esa la suma de sus partes, el sujeto se encarga de soldar las relaciones entre sus partes, o entre las partes de los estímulos recibidos.

En definitiva, la premisa básica de que el todo percibido es siempre superior a la suma de sus partes.

Desde la perspectiva de otras escuelas psicológicas alternativas, como la psicología cognitiva o la neurociencia computacional, se ha criticado la teoría gestaltista por no ofrecer un modelo del proceso de percepción. Por lo que ha sido considerada como una teoría redundante o no-informativa, en tanto que descriptiva y no-explicativa. Pero es precisamente su carácter descriptivo el que resulta útil para su aplicación en la práctica gráfica, especialmente en el diseño de interfaces. Si sabemos lo que pasa en el proceso perceptivo podemos tenerlo en cuenta al diseñar, aunque no tengamos una explicación de por qué pasa.

Los teóricos de la Gestalt y sus seguidores han usado cuatro principios clave para describir los procesos de percepción visual. Estos principios son:

a. Emergencia: reconocimiento global de objetos, emergiendo formas percibidas anteriormente de forma simultánea, no por la suma de sus partes constitutivas.

b. Reificación: construcción de nuevas formas partiendo de las existentes gracias a la propia experiencia visual.
c. Multiestabilidad: percepción ambigua entre fondo y figura que se dan en algunas imágenes.
d. Invariancia: prioridad en el reconocimiento y percepción de las formas y contornos de los elementos por encima de otras cualidades: color, textura, estilo, etc.

No tenemos que entenderlos como mecanismos que funcionen por separado, sino como aspectos distintos de un mismo sistema perceptivo dinámico. A continuación se explican los cuatro principios a partir de gráficos experimentales que sirven como demostración.

Ilustración 85. Prinicipio emergencia, figura dálmata.

El principio de la emergencia puede demostrarse por la percepción de la imagen del perro, que representa un perro dálmata olisqueando el suelo bajo la sombra de los árboles.

No reconocemos al perro a partir de la identificación previa de sus partes (patas, nariz, orejas, cola, etc.). Cuando reconocemos al perro lo hacemos globalmente, de una vez. La percepción del perro emerge en nuestra mente.

La reificación es un aspecto constructivo o generativo de la percepción a través del cual lo percibido mediante la experiencia visual contiene más información espacial explícita que el estímulo sensitivo en el que está basado.

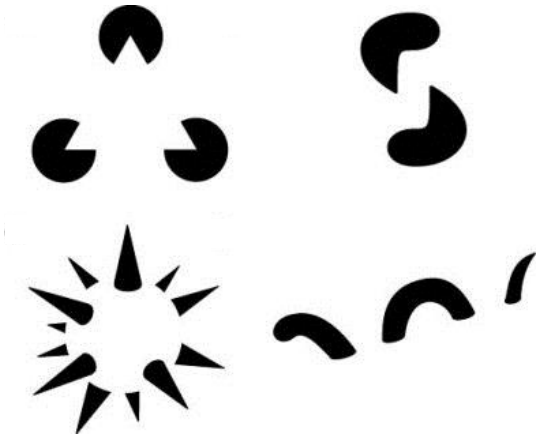


Ilustración 86. Imágenes efecto reitificación.

Por ejemplo, podemos percibir un triángulo en la figura de arriba a la izaquierda, y sin embargo no hay un triángulo dibujado en ella. En las imágenes de la derecha, la mente reconoce formas separadas como pertenecientes a una forma simple. En la imagen de abajo a la izquierda vemos una forma tridimensional completa, cuando en realidad no ha sido completamente dibujada.

La reificación puede ser explicada por el estudio de los contornos ilusorios, que son tratados por el sistema de percepción visual como contornos reales.

La multiestabilidad (o también percepción multiestable) es la tendencia que se da en las experiencias de percepción ambigua a saltar adelante y atrás de forma inestable entre dos o más interpretaciones alternativas. Podemos ver esta situación, por ejemplo, en el *Cubo de Necker* y en el *Vaso de Rubin* que se muestran a la izquierda. De nuevo la Gestalt no explica por qué esas imágenes son multiestables, sólo describe que el fenómeno ocurre.

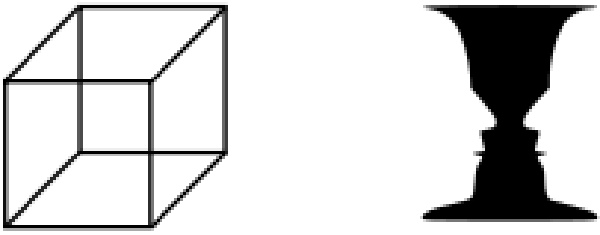
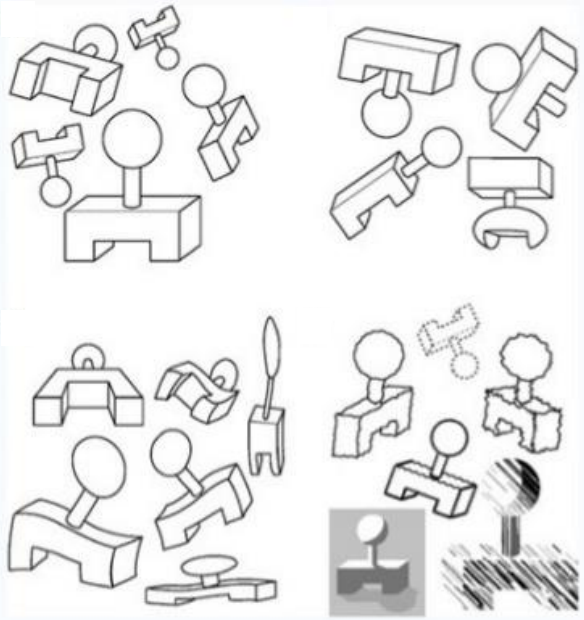


Ilustración 87. Cubo de Necker y Vaso de Rubin.

La invariancia es una propiedad de la percepción según la cual los objetos geométricamente simples son reconocidos independientemente de su rotación, traslación y escala. Incluso para otro tipo de variaciones como las deformaciones elásticas, diferencias de iluminación y cambios en las características de las partes que lo componen.



Por ejemplo, los objetos en la figura de arriba a la izquierda son reconocidos inmediatamente como la misma forma básica, los cuales podemos distinguir también de forma inmediata de las formas que hay a su derecha. Las reconocemos también a pesar de la perspectiva y las deformaciones elásticas en la imagen de abajo a la izquierda, e incluso cuando son representadas usando tratamientos gráficos distintos como en la última imagen.

2.3 Leyes de la visión aplicadas a la visualización de gráficos

Mediante estudios experimentales, los autores de la Gestalt llegaron a determinar una serie limitada de Leyes de la visión, un conjunto de principios descriptivos sobre el proceso de la percepción visual coherentes con la premisa básica de que toda nuestra experiencia perceptiva está estructurada y posee una tendencia a unir y a relacionar en grupos los estímulos percibidos.

A continuación se presentan las principales leyes que gobiernan la estructuración de los estímulos visuales según dichos autores.

2.3.1 Ley de la proximidad

Cuando los elementos están próximos en distancia, tendemos a agruparlos y formar 'todos' que separamos a su vez por la distancia.

Ilustración 88. Ley de proximidad.

En la primera figura vemos seis líneas paralelas, pero que a primera vista parecen tres bandas separadas por dos espacios mayores. En la segunda figura tendemos a relacionar los triángulos de manera horizontal, y nos es muy difícil percibir una relación vertical. En la tercera figura vemos que no influye la forma, el color, etc. sino que el único factor que logra esta agrupación es la proximidad; y no es intuitivo establecer otra relación entre los círculos, triángulos y cuadrados, más que tres conjuntos separados.

2.3.2 Ley de la semejanza o equivalencia

También tendemos a crear grupos cuando algunos de los elementos visualizados comparten ciertas características, como forma, color, tamaño, grosor o tipo de línea, etc. Para que se pueda apreciar esto, sin ser influido por la Ley de la proximidad, se han distribuido los elementos por igual distancia:

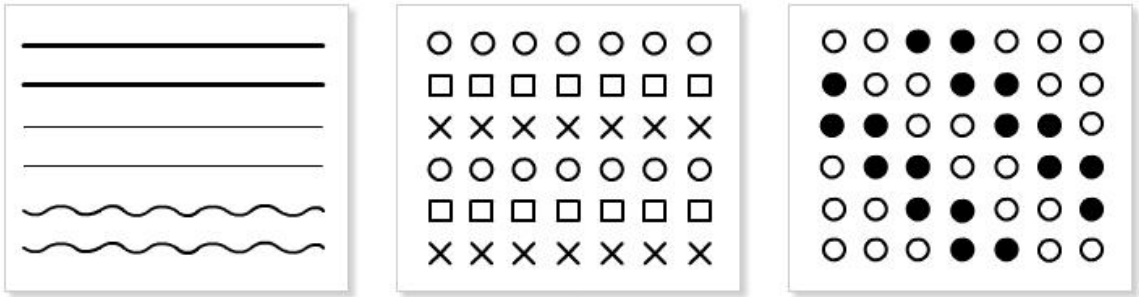


Ilustración 89. Ley de semejanza o equivalencia.

En la primera figura se observan seis líneas separadas por espacios iguales, pero que aun así agrupamos en tres conjuntos por la semejanza de grosor y tipo de línea. En la siguiente figura

lo que varían son las formas. Como son iguales en relación horizontal, tendemos a agruparlas de esa manera y nos resulta muy difícil establecer una relación vertical. En cambio, en la última figura las formas son idénticas, aunque percibimos semejanza por el tono, por lo que vemos dos bandas negras diagonales, una blanca central, y dos esquinas blancas separadas.

2.3.3 Ley de Prägnanz o de la buena forma

La palabra alemana Prägnanz -de difícil traducción-, es entendida como la tendencia a percibir las formas complejas de un modo más simple, simétrico, ordenado, resumido y esquematizado, de modo que sea más sencillo memorizarlas.

Nuestro sistema percibe siempre la mejor forma posible, considerando la totalidad de factores que coexisten, escogiendo la forma más sencilla, equilibrada, estable y regular.

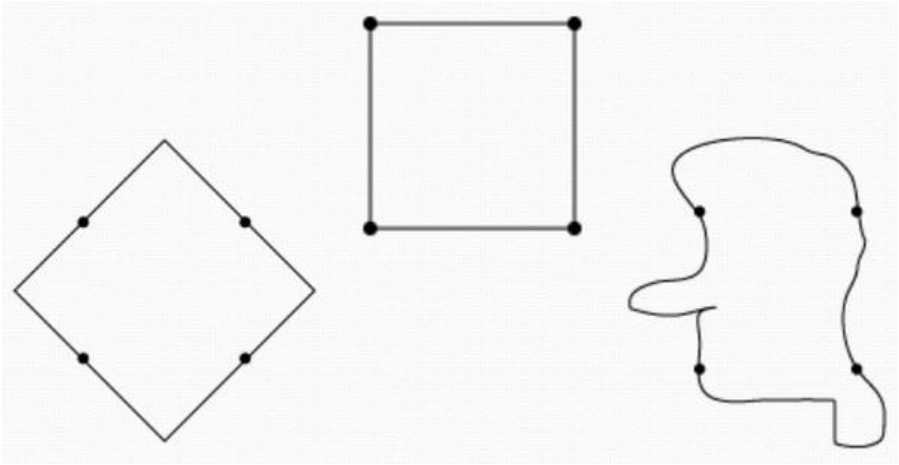


Ilustración 90. Ley de Prägnaz o de la buena forma.

Cuando observamos algo moverse, generalmente el móvil es un complejo conjunto de elementos, que agrupamos en un todo-simplista, gracias a que cada parte sigue un destino común. Y para referirnos a aquellas formas que ya de por sí son simples y fácilmente memorizables, se emplea la expresión "buena forma".

De esta manera, la Ley de Prägnanz hace notar la relevancia de tender hacia una estructura perceptual que sea lo más definida posible.

2.3.4 Ley de cierre

Esta ley es muy interesante, porque indica que la percepción es capaz de completar imágenes, con información que en ellas no hay. Veamos unos ejemplos.

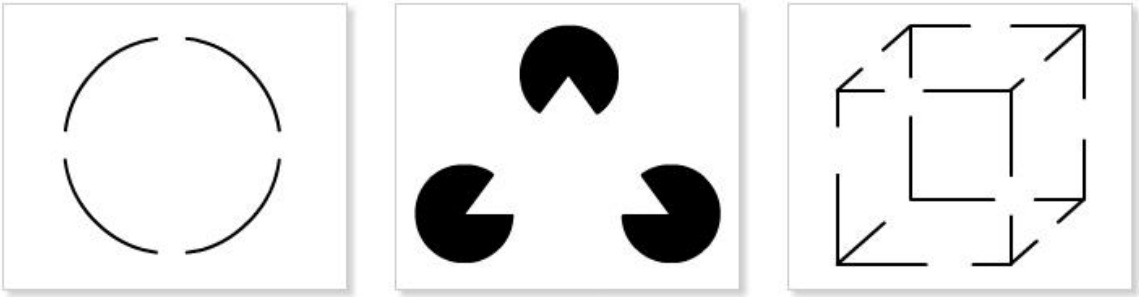


Ilustración 91. Ley de cierre.

La primera figura consta de cuatro líneas curvas, ajenas unas a las otras. Sin embargo, gracias a la buena forma de la figura que existiría si esas curvas se unieran -un círculo-, la percepción tiende a agregar la información que falta para que percibamos una figura coherente. Pero… ¿la información que añadimos, de dónde la obtenemos? Los estudios indican que la respuesta es la experiencia previa. Por ejemplo, la siguiente figura consiste en tres "pacman" que, como forman tres vértices de 60º, intuitivamente percibimos un triángulo equilátero como centro de la figura, que es una buena forma por excelencia, luego del círculo (Esta es una simplificación del triángulo de *Kanizsa*). Y la siguiente figura depende radicalmente de la experiencia previa. Alguien que nunca en su vida haya visto una imagen con perspectiva –con aparente profundidad–, percibirá un dibujo plano y desordenado, con seis flechas apuntando hacia fuera, una "L" en el centro, y una "Y" acostada. Pero la mayoría, seguramente verá un cubo, añadiendo la información que hace falta, proveniente su conocimiento de esta figura (*cubo de Necker*).

2.3.5 Ley de la simetría

¿Por qué todo es tan simétrico en el Universo? Es muy fácil notar que en la naturaleza encontramos simetrías por todas partes, como en nosotros mismos. Por ejemplo, casi todos los organismos tienen una mitad de su cuerpo exactamente igual -externamente- que la otra mitad. Hay teorías que indican que es causa de la gravedad, que seamos simétricos de izquierda y derecha pero no de arriba y abajo. Cuando una figura cumple con esta simetría de eje vertical, resulta ser mejor buena forma que otra con un eje distinto.

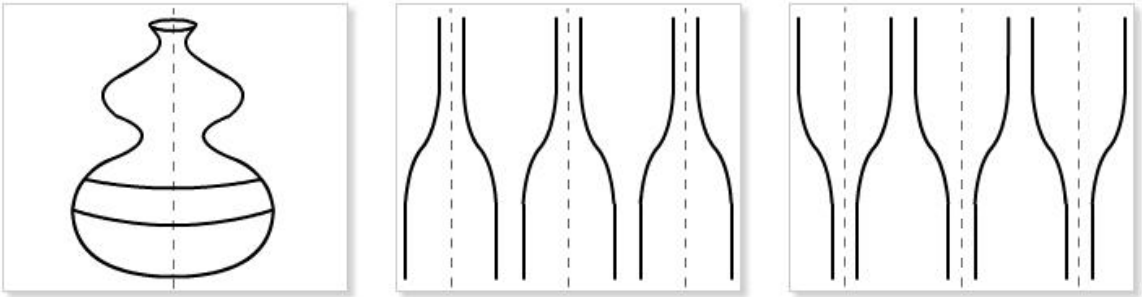


Ilustración 92. Ley de simetría.

En la figura de la izquierda vemos un caso de simetría con eje vertical, que es favorecida por la Prägnanz, es decir que es muchísimo más fácil de recordar que una figura no simétrica. En la figura del centro, en virtud de la ley de cierre, es intuitivo percibir tres botellas separadas. En cambio en la figura de la derecha, resulta fácil visualizar tres botellas invertidas. Pero mirando con atención se descubre que la forma (la línea curva) que se repite en ambas imágenes es exactamente la misma (por eso también se pueden ver en la figura del centro cuatro botellas invertidas, y en la figura de la derecha al revés).

2.3.6 Ley de la continuidad

Cuando observamos un paisaje, generalmente los elementos se hacen borrosos en dirección al horizonte. No sólo borrosos, sino también simples, resumidos y uniformes. Se trata de otra manifestación de la ley de Prägnanz.

Ilustración 93. Ley de continuidad. Fuente: Ciudad de São Paulo, Brasil. Copyright http://www.fotosedm.hpg.ig.com.br/

Vemos que las formas alejadas son casi indescifrables de por sí. Pero sabemos bien lo que son, gracias a que los elementos cercanos son nítidos, y esa nitidez se desvanece gradualmente. Para comprobarlo, se oculta la parte inferior de la fotografía, de modo que sólo se vean tres centímetros de la parte superior de la imagen. El contenido se vuelve indescifrable, pero al ver la imagen completa, se percibe que son edificios. El significado de esta ley es, entonces, que

podemos percibir cosas que no deberíamos poder, si no existiera una continuidad 'suave' entre los elementos.

2.3.7 Ley de la experiencia

Desde que fue propuesta, esta ley ha estado en debate psicológico, como también filosófico. Es seguro que, toda experiencia previa influye en la forma en que percibimos las figuras, pero lo que no sabemos es si lo modos de percibir, también los adquirimos de la experiencia, o son propios del ser humano. Es decir, se cuestiona si existen en el ser humano conocimientos nativos que no son influidos por la experiencia, que condicionan como debe percibir y qué elementos añadir. Es aquí en donde los diferentes autores de la psicología se empiezan a dividir. De esto se ocupó Immanuel Kant en el siglo XVIII, logrando una revolución en el rumbo del pensamiento en la humanidad.

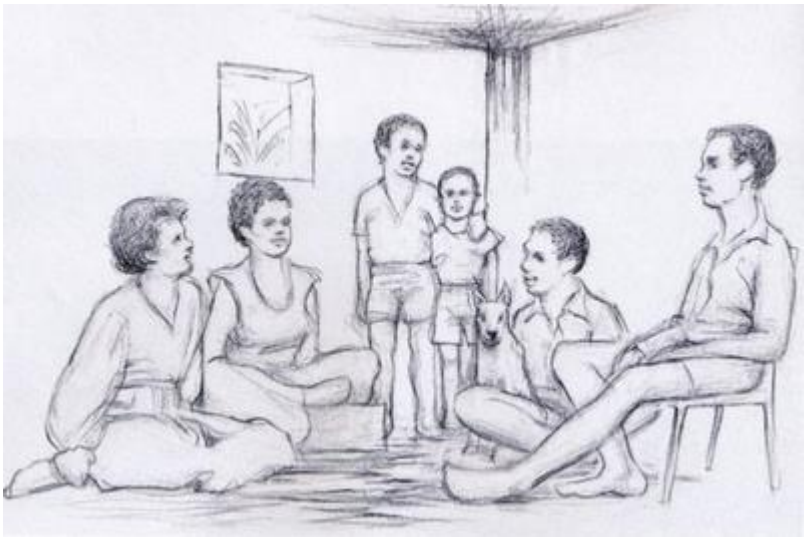


Ilustración 94. Ley de la experiencia.

Este dibujo, fue utilizado por Robert Laws (1851-1934) para poner a prueba esta ley. Cualquiera que esté habituado a vivir en ambientes angulares -como una casa-, percibirá que se trata de una habitación, cuyas paredes y techo están insinuados por el trazado central, en virtud de la ley de cierre, y por la ventana del lado izquierdo. Pero cuando este dibujo les fue mostrado a personas residentes en Malawi (sudeste de África), la mayor parte respondió que la familia que allí aparece se encuentra bajo un árbol, y que la "ventana" del lado izquierdo es una caja que reposa sobre la cabeza de la mujer. Por ello se puede afirmar que el ámbito social y cultural resulta ser un factor clave en la interpretación de las figuras.

2.3.8 Ley de la figura-fondo

Es la ley más utilizada de todas, a la hora de crear ilusiones ópticas. Se trata de que una figura pueda tomar el papel de fondo, mientas que el fondo pueda convertirse en

una figura. Pero ¿qué son en verdad los fondos y las figuras? En verdad, no son nada en sí son lo que queremos que sean. Intuitivamente, consideramos que los fondos se extienden por detrás de las figuras, y que suelen ser difusos y menos llamativos; mientras que las figuras son cuerpos definidos, mejor estructurados, y que prevalecen sobre el fondo. En el mundo del Arte, jugar con estos dos conceptos, haciendo que ambos tengan simultáneamente las propiedades nombradas, es un recurso valiosísimo, como en los siguientes ejemplos.

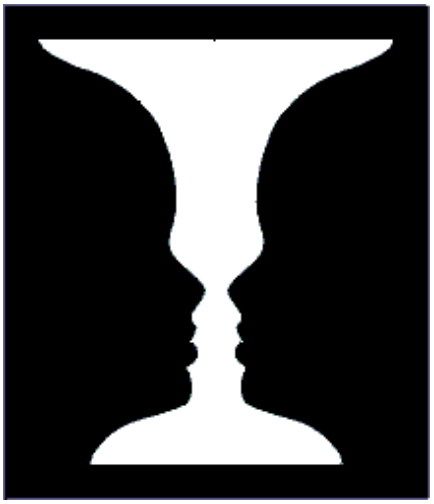


Ilustración 95. Jarrón de Rubin.

Este es el ejemplo más conocido de todos. Se trata del Jarrón de Rubin, de Edgar Rubin, en donde el fondo se convierte en figura y la figura en fondo. Nota cómo saca partido también de la Ley de simetría y de la buena forma del jarrón y los perfiles.

Otro ejemplo no menos impactante es "Mariposa a la Flor", de Octavio Ocampo. Allí podemos ver una planta y una mariposa como figura, y el resto rosado como fondo.

O bien, podemos percibir ese resto rosado como figura -el rostro de una mujer- y la planta y mariposa como condicionantes para esa figura, pero no como figuras en sí.

Ilustración 96. Mariposa a la Flor, de Octavio Ocampo.

2.4 El pensamiento visual

Lo que el pensamiento de los investigadores de la Gestalt nos ha dejado como legado es una mayor comprensión de la percepción visual como **actividad cognitiva**. Es decir, como una actividad que implica un grado de inteligencia y pensamiento. Todo pensamiento es de naturaleza perceptual, y la vieja dicotomía entre percepción y pensamiento resulta falsa.

Los procesos básicos de la visión implican mecanismos típicos del razonamiento. La visión, lejos de constituir una función inferior, resulta el medio fundamental por el que estructuramos los acontecimientos de los que derivamos las ideas y el lenguaje. En definitiva, los factores visuales determinan la formación de conceptos y, así, el material de nuestro pensamiento. Los estudios de los límites de la visión –itinerarios oculares, retentiva visual, agudeza visual y preferencias o perversiones de la mirada– señalan, a su vez, la posibilidad o imposibilidad de determinados actos de pensamiento:

a. **Retentiva visual**: no vemos todo lo que entra por nuestros ojos. La capacidad de retentiva visual está en 1/15 partes de segundo. Es decir, el cerebro solamente ve como independientes todos aquellos estímulos que aparecen ante los ojos a una velocidad igual o inferior. A velocidades superiores, el cerebro los funde produciendo la sensación de movimiento.
b. **Agudeza visual**: igual que en el caso anterior, el de la velocidad, existe también un umbral de la percepción referente al tamaño. Por debajo de un determinado umbral, el cerebro funde los estímulos recibidos.
c. **Preferencias de la mirada**: más allá de las distintas leyes básicas de la Gestalt, la visión humana presenta unas preferencias que determinan jerarquías visuales. Así, en igualdad de circunstancias, por ejemplo, existe una preferencia por la claridad frente a la oscuridad, por los colores de mayor intensidad frente a los más apagados, o por las figuras simples frente a las complejas.

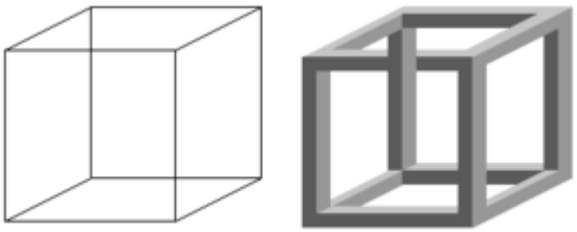


Ilustración 97. Preferencias de la mirada.

d. **Perversiones de la mirada**: asimismo, hay un conjunto de anomalías en la visión que pueden caracterizarse, a su vez, como trastornos del reconocimiento visual. Entre otras, destacaremos la ilusión de Müller-Lyer o el cubo de Necker. En ambos casos, el sujeto que mira sufre una discrepancia entre el precepto y la realidad objetiva que lo ha suscitado, es decir, una

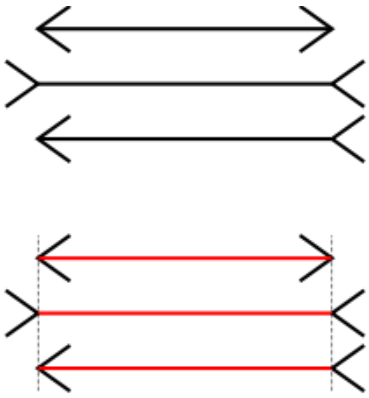


Ilustración 98. Perversiones de la mirada.

distorsión subjetiva (precepto) de un estímulo objetivamente presente (realidad).

Hay que tener en cuenta que algunas de las características de nuestro sistema perceptivo, que ponen de relieve las Leyes de la visión de la Gestalt y otros experimentos de la psicología de la percepción, han sido modeladas por la evolución a través de la selección natural. Son características que permiten la adaptación del sistema perceptivo al entorno, y por ello, nuestra adaptación al mismo. Sin embargo existen otras características que no son innatas, sino que se dan por la influencia de los factores culturales propios de las sociedades dotadas de un pensamiento abstracto. Esta posibilidad es la que pone en entredicho las Leyes de la Gestalt como leyes de carácter universal (Gubern, 1987).

En 1984, dos investigadores de los AT&T Bell Labs, William S. Cleveland y Robert McGill, publicaron un artículo fundamental en el Journal of the American Statistical Association. Se titulaba 'Graphical perception: theory, experimentation, and application to the development of graphical methods'. Sus contenidos continúan siendo relevantes para cualquier análisis racional de la infografía y la visualización. Es un texto poco conocido entre diseñadores y periodistas, por desgracia, aunque influyente en otros círculos: expertos como Stephen Few y Naomi Robbins continúan hoy la senda abierta por 'Graphical Perception' y otros escritos de los mismos autores. Lo más interesante del artículo es que propuso una guía para elegir la forma visual correcta dependiendo de la función que cada gráfico tenga. Cleveland y McGill listaron diez métodos para codificar datos visualmente (los llamaron elementary perceptual tasks) y los clasificaron, de mejor a peor, a partir de lo bien que el cerebro es capaz de interpretarlos

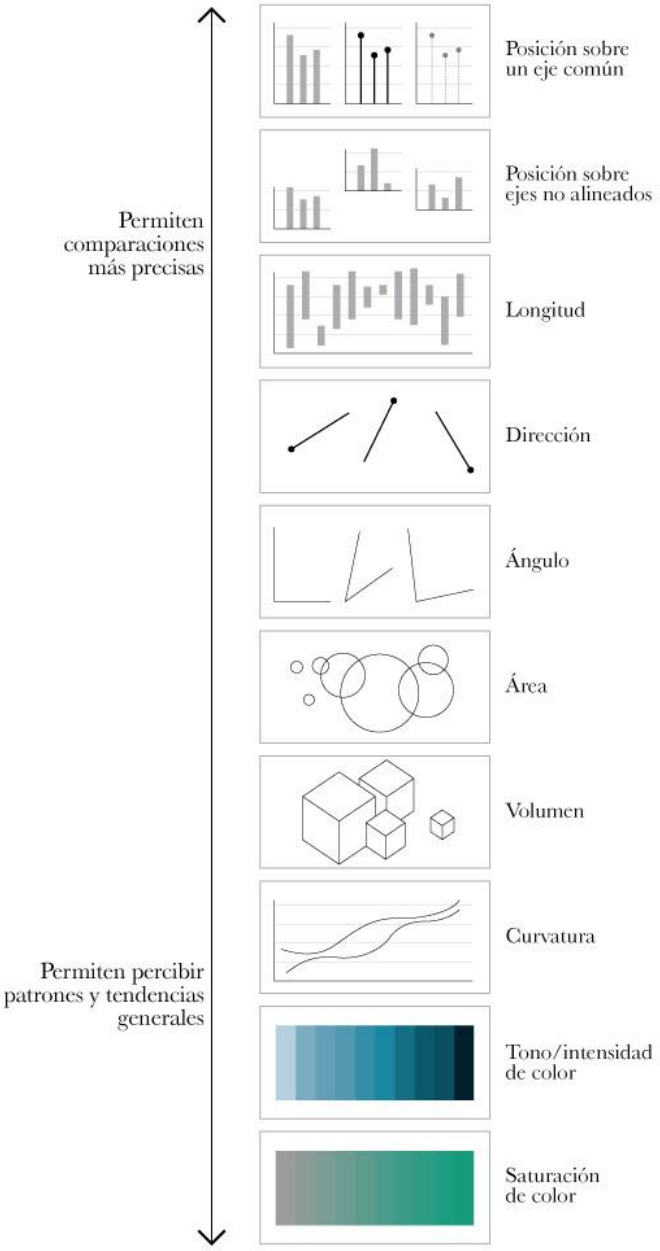


Ilustración 99. Teoría de la percepción gráfica, Cleveland & McGill.

y comparar sus elementos. Son:

a. Posición sobre un eje común.
b. Posición sobre ejes no alineados.
c. Longitud
d. Dirección
e. Ángulo.
f. Área.
g. Volumen, curvatura.
h. Intensidad (tono)
i. Saturación de color.

La jerarquía de Cleveland y McGill no se basa en meras intuiciones, sino en experimentos con usuarios. Una de las conclusiones del artículo fue que cuanto más detalladas y precisas sean las operaciones que los lectores deben completar con un gráfico, más alta en la jerarquía debe estar la forma de representar los datos utilizada. En otras palabras: un gráfico de barras es siempre superior a un gráfico de burbujas si el objetivo central del diseñador es facilitar comparaciones exactas entre números. La escala de Cleveland y McGill permite también razonar por qué los gráficos de tarta casi nunca son útiles.

Los tres ejemplos bajo este párrafo codifican las mismas cifras de tres formes diferentes: barras, burbujas y tonos de color; sabiendo los valores que corresponden al primer elemento y al último, y por ello podríamos preguntarnos cual de estos gráficos permite comparar mejor. La respuesta correcta es el de barras. De hecho, el cerebro humano subestima los tamaños relativos de los círculos: observe que la diferencia entre la segunda barra y la primera parece mucho mayor que la que existe entre la segunda burbuja y la primera. Las burbujas engañan.

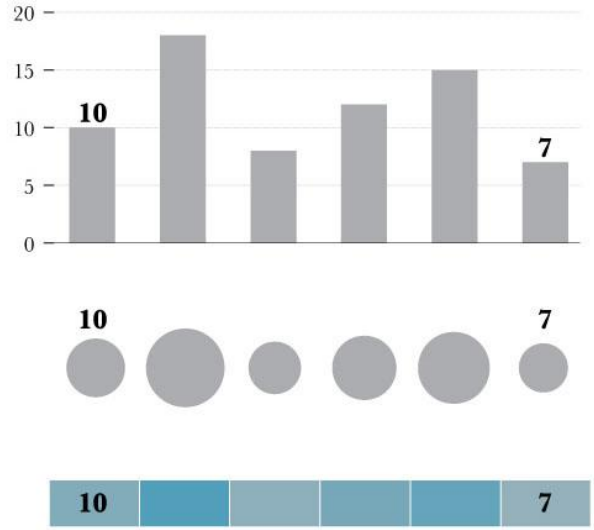


Ilustración 100. Comparar gráfico de barras, burbujas y tonalidades.

Otra aportación interesante, es la de Sthephen Kosslin, autor de diferentes libros y artículos académicos sobre comunicación visual.

Kosslin hace una distinción entre gráficos figurativos (depictive graphics) y gráficos simbólicos (symbolic graphics). Un gráfico figurativo es aquel que tiene una semejanza física con la realidad que representa, como el plano de un apartamento, la explicación de cómo funciona un aparato, el mapa de una región, etc.

En un gráfico simbólico, por el contrario, la relación con el fenómeno representado es formal: por ejemplo, en los gráficos estadísticos.

La forma de leer e interpretar esos dos tipos de gráfico depende de cada persona. La lectura de los primeros es sencilla pero, para entender los segundos, uno debe aprender ciertas convenciones: que existen ejes X e Y, que la altura de las barras es proporcional a las cantidades que codifican, etc. Hoy en día, los gráficos estadísticos son muy comunes, por lo que pensamos que su lectura es natural, pero en realidad, no lo es.

2.5 Conclusiones para una correcta representación de datos

De todo lo anteriormente analizado se muestran a continuación algunas conclusiones que deben tenerse en cuenta cuando se realizan representaciones visuales, para hacer comprensibles los datos:

a. El grupo – objetivo: Diferentes formas de presentación pueden ser adecuadas para diferentes públicos (por ejemplo para directivos, público general o clientes de una organización).
b. El papel de la gráfica en la presentación: analizar el cuadro completo o centrar la atención en puntos clave puede requerir de componentes de visualización diferentes.
c. Como y donde será presentado el dato: ¿un informe para imprimir o un cuadro de mando interactivo?
d. Aspectos contextuales que pueden distorsionar la comprensión. Grado de conocimiento de los datos por parte del usuario, ¿experto o novato?
e. Coherencia en las representaciones visuales de datos. Asegurar que las representaciones visuales se diseñan de forma coherente y usar las convenciones habituales cuando sea posible (rojo para dato en alerta, azul para representar el agua en un mapa, etc.)

De todos estos aspectos, se desprende la necesidad de definir un catálogo de componentes de visualización de datos que indique cuales son los objetivos y los usos adecuados de cada componente.

3 Definición de un catálogo de componentes para visualización

Ralph Lengler & Martin J. Eppler publican un estudio en el que evalúan más de 160 métodos de representación gráfica. Una vez determinados estos métodos de representación, reducen el número a un total de una centena, basándose en la aplicación de los siguientes criterios:

- El método debe de estar documentado en todos sus pasos.
- El método debe haber sido aplicado en la vida real, preferentemente en el mundo empresarial.
- El método debe ser indicado para la representación de conocimiento relevante.
- El método debe ser aplicable para no expertos.
- El método debe de haber sido evaluado anteriormente de alguna manera.

A cada uno de los métodos, aplican la siguiente clasificación:

a. Complejidad de visualización: Refiriéndose al número de reglas que deben aplicarse para la utilización del elemento
    - Se clasifica de baja a alta

b. Aplicación principal o área de contenido: Se distinguen seis tipos básicos de aplicaciones:
    - Datos
    - Información
    - Concepto
    - Metáfora
    - Estrategia
    - Conocimiento compuesto

c. Punto de vista:
    - Detalle (remarcando ítems individuales)
    - Visión de conjunto
    - Detalle y visión de conjunto a la vez

d. Tipo de ayuda al razonamiento:
    - Convergente : reduce la complejidad
    - Divergente: añade complejidad

e. Tipo de representación:
    - Proceso: cíclico o continuo
    - Estructura, por ejemplo jerarquía o redes causales.

Los autores organizaron los métodos en base a estas categorías mediante una tabla inspirada en la tabla de elementos químicos, como se muestra a continuación:

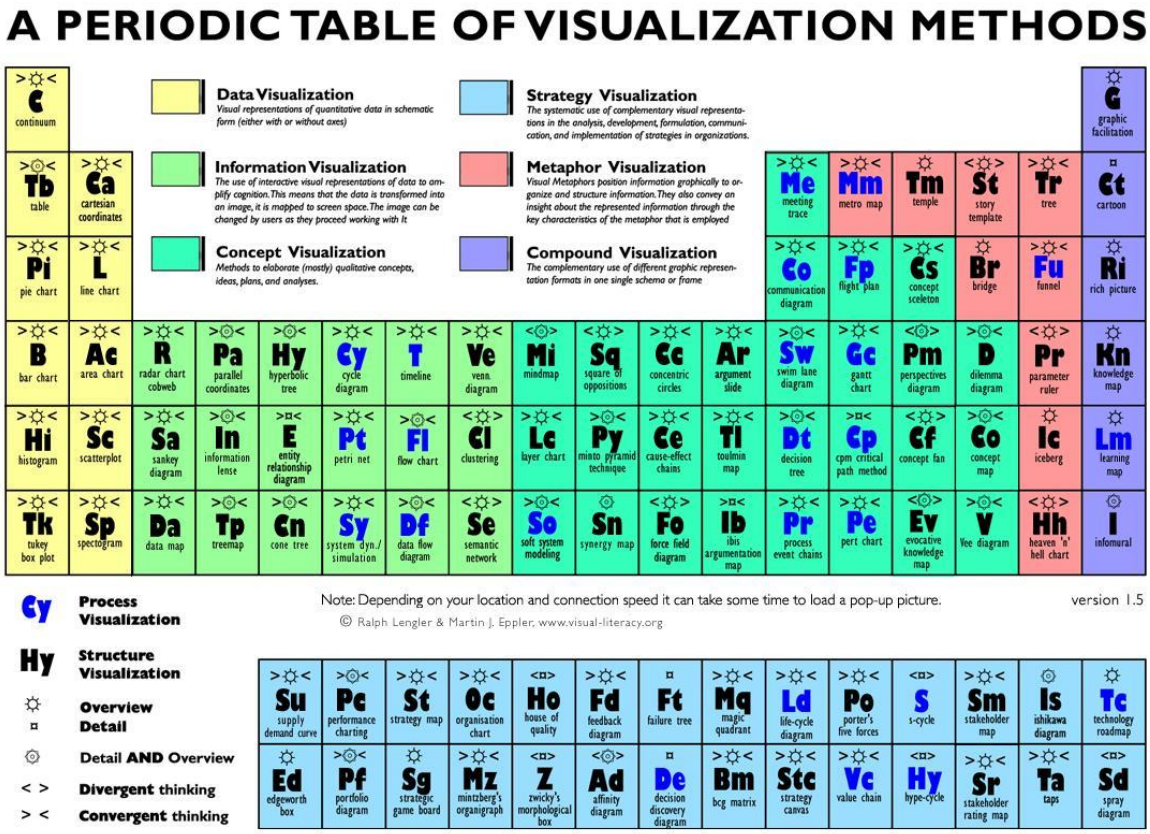


Ilustración 101. Tabla de elementos de visualización de Ralph Lengler & Martin J. Eppler.

Las clasificaciones que se usan en esta tabla para organizar los diferentes métodos de visualización son las siguientes:

Visualización de datos: Incluye los métodos tradicionales de visualización de datos cuantitativos, como por ejemplo gráficos de líneas, de tarta o de área. Son representaciones de datos cuantitativos de forma esquemática. Suelen responder a preguntas sobre el cuánto y sobre el cómo y el porqué, como metáforas visuales.

Visualización de información: Como por ejemplo los treemaps o un diagrama de flujos. Uso de representaciones visuales interactivas de datos para ampliar el conocimiento. El dato se transforma en una imagen. El usuario puede cambiar la imagen conforme va interactuando con ella.

Visualización de conceptos: Como por ejemplo en un mapa conceptual, un diagrama de Gantt. Son métodos usados principalmente para elaborar conceptos cualitativos, ideas, planes, y analizarlos mediante la ayuda de procedimientos estructurados de grafiado de ideas.

Visualización de metáforas: como por ejemplo en un mapa de las líneas del metro. Estos elementos dan respuesta a dos funciones, de una parte posicionan información geográficamente, y de otra, transmiten información acerca de las características clave de la metáfora que ha sido utilizada.

Visualización estratégica: Como por ejemplo con el modelo de negocio de Canvas, como herramientas de representación visual complementarias, para mejorar el análisis, desarrollo, formulación, comunicación, e implementación de estrategias en las organizaciones. Este es uno de los grupos de métodos más específico.

Visualización compuesta: Une varios de los métodos anteriores. Pueden ser desde complejos mapas que contienen diagramas y elementos metafóricos, a gráficos conceptuales unidos a gráficos cuantitativos, o cualquier otra combinación. Generalmente se usa esta clasificación cuando es preciso el uso complementario de diferentes métodos de representación en un único marco o esquema.

Para el modelo que se propone y que constituye el objetivo de este trabajo de investigación, se considera adecuado escoger los métodos de representación gráfica que son útiles para la representación de datos en el ámbito organizacional, de entre todos los presentados en la tabla de elementos de Lengler & Eppler.

Se han escogido fundamentalmente los métodos de representación de datos y algunos de los del grupo de representación de información.

Se clasifican los métodos en básicos y avanzados. Los métodos básicos los puede utilizar cualquier organización puesto que son métodos ampliamente conocidos y entendidos. Los métodos avanzados, permiten análisis más complejos, y requieren de cierto conocimiento o entrenamiento por parte de las personas de la organización.

Se ha añadido algún método que no consta en el estudio de Lengler & Eppler, puesto que este ámbito de conocimiento está evolucionando continuamente y se crean nuevos componentes de visualización muy frecuentemente.

4 Propuesta de un Catálogo de componentes para visualización de datos

A continuación se detalla el Catálogo de Componentes para visualización de datos para las organizaciones, que las organizaciones podrían de forma adecuada en función de los datos a presentar.

Tal y como se ha avanzado en el apartado anterior, se han clasificado los compontes en dos grupos: básicos y avanzados.

El grupo de componentes gráficos básicos los puede utilizar cualquier organización, y no precisan de una formación específica para su uso, pero si debe tenerse en cuenta la adecuada utilización en función del tipo de datos que quiere representarse.

El grupo de componentes avanzado requiere de cierto conocimiento para su interpretación y por ello deben utilizarse con cierto grado de acompañamiento y formación. Se aconseja su introducción de forma paulatina y con textos de apoyo para la interpretación.

Componentes gráficos básicos:

1. Tablas de datos
2. Gráfico de líneas o de áreas.
3. Gráfico de coordenadas. Gráfico de dispersión.
4. Gráfico de barras o columnas.
5. Gráfico de secciones o circular.
6. Mapa.

Componentes gráficos avanzados

1. Diagrama de caja
2. Treemap, Mosaic plot
3. Sankey. Gráfico de Chord chart. Gráfico de coordenadas paralelas
4. Heat map, Espectrograma
5. Sparklines
6. Nubes de palabras

4.1 Tablas de datos

Descripción: Una tabla de datos adecuada es parte esencial de la comunicación de un conjunto de datos. Utilizar tablas de forma efectiva evita tener que discutir sobre variables poco significativas que no son esenciales para la comprensión de los datos.

Miller (2004) da unas directrices sobre como diseñar buenas tablas:

- Que sea fácil para la audiencia encontrar y entender los números en las tablas.
- Diseñar de manera discreta y sencilla tanto el formato como el título de las tablas, de forma que la atención se centre en los puntos sustanciales expresados por los datos más que por la propia estructura de la tabla.

En las tablas de datos, los datos deben ser presentados de forma concisa y bien organizada para apoyar el análisis. Una tabla pequeña bien realizada puede proporcionar una gran cantidad de información que el usuario puede asimilar rápidamente.

Elementos a tener en cuenta para el buen diseño de una tabla:

- El título de la tabla, con una descripción clara y precisa de los datos.
- Los encabezados de las columnas, expuestos en la parte superior de la tabla, deben indicar qué datos hay presentes en cada columna de la tabla.
- Los encabezados de las filas
- Las notas al pie
- La línea que contiene la fuente y la fecha de actualización o carga de datos.

Tipos de datos: las tablas son adecuadas para presentar cualquier tipo de datos, pero fundamentalmente datos numéricos. Son relevantes por que facilitan la comparación de datos en las mismas filas o columnas. En algunos informes, las tablas se utilizan para explicar la segmentación de un dato en una serie temporal, por ejemplo.

Algunos consejos:

- Toda la información para entender la tabla debe de mostrarse en el título y subtítulo.
- Indicar la fuente de la que provienen los datos.
- Unificar criterios de alineación y número de decimales.
- Si se utilizan abreviaturas explicarlas al pie.
- Cuando en un informe se utilizan diferentes tablas, es recomendable utilizar siempre el mismo formato.
- Evitar recargar el formato con la utilización de bordes, sombreados, colores, etc.

Aplicaciones de este componente: Este componente es fuertemente utilizado en todas las organizaciones, y puede aplicarse siempre que quieran mostrarse datos en filas y columnas.

4.2 Gráfico de líneas

Descripción: Es una colección de observaciones de una variable (valores) recogidas secuencialmente en el tiempo. Estas observaciones se suelen recoger en instantes de tiempo equiespaciados, o bien acumular los valores sobre intervalos de tiempo. Es un componente eficaz para la visualización de tendencias.

Tipos de datos: Representar una variable y sus valores sobre un eje temporal. En este caso, podemos considerar la variable tiempo (del eje X) como una variable neutra.

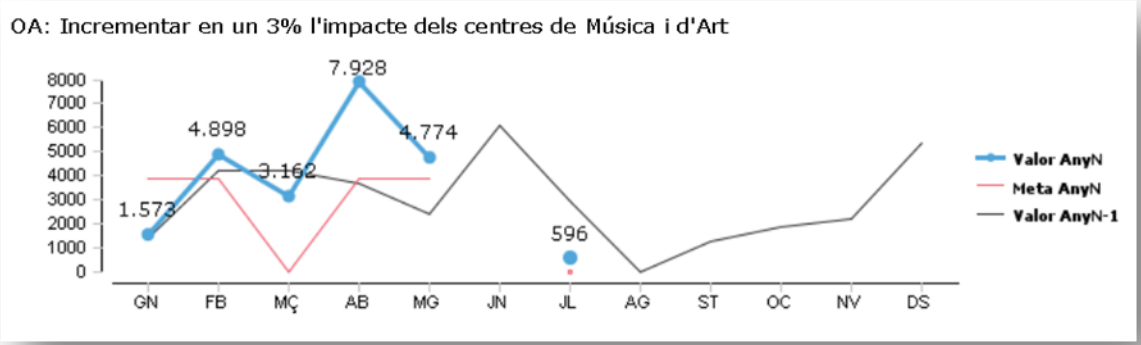


Ilustración 102. Gráfico de líneas. Fuente: Elaboración propia.

Variantes: Gráfico de área, que muestra colores distintos en las áreas situadas debajo de las líneas

Aplicaciones de este componente: representar la evolución en el tiempo de un dato cuando se muestra el tiempo en el eje x. Siempre que queramos representar la evolución temporal es aconsejable utilizar un gráfico de línea, no de barras o columnas.

4.3 Gráfico de coordenadas, gráfico de dispersión.

Descripción: Se utiliza para mostrar la relación entre dos variables. Es la manera más exacta de representar correlaciones entre dos variables. Algunos autores prefieren otro tipo de gráficos, puesto que los gráficos de dispersión pueden ser difíciles de interpretar para personas sin experiencia.

Tipos de datos: Representación de la relación entre dos variables que están representadas en los ejes X y Y.

Variantes: Existen muchas variaciones de este tipo de gráfico, pero las más comunes son:

- Bubble chart y Bubble chart con movimiento.
- Gráfico de clúster.

A este tipo de gráficos también se les llama "Scatterplots".

El Bubble chart es un gráfico de dispersión en el que se añade una tercera variable, que es el tamaño del punto de intersección entre las dos primeras variables, así mostramos el tercer valor con el tamaño del punto, convirtiéndolo en burbuja. En ocasiones, podemos añadir una cuarta variable incluyendo el color en la burbuja, o la graduación del color.

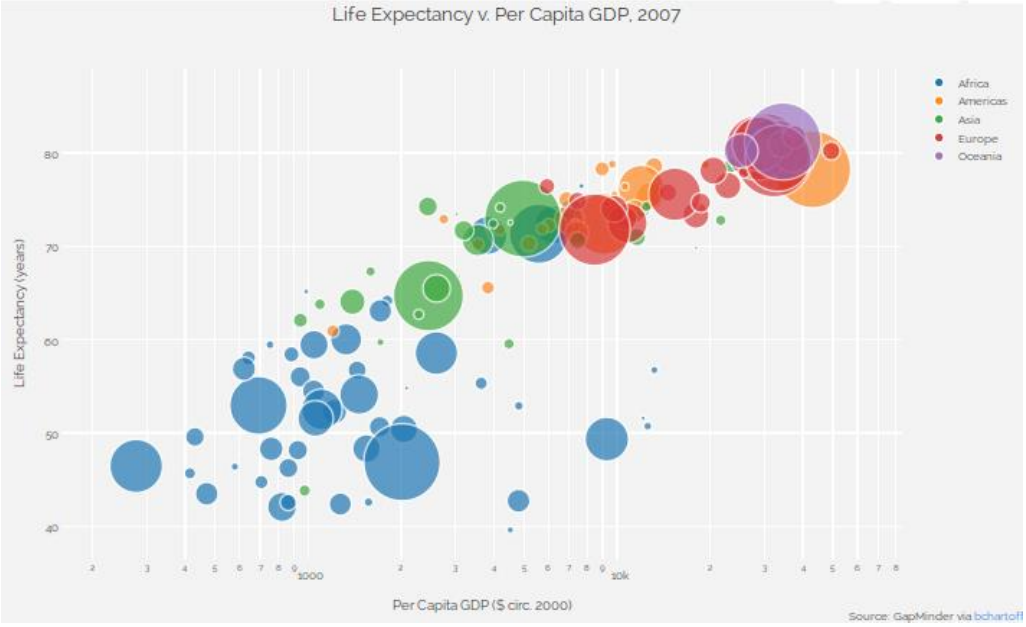


Ilustración 103. Bubble chart. Fuente: http://help.plot.ly/make-a-bubble-chart/ (2015)

Las presentaciones interactivas pueden incluso llegar a incluir una quinta variable que es la evolución temporal de los datos, mostrando con el movimiento de las burbujas como han ido evolucionando los datos (crecimiento o decrecimiento del tamaño de la burbuja, evolución de la intensidad del color, desplazamiento en el eje x y en el z). Ver un ejemplo en https://www.gapminder.org/ .

Aplicaciones de este componente: Puede aplicarse siempre que quiera conocerse la relación entre dos variables. Suele utilizarse para representar datos en matrices de relación, como en el ejemplo, la matriz Satisfacción/Importancia.

4.4 Gráfico de barras o columnas

Descripción: Un gráfico de barras o columnas es el tipo de gráfico más fácil de representar e interpretar junto con el de líneas. Se utiliza para representar y comparar las frecuencias o los valores de las distintas categorías o grupos de datos.

Tipos de datos: Representar las categorías de una variable o, también, de varias variables relacionadas. Muestran comparaciones.

Variantes: Existen muchas variaciones de este gráfico, pero las más comunes son:

- Columnas o barras simples
- Columnas o barras comparadas
- Columnas o barras apiladas, agrupadas o no por porcentaje.
- Diferenciales o win-lost
- Pirámide de población
- Clúster consolidado
- Histograma
- Pareto

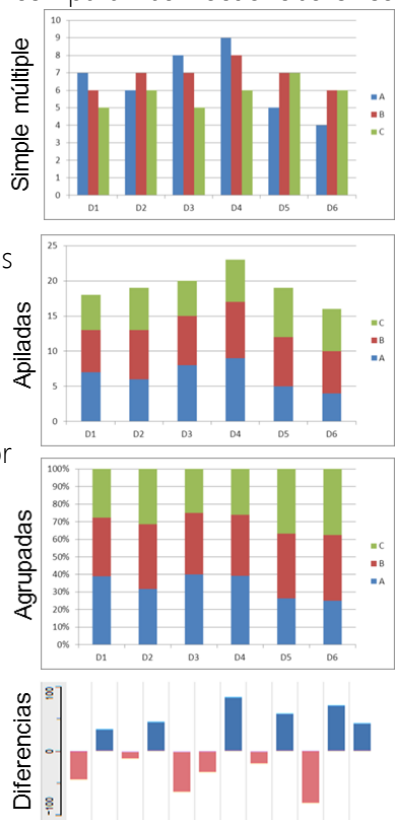


Ilustración 104. Variantes del gráfico de columnas.

Puede también variar la forma, siendo conos o pirámides en lugar de columnas o barras.

Aplicaciones de este componente: Es conveniente aplicar este componente, siempre que se quieran mostrar comparaciones, pero con las siguientes limitaciones:

- Pocos de los valores que desea trazar son valores cero.
- Tiene pocas categorías (no más de siete categorías).

Un gráfico a evitar es el diagrama de barras apiladas. Este tipo de gráfico se suele utilizar para representar porcentajes (o valores absolutos) en distintas unidades muestrales, pero es muy fácil que la representación gráfica no sea representativa, y no se distingan bien las diferencias entre porcentajes. A menudo, es preferible desglosar la información en varios diagramas de barras sencillos, uno sobre otro, o en un diagrama de barras más complejo, con varias barras por unidad muestral.

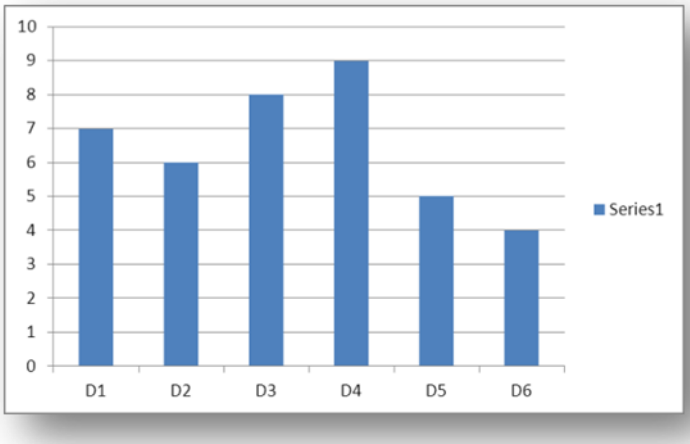


Ilustración 105. Gráfico de barras o columnas. Fuente: Elaboración propia.

4.5 Gráfico de secciones o circular

Descripción: Es un gráfico en el que los datos se muestran como una sección de un circulo. Muestra generalmente la distribución porcentual de una variable.

Tipos de datos: Por lo general, no debería usarse para mostrar más de seis categorías de datos. El uso de este tipo de gráfico no es recomendado por muchos estadísticos, ya que puede llegar a ser muy difícil de comparar los diferentes datos representados en un gráfico, y aún más entre dos gráficos. Para evitar este problema, se pueden indicar los valores porcentuales en cada segmento. Es importante ordenar los segmentos de mayor a menor, no intercalando segmentos pequeños entre los mayores.

En la mayoría de los casos, otro tipo de gráfico es más indicado, generalmente el de barras o columnas, pero los gráficos circulares no deben descartarse si el objetivo que se persigue es mostrar la importancia relativa de una categoría en el total, es decir para mostrar una visión general.

Variantes:

- Gráfico de radar, en el que las categorías se muestra en un gráfico de línea circular.
- Gráfico de anillo. En esta variante el círculo se substituye por un aro o anillo, que complica aún más su interpretación.

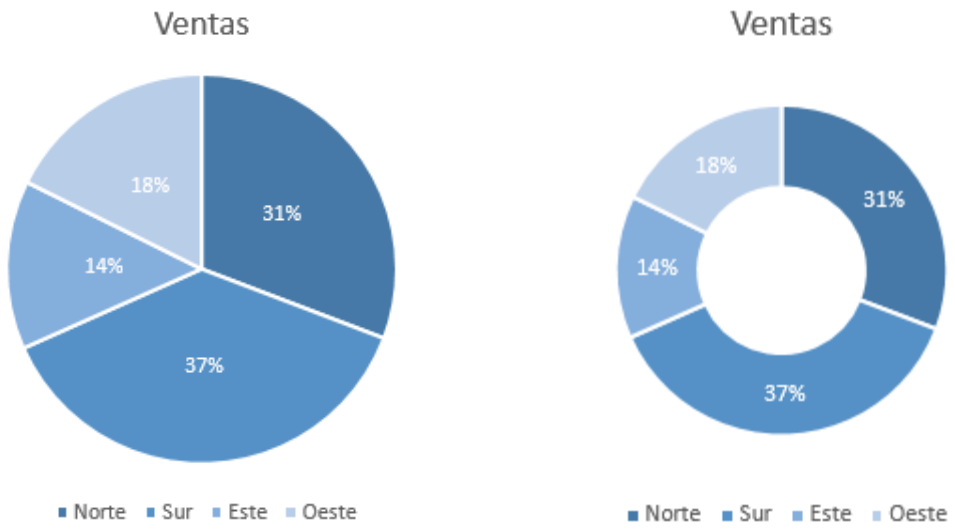


Ilustración 106. Gráfico de sección o circular.

Aplicaciones de este componente: Como ya se ha comentado, se suele justificar su uso para presentar datos porcentuales, por lo visual que resulta aparentemente. Sin embargo, este tipo de gráfico eso bien redundante o bien muy confuso. Si el círculo tiene sólo dos categorías, es redundante: el porcentaje puede presentarse directamente en el texto. Si la tarta tiene tres o más categorías, normalmente es confuso porque cuesta distinguir cuál de las porciones es mayor y cuál es menor. En ocasiones este problema se resuelve añadiendo al diagrama los porcentajes numéricos, al lado de cada porción. Eso vuelve a la gráfica derrochadora, pues la hace equivalente a una tabla con dichos porcentajes (una representación mucho más económica e igualmente clara de la información). Si es necesario representar información porcentual, es mucho más recomendable hacerlo o directamente en el texto, o en forma de tabla o en forma de diagrama de barras.

4.6 Mapa

Descripción: La información geográfica es una parte relevante de los datos estadísticos. Las áreas geográficas tienen límites físicos, nombres y otra información que permite su geolocalización.

Los mapas son las herramientas más útiles para visualizar los patrones espaciales. Ayudan a identificar y percibir las distribuciones y patrones que pueden no ser evidentes en tablas o gráficos. Sirven como valiosas herramientas para expertos, políticos y público en general.

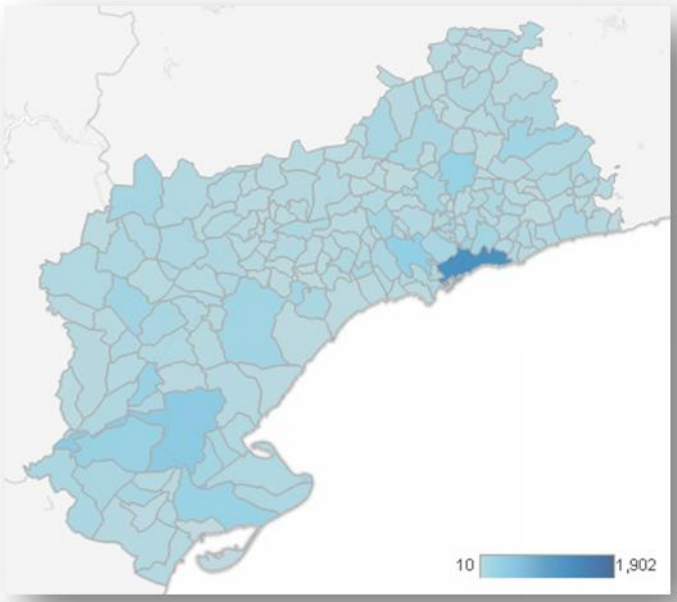
Tipos de datos: Es adecuado para mostrar la distribución espacial de datos y su localización geográfica. Permite comparar las diferentes áreas del mapa. Se reduce sensiblemente la complejidad de los datos, al mostrarlos gráficamente por zonas o áreas. Un mapa está bien diseñado si es fácil de entender.

No se pueden representar en este tipo de gráficos valores absolutos, si no ratios, rangos o frecuencias.

La clasificación de datos es una cuestión importante en la técnica de los mapas. Los patrones espaciales mostrados en el mapa, están representados por la agrupación de los valores de los datos. Se puede producir distintos mapas con los mismos datos mediante la aplicación de diferentes métodos de clasificación, dividiendo los datos en un número de clases distinto o seleccionando diferentes límites para las clases.

Variantes: Existen muchas variantes de este componente en función del tipo de mapa que se utilice, y del grado de simplificación o geometrización que se aplique. Este componente, como los anteriores, puede utilizarse en formato estático o interactivo, que clicando en cada zona del mapa nos muestre información detallada de la zona escogida.

El tipo más común de mapa utilizado en general es el mapa de coropletas, en el que las áreas se colorean en función del valor de la variable que se muestra. Con este tipo de mapas resulta fácil visualizar patrones, como se muestra en el ejemplo.



Otra variante frecuentemente utilizada es el mapa de puntos. Muestra la ubicación y la densidad de una variable en un mapa. Permite a los usuarios captar rápidamente la magnitud general de los datos, así como su concentración o dispersión. Cada punto representa un valor discreto, por lo general un gran número de entidades.

Ilustración 107. Mapa de coropletas.

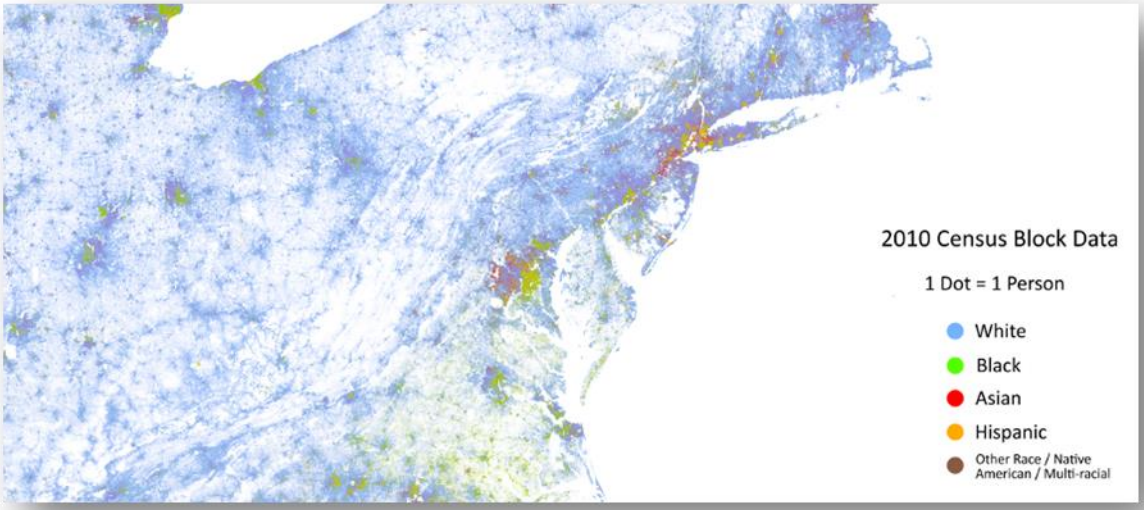


Ilustración 108. Mapa censo 1 punto = 1 persona. Fuente: http://demographics.coopercenter.org/DotMap/ (2016)

El mapa sobre estas líneas muestra cada ciudadano americano como un punto, diferenciando la raza por el color del punto.

Los mapas de símbolos proporcionales o graduados, se utilizan para mostrar valores absolutos. El tamaño del símbolo es proporcional al tamaño de la variable o fenómeno que se representa.

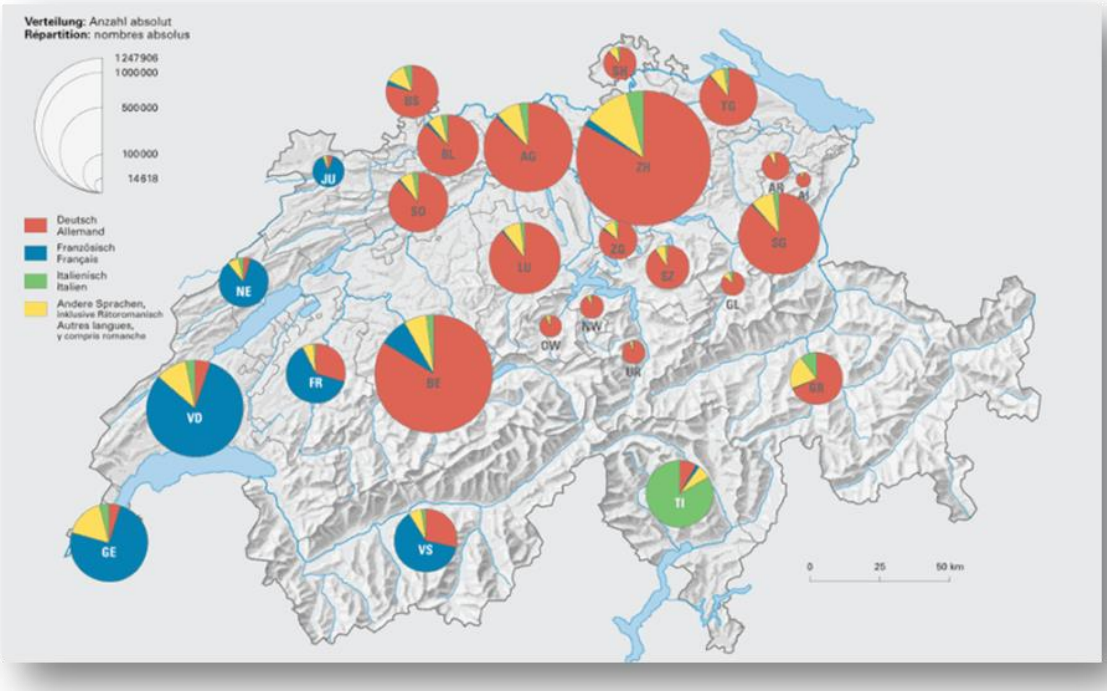


Ilustración 109. Mapa de símbolos proporcionales. Fuente: http://www.bfs.admin.ch/bfs/portal/en/index/regionen/thematische_karten/maps/bevoelkerung/sprachen_religionen.html (2016)

En este mapa se muestra el mapa de Suiza, y los gráficos en ella muestran de una parte la población de cada cantón, indicada por el tamaño del círculo, y de otra parte los idiomas que se hablan viéndose claramente la zona donde predomina el francés (azul), alemán (rojo) y el italiano (verde).

Se presenta a continuación una variante de mapa, llamado Voronoi, que muestra la partición de un mapamundi en regiones basadas en la distancia a puntos específicos de un conjunto determinado en el mapa. Este conjunto de puntos, llamados semillas, lugares o generadores, es determinado a priori y para cada semilla existe en una región del plano. Se determina todos aquellos puntos que están influidos por cada semilla.

Es una descomposición de un espació métrico en regiones, asociada a la presencia de objetos, de tal forma, que en dicha descomposición, a cada objeto se le asigna una región del espacio métrico formada por los puntos que son más cercanos a él que a ninguno de los otros objetos.

Por tanto, lo que se hace es dividir el espacio en tantas regiones como puntos u objetos tengamos de tal forma que a cada punto le asignemos la región formada por todo lo que está más cerca.

El ejemplo que se muestra a continuación toma como semillas los aeropuertos.

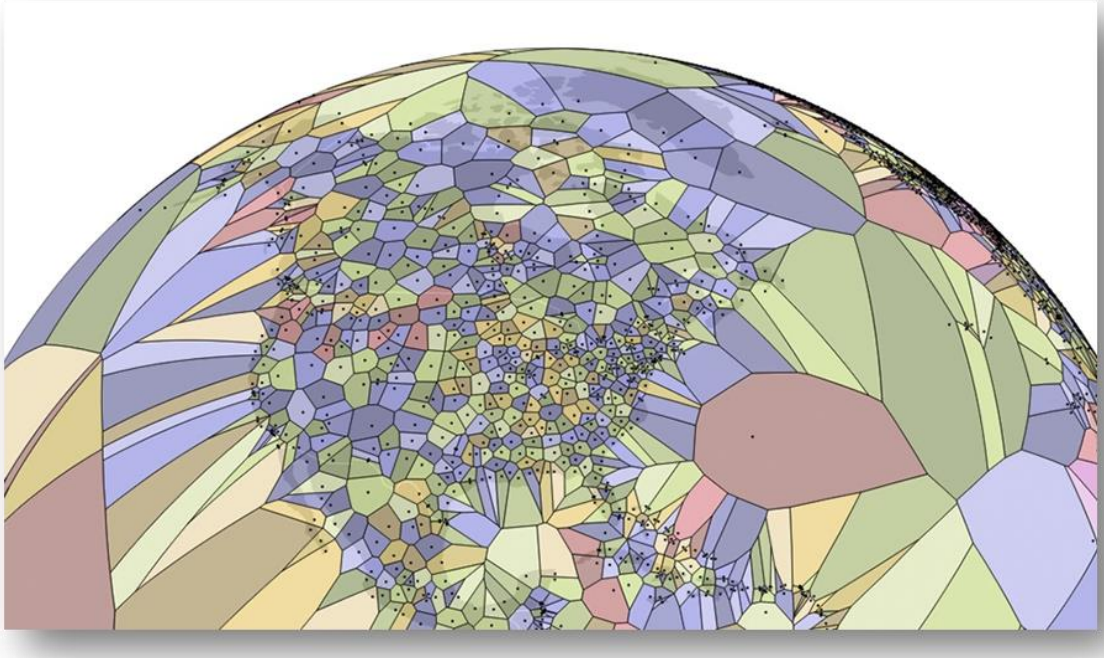


Ilustración 110. Gráfico de Voronoi.

Aplicaciones de este componente: Siempre que se quiera mostrar información geolocalizada y comparada en el territorio.

4.7 Diagrama de caja

Descripción: Un Diagrama de caja es un gráfico, basado en cuartiles, mediante el cual se visualiza un conjunto de datos. Está compuesto por un rectángulo, la "caja", y dos brazos, los "bigotes". También se le suele llamar Boxplot o Box and whiskers.

Es un gráfico que suministra información sobre los valores mínimo y máximo, los cuartiles Q1, Q2 o mediana y Q3, y sobre la existencia de valores atípicos y la simetría de la distribución. Primero es necesario encontrar la mediana para luego encontrar los 2 cuartiles restantes. El cuartil 1 corresponde al percentil 25%, el cuartil 2 al 50% y a la mediana, y el cuartil 3 al 75%.

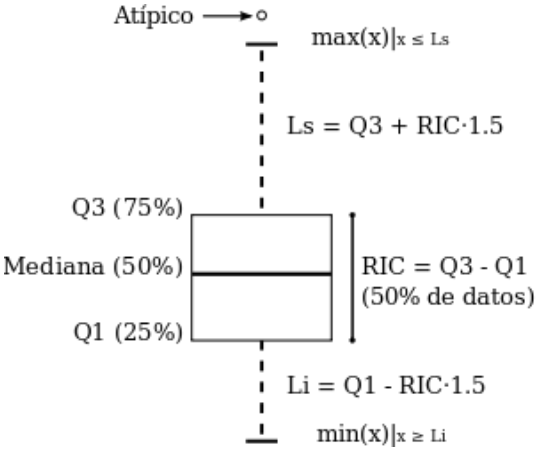


Ilustración 111. Diagrama de caja.

Tipos de datos: Cuando sea relevante mostrar la distribución estadística de datos, como por ejemplo en encuestas.

Variantes:

- Boxplot
- Gauge (Vu-meter) – Circular
- Bullet Graph (Vu-meter) – Lineal

Una variante de este gráfico con un componente enriquecido sería:

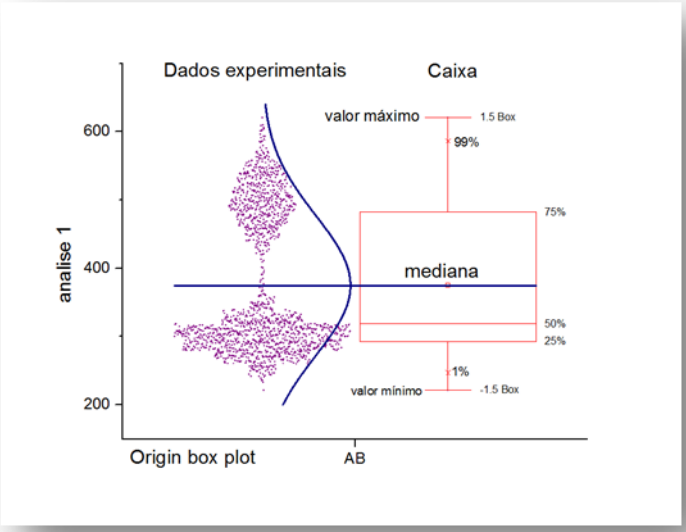


Ilustración 112. Diagrama de caja enriquecido.

Aplicaciones de este componente: Este componente permite ver de forma gráfica la distribución de los datos de una estadística. Además, es adecuado cuando quiere compararse datos o bien ver cómo ha evolucionado la distribución de un dato.

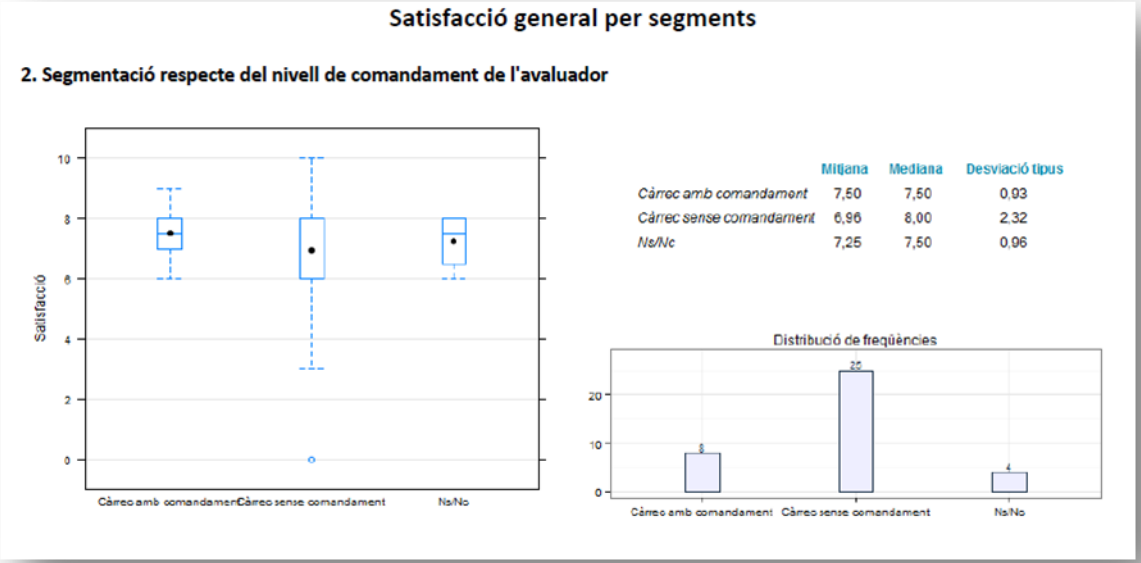


Ilustración 113. Gráfico de caja. Fuente: Diputación de Tarragona (2016)

4.8 Treemap, mosaic plot

Descripción: Consiste en un tipo de representación gráfica de datos jerárquicos en forma de rectángulos que ocupan el total del espacio de forma proporcional al valor de una variable. El resultado final es una visualización intuitiva y dinámica de un plano dividido en áreas proporcionales a los datos que representan, jerarquizado por tamaños y código de color, y optimizando el espacio.

Tipos de datos: Representación de una variable en jerarquía, una segunda variable en forma de superficie (agrupadas según su tamaño) y una tercera variable en forma de color de la superficie.

Aplicaciones de este componente: Es un componente muy interesante para mostrar jerarquías de datos. Tiene un impacto grande en el análisis de los datos, pero precisa de un conocimiento del componente y su correcta interpretación.

El Treemap a continuación, muestra el número de asistencias prestadas a los ayuntamientos de la demarcación por la Diputación de Tarragona en el año 2012. Los recuadros principales agregan las comarcas. El tamaño del rectángulo es en función del número de asistencias, y el color en función del porcentaje de incremento respecto el año anterior.

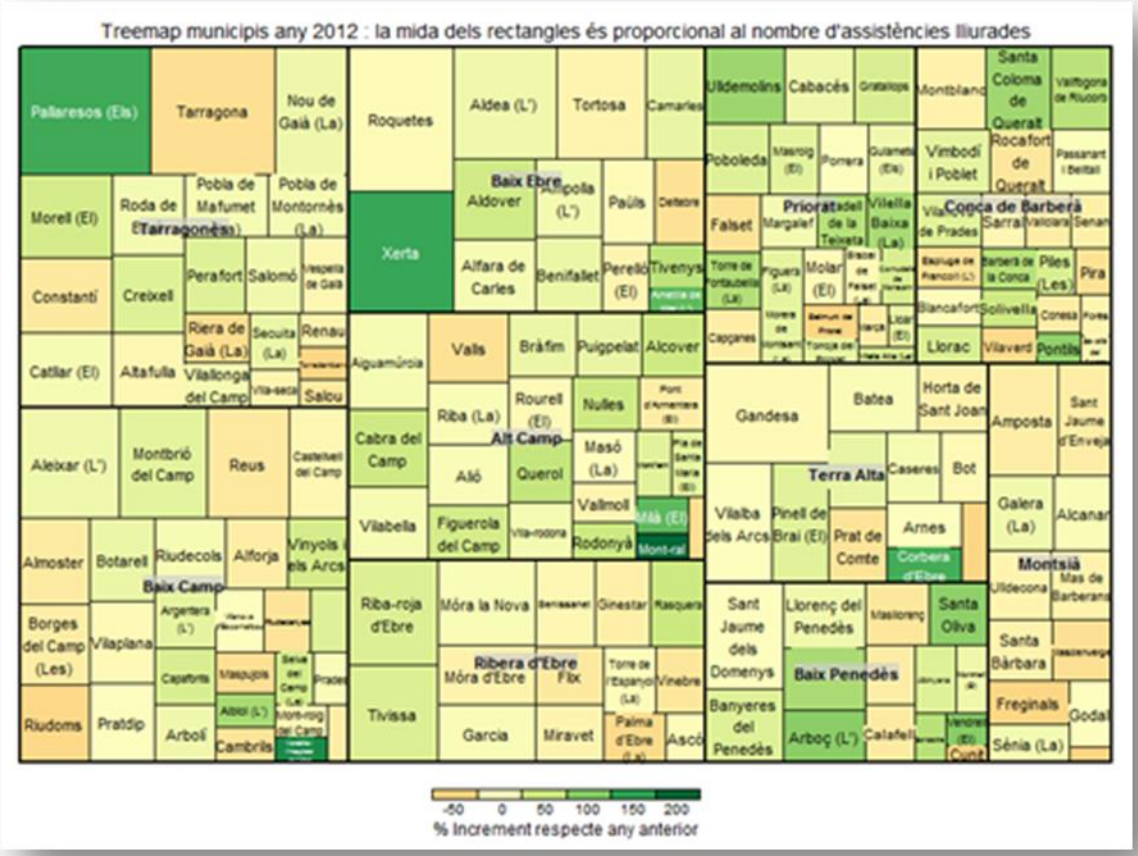


Ilustración 114. Treemap municipios 2012. Fuente: Diputació de Tarragona (2012).

4.9 Sankey

Descripción: Un diagrama que muestra los flujos entre varias variables. Generalmente muestran relaciones entre dos variables, incluyendo otras variables parciales en el recorrido de la variable inicial a la final. Puede mostrar segmentaciones de los flujos. El ancho de las bandas muestra el valor de la variable del flujo.

Tipos de datos: Después de fijar la posición horizontal de cada nodo, el algoritmo comienza a partir de las fuentes de la izquierda, posicionando los nodos corriente abajo para minimizar la distancia de enlace. Un pase inversa se realiza entonces de derecha a izquierda, y luego todo el proceso se repite varias veces. Nodos superpuestos se desplazan para evitar la colisión. El diseño totalmente automático es conveniente para un rápido posicionamiento de visualización puesto que de forma manual sería muy difícil de representar. Sin embargo, el algoritmo no es perfecto, los enlaces se dibujan con la transparencia parcial para resaltar cruces. Para facilitar la lectura y eliminar la ambigüedad, se permite cambiar la posición de los nodos de forma interactiva.

Variantes:

- Gráfico de Chord chart, en el que los flujos se muestran en una esfera.
- Gráfico de coordenadas paralelas, en el que los datos se muestran de forma paralela.

Aplicaciones de este componente: Este componente es muy útil para analizar las relaciones entre los datos y las categorías de datos. Por ejemplo, a continuación se muestra en la parte izquierda del gráfico los departamentos de la Diputación de Tarragona y en la parte derecha, las comarcas de la provincia. En el centro se muestran las familias principales del catálogo de servicios de las diputaciones: asistencias técnicas, subvenciones, etc. Puede detectarse claramente los flujos de actividad de los departamentos de la Diputación en relación con los servicios prestados y a qué comarca van dirigidos.

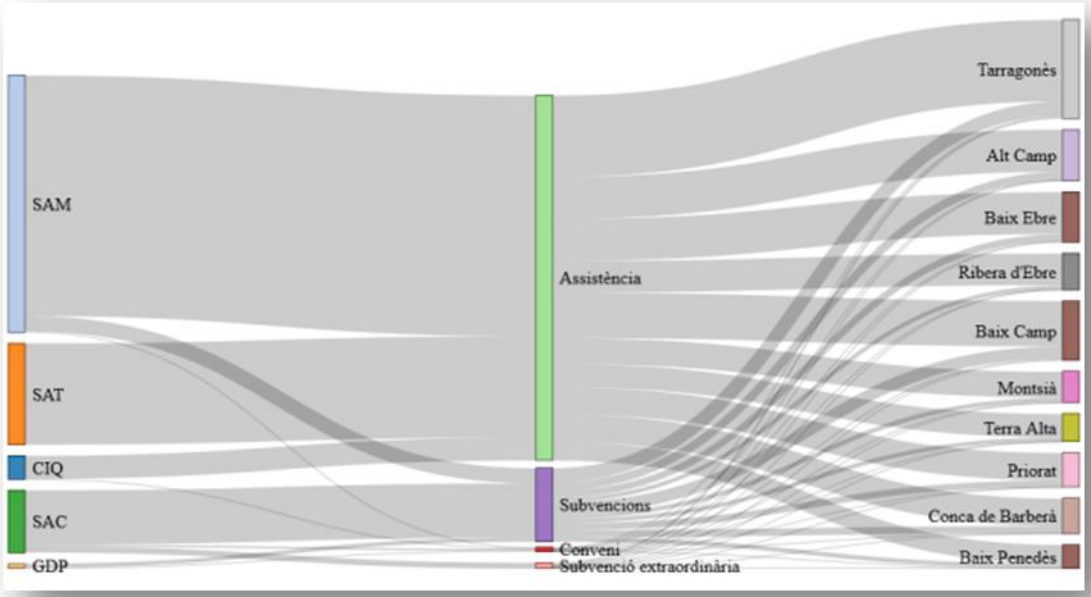


Ilustración 115. Sankey Servicios a municipios. Fuente: Diputación de Tarragona (2016)

A continuación se muestra otro ejemplo con una variante de este gráfico, en el que la información se muestra verticalmente, y que consiste en visualizar la composición por género, clase y edad de los pasajeros que sobrevivieron al hundimiento del Titanic.

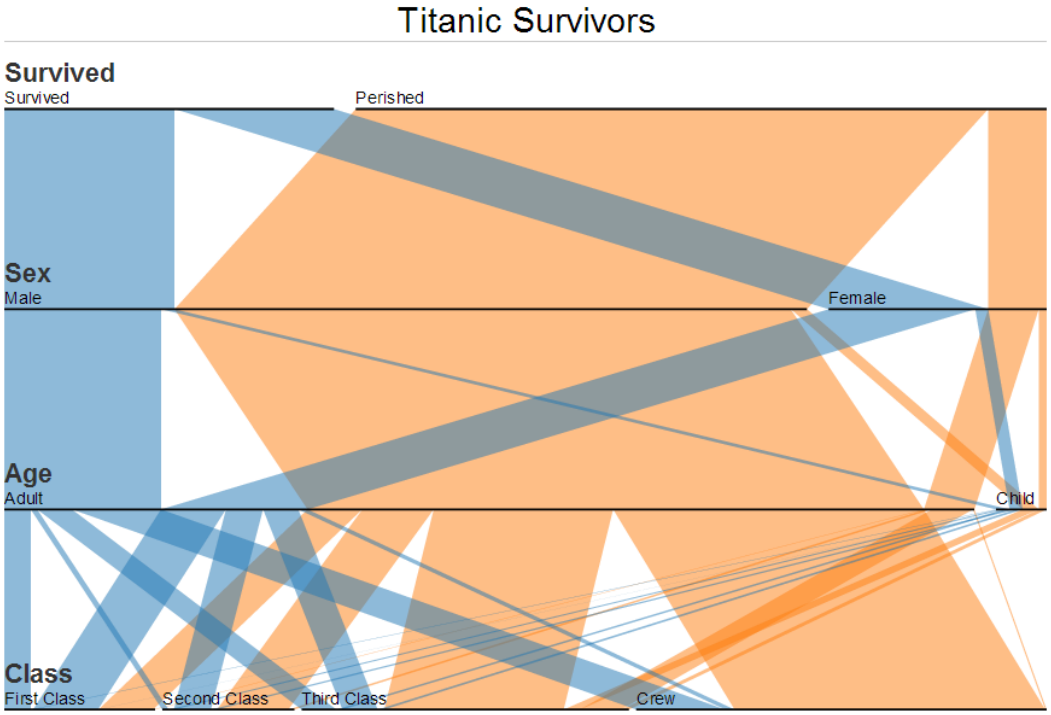


Ilustración 116. Ejemplo gráfico de paralelas. Fuente: https://www.jasondavies.com/parallel-sets/ (2016)

Una variante de este grafico sería el gráfico de coordenadas paralelas, debidas principalmente a Alfred Inselberg, entre otros, constituyen un sistema de representación relativamente reciente. Su objetivo es resolver el problema de la representación de conjuntos de datos multidimensionales. Se basan en representar cada dimensión como una escala vertical paralela a todas las demás. A cada elemento del conjunto de datos le corresponde una línea quebrada que une los valores que toman cada una de sus variables, el equivalente de un "punto" en representación bi o tridimensional.

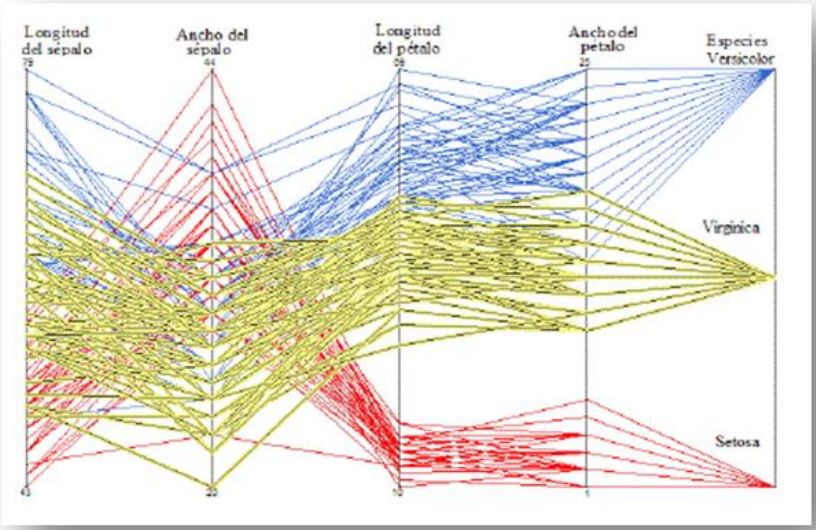


Ilustración 117. Gráfico de coordenadas paralelas.

4.10 Heat map

Descripción: Un mapa de calor (o heat map) es una representación gráfica de los datos, donde los valores individuales contenidos en una matriz se representan como colores y/o como graduaciones de color.

Tipos de datos: Fundamentalmente se usa para representar distribuciones.

Aplicaciones de este componente: Ver la distribución de una variable.

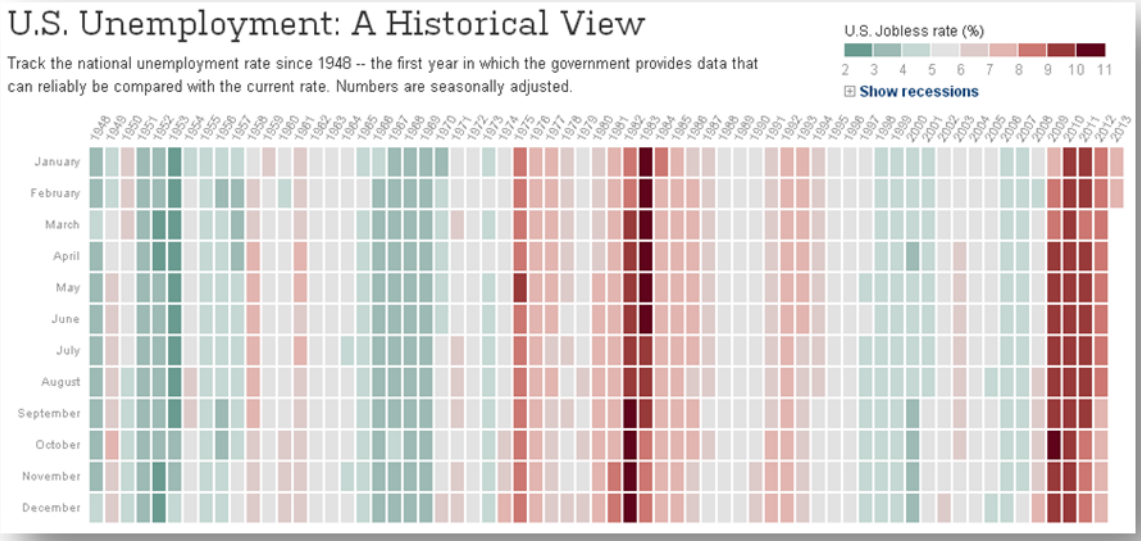


Ilustración 118. Ejemplo Heat Map. Fuente: http://www.wsj.com/ (2016)

4.11 Sparklines

Descripción: Los sparklines son pequeños gráficos del tamaño del texto generalmente que muestran información muy simplificada. Tienen la gran ventaja de mostrar mucha información en un solo componente y de ser integrables en cualquier texto o tabla para acompañar la información.

Tufte (2006) fue el primero en proponer las sparklines. Estos "directos y sencillos gráficos del tamaño de palabras", mejoran la presentación de los datos sin ocupar mucho espacio.

Los tipos de datos, las variantes y aplicaciones de este componente, son determinados por el tipo de gráfico seleccionado. Ver los apartados anteriores para cada componente.

A continuación se muestran ejemplos de uso de sparklines integrados en el texto o en una tabla.

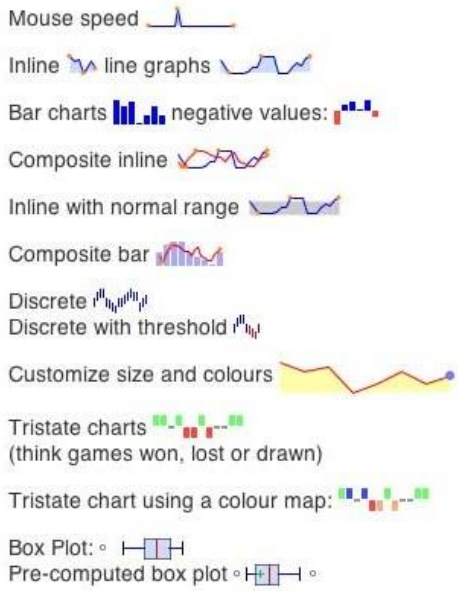


Ilustración 119. Muestra de sparklines. Fuente: Elaboración propia.

Ilustración 120. Muestra sparklines en tabla.

4.12 Nubes de palabras

Descripción: Una nube de palabras es una representación gráfica de la frecuencia de palabras o etiquetas en un texto o conjunto de datos. Existen infinidad de herramientas para preparar este tipo de componentes, que en general, muestran la lista de frecuencia de palabras y permiten eliminar proposiciones, artículos, conjunciones, etc. y dejar sólo las palabras significantes.

Tipos de datos: Las nubes de palabras o etiquetas son útiles para identificar los términos comunes de un texto y construir taxonomías de palabras clave.

Variantes: Las variantes de este componente vienen determinadas por el uso de fuentes, tamaños, formas, etc. de la nube creada y de su interactividad.

Aplicaciones de este componente: Mostrar frecuencias de palabras en textos.

Ilustración 121. Ejemplo Nube de palabras. Fuente: Elaboración propia (2016)

Dimensión Informes

1 Introducción al sistema de informes 178

2 El diseño de la información 181

2.1 Arquitectura de la información 183

3 Tipos de informes 184

3.1 Reporting 185

3.2 Dashboard y cuadro de mando 185

3.2.1 Concepto de cuadro de mando 188

3.2.2 Cuadro de mando: Tableau du Board 188

3.2.3 Cuadro de mando: Balance Score Card 188

3.3 OLAP 190

3.4 Infografía 194

4 Diseño de informes 202

4.1 Formato físico del informe 204

4.2 Componentes 205

4.3 Ejemplos de informes 206

5 El Sistema de Informes propuesto en el modelo SIEGO 209

5.1 Estructura 209

5.2 La Gestión de los Informes 213

5.2.1 Identificar el mensaje 213

5.2.2 Planificar el informe 213

5.2.3 Elaborar el informe 214

La cuadrícula 214

Tipografía 215

Colores 215

6 El proceso de forma gráfica 217

1 Introducción al sistema de informes

La dimensión informes es la que recoge los datos, los análisis y los compontes realizados y los prepara para su presentación.

La información medida y analizada debe de ser diseminada con agilidad haciéndola llegar en tiempo y espacio adecuado. De esta manera se favorece la eficiencia en la toma de decisiones.

El objetivo de este capítulo del trabajo de investigación es proponer un sistema de informes estructurado por niveles y que cada nivel tenga un número finito de Informes. En conjunto, el Sistema de Informes debería de dar respuesta al 95% de las peticiones de información en menos de 5 minutos, con la finalidad de dar apoyo fundamentado, basado en datos, al proceso de toma de decisiones.

Para justificar la necesidad y adecuación de este sistema para las organizaciones como parte del Modelo propuesto SIEGO, en el desarrollo metodológico de la tesis, se pregunta en el cuestionario de investigación acerca de cuál es el tiempo medio de respuesta a las peticiones de cualquier información del ámbito de responsabilidad del directivo. También sobre la opinión de si el hecho de disponer de un sistema de datos estructurado como el que se propone en este trabajo de investigación mejoraría o no la percepción de los resultados de la organización.

Conforme más tiempo necesita un directivo para obtener la información que le solicitan, más activa debe de ser su intuición, puesto que menos fundamentada en datos puede ser. Las decisiones de los directivos se podrían tomar con más fundamento si dispusieran de un Sistema de Informes que les permitiera obtener la información necesaria más frecuentemente, en menos tiempo. Al recibir una petición de una información nueva, que no se contempla en el Sistema de Informes, se debería buscar la información e incorporarla a uno de los Informes existentes, para desde ese momento en adelante tenerla disponible. El hecho de enriquecer el sistema de forma continua y dinámica forma parte del modelo de gestión presentado en el que se planifica, se hace, se revisa y se mejora de forma continua.

Se puede decir que un Sistema de Informes es un conjunto de instrumentos que generan, almacenan y muestran los Informes de una organización. Definiendo Informe como el elemento que recoge información de diferentes análisis realizados y los agrupa conceptual y geográficamente. También puede entenderse como el espacio, en papel o digital, con componentes gráficos y/o de texto, que destacan algunas técnicas analíticas, proporcionan una

imagen que de una manera intuitiva, casi automática para quien la recibe, comunique los datos en cuestión, resaltando aquellos aspectos de más relevancia, las comparaciones, tendencias, causalidades, etc.

El Informe en este trabajo, se entiende como una plantilla, un diseño de campos de texto, tablas y gráficos, que se guarda. Al ejecutar el informe, este formato se actualiza y rellena con los datos disponibles.

A continuación se muestra de ejemplo el diseño de un informe, en el que se pueden apreciar los nombres de los campos, la estructura de la información, etc.

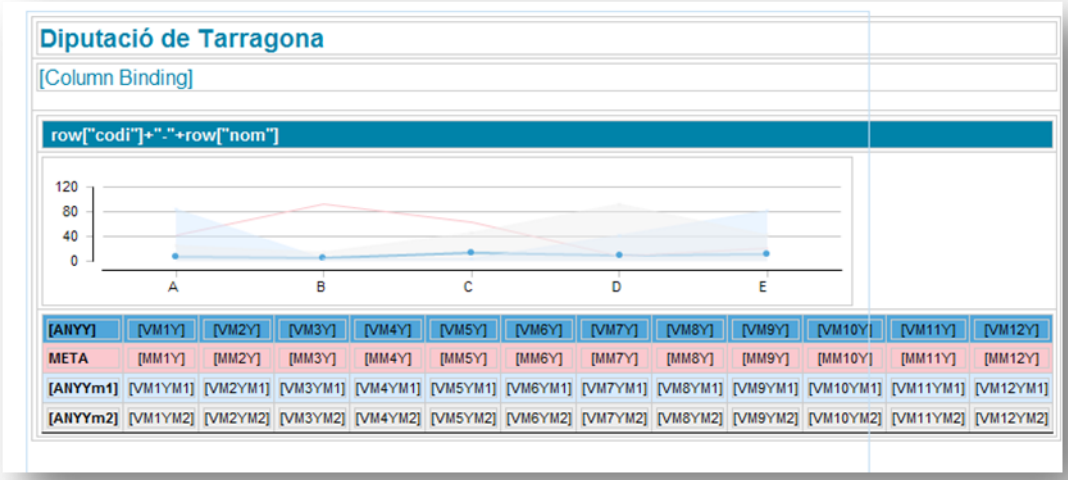


Ilustración 122. Diseño de un informe. Fuente: Diputació de Tarragona (2016)

A continuación mostramos una imagen del mismo informe una vez ejecutado y "rellenado" con los datos correspondientes:

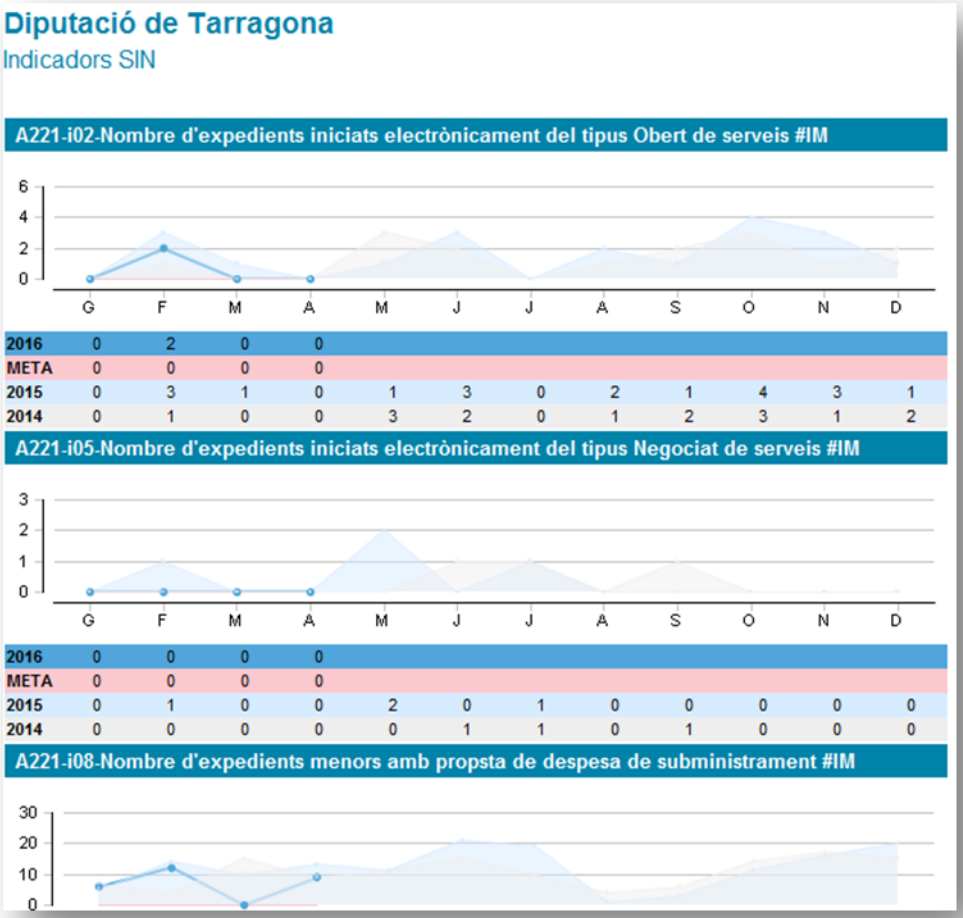


Ilustración 123. Informe ejecutado. Fuente: Diputació de Tarragona (2016)

Como se ha determinado en el capítulo anterior sobre Componentes de visualización de datos, el informe excelente es aquel que da al espectador el mayor número de ideas, en el menor tiempo posible, utilizando el mínimo de tinta y el espacio más pequeño, y que además, dice la verdad sobre los datos.

El propósito del informe es informar. Algunos informes como veremos más adelante, incluyen elementos persuasivos, como recomendaciones o sugerencias que habría que deslindar claramente. La finalidad del informe es proporcionar datos analizados agrupados y presentados de una manera determinada. El fruto del análisis puede ser una decisión ejecutiva u otro informe. Es relevante destacar que es claramente información y que es interpretación.

Los objetivos planteados para el Sistema de Informes es que en un tiempo límite se pueda obtener la información necesaria para el proceso de toma de decisiones o en cualquier proceso que cualquiera de los grupos de interés pueda requerir. El tiempo necesario para disponer y comprender la información de un informe se puede descomponer en diferentes microtiempos:

1. Tiempo transcurrido entre que surge la necesidad de información hasta la obtención de la información. Se puede subdividir a su vez en:
   a. Tiempo entre la detección de la necesidad de información hasta el acceso al Informe (t1)
   b. Tiempo entre el acceso al Informe y la obtención de la información (t2)
2. Tiempo transcurrido entre la obtención de la información y la interpretación adecuada de los datos (t3)

De este modo, el sistema de Informes debe permitir la optimización de cada uno de estos microtiempos del ciclo de peticiones de información, mediante una combinación de aplicación de tecnología y de técnicas de visualización de datos que optimizan la inversión de tiempo en la compresión de los datos, ya tratados.

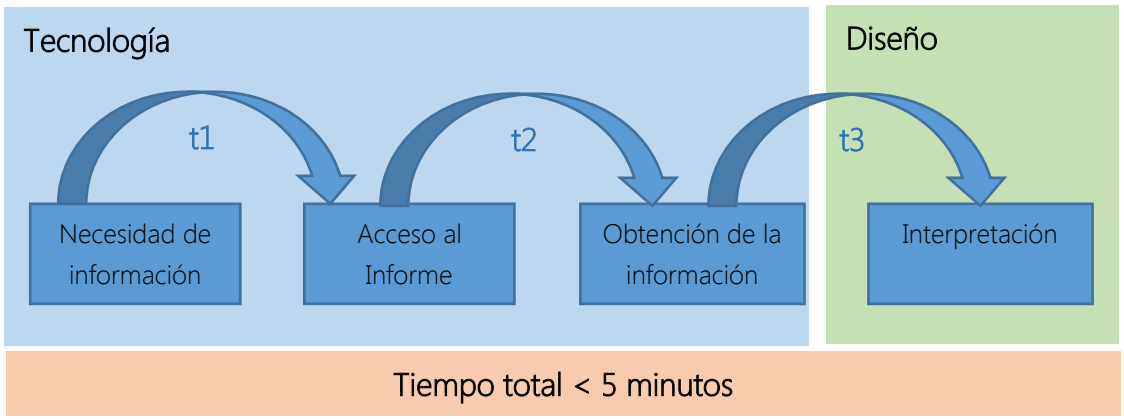


Ilustración 124. Tiempo de respuesta a una pregunta. Fuente: Elaboración propia (2016)

En el ámbito tecnológico es fundamental disponer de acceso a la información vía un servicio web, accesible desde cualquier navegador, que permita el trabajo desde cualquier tipo de plataformas: pc, tablets, smartphones, etc.

En el ámbito del diseño, además de la aplicación de componentes de visualización, es muy importante que el Sistema de Informes esté bien estructurado y organizado por niveles y tipos, de forma que el usuario sepa siempre dónde encontrar la información buscada en el menor tiempo posible.

2 El diseño de la información

Tufte, es defensor del minimalismo en la representación gráfica de datos y de la eliminación de todo tipo de atributo que entorpezca su comprensión. Propugna un estilo sobrio en que prima la información sobre la estética. Así, ha acuñado términos como "tablasura" para referirse a aquellos elementos no esenciales de una gráfica que complican su lectura. También es autor de la ratio tinta-información, que mide la cantidad de tinta que hay que utilizar para representar efectivamente un determinado conjunto de datos.

Los 8 principios de Tufte para el diseño de información son:

1. Mostrar comparaciones: Analizar significa comparar. El principal objetivo de cada presentación analítica es permitir comparaciones, y en última instancia, clasificar, evaluar y comprender.
2. Mostrar causalidad: Es la razón por la que miramos los datos, para buscar explicaciones y estructura. Queremos entender relaciones, mecanismos y efectos.
3. Mostrar complejidad. El mundo que intentamos comprender es siempre "multivariable". Cuantas más variables integra un informe mejor.
4. Integración de palabras, números e imágenes. El hecho de que el software estándar de hoy ha dividido la producción de texto, tablas, gráficos e imágenes en diferentes aplicaciones, no debe llevarnos a la tentación de separar estos medios de presentación de gran alcance.
5. El contenido es lo que cuenta: Si un gráfico influencia a alguien, se debe principalmente a la calidad, el significado y la integridad de contenido. Quien quiera mejorar un gráfico tiene que mejorar primero su contenido. Un buen diseño no puede compensar la escasez o la poca significación del contenido.
6. Toda la información a la vista: Si queremos comparar información, tenemos que ponerla junta. Es recomendable limitar el espacio a una página si el informe es en papel, o a una pantalla si el informe es digital.
7. Utilizar varios componentes gráficos. Ofrecen una presentación compacta y repetitiva de varias variables. Los principios generales que fundamentan los gráficos individuales tienen que ser claros y sencillos.
8. Los números son importantes. No reducirlos números a estados si/no o comparaciones simples.

Finalmente, a pesar de aparecer por defecto en las figuras producidas por muchos programas de ordenador, conviene prescindir de toda una serie de elementos gráficos como colores, líneas de referencia, volumen, rebordes, etc. Todo esto es lo que se conoce como 'chatarra gráfica'. La chatarra gráfica no añade absolutamente ninguna información a la gráfica, simplemente ruido, en el sentido que distrae la atención respecto a lo que es importante. El uso tan habitual de chatarra gráfica está basado en una visión errónea, de las gráficas como un elemento ornamental, y en una interpretación también errónea de que lo más ornamental es aquello más recargado. Al contrario, la función de las gráficas es informar y ello se consigue mejor con diseños simples, que maximizan la proporción entre información y tinta. La chatarra gráfica, obviamente, solo añade tinta, no información.

Otro referente relevante en lo que se refiere a la preparación de información es Williams Robin (1994) que plantea cuatro principios (CRAP) para el diseño de información:

Principio 1: Contraste. Da a un informe interpretabilidad rápidamente.

Se crea contraste cuando dos elementos son realmente distintos. Se debe enfatizar el elemento más interesante e importante del informe que esté diseñando de una forma clara y evidente.

Principio 2: Repetición. Da familiaridad al informe.

Se debe Repetir algún aspecto del diseño a lo largo de toda la pieza. Puede ser una fuente en negrita, un formato particular, relaciones espaciales, un color, un encabezado, etc.

Principio 3: Alineación. Da navegabilidad al informe.

Nada debe colocarse en la página de forma arbitraria, cada elemento debe tener una conexión visual con algo más de la página.

Principio 4: Proximidad. Da estructura al informe.

Se deben agrupar los elementos relacionados para que se vean como un grupo cohesionado, en lugar de informaciones sin relación.

2.1 Arquitectura de la información

El concepto de arquitectura de la información fue utilizado por primera vez por Richard Saul Wurman (1975), quien lo define como "El estudio de la organización de la información con el objetivo de permitir al usuario encontrar su vía de navegación hacia el conocimiento y la comprensión de la información".

No fue hasta la segunda mitad de la década de los 90 (Ronda, 2004: 3) que el concepto de arquitectura de la información se popularizó. Otras definiciones nos las dan:

"La disciplina que se ocupa del diseño estructural de los sistemas de información, su problema central es la organización, recuperación y presentación de información mediante el diseño de ambientes intuitivos"(Baeza, Rivera y Velasco, 2004: 169).

"Disciplina (arte y ciencia) encargada de estructurar, organizar y etiquetar los elementos que conforman los entornos informacionales para facilitar de esta manera la localización (o el acceso) de la información contenida en ellos y mejorar, así, su utilidad y su aprovechamiento por parte de los usuarios "(Pérez-Montoro, 2010).

Este enfoque es especialmente relevante cuando diseñamos informes interactivos.

Como decía David Robins, la Arquitectura de la Información es un término difícil de definir, y para cada persona puede tener una acepción particular. Este término es usado para describir la planificación de la experiencia del usuario frente a la información que se presenta en una pantalla. Algunos de estos elementos son más tangibles, como el impacto del diseño gráfico. Estos elementos combinados permiten organizar la información de forma que el usuario pueda obtener adecuadamente y en forma rápida los datos de una pantalla, para el caso de los informes digitales.

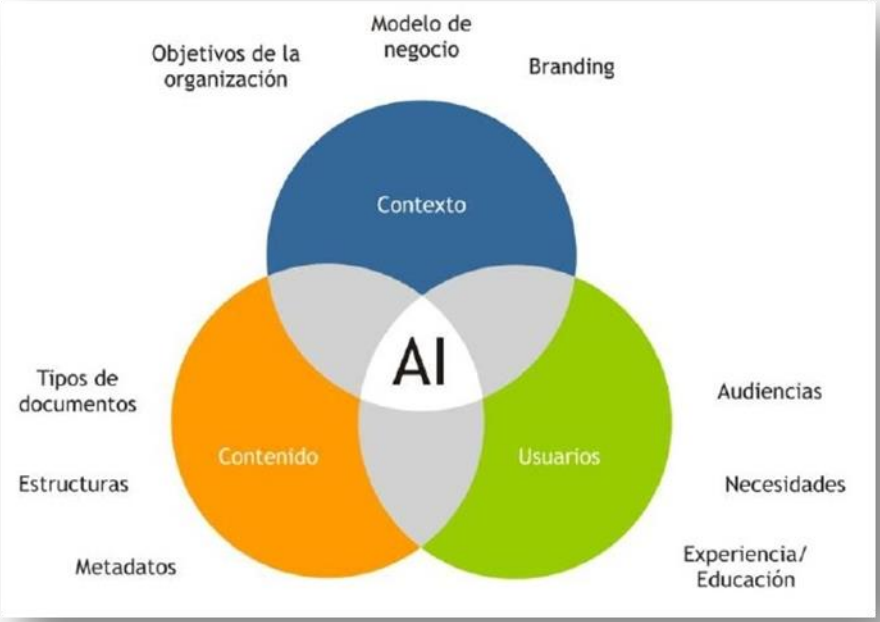


Ilustración 125. Arquitectura de la información.

Se define entonces la arquitectura de la información como el espacio en el que confluye la experiencia del usuario (que tipo de usuario es, que necesidades tiene, que experiencia y conocimiento), con el contexto (los objetivos de la organización, el modelo de negocio) y con el contenido que quiere comunicarse (estructura, metadatos, documentos).

3 Tipos de informes

Los informes pueden ser muy simples y estáticos en su naturaleza, como una lista de ventas por un determinado periodo de tiempo, o un informe más avanzado con agrupaciones, resúmenes, dinámica de exploración de datos interactiva o vinculación. Los informes son más utilizados cuando el usuario necesita ver los datos "en bruto" en un formato más fácil de leer. Cuando se combinan datos, tablas y gráficos, los informes ofrecen un medio ideal para permitir a los usuarios el análisis de los datos concretos destacando las medidas y los indicadores clave de rendimiento (KPI's).

Existen infinidad de tipos y clasificaciones de informes, además últimamente la tecnología y las herramientas permiten hacer múltiples combinaciones entre ellos, por lo cual en este trabajo se ha escogido la siguiente clasificación que será desarrollada posteriormente:

Informes estáticos

- Informes de Reporting, que suelen reproducirse en papel o bien en un fichero digital en Word, Excel, o pdf. En su estado básico no son interactivos.

Informes dinámicos

- Dashboards
- OLAP (incluido HOLAP,ROLAP y MOLAP)

Estos informes se visualizan en pantalla y son interactivos.

- Infografía (puede ser dinámica o estática)

3.1 Reporting

El reporting (o la elaboración de informes y reportes) es una actividad fundamental en la gestión de las organizaciones. Previene, entre otros riesgos, de los asociados a la espontaneidad y la intuición en la toma de decisiones, aportando conocimiento real y actualizado del estado y el desempeño de las actividades, la estrategia, las operaciones y los procesos clave de la organización.

Existen muchos tipos de reportes en función de los fines concretos para los cuales se realizan, y no cabe duda de que, cuanto mayor sea la flexibilidad del sistema de datos e información que se adopte, mayores serán las posibilidades de sacar el máximo partido a los sistemas de información implementados en la organización.

Sin embargo, y pese a la variedad de reportes posibles, podemos clasificar grandes grupos según la tipología de los fines que persiguen:

- Reportes históricos: se basan en el análisis de los datos y las informaciones históricas relevantes. Sirven para saber qué sucedió en áreas, departamentos o con el desempeño de determinadas operaciones, identificando patrones e irregularidades y visibilizando las consecuencias de las decisiones tomadas en el pasado.
- Reportes operativos: remiten al presente de la organización, al estado actual de las operaciones que se llevan a cabo y permiten saber qué está sucediendo, en tiempo real, en la organización.
- Reportes analíticos: Nos permiten descubrir las relaciones de causalidad en la información mostrada en los reportes anteriores, traduciendo esta información en conocimiento relevante para guiar la toma de decisiones.
- Reportes predictivos: fundamentales para el aprovechamiento de las nuevas oportunidades. Prevén las consecuencias que acarreará la toma de decisiones en un sentido u otro mediante el diseño de escenarios posibles, mostrando qué podría suceder en el futuro y el modo más adecuado para evitar riesgos innecesarios.

3.2 Dashboard y cuadro de mando

Un dashboard es un concepto integrado de información utilizado fundamentalmente en el ámbito del Business Intelligence. El nombre dashboard proviene del inglés, que lo utiliza para el

tablero de mandos de un coche, que muestra al conductor la información esencial del funcionamiento del vehículo. De forma análoga, un dashboard muestra los indicadores más relevantes de la organización en el mundo empresarial.

Así, un dashboard es una página desarrollada en base a tecnología web mediante la cual se despliega en tiempo real información de la empresa extraída de diferentes fuentes o bases de datos. Su característica de tiempo real otorga a los usuarios un conocimiento completo sobre la marcha de la organización y permite hacer el análisis instantáneo y de inteligencia de negocio.

Un dashboard es un informe que está diseñado para ser presentado en una pantalla de ordenador, tableta o Smartphone y que, la característica principal es poder navegar y profundizar en los datos en tiempo real. De esta manera, con un solo dashboard, se puede acceder a infinidad de informes.

Ilustración 126. Ejemplo de dashboard. Fuente: Diputació de Tarragona (2016)

Un dashboard suele tener diferentes áreas sensibles donde, cada área, tiene un elemento (grafica de líneas, de sectores, mapa geográfico, listado, etc.). Cuando se hace el clic sobre una parte de un área, se activa otra área presentado información asociada. Este despliegue puede tener varios niveles. Un dashboard está enfocado a mostrar información sobre el cumplimiento de una organización y muestra la información por niveles.

Nivel superior: Monitorear y mostrar excepciones. El nivel superior de un dashboard es un informe visual de excepciones. Permite al usuario visualizar de forma fácil y clara las métricas fundamentales (algunas veces no sólo métricas, sino informes, documentos, textos de análisis, resultados de un proceso de minería de datos, etc.). Normalmente las métricas están representadas por componentes de visualización (semáforos, vu-meters, termómetros, etc.), gráficos o tablas según sea las necesidades del usuario. Las métricas dan a los usuarios una rápida perspectiva del cumplimiento de los procesos, objetivos, metas o personas que los gestionan. Cuando el dashboard está diseñado como un portal de conocimiento, este nivel superior es posible combinarlo con algunos de los indicadores KPI del modelo Balanced Score Card (BSC) implementado en la organización.

Nivel medio: Análisis y exploración. En este nivel el usuario trata de responder las preguntas del porqué de las excepciones que se presentaron en una o diversas de las métricas que se mostraron en el nivel superior. En este nivel el usuario tendrá la posibilidad de hacer un análisis más avanzado, arrastrando dimensiones y atributos hasta llegar a responder sus preguntas de las excepciones. Normalmente estas funcionalidades avanzadas se ponen en práctica con cubos OLAP. Es importante vincular las medidas del nivel superior con la vista del cubo multidimensional, de tal forma que se entregue de forma ágil al usuario que lo pueda necesitar.

Nivel inferior: Informes de reporting y datos operacionales. Presenta informes planos con la información detallada (un listado de facturas, un listado de órdenes de compra, una operación bancaria, etc.). A veces no se llega a este nivel de detalle, puesto que los usuarios de la parte estratégica y táctica tienen toda la información que para ellos es relevante en el nivel medio.

Los dashboards son parte de las características de la nueva generación de herramientas de Business Intelligence descritas por Eckerson (2005) y que a continuación se relacionan:

- Basadas en web. Proveen facilidad de acceso, simplifica la navegación al usuario y centraliza la administración de actualizaciones y permisos de acceso.
- Apariencia de portal. Se convierten en el lugar único de entrega de información y uno de los más relevantes para la generación de conocimiento.
- Visualización tipo dashboard. Va del ámbito micro al macro, inicialmente muestra unas pocas métricas y dependiendo de las excepciones, va mostrando el camino para encontrar las respuestas.
- Integración en los procesos de negocio que se haga el seguimiento de su cumplimiento.
- Parametrización de la información por niveles de tal manera que la parte estratégica, táctica y operativa sea cada nivel del sistema.

Algunos informes de tipo reporting se pueden convertir en dashboard cuando se añaden funciones de interactividad.

3.2.1 Concepto de cuadro de mando

Se considera un cuadro de mando a un tipo de informe dashboard particular. Se puede definir como:

"Aquel soporte de información periódica para la dirección de la organización capaz de facilitar la toma de decisiones oportuna, conociendo el nivel de cumplimiento de los objetivos definidos previamente a través de los indicadores de control y de otras informaciones que lo soportan." (AECA, 1999:42)

Hay diferentes sistemas de cuadros de mando y cada uno tiene su propio grupo de partidarios. Por ejemplo, el Balance Score Card de Kaplan y Norton (1997), el Performance Prism de Neely (2001), y el Cambridge Performance Process, que están diseñados para ser implantados en un amplio tipo de organizaciones; y las metodologías de TPM process, Jones y Chelin (2000), 7-sept TPM process de Zigon (1999) y Total Measurement Development Method (TMDM) de Tarkenton Productivity Group (2000) son para estructuras específicas más basadas en equipos de trabajo.

A continuación se presentan dos de las metodologías de cuadro de mando a modo de ejemplo:

3.2.2 Cuadro de mando: Tableau du Board

Es el primer cuadro de mando documentado en la literatura. Según Burguignon et al. (2004), los orígenes del Tableau du Board se pueden remontar a 1932. Fue desarrollado por los ingenieros de proceso para mejorar los procesos internos y controlar el funcionamiento, Epstein y Manazoni (1998). Des de sus orígenes, el Tableau du Board ha ido experimentando cambios en su arquitectura y metodología. Al principio, fue utilizado principalmente como herramienta de divulgación y de control, para dar a los gestores la información financiera y no financiera relevante sobre las actividades de la organización. En los años 90, el Tableau du Board se alineó cada vez más con la estrategia. El Tableau du Board se basa en la premisa de la existencia de relaciones causales entre la acción estratégica y los objetivos estratégicos fijados en el Plan Estratégico correspondiente.

3.2.3 Cuadro de mando: Balance Score Card

Aunque no hay ningún cuadro de mando que se haya aceptado universalmente, el Balance Score Card es el cuadro de mando más ampliamente utilizado.

Fue desarrollado por los académicos Kaplan y Norton después de diferentes investigaciones en diversas empresas americanas y fue presentado en el año 1992 Kaplan y Norton (1992). Una serie de publicaciones subsiguientes perfilaron el concepto de BSC y lo colocaron en el centro de la formulación de la estrategia. Kaplan y Norton (1996a, 2001b, 2004).

Según Kaplan y Norton (2001b, p.65) "la estrategia no puede ser ejecutada si no puede ser entendida y no puede ser entendida si no se ha descrito". Si una organización puede describir de manera comprensiva su estrategia y puede comunicarla con eficacia a través de todos los niveles de la organización, la probabilidad de poder ejecutarla con éxito aumentará considerablemente. Así pues, según Kaplan y Norton (2001b) una vez que una empresa defina su visión estratégica deberá de traducirla a un mapa estratégico. Un mapa estratégico visualiza, comunica la estrategia, vía hipotéticos acoplamientos de causa efecto entre las actividades de valor añadido y los resultados estratégicos deseados.

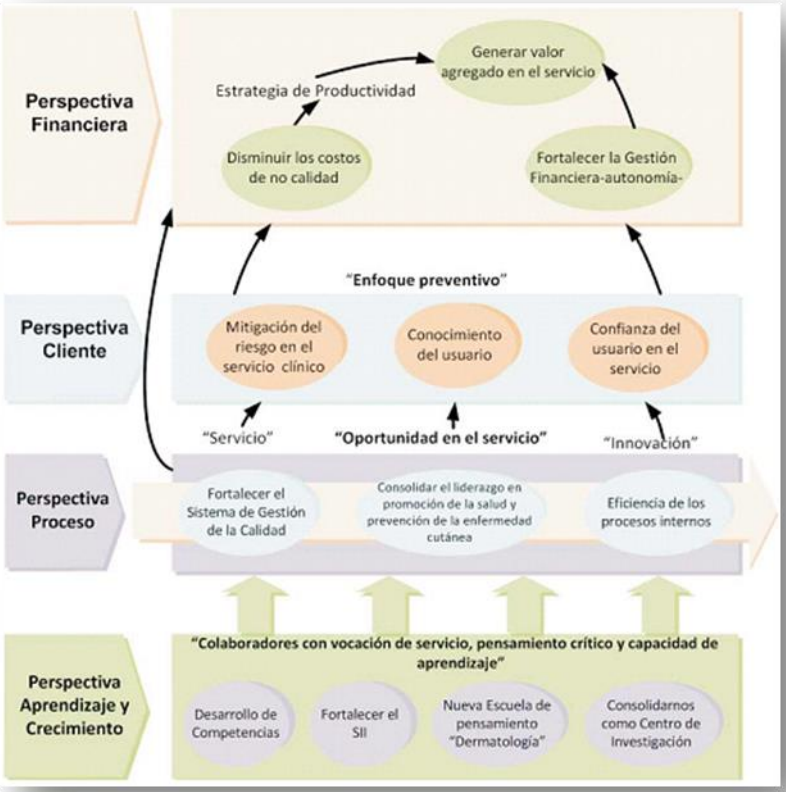


Ilustración 127. Ejemplo de mapa estratégico BSC. Fuente: Internet (2016)

Las investigaciones recientes sobre la aplicación práctica del Balance Score Card han revelado que, aproximadamente, la mitad de las empresas más grandes de Estados Unidos y de Europa occidental utilizan, en cierta medida, el Balance Score Scard, Rusel (2003).

Al Balance Score Card de Kaplan y Norton se le suele llamar Cuadro de Mando Integral, e incluye la información sobre las cuatro perspectivas que se muestran en la Ilustración 127. Ejemplo de mapa estratégico BSC. Fuente: Internet (2016).

Un cuadro de mando, del tipo que sea, se puede monitorear con cualquier tipo de Informe, bien se mediante un reporting, bien con un dashboard o bien con un OLAP. A veces se confunde el informe con la metodología y se le llama cuadro de mando integral al dashboard

que muestra los indicadores de las cuatro perspectivas del modelo de Kaplan & Norton, como el que se muestra a continuación.

Ilustración 128. Ejemplo de Dashboard de CMI. Fuente: http://www.cmigestion.es/2014 (2016)

3.3 OLAP

Las herramientas OLAP nos proporcionan análisis interactivo por las diferentes dimensiones de los datos (por ejemplo, tiempo, producto, cliente, criterios geográficos, etc.) y por los diferentes niveles de detalle (para la dimensión tiempo, habrá nivel de detalle año, trimestre, mes, día).

Ello significa pasar de la visión estática de los datos a una visión dinámica, donde se puede navegar por los datos, bajando en el nivel de detalle, cambiando la dimensión por la cual analizamos la información. El típico ejemplo sería una tabla con los datos de ventas y márgenes por delegación de una empresa, y cuando observamos un indicador de rentabilidad negativa, buceamos en los datos de esta delegación hasta dar con el producto que se está vendiendo a precios con margen negativo. Este sería el típico ejemplo de los Cubos OLAP y los visores multidimensionales que nos permiten "profundizar en los datos". Nos permitirían contestar a la pregunta: ¿qué sucedió y por qué?

On-Line Analytical Processing (OLAP) es una tipo de informe basado en una categoría de software que permite a los analistas, controllers y ejecutivos obtener información útil de los datos a través de un acceso rápido, consistente e interactivo a una amplia variedad de posibles vistas de datos.

Los informes OLAP, facilitan el análisis intensivo y permiten a los directores generar sus análisis interactivamente según navegan por sus datos. Las herramientas OLAP (On-Line Analytical Processing - Procesamiento Analítico On-Line) son descendientes de paquetes de generación de búsquedas que, a su vez, descienden de programas de generación de informes por lotes. Al igual que sus antecesoras, estas herramientas están diseñadas para responder a búsquedas "de arriba a abajo" de los datos o planificar escenarios para los analistas. Recientemente, las herramientas OLAP se han hecho muy populares como métodos primarios de acceso a bases de datos, Data Marts, y almacenes de datos (Data Warehouse).

Las herramientas OLAP fueron diseñadas para permitir a los analistas de datos informes personalizados y en forma de "construcción de cubos". La estructura de datos OLAP es similar a un Cubo de Rubik de datos que un analista puede girar y visualizar diferentes formas de trabajar a través de informes multidimensionales y escenarios "what-would-happen"" (qué ocurriría si). Las herramientas OLAP proporcionan fundamentalmente un análisis multidimensional de los datos, es decir, permiten resumir y segmentar los datos.

Ilustración 129. Ejemplo de dimensiones en un cubo OLAP. Fuente: https://technet.microsoft.com/es-es/library/hh916536(v=sc.12).aspx (2016)

Existen distintos sistemas OLAP que se diferencian fundamentalmente entre sí por el tipo de base de datos sobre los que se construyen y que dan lugar, entre otras y principalmente, a las siguientes categorías o sistemas:

- Sistemas ROLAP: construidos sobre bases de datos relacionales que emplean, principalmente, esquemas de copo de nieve o de estrella. Es un sistema apto para analizar grandes volúmenes de datos.
- Sistemas MOLAP: las bases de datos sobre las que trabaja el motor OLAP son de tipo multidimensional. Presenta algunas ventajas frente al modelo anterior, como la mejora de la rapidez en el almacenamiento de datos, la optimización del rendimiento de la memoria caché, o la eficiencia en la extracción de datos (debido a la necesidad de que estos estén previamente tratados o estructurados); sin embargo, también presenta algunas desventajas, entre otras el riesgo de duplicidad en los datos analizados (especialmente cuando se trabaja con cierto número de dimensiones), o la complejidad del proceso de carga de los datos en las bases debido a lo que los datos deben ser tratados o estructurados en el momento de la carga.

- Sistemas HOLAP: combinan los sistemas ROLAP (relacionales) con los MOLAP (multidimensionales). Se recomienda almacenar los datos más recientes en MOLAP para mejorar la rapidez de los análisis, y los más antiguos o de menor uso en ROLAP, dada la facilidad que presenta para su almacenamiento.
- Sistemas Dynamic OLAP (DOLAP): generan una pequeña caché (memoria) multidimensional cuando los usuarios ejecutan las consultas contra la base de datos.

Los cubos OLAP son bases de datos multidimensionales compuestas de medidas y dimensiones, y creadas a partir de los esquemas de las tablas empleadas en las bases de datos relacionales (las tablas, en concreto sus registros, proporcionan las medidas de los cubos, y sus dimensiones vienen determinadas por las dimensiones de los cuadros con los que cuenta cada tabla).

Se debe tener muy en cuenta que los cubos no sustituyen a las tablas relacionales; simplemente ayudan a mejorar la eficiencia, la velocidad de respuesta y la simplicidad de las consultas, pero deben apoyarse siempre en un modelo relacional con los datos previamente normalizados. (Ver Dimensión Datos).

Para entender que se analiza mediante los cubos OLAP, debemos saber que la información de gestión se compone de conceptos de información (dimensiones) y coeficientes de gestión (indicadores), que los responsables de la organización pueden consultar según sus necesidades. Dichas dimensiones de negocio se estructuran a su vez en distintos niveles de detalle (por ejemplo, la dimensión geográfica puede constar de los niveles nacional, provincial, ayuntamientos, etc.).

En general, los informes o cubos OLAP deben:

- Soportar requerimientos complejos de análisis.
- Analizar datos desde diferentes perspectivas.
- Soportar análisis complejos contra un gran volumen de datos.

La funcionalidad de los sistemas OLAP se caracteriza por ser un análisis multidimensional de datos corporativos, que soportan los análisis del usuario y unas posibilidades de navegación, seleccionando la información a obtener. Normalmente este tipo de selecciones se ve reflejada en la visualización de la estructura multidimensional, en unos campos de selección que nos permitan elegir el nivel de agregación (jerarquía) de la dimensión, y/o la elección de un dato en concreto, la visualización de los atributos del sujeto, frente a una(s) dimensiones en modo tabla, pudiendo con ello realizar, entre otras las siguientes acciones:

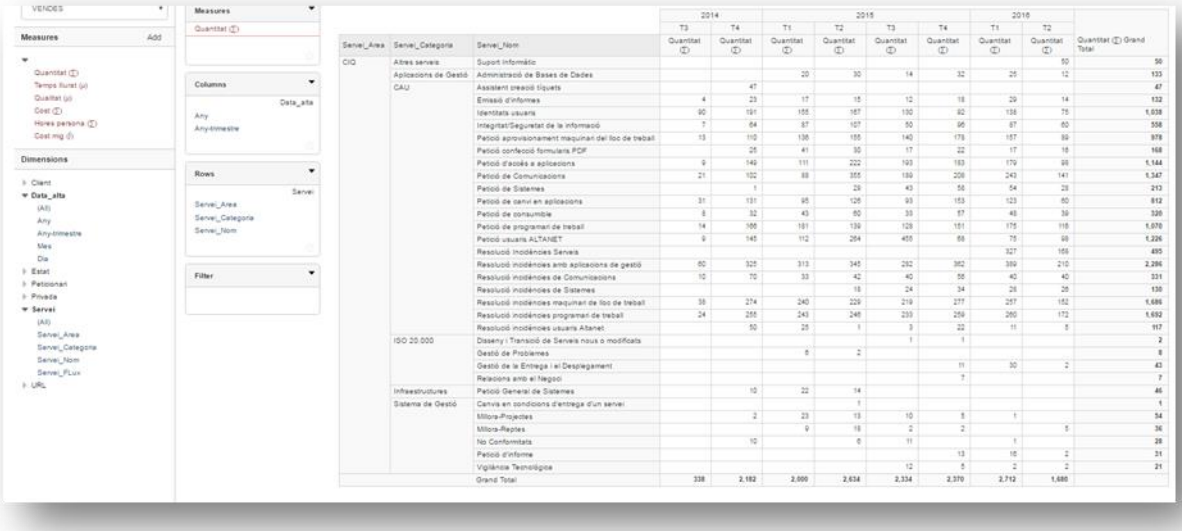
- Rotar (*Swap*): alterar las filas por columnas (permutar dos dimensiones de análisis)
- Bajar (*Down*): bajar el nivel de visualización en las filas a una jerarquía inferior.
- Detallar (*Drilldown*): informar para una fila en concreto, de datos a un nivel inferior.

- Expandir (*Expand*): id. anterior sin perder la información a nivel superior para éste y el resto de los valores.
- Colapsar (*Collapse*): operación inversa a la anterior.

La multidimensionalidad es una característica de los informes OLAP que convierte los datos bidimensionales (orientados en columnas y filas) en un cubo multidimensional. Las caras del cubo representan las dimensiones.

Una dimensión es un atributo estructural de un cubo identificado con una etiqueta de metadata. La etiqueta refleja la percepción de los datos del usuario, es decir el nombre con el que el usuario reconoce los datos.

En un cubo multidimensional se pueden examinar los datos por varias dimensiones simultáneamente (ventas por mes, ventas por producto, ventas para todos los mercados…). Por ejemplo, todos los meses, trimestres y años componen la dimensión de tiempo. Todas las ciudades, regiones, países, etc. componen la dimensión geográfica. Las dimensiones ofrecen una forma concisa e intuitiva de organizar y seleccionar datos para la exploración y el análisis.



Servei_Area	Servei_Categoria	Servei_Nom	2014 T3 Quantitat	2014 T4 Quantitat	2015 T1 Quantitat	2015 T2 Quantitat	2015 T3 Quantitat	2015 T4 Quantitat	2016 T1 Quantitat	2016 T2 Quantitat	Quantitat Grand Total
CIQ	Altres serveis	Suport Informàtic								50	50
	Aplicacions de Gestió	Administració de Bases de Dades			20	30	14	32	25	12	133
	CAU	Assistent creació tiquets		47							47
		Emissió d'informes	4	23	17	15	12	18	29	14	132
		Identitats usuaris	90	191	155	167	130	92	138	75	1,038
		Integritat/Seguretat de la informació	7	64	87	107	50	95	87	60	558
		Petició aprovisionament maquinari del lloc de treball	13	110	136	155	140	179	167	89	979
		Petició confecció formularis PDF		25	41	30	17	22	17	16	168
		Petició d'accés a aplicacions	9	149	111	222	193	183	179	98	1,144
		Petició de Comunicacions	21	102	88	355	189	206	243	141	1,347
		Petició de Sistemes		1		29	43	58	54	28	213
		Petició de canvi en aplicacions	31	131	95	126	93	153	123	60	812
		Petició de consumible	8	32	43	60	33	57	48	39	320
		Petició de programari de treball	14	166	181	139	129	151	175	116	1,070
		Petició usuaris ALTANET	9	145	112	264	455	68	75	98	1,226
		Resolució incidències Serveis							327	168	495
		Resolució incidències amb aplicacions de gestió	60	325	313	345	292	362	389	210	2,296
		Resolució incidències de Comunicacions	10	70	33	42	40	56	40	40	331
		Resolució incidències de Sistemes				18	24	34	26	28	130
		Resolució incidències maquinari de lloc de treball	38	274	240	229	219	277	257	152	1,686
		Resolució incidències programari de treball	24	255	243	246	233	259	260	172	1,692
		Resolució incidències usuaris Altanet		50	25	1	3	22	11	5	117
	ISO 20.000	Disseny i Transició de Serveis nous o modificats					1	1			2
		Gestió de Problemes			5	2					8
		Gestió de la Entrega i el Desplegament						11	30	2	43
		Relacions amb el Negoci						7			7
	Infraestructures	Petició General de Sistemes		10	22	14					46
	Sistema de Gestió	Canvis en condicions d'entrega d'un servei				1					1
		Millora-Projectes		2	23	13	10	5	1		54
		Millora-Repàs			9	18	2	2		5	36
		No Conformitats		10		6	11		1		28
		Petició d'informe						13	16	2	31
		Vigilància Tecnològica					12	5	2	2	21
		Grand Total	338	2,182	2,000	2,634	2,334	2,370	2,712	1,680	

Ilustración 130. Vista de un cubo OLAP de la Diputación de Tarragona. Servicios por año y trimestre. Fuente: Diputación de Tarragona (2016).

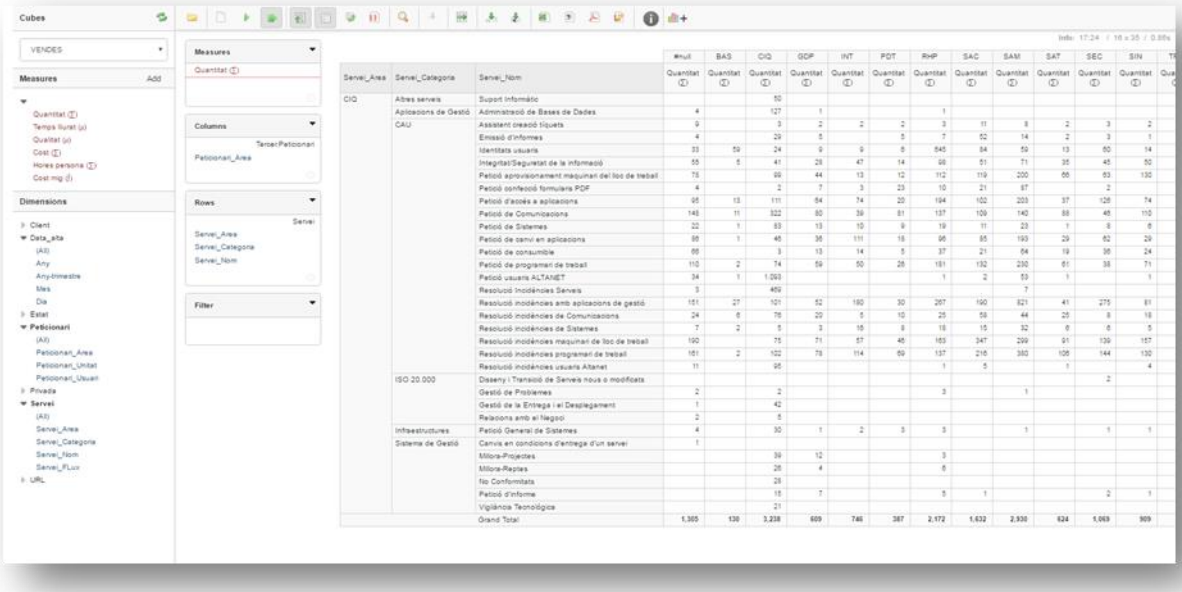


Ilustración 131. . Vista del mismo cubo OLAP de la Diputación de Tarragona. Servicios por Área a la que se le prestaron. Fuente: Diputación de Tarragona (2016).

En general, en todo tipo de informes OLAP, para cada dimensión se puede escoger de forma interactiva los campos que se quieren para las filas y para las columnas, y escoger qué métrica se quiere mostrar. Se pueden crear campos calculados para las métricas y navegar en la profundidad de los campos escogidos, con la funcionalidad de drill-down.

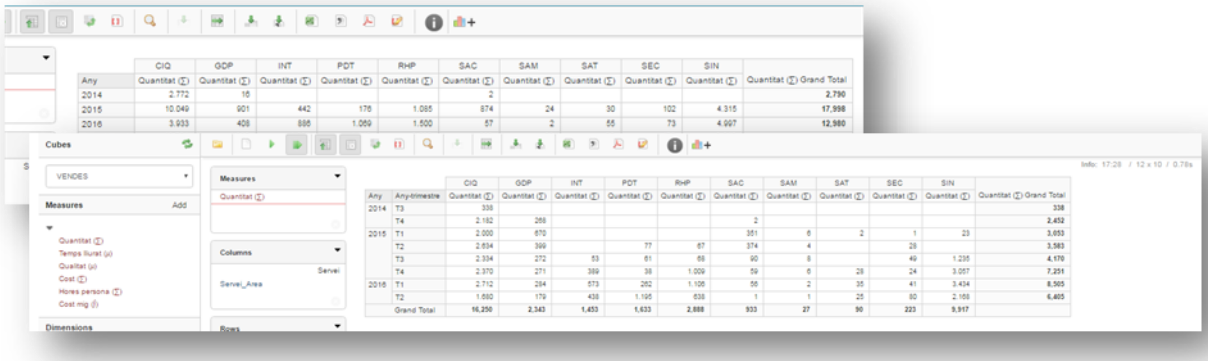


Ilustración 132. Cubo OLAP de la Diputación de Tarragona. Muestra de drill-down.

3.4 Infografía

Ya hemos visto en el capítulo de Visualización de datos la relevancia que tiene para la interpretación de la información la presentación gráfica de los datos. El siguiente tipo de informe, la Infografía, surge fundamentalmente para la información periodística, pero no deja de ser un recurso más para presentar información.

Como ya se ha comentado en otros apartados de este trabajo, la riqueza y calidad en la elaboración de informes nace muchas veces de la combinación de diferentes técnicas y componentes.

Aunque los gráficos tienen una existencia que data de muchísimos años atrás, desde la aparición del diario norteamericano USA Today en 1982 estos elementos ahora llamados infográficos (del inglés informational graphics) han revolucionado el diseño, especialmente el periodístico y editorial. De acuerdo con Richard Curtis, director de fotografía y diseño de dicho diario, "la gente lee los gráficos primero; algunas veces es la única cosa que leen".

Los infográficos son útiles para presentar la información que es complicada de entender a través del puro texto. El lector lee menos cada día y de acuerdo con Curtis, "los infográficos establecen la diferencia entre uno y otro diario" al convertir la publicación más visual, clara, directa y fácil de entender. Según Jeff Goertzen, ex asesor de gráficos de 'El periódico de Catalunya', "la infografía combina las habilidades del dibujo y diseño de un artista con las habilidades periodísticas de un reportero".

Peggie Stark, del *Poynter Institute for Media Studies* de St. Petersburg, Florida, nos dice que los infográficos son "una combinación de palabras y elementos visuales que explican los acontecimientos descritos en el texto y sitúan a la historia o a sus protagonistas en un contexto determinado". Pero muchas veces estos elementos mostrados en el gráfico no son descritos con lujo de detalles en el artículo; por eso se necesita un gráfico para hacer más clara y atractiva la información. El gráfico puede ser al final de todo el último recurso para contar una historia o para mostrar los datos.

Algunas definiciones de infografía:

- Josep María Serra: 'La infografía es un género periodístico que consiste en la transmisión de una información o noticia por medio de gráficos, dibujos, ilustraciones, fotografías… y que se realiza de forma manual o por medios informáticos'.
- José Manuel de Pablos: 'La infografía es la dualidad del dibujo analógico con el mensaje del texto, con el objeto de culminar una noticia'.
- Gonzalo Peltzer: 'La infografía es la expresión gráfica más compleja de información cuyos contenidos son hechos o acontecimientos, la explicación de cómo algo funciona, o la información de cómo es una cosa'.
- José Luis Valero Sancho: 'Aportación informativa, realizada con elementos icónicos y tipográficos, que permite o facilita la compresión de los acontecimientos, acciones o cosas de actualidad o algunos de sus aspectos más significativos, y acompaña o sustituye al texto informativo'.

Se concluye que la infografía es una representación más visual que textual, en la que intervienen descripciones, narraciones o interpretaciones, presentadas de manera gráfica normalmente figurativa, que pueden o no coincidir con grafismos abstractos y / o sonidos. La infografía nació como un medio de transmitir información gráficamente. Los mapas, gráficos,

viñetas, etc. son infogramas, es decir unidades menores de la infografía, con la que se presenta una información completa aunque pueda ser complementaria o de síntesis.

Las infografías transmiten hechos, procesos, noticias, eventos o datos de manera amena, sintética y visual, facilitando la compresión de información árida o compleja y estimulando el interés del lector que, de un golpe de vista, puede seleccionar en ellas lo que le interesa, lo que ya conoce y lo que no.

Los infográficos pueden incluir en las categorías de gráficos, mapas, tablas y diagramas.

Los gráficos son los más comúnmente utilizados y presentan información numérica y estadística. Se dividen, a su vez, en gráficos de barra, circulares y de líneas.

El gráfico de barras funciona preferentemente con unidades o porcentajes y lo que hace es establecer una comparación. Las barras presentan el mismo grueso y el largo depende de la cantidad que representen.

Ilustración 133. Ejemplo de Infografía Gráfico de barras. Fuente: Internet (2016)

El gráfico circular (tarta, pastel o queso) indica la división de partes de un todo y sus proporciones, especialmente en porcentajes. Está representado por un círculo que supone un todo y se encuentra dividido en partes. Estas partes no deben ser muchas, especialmente cuando suponen pequeñas partes del todo que se presenta, pues el gráfico se vuelve confuso y la información se ve desordenada.

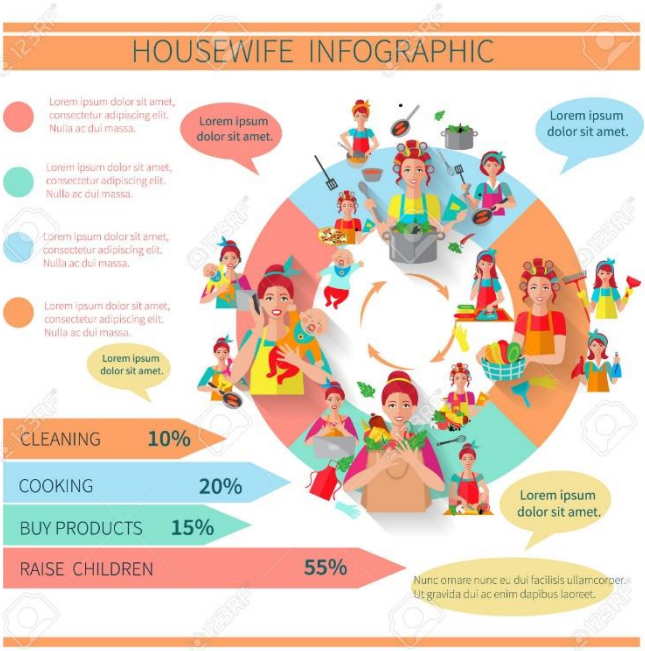


Ilustración 134. Ejemplo de Infografía gráfico circular. Fuente: Internet (2016)

El gráfico de línea muestra los cambios, expresados en números, a través del tiempo. Los gráficos de líneas funcionan si:

1. la línea que traza el cambio de cantidades representa un período de tiempo y
2. si cada cantidad establecida dentro de la línea representa incrementos por igual del tiempo indicado. Lo que ocurre es que algunas veces se quiere comparar incrementos o caída de cantidades entre lapsos de tiempo que no sin iguales, lo cual es engañoso y confunde al lector.

Ilustración 135. Ejemplo Infografía de líneas. Fuente: www.elpais.com (2016)

Los anteriormente descritos "son los gráficos más utilizados, pero los buenos diseñadores crean las formas apropiadas para la información numérica específica" y a veces decoran los cuadros. Esta decoración sólo será útil cuando tales elementos sirvan para que la información tratada sea más clara, precisa, entendible.

El mapa es necesario para mostrar la ubicación de un acontecimiento. El público lector está siempre interesado en conocer dónde ha ocurrido un determinado hecho. Cuando se realiza un mapa de una determinada zona de una ciudad, por ejemplo, a veces es sólo necesario ubicar las calles más importantes que circundan en suceso; dibujar cada calle puede ser confuso. Colocando zonas vecinales fáciles de identificar es suficiente, en indicar los puntos cardinales añade utilidad a la información. Para resaltar la ubicación se pude utilizar tramas grises o de colores.

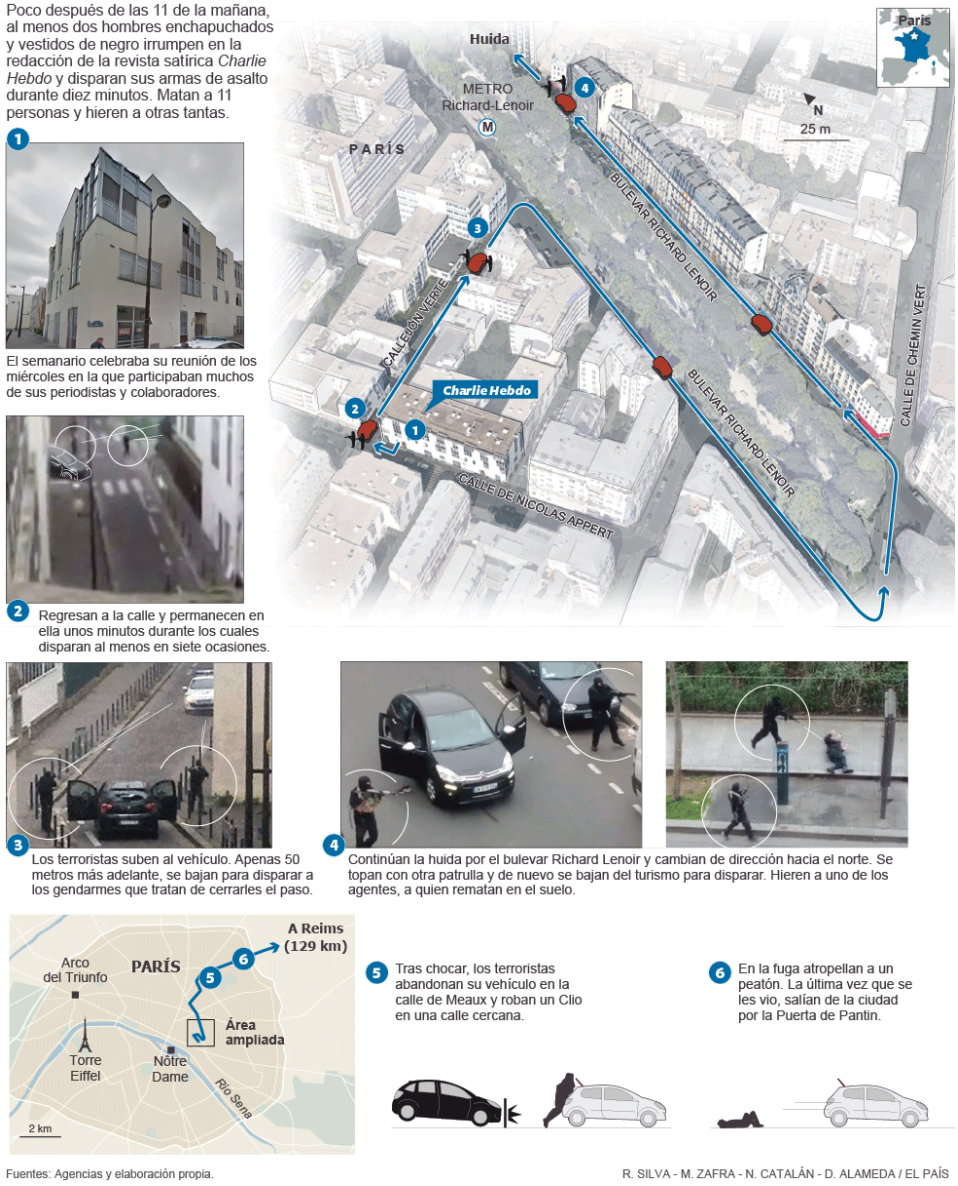


Ilustración 136. Ejemplo Infografía Mapa. Fuente: www.elpais.com

Para indicar la evolución de las condiciones climatológicas de un país, el mapa ha dado muy buenos resultados. Si bien un mapa no permite realizar nada emocionante en el gráfico, su función localizadora del hecho es muy importante para el usuario.

La tabla es un cuadro sencillo en el que se presentan datos descriptivos que, a veces, no son fáciles de cruzar y no se pueden comparar con facilidad. Puede aparecer como una simple lista de datos que se colocan en varias columnas, una al lado de la otra. Generalmente es buena cuando organiza información compleja que no puede presentarse utilizando, por ejemplo, un gráfico de barra o de líneas. Ejemplo de ello podemos ver en tablas que presentan horarios, distancias, encuestas, etc.

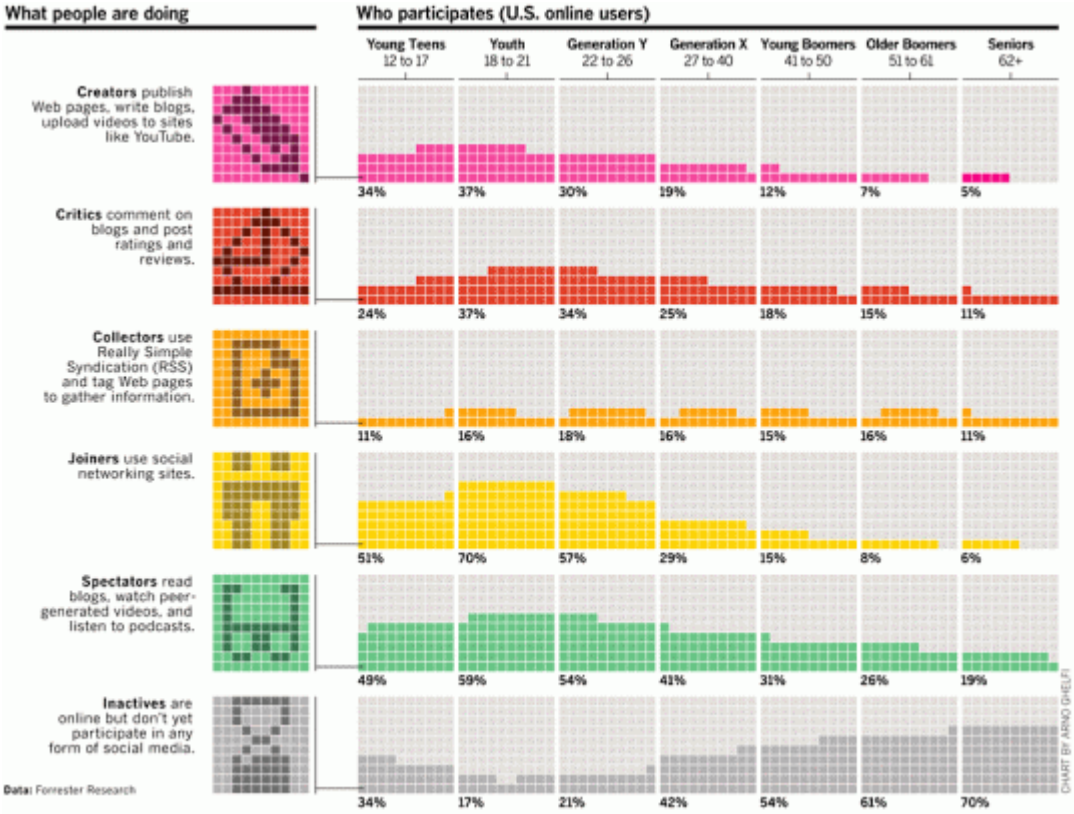


Ilustración 137. Ejemplo Infografía de Tabla. Fuente: Arno Gihelfi (2015)

Existen unas tablas de datos que sólo presentan información sumaria del texto. Estas tablas se utilizan para resumir la información del artículo y usualmente van acompañadas de algunos pictogramas que ayudan a la fácil identificación de la información allí tratada. Puede tratarse también de una cronología de hechos, de antecedentes o a manera de un cuadro sinóptico.

Los infográficos pueden volverse más elaborados de acuerdo a la complejidad del artículo. De esta manera, un gráfico de barras puede combinarse con un diagrama y un gráfico de tarta, por ejemplo, para expresar una sola información.

El diagrama es un gráfico que puede precisar de mayores habilidades artísticas. "Cuando el propósito del informe es mostrar cómo se ve o funciona algo, un diagrama es más apropiado

que los números o la prosa". Los objetos o sucesos pueden mostrarse con leyendas o pueden ser graficados de diversos ángulos, su interior, o cómo un objeto ha evolucionado. De esta manera, podemos graficar un accidente, el interior de un edificio o cómo un objeto ha evolucionado, el funcionamiento de una cámara de televisión debajo del agua o la caída de un niño en un pozo.

Ilustración 138. Ejemplo de Infografía Diagrama. Fuente: www.lavanguardia.com (2016)

Al final, el dilema clave con las infografías es la confusión con el arte. Alberto Cairo, en su libro Infografía 2.0 dice: "(Infografía estetizante es) aquella corriente que enfatiza el aspecto de la presentación, el "peso" visual del gráfico, el poder que tiene para hacer las páginas más atractivas, ligeras, dinámicas. La concepción estetizante concibe la infografía como un elemento ornamental e informativo *al mismo tiempo* a pesar de que en muchos casos se permita que los elementos estéticos obstaculicen la comprensión de las historias. Ésta es la corriente dominante

en la visualización de información en prensa en la actualidad y conduce, en sus manifestaciones más extremas (pero no por ello poco comunes), a *hermanar infografía con ilustración*".

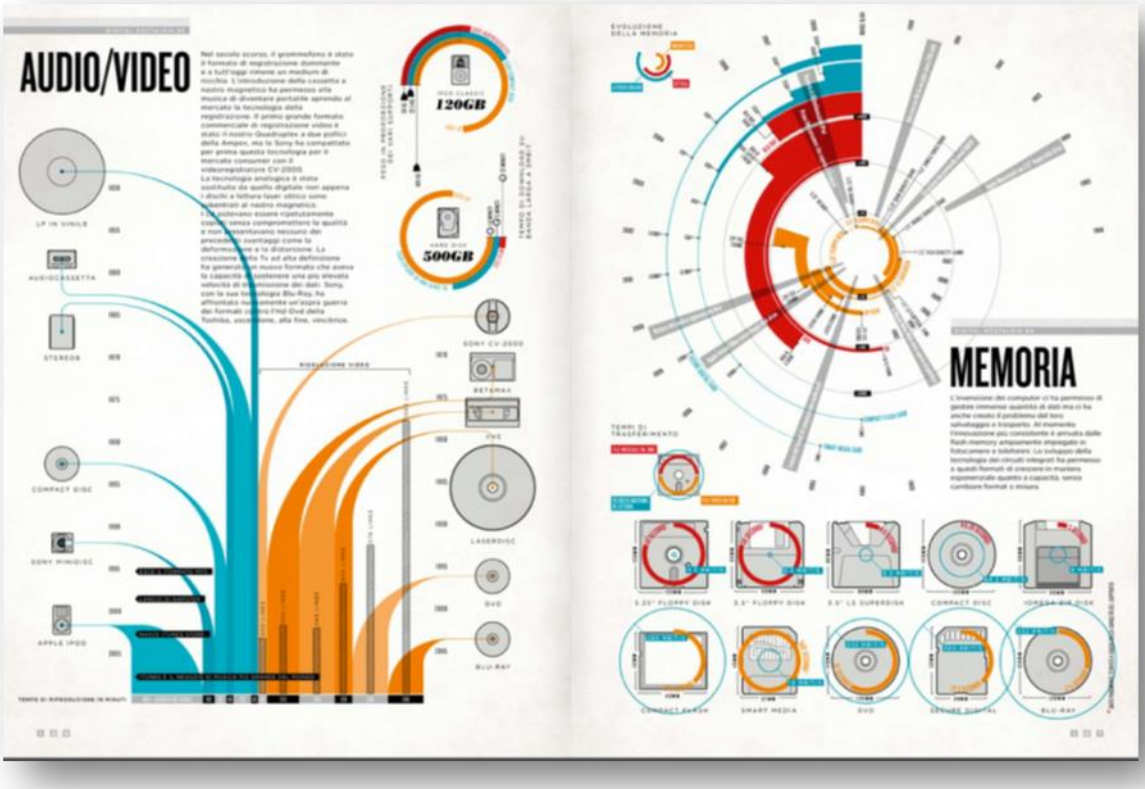


Ilustración 139. Ejemplo de Infografía no entendible. Fuente: http://sectiondesign.co.uk/wp-content/uploads/audiovideo.jpg

En la imagen superior puede observase este fenómeno, que muestra como una bella imagen no aporta comprensión a la información mostrada, al contrario, la obstaculiza.

Aunque la infografía se creó para dar comprensibilidad a los artículos periodísticos, actualmente existe la tendencia a "estetizar" y personalizar los informes de dashboard mediante ilustraciones.

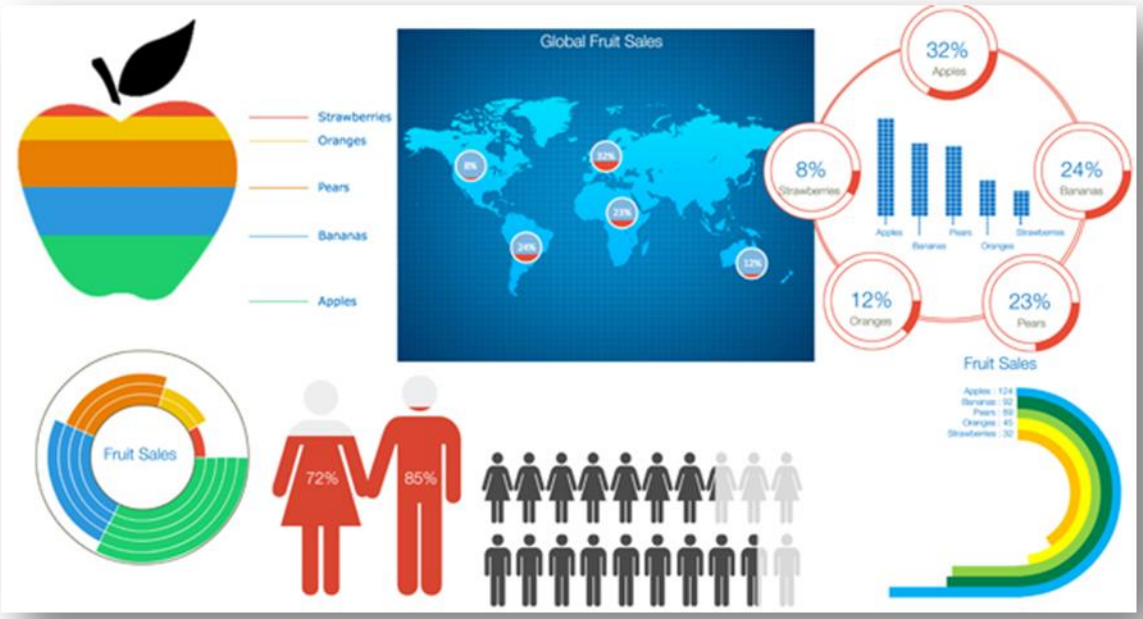


Ilustración 140. Ejemplo de dashboard con infografías. Fuente: https://blog.retargeter.com/more/infographic/online-advertising-dashboard-driving-growth-in-display-ads (2016)

Es conveniente entonces recordar que los elementos estéticos no deben obstaculizar la comprensión de la información, sino facilitarla.

4 Diseño de informes

El fruto del análisis de los datos deberá servir de base para la toma de decisiones. Es útil que los diferentes análisis se configuren en una serie de informes 'estándar'. El conjunto de informes 'normalizados' de una organización forman el Catálogo de Informes. Los informes le permitirán comprender el grado de ajuste entre lo planificado y lo ejecutado.

Un informe es el resultado de ubicar 'geográficamente' el resultado de una o varias técnicas analíticas. No es cuestión de embutir muchos análisis en un informe: hay que determinar aquellos que permitan comprender mejor el comportamiento de la organización.

Un informe puede ser interpretado (comentado) o no. En el caso de ser interpretado, es conveniente que quede claro qué tipo de información es el resultado del análisis técnico y qué tipo de información es el resultado del análisis humano.

De todos los referentes expuestos anteriormente, se considera deseable buscar la excelencia gráfica de acuerdo con Tufte, un informe excelente:

a. muestra los datos
b. hace que el espectador piense sobre el tema del informe
c. no altera los datos

d. ayuda a hacer comparaciones
e. es visualmente eficaz: gran cantidad de información con poca tinta (concepto data-ink ratio).

"Es el que da al espectador el mayor número de ideas, en el menor tiempo y usando la menor cantidad de tinta, el espacio más pequeño, y que dice la verdad acerca de los datos".

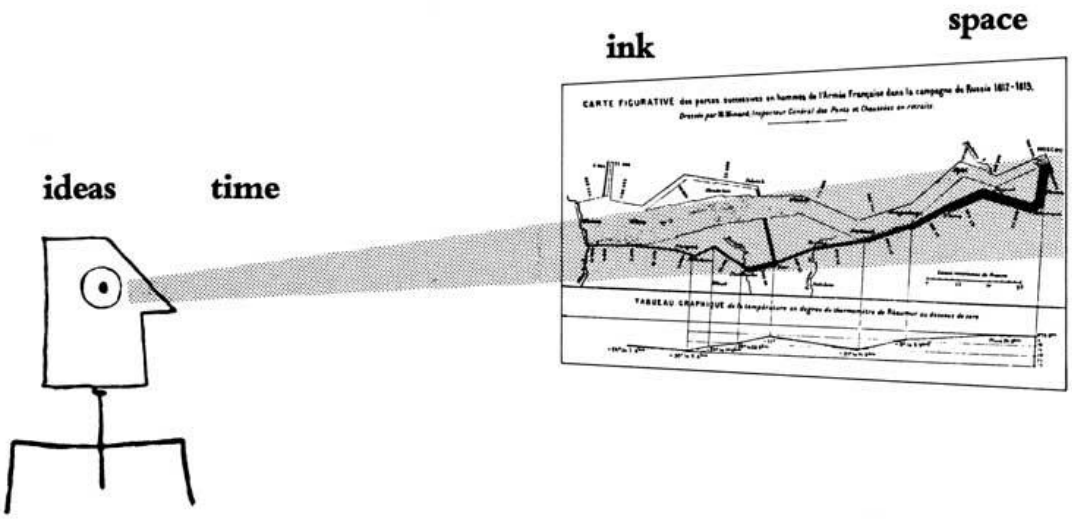


Ilustración 141. El informe perfecto según Tufte. Fuente: ed-informatics.org (2016)

Para elaborar informes íntegros:

- La representación física de los números medidos en la superficie de la gráfica debe ser directamente proporcional a los números representados.
- Cualquier posible distorsión o ambigüedad debe ser mostrada con un etiquetado claro.
- La variación visual debe representar la variación de datos real, no fruto del diseño
- Los valores de dinero en el tiempo deben estar ajustados (desinflacionados y estandarizados).
- No mostrar más dimensiones que datos.
- Los gráficos no deben mostrar datos fuera de contexto.

Debemos profundizar en el conocimiento sobre el diseño de informes. En un DIN A4, por ejemplo, podemos poner gran cantidad de información y no es trivial ponerla de una o de otra manera. Una serie temporal, por ejemplo, podemos representarla mediante una tabla numérica, mediante un gráfico de líneas, mediante un gráfico de áreas, etc. Si un informe contiene diversos análisis, la forma de representación de cada uno de ellos y la situación 'geográfica' dentro del DIN A4 (o de una pantalla de ordenador) es muy importante. El paso de un informe entendible a un informe no entendible es muy fácil.

4.1 Formato físico del informe

Un informe se diseña mediante la definición de las Áreas, Subáreas y la elección de componentes.

En nuestro modelo, un informe ocupa un área que está sobre una cuadrícula. En el caso de un DIN A4 la cuadrícula será:

- Horizontal: 22+168(21x8)+20
- Vertical: 25,8+257(15,7x16)+20

Que da un total de 8x16=128 cuadrículas o módulos, donde cada módulo tiene una dimensión de 21x15,7 que significa una proporción muy cercana a ¾.

En el caso de una pantalla de ordenador, la cuadricula se obtiene de la horizontalización de la del DIN A4.

- Vertical: 22+168(21x8)+20
- Horizontal: 25,8+257(15,7x16)+20

Estos módulos se pueden agrupar formando subáreas. Así, un informe está estructurado en 'Subáreas', desde 1 hasta un máximo teórico de 128. Las estructuras básicas de un informe son:

- 1
- 2 verticales
- 2 horizontales
- vertical + 2 horizontales
- horizontales + 1 vertical
- …

En la figura siguiente podemos ver diferentes ejemplos de estructuras de informes.

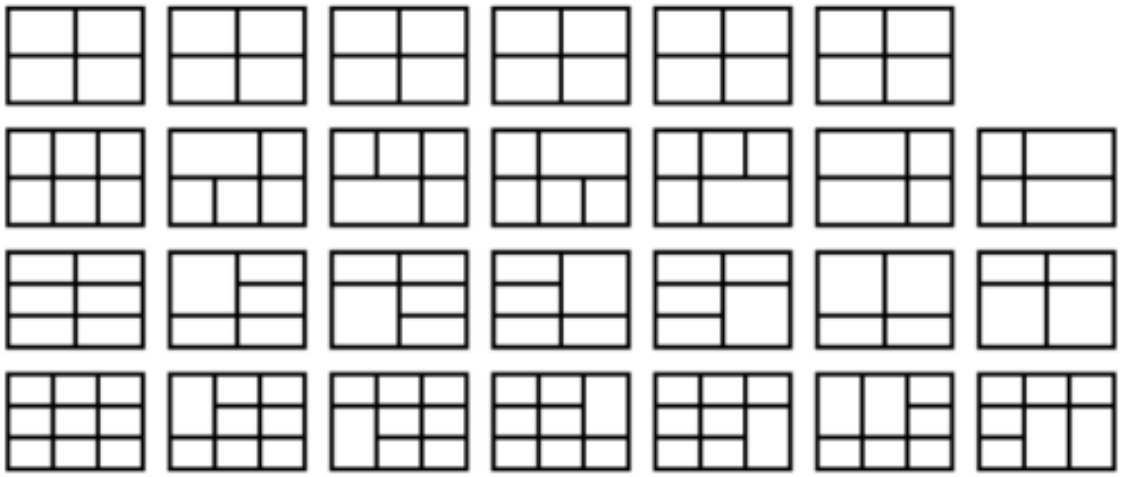


Ilustración 142. Ejemplos de configuraciones de informes. Fuente: Elaboración propia (2016)

Cada subárea incluirá un componente de los citados en el apartado anterior Catálogo de Componentes de visualización. Es importante destacar, que la posición en la estructura general del informe, es relevante, pues cuanto más superior y hacia la izquierda se coloca una información se percibe como más importante, y cuanto más a la derecha inferior, como menos, siguiendo el siguiente esquema, Kirk (2012):

+Importante **-Importante**

Superior izquierda	**Superior derecha**	**Inferior izquierda**	**Inferior derecha**

4.2 Componentes

En cada área o subárea de un informe se ubica un componente. Los componentes, si se trata de un informe digital, pueden permitir interactividad entre los otros componentes (situados en otras áreas) de un informe. Por ello es relevante planificar los saltos o vínculos entre componentes, de forma que los saltos sean lógicos y sigan el pensamiento natural del usuario en busca de la información.

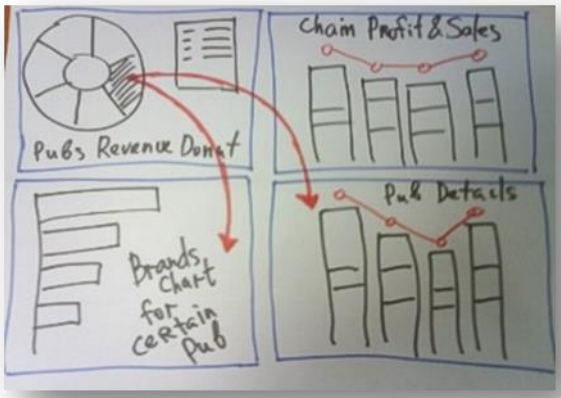


Ilustración 143. Planificar la interactividad. Fuente: Internet (2016)

Es conveniente mencionar la aportación de JeffHeer y Shneiderman con respecto a las dinámicas interactivas para el análisis visual.

Para mejorar la usabilidad de un informe es importante contar con una interfaz bien definida. "Ocho reglas de oro del diseño de interfaces" de Shneiderman son una buena guía para el diseño de interacción.

1 Buscar la consistencia. Secuencias coherentes de acciones deben ser requeridas en situaciones similares; terminología idéntica se debe utilizar en avisos, menús y pantallas de ayuda.

2 Permitir a los usuarios frecuentes usar "atajos": A medida que la frecuencia de uso se incrementa, también lo hacen los deseos del usuario para reducir el número de interacciones y

para aumentar el ritmo de la interacción. Abreviaturas, teclas de función, comandos ocultos, y las instalaciones de macro son muy útiles para un usuario experto.

3 Ofrecer retroalimentación informativa. Por cada acción del usuario, debe haber una cierta regeneración del sistema. Para las acciones frecuentes y de menor importancia, la respuesta puede ser menor, mientras que para las acciones poco frecuentes e importantes, la respuesta debería ser más sustancial.

4 Diseñar el cierre de la información. Secuencias de acciones deben organizarse en grupos con un principio, medio y final. La retroalimentación informativa en la realización de un conjunto de acciones da a los usuarios la satisfacción del deber cumplido, una sensación de alivio, la señal para abandonar su plan y las opciones de contingencia de sus mentes, y una indicación de que el camino está libre para prepararse para el siguiente grupo de acciones.

5 Prever la gestión de los errores. Tanto como sea posible, diseñar el sistema para que el usuario no pueda cometer un error grave. Si se comete un error, el sistema debe ser capaz de detectar el error y ofrecer mecanismos simples, comprensibles para el manejo del error.

6 Facilitar la reversión de acciones. Esta característica alivia la ansiedad, ya que el usuario sabe que los errores pueden ser deshechos; Por lo tanto, facilita la exploración de opciones desconocidas. Las unidades de la reversibilidad pueden ser una sola acción, una entrada de datos, o un grupo completo de las acciones.

7 Enfoque a usuario. Los usuarios avanzados desean sentir que controlan el sistema, diseñe el informe para dar respuesta a la experiencia de usuario, no a las necesidades de comunicación del diseñador.

8 Reducir la carga de la memoria a corto plazo. La limitación del procesamiento humano de la información en la memoria a corto plazo requiere que las pantallas se mantengan, y puedan visualizarse la información en páginas simples, múltiples.

4.3 Ejemplos de informes

Definimos el modelo propuesto para la Dimensión Informes, como el conjunto de dos o más componentes de visualización.

Informes univariables Recogen técnicas de análisis de, solamente, una variable. Sugerimos que todos los indicadores que se hayan diseñado en los procesos, sean susceptibles de ser representados en un informe. Será el nivel de análisis más bajo.

Un ejemplo de informe univariable muy básico podría ser el siguiente:

Ilustración 144. Ejemplo de informe con serie temporal y bullet graph

Este informe presenta los componentes serie temporal y bullet graph. Es decir, aunque sea una presentación muy simple, consideramos que constituye un informe porqué une varios componentes de visualización: Muestra en una tabla el nombre del indicador, la gráfica de línea con la serie temporal, la línea central oscura representa el valor real, la línea vertical oscura representa un valor objetivo y las bandas de colores representan rangos, como inferior, promedio y bueno. Además también se muestra el número total.

Este otro ejemplo informe combina la serie temporal con boxplots.

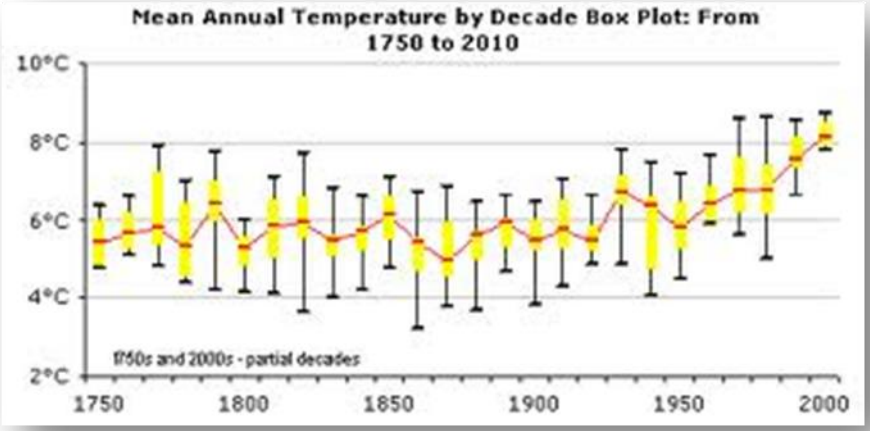


Ilustración 145. Ejemplo de informe con serie temporal.

En cada punto de la serie temporal se ha incluido un diagrama de caja, que muestra los valores máximos y mínimos, la dispersión, etc.

Otro ejemplo de informe que recoge diversas técnicas de análisis univariable, lo mostramos a continuación:

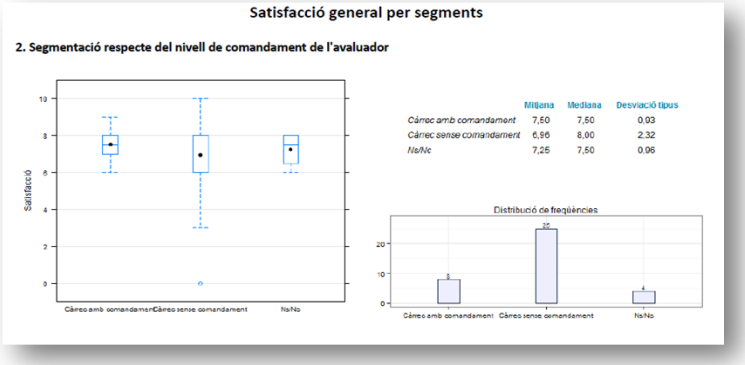


Ilustración 146. Ejemplo informe univariable. Fuente: Diputación de Tarragona (2014)

Se observa que el informe incluye una gráfica de caja; análisis de centralidad (media, mediana y desviación); análisis de dispersión (máximo, mínimo y cuartiles) y la tabla de frecuencias en un gráfico de columnas.

Informes multivariables: Otro informe, algo más elaborado, que incorpora diferentes técnicas y, además, sobre diferentes indicadores, se muestra a continuación.

En el siguiente informe de ejemplo se puede observar que se ha incluido una primera área con los títulos, valores y metas de los seis indicadores más relevantes de la Unidad de Gestión de Personas.

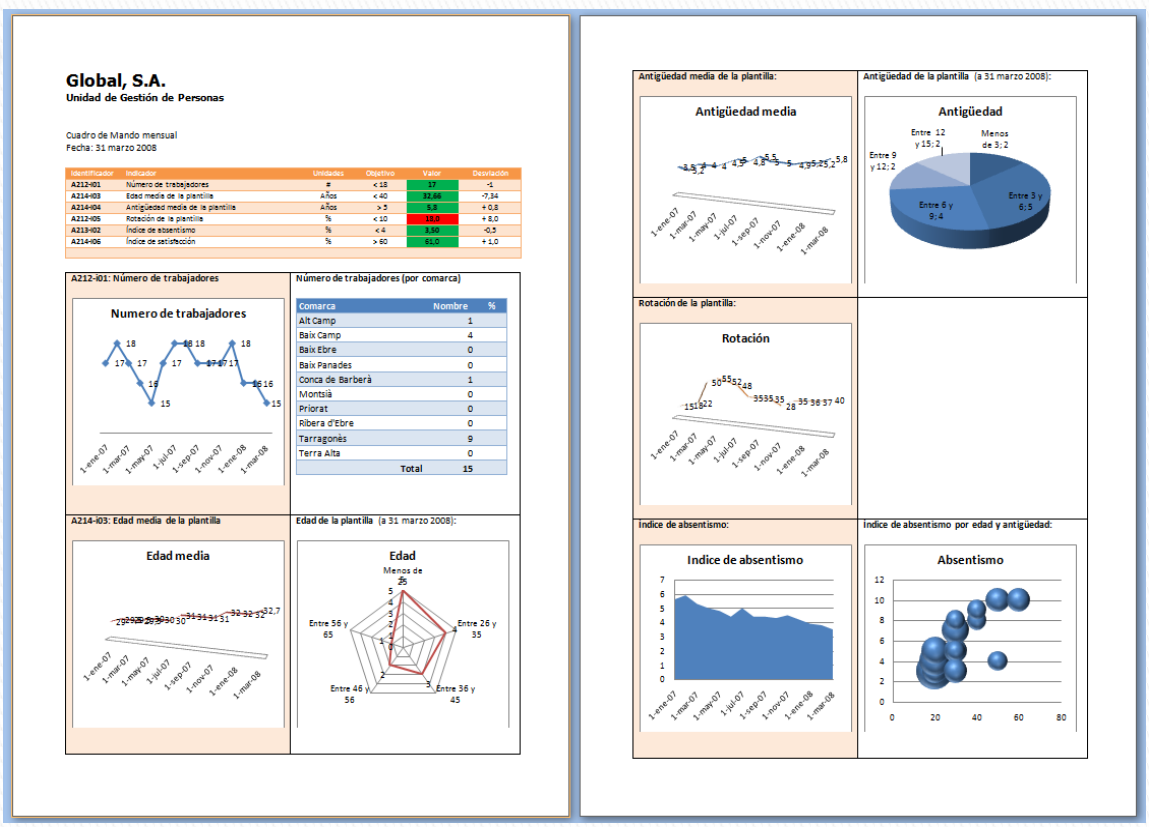


Ilustración 147. Ejemplo informe multivariable. Fuente: Enric Brull (2012)

Este informe incluye como puede verse, las series temporales (con gráficos de diferentes tipos) de los seis indicadores. Al lado de cada serie temporal, hemos incluido una taxonomía del valor de la última medida del indicador (la que corresponde al mes del informe). En el primer caso, el indicador de 'Número de trabajadores', al lado de la serie temporal correspondiente, hemos incluido una tabla que refleja la comarca donde viven los trabajadores. En el resto de los indicadores hemos incluido otros gráficos de detalle del último valor de la serie temporal.

Un tipo particular de informe bivariable son las Tablas de Contingencia o, dicho de otra forma, Cuadros OLAP.

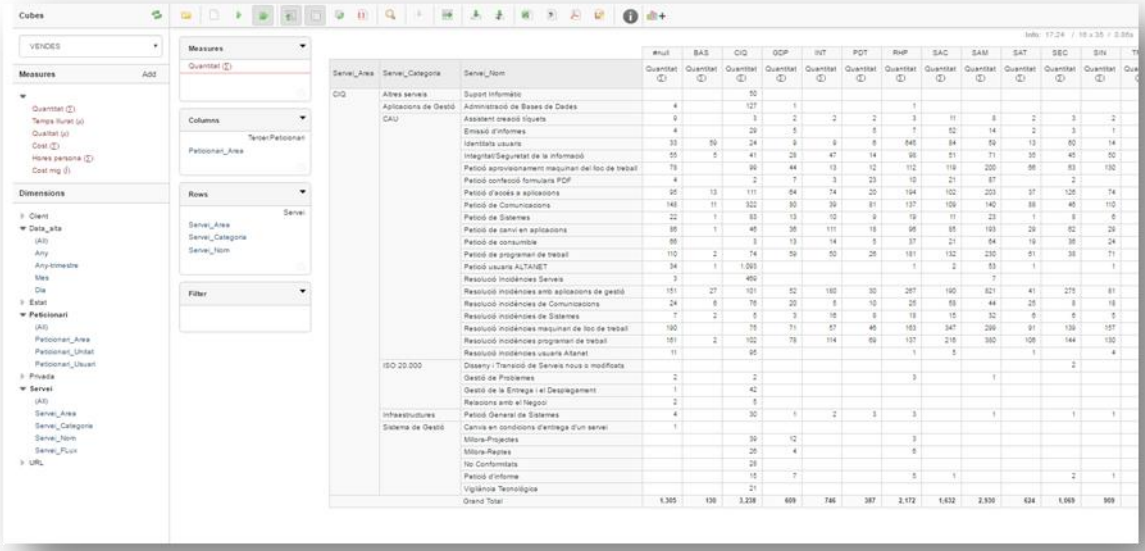


Ilustración 148. Ejemplo informe OLAP. Fuente: Diputació de Tarragona

5 El Sistema de Informes propuesto en el modelo SIEGO

En el siguiente apartado definimos como debería de ser la estructura y contenido del Sistema de Informes de una organización, correspondiente a la Dimensión Informes del modelo propuesto SIEGO.

5.1 Estructura

Un adecuado diseño debe de permitir identificar pocos informes pero que den la mayoría de la información necesaria. Un usuario debe conocer y recordar fácilmente la estructura de este sistema en los diferentes niveles, e identificar de forma clara dónde va a encontrar la información que busca, en un número finito de informes.

Los informes del Sistema de Informes que se propone responden a tres ámbitos básicos de la organización:

a. Estrategia. Una organización, por definición, existe para conseguir unos resultados: su estrategia. Por ello, los primeros informes, y los más relevantes, deben estar destinados a dar información del grado de cumplimiento de la estrategia, de cuál es el encaje entre la planificación que se realizó y su ejecución real.
b. Servicios o Productos. Para alcanzar la estrategia, las organizaciones deben tener definido un catálogo de servicios o productos, que ofrecen o venden. El segundo grupo de informes deben de estar destinados a dar información sobre la actividad del catálogo de servicios o productos. En el ámbito de la gestión de las organizaciones, para prestar un servicio debe realizarse un proceso, por lo que este apartado se refiere

al sistema de **procesos** de la organización. Los servicios o productos se prestan o venden a clientes, por lo que la información sobre **clientes** también se muestra en este bloque de informes.

c. **Recursos**. El tercer grupo de informes están destinados a dar información de los recursos consumidos en la prestación de los servicios o elaboración de los productos.

Se muestra a continuación un ejemplo de cómo estructura un informe mensual básico la Diputación de Tarragona, lo cual puede verse claramente en el índice:

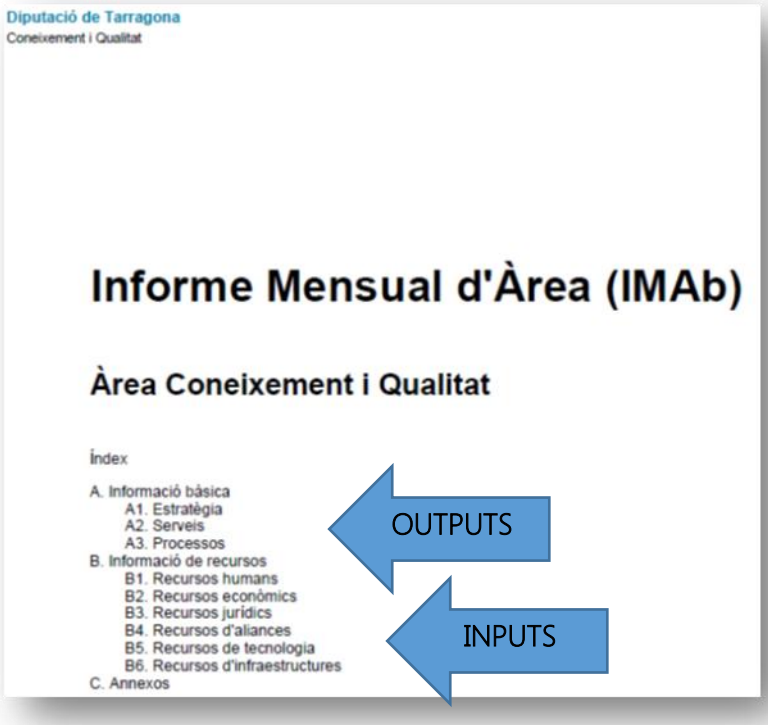


Ilustración 149. Ejemplo Informe outputs e inputs. Fuente: Diputación de Tarragona (2016)

Esta clasificación de informes responde a distintas inspiraciones en el ámbito de la gestión de las organizaciones como el modelo IPO (Inputs, procesos, outputs) o el modelo EFQM. Si la organización utiliza otro sistema de gestión, la estructura del sistema de informes debería de adaptarse. Por este motivo en esta dimensión del modelo SIEGO no establecemos una única clasificación de los informes, si no que la dejamos abierta al modelo de gestión de la organización.

Además de la clasificación temática de los informes, se propone también una clasificación por nivel estructural (jerárquico) de la organización y por nivel de detalle de la información.

En función del tamaño de la organización puede determinarse más profundidad del nivel estructural, pero se propone una clasificación en tres niveles:

a. Dirección general / Propiedad
b. Área / Departamento de primer nivel
c. Unidad / Sección / Departamento de segundo nivel

En lo referente al detalle del contenido, se propone dos niveles de profundidad:

a. Básico
b. Complementario

Atendiendo a esta clasificación, el sistema se estructuraría de la siguiente forma:

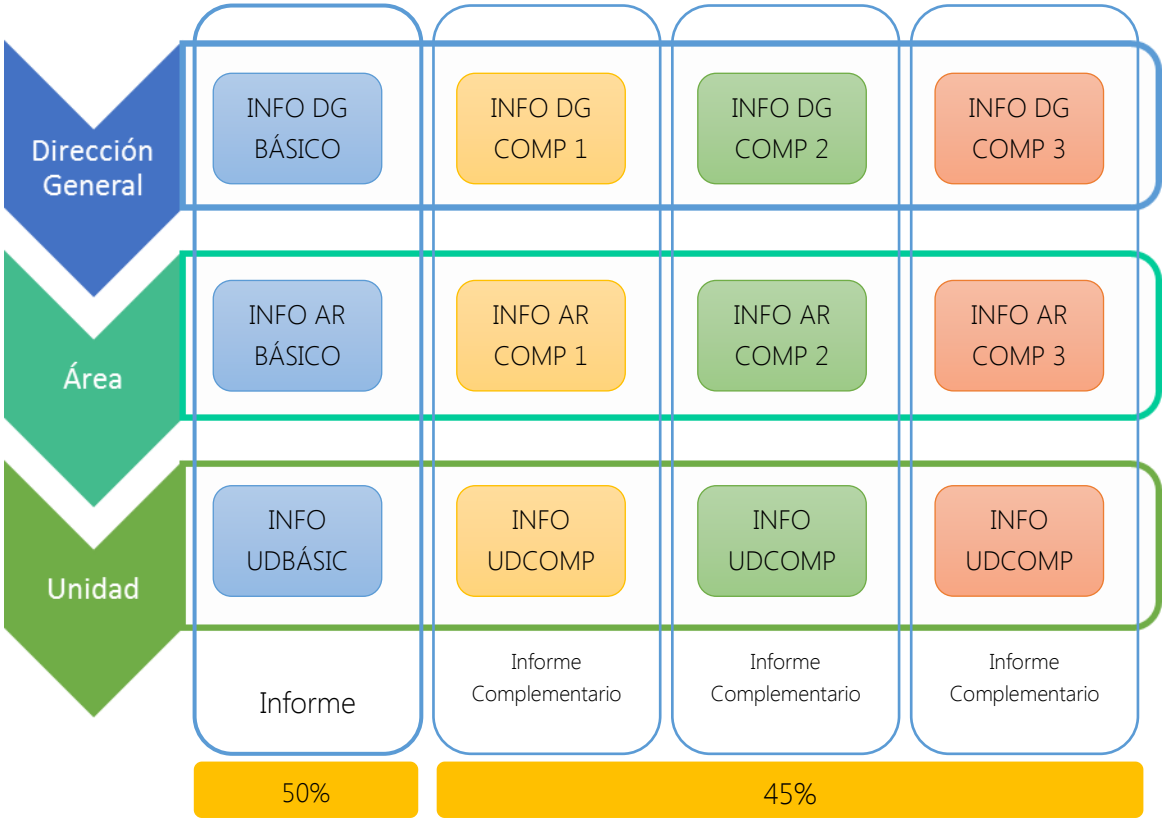


Ilustración 150. Sistema de Informes. Fuente: Elaboración propia

Con los informes básicos deberíamos poder dar respuesta a un 50% de las necesidades de información y con los complementarios al 45% restante, dejando un 5% de la información que no estaría disponible y habría que elaborar a demanda.

Un informe básico muestra información fundamental agregada, de forma que el gestor dispone de la información más relevante en un solo informe. Si precisa ampliar información sobre un aspecto concreto, por ejemplo puede ver el nivel de ejecución económica de su área, pero precisa ver el de un proyecto en concreto, deberá buscarlo en el complementario de Economía y Finanzas, en el que encontrará mayor nivel de detalle de esta información.

En general, se considera que los informes complementarios estáticos cubren el 15 % de la información y el 30 % restante se cubre mediante los informes complementarios dinámicos o interactivos, en los que el usuario interactúa con los datos para elaborar su propia vista de la información.

Cada nivel organizativo en la estructura de la organización debe disponer de un informe básico como apoyo fundamental al seguimiento y toma de decisiones de los responsables.

Se propone además que el informe básico para cada unidad organizativa se elabore de forma periódica fija, y se guarde un histórico para el seguimiento. Una buena periodicidad podría ser mensual o trimestral, pero ello va a depender mucho del tipo de organización y de su actividad.

También se considera adecuado que los informes tengan nombre propio en la organización, por ejemplo el informe DGB Mensual, el ARB Mensual, el UDC de Recursos Humanos, o el ARC de Economía y Finanzas. Ello facilita la comunicación en la organización y el seguimiento por parte de la dirección, además de reducir la curva de aprendizaje.

El contenido del informe básico se propone que sea el de la cadena de valor de la organización, como se muestra a continuación:

Parte A: OUTPUTS

- Estrategia / Proyectos
- Productos / Servicios
- Procesos

Parte B: INPUTS

- Economía
- Personas
- Infraestructuras (Equipos, Edificios, etc.)
- Proveedores
- ….

Otra posible estructura del informe básico, seria basado en los criterios del modelo EFQM:

- Liderazgo
- Estrategia
- Personas
- Procesos/ Productos / Servicios
- Alianzas y Recursos
- Resultados

5.2 La Gestión de los Informes

Una metodología fundamentada para gestionar las presentaciones de información será la que integre las diferentes recomendaciones hechas por las normas y estándares comentados.

La metodología que proponemos incluye las fases:

a. Identificar el mensaje
b. Planificar el informe
c. Elaborar el informe
d. Publicar el informe (que se trata en el apartado siguiente, Dimensión Publicar)

A continuación se describen, con más detalle, estos procesos.

5.2.1 Identificar el mensaje

Una organización, por definición, existe para conseguir unos resultados: su estrategia. Para ello, los primeros informes deben estar destinados a dar información del cumplimiento de su Estrategia.

Para conseguir su estrategia las organizaciones tienen diseñado un Catálogo de productos (o servicios) que ofrecen o venden. El segundo grupo de informes deben estar destinados a dar información de la actividad del Catálogo de Servicios.

Para poder obtener beneficios, las organizaciones deben tener equilibrados los ingresos y los gastos. El tercer grupo de informes deben estar destinados a dar información de los Recursos utilizados.

Hay que identificar claramente la finalidad, el uso, de la información que se va a elaborar. Dado que el mensaje puede ser desde muy general hasta muy detallado, la complejidad del informe elaborado puede ser muy distinta.

En esta fase se debe determinar el título del informe y los componentes que lo conforman. De cada componente se identificará no solo el tipo, sino sus ejes y unidades.

5.2.2 Planificar el informe

Como se ha comentado en los capítulos anteriores, disponemos de unos componentes básicos y de unas técnicas analíticas que nos van a permitir elaborar los informes.

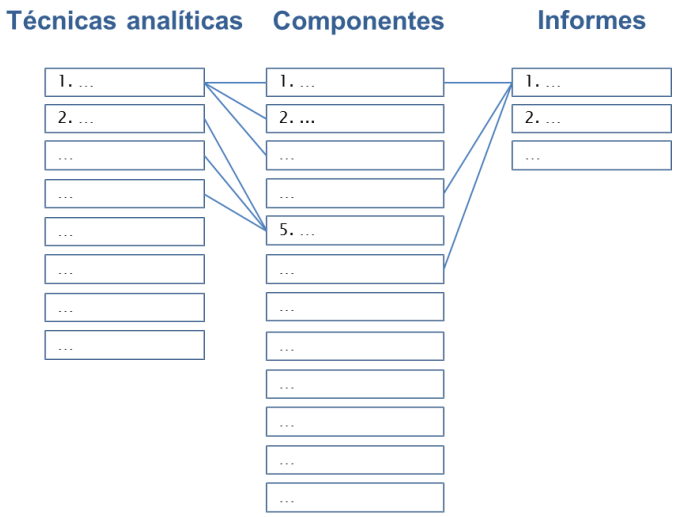


Ilustración 151. Elaboración de informes. Fuente: Elaboración propia.

Con esta base, se escoge el tipo (o tipos) de componentes que conformaran el informe y que se adaptan mejor a la naturaleza del mensaje. Como se ha comentado en capítulos anteriores, una regla para esta decisión es que **la forma depende de la función**. Si el objetivo del informe es:

- observar la tendencia de una variable, se deben utilizar gráficos de líneas o series temporales.
- facilitar la comparación entre cifras, se deben utilizar gráficos de barras
- identificar las localizaciones de ciertos fenómenos, se deben utilizar mapas
- mostrar un proceso, se deben utilizar diagrama de relaciones.

5.2.3 Elaborar el informe

Un informe gráfico tiene tres partes fundamentales: La cuadricula, la tipografía y las ilustraciones. Como elementos complementarios destacamos los colores y los componentes.

La cuadrícula

La composición de un informe permite conseguir la disposición perfecta de los elementos, proporcionada a un marco dentro del cual debe operarse para conseguir una sensación.

El elemento que permite la composición es la retícula o cuadrícula.

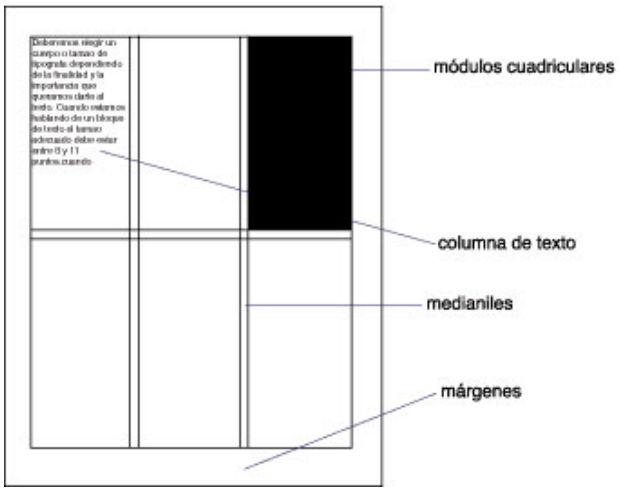


Ilustración 152. Formato de un informe.

Aunque podemos utilizar cualquier tipo de cuadrícula, si atendemos a las recomendaciones de Dan Pedoe, deberíamos tener tendencia a utilizar las proporciones puesto que se ha demostrado que las formas que cumplen con alguna de las normas de proporcionalidad (Vitrubio, número áureo, etc.) se perciben como más bellas y agradables.

Tipografía

La tipografía es el elemento que soporta el texto del informe y debe diseñarse pensando en el efecto que se pretenda conseguir. La tipografía es un elemento comunicador.

Ilustración 153. Tipografía.

Las serif se clasifican en las siguientes clases: Romanas Antiguas, Romanas de Transición, Romanas Modernas y egipcias.

Las sans serif se clasifican como: Grotescas, Neogóticas, Geométricas y humanísticas.

Colores

El color también es un elemento comunicador. Algunos colores pueden resultar inapropiados o ilegibles. Mediante el color podemos crear la atmósfera adecuada al mensaje del informe.

Ilustración 154. Escala de colores.

a. Matiz: Es el estado puro del color, sin el blanco o negro agregados.

b. Primarios: Rojo, Amarillo y Azul
c. Secundarios: Naranja, Verde y Violeta
d. Complementarios: se encuentran simétricos respecto al centro.
e. Complementarios cercanos: se encuentra al lado del complementario.
f. Triadas complementarias: se encuentran separados 120º

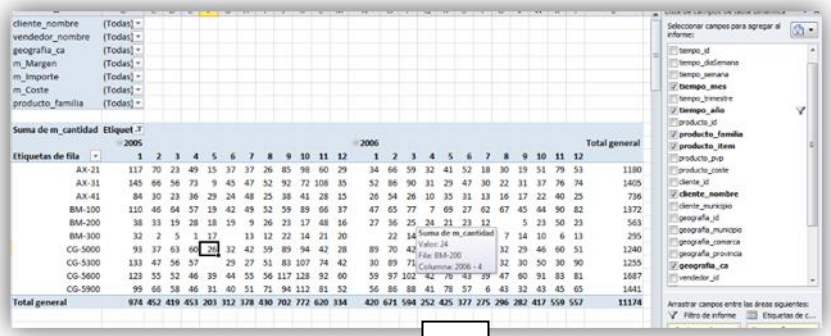
6 El proceso de forma gráfica

A continuación, a modo de resumen esquemático se muestra el proceso de elaboración de un informe.

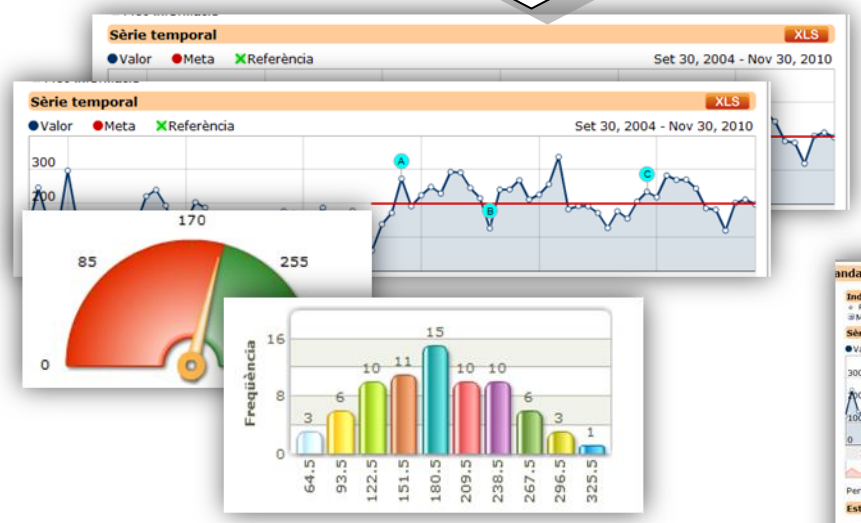
a) Recoger datos (Dimensión Datos)

	Tiempo		Producto			Cliente			Métricas			
venta_id	tiempo_id	...	producto_id	producto_nombre	...	cliente_id	cliente_nombre	...	Cantidad	Importe	Coste	Margen
1	05/01/2008		23	Interruptor		34	Soltex, S.L.		7	23,10	17,5	5,60
2	15/01/2008		41	Tubo fluorescente		78	Agriram, S.L.		3	6,30	3,6	2,70
3	22/01/2008		17	Enchufe		18	Quimiclar, S.A.		24	74,40	45,6	28,80

b) Aplicar técnicas de análisis (Dimensión Análisis)



c) Seleccionar y aplicar componentes de visualización (Dimensión Visualización)



d) Preparar Informes (Dimensión Informes)

Dimensión Publicar

1 Publicar los informes 221

1 Publicar los informes

En este capítulo mostramos la última de las dimensiones del modelo SIEGO, la dimensión publicar, que se realiza mediante el Portal de Informes.

Una vez preparados los informes como hemos visto en el capítulo anterior, estos deben ponerse a disposición de los diferentes tipos de usuarios y roles. Es conveniente, que cada usuario pueda navegar y parametrizar los informes de acuerdo a sus necesidades y conocimiento, y es por ello que se considera adecuado que la puesta a disposición de los informes se haga mediante un portal web interactivo. El usuario navegará por las carpetas del portal hasta encontrar el informe requerido, y a continuación irá construyendo de forma interactiva el contenido que precisa.

Un portal web no deja de ser una página web de la organización, que generalmente no es visible para usuarios externos.

Un sitio web es un conjunto de archivos electrónicos y páginas web referentes a un tema en particular, que incluye una página inicial de bienvenida, generalmente denominada home page, con un nombre de dominio y dirección en Internet específicos. Una página web se define como un documento electrónico el cual contiene información textual, visual y/o sonora que se encuentra alojado en un servidor y puede ser accesible mediante el uso de navegadores

Un portal web se puede definir como la evolución del concepto "sitio web", en el cual el sitio web se ha convertido en el punto de entrada a un conjunto de servicios de información a los que se accede de forma sencilla, unificada y segura.

Nuestra propuesta, en el modelo SIEGO es un portal de Informes, un portal con tecnología web temático, que ofrece al usuario, de forma fácil e integrada, el acceso a una serie informes de reporting, dashboards, cubos OLAP, etc. Este portal de informes es en realidad un sitio web privado de la organización, para el que se requiere autenticación para poder acceder.

El portal de informes está orientado a resolver las necesidades de información de cualquier tipo para cualquier persona de la organización.

El portal es el punto de entrada. Al acceder a él, el usuario accede a un conjunto de carpetas con todos los informes disponibles para el usuario en concreto o para el rol establecido, que es la forma de establecer permisos y funcionalidades para grupos de usuarios. Cada usuario

visualiza las carpetas para las que tiene permiso de acceso, y una vez clica en alguna de ellas, visualiza los informes disponibles que puede ejecutar. Además, al ejecutar el informe solicitado, la información que se muestre también podrá tener diferentes niveles de acceso, según sean sus permisos personales.

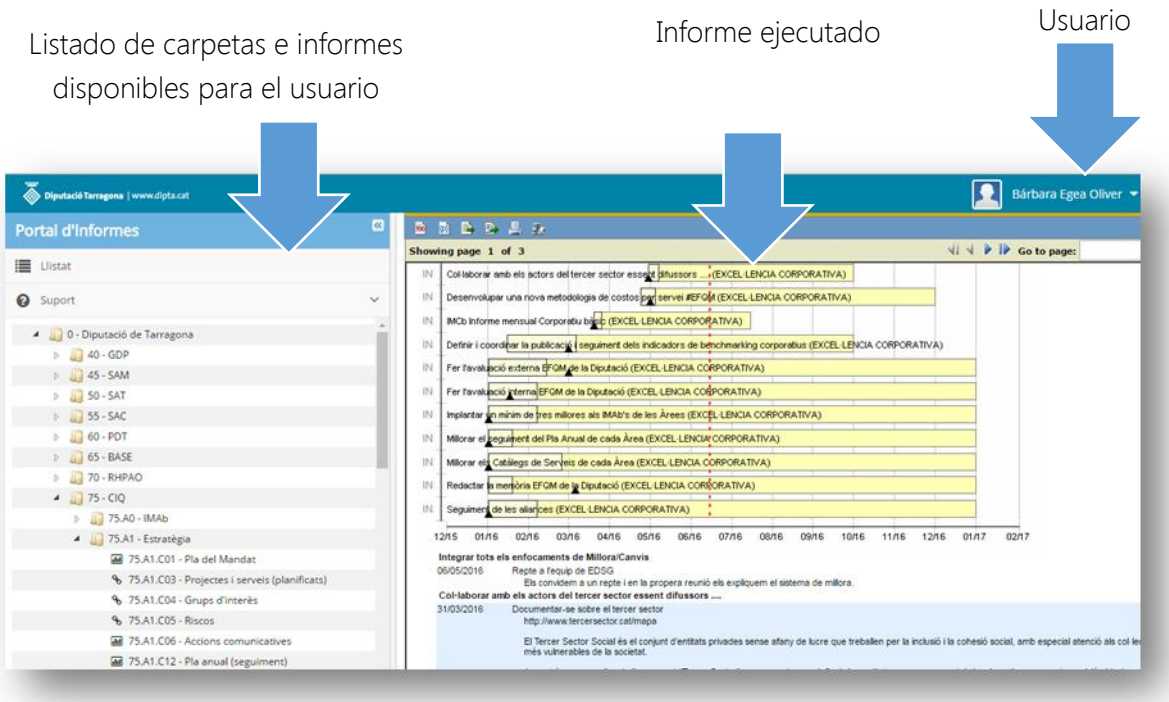


Ilustración 155. Portal de informes de la Diputació de Tarragona (2016)

Los informes deben de ponerse al alcance de los interesados de manera rápida y sencilla. Por ello se propone habilitar un portal de informes que se estructure en carpetas (parte izquierda de la imagen) que se despliegan en diferentes subniveles. En cada carpeta podemos alojar una o varias carpetas, y a su vez, en cada carpeta uno o varios informes.

Con este concepto de portal, permitimos a los interesados obtener los informes en una sola interface: no hace falta ir a la aplicación de nóminas para obtener el informe de personas, a la de contabilidad para obtener el informe de finanzas, etc. Pueden encontrarlos en un solo sitio, organizados por carpetas. Además, pueden consultarse informes que cruzan diferentes fuentes de información que de otra forma no sería posible.

El portal permite el almacenamiento centralizado de los informes, su clasificación, ordenación, etc. y la generación bajo petición del usuario de un nuevo informe a partir de la unión de la plantilla del informe con los datos actualizados. Esta generación es visible en pantalla y exportable a la mayoría de los formatos electrónicos. El usuario puede imprimir el informe o bien exportarlo y guardarlo.

Tal y como se ha planteado en el anterior capítulo, el tiempo necesario para disponer y comprender la información de un informe se puede descomponer en diferentes microtiempos:

3. Tiempo transcurrido entre que surge la necesidad de información hasta la obtención de la información. Se puede subdividir a su vez en:
   c. Tiempo entre la detección de la necesidad de información hasta el acceso al Informe (t1)
   d. Tiempo entre el acceso al Informe y la obtención de la información (t2)
4. Tiempo transcurrido entre la obtención de la información y la interpretación adecuada de los datos (t3)

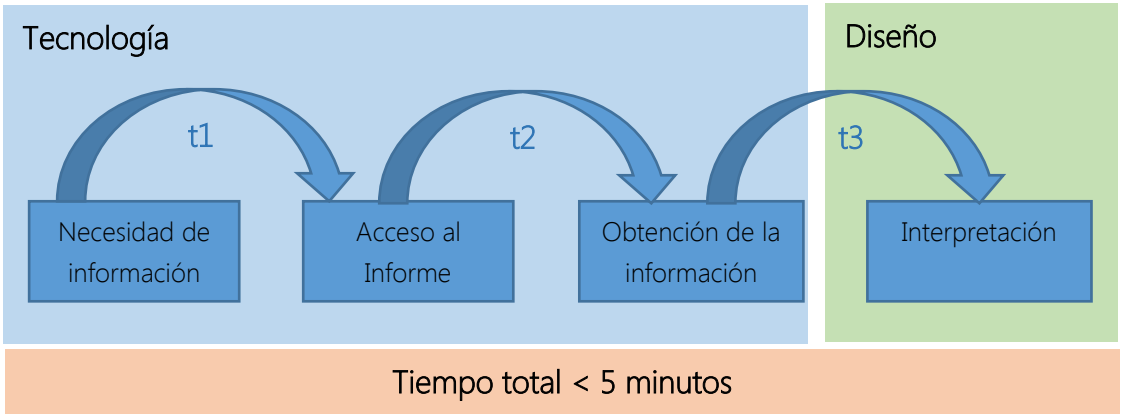


Ilustración 156. Tiempo de respuesta a una pregunta. Fuente: Elaboración propia (2016)

El portal de informes, permite reducir considerablemente el microtiempo (t2) entre el acceso al Informe y la obtención de la información, puesto que el usuario sabe perfectamente dónde encontrar cada tipo de información requerida.

En el ámbito tecnológico es fundamental disponer de acceso a la información vía un servicio web, accesible desde cualquier navegador, que permita el trabajo desde cualquier tipo de plataforma: pc, tablets, smartphones, etc. Ello tiene unas connotaciones técnicas relevantes en lo relativo a la elección de la tecnología.

También cabe mencionar, que el acceso al portal de informes debería realizarse con un sistema de validación de usuarios corporativo, que se integre con el sistema de permisos establecido, de forma que no sea necesaria una segunda validación al acceder a él, sin necesidad de autenticarse múltiples veces, con las tecnologías llamadas mecanismos de Single Sign On.

Los usuarios pueden ser agrupados de forma intuitiva mediante una jerarquía en la organización, o utilizando "grupos de usuarios" transversales, ofreciéndose así una enorme flexibilidad y facilidad de administración. Resulta posible, por ejemplo, agrupar mediante grupos a miembros situados en cada ubicación geográfica, o bien por tipo de rol (usuario, gestor, director, etc.) y simultáneamente articular a través de grupos de usuarios de diferentes ubicaciones equipos departamentales o basados en proyectos. La tecnología permite asignar carpetas web tanto a organizaciones como a grupos de usuarios, disfrutando cada uno de ellos

de su propio conjunto de informes, su sistema de gestión de contenidos, su calendario compartido y su sistema de autorizaciones. Un usuario puede pertenecer a diferentes sitios y navegar fácilmente entre ellos.

En el ámbito del diseño, además de la aplicación de componentes de visualización, es muy importante que el Portal de Informes esté bien estructurado y organizado por niveles y tipos, de forma que el usuario sepa siempre dónde encontrar la información buscada en el menor tiempo posible.

Recuperamos en este capítulo la estructura de informes propuesta en el capítulo de Sistema de Informes para ver como el Portal de informes debería tener una estructura paralela.

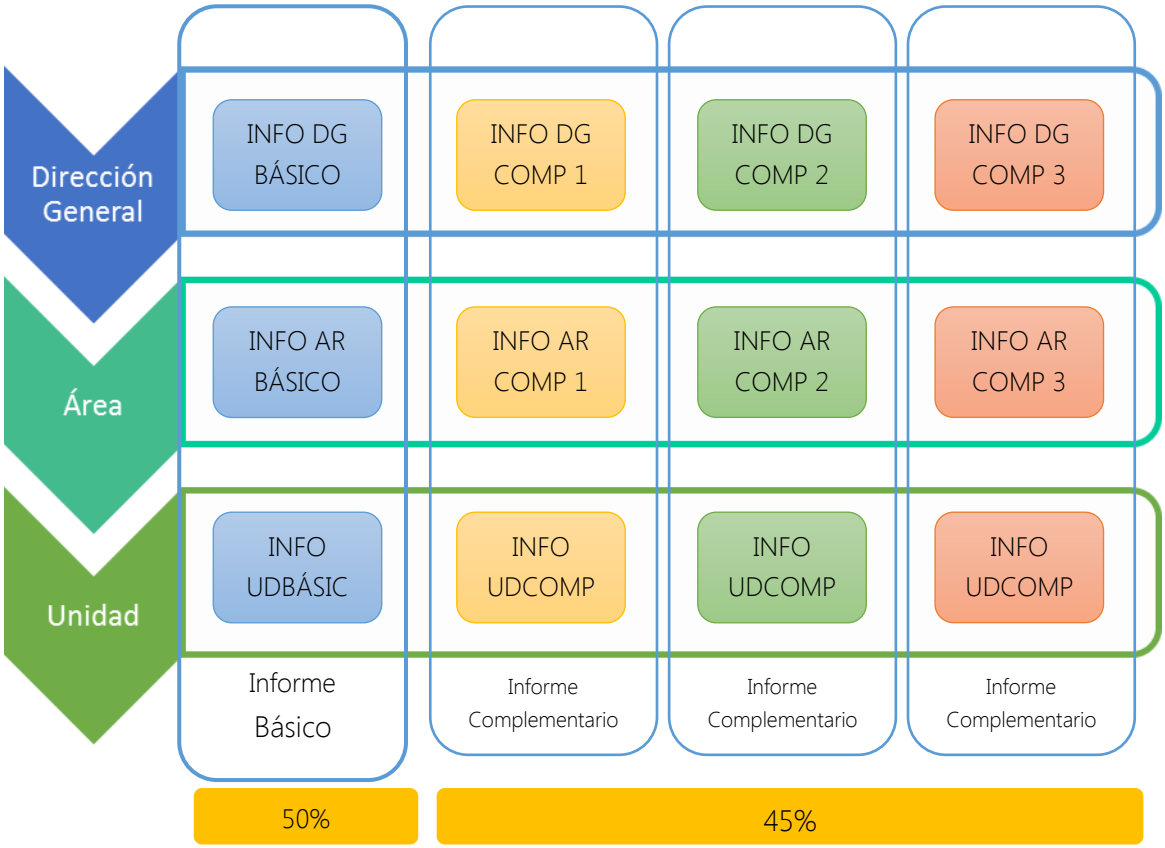


Ilustración 157. Sistema de Informes. Fuente: Elaboración propia

Es decir, el primer nivel de carpeta seria el del nivel jerárquico superior, como por ejemplo Dirección General y en el siguiente nivel de carpetas se mostraría la categorización por temas, como se muestra en el ejemplo:

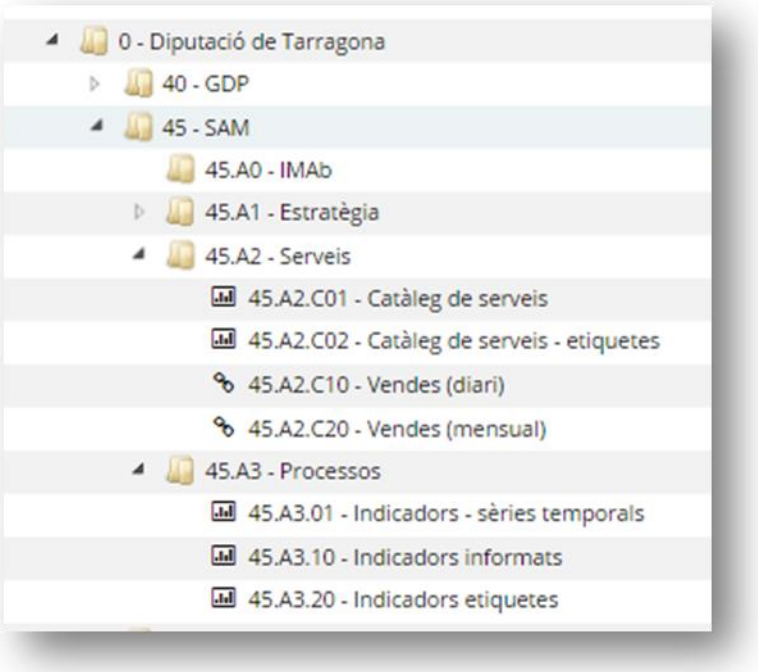


Ilustración 158. Sistema de carpetas en el portal de informes de la Diputació de Tarragona (2016)

El portal de Informes puede ser una herramienta muy sencilla, que se conciba simplemente como el repositorio único de informes de la organización, pero puede irse sofisticando y evolucionando en función de las necesidades y el mayor conocimiento de la organización.

Por ello, por ejemplo puede establecerse una pantalla de inicio con los últimos informes accedidos y una lista con los informes preferidos, que el usuario pueda determinar y cambiar según su criterio.

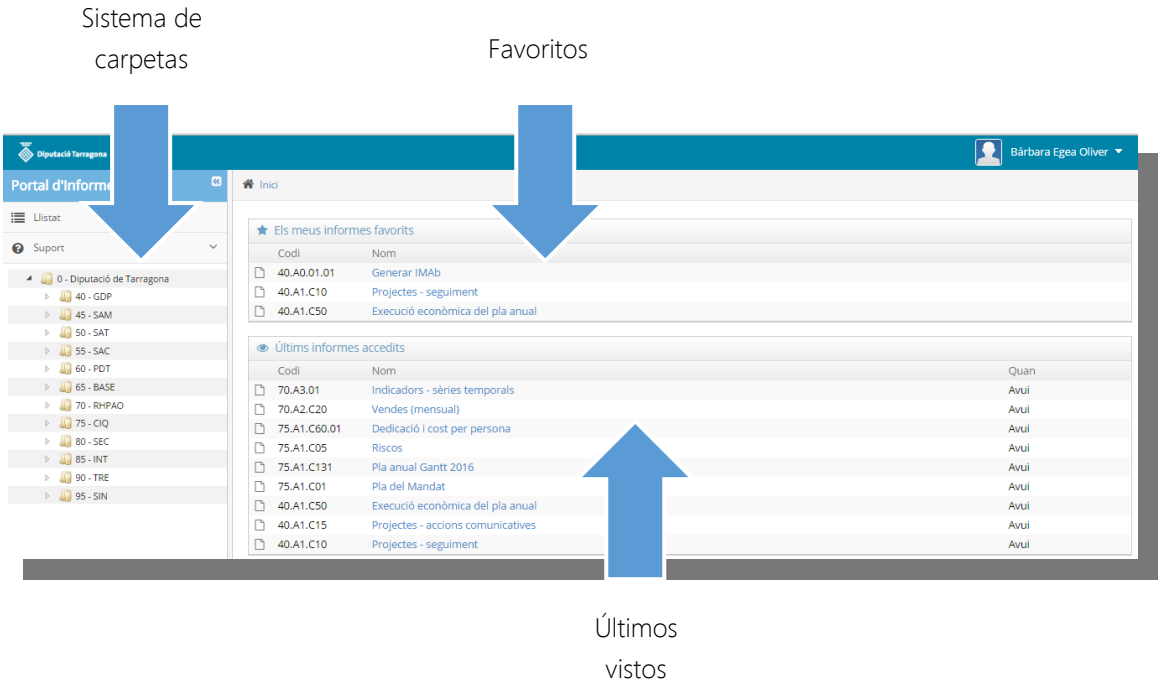


Ilustración 159. Portal de informes Diputació de Tarragona. Fuente: Diputació de Tarragona (2016).

También se puede establecer un sistema de alarmas y correos electrónicos que avisen al usuario si ya se ha generado un informe o si determinado indicador de un informe dashboard sobrepasa la meta o cumple con alguno de los parámetros establecidos como alarma.

De otra parte se podría integrar con un buscador de contenidos, incluir un servicio wiki, etc.

En resumen estas serían las características del Portal de informes:

- Recursos informativos (informes de todo tipo) organizados en carpetas por temas
- Interacción personalizada con los servicios del portal, el usuario puede interactuar con los informes y confeccionar sus propias peticiones de información siempre en el ámbito de contenidos al que tiene acceso
- Acceso a información de fuentes diversas, agregada y estructurada.
- Integración de herramientas de colaboración como grupos de trabajo, comunidades, foros, etc.
- Búsqueda, favoritos, alarmas…

Debido a que cada vez más organizaciones se ven en la necesidad de hacer pública información relativa a sus operaciones y actividades, esta podría ser una nueva funcionalidad añadida al portal de informes. Bien por ser una administración pública, bien por buscar un reconocimiento de responsabilidad social corporativa, bien por rendición de cuentas, el portal podría incluir informes de carácter público.

También cada vez más organizaciones quieren establecer servicios interactivos con sus clientes para informarles de los productos o servicios prestados, plazos, estado de los pedidos, calidades, nivel de satisfacción, etc.

Es por ello que podría plantearse la posibilidad de que los informes preparados en el portal de informes pudieran publicarse de forma abierta en otros portales de la organización, como sería la web corporativa, la web de RSC o en el caso de las administraciones públicas, directamente en el Portal de Transparencia o la Sede Electrónica.

La novedad de este portal estriba en el desacoplamiento del resto de dimensiones del modelo, los cual permite integrar en el todo tipo de informes de todo tipo de sistemas de la organización. Las soluciones de Business Intelligence en general trabajan con un sistema de licenciamiento que limita el acceso a la visualización de dashboards, cubos de datos o reporting al número de licencias obtenidas, y no permite la integración de otros sistemas de reporting de la organización.

Capítulo IV. Estudio Empírico

1. Preguntas de tesis y metodología

1 Objetivos del estudio empírico 231

2 Hipótesis 232

3 Posicionamiento epistemológico 233

4 Elección del método 235

5 Unidad de análisis y población 236

6 Muestra 238

6.1 Cálculo del tamaño muestral 238

7 Cuestionario y preguntas 239

7.1 Preguntas sobre resultados 243

7.2 Herramienta utilizada 245

7.3 Tipo de Escala y propiedades socio métricas asociadas 246

7.4 Fiabilidad de las escalas seleccionadas en la encuesta 247

7.5 Validez 250

1 Objetivos del estudio empírico

Más allá de la propuesta de este trabajo de investigación, y del modelo presentado de sistema de información, el objetivo de la parte empírica de este trabajo es evaluar las prácticas de las organizaciones en referencia a los sistemas de datos, la aplicación de técnicas de análisis, componentes de visualización, elaboración de informes y su publicación.

Respondiendo pues a estos objetivos, las preguntas de esta investigación son:

A. ¿Las organizaciones que recogen datos, aplican técnicas analíticas, componentes de visualización, preparan con ello informes y los publican, perciben mejores resultados?
B. ¿Qué tipo de datos recogen las organizaciones de sus actividades y procesos?

Sobre la pregunta A:

No se dispone de evidencia sobre la realidad práctica de las organizaciones con respecto a la recogida de datos y su uso para sistemas de información. Tampoco sobre la evidencia de hasta qué punto estos datos son utilizados para el análisis de la actividad de la organización y para la toma de decisiones.

Debido a que los estudios científicos evidencian la conveniencia de medir ciertos aspectos fundamentales de la actividad de las organizaciones, se pretende conocer si las organizaciones recogen efectivamente datos de estos aspectos.

Esta pregunta se investiga mediante las hipótesis:

H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados.

H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas.

Sobre la pregunta B:

Una consecuencia clara de la mejor toma de decisiones en una organización, sería la de obtener mejores resultados. Por ello se pretende determinar si existe relación o no entre la percepción de resultados y la disposición o no de un sistema de información parecido al que se plantea en este trabajo.

Debido a que la finalidad de este trabajo de investigación no es conocer el grado concreto de mejora de los resultados obtenidos, si no la percepción que sobre ellos tiene la persona que responde al cuestionario, no se formulan preguntas concretas sobre los resultados de la organización, sino preguntas sobre percepción de resultados en general y sobre el clima laboral.

Se da respuesta a esta pregunta mediante las hipótesis:

H3 Las organizaciones que disponen de datos de:

- Planificación y seguimiento de la estrategia
- Personas (nómina, altas, bajas, formación, etc.)
- Economía y finanzas (contabilidad, costes, etc.)
- Gestión del conocimiento y TI
- Compras / Proveedores
- Clientes / CRM
- Procesos / Productos / Servicios

Perciben mejores resultados.

H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados.

2 Hipótesis

Para poder realizar el tratamiento de la encuesta y medir las preguntas de tesis presentadas en el apartado anterior, a continuación se plantean un conjunto de hipótesis, en función del objetivo que plantea cada pregunta.

En este trabajo de investigación se plantean cuatro hipótesis relacionadas con las preguntas de investigación:

H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados.

Con esta hipótesis queremos responder a la pregunta de si las organizaciones que disponen de un sistema como el que se propone en este trabajo de investigación perciben mejores resultados, y qué correlación existe entre cada uno de los componentes o dimensiones del modelo propuesto (medir, analizar, visualizar, informar y publicar) y la percepción de los resultados.

H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas.

Interesa también en la investigación conocer en qué grado, el hecho de disponer de un modelo de sistema de información similar al propuesto permite al directivo responsable responder más rápidamente a las peticiones de información sobre su ámbito de responsabilidad.

H3 Las organizaciones que disponen de datos de:

- Planificación y seguimiento de la estrategia
- Personas (nómina, altas, bajas, formación, etc.)
- Economía y finanzas (contabilidad, costes, etc.)
- Gestión del conocimiento y TI
- Compras / Proveedores
- Clientes / CRM
- Procesos / Productos / Servicios

Perciben mejores resultados.

En esta hipótesis nos planteamos si la adopción de un modelo de excelencia, que nos determina que debemos recoger datos sobre determinados criterios fundamentales, llevará a la organización a percibir mejores resultados.

H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados

En esta hipótesis planteamos directamente la relación entre los resultados de una organización y el hecho de que la organización haya adoptado un sistema de excelencia o calidad.

3 Posicionamiento epistemológico

Easterby-Smith et al. (2002:27) resaltan las tres razones por las que consideran que comprender los aspectos epistemológicos y filosóficos de una investigación resulta beneficioso:

- Porque ayuda a clarificar el diseño de una investigación (qué datos recoger, de donde, y como pueden ser interpretados para proporcionar respuesta a las preguntas inicialmente planteada).
- Porque ayuda a comprender las limitaciones prácticas de determinados enfoques, y ver qué diseños funcionarían y cuáles no.
- Porque puede ayudar al investigador a identificar, y hasta crear, diseños que pueden estar fuera de su experiencia pasada, así como sugerir la forma en que adaptar diseños a otras investigaciones.

La adopción del enfoque epistemológico y metodológico permite obtener guías de investigación más concretamente, en el momento de determinar las cuestiones operativas. Dado que el tipo de preguntas de este trabajo de investigación, en el que se considera posible y esencial para el investigador adoptar una postura distante y no interactiva, es decir, que el sujeto y el objeto son entidades independientes, las posición epistemológica ha sido la de adoptar el paradigma positivista.

El positivismo es una corriente de pensamiento que se puede situar entre los planteamientos de Auguste Compte, que sólo admite como válidos los conocimientos que proceden de las ciencias empíricas. Tan importante es la influencia de Compte que algunos autores hacen coincidir el inicio del positivismo con la publicación de su obra "Curso de filosofía positiva" (1842). No obstante, otros autores sugieren que algunos conceptos positivistas se remontan al filósofo británico David Hume (1711-1776) y al filósofo francés Saint-Simon (1760-1825).

Para Kolakowski (1988), el positivismo es un conjunto de reglamentaciones que rigen el saber humano y que muestra tendencia a reservar el nombre de "ciencia" a las operaciones observables en la evolución de las ciencias modernas de la naturaleza. Durante su historia, el positivismo ha dirigido en particular sus críticas contra los desarrollos metafísicos de toda clase, por tanto, contra las reflexiones que no pueden fundamentar totalmente sus resultados sobre datos empíricos, o que formulan sus juicios de forma que los datos empíricos no puedan refutarlos nunca.

De acuerdo con Dobles et al. (1998), la teoría de la ciencia que sostiene el positivismo se caracteriza por afirmar que el único conocimiento cierto es aquel que es producido por la ciencia, particularmente con el uso de sus método. En consecuencia, el positivismo asume que solo las ciencias empíricas son fuente aceptable de conocimiento.

Otra de las características relevantes del positivismo tiene que ver con su posición epistemológica central. En efecto, el positivismo supone que la realidad está dada y que puede ser conocida de forma absoluta por el sujeto cognoscente, y que por tanto, de lo único que nos hemos de preocupar, indican Dobles et al. (1998), es de encontrar el método adecuado y valido para "descubrir" esta realidad. En particular, asume la existencia de un método específico para conocer esta realidad y propone el uso de este método como garantía de verdad y legitimidad para el conocimiento. Por tanto, la ciencia positivista se basa en el supuesto de que el sujeto tiene una posibilidad absoluta de conocer la realidad mediante un método específico.

Otro aspecto importante del positivismo es el supuesto de que tanto las ciencias naturales como las sociales pueden hacer uso del mismo método para desarrollar la investigación. De acuerdo con Tejedor (1986), citado por Dobles et al. (1998), los científicos positivistas suponen que se puede obtener conocimiento objetivo del estudio del mundo natural y social. Para ellos, las ciencias naturales y las ciencias sociales utilizan una metodología básica similar para utilizar la misma lógica y procedimientos de investigación similares. Desde esta perspectiva, se considera que el método científico es único y el mismo en todos los campos del saber, para lo cual la unidad de todas las ciencias se fundamenta en el método: lo que hace la ciencia es el método con el que se tratan los "hechos".

A consecuencia de lo que se ha dicho anteriormente, se puede indicar que el positivismo busca los hechos o causas de los fenómenos sociales con independencia de los estados subjetivos de los individuos. De acuerdo con Dobles et al. (1998) el positivismo se caracteriza por postular los aspectos siguientes:

a. El sujeto descubre el conocimiento
b. El sujeto tiene acceso a la realidad mediante los sentidos, la razón y los instrumentos que utilice.
c. El conocimiento válido es el científico.
d. Existe una realidad accesible a cualquier sujeto mediante la experiencia. El positivismo supone la existencia independiente de la realidad con respecto al ser humano que la conoce.
e. Lo que es dado a los sentidos puede ser considerado como real.
f. La verdad es una correspondencia entre lo que el ser humano conoce y la realidad que descubre.
g. El método de la ciencia es lo único válido.
h. El método de la ciencia es descriptivo. Ello significa, según Abagnaro, que la ciencia describe los hechos y muestra las relaciones constantes entre los hechos, que se expresan mediante leyes y permiten la previsión de los hechos.
i. Sujeto y objeto de conocimiento son independientes: se plantea como un principio la neutralidad valorativa. Es decir, que el investigador se ubique en una posición neutral respecto de las consecuencias de sus investigaciones.

Al tomar esta posición epistemológica, se derivan unas consecuencias metodológicas que se desarrollan a continuación.

4 Elección del método

A continuación se exponen las diferentes fases que se han seguido para el diseño y realización del trabajo empírico, con el objetivo de recoger la información necesaria que permitiera responder a las preguntas planteadas, y al mismo tiempo, validar el modelo propuesto.

Fenómeno	Realidad
Definición conceptual	Conceptualización
Dimensionalización Indicadores, índice y tipologías	Operacionalización
Fenómeno construido	Realidad, observación sistematizada y estructurada

Ilustración 160. Elección del método. Fuente: Domínguez Amorós et al. (2000:93)

El contraste de la existencia de un sistema de información parecido al modelo SIEGO planteado en este trabajo de investigación, se ha planteado mediante seis dimensiones, que son:

- Dimensión Datos
- Dimensión Analizar
- Dimensión Visualizar
- Dimensión Informar
- Dimensión Publicar

- Dimensión Resultados

Las dimensiones del modelo SIEGO se han descrito en los correspondientes apartados de este documento, en el Capítulo III.

Adoptar un correcto método de investigación debe de permitir recoger la información adecuada para delimitar el grado en que las empresas están implantando prácticas de gestión de la información y ver la relación de estas prácticas con la percepción de resultados. Para dar cumplimiento al requisito de que las dimensiones puedan ser medidas empíricamente, debemos usar una técnica de recogida de información. Para conseguir este objetivo, dada la posición epistemológica tomada y explicada anteriormente, el método de investigación que se considera más conveniente es el estudio de campo, el modelo de observación utilizado al obtención de información a través de encuestas por medio de la utilización de un cuestionario como única fuente de recogida de información.

García Ferrando (1992:141) expresa: "una encuesta es una investigación realizada sobre una muestra de individuos, representativa de un colectivo más amplio, que se lleva a término en el contexto de la vida cotidiana, utilizando procedimientos estandarizados de interrogación, con la finalidad de obtener medidas cualitativas de gran variedad de características objetivas y subjetivas de la población".

El cuestionario es un instrumento ampliamente utilizado, y según Cea (1999), es la estrategia de recogida de datos más conocida y practicada por los investigadores de ciencias sociales. El cuestionario da cumplimiento a la función de enlace entre los objetivos de la investigación y la realidad de la población de las organizaciones observadas. Sierra (1998:306-307) comenta que: "Las condiciones fundamentales que debe de reunir un cuestionario, dependen de la investigación y de la población. Se pueden sintetizar, de una parte, en traducir los objetivos de la investigación en preguntas concretas sobre esta realidad, y, de otra parte, en ser capaz de suscitar en los encuestados respuestas sinceras y claras a las preguntas, que puedan luego ser tratadas científicamente, es decir, clasificadas y analizadas". Por tanto, "la finalidad del cuestionario es obtener de forma sistemática y ordenada, la información de la población investigada sobre las variables objeto de la investigación".

5 Unidad de análisis y población

Una vez decidida la técnica de recolección de datos, es necesario tomar decisiones correspondientes a la unidad de análisis e identificar la población a estudiar.

La descripción de las unidades de análisis determina cuales son las realidades que se desea observar y qué datos debemos de obtener de las mismas para desarrollar la investigación empírica: Sierra, (1994:351) citado por Escrig (2001:185). La delimitación de las unidades de análisis debe de estar de acuerdo con los objetivos y las preguntas derivadas del modelo propuesto. Los objetivos planteados en esta investigación requieren analizar en qué medida las

organizaciones recogen datos de sus actividades, los analizan, aplican componentes de visualización, preparan informes y los publican, y en qué medida ello influye en la percepción de los resultados de la organización.

Partiendo de este propósito, el ámbito de estudio debe proporcionar una muestra en la que sea posible encontrar organizaciones con un grado diferente de aplicación de este tipo de sistemas.

La identificación de la población se ha hecho considerando los siguientes razonamientos:

a. La población debería de estar compuesta por organizaciones que en la actualidad dispongan de algún tipo de sistema de información. Ello se justifica porque lo que queremos determinar es cómo son estos sistemas y no cuántas organizaciones disponen de ellos. No fijar esta limitación comporta iniciar las consultas a una parte de las organizaciones que no disponen de sistemas de información, y por tanto, no se podría continuar la consulta. Al no poder determinar la existencia o no de sistemas de información en las bases de datos de organizaciones, se decidió optar por fijar un tamaño mínimo de organización y de este modo garantizar la existencia de sistemas de datos. Se fijó el tamaño en más de 50 trabajadores.
b. Debemos garantizar que la encuesta llegue a cierto número de organizaciones que dispongan de sellos de calidad o excelencia, para de este modo contrastar que este tipo de organizaciones perciben mejores resultados. Por ello se contactó el Club de Excelencia de Gestión, licenciatario del modelo EFQM para España y se pidió su colaboración para la difusión del estudio.

Partiendo de estos razonamientos, las bases de datos utilizadas fueron:

SABI: Base de Datos que incluye análisis financieros de más de 1.300.000 empresas españolas y 500.000 empresas portuguesas con un histórico de cuentas anuales de hasta 10 años. Producto de Bureau van Dijk. La información se obtiene de distintas fuentes oficiales: Registros Mercantiles, BORME, prensa, etc. Se seleccionaron las empresas de más de 50 trabajadores y que disponían de dirección de correo electrónico. El fichero manejado en la encuesta contenía 5.000 empresas.

Alimarket: Incluye datos de más de 45.000 empresas españolas y sus directivos. Se seleccionaron las empresas con más de 50 trabajadores y que disponían de correo electrónico. El fichero manejado en la encuesta contenía 2.500 empresas.

Club Excelencia de Gestión: El Club es el licenciatario del modelo EFQM para España. Se pidió colaboración al Club para garantizar la inclusión en la muestra de empresas que tuvieran implantado el modelo EFQM. El Club hizo difusión del estudio en la comunidad autónoma de Catalunya, aproximadamente 40 organizaciones, de las cuales respondieron 28.

6 Muestra

Tal y como se ha comentado, con la finalidad de poder enfocar la elección de la muestra que se utilizará como base del trabajo de investigación, se entiende conveniente escoger organizaciones de más de 50 trabajadores y que dispongan de dirección de correo electrónico.

El procedimiento utilizado fue el siguiente:

1. Envío de un mensaje de correo electrónico dirigido a un representante de las organizaciones de la muestra, en el cual se explicaba el motivo por el que se solicitaba su colaboración. En este mensaje se les proporcionaba el link a la página web del cuestionario, que se podía realizar online desde cualquier navegador web, tanto en pc, como en tableta como en Smartphone.

   A continuación se muestra el texto del correo electrónico:

   Solicitamos colaboración para la realización del ***trabajo de investigación*** *denominado "UN MODELO DE SISTEMA DE INFORMACIÓN PARA LA EFICIENCIA EN LA GESTIÓN DE LAS ORGANIZACIONES" dentro del marco de los estudios doctorales desarrollados en la Universidad de Oviedo.*

   La colaboración se lleva a cabo a través de la participación en una encuesta on-line de no más de 10 min. de duración. Las organizaciones que completen la encuesta recibirán un resumen de los resultados de la investigación una vez finalizados.

   Acceda aquí a la encuesta. Muchas gracias por su colaboración

2. Respuesta al cuestionario por parte de los responsables de las organizaciones.
3. Con posterioridad a la publicación de este trabajo de investigación, se enviará a los participantes un resumen con los resultados de la investigación y las conclusiones de la investigadora.

6.1 Cálculo del tamaño muestral

Con el fin de determinar si existe o no una relación significativa entre dos variables numéricas X e Y, se recoge una muestra de individuos en donde de cada uno de ellos se determina el valor que toma cada una de las dos variables. A continuación se muestra cómo calcular el tamaño de muestra necesario para contrastar la hipótesis de que el correspondiente coeficiente de correlación sea significativamente diferente de 0.

La distribución muestral del coeficiente de Pearson no es normal, pero bajo la suposición de que las dos variables de estudio presentan una distribución gaussiana, el coeficiente de Pearson puede transformarse para conseguir un valor de z que sigue una distribución normal.

Para la estimación del tamaño muestral para llevar a cabo el estudio que analiza las diferentes hipótesis respecto la relación entre la percepción de resultados y el hecho de publicar, informar, analizar y visualizar, se utilizó la fórmula detallada a continuación, considerando un nivel de confianza del 95%, una potencia del 80% y una estimación del coeficiente de correlación de 0,3.

$$n = \left(\frac{z_{1-\frac{\alpha}{2}} + z_{1-\beta}}{\frac{1}{2} \cdot ln(\frac{1+r}{1-r})} \right)^2 + 3$$

Ecuación 3. Tamaño muestral

Substituyendo los valores antes indicados en la fórmula, obtenemos que necesitamos considerar como mínimo un tamaño de 85:

$$n = \left(\frac{1,96 + 0{,}842}{\frac{1}{2} \cdot ln(\frac{1+0{,}3}{1-0{,}3})} \right)^2 + 3 = 84{,}95$$

Ecuación 4. Tamaño muestral, con valores

Podríamos considerar un porcentaje de posibles pérdidas o no respuesta del 15%, con lo que necesitaríamos 100 datos:

$$n' = n \left(\frac{1}{1-p} \right)$$

Ecuación 5. Proporción de pérdidas.

Siendo n el tamaño muestral sin pérdidas ya calculado (0,85), y p la proporción esperada de pérdidas (0,15).

7 Cuestionario y preguntas

Para poner en práctica el modelo propuesto en este trabajo de investigación, se ha diseñado una encuesta que integra una escala que permite medir los diferentes conceptos. Siguiendo el mismo esquema utilizado el Modelo SIEGO, la encuesta se estructura en diferentes preguntas que están relacionadas con las dimensiones que se han definido en relación a las diferentes partes del modelo.

En el inicio del cuestionario, se pregunta sobre los datos básicos que permitirán segmentar la muestra. La recogida de datos pretende obtener información sobre la organización

encuestada, en lo que respecta a los sectores de actividad, el territorio de la organización, su tamaño y sobre la formación y tamaño del equipo del directivo.

1 Preguntas de segmentación

1.1 Soy hombre / mujer
1.2 Mi edad es:
1.3 Mi nivel de estudios es:
1.4 Gestiono, directa o indirectamente, un equipo de…
1.5 Mi organización pertenece al sector económico
1.6 El tamaño de mi organización es:
1.7 La sede en España de mi organización se encuentra en la comunidad de:
1.8 Mi organización dispone de los siguientes sellos o certificados:

A continuación se contrasta el modelo con preguntas agrupadas para cada una de las dimensiones del modelo presentado:

2 SIEGO Dimensión Datos

SIEGO Dimensión Datos - Recoger datos

2.1 En mi organización recogemos datos de (Sistema de datos existente):

Planificación y seguimiento de la estrategia
Personas (nómina, altas, bajas, formación, etc.)
Economía y finanzas (contabilidad, costes, etc.)
Gestión del conocimiento y TI
Compras / Proveedores
Clientes / CRM
Procesos / Productos / Servicios

2.2 En mi organización disponemos de bases de datos (Data Warehouse) con los datos organizados por tipos y materias (Data Marts).

SIEGO - Dimensión Datos - Procesos /Productos / Servicios

9.1 Mi organización dispone de datos sobre el tiempo medio de servicio a un cliente, para cada producto vendido o servicio prestado.

SIEGO Dimensión Datos - Clientes / CRM

8.1 Recibo o tengo acceso a información sobre cuáles han sido las ventas de productos o servicios en el período
8.2 Sé en todo momento cual es el nivel de satisfacción de mis clientes

SIEGO Dimensión Datos - Compras / Proveedores

2.1 En mi organización recogemos datos de (Sistema de datos existente):

SIEGO Dimensión Datos - Gestión del conocimiento y TI

2.1 En mi organización recogemos datos de (Sistema de datos existente):

SIEGO Dimensión Datos - Economía y Finanzas

2.1 En mi organización recogemos datos de (Sistema de datos existente):

SIEGO Dimensión Datos - Personas

10.1 Tengo acceso a los datos sobre la composición de mi equipo: número de personas, género, edad, nivel de estudios…
10.2 Mi organización me ofrece datos sobre el potencial de las personas de mi equipo, su situación en el plan de carrera o su plan de formación

SIEGO Dimensión Datos - Planificación y seguimiento de la estrategia

7.1 Conozco cuales son mis objetivos estratégicos (o líneas, o áreas, etc.) y cuál es su estado de ejecución

7.2 Dispongo de un cronograma de los proyectos estratégicos y su estado de ejecución

3 SIEGO Dimensión Análisis

3.1 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas univariables

3.2 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas bivariables

3.3 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas multivariables

4 SIEGO Dimensión Visualización

4.1 En mi organización vemos datos presentados con este aspecto: Tablas Excel

4.2 En mi organización vemos datos presentados con este aspecto: Gráfico de líneas

4.3 En mi organización vemos datos presentados con este aspecto: Gráfico de columnas

4.4 En mi organización vemos datos presentados con este aspecto: Treemap

4.5 En mi organización vemos datos presentados con este aspecto: Bubble

5 SIEGO Dimensión Informes

5.1 En mi organización se utiliza un sistema de informes para la presentación periódica de datos en papel o bien en formato digital (cuadros de mando, reporting, mapas, etc.)

5.2 En mi organización disponemos de herramientas de análisis de datos como por ejemplo tablas dinámicas de Excel o cubos multidimensionales para análisis (OLAP).

6 SIEGO Dimensión Publicar

6.1 En mi organización existe un espacio común (portal web, intranet, etc.) en el que se publican los informes.

En esta parte del cuestionario, también se valida la existencia de un sistema de datos basado en los criterios definidos en los modelos de excelencia (Estrategia, Clientes, Productos/Servicios/Procesos y Personas).

Tal y como se ha comentado en la motivación de este trabajo de investigación, se pretende contrastar que las organizaciones toman datos de los criterios que los modelos de excelencia han determinado que son los más relevantes. Estos criterios son:

- Recoger datos en general (preguntas 2.1 y 2.2)
- Procesos /Productos / Servicios (preguntas 2.2 y 9.1)
- Clientes / CRM (pregunta (preguntas 2.2, 8.1 y 8.2)
- Compras / Proveedores (pregunta 2.2)
- Gestión del conocimiento y TI (pregunta 2.2)
- Economía y Finanzas (pregunta 2.2)
- Personas (preguntas2.2, 10.1 y 10.2)
- Planificación y seguimiento de la estrategia (preguntas 2.2, 7.1 y 7.2)

Además de los datos de validación del modelo y de los criterios de los modelos de gestión, también se ha incluido un grupo de preguntas para medir la percepción de resultados de la organización. Esta percepción de los resultados es la que se usará para contrastar las hipótesis:

H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados.

H3 Las organizaciones que disponen de datos de:

- Planificación y seguimiento de la estrategia
- Personas (nómina, altas, bajas, formación, etc.)
- Economía y finanzas (contabilidad, costes, etc.)
- Gestión del conocimiento y TI
- Compras / Proveedores
- Clientes / CRM
- Procesos / Productos / Servicios

Perciben mejores resultados.

H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados.

Las preguntas relacionadas con la Dimensión Resultados son las siguientes:

7 Dimensión Resultados

11.1 Mi percepción sobre el clima laboral en mi organización:

11.2 Mi percepción sobre los resultados de mi organización es:

13.1 En su opinión, ¿está de acuerdo en que la implantación de un sistema de datos estructurados, como el que hemos presentado en esta encuesta, mejoraría los resultados de su organización?

La hipótesis: H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas, pone en relación el tiempo de respuesta a las peticiones de información con el modelo propuesto SIEGO. Por ello, el cuestionario recoge una pregunta sobre el tiempo de respuesta:

8 Dimensión tiempo de respuesta

12.1 Puedo responder al 95% de preguntas sobre cualquier tema de mi ámbito de responsabilidad con datos concretos en…

Finalmente, una pregunta cualitativa recoge las opiniones de las personas que responden al cuestionario con respecto a la conveniencia o no de implantar un sistema de información como el que se propone en este trabajo, y de la que también se hará un análisis cualitativo.

9 Comentarios

14.1 Para finalizar, nos gustaría darle la oportunidad de dejar los comentarios que crea oportunos. Nos interesa mucho su opinión, experiencia o impresión sobre las barreras o las ventajas de la implantación de un sistema de análisis y presentación de datos en su organización.

7.1 Preguntas sobre resultados

En relación con una de las hipótesis planteadas en esta investigación, se quiere conocer la percepción de los resultados de la organización en dos ámbitos, resultados en general, y resultados en el ámbito de la gestión de personas (clima laboral). Para tener información en relación con estas hipótesis se plantean las siguientes preguntas:

11.1 Mi percepción sobre el clima laboral en mi organización:

11.2 Mi percepción sobre los resultados de mi organización es:

13.1 En su opinión, ¿está de acuerdo en que la implantación de un sistema de datos estructurados, como el que hemos presentado en esta encuesta, mejoraría los resultados de su organización?

Estas preguntas, no tienen la finalidad de obtener datos cuantitativos sobre la influencia del modelo presentado en los resultados organizacionales, sino de obtener una medida subjetiva que permita tener una primera aproximación para valorar los resultados de la organización.

Después de la revisión de los trabajos que han investigado la relación entre las prácticas de gestión y los resultados de las organizaciones y la dificultad que plantea conseguir información al respecto, podemos destacar Adam et al. (1997), Anderson y Sohal (1999), Terziosvski y Samsom (1999) o Zhang (2000), que indican que son dos las estrategias que se llevan a cabo usualmente:

La utilización de datos objetivos, derivados de:

- Fuentes secundarias: Una solución que plantea Lemak y Reed (1997) y Easton y Jarrell (1998) que utilizan dos medidas, unas basadas en el valor de la empresa en el mercado de valores y otras basadas en sus datos contables. Chapman et al. (1997) utilizan únicamente indicadores financieros para medir los resultados, como, por ejemplo, los beneficios obtenidos por los accionistas, la rentabilidad sobre activos o la productividad del trabajo.
- Percepciones de las propias organizaciones, por medio de respuestas a un cuestionario. Adam et al. (1997) y Lee et al. (1999) que miden los resultados diferenciando entre resultados de calidad, operaciones y financieros.

La utilización de datos subjetivos, basados en valoraciones emitidas por los directivos de las propias organizaciones. Powell (1995) y Leal (1997) utilizan escalas subjetivas para valorar los resultados financieros medidos en función de la rentabilidad, el crecimiento de las ventas o los

resultados financieros globales. En estas investigaciones se pide a los encuestados que valoren su grado de acuerdo o desacuerdo, en una escala tipo Likert, en relación a si los resultados de la organización han sido excelentes en los últimos años. Anderson y Sohal (1999) utilizan una escala Likert de 5 puntos donde los investigados valoran si el impacto de las prácticas sobre cada uno de los indicadores apuntados ha sido alto o bajo. Estos indicadores de resultados comprenden elementos relativos a la calidad de los productos, la flexibilidad de producción, ventas o cuota de mercado. Whang (2000), que sigue las aportaciones de Mann y Keohe (1994), utiliza las opiniones de la organización recogidas mediante un cuestionario y distingue entre resultados estratégicos (rentabilidad, cuota de mercado o ventas) y resultados operativos (relativos a tiempo del ciclo de los procesos, relaciones con proveedores, con clientes o con empleados). Ittner y Larcker (1997) o Terziovski y Samson (1999) utilizan una combinación de medidas objetivas basadas en datos financieros y otras subjetivas basadas en opiniones de los entrevistados sobre elementos como la satisfacción de los empleados.

La medida de los resultados en base exclusivamente a datos objetivos presenta destacadas insuficiencias:

a. Las organizaciones son reticentes a aportar datos cuantitativos de datos contables.
b. Los datos secundarios (Central de balances del Banco de España, Registro mercantil, etc.) no presentan el nivel de fiabilidad necesario para investigaciones precisas.
c. El registro contable presenta cierta falta de homogeneidad.
d. La medida de la rentabilidad por fracciones presenta el defecto de que los resultados pueden aumentarse reduciendo los denominadores o aumentando los numeradores. Así, puede darse el caso que se reduzca la rentabilidad de la inversión, aunque crezcan los beneficios después de impuestos, cuando los fondos propios aumentan más rápidamente.
e. El cálculo agregado esconde la causa de algunas variaciones.
f. Obtenemos datos de cierta antigüedad.
g. Sólo vemos datos contables, así pues, perdemos la información de los intangibles.
h. Este tipo de datos nos son aplicables en el caso de organizaciones del sector público.

Diferentes trabajos indican que la medida de resultados por medio de cuestionarios subjetivos ofrece evaluaciones equiparables a los realizados con medidas de tipo objetivo (Venkatraman y Ramanujam, 1986; Smith et al., 1989; Camisón, 1999a).

Asimismo, las conclusiones del trabajo de Martin (1997), indican que medidas objetivas como las medidas financieras son engañosas dado que no recogen todas las mejoras atribuibles a las iniciativas de gestión. Además, se han recogido apreciaciones de Chakravarthy (1986) para el cual los tradicionales indicadores económico-financieros de rentabilidad y ventas no son suficientes para explicar las diferencias entre empresas excelentes y las no excelentes, sugiriendo el uso de medidas que contemplen el grado de satisfacción de todos los grupos de interés de la organización.

Del mismo modo, importantes elementos decisivos para la competitividad, como son los intangibles o las competencias distintivas, (Prahalad y Hamel, 1990; Stalk et al., 1992; Hall, 1992, 1993) medidas claramente subjetivas, no se recogen en los sistemas de contabilidad, por lo cual la medida de resultados no puede ser reducida a medidas cuantitativas, si no que debería de incluir medidas de resultados cualitativos derivados de la posesión de estas competencias.

Kaplan y Norton (1997) propusieron la medida de resultados por medio de su balance scorecard, que identifica un conjunto equilibrado de medidas financieras y no financieras. En su trabajo consideraron cuatro grupos de interés: accionistas, clientes, organización y empleados. Otros trabajos que han considerado su enfoque de la medida orientada a la satisfacción de los grupos de interés son Freeman (1983) y Atkinson, Waterhouse y Wells (1997). Provost y Leddick (1993) escriben que la valoración de la salud y el rendimiento de una organización requieren de una familia de medidas del impacto de los cambios en todos los grupos de interés de la organización: clientes, empleados, finanzas, operaciones y entorno. El modelo EFQM de excelencia indica cuatro dimensiones de resultados: Clientes, Empleados, Sociedad y Resultados clave.

Considerando todas estas aportaciones, se ha estimado adecuado que para la obtención de la información sobre los resultados podemos realizar una medida de tipo subjetivo.

7.2 Herramienta utilizada

Se utilizó la herramienta online Typeform®, que permite la creación y gestión de encuestas digitales. La encuesta se manda en un link mediante correo electrónico y permite recoger las respuestas desde cualquier plataforma (portátil, pc, tableta o Smartphone).

Es una herramienta con una interfaz de uso y manejo muy sencilla que permite crear cuestionarios atractivos, interactivos y dinámicos.

Consigue hacer unas encuestas ágiles con un diseño muy profesional y adaptable a cualquier dispositivo móvil y/o táctil, que incrementan el índice de respuesta habitual en este tipo de cuestionarios.

A la encuesta creada mediante Typeform® se puede acceder a través de su dirección web:

https://begea.typeform.com/to/OfFgln

Las respuestas recogidas en el cuestionario quedan guardadas en un fichero Excel, que es el que se utilizó posteriormente para el análisis empírico.

7.3 Tipo de Escala y propiedades socio métricas asociadas

Previamente a la justificación de la elección de las diferentes escalas utilizadas, cabe argumentar la conveniencia del uso de escalas multi-ítem. En esta línea, Churchill (1979), defiende el uso de las escalas de un único ítem, indicando como inconvenientes los siguientes:

a. Los ítems individuales tienden a poseer bajas correlaciones con el atributo objeto de medición; al mismo tiempo, tienden a relacionarse con los otros atributos.
b. Los ítems individuales tienen un considerable error de medición.

Estas limitaciones hacen aconsejable el uso de escalas multi-ítem, porque:

a. La especificidad de los ítems permite el cálculo de los mismos.
b. Mediante la combinación de los ítems se pueden determinar claras distinciones entre los individuos.
c. La fiabilidad tiene a incrementarse y el error de medición decrece conforme el número de ítems se incrementa (Churchill, 1979).

Así pues, el cuestionario se evalúa con escalas multi-ítem tipo Likert de 0 a 10 puntos. Con ello se pretende medir las diferentes dimensiones a través de múltiples ítems.

A diferencia de las preguntas dicotómicas con respuesta sí/no, la escala de Likert nos permite medir actitudes y conocer el grado de conformidad del encuestado con cualquier afirmación que le propongamos.

Resulta especialmente útil emplearla en situaciones en las que queremos que la persona matice su opinión. En este sentido, las categorías de respuesta nos servirán para capturar la intensidad de la respuesta del encuestado hacia dicha afirmación.

Todas las preguntas del cuestionario se responden con una escala de 0 a 10, excepto las preguntas de segmentación que utilizan una lista de valores seleccionable:

1.1 Soy hombre / mujer
1.2 Mi edad es:
1.3 Mi nivel de estudios es:
1.4 Gestiono, directa o indirectamente, un equipo de…
1.5 Mi organización pertenece al sector económico
1.6 El tamaño de mi organización es:
1.7 La sede en España de mi organización se encuentra en la comunidad de:
1.8 Mi organización dispone de los siguientes sellos o certificados:

La escala de 0 a 10 se graduó con un texto para dar información adicional sobre el valor de la respuesta, que en todo caso se indicaba cerca de los valores 0, 5 y 10, como se muestra a continuación:

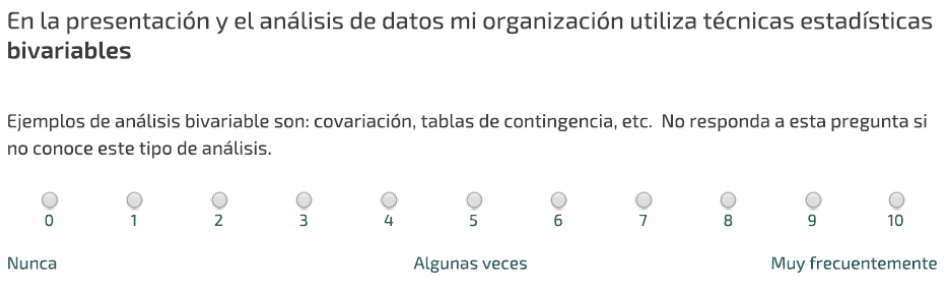


Ilustración 161. Imagen cuestionario, respuesta múltiple.

Al hacer el análisis cualitativo de las frecuencias, se ha sumado los valores de 0 a 2 para la primera categoría, de 3 a 7 para la segunda categoría y de 8 a 10 para la tercera.

Para que cualquier escala de medida sea rigurosa y académicamente adecuada a su propósito, esta debe de cumplir las siguientes propiedades socio métricas:

a. Fiabilidad. Está relacionada con los errores aleatorios, con la precisión de la medida.
b. Validez. Está relacionada con los errores sistemáticos, con la adecuación entre lo que se mide y lo que se quiere medir.

Una vez obtenidas las respuestas al cuestionario, se ha comprobado que las principales escalas utilizadas en la investigación cumplen las propiedades socio métricas, es decir son fiables y válidas.

De esta manera, se han evaluado las escalas de medida para asegurarnos que son útiles para la recogida de información de los constructos que se pretende medir. Asimismo, también se ha corroborado que esta información ha sido obtenida a partir del procedimiento más exacto y fiel posible a la realidad.

A continuación se evalúan estas escalas para comprobar si realmente cumplen las dos propiedades requeridas en este estudio: fiabilidad y validez.

7.4 Fiabilidad de las escalas seleccionadas en la encuesta

César (2001:163) afirma que la fiabilidad de una escala se refiere "al grado en que ésta dará resultados consistentes si se llevan a término mediciones repetidas. En otras palabras, es la proporción de varianza atribuible a la puntuación real de la variable, siendo un buen indicador de la cantidad de error de medida inherente a un modelo. Esto equivale a considerar que la estructura de las escalas creadas está correctamente diseñada". Hayes (1992:50) define la fiabilidad como "el grado en el cual las medidas están libres de desviaciones producidas por los errores causales". Acostumbra a ser bastante habitual interpretar la fiabilidad como el grado de consistencia interna de la escala, y la forma más frecuente de estimarla es a partir de la α de Cronbach (Cronbach, 1951). La principal ventaja de su uso es que permite determinar la fiabilidad habiendo aplicado las escalas únicamente sobre un grupo de individuos.

Se puede medir la fiabilidad de una escala de diferentes formas, que dependen del enfoque de la investigación y de la naturaleza de los datos. De entre los métodos más utilizados encontramos la α de Cronbach (Peterson, 1951). Otros métodos son el de Spearman-Brown (dividir por la mitad) para medir la consistencia interna (Sánchez y Sarabia, 1999), la prueba de test reset para medir la estabilidad (Malhortra, 1997) y el uso de cuestionarios alternativos para medir la equivalencia (Miquel et al., 1997).

En el caso de este trabajo, se procede a examinar la fiabilidad de la escala a través de la medida de la consistencia interna más utilizada como es el coeficiente α de Cronbach (Cronbach, 1951). Así, para analizar la fiabilidad de las escalas utilizadas se aplica el estadístico α de Cronbach a todas y cada una de ellas; este estadístico analiza la consistencia interna de la escala a través de la correlación mediana de un ítem con todos los otros ítems. De esta manera, un valor alto pone de manifiesto que la escala es fiable.

Tal y como se ha dicho, la α de Cronbach es el índice más utilizado para calcular la fiabilidad. Cabe comentar pero, que no existe unanimidad a la hora de determinar el valor a partir del cual la fiabilidad se considera adecuada o aceptable (Peterson, 1994). Los valores cercanos a 0 indican ausencia de consistencia interna y los valores cercanos a 1 indican consistencia interna entre los ítems de la escala. No obstante, el valor de 1 no es el más adecuado, ya que indica repetición o redundancia entre los ítems, cuando lo óptimo es que cada ítem tenga su propia aportación a la escala. Son aconsejables valores alfa comprendidos entre 0,75 y 0,89 (Nunnally, 1978).

Al mismo tiempo, Nurosis (1993) establece que deben eliminarse los ítems que presenten correlación ítem total inferior a 0,3.

Atendiendo a estas recomendaciones, en este trabajo de investigación se eliminaran los ítems que presenten una correlación inferior a 0,3. También serán susceptibles de ser eliminados los que, con su supresión permitan obtener una α de Cronbach superior.

A continuación se realiza el análisis de la fiabilidad de las escalas utilizadas aplicando el estadístico α de Cronbach a todas las variables, agrupadas en su correspondiente dimensión.

Se analiza, en primer lugar, la fiabilidad de la Escala del Modelo SIEGO en su conjunto, que incorpora todas las variables sobre la información relacionada con las dimensiones del modelo: Recoger datos, analizarlos, aplicar componentes de visualización, preparar informes y publicarlos. En esta escala se evalúan así las 26 variables o elementos que participan en las diferentes dimensiones del modelo propuesto.

Estadísticos de fiabilidad para las variables del modelo SIEGO

Alfa de Cronbach	Alfa de Cronbach basada en los elementos tipificados	N de elementos
,858	,862	26

El estudio de la fiabilidad del cuestionario resulta con un α de Cronbach tan elevada que contrasta que el modelo está bien explicado con estas variables.

Una vez evaluada la fiabilidad del modelo en su conjunto, se procede a continuación a evaluar la fiabilidad de cada una de las variables.

Estadísticos de fiabilidad de las variables del modelo

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si se elimina el elemento
planisegui	116,90	890,393	,302	,688	,858
personas	116,67	897,333	,102	,696	,859
ecoyfinanzas	116,65	897,638	,099	,653	,859
gestionconoTI	117,04	887,275	,380	,609	,857
compras	116,85	892,085	,253	,705	,858
clientes	116,85	890,638	,308	,617	,858
procesos	116,85	893,531	,197	,593	,858
datawarehouse	111,54	794,168	,667	,895	,844
univariables	112,67	785,844	,565	,801	,847
bivariables	114,46	787,062	,602	,970	,845
multivariables	114,60	785,351	,583	,960	,846
vistablas	109,81	871,943	,184	,643	,859
vislineas	110,19	862,156	,270	,934	,856
visbarras	110,46	858,551	,247	,844	,857
vistreemap	114,46	852,296	,192	,729	,862
visburbuja	113,48	816,468	,395	,761	,854
informes	110,02	825,468	,553	,636	,849
herramientasanalisis	111,77	783,670	,561	,846	,847
espaciopublicar	111,35	765,042	,591	,721	,846
objetivosestrategia	110,54	797,615	,604	,906	,846
cronogramaproyectos	111,65	794,617	,651	,927	,844
venta	109,71	835,020	,364	,788	,854
nivelsatisfaccion	111,44	811,528	,549	,909	,848
tiempomedio	111,33	812,312	,443	,771	,852
datosequipo	110,27	823,946	,388	,867	,854
potencialequipo	112,54	818,764	,412	,854	,853

Como se observa en el cuadro superior, ningún elemento incrementa la fiabilidad substancialmente si se elimina como indica la α de Cronbach si se elimina el elemento, por lo que se mantienen todas las variables en el análisis empírico posterior.

Se analiza así mismo la fiabilidad de la Dimensión Datos: En esta escala se incluyen la parte de las variables del modelo SIEGO que se refieren a la Dimensión Datos. Este subconjunto de 15 variables son las que se analizan a continuación:

Estadísticos de fiabilidad

Alfa de Cronbach	Alfa de Cronbach basada en los elementos tipificados	N de elementos
,815	,823	15

Como se ha comentado anteriormente, se contrasta que una α de Cronbach que en este caso es superior a 0,8 y por tanto:

Como criterio general, George y Mallery (2003, p. 231) sugieren las recomendaciones siguientes para evaluar los coeficientes de alfa de Cronbach: Coeficiente alfa >0,8 es bueno.

Loo (2001, p. 223): el valor de consistencia que se considera adecuado es de 0,8 o más.
Gliem & Gliem (2003): un valor de alfa de 0,8 es probablemente una meta razonable.

A continuación se procede a valorar la fiabilidad de la Dimensión si se eliminara alguna de las quince variables que constituyen la Dimensión Datos:

Estadísticos total-elemento

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si se elimina el elemento
planisegui	57,69	251,432	,346	,403	,814
personas	57,49	255,929	,065	,305	,818
ecoyfinanzas	57,49	255,226	,153	,399	,818
gestionconoTI	57,81	253,316	,190	,345	,817
compras	57,64	252,855	,266	,538	,816
clientes	57,67	251,523	,353	,543	,814
procesos	57,61	253,538	,226	,323	,816
datawarehouse	52,15	210,235	,534	,494	,796
objetivosestrategia	51,11	207,718	,565	,634	,793
cronogramaproyectos	52,23	202,853	,624	,668	,787
venta	50,71	208,264	,523	,488	,797
nivelsatisfaccion	52,08	197,912	,740	,663	,777
tiempomedio	52,15	200,559	,560	,576	,794
datosequipo	51,01	202,797	,543	,557	,796
potencialequipo	53,13	192,063	,657	,648	,784

Como puede observarse en la tabla, no se producen valores α inferiores a 0,75, por lo que no se descarta ninguno de los elementos de la escala para el posterior análisis.

7.5 Validez

Una vez determinada la fiabilidad de las escalas, se pasa a evaluar la validez de las mismas.

Una escala es válida cuando realmente mide aquello que el investigador pretende medir, es decir, el constructo o variable de estudio. En otras palabras, la medida es válida cuando las diferencias entre los valores observados derivados de la escala reflejan diferencias reales alrededor de la variable objeto de la medida, y no otros factores (Churchill, 1979). Se pretende por tanto, que la medida sea libre de error sistemático y de error aleatorio.

Una escala tiene validez de contenido si existe acuerdo general respecto a que por medio de los indicadores considerados se han medido todos los aspectos que comprende el concepto. Así, depende de la medida en que el investigador haya creado los indicadores adecuados para cubrir todo el dominio del concepto (Nunnally, 1979; De Vellis, 1991:43). Entonces, la validez de contenido no se mide numéricamente, si no subjetivamente por los investigadores (Bollen, 1989:185).

A diferencia de la fiabilidad, la validez no puede ser estimada de manera directa, si no que únicamente puede ser inferida:

- De la forma que ha sido creada
- De su relación con las medidas de otras variables
- De su habilidad para predecir cuestiones específicas

Existen pues tres tipos básicos de validez que se corresponden a estas tres formas de determinarla:

- Validez de contenido
- Validez de constructo, concepto, significado o validez interna
- Validez relacionada con criterios

Comúnmente se utiliza una sola forma de validación para un instrumento (Rodríguez, 2009).

A continuación se determina la validez de contenido para las escalas utilizadas en este trabajo de investigación, en concreto la validez del constructo. La validez del constructo trata de reflejar la existencia de una relación teórica entre la variable objeto de medida y las otras variables (Bollen, 1989). Una forma de operativizar la validez del constructo se puede cuantificar a través de (Sánchez y Sanabria, 1999):

- La matriz multiconcepto-multimétodo
- El análisis de correlaciones
- El modelo causal de Bagozzi
- El análisis factorial

Para comprobar la validez de las escalas utilizadas en este trabajo de investigación optaremos por la metodología de McColl-Kennedy et al. (1999), que estudia si el coeficiente de correlación entre una determinada variable y la escala se comporta según lo que la literatura postula. En este caso se pretende comprobar si las escalas de medida de las dimensiones propuestas se relacionan positivamente con la mejora de los resultados expresados en el cuestionario.

A continuación se presenta el coeficiente de correlación de Pearson y la significación del test asociado, utilizados para estudiar la correlación entre la percepción de resultados y las dimensiones propuestas en el modelo.

	Coef. Correlación r	p-valor
Dim Datos	0,548	<0,001
DimAnálisis	0,295	<0,01
Dim Visualización	0,117	0,24
Dim Informar	0,211	0,03
Dim Publicar	0,239	0,02

Únicamente la dimensión visualización no alcanza significación estadística por lo cual habrá que prestar atención a los resultados que impliquen a esta variable.

2. Análisis de los resultados

1 Datos obtenidos 255

2 Metodología 256

2.1 Tabla descriptiva de variables, preguntas y dimensiones 257

2.2 Fiabilidad 259

2.3 Segmentación 260

2.4 Organizaciones con sellos de excelencia y/o certificados de calidad 262

3 Resultados de las dimensiones del modelo propuesto SIEGO: 263

3.1 Dimensión Datos 263

3.1.1 Utilización de bases de datos (Data Warehouse) con los datos organizados por tipos y materias (Data Marts). 263

3.1.2 Datos del Sistema de Gestión 264

3.2 Datos sobre Estrategia 265

3.2.1 Objetivos estratégicos (o líneas, o áreas, etc.) y estado de ejecución. 265

3.2.2 Datos sobre Clientes y Ventas 267

3.2.3 Nivel de satisfacción de los clientes. 268

3.2.4 Datos sobre Procesos / Productos / Servicios 269

3.3 Datos sobre Personas 270

3.3.1 Datos sobre la composición del equipo (número de personas, género, edad, nivel de estudios...) 270

3.3.2 Datos sobre el potencial de las personas, plan de carrera y plan de formación. 271

3.4 Dimensión Análisis 272

3.4.1 Técnicas estadísticas univariables 273

3.4.2 Técnicas estadísticas bivariables 274

3.4.3 Técnicas estadísticas multivariables 274

3.5 Dimensión Visualización 276

3.5.1 Utilización de tablas de datos 276

3.5.2 Gráficos de líneas 277

3.5.3 Uso de gráficos de columnas 278

3.5.4 Uso de gráficos treemap 278

3.5.5 Gráficos bubble o de burbuja 279

3.6 Dimensión Informes 281

3.6.1 Utilización de un sistema de informes para la presentación periódica de datos en papel o bien en formato digital (cuadros de mando, reporting, mapas, etc.). 281

3.6.2 Herramientas de análisis de datos como por ejemplo tablas dinámicas de Excel o cubos multidimensionales (OLAP). 282

3.7 Dimensión Publicación 283

3.7.1 Utilización de un espacio común (portal web, intranet, etc.) en el que se publican los informes. 283

3.8 Dimensión Resultados 284

3.8.1 Percepción sobre el clima laboral de la organización 284
3.8.2 Percepción sobre los resultados de la organización 285
3.8.3 Un sistema de información como el propuesto, SIEGO, mejoraría los resultados de la organización. 286
3.9 Dimensión Tiempo de respuesta 287
3.10 Comentarios de los encuestados 288
4 Análisis de las relaciones entre las dimensiones 289
4.1 Descriptivo por dimensiones del modelo propuesto y para resultados 290
4.2 Análisis de correlaciones entre las dimensiones 290
5 Comprobación de las hipótesis 292
5.1 H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados 292
5.2 H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas. 295
5.3 H3 Las organizaciones que disponen de datos perciben mejores resultados 296
5.4 H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados 298
6 Análisis multivariante: modelo de regresión 299
Capítulo V. Conclusiones 301

1 Datos obtenidos

Una vez completada la encuesta se procede al análisis de los datos obtenidos con la finalidad de comprobar las hipótesis planteadas en este trabajo de investigación. Puede consultarse el cuestionario en el Anexo I.

La encuesta realizada se compone de un total de 33 preguntas, distribuidas en nueve bloques:

Bloque	Preguntas	nº
1. Preguntas de segmentación	1.1 a 1.8	8
2. SIEGO Dimensión Datos	2.1 y 2.2 y de la 7.1 a la 10.2	9
3. SIEGO Dimensión Análisis	3.1 a 3.3	3
4. SIEGO Dimensión Visualización	4.1 a 4.5	5
5. SIEGO Dimensión Informes	5.1 y 5.2	2
6. SIEGO Dimensión Publicar	6.1	1
7. Percepción de resultados de la organización	11.1, 11.2 y 13.1	3
8. Tiempo de Respuesta	12.1	1
9. Comentarios	14.1	1

A continuación se describen las preguntas incluidas en cada bloque.

Las preguntas iniciales, preguntas 1.1 a 1.8 incluidas, corresponden al bloque de **segmentación**, y nos permitirán conocer la composición de la muestra. También nos permitirán establecer relaciones de segmentación de datos con las variables propias del modelo. La pregunta 1.8 pide datos sobre los sistemas de calidad o sellos de excelencia de los que dispone la organización, información que se relaciona con la percepción de resultados en la hipótesis H4 en la que se plantea directamente la relación entre los resultados de una organización y el hecho de que la organización haya adoptado un sistema de excelencia o calidad.

Las preguntas de la 3.1 a la 6.1 se refieren a la existencia o no de un modelo de sistema de información similar al que se propone en este trabajo de investigación, SIEGO, en el cual las organizaciones recogen datos, analizan, aplican técnicas de visualización, preparan informes y los publican. Los bloques de preguntas son los siguientes:

- **Dimensión Análisis** - Aplicar técnicas de análisis (preguntas 3.1 a 3.3)
- **Dimensión Visualización** - Componentes de visualización de datos (preguntas 4.1 a 4.5)
- **Dimensión Informes** - Elaborar informes (preguntas 5.1 y 5.2)
- **Dimensión Publicar** - Publicar informes (pregunta 6.1)

Para el bloque de la **Dimensión Datos**, corresponden las preguntas 2.1 y 2.2 y de la 7.1 a la 10.2, en total 9 preguntas, que se refieren a la existencia de datos en las organizaciones como los que propone **Talwar** en su estudio, en el cual se identifican cuáles son los elementos o

criterios más utilizados para evaluar las organizaciones excelentes. Mayoritariamente los modelos coinciden en la evaluación de nueve criterios, de los cuales se ha preguntado en el cuestionario sobre:

- Recoger datos en general (preguntas 2.1 y 2.2)
- Procesos /Productos / Servicios (preguntas 2.2 y 9.1)
- Clientes / CRM (pregunta (preguntas 2.2, 8.1 y 8.2)
- Compras / Proveedores (pregunta 2.2)
- Gestión del conocimiento y TI (pregunta 2.2)
- Economía y Finanzas (pregunta 2.2)
- Personas (preguntas 2.2, 10.1 y 10.2)
- Planificación y seguimiento de la estrategia (preguntas 2.2, 7.1 y 7.2)

El bloque de **resultados** se cuestiona en las preguntas 11.1, 11.2 y 13.1, que inquieren sobre la percepción de resultados de la organización.

La pregunta 12 pide la percepción del usuario para determinar el **tiempo de respuesta** medio a las peticiones de información de cualquier tipo del ámbito de responsabilidad.

La pregunta 14 es una pregunta cualitativa sobre las ventajas de la implantación de un sistema de información como el modelo propuesto, SIEGO.

Se realiza un análisis estadístico de las respuestas obtenidas del cuestionario para las cinco dimensiones del modelo SIEGO así como para el resto de bloques. A continuación se muestra en primer lugar el análisis de cada pregunta del cuestionario, que incluye la tabla de frecuencias y la gráfica de la distribución de frecuencias. Asimismo, se comenta su comportamiento. En segundo lugar, el conjunto de respuestas de una misma dimensión son objeto de análisis bivariante y multivariante para analizar las posibles interrelaciones.

2 Metodología

Se ha realizado un análisis estadístico descriptivo de cada ítem del cuestionario, proporcionando medidas de posición y dispersión en el caso de variables cuantitativas y distribución de frecuencias para las variables cualitativas. La validez del cuestionario, y de cada dimensión, se evaluó a través del coeficiente Alfa de Cronbach.

El nivel de significación empleado en todo el análisis fue de 0,05. Cuanto menor sea el valor de p, menor será la probabilidad de que los resultados obtenidos se deban al azar y mayor evidencia habrá a favor de la hipótesis planteada. Si dicha probabilidad es menor que un valor de p fijado previamente (en este caso se toma $p<0,05$), la hipótesis nula se rechazará. Así, cuando el valor de p está por debajo de 0,05, se dirá que el resultado es *estadísticamente significativo* y será *no significativo* en cualquier otro caso. Se ha escogido este valor por ser el valor de referencia en la bibliografía consultada.

El análisis estadístico se efectuó mediante el programa R (R Development Core Team), versión 3.2.0. R: A language and environment for statistical computing [Manual de software informático]. Vienna, Austria. Disponible en http://www.r-project.org/ (ISBN 3-900051-07-0).

2.1 Tabla descriptiva de variables, preguntas y dimensiones

A continuación se muestran en una tabla las variables utilizadas en el análisis estadístico que se ha realizado, en relación a los bloques y dimensiones establecidos en el cuestionario:

Bloque	Pregunta	Variable
1	Preguntas de segmentación	1.1 a 1.8
	1.1 Soy hombre / mujer	Sexo
	1.2 Mi edad es:	Edad
	1.3 Mi nivel de estudios es:	Nivelestudio
	1.4 Gestiono, directa o indirectamente, un equipo de...	Equipo
	1.5 Mi organización pertenece al sector económico	Sector
	1.6 El tamaño de mi organización es:	Sectorecono
	1.7 La sede en España de mi organización se encuentra en la comunidad de:	Tamano Sedeespana
	1.8 Mi organización dispone de los siguientes sellos o certificados:	sello300 sello400 sello500 iso9001 otrosello
2	SIEGO Dimensión Datos	2.1 y 2.2 y de la 7.1 a la 10.2
	SIEGO Dimensión Datos - Recoger datos	2.1 y 2.2
	2.1 En mi organización recogemos datos de (Sistema de datos existente):	
	Planificación y seguimiento de la estrategia	Planisegui Personas
	Personas (nómina, altas, bajas, formación, etc.)	EcoyFinanza
	Economía y finanzas (contabilidad, costes, etc.)	Gestioncono Compras
	Gestión del conocimiento y TI	Clientes
	Compras / Proveedores	Procesos
	Clientes / CRM	Datawarehouse
	Procesos / Productos / Servicios	
	2.2 En mi organización disponemos de bases de datos (Data Warehouse) con los datos organizados por tipos y materias (Data Marts).	
	SIEGO - Dimensión Datos - Procesos /Productos / Servicios	2.1 y 9.1
	9.1 Mi organización dispone de datos sobre el tiempo medio de servicio a un cliente, para cada producto vendido o servicio prestado.	Tiempomedio
	SIEGO Dimensión Datos - Clientes / CRM	2.1, 8.1 y 8.2
	8.1 Recibo o tengo acceso a información sobre cuáles han sido las ventas de productos o servicios en el período	Venta

	8.2 Sé en todo momento cual es el nivel de satisfacción de mis clientes	Nivelsatisfaccion
	SIEGO Dimensión Datos - Compras / Proveedores	2.1
	SIEGO Dimensión Datos - Gestión del conocimiento y TI	2.1 GestionconoTI
	SIEGO Dimensión Datos - Economía y Finanzas	2.1 EcoyFinanzas
	SIEGO Dimensión Datos - Personas	2.1, 10.1 y 10.2
	10.1 Tengo acceso a los datos sobre la composición de mi equipo: número de personas, género, edad, nivel de estudios…	Datosequipo
	10.2 Mi organización me ofrece datos sobre el potencial de las personas de mi equipo, su situación en el plan de carrera o su plan de formación	Potencialequipo
	SIEGO Dimensión Datos - Planificación y seguimiento de la estrategia	2.1, 7.1 y 7.2
	7.1 Conozco cuales son mis objetivos estratégicos (o líneas, o áreas, etc.) y cuál es su estado de ejecución	Objectivosestrategia
	7.2 Dispongo de un cronograma de los proyectos estratégicos y su estado de ejecución	Cronogramaproyectos
3	SIEGO Dimensión Análisis	3.1 a 3.3
	3.1 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas univariables	Univariables
	3.2 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas bivariables	Bivariables
	3.3 En la presentación y el análisis de datos mi organización utiliza técnicas estadísticas multivariables	Multivariables
4	SIEGO Dimensión Visualización	4.1 a 4.5
	4.1 En mi organización vemos datos presentados con este aspecto: Tablas Excel	Vistablas
	4.2 En mi organización vemos datos presentados con este aspecto: Gráfico de líneas	Vislineas
	4.3 En mi organización vemos datos presentados con este aspecto: Gráfico de columnas	Visbarras
	4.4 En mi organización vemos datos presentados con este aspecto: Treemap	Vistreemap
	4.5 En mi organización vemos datos presentados con este aspecto: Bubble	Visburbuja
5	SIEGO Dimensión Informes	5.1 y 5.2
	5.1 En mi organización se utiliza un sistema de informes para la presentación periódica de datos en papel o bien en formato digital (cuadros de mando, reporting, mapas, etc.)	Informes
	5.2 En mi organización disponemos de herramientas de análisis de datos como por ejemplo tablas dinámicas de Excel o cubos multidimensionales para análisis (OLAP).	Herramientasanalisis
6	SIEGO Dimensión Publicar	6.1
	6.1 En mi organización existe un espacio común (portal web, intranet, etc.) en el que se publican los informes.	Espaciopublicar
7	Dimensión Resultados	11.1, 11.2 y 13.1

11.1 Mi percepción sobre el clima laboral en mi organización:	Clima
11.2 Mi percepción sobre los resultados de mi organización es:	Resultados
13.1 En su opinión, ¿está de acuerdo en que la implantación de un sistema de datos estructurados, como el que hemos presentado en esta encuesta, mejoraría los resultados de su organización?	Opinión
8 Dimensión tiempo de respuesta	12.1
12.1 Puedo responder al 95% de preguntas sobre cualquier tema de mi ámbito de responsabilidad con datos concretos en…	Tiemporespuesta
9 Comentarios	14.1
14.1 Para finalizar, nos gustaría darle la oportunidad de dejar los comentarios que crea oportunos. Nos interesa mucho su opinión, experiencia o impresión sobre las barreras o las ventajas de la implantación de un sistema de análisis y presentación de datos en su organización.	

Ilustración 162. Tabla descriptiva variables, bloques, dimensiones y preguntas.

En total se han analizado 43 variables, organizadas en 9 bloques, 6 dimensiones y 33 preguntas.

2.2 Fiabilidad

Se estudió la fiabilidad del cuestionario mediante el estadístico α de Cronbach en el apartado Metodología y preguntas. Se hace necesario al objeto de estudiar la fiabilidad de cada una de las dimensiones que se corresponden con los componentes del modelo SIEGO, aplicar este estadístico a cada una de las dimensiones.

El coeficiente Alfa de Cronbach es un modelo de consistencia interna, basado en el promedio de las correlaciones entre los ítems.

El método de consistencia interna basado en el alfa de Cronbach permite estimar la fiabilidad de un instrumento de medida a través de un conjunto de ítems que se espera que midan el mismo constructo o dimensión teórica. La validez de un instrumento se refiere al grado en que el instrumento mide aquello que pretende medir. Y la fiabilidad de la consistencia interna del instrumento se puede estimar con el alfa de Cronbach. La medida de la fiabilidad mediante el alfa de Cronbach asume que los ítems (medidos en escala tipo Likert) miden un mismo constructo y que están altamente correlacionados (Welch & Comer, 1988). Cuanto más cerca se encuentre el valor del alfa a 1 mayor es la consistencia interna de los ítems analizados. Como criterio general, George y Mallery (2003, p. 231) sugieren las recomendaciones siguientes para evaluar los coeficientes de alfa de Cronbach:

- Coeficiente alfa > 0.9 es excelente
- Coeficiente alfa > 0.8 es bueno

- Coeficiente alfa > 0.7 es aceptable
- Coeficiente alfa > 0.6 es cuestionable
- Coeficiente alfa > 0.5 es pobre

En la tabla siguiente se indican los coeficientes de α de Cronbach calculados para todas las variables de cada una de las dimensiones del modelo propuesto en este trabajo.

Dimensión	Alfa de Cronbach	Alfa de Cronbach basada en elementos tipificados
Dimensión Datos	0.799	0.823
Dimensión Estadística	0.847	0.850
Dimensión Visualización	0.656	0.666
Dimensión Informes	0.587	0.618
Dimensión Resultados	0.569	0.569
TOTAL	**0.868**	**0.874**

Una vez estudiada la fiabilidad del cuestionario de las dimensiones del modelo, se través del coeficiente α de Cronbach, en nuestro caso, se aprecia el coeficiente global de 0,868 que como anteriormente se ha indicado se considera que es bueno, ya que está muy próximo a 1.

Respecto a la fiabilidad de cada una de las dimensiones estudiadas, si bien hay algún valor más cuestionable (dimensión visualización y dimensión informes), en conjunto podríamos decir que estaríamos en un nivel aceptable de fiabilidad.

2.3 Segmentación

En este apartado se describe las variables de este trabajo de investigación que se han caracterizado como variables de segmentación. Estas variables nos dan información sobre los sectores de actividad, el territorio de la organización, su tamaño y sobre la formación y tamaño del equipo del directivo.

Se dispone de 101 respuestas, a las cuales respondieron 61 hombres (60.4 %) y 40 mujeres (39.6 %). Respecto a su edad, casi la mitad tiene entre 40 y 49 años (49.5 %), seguidos por el grupo de 50 a 59 años (22.77 %) y por el de 30 a 39 años (15.84 %). El 8.91% tiene más de 60 años y menos de 29 solo un 2.97 %.

Respecto al nivel de estudios, se aprecia como la mayoría son Licenciados-Graduados (41.58 %) o poseen estudios de Postgrado (30.69 %). Los Diplomados se corresponden con el 9.9% de la muestra, los Doctores con el 7.92%, mientras que el 7.92% tiene estudios secundarios y un 1.8% otros estudios.

En lo que se refiere al tamaño del grupo de persones que gestionan, directa o indirectamente, se obtiene la siguiente prelación: de 5 a 9 personas (27.72 %), de 10 a 49 personas (22.77 %),

más de 200 personas (17.82 %), No gestiono equipo de personas (17.82 %) y de 50 a 199 personas (13.86 %).

Analizando el sector al que pertenece su organización, se parte de 100 respuestas, obteniendo la siguiente distribución de frecuencias -en la tabla aparece como NA los valores perdidos, y %(NA+) y %(NA-) representan la distribución porcentual incluyendo o no los casos perdidos-.

Cuando se estudia la distribución de frecuencias por sector de actividad, cabe destacar la participación en el ámbito del sector público, debido a la difusión que se hizo al ser un trabajo de investigación realizado desde una administración pública.

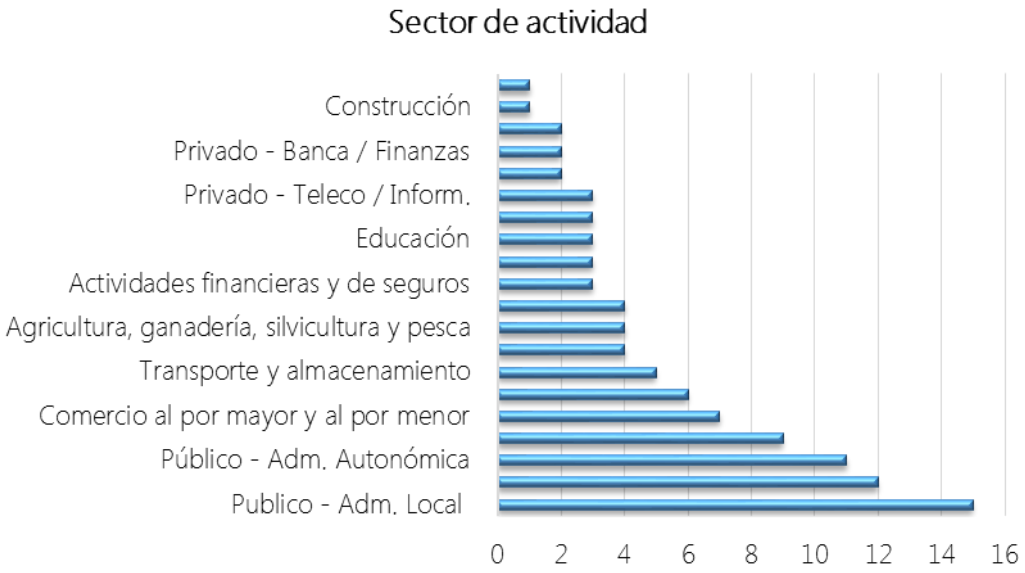


Ilustración 163. Segmentación por Sector de Actividad. Fuente: Elaboración propia (2016)

Respecto al tamaño de la organización, el 55.45% tiene más de 200 trabajadores, el 25.74% entre 50 y 199, y entre 10 y 49 el 14.85 %. Solamente un 2.97% no tiene trabajadores asalariados y el 0.99% tiene entre 1 y 9 personas.

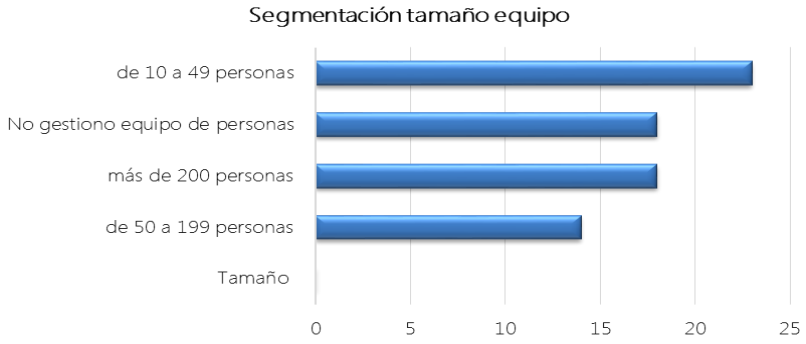


Ilustración 164. Segmentación por tamaño del equipo. Elaboración propia (2016)

Por último, en cuanto a la sede en España de la organización se obtiene la prelación: Catalunya (67.33 %), Madrid, Comunidad de (9.9 %), País Vasco (5.94 %), Aragón (2.97 %), Andalucía

(1.98 %), Castilla - La Mancha (1.98 %), Castilla y León (1.98 %), Comunidad Valenciana (1.98 %), Rioja, La (1.98 %), Canarias (0.99 %), Galicia (0.99 %), Murcia, Región de (0.99 %) y No dispone de sede en España (0.99 %).

Se concentra por tanto un 68% de respuestas en la comunidad autónoma de Catalunya, comunidad en la que se realiza el estudio y que ha recibido mayor difusión.

2.4 Organizaciones con sellos de excelencia y/o certificados de calidad

Una de las hipótesis que se plantea en este estudio es si las organizaciones que aplican modelos de excelencia, que comprometen a la organización en la medición y análisis de datos en diferentes ámbitos o criterios, tienen una mejor percepción de los resultados.

Es por ello que se pide este dato en el cuestionario, incorporándose este valor de segmentación, obteniéndose los siguientes resultados:

EFQM 300+	EFQM 400+	EFQM 500+	Total EFQM	ISO 9001
4	9	15	28	65

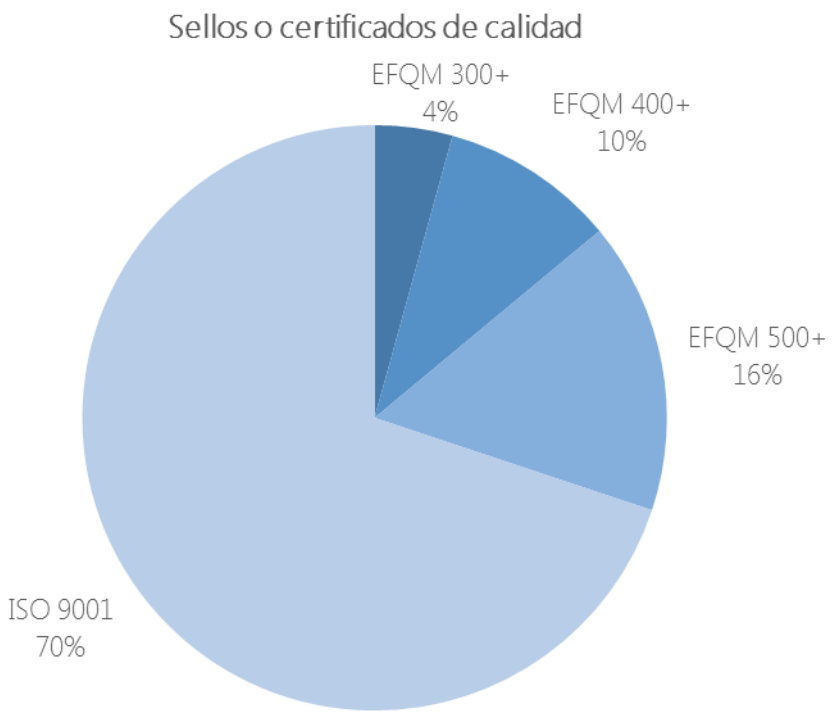


Ilustración 165. Segmentación por sello de calidad o excelencia. Elaboración propia (2016)

3 Resultados de las dimensiones del modelo propuesto SIEGO:

En este apartado se hace un análisis de las variables agrupadas por dimensiones, que en este trabajo de investigación se han caracterizado como las dimensiones del modelo SIEGO: Recoger datos, analizarlos, aplicar componentes de visualización, preparar informes y publicarlos.

3.1 Dimensión Datos

En esta dimensión se pretende analizar como son los sistemas de datos de las organizaciones. En primer lugar, conocer si las organizaciones disponen efectivamente de Data Warehouse y Data Marts. En segundo lugar ver si estos datos responden a los criterios de un modelo de datos estructurado como por ejemplo un sistema de excelencia.

3.1.1 Utilización de bases de datos (Data Warehouse) con los datos organizados por tipos y materias (Data Marts).

A continuación se muestra la tabla de frecuencias de las respuestas y el gráfico correspondiente, así como los datos estadísticos básicos de la variable:

Muestra	Media	Mediana	Moda	Varianza
77	6,22	7,00	8	6,87

Se observa que un 71,43% de la muestra se encuentra concentrada en las categorías 5 a 8, que representan las organizaciones que disponen sistemas de almacenamiento de datos "Sólo para algunos datos de la organización".

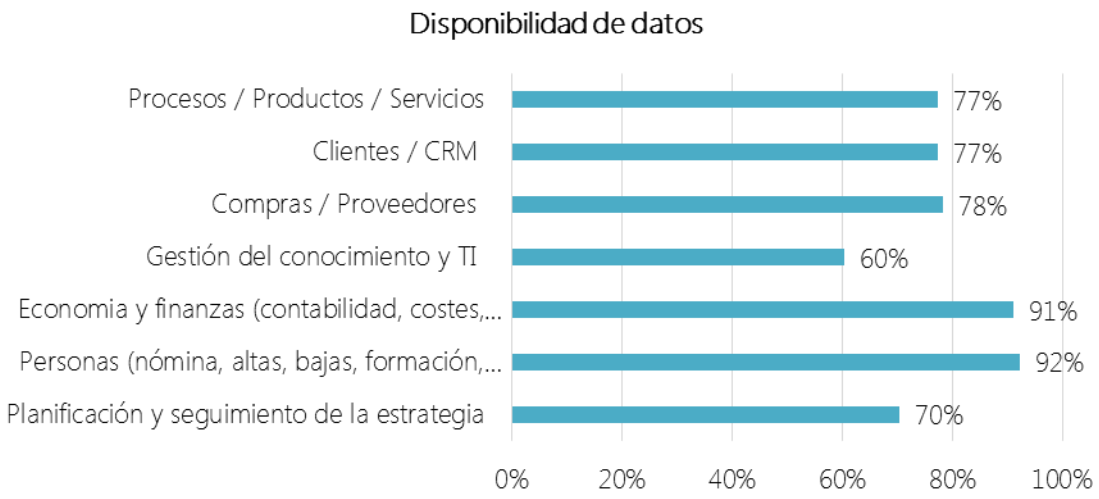
Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No disponemos de datos	0	7	9,09%
	1	1	1,30%
	2	0	0,00%
	3	1	1,30%
	4	2	2,60%
Sólo para algunos datos de la organización	5	16	20,78%
	6	8	10,39%
	7	14	18,18%
	8	17	22,08%
	9	5	6,49%
Para todos los datos de la organización	10	6	7,79%
		77	100,00%

A destacar que solamente un 7,79% de las organizaciones declara disponer de este tipo de sistemas para todos los datos de la organización.

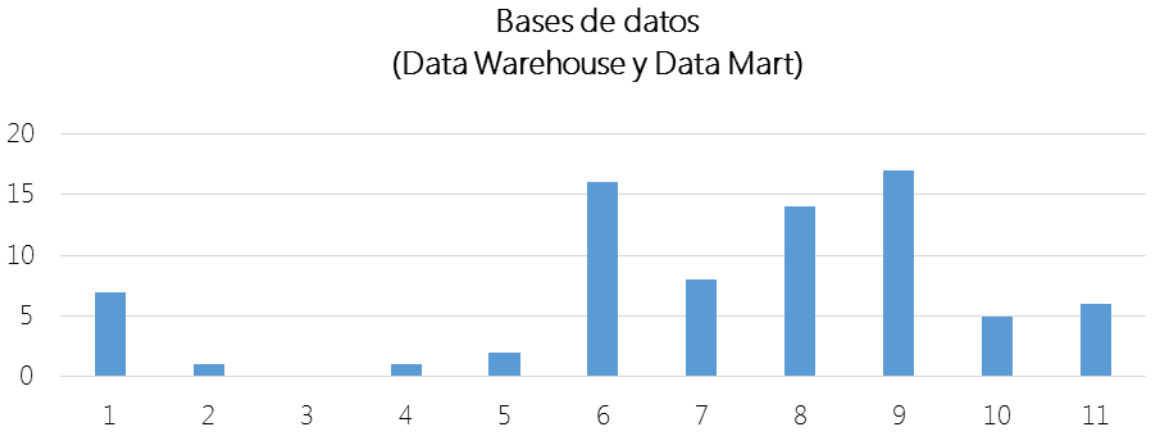
3.1.2 Datos del Sistema de Gestión

En relación con el tipo de datos que la organización incorpora al sistema de gestión se analizan las respuestas según una primera clasificación de datos que incluye las siguientes categorías:

- Planificación y seguimiento de la estrategia
- Personas (nómina, altas, bajas, formación, etc.)
- Economía y finanzas (contabilidad, costes, etc.)
- Gestión del conocimiento y TI
- Compras / Proveedores
- Clientes / CRM
- Procesos / Productos / Servicios



Relativa a la recogida de datos, la primera de las tablas muestra la frecuencia de organizaciones que registran datos de las operaciones o procesos indicados.



En cuanto a la disposición de bases de datos (Data Warehouse) con los datos organizados por

tipos y materias se disponen de 77 casos válidos. La puntuación media es de 6.22 puntos, la moda de 8 y la mediana de 7, lo que significa que más de la mitad de las organizaciones sí disponen de este tipo de información.

					Percentiles (%)					Atípicos (límites)		
	n	Media	Moda	D.típ.	0	25	50	75	100	n	Inf.	Sup.
Datawarehouse	77	6.22	8	2.64	0,00	5,00	7,00	8,00	10,00	7	1,00	10,00

La menor puntuación fue de 0 y la mayor de 10. Se observan 7 datos atípicos. Si denotamos por Q1 o Percentil25 al primer cuartil, por Q3 al tercer cuartil o Percentil75 y por RI al Recorrido Intercuartílico (Q3-Q1), se considera dato atípico aquel que queda fuera del intervalo (Q1-1.5RI, Q3+1.51.5RI).

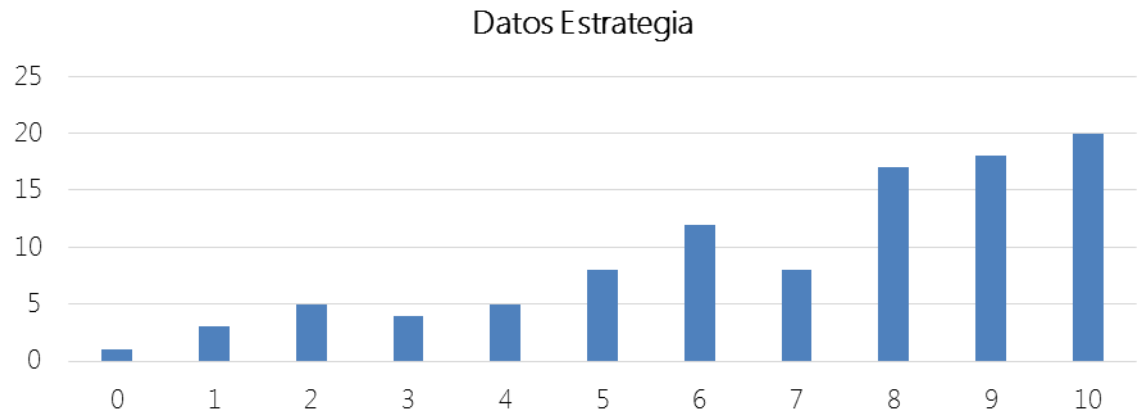
3.2 Datos sobre Estrategia

En este punto se trata de detectar si las organizaciones tienen claramente explicitados sus objetivos estratégicos, con calendarios o programación definida y con seguimiento de su ejecución como elemento básico de un sistema de gestión excelente.

3.2.1 Objetivos estratégicos (o líneas, o áreas, etc.) y estado de ejecución.

Preguntadas las empresas u organizaciones sobre sus objetivos estratégicos (o líneas, o áreas, etc.) y cuál es su estado de ejecución, las respuestas ponen de manifiesto que las organizaciones conocen sus objetivos y su estado de ejecución, estando la moda en 10 y la mediana en 8.

Muestra	Media	Mediana	Moda	Varianza
101	7,04	8,00	10	6,97

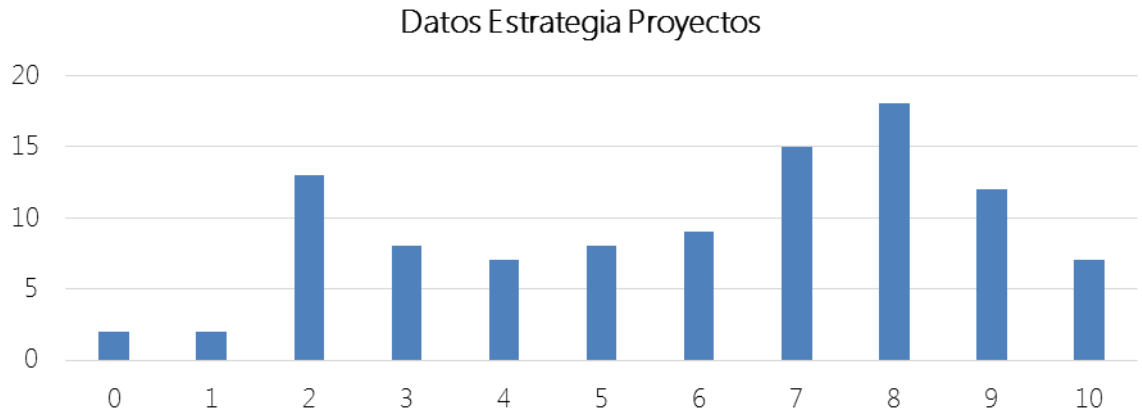


El 74, 25% de las respuestas corresponden a las categorías 6 a 10, "Dispongo de información detallada".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No existe o no dispongo de estos datos	0	1	0,99%
	1	3	2,97%
	2	5	4,95%
	3	4	3,96%
	4	5	4,95%
	5	8	7,92%
	6	12	11,88%
	7	8	7,92%
	8	17	16,83%
	9	18	17,82%
Dispongo de información detallada	10	20	19,80%
		101	100,00%

Sobre el cronograma de proyectos y su estado de ejecución, decir que aunque para esta pregunta la moda baja de 10 a 8 respecto de la anterior pregunta también de estrategia, se refleja que las organizaciones en general sí disponen de cronograma de los proyectos estratégicos y su estado de ejecución.

Muestra	Media	Mediana	Moda	Varianza
101	5,95	7,00	8	7,37



El 60,39% responden a esta pregunta con los valores correspondientes a las categorías 6 a 10, "Dispongo de información detallada".

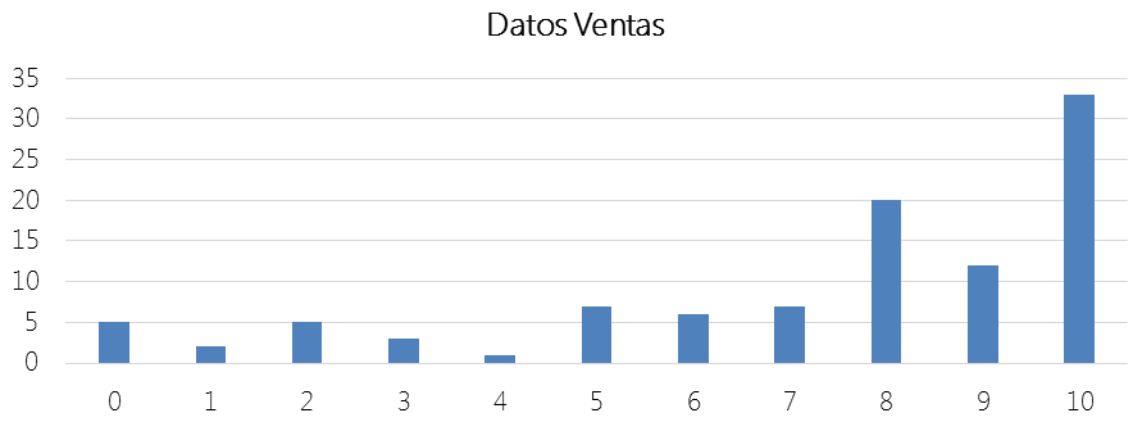
Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No existe planificación	0	2	1,98%
	1	2	1,98%
	2	13	12,87%
	3	8	7,92%
	4	7	6,93%
	5	8	7,92%
	6	9	8,91%
	7	15	14,85%
	8	18	17,82%
	9	12	11,88%
Dispongo de información detallada	10	7	6,93%
		101	100,00%

En general, se detecta que sí se dispone de datos estratégicos generales, pero cuando preguntamos sobre si disponemos de los datos más operativos para proyectos concretos, la disponibilidad de datos disminuye.

3.2.2 Datos sobre Clientes y Ventas

Acceso a información sobre cuáles han sido las ventas de productos o servicios en el período. La moda de la respuesta a esta pregunta se sitúa en el valor máximo, 10, siendo la mediana 8.

Muestra	Media	Mediana	Moda	Varianza
101	7,36	8,00	10	8,76



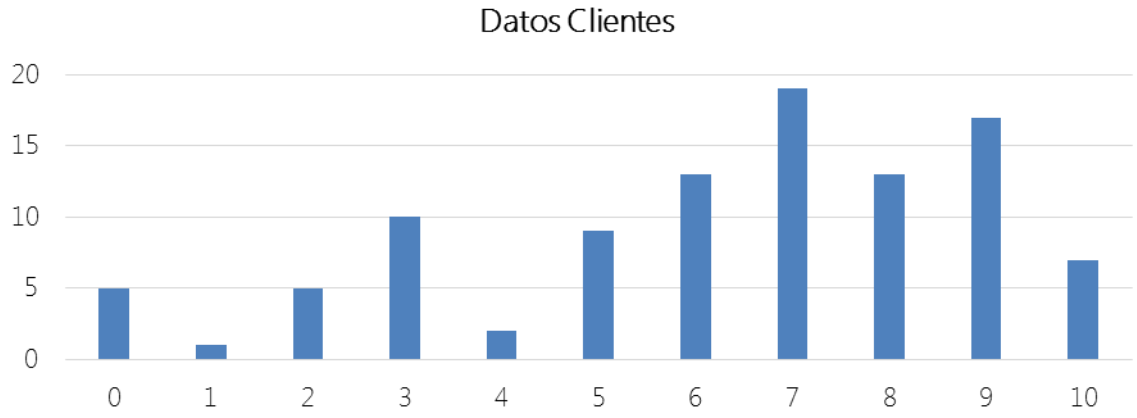
Los datos de ventas son conocidos por el 71,28% de los encuestados, puesto que han respondido a esta pregunta con las categorías 7 a 10, que representan el valor "Dispongo de datos de ventas detallados".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No existe o no conozco esta información	0	5	4,95%
	1	2	1,98%
	2	5	4,95%
	3	3	2,97%
	4	1	0,99%
	5	7	6,93%
	6	6	5,94%
	7	7	6,93%
	8	20	19,80%
	9	12	11,88%
Dispongo de datos de ventas detallados	10	33	32,67%
		101	100,00%

3.2.3 Nivel de satisfacción de los clientes.

Cuando preguntamos sobre el conocimiento que tiene la organización del nivel de satisfacción de los clientes, vemos que la moda se sitúa en el valor 7 al igual que la mediana.

Muestra	Media	Mediana	Moda	Varianza
101	6,26	7,00	7	7,14



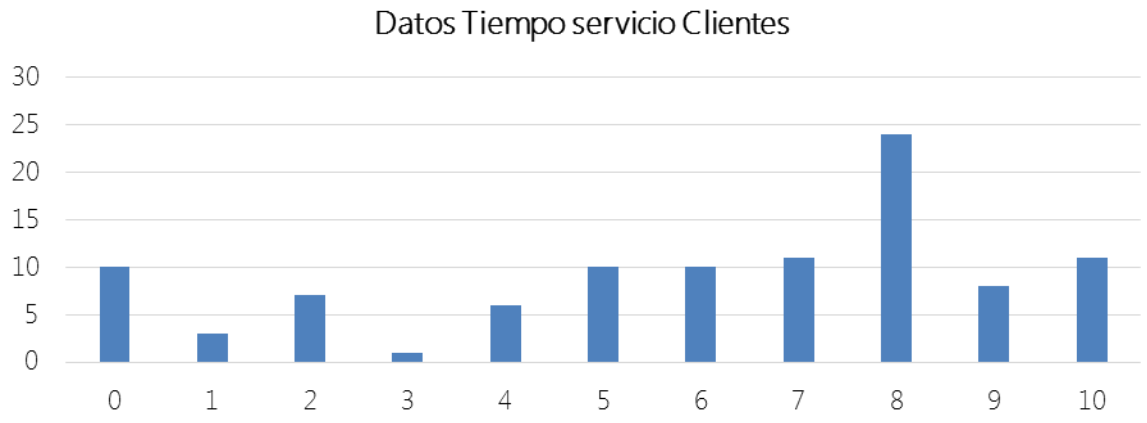
Los datos de satisfacción de los clientes son conocidos por el 68,31% de las organizaciones que respondieron a esta pregunta entre los valores 1 y 10, correspondientes a la categoría "Dispongo de datos de clientes detallados".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No existe o no conozco esta información	0	5	4,95%
	1	1	0,99%
	2	5	4,95%
	3	10	9,90%
	4	2	1,98%
	5	9	8,91%
	6	13	12,87%
	7	19	18,81%
	8	13	12,87%
	9	17	16,83%
Dispongo de datos de clientes detallados	10	7	6,93%
		101	100,00%

3.2.4 Datos sobre Procesos / Productos / Servicios

El tiempo medio de venta o de servicio a un cliente, para cada producto vendido o servicio prestado es una medida bastante conocida por las organizaciones, puesto que la moda de la respuesta se encuentra en 8 y la mediana en 7.

Muestra	Media	Mediana	Moda	Varianza
101	5,99	7,00	8	9,35



Vemos que la mayor parte de las respuestas se concentran en las categorías 5 a 8, siendo la más respondida, la 8, que concentra casi 1 de cada 4 respuestas, el 23,76%.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No recogemos esta información	0	10	9,90%
	1	3	2,97%
	2	7	6,93%
	3	1	0,99%
	4	6	5,94%
Sólo en algunos casos	5	10	9,90%
	6	10	9,90%
	7	11	10,89%
	8	24	23,76%
	9	8	7,92%
Dispongo de datos detallados de tiempos	10	11	10,89%
		101	100,00%

Se desprende del análisis estadístico que las organizaciones disponen de datos detallados de clientes y ventas.

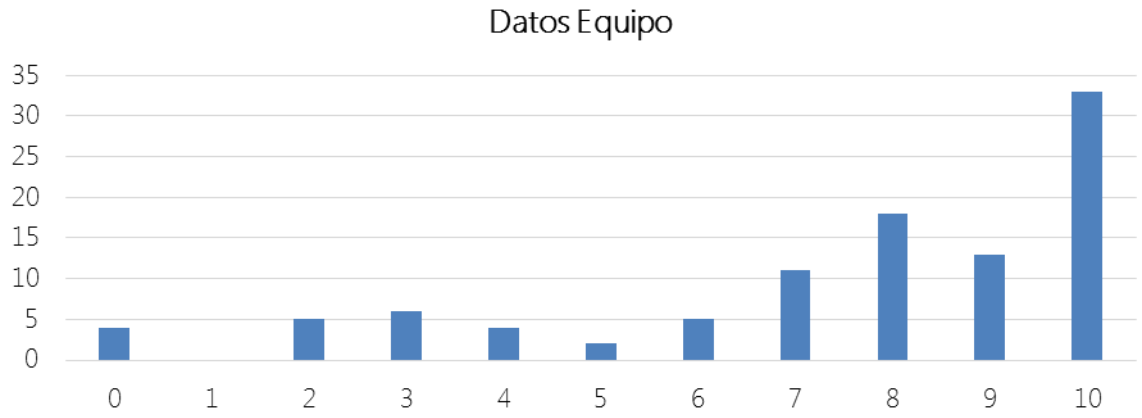
3.3 Datos sobre Personas

Las organizaciones son grupos de personas que realizan procesos o prestan servicios, por ello es relevante el disponer de datos sobre la composición de los equipos de personas así como de todo tipo de datos de segmentación o potencial, que pueden permitir al directivo tomar decisiones sobre la creación de grupos, promoción de personas, etc. y en definitiva, gestionar los equipos de forma eficiente.

3.3.1 Datos sobre la composición del equipo (número de personas, género, edad, nivel de estudios…)

Se encuentra el valor de la moda de la respuesta a esta pregunta en 10 y en 8 la mediana, indicando que las organizaciones conocen mayoritariamente la composición y segmentación de sus equipos.

Muestra	Media	Mediana	Moda	Varianza
101	7,45	8,00	10	8,05



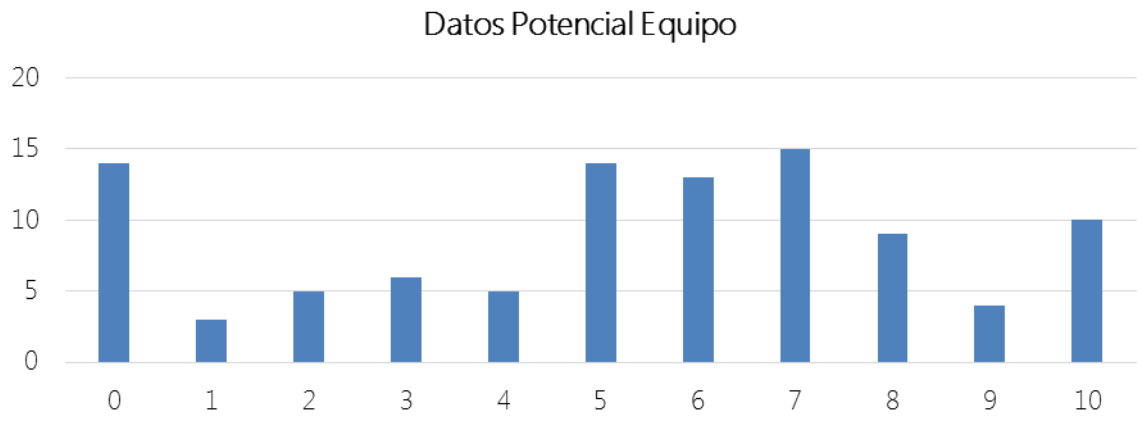
Se concentran en las respuestas 7 a 10 un 74,25% de las respuestas, confirmando que una gran mayoría de las organizaciones disponen de datos detallados de personas.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No recogemos esta información	0	4	3,96%
	1	0	0,00%
	2	5	4,95%
	3	6	5,94%
	4	4	3,96%
Sólo en algunos casos	5	2	1,98%
	6	5	4,95%
	7	11	10,89%
	8	18	17,82%
	9	13	12,87%
Dispongo de datos detallados de personas	10	33	32,67%
		101	100,00%

3.3.2 Datos sobre el potencial de las personas, plan de carrera y plan de formación.

Cuando movemos la pregunta a una mayor profundidad del conocimiento de los equipos de las organizaciones, vemos que la moda se desplaza hacia el valor 7 inferior a la anterior pregunta, mostrando que ya no tantas organizaciones disponen de estos datos.

Muestra	Media	Mediana	Moda	Varianza
98	5,22	6,00	7	9,54



La mayor concentración de respuestas se encuentra en los valores 5 a 7 de las respuestas, que equivalen al 42,87% del total.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No recogemos esta información	0	14	14,29%
	1	3	3,06%
	2	5	5,10%
	3	6	6,12%
	4	5	5,10%
Sólo en algunos casos	5	14	14,29%
	6	13	13,27%
	7	15	15,31%
	8	9	9,18%
	9	4	4,08%
Dispongo de datos detallados de personas	10	10	10,20%
		98	100,00%

De acuerdo al análisis estadístico de las preguntas sobre datos de personas, es habitual disponer de datos del equipo, sobre su composición, su formación, etc. pero ya es más difícil para las organizaciones disponer de información sobre su potencial desarrollo y formación.

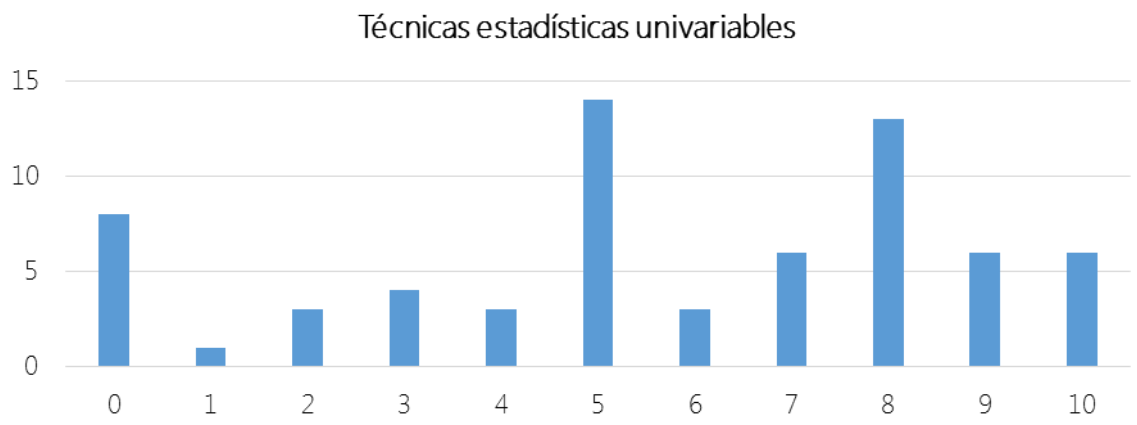
3.4 Dimensión Análisis

En las preguntas relacionadas con esta dimensión, se trata de estudiar el uso de técnicas estadísticas en la presentación y análisis de los datos por parte de las organizaciones. Se analiza si las organizaciones utilizan técnicas estadísticas univariables, bivariables o multivariables.

3.4.1 Técnicas estadísticas univariables

Del estudio estadístico de los datos se deduce que las técnicas estadísticas univariables son las más conocidas por las organizaciones, aunque la moda se sitúa en 5, lo cual corresponde a la respuesta "algunas veces".

Muestra	Media	Mediana	Moda	Varianza
67	5,66	6,00	5	9,60



La mayor concentración de respuestas se encuentra entre las respuestas 5 y 8, que componen el 53,73% del total, y que corresponden a la categoría "Algunas veces".

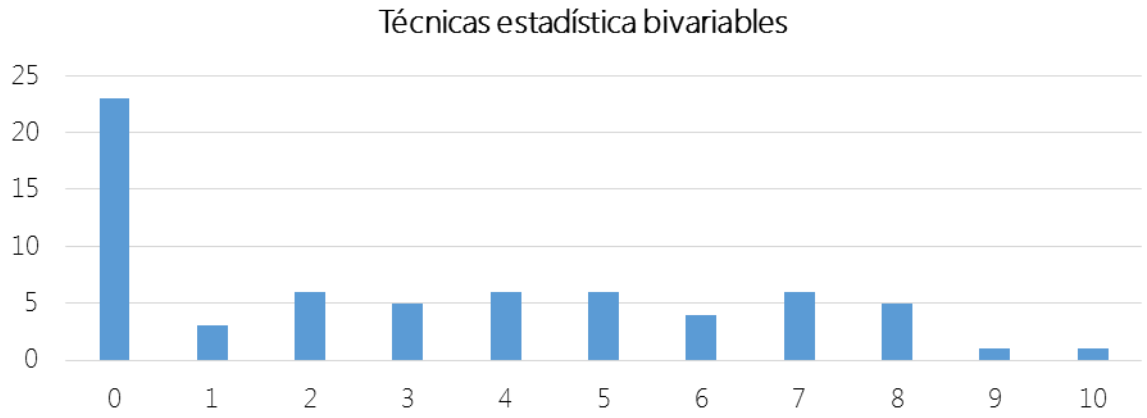
Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	8	11,94%
	1	1	1,49%
	2	3	4,48%
	3	4	5,97%
	4	3	4,48%
Algunas veces	5	14	20,90%
	6	3	4,48%
	7	6	8,96%
	8	13	19,40%
	9	6	8,96%
Muy frecuentemente	10	6	8,96%
		67	100,00%

Cerca de un 20% de las organizaciones considera que aplica técnicas estadísticas univariables muy frecuentemente.

3.4.2 Técnicas estadísticas bivariables

Se observa en los resultados de esta variable, que la moda disminuye al preguntar sobre las técnicas estadísticas bivariables de 5 a 0, lo cual denota que estas técnicas son muy poco utilizadas por las organizaciones, ya que las moda de 0 corresponde a la respuesta "Nunca".

Muestra	Media	Mediana	Moda	Varianza
66	3,17	3,00	0	9,02



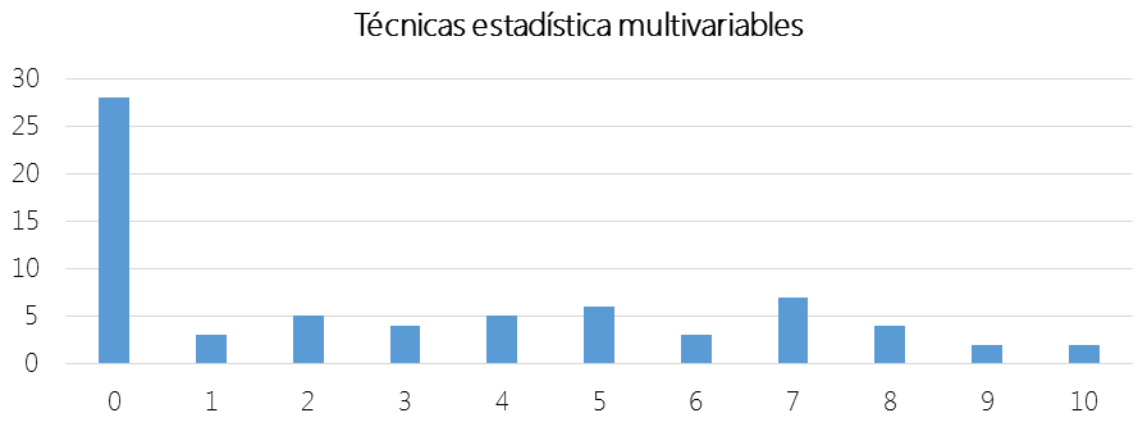
En la tabla de frecuencias se puede ver que el 48,49% de las respuestas corresponden a la categoría "Algunas veces", que comprende las respuestas de la 3 a la 8.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	23	34,85%
	1	3	4,55%
	2	6	9,09%
	3	5	7,58%
	4	6	9,09%
Algunas veces	5	6	9,09%
	6	4	6,06%
	7	6	9,09%
	8	5	7,58%
	9	1	1,52%
Muy frecuentemente	10	1	1,52%
		66	100,00%

3.4.3 Técnicas estadísticas multivariables

De forma muy parecida a la pregunta anterior, las respuestas de las organizaciones a esta pregunta se encuentran mayoritariamente en el grupo "Nunca", que corresponde a las respuestas 0 a 2, y sitúan la moda en 0.

Muestra	Media	Mediana	Moda	Varianza
69	3,07	2,00	0	10,24



Un 40,58% de las respuestas pertenecen a la categoría 0, correspondiente a la respuesta "Nunca".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	28	40,58%
	1	3	4,35%
	2	5	7,25%
	3	4	5,80%
	4	5	7,25%
Algunas veces	5	6	8,70%
	6	3	4,35%
	7	7	10,14%
	8	4	5,80%
	9	2	2,90%
Muy frecuentemente	10	2	2,90%
		69	100,00%

Se concluye que las organizaciones conocen en algunos casos las técnicas estadísticas univariables, que son las que requieren de menos conocimiento técnico y que pueden aplicarse mediante herramientas básicas como una hoja de cálculo. Para las técnicas estadísticas bivariables, en las que se debe poner en relación dos variables, los conocimientos ya deben de ser más avanzados y por ello disminuye el conocimiento que las organizaciones hacen de ellas. Cuando pasamos a técnicas estadísticas multivariables, además del conocimiento de la técnica es necesario disponer de una herramienta estadística avanzada para cálculo específica del tipo R, SPSS o similar, lo cual puede influir en el poco uso que hacen las organizaciones de este tipo de análisis por lo que se desprende de las respuestas.

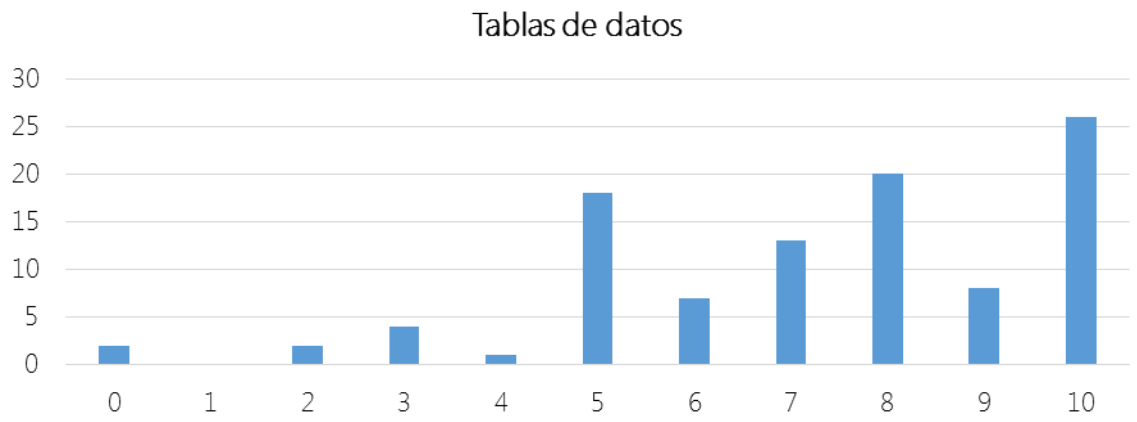
3.5 Dimensión Visualización

El análisis de las respuestas que se agrupan bajo la dimensión de visualización proporciona información sobre la utilización de tablas y gráficos para mostrar los datos de la organización. Se analiza el uso de tablas, gráficos de líneas, gráficos de columnas, gráficos treemaps y gráficos de burbuja o bubble.

3.5.1 Utilización de tablas de datos

En la pregunta sobre la utilización de tablas de datos, la moda se encuentra en 10 correspondiente a la respuesta "Muy frecuentemente". Así mismo, la mediana se encuentra en el valor 8, correspondiente al mismo grupo de respuestas.

Muestra	Media	Mediana	Moda	Varianza
101	7,27	8,00	10	5,72



La frecuencia relativa de las respuestas desde "Algunas veces" a "Muy frecuentemente" (de la opción 5 a la 10) concentra el 91,09% de las respuestas de las organizaciones, lo cual indica que este tipo de representaciones gráficas son muy utilizadas.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	2	1,98%
	1	0	0,00%
	2	2	1,98%
	3	4	3,96%
	4	1	0,99%
Algunas veces	5	18	17,82%
	6	7	6,93%
	7	13	12,87%
	8	20	19,80%
	9	8	7,92%
Muy frecuentemente	10	26	25,74%
		101	100,00%

A continuación se analiza el uso de los distintos tipos de gráficos: de líneas, columnas, tipo treemap o tipo burbuja o bubble.

3.5.2 Gráficos de líneas

Al igual que en el caso de las tablas, las respuestas registradas muestran que este tipo de componente (gráfico de líneas) es muy utilizado por las organizaciones, siendo el valor de la moda 8 y el de la mediana 7.

Muestra	Media	Mediana	Moda	Varianza
101	7,01	7,00	8	5,13

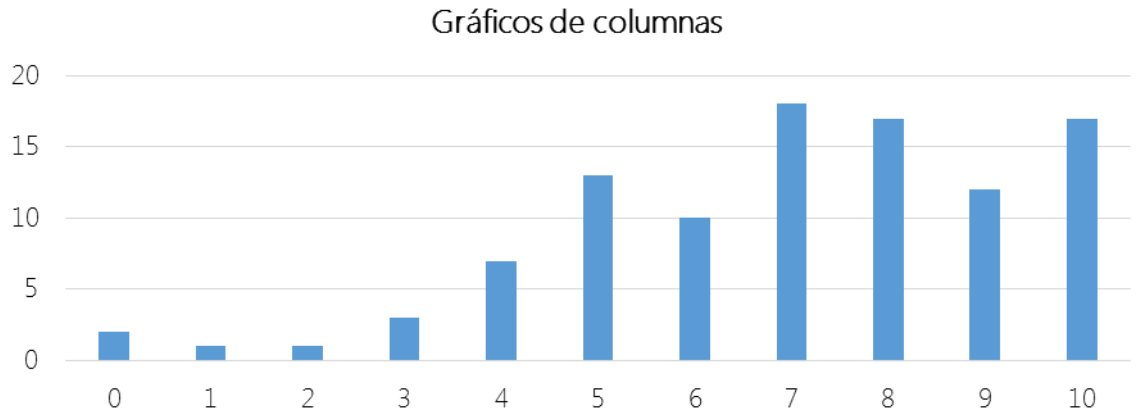
El 87,12% de las 101 respuestas obtenidas se concentran en los grupos "Algunas veces" y "Muy frecuentemente", categorías de la 5 a la 10.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	2	1,98%
	1	0	0,00%
	2	1	0,99%
	3	3	2,97%
	4	7	6,93%
Algunas veces	5	17	16,83%
	6	5	4,95%
	7	18	17,82%
	8	20	19,80%
	9	12	11,88%
Muy frecuentemente	10	16	15,84%
		101	100,00%

3.5.3 Uso de gráficos de columnas

Al igual que las dos preguntas anteriores, las repuestas a esta preguntan denotan que este tipo de componente (Gráfico de columnas) es fuertemente usado por las organizaciones y sus directivos, estando la mediana en 7 al igual que la moda.

Muestra	Media	Mediana	Moda	Varianza
101	6,98	7,00	7	5,44



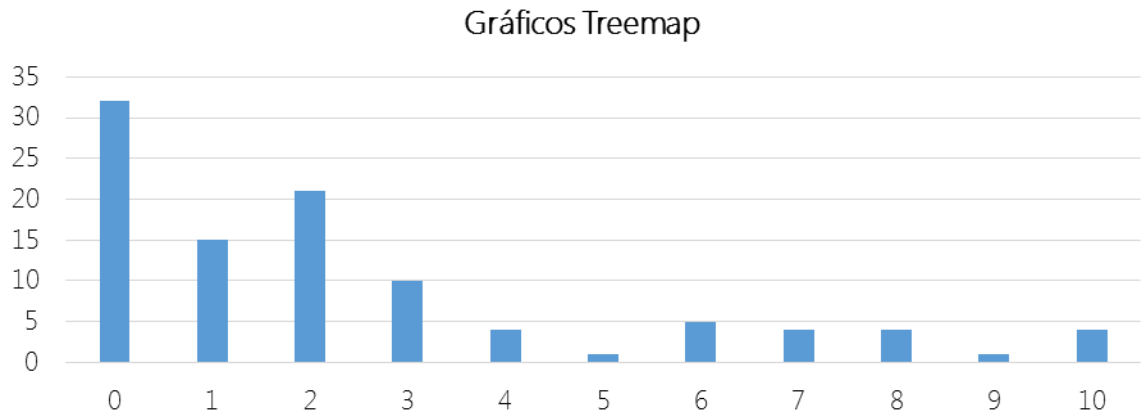
El 28,71% de las respuestas corresponden a los valores 9 y 10, de la categoría "Muy frecuentemente".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	2	1,98%
	1	1	0,99%
	2	1	0,99%
	3	3	2,97%
	4	7	6,93%
Algunas veces	5	13	12,87%
	6	10	9,90%
	7	18	17,82%
	8	17	16,83%
	9	12	11,88%
Muy frecuentemente	10	17	16,83%
		101	100,00%

3.5.4 Uso de gráficos treemap

En este caso, se observa claramente que las organizaciones no usan este tipo de componente gráfico, puesto que la moda se encuentra en 0 y la mediana en 2.

Muestra	Media	Mediana	Moda	Varianza
101	2,44	2,00	0	7,79



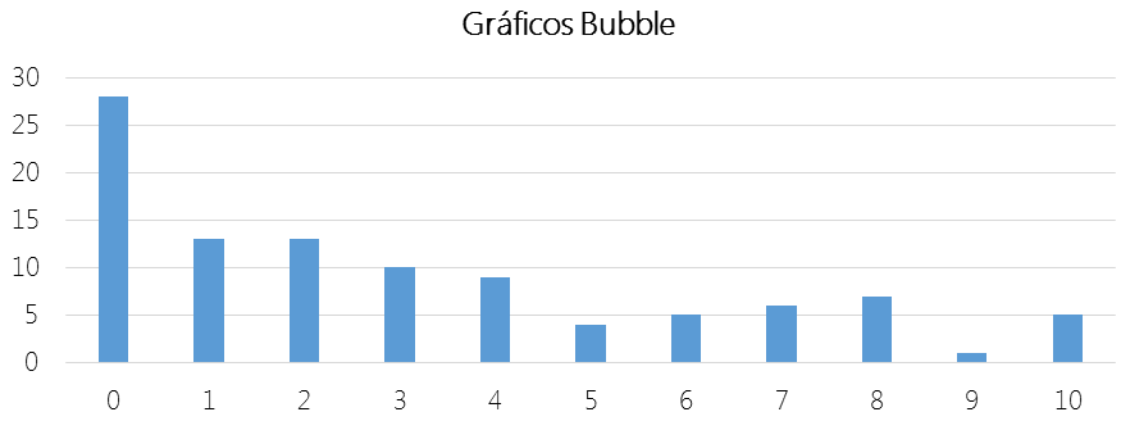
Las respuestas de la 0 a la 2, correspondientes al grupo "Nunca", corresponden al 67,32% de la muestra, lo que claramente indica que la mayoría de las organizaciones no ven este tipo de componente en sus informes.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	32	31,68%
	1	15	14,85%
	2	21	20,79%
	3	10	9,90%
	4	4	3,96%
Algunas veces	5	1	0,99%
	6	5	4,95%
	7	4	3,96%
	8	4	3,96%
	9	1	0,99%
Muy frecuentemente	10	4	3,96%
		101	100,00%

3.5.5 Gráficos bubble o de burbuja

Al igual que en la pregunta anterior, se ve claramente que las organizaciones no utilizan o no conocen este tipo de gráficos, siendo la moda 0 y la mediana 2.

Muestra	Media	Mediana	Moda	Varianza
101	3,08	2,00	0	9,28



Se observa claramente que el 72,27% de las organizaciones concentran sus respuestas en la categoría "Nunca" reflejando que nunca ven este tipo de gráficos en su organización.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Nunca	0	28	27,72%
	1	13	12,87%
	2	13	12,87%
	3	10	9,90%
	4	9	8,91%
Algunas veces	5	4	3,96%
	6	5	4,95%
	7	6	5,94%
	8	7	6,93%
	9	1	0,99%
Muy frecuentemente	10	5	4,95%
		101	100,00%

Se confirma por el estadístico que las organizaciones a menudo utilizan sistemas de visualización de información básicos como los gráficos de líneas, tablas o columnas, pero que no son muy frecuentes las organizaciones que utilizan componentes más avanzados como los treemaps o los bubble.

El uso de este tipo de gráficos más complejos requiere de dos factores esenciales: el conocimiento de métodos estadísticos bi y multivariables para calcularlos, y del conocimiento de herramientas de visualización que permiten generarlos. No se puede separar la necesidad de cierto conocimiento de la generación o uso de estos componentes, ya que aunque actualmente muchas aplicaciones transaccionales incluyen módulos de preparación de dashboards y/o informes, ello no significa que los usuarios los entiendan y sepan interpretarlos.

3.6 Dimensión Informes

La información de una organización debe de ser diseminada con agilidad haciéndola llegar en tiempo y espacio adecuado al usuario que la precisa. De esta manera se favorece la eficiencia en la toma de decisiones.

Conforme más tiempo necesita un directivo para obtener la información que le solicitan, más activa debe de ser su intuición, puesto que menos fundamentada en datos puede ser. Las decisiones de los directivos se podrían tomar con más fundamento si dispusieran de un Sistema de Informes que les permitiera obtener la información necesaria más frecuentemente, en menos tiempo.

El Sistema de Informes, que constituye una de las dimensiones del modelo SIEGO, está compuesto por un conjunto de informes mediante los cuales se debe poder obtener, en menos de 5 minutos, hasta un 95% de la información necesaria para la toma de decisiones.

3.6.1 Utilización de un sistema de informes para la presentación periódica de datos en papel o bien en formato digital (cuadros de mando, reporting, mapas, etc.).

De las respuestas se desprende que en general las organizaciones sí disponen de un sistema de informes, sean en papel o digitales, siendo tanto la moda como la mediana de 8.

Muestra	Media	Mediana	Moda	Varianza
101	7,40	8,00	8	5,17

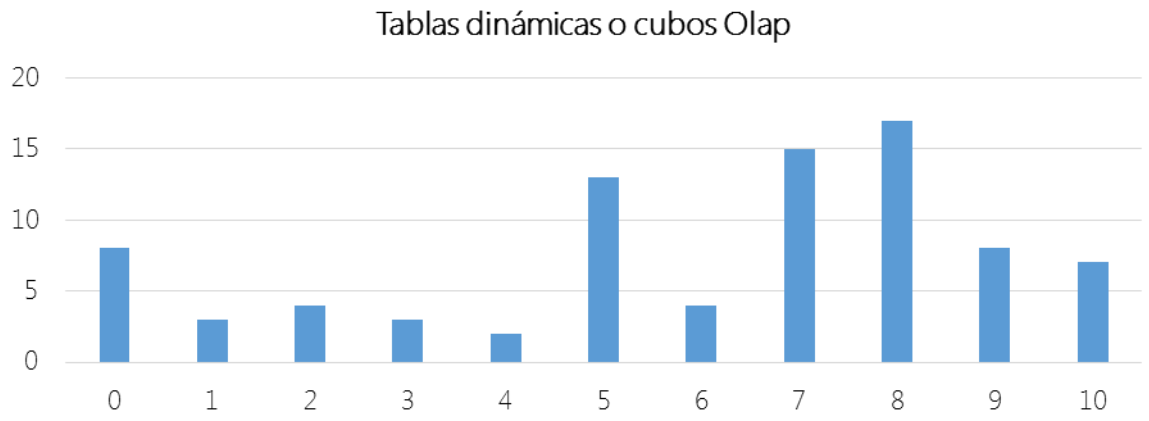
El 56,43% de las organizaciones han respondido la categoría 8 a 10, correspondiente a la respuesta que disponen de un sistema de informes definido.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No disponemos de informes	0	3	2,97%
	1	0	0,00%
	2	1	0,99%
	3	2	1,98%
	4	1	0,99%
Sólo para algunos datos	5	14	13,86%
	6	6	5,94%
	7	17	16,83%
	8	23	22,77%
	9	13	12,87%
Sistema de informes definido	10	21	20,79%
		101	100,00%

3.6.2 Herramientas de análisis de datos como por ejemplo tablas dinámicas de Excel o cubos multidimensionales (OLAP).

Las herramientas OLAP o las tablas dinámicas son muy utilizadas por las organizaciones, estando el valor 7 en la mediana y la moda en 8.

Muestra	Media	Mediana	Moda	Varianza
84	5,95	7,00	8	8,83



El 55,95% de las organizaciones ha respondido las categorías 8 a 10, que corresponde al valor "Para todo tipo de datos".

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No disponemos	0	8	9,52%
	1	3	3,57%
	2	4	4,76%
	3	3	3,57%
	4	2	2,38%
Sólo para algunos tipos de datos	5	13	15,48%
	6	4	4,76%
	7	15	17,86%
	8	17	20,24%
	9	8	9,52%
Para todo tipo de datos	10	7	8,33%
		84	100,00%

Según se desprende de estos datos, claramente las organizaciones disponen de sistemas de informes para todo tipo de datos, sean del tipo que sean, dinámicos o estáticos.

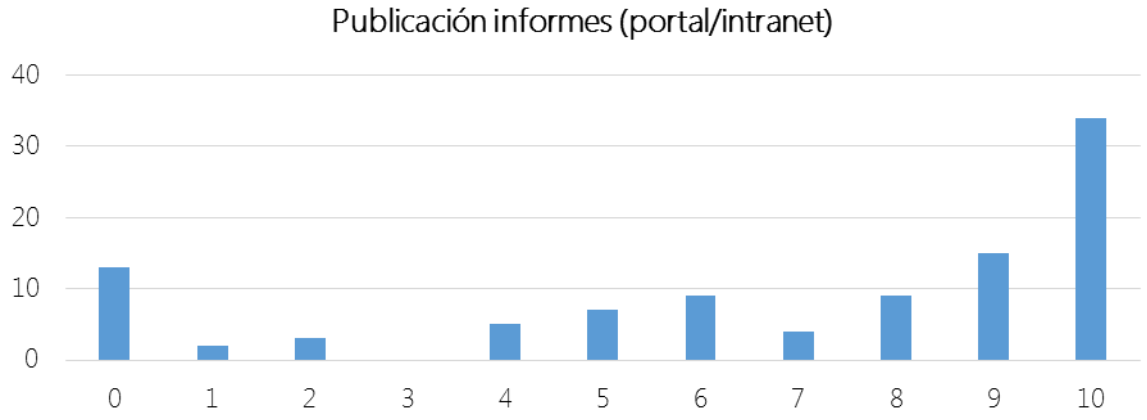
3.7 Dimensión Publicación

El modelo SIEGO propone como última de sus dimensiones la de Publicación. En esta dimensión se evalúa si las organizaciones disponen de un espacio común en el que compartir el conocimiento explícito de la organización que se genera con el Sistema de Informes de la organización.

3.7.1 Utilización de un espacio común (portal web, intranet, etc.) en el que se publican los informes.

Se encuentra la moda en el valor máximo 10, por lo que se puede afirmar que las organizaciones sí disponen de un espacio para la publicación de informes.

Muestra	Media	Mediana	Moda	Varianza
101	6,85	8,00	10	12,26



Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No existe este espacio	0	13	12,87%
	1	2	1,98%
	2	3	2,97%
	3	0	0,00%
	4	5	4,95%
	5	7	6,93%
	6	9	8,91%
	7	4	3,96%
	8	9	8,91%
	9	15	14,85%
Existe un portal / intranet	10	34	33,66%
		101	100,00%

Se observa que no existen prácticamente valores intermedios en esta pregunta. La varianza se sitúa en un 12.26, pues las organizaciones disponen en su mayoría de este espacio o no disponen de él.

3.8 Dimensión Resultados

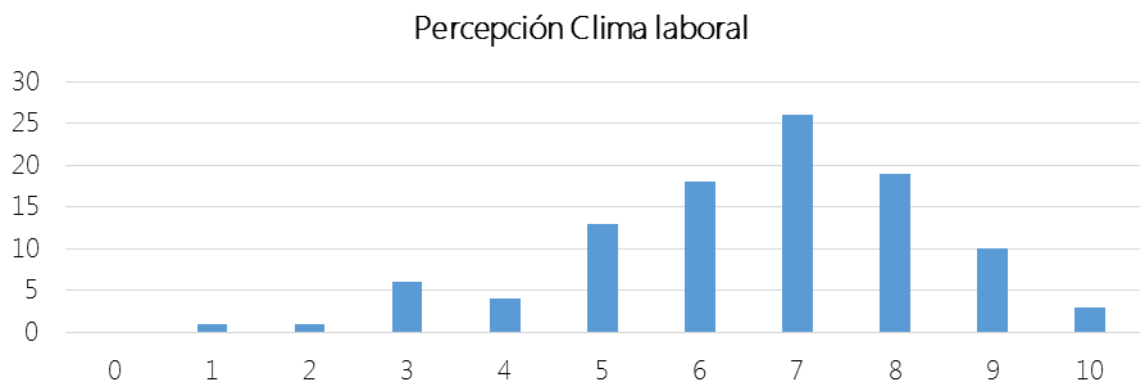
Se analiza a continuación las respuestas relativas a la percepción de los resultados de la organización por parte de los directivos.

Estas preguntas, no tienen la finalidad de obtener datos objetivos sobre la influencia del modelo presentado en los resultados organizacionales, sino de obtener una medida subjetiva del usuario que responde al cuestionario que permita tener una aproximación subjetiva para valorar los resultados de la organización.

3.8.1 Percepción sobre el clima laboral de la organización

Se pregunta a las organizaciones sobre cuál es su percepción sobre el clima laboral, y se calcula tanto la media como la mediana de las respuestas se encuentra en el valor 7.

Muestra	Media	Mediana	Moda	Varianza
101	6,57	7,00	7	3,27



El 75,24% de las respuestas se concentran en los valores 5 a 8, que corresponderían a la respuesta intermedia entre pésimo e inmejorable.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Pésimo	0	0	0,00%
	1	1	0,99%
	2	1	0,99%
	3	6	5,94%
	4	4	3,96%
	5	13	12,87%
	6	18	17,82%
	7	26	25,74%
	8	19	18,81%
	9	10	9,90%
Inmejorable	10	3	2,97%
		101	100,00%

3.8.2 Percepción sobre los resultados de la organización

Las organizaciones tienen una buena percepción de sus resultados. Se encuentra el valor de la mediana en 7 y la moda en 8.

Muestra	Media	Mediana	Moda	Varianza
101	6,80	7,00	8	3,36

La mayor parte de los resultados se encuentran en las respuestas correspondientes a los valores 6 a 9, que conforman el 76,23% de las respuestas. Tres de cada cuatro encuestados perciben buenos resultados en su organización.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Pésimos	0	0	0,00%
	1	0	0,00%
	2	1	0,99%
	3	7	6,93%
	4	8	7,92%
	5	6	5,94%
	6	12	11,88%
	7	23	22,77%
	8	29	28,71%
	9	13	12,87%
Inmejorables	10	2	1,98%
		101	100,00%

3.8.3 Un sistema de información como el propuesto, SIEGO, mejoraría los resultados de la organización.

La gran mayoría de las organizaciones están de acuerdo en que implantar un sistema como el Modelo SIEGO que se plantea, tendría un impacto positivo en los resultados de la organización mejorándolos. Se encuentra la moda en el valor 10 y la mediana en 8.

Muestra	Media	Mediana	Moda	Varianza
101	7,79	8,00	10	4,04

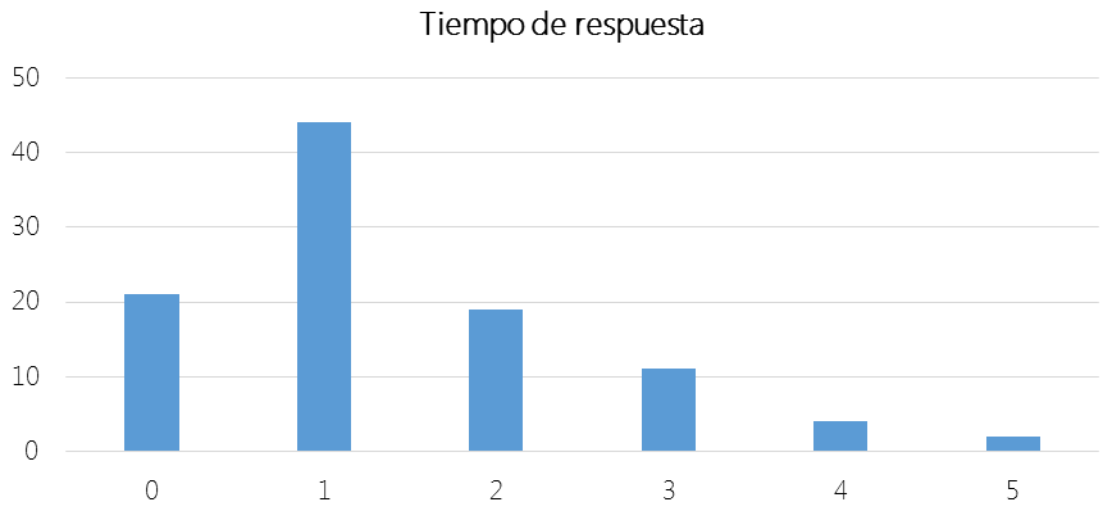
Ninguna organización responde que no está de acuerdo con el hecho que de disponer de un sistema como el que se plantea en este trabajo mejorase los resultados de la organización.

Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
No estoy de acuerdo	0	0	0,00%
	1	0	0,00%
	2	1	0,99%
	3	0	0,00%
	4	2	1,98%
Es probable	5	20	19,80%
	6	4	3,96%
	7	14	13,86%
	8	15	14,85%
	9	15	14,85%
Estoy seguro	10	30	29,70%
		101	100,00%

Prácticamente todas, el 97,02% de las organizaciones, creen que es probable o están seguros (categorías de la 5 a la 10), que un sistema como el planteado en este trabajo de investigación tendría un impacto positivo en los resultados de la organización.

3.9 Dimensión Tiempo de respuesta

Se plantea en esta dimensión una pregunta des de la perspectiva que hay un porcentaje finito de demandas de información para las cuales no existe respuesta o no se conoce. Se considera que este tipo de preguntas no deberían representar más del 5% de las demandas de información que un directivo puede recibir. Se evalúa entonces cual es el tiempo medio que subjetivamente considera el directivo que puede tardar en responder a una solicitud de información de su ámbito de responsabilidad.



Respuesta	Categoría	Frecuencia por categoría	Frecuencia relativa por categoría
Menos de 5 minutos	0	21	20,79%
Entre 5 minutos y 60 minutos (1 hora)	1	44	43,56%
Entre 1 hora y 8 horas (1 día)	2	19	18,81%
Entre 1 día y 5 días (1 semana)	3	11	10,89%
Entre 1 semana y 4 semanas (1 mes)	4	4	3,96%
No puedo responder al 95% de las peticiones de información	5	2	1,98%
		101	100,00%

Destaca que para esta pregunta, un 43,56% de las organizaciones creen que pueden responder en un tiempo inferior a una hora a cualquier petición de información.

Sólo el 20% de las organizaciones cree que puede responder en menos de 5 minutos.

3.10 Comentarios de los encuestados

Se pedía a los participantes en la encuesta, que aportaran sus comentarios sobre los sistemas de información como los que se presentan en este trabajo. El usuario podía dejar en texto libre los comentarios que creyera oportunos. También se le solicitaba su opinión sobre las ventajas e inconvenientes de la implantación de un sistema de análisis y presentación como el modelo SIEGO.

Para hacer el análisis de esta pregunta, se utilizó una nube de palabras:

Ilustración 166. Opinión encuestados. Fuente: Elaboración propia (2016)

Una nube de palabras es una representación gráfica de frecuencias en un texto. El tamaño de cada palabra se determina en función de las veces que se repite ésta en el texto seleccionado. Actualmente, existen muchas páginas web en las que al pegar el texto se genera el gráfico nube de palabras de forma automática. Posteriormente pueden eliminarse palabras en concreto de la lista, como por ejemplo las preposiones, artículos, pronombres, etc. Y volver a generar de nuevo la nube que finalmente se convierte en un gráfico que puede insertarse en cualquier documento.

En esta nube destacan principalmente las palabras datos, información, análisis, decisiones. No aparecen palabras como coste, dificultad, etc.

4 Análisis de las relaciones entre las dimensiones

Para analizar las relaciones entre las dimensiones, se calculó el coeficiente de correlación de Pearson, junto con la significación asociada al test correspondiente, para constatar la relación lineal entre obtener mejores resultados y disponer de un sistema de datos, analizarlos, aplicar técnicas de visualización, preparar informes y publicarlos. Se realizó también dicho análisis para corroborar las relaciones entre las dimensiones del modelo propuesto.

El test de Kruskal-Wallis fue utilizado para estudiar la relación entre el tiempo de respuesta a peticiones de información y el hecho de disponer de un sistema de información como el que se propone.

Para valorar la tercera y la cuarta hipótesis de la investigación, se empleó el test t de Student o el test de Wilcoxon para muestras independientes, según se verificase o no la hipótesis de normalidad.

El nivel de significación empleado en todo el análisis fue de 0,05. Cuanto menor sea el valor de p, menor será la probabilidad de que los resultados obtenidos se deban al azar y mayor evidencia habrá a favor de la hipótesis planteada. Si dicha probabilidad es menor que un valor de p fijado previamente (en este caso se toma p<0,05), la hipótesis nula se rechazará. Así, cuando el valor de p está por debajo de 0,05, se dirá que el resultado es *estadísticamente significativo* y será *no significativo* en cualquier otro caso.

4.1 Descriptivo por dimensiones del modelo propuesto y para resultados

Se calculó la suma de puntuaciones de los ítems que conforman cada dimensión propuesta. La tabla muestra así para cada dimensión las medidas de posición central y percentiles, junto con la desviación típica.

					Percentiles (%)					Atípicos (límites)		
	n	Media	Moda	D.típ.	0	25	50	75	100	n	Inf.	Sup.
Dim Datos	101	55.32	69	16.77	5,00	44,00	59,00	68,00	82,00	1	12,00	82,00
Dim Análisis	85	10.27	0	7.65	0,00	4,00	10,00	15,00	30,00	0	0,00	30,00
Dim Visualización	101	26.80	28	8.42	10,00	21,00	27,00	31,00	50,00	2	10,00	46,00
Dim Informar	101	12.36	9	4.89	0,00	9,00	13,00	16,00	20,00	0	0,00	20,00
Dim Publicar	101	6.85	10	3.52	0,00	5,00	8,00	10,00	10,00	0	0,00	10,00
Dim Resultados	101	13.38	14	3.06	3,00	12,00	14,00	15,00	20,00	6	8,00	19,00

Los valores de los estadísticos de centralidad no se corresponden con los valores de 0 a 10, puesto que para construir las dimensiones se ha tomado la suma de los valores de las variables incluidas en cada dimensión.

4.2 Análisis de correlaciones entre las dimensiones

En este apartado se relacionan una dimensión con todas las demás, para obtener las distintas correlaciones. Como puede apreciarse en la tabla, todas las relaciones son positivas lo cual significa que si mejora una de las dimensiones también lo hará la dimensión relacionada.

Las organizaciones que trabajen en la mejora de una de las dimensiones del modelo SIEGO mejorarán todas las demás dimensiones del modelo, debido a la correlación positiva que se da entre todas ellas.

Las correlaciones más fuertes se dan entre la dimensión Informar y la Dimensión Datos y Visualización, y entre la Dimensión Publicar y la de Informar, como se muestra en la tabla siguiente.

También se indica en la tabla el p-valor de la relación, indicando con un asterisco cuando p<0,05, con dos asteriscos cuando p<0'001 y con tres asteriscos cuando p<0,0001. Nótese que la relación Informar-Datos, Informar-Visualizar y Visualizar-Análisis muestran una alta significación estadística, lo cual indica que la probabilidad de que este resultado se deba al azar es altamente improbable.

	Dim Datos	Dim Análisis	Dim Visualización	Dim Informar	Dim Publicar
Dim Datos	1	0,382***	0.129	0,428***	0,299**
DimAnalisis		1	0,366***	0,339**	0,263*
Dim Visualización				0,424***	0,287**
Dim Informar				1	0,308**
Dim Publicar					1

Matriz de correlaciones: *:p<0,05; **:p<0,001; ***:p<0,001

Analizadas las relaciones entre las dimensiones del modelo, se aprecia que la relación más fuerte se da entre la dimensión asociada a los Datos y la de Informar muy cercana a la de Visualizar con Informar, mientras que la más débil se encuentra entre el hecho de disponer de Datos y la aplicación de técnicas de Visualización.

Parece que del análisis estadístico se desprende que para informar lo más relevante es disponer de datos y aplicar componentes de visualización, y para publicar lo relevante es disponer del Informe, como indican los valores señalados en azul.

Debido a que las técnicas de visualización llevan implícitas las técnicas estadísticas, se entiende que la correlación más fuerte salga con los datos y con la aplicación de componentes de visualización, y no tanto con las técnicas estadísticas. Es decir, el tipo de análisis estadístico condiciona el tipo de componente a utilizar. Se considera además que en el Informe, lo más relevante es como se visualiza la información, y no se estima tanto como se ha analizado estadísticamente la información.

Todas las relaciones resultaron estadísticamente positivas, obteniéndose coeficientes de correlación positivos. Dichos coeficientes, junto con los diagramas de dispersión e histogramas se combinan en el gráfico siguiente:

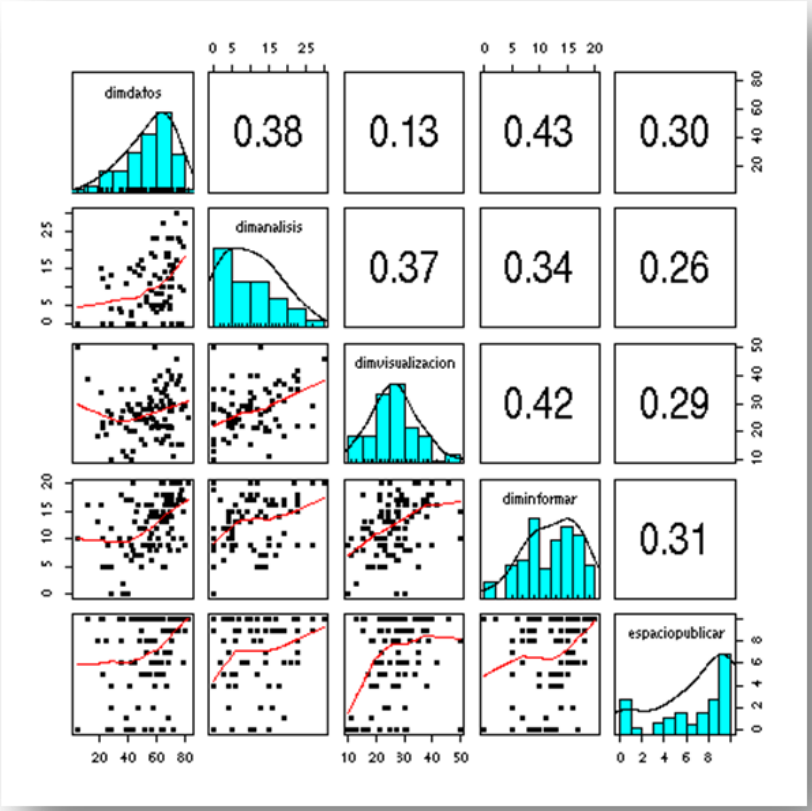


Ilustración 167. Correlación entre dimensiones

5 Comprobación de las hipótesis

5.1 H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados

A continuación se presenta el coeficiente de correlación de Pearson y la significación del test asociado, utilizados para estudiar la correlación entre la percepción de resultados y las dimensiones propuestas en el modelo SIEGO: datos, análisis, visualización, informe y publicación.

En todos los casos los coeficientes de correlación son positivos, indicadores de correlaciones positivas entre las variables consideradas, siendo todas ellas estadísticamente significativas ($p<0.05$) salvo para la dimensión visualización. Por tanto **se contrasta que las organizaciones que dispongan de un sistema de información como el propuesto percibirán mayores resultados**, excepto para la aplicación de componentes de visualización que no es estadísticamente significativa.

	Coef. Correlación r	p-valor
Dim Datos	0,548	<0,001
Dim Análisis	0,295	<0,01
Dim Visualización	0,117	0,24
Dim Informar	0,211	0,03
Dim Publicar	0,239	0,02

En la primera fila del siguiente gráfico se muestra el coeficiente de correlación entre la dimensión resultados y el resto de dimensiones, y en la primera columna los diagramas de dispersión asociados.

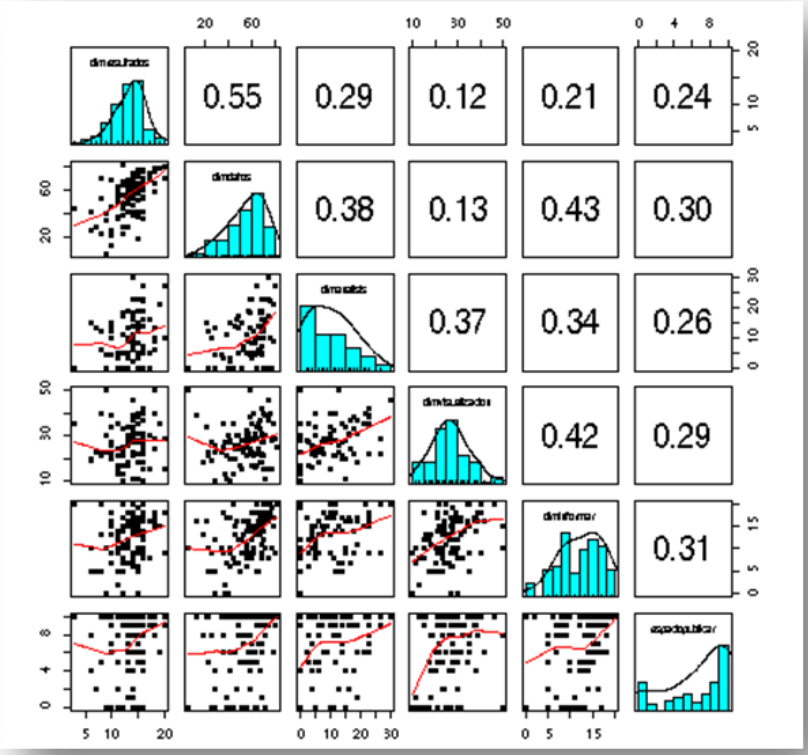


Ilustración 168. Correlación entre dimensiones y resultados.

De ello **se podría afirmar que las organizaciones que utilizan un modelo similar al que se propone, SIEGO**:

- Disponen de un sistema de datos, porque presenta un p-valor <0,001
- Aplican técnicas de análisis, con un p-valor <0,01
- Disponen de un Sistema de Informes, con p-valor de 0,03
- Publican los informes, con p-valor de 0,02

Percibirán mejores resultados en la organización.

La comprobación de esta hipótesis es fundamental porque da validez prácticamente a la totalidad del modelo SIEGO presentado en este trabajo de investigación.

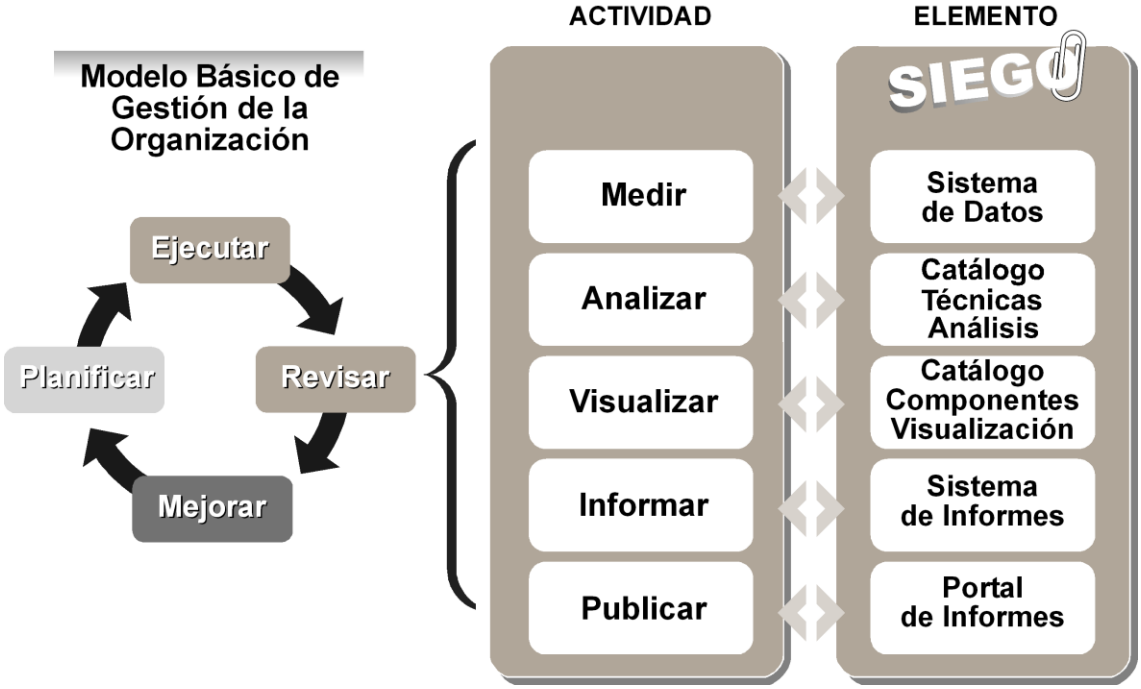


Ilustración 169. Modelo SIEGO. Fuente: Elaboración propia (2016)

5.2 H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas.

Para contrastar esta hipótesis se realiza un análisis entre las dimensiones del modelo y las diferentes franjas de tiempo de respuesta establecidas en el cuestionario.

Al poner en relación las dimensiones con el tiempo de respuesta se encuentran diferencias significativas en cuanto a la dimensión Informar y la dimensión Datos, no siendo así para el resto de las dimensiones.

Se aprecian diferencias significativas en el tiempo de respuesta debido a que la puntuación obtenida en el grupo que tarda un día en responder es muy baja comparada con el resto. En este grupo solo hay 4 respuestas, lo que provoca también esta descompensación con el resto.

La dimensión Informar presenta un p-valor de 0,04 y la dimensión Datos de 0,049, ambas menores de 0,05 que es el valor de significación estadística establecida para este estudio.

	< 5 min				1 hora				1 día			
	n	Media	Mediana	DT	n	Media	Mediana	DT	n	Media	Mediana	DT
Dim Datos	19	54,53	58,00	11,26	44	59,68	61,15	14,26	4	44,50	45,00	6,61
Dim Análisis	16	11,50	9,00	7,34	37	10,51	10,00	7,76	3	15,67	15,00	1,15
Dim Visualización	19	27,00	29,00	8,96	44	26,39	26,50	8,41	4	24,75	25,50	3,59
Dim Informar	19	12,00	13,00	4,62	44	13,27	14,50	4,87	4	6,50	8,00	4,51
Dim Publicar	19	6,37	6,00	3,37	44	7,05	8,50	3,51	4	8,25	9,50	2,87

	1 semana				1 mes				
	n	Media	Mediana	DT	n	Media	Mediana	DT	pvalor
Dim Datos	11	47,18	44,00	15,33	21	54,90	65,00	22,16	0,049
Dim Análisis	10	7,50	6,50	7,00	17	11,50	10,47	8,28	0,37
Dim Visualización	11	26,55	27,00	6,49	21	27,00	26,86	8,88	0,98
Dim Informar	11	10,36	9,00	4,61	21	12,00	13,05	4,90	0,04
Dim Publicar	11	6,36	8,00	4,11	21	6,37	7,10	3,48	0,88

Por ello podemos decir que **se contrasta estadísticamente la hipótesis para Informar y para el Sistema de datos debido a que se muestra un p-valor inferior a 0,05 para estas dimensiones**, pero no para el resto de las dimensiones del modelo.

Podría interpretarse de este resultado que disponer de informes es lo que influye más en el tiempo de respuesta, así como disponer del sistema de datos estructurados crudos con los que si es necesario se puede intentar dar respuesta. No parece tan claro que influya en los resultados por ejemplo el disponer de un portal de publicación de informes, aplicar técnicas estadísticas o componentes de visualización. Ello puede ser debido a que lo verdaderamente

relevante para poder responder rápidamente a las peticiones de información es que se disponga de un informe que contenga los datos correctos en el momento que se precisa o bien se tenga acceso directo a los datos.

Parece que lo que el directivo valora fundamentalmente para la toma de decisiones es el hecho de disponer del informe correspondiente, más que las técnicas de análisis o los componentes de visualización que se hayan aplicado.

El modelo SIEGO es un modelo incremental, en el que cada dimensión se construye sobre la anterior, por tanto no sorprende este resultado, al ser el fruto final del modelo el Sistema de Informes.

5.3 H3 Las organizaciones que disponen de datos perciben mejores resultados

Respecto al registro de datos, y su relación con la percepción de resultados, se aprecian diferencias significativas en la disposición de datos de planificación y seguimiento de la estrategia (mayor puntuación en el grupo que sí dispone de este tipo de datos), si bien los datos relativos a personas (nóminas, altas, bajas, formación, etc.) y a procesos/productos/servicios, se encuentran en el límite de la significación estadística.

	No disponen				Sí disponen				
	n	Media	Mediana	DT	n	Media	Mediana	DT	pvalor
Plan y seg. Estrategia	30	12,20	12,00	3,40	71	13,87	14,00	2,79	0,01
Personas	8	11,38	12,00	4,41	93	13,55	14,00	2,89	0,05
Eco y Fin	9	12,89	13,00	3,06	92	13,42	14,00	3,08	0,71
Gestión Cono y TI	40	12,78	13,00	3,44	61	13,77	14,00	2,75	0,11
Compras	22	12,36	12,50	2,75	79	13,66	14,00	3,10	0,07
Clientes	23	12,43	13,00	2,94	78	13,65	14,00	3,06	0,11
Procesos	23	12,26	12,00	3,15	78	13,71	14,00	2,98	0,05

En la tabla se aprecian diferencias significativas (p-valor=0.01<0.05) en la percepción de resultados si la empresa dispone de un registro de datos relativo a la planificación y seguimiento de la estrategia, obteniéndose resultados mejores si se dispone de tales datos (mayor puntuación media). En el resto de los casos sí se aprecian diferencias a nivel de muestra a favor siempre de disponer del registro de datos correspondiente, pero no se alcanza significación estadística, aunque para los datos sobre personas y sobre procesos el p-valor se encuentra en el límite de la significación. Por tanto, **se contrasta la hipótesis para las organizaciones que recogen datos sobre Planificación y Seguimiento de la Estrategia, sobre Personas y Sobre Procesos.**

A continuación se muestran los gráficos de caja que muestran la distribución de las respuestas para cada variable (diferentes tipos de datos) por cuartiles, diferenciándose si la organización recoge este tipo de datos o no. Se puede apreciar en todos los gráficos que la distribución de respuestas de los cuartiles 2 y 3 para las organizaciones que sí recogen datos, es en todos los casos más alta que las organizaciones que no los recogen.

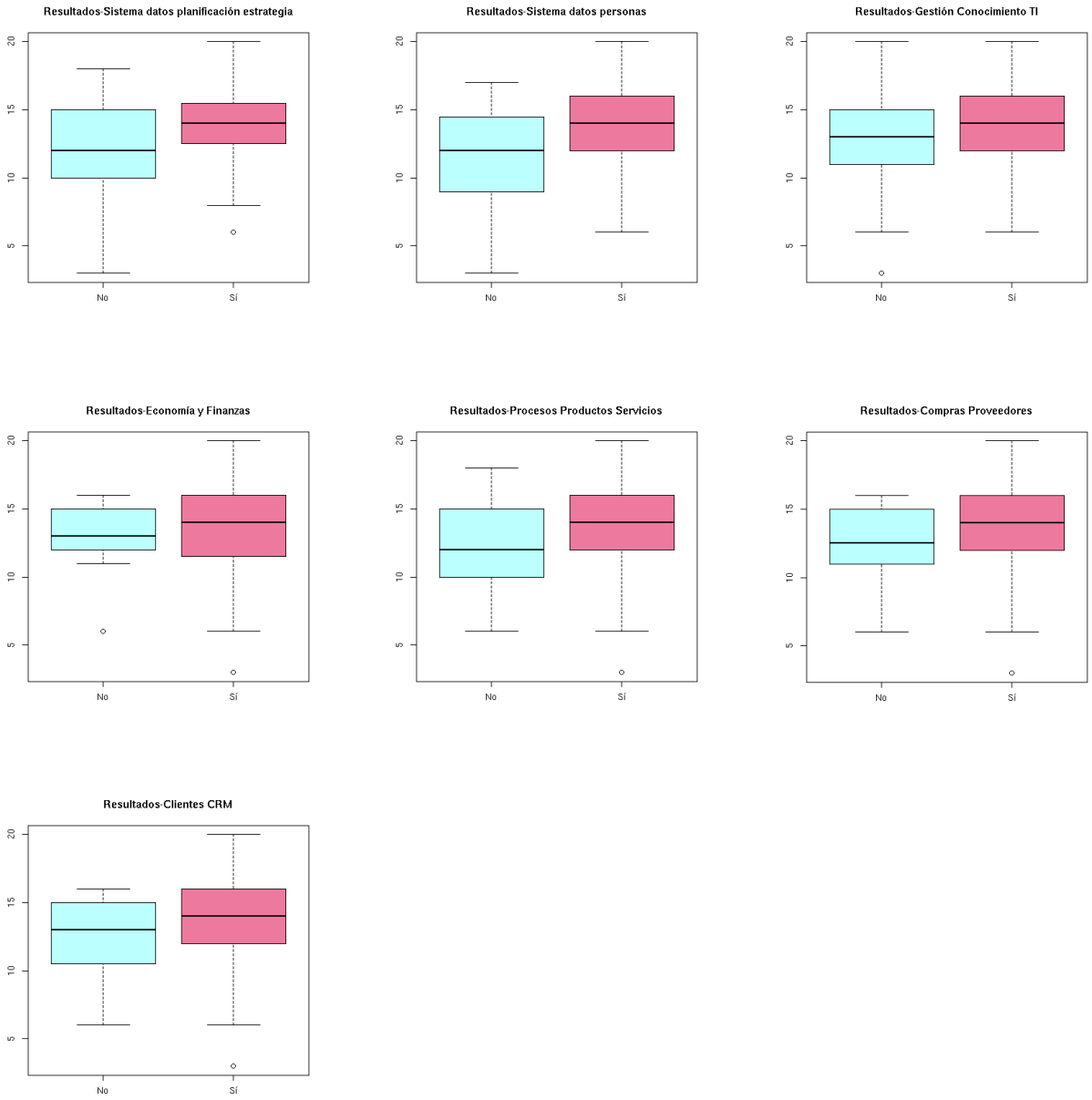


Ilustración 170. Gráficos de caja relación datos con resultados.

5.4 H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados

En esta hipótesis se pretende contrastar si se detectan diferencias entre la percepción de los resultados obtenidos según se tenga o no alguno de los sellos de calidad o excelencia: cuando se tiene, ¿se perciben mejores resultados?

En la tabla siguiente se calcula el p-valor para la dimensión resultados, diferenciando las organizaciones que han respondido que no disponen de ningún sello de calidad (Calidad No) y las que han declarado sí disponer de algún sello de calidad o excelencia (Calidad Sí).

	Calidad No				Calidad Sí				
	n	Media	Mediana	DT	n	Media	Mediana	DT	pvalor
Dim resultados	24	12,50	12,50	3,01	77	13,65	14,00	3,05	0,04

Se contrasta que para las organizaciones que Sí disponen de algún sello de calidad o excelencia la correlación presenta un p-valor inferior a 0,05 lo cual es estadísticamente significativo y por tanto **contrasta la hipótesis que las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados.**

A continuación se muestra el gráfico de cada de la distribución de las respuestas obtenidas para la correlación con la dimensión resultados, en el que se aprecia que la mediana para las que sí disponen de sello de calidad o excelencia es 14 mientras que para las que no disponen es de 12,50.

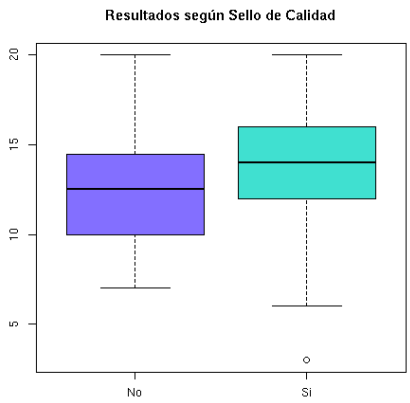


Ilustración 171. Gráfico Resultados - Sello de calidad

Si concretamos más la variable de segmentación, escogiendo únicamente las organizaciones que disponen de un sello EFQM (sea 300+, 400+ o 500+), vemos que se acentúa el p-valor siendo todavía más fuerte la significación estadística, presentando un p-valor de 0,001.

	EFQM No				EFQM Sí				
	n	Media	Mediana	DT	n	Media	Mediana	DT	pvalor
Dim resultados	77	12,86	13,00	3,18	24	15,04	15,00	1,88	0,001

Se verifica que las organizaciones que disponen de alguno de los sellos de Excelencia del modelo EFQM (300+, 400+ o 500+) perciben todavía mejores resultados que las

organizaciones que disponen de otros sellos de calidad, por ello el p-valor es todavía más significativo.

A continuación se muestra el gráfico de distribución segmentando sólo las organizaciones que disponen de alguno de los sellos de excelencia, en el que se aprecia que los resultados se concentran mucho más cerca de la mediana que se incrementa a 15.

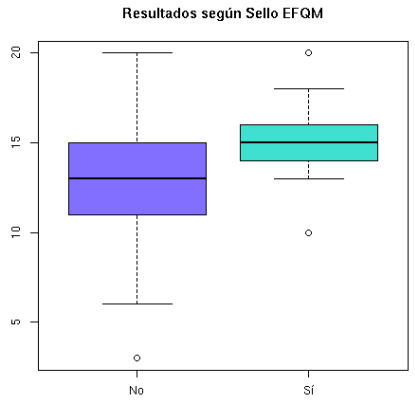


Ilustración 172. Resultados - Sello EFQM.

6 Análisis multivariante: modelo de regresión

Se construyó un modelo multivariante de datos, en concreto, un modelo de regresión lineal para identificar las variables asociadas a la percepción de resultados. Se consideró por tanto la percepción de resultados como variable dependiente en dicho modelo, y las variables asociadas a la dimensión de datos, análisis, visualizar, informar y publicar como variables independientes o predictoras. La tabla siguiente muestra los coeficientes de regresión, junto con su significación, asociados a las variables resultantes de aplicar un método de selección paso a paso. El modelo obtenido resultó significativo, y explica en torno a un 35% de la variabilidad de la percepción de los resultados ($R^2_{ajustado}$= 0,3474; p – valor < 0,001).

	Coeficiente	p-valor
Constante	6,74	<0,001
Dimensión Datos	0,12	<0,001

De esa forma, construiríamos la ecuación de regresión que podría servir para predecir los resultados de la organización en función de las variables obtenidas tras realizar el stepwise:

Dimensión Resultados = 6.74 + 0,12 Dimensión Datos

Este resultado indica que la dimensión más relevante y predictora para que la organización perciba mejores resultados es disponer de un sistema de datos estructurado.

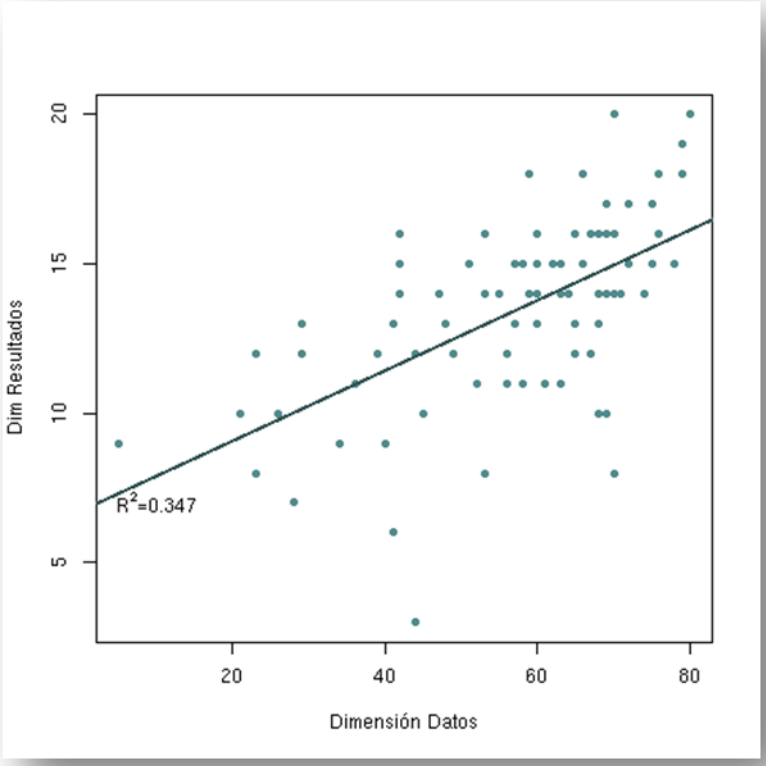


Ilustración 173. Gráfico de la función. Fuente: Elaboración propia (2016)

Ello comporta que el esfuerzo que la organización invierta en estructurar los datos, en tener un sistema de datos organizado y que muestre información de los criterios relevantes que se han definido en este trabajo, implicará una mejor percepción de los resultados de la organización.

Según este estudio, si la organización quiere mejorar la percepción de los resultados, puede hacerlo mediante la mejora del sistema de datos de que dispone.

Este resultado corrobora la experiencia en implantación de proyectos de Business Intelligence en los cuales la parte de definición de la estructura de los datos, Data Warehouse, Data Marts, esquemas, etc., a menudo es subestimada, pero finalmente consume una parte importante de los recursos y del tiempo del proyecto, y es la parte imprescindible para la calidad de los análisis e informes finales.

Capítulo V. Conclusiones

1 Objetivos

El trabajo realizado trataba de responder a la pregunta básica de cómo deberían ser los sistemas de información para ser útiles a las organizaciones y permitirles aumentar la eficiencia en su gestión. De la intensa revisión bibliográfica, así como de la experiencia personal de la doctoranda, se construyó un modelo formado por cinco elementos fundamentales: un modelo estructurado de datos que presenten la información y las relaciones entre los criterios de excelencia, un catálogo de técnicas analíticas para aplicarlas en la revisión y estructuración de los datos así como un inventario de componentes de visualización de datos. Con todo ello, se propone un sistema de informes corporativo para presentar el sistema de información a los grupos de interés de la organización que muestra los datos de la forma más eficiente e intuitiva para generar el mayor conocimiento, en el menor tiempo posible. Finalmente, un espacio común y único para publicar los informes, en el que las personas de la organización encuentren toda la información que precisan para su gestión, de forma estructura y organizada por temas, con una política de permisos y contenidos adecuados a cada perfil y cada grupo de interés.

El objetivo principal de la tesis doctoral, una vez planteado el modelo de sistema de información SIEGO, ha sido validar la adecuación de su estructura y de los contenidos que incorpora para las organizaciones que decidan implantarlo. Para ello se ha desarrollado una investigación empírica en la que se intenta contrastar si las organizaciones disponen de un modelo parecido, y de ser así qué relación existe con la percepción de resultados.

Igualmente, se proponía relacionar la rapidez en la respuesta a las solicitudes de información, con la existencia o no de un sistema de información como el que se ha propuesto, ya que esta cuestión se considera básica para que los procesos de toma de decisiones de la organización se ejecuten de forma eficiente.

Las preguntas de tesis que la investigadora se ha planteado para ello fueron las siguientes:

A. ¿Las organizaciones que recogen datos, aplican técnicas analíticas, componentes de visualización, preparan con ello informes y los publican, perciben mejores resultados?

   No se dispone de evidencia sobre la realidad práctica de las organizaciones con respecto a la recogida de datos y su uso para sistemas de información. Tampoco sobre la evidencia de hasta qué punto estos datos son utilizados para el análisis de la actividad de la organización y para la toma de decisiones.

   Debido a que los estudios científicos evidencian la conveniencia de medir ciertos aspectos fundamentales de la actividad de las organizaciones, se pretende conocer si las organizaciones recogen efectivamente datos de estos aspectos.

   Esta pregunta se ha investigado mediante el planteamiento de dos hipótesis:

H1 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, perciben mejores resultados.

H2 Las organizaciones que disponen de un sistema de datos, analizan, visualizan, informan y publican, pueden responder más rápido a las preguntas.

La siguiente pregunta de investigación se refiere a:

B. ¿Qué tipo de datos recogen las organizaciones de sus actividades y procesos?

Una consecuencia clara de la mejor toma de decisiones en una organización, sería la de obtener mejores resultados. Por ello se pretende determinar si existe relación o no entre la percepción de resultados y la disposición o no de un sistema de información parecido al que se plantea en este trabajo.

Debido a que la finalidad de este trabajo de investigación no es conocer el grado concreto de mejora de los resultados obtenidos, sino la percepción que sobre ellos tiene la persona que responde al cuestionario, no se formulan preguntas concretas sobre los resultados de la organización, sino preguntas sobre percepción de resultados en general y sobre el clima laboral.

Se da respuesta a esta pregunta mediante las hipótesis:

H3 Las organizaciones que disponen de datos de:

- Planificación y seguimiento de la estrategia
- Personas (nómina, altas, bajas, formación, etc.)
- Economía y finanzas (contabilidad, costes, etc.)
- Gestión del conocimiento y TI
- Compras / Proveedores
- Clientes / CRM
- Procesos / Productos / Servicios

Perciben mejores resultados.

H4 Las organizaciones que disponen de un sello de calidad o excelencia perciben mejores resultados.

2 Resultados contrastados

La evaluación del modelo se ha contrastado con las 101 respuestas recibidas al cuestionario lanzado, de organizaciones de distintos sectores económicos y tamaños, tanto públicas como privadas.

En relación con la calidad y la excelencia, se ha evaluado qué impacto tiene para los resultados de la organización la implantación de un sistema de información y se ha contrastado que además de ser relevante, esta importancia se acentúa si el modelo de gestión escogido es el modelo EFQM de excelencia.

Todas las partes del modelo muestran relaciones positivas de correlación, lo cual significa que si mejora una de las dimensiones también lo hará la dimensión relacionada. Las organizaciones que trabajen en la mejora de alguna de las dimensiones del modelo SIEGO mejorarán todas las demás dimensiones del modelo.

Con todo ello se ha validado que las organizaciones que dispongan de un sistema de información como el propuesto percibirán mejores resultados, excepto para la dimensión visualizar que no alcanza significación estadística.

También se puso en relación el modelo con la velocidad de respuesta de los responsables de la organización a cualquier petición de información de su ámbito de responsabilidad, viendo que son especialmente significativas la dimensión Informar y la del sistema de datos.

Otro resultado basado en las respuestas recibidas de las organizaciones, es que aquellas que midan su estrategia, registren datos sobre personas y sobre sus procesos, van a lograr percibir mejores resultados. Se ha visto que las organizaciones disponen de datos, sobretodo económico-financieros y de personas, pero otro tipos de datos (estratégicos o de procesos, por ejemplo) no son tan habituales. Por ello, se considera que existe cierto déficit de datos, estructurados y ordenados como plantea este trabajo. Así como se dispone de datos de ventas a clientes mayoritariamente, ya no es tan habitual disponer de datos de satisfacción de clientes, por ello se considera que existe un margen razonable de mejora de los sistemas de datos de las organizaciones.

Se ha validado que para mejorar la percepción de resultados de la organización, se percibe como positivo un sistema de información cuyas variables estén alineadas con los modelos de excelencia.

Sobre el impacto en los resultados de las organizaciones, este estudio ha evaluado qué prácticas relacionadas con el sistema de información son relevantes. En este sentido se toman como fundamentales las tareas de implantación de un Sistema de Datos, que va a tener un impacto positivo sobre el resto de dimensiones del modelo y sobre los resultados de la organización. Ello comporta que el esfuerzo que la organización invierta en estructurar los datos, en tener un sistema de datos de calidad y que muestre información de los criterios relevantes que se han definido en este trabajo, implicará una mejor percepción de los resultados de la organización

Según este estudio, si la organización quiere mejorar la percepción de los resultados, puede hacerlo mediante la mejora del sistema de datos de que dispone.

Este resultado corrobora la experiencia en implantación de proyectos de Business Intelligence en los cuales la parte de definición de la estructura de los datos, Data Warehouse, Data Marts, esquemas, etc., a menudo es subestimada, pero finalmente consume una parte importante de los recursos y del tiempo del proyecto, y es la parte imprescindible para la calidad de los análisis e informes finales.

Sobre todo lo mencionado anteriormente, se quiere resaltar la relevancia que tiene en los resultados de este estudio el perfil profesional de las personas que han respondido al cuestionario, que tienen entre 40 y 59 años en un 72% de los casos, son licenciados o tienen estudios de postgrado también en un 72% de los casos y gestionan un equipo de entre 10 y más de 200 personas en el 54% de los casos.

3 Conclusiones y Contribuciones

La principal contribución de este trabajo de investigación es el modelo de sistema de información para las organizaciones, que debería contribuir a la mejora de sus resultados y a la competitividad de la organización, puesto que facilita tomar decisiones con mayor agilidad.

Cuando una organización, sea del tipo que sea, se plantee la implantación o mejora de su sistema de información, podrá evaluar sus necesidades desde la perspectiva del modelo SIEGO. Debido a que el modelo se ha concebido como un conjunto de dimensiones o partes que permiten su funcionamiento de forma desacoplada pero a la vez integrada, la organización podrá valorar la implantación del sistema de forma total o parcial. Asimismo podrá realizar la implantación de forma asíncrona e independiente de soluciones propietarias, ya que existen en el mercado soluciones gratuitas que pueden cumplir a la práctica con los requerimientos de cada parte.

La información disponible sobre sistemas de información es muy variada y especializada, por lo que creemos que poder definir las dimensiones y las funcionalidades que deben analizarse para este tipo de soluciones es una aportación relevante.

Los sistemas de información, y por ello la recopilación de datos en la fase de medición, forman parte del capital de conocimiento explícito de la organización, que a su vez es capital estratégico, y por ello siempre que se sea posible no debería externalizarse. Es mucho más importante para llevar al sistema de información a un nivel de madurez mayor el conocimiento de este tipo de herramientas por parte del equipo de personas de la organización que lo van a gestionar, que el buscar una solución todo en uno, que la vinculará a una empresa externa mientras perviva el sistema.

Por lo que se refiere a las técnicas de análisis, se concluye que además de saber con qué objeto se analizan los datos, también es relevante quien va a realizar el análisis. Los análisis univariables y la mayoría de los bivariables pueden realizarlos personas con cierto conocimiento estadístico, pero no es necesario que sean personas especializadas. Cuando

hablamos del análisis multivariable, ya precisamos de un perfil profesional muy especializado y con un conocimiento estadístico profundo. La organización que se plantee evolucionar el conocimiento de las personas de su organización deberá vincular el uso de determinadas técnicas estadísticas con un plan de formación o selección debidamente alineado, o bien buscar apoyo externo en estas tareas.

Lo mismo sucede si hablamos de los recursos técnicos necesarios para ejecutar los análisis. En el caso de los análisis univariables y bivariables, estos análisis pueden realizarse con una herramienta al alcance de todas las organizaciones, como es MS Excel, pero para el análisis multivarible, en general es precisa una herramienta específica de análisis estadístico como R o SPSS.

En lo referente a los componentes de visualización, podría hacerse un paralelismo con las técnicas de análisis. Los componentes básicos de visualización de datos puede realizarlos cualquier perfil profesional con unos mínimos conocimientos y con un programa de uso sencillo. Cuando queremos aplicar componentes de visualización avanzados, ya precisamos de perfiles profesionales especializados y de herramientas específicas, menos habituales en las organizaciones.

Según se desprende de estos datos, claramente las organizaciones disponen de sistemas de informes para todo tipo de datos, sean del tipo que sean, dinámicos o estáticos. Las organizaciones han respondido que disponen de un sistema estructurado de informes y que utilizan sistemas de informes dinámicos.

Creemos que también es muy relevante la aportación del portal de informes, en el que cualquier persona de la organización pueda encontrar de forma rápida y estructurada los informes de cualquier tipo. En este portal deberíamos encontrar toda la información que precisemos para nuestra actividad así como aquella que nos permita responder a la mayoría de las peticiones de información. También es la dimensión que representa el final de la fase de revisión en un sistema de gestión básico, y que permitirá por tanto usar los informes del portal para planificar la mejora de la organización.

Se considera deseable que las herramientas utilizadas en cualquiera de los ámbitos del modelo SIEGO sean independientes entre sí; es decir, que cada dimensión pueda funcionar de forma totalmente autónoma y desacoplada del resto. Asimismo, cada elemento debe de estar integrado en el sistema de información de la organización adecuadamente. El motivo de esta propuesta es que el ritmo de avance y desarrollo de las herramientas que conforman un sistema de información como éste es desigual y el grado de substitución o agregación de otras herramientas debería poder hacerse con facilidad.

Se considera que la mejor solución es un modelo como SIEGO, que permitirá disponer de un sistema de información altamente flexible pero integrable con los sistemas de la organización. El inconveniente que presenta tomar esta opción es que precisa disponer de conocimientos técnicos más avanzados. Aun así, se entiende que un proyecto de desarrollo de un sistema de

información, por la importancia estratégica que tiene, debe diseñarse de manera que el control total sobre el sistema lo tenga la propia organización y no esté gobernado por un solo producto o proveedor, ni tampoco por una única empresa consultora especialista.

Si se escoge un solo producto, lo que se suele llamar una solución vertical, no seremos partícipes de las decisiones de su evolución, y puede darse el caso que nuestras necesidades como organización se desarrollen de forma diferente a las del mercado, y que en determinado momento nos encontremos con una aplicación que no cumple con nuestras necesidades ni expectativas.

Como conclusión final se puede afirmar que con esta investigación se han conseguido cumplir los objetivos planteados y se ha podido contrastar que el modelo propuesto puede ayudar a las organizaciones a definir su sistema de información para mejorar la percepción de sus resultados y para mejorar el tiempo de respuesta a las necesidades de información, lo cual les puede permitir disponer de un sistema de gestión más eficiente, y por tanto poder tomar decisiones más rápidas y efectivas.

Se ha contrastado también que existe una relación estrecha entre los datos que se recogen y la excelencia de las organizaciones, siendo por ello relevante qué datos se miden.

4 Líneas de futuro

Para la **dimensión datos**, se considera que puede investigarse en la búsqueda de aplicaciones que faciliten la creación de los Data Warehouses y los Data Marts, mediante soluciones informáticas inteligentes que seleccionen y obtengan los datos en las aplicaciones transaccionales y otras fuentes y sean capaces de proponer una primera estructura de datos. Si además esta solución tuviera una interface gráfica, para usuarios no expertos, podría conseguir que la construcción del Data Warehouse fuera más fácil, y no requiriera la implicación de técnicos especializados. Esta fase de implantación de los sistemas de información, suele ser infravalorada, y a la práctica consumir grandes recursos y alargarse mucho en el tiempo.

Para la **dimensión análisis**, creemos que puede la evolución debería producirse hacia soluciones tecnológicas más amigables que R o SPSS y hacia soluciones para todo tipo de usuarios, que de forma gráfica y muy asistida, elaboren análisis complejos que actualmente no son realizables si no es por perfiles muy especializados y herramientas concretas. Podría valorarse también, en función de la madurez del sistema de información de la organización, la implantación de nuevos análisis, en los que se englobarían los trabajos de predicción, prescripción y prevención.

Las limitaciones observadas en la implementación de la **dimensión visualización y la dimensión informes** del modelo, sugiere como una línea interesante de mejora futura la estandarización, que llevaría a una mejor compresión de la información. Es decir, ¿podría investigarse un

modelo de informe sobre personas con sus componentes adecuados, que sea válido para cualquier tipo de organización?, y para ventas?, y para procesos ?

Para la dimensión publicación, se cree conveniente investigar en un protocolo o sistema de normalización de la visualización de informes, de forma que cuando desde un informe del portal se quiera acceder a más detalle de datos, se pase al informe en la aplicación transaccional sin que el usuario ni siquiera lo perciba.

Queda también abierta una línea de normalización de los términos que se utilizan en este ámbito, los sistemas de información, puesto que términos como "informe", "componente", "cuadro de mando", etc., pueden significar cosas muy diferentes en cada organización. Por ejemplo, nos genera cierta duda que la mayoría de organizaciones respondan que disponen de un portal de informes como el que planteamos en este trabajo, y a lo mejor ello proviene de la diferente comprensión de lo que para el modelo SIEGO es un portal de informes y lo que las organizaciones entienden.

Finalmente, una de las líneas de investigación más importantes que creemos que debería abordarse en una investigación posterior es la valoración de la aportación del modelo SIEGO a la siguiente fase del ciclo de gestión básico: la mejora. El objetivo sería desarrollar metodologías para valorar las mejoras que genere el modelo y medir el valor que aportará a la organización esta mejora.

Bibliografia

AENOR (2006). ISO 15489-1 Información y documentación. Gestión de documentos. Parte 1: Generalidades. Madrid: AENOR.

AENOR (2008). ISO 9001. *Sistema* de *Gestión de la calidad. Requisitos.* Madrid: AENOR.

AENOR (1997). Norma ISO 10006: *Dirección de Proyectos*. AENOR.

AENOR (2004). Norma ISO 10017: *Orientación sobre las técnicas estadísticas para la Norma ISO 9001:2000.* AENOR.

AENOR (2003). Norma UNE 66175: *Sistemas de gestión de la calidad. Guía para la implantación de sistemas de indicadores*. AENOR.

AENOR (2009). Norma UNE-ISO/IEC 15939:2009 Ingeniería del software y sistemas. Procesos de medición (ISO/IEC 15939: 2007)

AENOR (2008). Norma ISO 9001: *Sistema de Gestión de la calidad. Requisitos.* Aenor.

Abbott, Lawrence (1955). *Quality and Competition: An Essay in Economic Theory*. New York: Columbia University Press.

Abraham, B. y Ledolter, J. (1983). *Statistical Methods for Forecasting*, 1a edición. New York: JohnWiley& Sons.

Acuña Echeverría, Hernán y Konow Hott, Hernán (1990). *Métodos y técnicas de investigación prospectiva para la toma de decisiones*. Santiago: FUNFUTURO.

Adam, E.; Corbert, L.; Flores, B.; Harrison, N.; Lee, T.S.; Rho, B.; Ribera, K.; Samson, D. y Wetbrook, R. (1997). "An International Study of Quality Improvement Approach and Firm Performance". *International Journal of Operations and Production Management*. Vol.17, No.9, pp.842-874.

Adriaans, P., & Zantinge (1996) D. Data mining. *Addision-Wesley, Harlow*.

Agrawal, R., & Srikant, R. (1994, September). Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB* (Vol. 1215, pp. 487-499).

Al-Mashari, Majed (2002). "Enterprise resource planning (ERP) systems: A research agenda". *Industrial Management & Data Systems*. Vol.103, No.1, pp.165-170.

Al-Mashari, Majed; Al-Mudimigh, Abdullah y Zairi, Mohamed (2002). "Enterprise resource planning: A taxonomy of critical factors". *European Journal of Operational Research*. Vol.146, No.3, pp.352-364.

Alvira Martín, Francisco Ricardo (1982). "La perspectiva cualitativa y cuantitativa en las investigaciones sociales". *Estudios de Psicología.* ll, pp.34-36.

Alvira Martín, Francisco Ricardo (1983). "Perspectiva cualitativa-Perspectiva cuantitativa en la metodología sociológica". *Revista Española de Investigaciones Sociológicas.* No.22, pp.53-75.

Amaya Amaya, Jairo (2007). *Gerencia: Planeación & Estrategia. Fundamentos, Modelo y software de Planeación*. Universidad Santo Tomas.

Anderson, M. y Sohal, A.S. (1999). "A study of the relationship between quality management practices and performance in small business". *International Journal of Quality and Reliability Management.* Vol.16, No.9, pp.859-877.

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*(1), 17-21.

Armistead, Colin; Pritchard, J. P. y Machin, S. (1999). "Strategic business process management for organizational effectiveness". *Long Range Planning*. Vol.32, No.1, pp.96-106.

Arriaza Gómez, A.J., Fernandez Palacín, F., López Sánchez, M.A., Muñoz Márquez, M., Pérez Sala, S., Sánchez Navas, A. (2008). *Estadística básica con R y R-commander*. Universidad de Cádiz.

ASQ, American Society for Quality. http://www.asq.org/

Atkinson, A. A.; Waterhouse, J. H. y Wells, R. B. (1997). "A stakeholder approach to strategic performance measurement". *Sloan Management Review*. Spring, pp.25-37. Cambridge.

Aznar, Antonio y Trivez, Francisco Javier (1993). *Métodos de predicción en economía. I: Fundamentos, Input-output, Modelos econométricos y métodos no paramétricos de series temporales*. Barcelona: Ariel.

Babbie, Earl (1996). Manual para la práctica de la investigación social. Bilbao: Desclee De Brouwer

Baeza-Yates, R., Rivera-Loaiza, C., & Velasco-Martín, J. (2004). Arquitectura de la información y usabilidad en la web. *El profesional de la información*,*13*(3), 168-178.

Bagozzi, R.P. y L.W. Phillips (1982). "Representing and testing organizational theories: a holistic construal". *Administrative Science Quarterly. Vol.* 27, No.3, pp.459-489.

Baranger, Denis (1992). *Construcción y análisis de datos*. Editorial Universitaria.

Barnes. Ralph M. (1980). *Motion and Time Study Design and Measurement of Work.* New York: Wiley.

Bas, E. (1999). *Prospectiva. Herramientas para una gestión estratégica del cambio*. Barcelona: Ariel.

Bass, Bernard M. (1981). *Stogdill's Handbook of Leadership: A Survey of Theory and Research*. New York: Free Press.

Bass, Bernard M. (1985). *Leadership and Performance Beyond Expectations*. New York: Free Press.

Bass, Bernard M. (1987). "Policy implications of a new paradigm of leadership". En *Military Leadership: traditions and Future Trends*. Pp.155-164. Conferencia en la *United States Naval Academy*.

Bass, Bernard M. y Avolio, B.J. (1990). "The implications of transactional and transformational leadership for individual, team and organizational development". *Organizational Change and Development*. Vol.4, pp.231-272.

Batista Foguet, Joan Manuel y Coenders Gallart, Germà (2000). *Modelos de ecuaciones estructurales*. Cuadernos de Estadística. Madrid: La Muralla.

Bearden, W.O.; Netemeyer, R.G. y Mobley, M.F. (1993). *Handbook of Marketing Scales. Multi-item Measures for Marketing and Consumer Behavior Research.* Newbury Park: Sage Publications.

Beer, S. (1966). *Decision and Control.* Chichester: Wiley.

Belbin, M. (1993). *Equipos directivos: El porqué de su éxito o fracaso*. Editorial Associates.

Benavides Velasco, Carlos Ángel y Quintana García, Cristina (2003). *Gestión del conocimiento y la calidad total*. Madrid: Díaz de Santos.

Bennis, W.G. y Schein, E.H. (1965). *Personal and Organizational Change Through Group Methods*. New York: John Wiley.

Berelson, Bernard (1952). "*Content Analysis". Handbook of Social Psychology.* Vol I. Nueva-York.

Berger, Gaston (1964). *Phénoménologie du temps et prospective*. Presses Universitaires de France.

Berné, C.; Pedraja, M. y Rivera, P. (1994). "Los determinantes de la actitud del consumidor ante la publicidad: un estudio exploratorio". *VI Encuentros de Profesores de Marketing.* San Sebastián, pp.191-202.

Bertin, J. (1973). Sémiologie graphique: Les diagrammes-Les réseaux-Les cartes.

Bettencourt, Lance A. y Ulwick, Anthony W. (2008). "El mapa Cliente-Centrado de la innovación. Revisión del negocio". *Harvard Business Review.* Vol.86, Fascículo 5.

Blaikie, Norman (1993). *Approaches to Social Enquiry.* Cambridge: Polity.

Bloom, B. *et al.* (1971). *Taxonomía de los objetivos de la educación: la clasificación de las metas educacionales: manuales y y II.* Traducción de Marcelo Pérez Rivas; prólogo del Profesor Antonio F. Salonía. Buenos Aires: Centro Regional de Ayuda Técnica: Agencia para el Desarrollo Internacional (A.I.D).

Bluet, J,C, y Zemor, J. (1970). *Prospective géographique: méthode et direction de recherches*. RevistaMetra. Vol.IX, No.1

Boehm, B. (1981). *Software Engineering Economics*. Prentice Halls.

Bodwell, Donald J. (2009). *High Performance Teams. Process Owner*. http://www.highperformanceteams.org/hpt_pown.htm. 6/8/2009.

Boland, Tony y Fowler, Alan (2000). "A systems perspective of performance management in public sector organizations". *The International Journal of Public Sector Management*. Vol.13, No.5, pp.417-446.

Bollen, K.A. (1989). *Structural Equations with Latent Variables.* Nueva York: John Wiley & Sons.

Booms, B. H. y Bitner, M. J. (1981). "Marketing strategies and organizations structure for services firms", en Donelly, J. y George, W.R. (eds) Marketing for services. American Marketing Association. Chicago. pp. 41-51.

Borysowich, Craig (2006). *Process Management vs Project Management*. http://it.toolbox.com/blogs/enterprise-solutions/process-management-vs-project-management-10579 http://it.toolbox.com/blogs/enterprise-solutions/process-management-vs-project-management-10579. 2/8/2009.

Boudon, R. y Lazarsfeld, P. (1985). *Metodologías de las ciencias sociales*. Vol.1. Barcelona: Laia.

Bourguignon, A., Mallerret, V., & Nørreklit, H. (2002). L'irréductible dimension culturelle des instruments de gestion: l'exemple du tableau de bord et du balanced scorecard. *Comptabilité-Contrôle-Audit*, *8* (3), 7-32.

Bourguignon, Annick y Jenkins, Alan (2004). "Management Accounting Change and the Construction of Coherence in Organizations: a Case Study". *ESSEC Working Papers.* DR 04003, ESSEC Research Center, ESSEC Business School.

Brignall, T.J. y Ballantine, J. (1996). "Interactions and trade-offs in multi-dimensional performance measurement". University of Alberta, Edmonton, paper presented at a Conference on Strategic Management Accounting.

Brown, M. (1996). *Keeping score: Using the right metrics to drive world class performance.* EEUU: Quality Resources.

Brull Alabart, E. (2011). Gestión de procesos en las organizaciones. *Arola editors*.

Brunet, Ignasi; Pastor, Inma y Belzunegui, Ángel (2002). *Tècniques d'investigació social. Fonaments epistemològics y metodològics*. Barcelona: Pòrtic.

Burrell, G. y Morgan, G. (1979). *Sociological Paradigms and Organizational Analysis*. Heinemann.

Bueno, E., & Azúa, S. (1998). Medición del capital intelectual: modelo Intelect.Madrid: Instituto Universitario Euroforum Escorial.

Bueno, E. (2005) «Génesis, evolución y concepto del Capital Intelec: Enfoques y modelosprincipales», Capital Intelectual nº 1, 8-19.

Bueno, **E.** (2002): Dirección estratégica basada en conocimiento: teoría y práctica de la nueva perspectiva. En Morcillo, P., Fernández, J. (Coords.): Nuevas claves para la dirección estratégica. Ariel Economía, Madrid.

Bueno, E. (2003): Modelo Intellectus: Medición y Gestión del Capital Intelectual, Documento Intellectus, nº5, CIC-IADE (UAM), Madrid.

Buskard, D., Mollot., & Mollot, M. (2000). Business Intelligence made easy.

Caballero, I., et al. (2007) A Data Quality Measurement Information Model based on ISO/IEC 15939. in 12th International Conference on Information Quality. 2007. MIT, Cambridge, MA.

Cadogan, J.W., Diamantopoulos, A. y Pahud, C. (1997). "Developing a measure of export market orientation: scale construction and cross-cultural validation". *26 EMAC Conference*, 232-251, Warwick.

Cairo, A. (2012). The Functional Art: An introduction to information graphics and visualization. New Riders.

Cairo, A. (2006). What should you show in a graphic. *Design Journal*, *99*, 30-33.

Camacho, Salvador (1997). *Planeación de Recursos Empresariales*. http://www.raldia.com/texto17.htm

Camisón, C. (1999a). "La Medición de los Resultados Empresariales desde una óptica estratégica: Construcción de un Instrumento de Medida a partir de un estudio Delphi y Aplicación a la Empresa Industrial Española en el período 1983.96". *Revista de Contabilidad y Tributación*. No.199, pp.201-266.

Camisón, C. (1999b). "Sobre cómo medir las competencias distintivas: un examen empírico de la fiabilidad y validez de los modelos multi-item para la medición de los activos intangibles". *The Iberoamerican Academy of Management, First International Conference.* Madrid, Diciembre.

Camisón, César; Cruz, Sonia y González, Tomás (2006). *Gestión de la calidad: conceptos, enfoques, modelos y sistemas*. Madrid: Pearson-Prentice Hall.

Campbell, D.T. y D.W. Fiske (1959). "Convergent and discriminant validation by the multitrait-multimethod matrix". *Psychological Bulletin. Vol.*56, pp.81-105.

Canales Cerón, Manuel (2008). *Metodologías de investigación social: Introducción a los oficios*. Chile: LOM, Ediciones

Canals, Agustí (2003). "La gestión del conocimiento". En: *Acto de presentación del libro* Gestión del conocimiento (2003: Barcelona) [en línea]. UOC. [Fecha de consulta: 20/05/16].<http://www.uoc.edu/dt/20251/index.html>

Cano, C. (1999). *Business Intelligence, decisiones de negocio basadas en tecnología: ruta crítica del negocio moderno*.

Carmines, E.G. y Zeller, R.A. (1979). *Reliability and validity assessment*. Beverly Hills: Sage.

Carmona Guillen, José Antonio (1977). *Los indicadores sociales hoy.* Madrid: Centro de investigaciones sociológicas.

Cauchick Miguel, Paulo A. (2005) "Evidence of QFD best practices for product development:a multiple case study", International Journal of Quality & Reliability Management, Vol. 22 Iss: 1, pp.72 – 82

Cea d'Ancona, Mº Ángeles (1999). *Metodología cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Síntesis Sociología.

Chakravarthy, B.S. (1986). "Measuring Strategic Performance". *Strategic Management Journal*. Vol.7, pp.437-458.

Chapman, R.; Murray, P.C. y Mellor, R. (1997). "Strategic Quality Management and Financial Performance Indicators". *International Journal of Quality & Reliability Management*. Vol.14, No.4, pp.432-448.

Child, J. (1972). "Organizational Structure, Environment and Performance: The Role of Strategic Choice". *Sociology*, 6, pp.1-21. C. L. en Cooper (Ed.) (2002), Ch. 35; In S. Clegg (Ed.) (2002), Ch.31.

Churchill, G.A. (1979). "A paradigm for developing better measures of marketing constructs". *Journal of Marketing Research*. Vol.XVI, febrero, pp.64-73.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods.*Journal of the American statistical association*, *79*(387), 531-554.

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, *229*(4716), 828-833.

Cleveland, W. C., & McGill, M. E. (1988). *Dynamic graphics for statistics*. CRC Press, Inc..

Cole, Robert E. (1994). "Reengineering the corporation: a review essay". *Quality Management Journal*. Vol.1, No.4, pp.77-85.

Comte, Augusto (1842). *Curso de filosofía positiva*. Madrid: Magisterio español, edición 2002.

Conant, J.S., M.P. Motwa y P.R. Varadarajan (1990). "Strategic types, distinctive marketing competences and organizational performance: A multiple measures-based study". *Strategic Management Journal.* Vol. 11, No.5, pp.365-383.

Corbetta, Piergiorgio (2003). *Metodología y técnicas de investigación social*. Madrid: McGrawHill.

Cos Castillo, Manuel de (1995). *Teoría general del Proyecto: Gestión de Proyectos*. Madrid: Sintesis.

Crawford, K.M y Cox, J.F. (1990). "Designing performance measurement systems for just-in-time operations". EEUU: *International Journal of Production Research*. Vol. 28, No.11.

Cronbach, L. J. (1943). On estimates of test reliability. *The Journal of Educational Psychology.* Vol.34, No.4, pp.485-494.

Cronbach, L. J. (1946). "A case study of the split-half reliability coefficient". *The Journal of Educational Psychology.* Vol.37, No.3, pp.473-480.

Cronbach, L. J. (1947). "Test "reliability": its meaning and determination". *Psychometrika.* Vol.12, No.1, pp.1-16.

Cronbach, L. J. (1951). "Coefficient alpha and the internal structure of tests". *Psychometrika.* Vol.16, No.2, pp.297-334

Crosby, Philip B. (1991). *Calidad sin lágrimas. El arte de administrar sin problemas.* Editorial Continental.

Cruz Ros, Sonia (2001). *Relación entre el enfoque de gestión de la calidad y el desempeño organizativo. Una aproximación a la perspectiva basada en los recursos.* Valencia: Universidad de Valencia. Tesis Doctoral.

Cuatrecases Arbos, Lluis (2006). *Claves De Lean Management: Un Enfoque Para La Alta Competitividad En Un Mundo Globalizado*. Gestión 2000.

Cuenca González, Llanos y Boza García, Andrés (2006). *Estudio comparativo de paquetes ERP*. X Congreso de Ingeniería de Organización. Valencia, 7 y 8 de septiembre.

Dale, Barrie G. (1994). *Managing Quality* (Second Edition). Hertfordshire: Prentice Hall.

Dale, Barrie G. y Bunney, Heather (1999). *Total Quality Management*. Blueprint. Blackwell Business.

Davenport, Thomas H. y Short, James (1990). "The New Industrial Engineering: Information Technology and Business Process Redesign". *Sloan Management Review*. Vol.31, n.4, pp.11-27.

Davenport, Thomas H. (1993). *Process Innovation. Reengineering Work through Information Technology.* Cambridge: Harvard Business School Press.

Davenport, Thomas H. (1998). "Putting the enterprise into the system". *Harvard Business Review*. Pp.121-131.

Davenport, T. H. y Prusak, L. (1998): Working Knowledge: How Organizations Manage What They Know. Cambridge, MA: Harvard Business School Press.

De Cos Castillo, Manuel (1995). *Teoría general del Proyecto: Gestión de Proyectos*. Madrid: Síntesis.

De Saa Perez, Petra (2000). "El valor estratégico de los recursos humanos según la visión de la empresa basada en los recursos". *Revista Europea de Dirección y Economía de Empresa*. Vol.9, No.2, pp. 97-116.

De Jouvenel, Bertrand (1964). *El arte de prever el futuro político*. Madrid: Ediciones RIALP (Edició consultada, 1966).

De Vellis, Robert F. (1991). *Scale development: Theory and applications*. Newbury Park,CA:Sage.

Deming, E. (1982). *Out of the crisis*. Cambridge University Press.

Deming, E. (1986). *Out of the crisis*. MIT Press.

Deming, W. Edwards (1989). *Calidad, Productividad y Competitividad. La salida de la crisis.* Madrid: Diaz de Santos.

Deming, E. (1993). *The Essential Deming Videotapes*. Cambridge MA: Center for Advanced Engineering Study. Massachusetts Institute of Technology.

Denzin, N.K. y Lincoln, Y. (1944). *Handbook of qualitative research*. Thousand Oaks, California: Sage.

Department of Management & Budget (2002). *Measures Up. Fairfax County.*

DeToro, Irving y McCabe, Thomas (1997). "How to Stay Flexible and Elude Fads". *Quality Progress*, March, pp.55-60.

Diez Medrano, Juan (1992). Métodos de análisis causal. Madrid: CIS.

Dixon, J.Robb; Arnold, Peter; Heineke, Janelle; Kim, Jay S. y Mulligan, Paul (1994). "Business Reengineering: Improving in New Strategic Directions". *California Management Review*, Summer, pp.93-108.

Dobles, C., Zúñiga, M. y García, J. (1998). *Investigación en educación: procesos, interacciones y construcciones*. San José: EUNED.

Doménech Massons, Josep M. (1982). *Bioestadística. Métodos estadísticos para investigadores.* Barcelona: Herder.

Domenech Massons, Josep M. y Riba. M.D. (1987). *Una síntesis de los métodos estadísticos bivariantes.* Herder.

Domínguez Amorós, Màrius y Coco Prieto, Andrés (2000). *Tècniques d'investigació social I*. Barcelona: Universitat de Barcelona.

Donati, Pierpaolo (1994). *Manual de sociología en la salud*. Díaz de Santos.

Dosi, G., Pavitt, K. y Soete, L. (1990). *The economics of technological change and international trade*. Brihton: Wheatsheaf/Harvester Press.

Druker, Peter F. (1996). *La gestión en un tiempo de grandes cambios*. Edhasa.

Dumond, E. J. (1994). "Making Best Use of Performance-Measures and Information". *International Journal of Operations & Production Management. Vol.*14, No.9, pp.16-32.

Duncan, J. (1984). 'Selective attention and the organization of visual information'. *Journal of Experimental Psychology: General,* No.114, pp.501-517.

Earl, Michael J. (1994). *Information Management. Cap 4: Business Process Reengineering: a phenomenon of organization*. Oxford Institute of Information Management.

Easterby-Smith, M., Thorpe, R. y Lowe, A. (2002). *Management research: An introduction*. London: SAGE Publications.

Easton, G.S. y Jarrell, S.L. (1998). "The Effects of Total Quality Management on Corporate Performance: an Empirical Investigation". *Journal of Business.* Vol.71, No.2, pp.253-307.

Eccles, R.G. (1991). "The performance measurement manifesto". *Harvard Business Review*. Vol.69, January-February, pp.131-137.

Eccles R.G. y Pyburn, P.J. (1992). "Creating a comprehensive system to measure performance". Management accounting. Vol.74, pp.41-45.

Eckerson, Wayne W. (2005). "Next-generation". Business Intelligence. February, 8.

Eckerson, W. W. (2005). What Are Performance Dashboards?. Information Management, 15(11), 26.

Edel Mendioca, Gloria (1998). Manual teórico-práctico de investigación social (Edición 2000). Buenos Aires: Espacio.

EFQM , European Foundation for Quality Management. www.efqm.org

Epstein, M, y Manazoni, J. (1998). "Implementing corporate strategy: from Tableaux de Bord to Balanced Scorecard". *European Management Journal. Vol.*16, No.2, pp.190-203.

Epstein, M. J., Kumar, P. y Westbrook, R.A. (2000). "The drivers of customer and corporate profitability: modeling, measuring, and managing the causal relationships". *Advances in Management Accounting*. Vol.9, pp.43-72.

Epstein, M. y Westbrook R. (2001). "Linking Actions to Profits in Strategic Decision Making". *MIT Sloan Management Review. Vol.*42, No.3, pp.39–49.

Escrig Tena, Ana Belén (2001). *Efectos de la dirección de calidad en los resultados: el papel mediador de las competencias distintivas*. Tesis Doctoral. Universidad Jaume I, Castellón, U.P. de Administración de Empresas y Marketing. Dirección: César Camisón Zornoza y Juan Carlos Bou Llusar.

Esteves J. y Pastor J. (2001). "Enterprise Resource Planning Systems Research: An Annotated Bibliography". *Communications of the Association for Information Systems* (CAIS). Vol.7, Article 8, August.

Etxeberría, Juan y Garcia, Eduardo (1995). *Análisis de datos y textos.* Madrid: Ra-ma.

Evans, James R. y Lindsey, William (2000). *Administración y control de la calidad*. Thomson.

Exteberria, Juan (1999). *Regresión Múltiple* (Ed. 2007). Cuadernos de Estadística. Madrid: La Muralla.

Frank, E., Chui, C., & Witten, I. H. (2000). Text categorization using compression models.

Fayol, Henry (1916). *Administración Industrial y General*. (Extraído del "Bulletin de la Société de l'Industrie Minérale"). Dunod Editeur, Paris.

Fayyad, U.; Haussuler, D. y Stolorz, P. (1996). "Mining scientific data". *Communications of the ACM*. Vol.39, pp.51-57.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, *17*(3), 37.

Feigenbaum, Armand V. (1951). *Quality Control: Principles, Practice and Administration*. McGraw-Hill.

Feigenbaum, Armand V. (1956). "Total Quality Control". *Harvard Business Review. Vol.*34, No.6, pp.93-101.

Feigenbaum, Armand V. (1961). *Total Quality Control. Engineering and Management*. New York: McGraw-Hill.

Fenton, Evelyn M. (2007*). Visualizing Strategic Change: The Role and Impact of Process Maps as Boundary Objects in Reorganization*. The University of Reading Business School.

Few, S., & EDGE, P. (2007). Data visualization: past, present, and future. *IBM Cognos Innovation Center*.

Few, S. (2007). Save the pies for dessert. *Visual Business Intelligence Newsletter*, 1-14.

Findibec¸ Fundación Iberoamericana para la Gestión de la Calidad. http://www.fundibeq.org/

Fitzgeral, L.; Johnston, R.; Brignall S.; Silvestro R. y Voss C. (1991). *Performance measurement in service business*. The Chartered Institute of Management Accountants (CIMA).

Flapper, Simme Douwe P., Fortuin, Leonard, Stoop, Paul P.M. (1996). "Towards consistent performance management systems". *INTERNATIONAL JOURNAL OF OPERATIONS & PRODUCTION MANAGEMENT.* Vol.16, No.7, pp.27-37.

Freeman, Ch. (1982). *The economics of industrial innovation*. Londres: Pinter.

Freeman, Ch. (1983). *Long Waves in the world economy*. Londres: Frances Printer.

Ghalayini, A.M. y Noble, J.S. (1996). The Changing Basis of Performance Measurement". *International Journal of Product Innovation Management*. Vol.16, No.8, pp.63-80.

Garcia Ferrando, Manuel; Ibañes, Jesús y Alvira, Francisco (1986). *El análisis de la realidad social. Métodos y técnicas de investigación* (3a edición, 2002). Madrid: Alianza.

García Ferrando, Manuel (1992). *Socio estadística. Introducción a la estadística en sociología*. 4ta. Reimpresión. Alianza.

Galtung, Johan (1966). *Teoría y Método de la Investigación Social*. Buenos Aires: EUDEBA.

Gangadharan, G. R. andSundaravalli N. Swami. (2004) "Business intelligence systems: design and implementation strategies." Information Technology Interfaces, 2004. 26th International Conference on Information Technology Interfaces ITI, June 2004.

Garcia Giménez, Eduardo; Gil Flores, Javier y Rodríguez Gómez, Gregorio (2000). *Análisis factorial*. Cuadernos de Estadística. Madrid: La Muralla.

Gartner. (2016). Josh Parenteau, Rita L. Sallam, Cindi Howson, JoaoTapadinhas, Kurt Schlegel, Thomas W. Oestreich Magic Quadrant for Business Intelligence and Analytics Platforms.https://www.gartner.com/doc/reprints?id=1-2XXET8P&ct=160204&st=sb

Gartner. (2006) Essential Components and Success Factors of Business Intelligence and Performance Management. Cannes, France: Gartner Symposium IT Expo 2006.

Garvin, D.A. (1988). *Managing Quality. The Strategic and Competitive Edge.* New York: The Free Press.

Georgoff, David M. y Murdick, Robert G. (1986). "Manager's guide to forecasting". *Harvard Business Review,* 64, Jan-Feb, pp.110-120.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.

George, L.K. (1992). *Economic Status and Subjective Well-being: a review of de Literature and an Agenda for Future Research*. New York: Springer.

Gil Flores, Javier; García Giménez, Eduardo y Rodríguez Gómez, Gregorio (2001). *Análisis discriminante*. Cuadernos de Estadística. Madrid: La Muralla.

Gilmore, Harold L. y Schwartz, H.C. (1986). *Integrated Product Testing and Evaluation*. New York: Marcel Dekker.

Giner de la Fuente, Fernando (2004). *Los sistemas de información en la sociedad del conocimiento*. Madrid: ESIC.

Gliem, J. A. and Gliem, R.R, (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales

Globerson, Shlomo (1985). "Issues in developing a performance criteria system for an organization". *International Journal of Production Research. Vol.*23, No.4, pp.639-646.

Godet, Michel (1979). *The crisis in forecasting and the emergence of the prospective approach*. New York: Pergamon Press.

Godet, Michel (1993). *De la anticipación a la acción. Manual de Prospectiva y Estrategia.* Barcelona: Editorial Marcombo.

Gopal, C., & Gagnon, J. (1995). Knowledge, information, learning and the IS manager. *Computerworld, 29*(25), 1–7.

Grant, Robert (1991). "The resourced-based theory of competitive advantage: implications for strategy formulation". *California Management Review.* Vol.33 No.3, Spring.

Grant, Robert M. (1996). *Dirección Estratégica. Conceptos, técnicas y aplicaciones*. Madrid: Cívitas. Primera edición, 1995.

Grapentine, T. (1994). "Problematic scales: when measuring quality, expectations scales exhibit several drawbacks". *Marketing Research. Vol.*6, No.4, pp.8-12.

Guba, E. G. (1978). *Toward a methodology of naturalistic inquiry in educational evaluation. Monograph 8*. Los Angeles: UCLA Center for the Study of Evaluation.

Guerra-López, Ingrid (2007). *Evaluación y mejora continua: conceptos y herramientas para la medición y mejora del desempeño*. Authorhouse.

Gulledge, T.R. y Sommer, R.A. (2003). "Public Sector Enterprise Resource Planning". *Industrial Management and Data Systems.* Vol.103, n.7, pp.471-483.

Gutiérrez, L. (1996). "Paradigmas cuantitativo y cualitativo en la investigación socio-educativa: proyecciones y reflexiones". *Revista PARADIGMA.* Vol. XIV al XVII.

Hackney, Douglas K. (2000). *Your Business Intelligence arsenal. Telephony.*

Hair, Joseph F. Jr; Anderson, Rolph E.; Tatham, Ronald L. y Black, William C. (1999). *Análisis multivariante* (Consultada la 5a edición). Madrid: Prentice Hall.

Hall, R. (1992). "The Strategic Analysis of Intangible Resources". *Strategic Management Journal.* Vol.13, pp.135-144.

Hall, R. (1993). "A Framework Linking Intangible Resources and Capabilities to Sustainable Competitive Advantage". *Strategic Management Journal.* Vol.14, pp.607-618.

Hamel, G. y Prahalad, C. (1994). *Competing for the Future*. Boston: Harvard Business School Press.

Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. MIT press.

Harrington, H. James y Harrington, James S. (1995). *Total Improvement Management. The next generation in performance improvement.* New York: McGraw Hill.

Harrington, H. James (1995). "The new model for improvement – Total improvement management". *Business process re-engineering and management journal*. Vol.1, No.1, pp.31-43.

Harrington, H. James (1998). "Performance improvement: the rise and fall of reengineering". *The TQM Magazine*. Vol.10, No.2, pp.69-71.

Hernández, José Antonio. (1998). *SAP R/3*. Santafé de Bogotá: Ed. Mc GrawHill.

Hibbard, J. (1997). "Knowing What We Know". *Information Week.* Oct,20.

Hilson, G. (2001). *BI market fraught with instability*. Computing Canada.

Hindle, John (1997). Process improvement and information management. *Health Manpower Management*. Vol.23, No.5, pp.184-186.

Hope, Jeremy (2004). *Más allá del sistema presupuestario: Beyond budgeting*. Barcelona: Deusto.

Hope, Jeremy (2006). *El nuevo papel de los directores financieros*. Barcelona: Deusto.

Horine, Gregory M. (2005). *Gestión de proyectos*. Madrid: Anaya.

Hosotani, Katsuya (1993). *Japanese Quality Concepts: An Overview*. Quality Resources.

Hronec, Steven M. (1993). *Vital Signs: Using Quality, Time, and Cost Performance Measurements to Chart Your Company's Future*, Arthur Andersesn/AMACOM (version castellano McGraw Hill, 1995).

Inmon, William. H. (1992). *Building the data warehouse*. Wiley-QED.

Inselberg, A. (1985). The plane with parallel coordinates. *The visual computer*,*1*(2), 69-91.

Ittner, C.D. y Larcker, D.F. (1997). "Quality Strategy, Strategic Control Systems and Organizational Performance". *Accounting, Organizations and Society*. Vol.22, No.3/4, pp.293-314.

Juran, J.M. (1990). *Manual del Control de la Calidad*. Barcelona: Ed. Reverte.

JUSE, Union of Japanese Scientists and Engineers. http://www.juse.or.jp/e/

Kaidos, Will (1991). *Sistemas de mediciones para maximizar la productividad y la calidad.* Oregon: Productivity. (TGO Hoshin, Madrid, 2002).

Kald, M. y Nilsson, F. (2000). "Performance Measurement at Nordic Companies". *European Management Journal.* Vol.18, No.1, pp.113-127.

Kaplan, Robert S. y Norton, David P. (1997). *El Cuadro de Mando Integral.* Barcelona: Gestión 2000.

Kaplan, Robert S. y Norton, David P. (2000). "Having Trouble with Your Strategy? Then Map It". *Harvard Business Review*, September-October, pp.167-176.

Kaplan, Robert S. y Norton, David P. (2001a). "Transforming the Balance Scorecard from Performance Measurement to Strategic Management: Part I". *American Accounting Association. Accounting Horizons*. Vol.15, No.1, pp.87-104.

Kaplan, Robert S. y Norton, David P. (2001b). "Transforming the Balance Scorecard from Performance Measurement to Strategic Management: Part II". *American Accounting Association. Accounting Horizons*. Vol.15, No.2, pp.147-160.

Kasperskaya, Yulia (2006). *Essays on Causal Performance Measurements Models*. Barcelona: UPF.

KDnuggets Polls, K.(2008) Industries/Fields where you applied data mining, June 3-16, 2008.[en línea][Consulta: 4 de noviembre de 2009]. *Disponible en Web: http://www. kdnuggets. com/polls/2008/data_mining_applications. htm*.

Kennerley, M., Neely, A. (2002). "A framework of the factors affecting the evolution of performance measurement systems". *International Journal of Operations & Production Management*. Vol.22, No.11, pp.1222-1245.

Kilman, Ralph H. (1995). "A Holistic Program and Critical Success Factors of Corporate Transformation". *European Management Journal.* Vol.13, No.2, pp.175-186.

Kimball, Ralph et al. (1998). *The Data Warehouse Lifecycle Toolkit.* New York: Wiley.

Kimball R, J. Caserta, (2011) "The Data Warehouse ETL Toolkit: Practical Tecniques for Extrating, Cleaning, Conforming, and Delivering Data, WILEY, VERSIÓN Ebook.

Kimball, R. (2008) The Data Warehouse Lifecycle Toolkit (2nd ed.)

Kochan, T.A. y Osterman (1994). "The Mutual Gains Enterprise: Forging a Winning Partnership among Labour". *Management and Government.* Boston: Harvard University Press.

Kolakowski, L. (1988). *La filosofía positiva*. Madrid: Cátedra.

Korn, Francis (1984). *Conceptos y variables en la investigación social*. Buenos Aires: Nueva visión.

Kourtit, Karima y de Waal, Andre A. (2006). *Strategic performance management in practice: advantages, disadvantages and reasons for use*. VrigeUniversiteitAmsterdam&Maastricht School of Management, the Netherlands. E-mail: andredewaal@planet.nl

Kosslyn, S. M. (1980). *Image and mind*. Harvard University Press.

Kosslyn, S. M. (1996). *Image and brain: The resolution of the imagery debate*. MIT press.

Krausse *(1995).* "La investigación Cualitativa: Un Campo de Posibilidades y Desafíos". *Revista Temas de Educación.* Nº 7, pp.19-38.

Kuhn, Tomas S. (1971). *La estructura de las revoluciones científicas*. México: Fondo de Cultura Económica.

Lado, N. (1996). "Tipos estratégicos, orientación al mercado y resultados económicos". *Cuadernos de la Fundación Mapfre Estudios.* No.31.

Langley; Nolan; Nolan, Norman y Provost (1996). *The improvement guide*. San Francisco: Jossey-Bass.

Lapsley, I. y Pallet, J. (2000). "Accounting, management and organizational change: a comparative study of local government". *Management Accounting Research*. Vol.11, pp.213-229.

Lawer, Edward E. (1996). *Ventaja creando organizaciones participativas e innovadoras*. Granica.

Leal, A. (1997). "Gestión de la Calidad Total en Empresas Españolas: Un análisis Cultural y de Rendimiento". *Revista Europea de Dirección y Economía de la Empresa.* Vol.6, No.1, pp.37-56.

Lee, R.G. y Dale, B.G. (1998). "Business process management: a review and evaluation". *Business Process Management Journal*. Vol.4, No.3, pp.214-225.

Lee, T.S.; Adam, E. y Tuan, C. (1999). "The convergent and predictive validity of quality and productivity practices in Hong Kong Industry". *Total Quality Management*. Vol.10, No.1, pp.73-84.

Lengler, R., & Eppler, M. J. (2007, January). Towards a periodic table of visualization methods for management. In IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007), Clearwater, Florida, USA.

Lemak, D.J. y Reed, R. (1997). "Commitment to Total Quality Management: Is There a Relationship with Firm Performance?".*Journal of Quality Management*. Vol.2, No.1, pp.67-87.

Levy Mangin, Jean-Pierre (2005). *Análisis multivariable para las ciencias sociales*. Madrid: Pearson Prentice Hall.

Levis, M., M. Helfert, and M. Brady. (2007) Information Quality Management: Review of an Evolving Research Area. ICIQ'07. MIT, Cambridge, MA, USA.

Loo, R. (2000). A psychometric evaluation of the general decision-making style inventory. *Personality and Individual Differences*, *29* (5), 895-905.

Luhn, Hans Peter. (1958) "A business intelligence system." IBM Journal of Research and Development Vol. 2.4, pp. 314-319.

Lynch, Richard L. y Cross, Kelvin F. (1991). *Measure Up!: Yardsticks for continuous improvement*. Cambridge: Blackwell Publishers. Consultada la versión en castellano, 'La mejora continua. Patrones de medida' editada per Deusto, el año 1993.

Makridakis, Spyros y Wheelwright, Steven C. (1989). *Forecasting methods for management*. Fifth Edition New York: Wiley.

Malhotra, N.K. (1981). "A scale to measure self-concepts, person concepts, and product concepts". *Journal of Marketing Research. Vol.*18, No.4, pp.456-464.

Malhotra, N.K. (1997). *Investigación de Mercados. Un enfoque práctico.* 2ª ed., México: Prentice Hall.

Manheim, Jarol B. y Rich, Richard C. (1988). *Análisis político empírico. Métodos de investigación en ciencia política*. Madrid: Alianza Universidad Textos.

Mann, R. y Kehoe, D. (1994). "And Evaluation of the Effects of Quality Improvement Activities on Business Performance". *International Journal of Quality & Reliability Management.* Vol.11, No.4, pp.29-44.

Marshall, C. y Rossman, B. (1989). Designing qualitative research. Newbury Park, California: Sage.

Martín, J. (2001). Alien Intelligence. The Journal of Business Strategy.

Martín, R. (1997). "Do we practice Quality Principles in the Performance Measurement of Critical Success Factors? *"Total Quality Management*. Vol.8, No.6, pp.429-444.

Martinez Arias, Rosario (1999). *El análisis multivariante en la investigación científica*. (Consultada edición 2008). Madrid: La Muralla.

Mayntz, Renate; Holm, Kurtz y Hubner, Peter (1988). *Introducción a los métodos de la sociología empírica*. Madrid: Alianza.

McColl-Kennedy, J.R. y Fetter, R.E., jr.(1999). "Dimensions of consumer search behavior in services". *The Journal of Services Marketing*. Vol.13, No.3, pp.242-265.

McGeever, C. (2000). *Business Intelligence*. Computer World.

Mensch, Gerard (1979). *Stalemate in technology: Innovations overcome the depression*. Cambridge, MA: Ballinger.

Mintzberg, Henry (1984). *La Estructuración de las Organizaciones* (4a reimpresión, 1995). Barcelona: Ariel.

Michel, R. (2000). Business Intelligence solutions. Manufacturing Systems.

Miquel, S.; Bigné, E.; Lévy, J.P.; Cuenca, A.C. y Miquel, M.J. (1997). *Investigación de Mercados*. Madrid: McGraw-Hill.

Merello, Agustín (1973). *Prospectiva. Teoría y práctica.* Buenos Aires: Ed. Guadalupe.

Molina Félix, Luis Carlos (2002). *Data mining: torturando a los datos hasta que confiesen*. FUOC.

Morales Vallejo, Pedro; Urosa Sanz, Belén y Blanco Blanco, Ángeles (2003). *Construcción de escalas de actitudes tipo Likert*. Cuadernos de Estadística. Madrid: La Muralla.

Morales Vallejo, Pedro (2006). *Medición de actitudes en psicología y educación*. Madrid: Universidad Pontifica Comilla.

Morcillo Morcillo, P. (1997). *La Dirección Estratégica de la Tecnología e Innovación.* Madrid: Civitas.

Moreno Albarracin, Antonio Luis (2004). "La implantación de indicadores de gestión para optimizar la gestión de las entidades no lucrativas: el modelo de la ONCE". En Revesco, *Revista de Estudios Cooperativos*. AECOOP.

Muñiz, J. (1996). *Teoría clásica de los test*. Madrid: Editorial Pirámide.

Muñiz, José; Fidalgo, Ángel M.; García-Cueto, Eduardo; Martínez, Rafael y Moreno, Rafael (2005). *Análisis de los ítems*. Cuadernos de Estadística. Madrid: La Muralla.

Nanni, Alfred L. Jr., Dixon, J.Robb y Vollmann, Thomas E. (1992). "Integrated Performance Measurement: Management Accounting to Support the New Manufacturing Realities". *Journal of Management Accounting Research. Vol.* 4, pp.1-19.

Navarro Galera, A. (1998). *El control económico de la gestión municipal: un modelo basado en indicadores*. Valencia: Sindicatura de comptes.

Neely, Andy D., Gregory, Mike y Platts, Ken. (1995). "Performance measurement system design: A literature review and research agenda". *International Journal of Operations & Production Management.* Vol.15, No.4, pp.80-116.

Neely, Andy D.; Huw, Richards; Mills, John; Platts, Ken y Bourne, Mike (1997). "Designing performance measures: a structured approach". *International Journal of Operations & Production Management. Vol.*17, No.11, pp.1131-1152.

Neely, Andy D. (1999). "The performance measurement revolution: why now and what next?" *International Journal of Operations & Production Management*. Vol.19, No.2, pp.205-228.

Neely, Andy D., Adams, C. (2000). *Perspectives on Performance: The Performance Prism*. London: Gee Publishing.

Neely, Andy D., Adams, C., Crowe, P. (2001). "The Performance Prism in Practice". *Measuring Business Excellence. Vol.*5, No.2, pp.6-12.

Neely, Andy D. (2002). *Business Performance Measurement: Theory and Practice*. Interfaces.

Neely, Andy; Adams, C. y Kennerly, M. (2002). *The Performance Prism: The Scorecard for Measuring and Managing Business Success*. London: FT Prentice-Hall.

Neely, Andy D. (2008). *Business performance Measurement. Unifying Theory and Integrating Practice*. Cambridge.

Nelson, Richard y Winter, Sydney G. (1982). *An Evolutionary Theory of Economic Change*. Belknap/Harvard University Press.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J.C. y Berstein, I.H. (1994). *Psychometric Theory.* 3ª ed. USA: McGraw Hill

Nurosis, M.J. (1993). *SPSS. Statistical Data Analysis*. SPSS Inc.

Olias de Lima, Blanca (2001). *La nueva gestión pública*. Madrid: Prentice Hall.

Oliveira, P., et al. (2005). A Taxonomy of Data Quality Problems. in Second International Workshop on Data and Information Quality (in conjunction with CAISE'05).Porto, Portugal.

Ortigueira Bouzada, M (1987). *Administraciones públicas: teoría básica de las auditorias de gestión*. Granada: Publicaciones del CUR.

Ortiz de Urbina, Gimeno, Iñigo (2007). *La excusa del positivismo: la presunta superación del positivismo y el formalismo por la dogmática penal contemporánea*. Cívitas.

Oakland, J.S. (1989). *Total Quality Management*. London: Heinemann Professional.

O'Donovan, T.M. (1983). *Short term Forecasting. An introduction to the Box-Jenkins Approach*. New York: John Wiley and Sons.

Padilla García, J.L.; Gómez Benito, J ; Hidalgo Montesinos, M.D. y Muñiz Fernández, J. (2007). "Esquema conceptual y procedimientos para analizar la validez de las consecuencias del uso de los test". *Psicothema*. No.19, pp.173-178.

Parmenter, David (2007). *Key Performance Indicators: Developing, implementing and Using Winning KPIs*. John Wiley & Sons.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods* (2nd ed.). Newbury Park, CA: Sage Publications, Inc.

Pedoe, Dan. (1979) La geometría en el arte, G. Gili.

Pérez-Montoro, M. (2010). Arquitectura de la información en entornos web. *El profesional de la información*, *19*(4), 333-337.

Peter, J.P. y G.A. Churchill (1986). "Relationship among research design choices and psicometric properties of rating scales: a meta-analysis". *Journal of Marketing Research. Vol.*23, No.1, pp.1-10.

Peters, Tom (2005). *50 Claves para hacer de usted una Marca*. Barcelona: Deusto.

Peterson, R.A. (1994). "A meta-analysis of Cronbach's Coefficient Alpha". *Journal of Consumer Research.* 21 (September), pp.381-391.

Pfeffer, J. (1994). *Competitive Advantage through People*. Boston: Harvard Business, School Press.

Piateski, G., & Frawley, W. (1991). *Knowledge discovery in databases*. MIT press.

Playfair, W., Wainer, H., & Spence, I. (2005). *Playfair's commercial and political atlas and statistical breviary*. Cambridge University Press.

Plaz Landaeta, R. (2005). Gestión del conocimiento: una visión integradora del aprendizaje organizacional.

PM Institute (1998). *Guía de los Fundamentos de la Dirección de Proyectos.* AEIPRO. Zaragoza.

Popper, K.R. (1967). *La Lógica de la Investigación Científica*. Madrid: Tecnos.

Popper, Karl (1997). *El Mito del Marco Común.* España: Editorial PAIDOS, 1ª edición.

Porter, Michael (1987). *Ventaja competitiva. Creación y sostenimiento de un desempeño superior.* Traducción: de la Campa de la Pérez, María Ascensión. México: Compañía Editorial Continental.

Powell, T.C. (1995). "Total Quality Management as Competitive Advantage: a Review and Empirical Study". *Strategic Management Journal*. Vol.16, pp.15-37.

Prahalad, C. y Hamel, G. (1990). "The Core competencies of the organization". *Harvard Business Review*. Vol.63, n.3, pp.79-91.

Pritchard, Robert D. (1990). *Measuring and Improving Organizational Productivity. A practical guide*. Greenwood Publishing Group.

Provost, L. y Leddick, S. (1993). "How to take multiple measures to get a complete picture of organizational performance". *National Productivity Review.* Vol.1993, pp.477-490.

Pulido San Roman, Antonio (1989). *Predicción Económica y Empresarial*. Madrid: Pirámide.

Putterman, Luis (1994). *La naturaleza económica de la empresa*. Madrid: Alianza.

Quick, J.H.; Shea, W.J. y Koehler, R.E. (1945). *Motion-Time Standards*. en: Factory, Management and Maintenance. Pp.97-108.

Ramió, C. y Ballart, X. (Comps.) (1993). *Lecturas de teoría de la organización* (2 vols). Madrid: M.A.P.

Ramirez Correa, Patricio y García Cruz, Rosario (2005). "Meta-análisis sobre la implantación de sistemas de planificación de recursos empresariales (ERP)". *Revista de Gestao da Tecnologia e Sistemas de Informaçao*. Vol.2, No.3, pp.245-273.

Rao, A., *et al.*, (1996). *Total Quality Management. A Cross-Functional Perspective.* New York: John Wiley.

Riedman, Wolfgang (1981). *Técnicas de dirección*. Madrid: Paraninfo.

Robbins, N. B. (2006). Dot Plots: A Useful Alternative to Bar Charts. *Business Intelligence Network Newsletter*.

Rockart, John y Short, James (1989). "IT in the 1990's: Managing Organizational Interdependence". *Sloan Management Review*, pp.7-18.

Rodriguez Cano, J. (2009). *Técnicas cuantitativas*. Universidad Viña del Mar.

Rodríguez Gómez, David. (2006). Modelos para la creación y gestión del conocimiento: una aproximación teórica Universitat Autònoma de Barcelona. Educar 37, 25-39.

Rodríguez Jaume, María José (2000). *Modelos sociodemográficos: Atlas Social de la Ciudad de Alicante*. Tesis doctoral. Universidad de Alicante.

Rodríguez Rodríguez, Raúl; Alfaro Saiz, Juan José y Ortiz Bas, Ángel (2006). *Matriz dinámica de rendimiento para PYMES*. X Congreso de Ingeniería de Organización.

Roos, J. et Al. (1997): *Intellectual capital: navigating in the new business landscape,* Macmillan, Houndsmills.

Roos, G., Dragonetti, N. (1998): La evaluación de Ausindustry y el business network programme: una perspectiva desde el capital intel·lectual Boletín de Estudios Económicos 53.

Ronda León, R. (2005). La Arquitectura de la Información y las Ciencias de la Información. *No Solo Usabilidad*, (4).

Rosenberg, N. (1994). "Incertidumbre y cambio tecnológico". *Revista de Historia Industrial* 6.

Rouse, Paul y Putterill, Martin (2003). "An integral framework for performance measurement". *Management Decision. Vol.*41, No.8, pp.791-805.

Rummler, Geary A. y Brache, Alan P. (1990). *Improving Performance: How to manage the white space in the organizational chart*. San Francisco: Jossey-Bass Publishers. (Versión castellano de 1997, Deusto-Biblioteca Deusto).

Russel, R. (2003). *The international Perspective: Balanced Scorecard Usage in Europe.* Harvard Business School Publishing.

Sánchez, M. y Sarabia, F.J. (1999). "Validez y fiabilidad de las escalas", en Sarabia, F.J. (coord). *Metodología para la investigación en marqueting y dirección de empresas.* Madrid: Pirámide. pp.363-394.

Sarabia, F.J. (1999). "Construcción de escalas de medida", en Sarabia, F.J. (coord). *Metodología para la investigación en marqueting y dirección de empresas*. Madrid: Pirámide. pp.331-355.

Schumpeter, Joseph A. (1912). *Theory of Economic Development*. (Traducción al inglés de 1934, Harvard. Reimpreso 1961 y 1978 en Oxford).

Senge, Peter M. (1990). *The fifth Discipline.* Edición 1999. Barcelona: Granica.

Shaw (1994). "Prospects for an engineering discipline of software". *Software Engineering Encyclopedia*. IEEE, 1994, pp.930-940.

Shewhart, Walter Andrew (1931). *Economic control of quality of manufactured product*. New York: D. Van Nostrand Company, pp.501.

Shewhart, Walter Andrew (1939). *Statistical method from the viewpoint of quality control.* Washington: The Graduate School, the Department of Agriculture, pp.155.

Sierra Bravo, Restituto (1994). *Análisis estadístico multivariable. Teoría y ejercicios*. Madrid: Paraninfo.

Sierra Bravo, Restituto (1998). *Técnicas de investigación social. Teoría y ejercicios*. Madrid: Paraninfo.

Sierra Bravo, Restituto (1998). *Tesis doctorales y trabajos de investigacióncientífica.* Madrid: Paraninfo.

Silverman, David (1993). *Interpreting qualitative data. Methods for analyzing talk, text and interaction.* Thousand Oaks: Sage.

Simon, Herbert (1947). *El comportamiento administrativo*. España: Aguilar.

Sink, D. Scott y Tuttle, Thomas C. (1990). "The performance management question in the organization of the future". *Industrial Management.* January-February, pp.4-12.

Sink, D. Scott y Porter, David E. (1999). "Get best results". *IIE Solutions,* October, pp.22-27, http://solutions.iienet.org

Smith, Adam (1776). (*Ensayo sobre la naturaleza y la causa de) La riqueza de las naciones*. Alianza Editorial, edición 2001.

Smith, Howard y Fingar, Peter (2003). *Business Process Management: the third wave. The breakthrough that redefines competitive advantage for the next fifty years*. Tampa, Florida, USA: Megahan-Kiffer Press.

Smith, H.; Neal, D.; Ferrara, L. y Hayden, F. (2002). *The Emergence of Business Process Management,* CSC's Research Services, Enero.

Smith, K.G.; Guthrie, J.P. y Chen, M.J. (1989). "Strategy, Size and Performance". *Organization Studies. Vol.10, No.1, pp.63-81.*

Smith, Michael Lee (2009). *Do you have a Six Sigma tip, learning or case study?*. www.isixsigma.com. 9/8/2009.

Stalk, George Jr. (1988). "Time- the next source of competitive advantage". *Harvard Business Review,* July-August, pp.41-51.

Stalk, G.; Evans, P. y Shulman, L. (1992). "Competir en habilidades, clave de la nueva estrategia empresarial". *Harvard-Deusto Business Review.* 4º trimestre, pp.44-56.

Stevens, S. S. (1951). *Handbook of Experimental Psychology*. New York: J. Wiley and Sons Inc.

Storey, Chris y Easingwood, Christopher (1996). "Determinants of new product performance. A study in the financial service sector". EEUU: *International Journal of Service Industry Management*. Vol.7, No.1.

Storey, Chris y Easingwood, Christopher (1998). "The augmented service offering: a conceptualization and study of its impact on new service success". *Journal of Product Innovation Management*. Vol. 15.

Strong, D., Y. Lee, and R. Wang, Ten Potholes in the Road to Information Quality. IEEE Computer, 1997: p. 38-46.

Sullivan, T. (2001). *Business Intelligence keeps tabs on the Net*. InfoWorld. http://www Infoworld.com

Sveiby, Karl Erik (2000). *Capital intelectual. La nueva riqueza de las empresas*. Barcelona: Gestión 2000.

Swanson. Roger C. (1995). *The quality improvement handbook: team guide to tools and techniques*. USA: CRC Press.

Taguchi, Genichi (1986). *Introduction to Quality Engineering*. New York: Asian Productivity Organization.

Taguchi, Genichi; Elsayed A. Elsayed y Thomas C. Hsiang (1989). *Quality Engineering in Production Systems*. New York: McGraw-Hill, Inc.

Takeuchi, Hirotaka and Nonaka, Ikujiro (1986). "The new new product development game". *Harvard Business Review. Vol*.64, No.:1, pp.137-146, Jan/Feb.

Taylor, Frederick Winslow (1911). *Principles of scientific management*. New York: Harper and Row, Publishers.

Tissen, R., Andriessen, D. y Deprez, F.L. (1998): Value-based knowledge management, Addison – Wesley, Longman.

Tissen, R., Andriessen, D. y Lekanne, F. (2000). *El valor del conocimiento para aumentar el rendimiento en las empresas.* Madrid: Prentice Hall.

Tissen, R., Andriessen, D. & Lekanne Deprez, F. (2000). El valor del conocimiento: Para aumentar el rendimiento en las empresas. (E.de Paz Ureña, Trad.). Madrid, España: Pearson Educación.

Terziovski, M. y Samson, D. (1999). "The link between Total Quality Management Practice and Organizational Performance". *International Journal of Quality & Reliability Management*. Vol.16, No.3, pp.226-237.

Tesch, Renata (1990). *Qualitative research: analysis types and software tools.* New York: The Falmer Press.

Thompson, Artur A. Jr; Gamble, John y Strickland, A.J. (2009). *Crafting & Executing Strategy: The Quest for Competitive Advantage: Concepts and Cases.* McGraw-Hill Companies.

Thompson, P. y McHugh, D. (1995). *Work organizations: a critical introduction*. London: Mcmillan.

Torgerson, Warren S. (1958). *Theory and methods of scaling*. New York: Wiley.

Tufte, E. R., & Weise Moeller, E. (1997). *Visual explanations: images and quantities, evidence and narrative* (Vol. 36). Cheshire, CT: Graphics Press.

Tufte, E. R. (2006). Beautiful evidence. *New York*.

Tufte, E. R., & Graves-Morris, P. R. (1983). *The visual display of quantitative information* (Vol. 2, No. 9). Cheshire, CT: Graphics press.

Tukey, J. W. (1977). Exploratory data analysis.

Vallés Martínez, Miquel S. (2007). *Técnicas cualitativas de investigación social. Reflexión metodológica y práctica profesional* (1a edición, 1999). Madrid: Síntesis sociología.

Van der Stede, Wim A.; Chee W. Chow y Thomas W. Lin (2006). 'Strategy, Choice of Performance Measures, and Performance'. *Behavioral Research in Accounting.* Vol.18, pp.185-205.

Venkatraman, N. y Ramanujan, V. (1986). "Measurement of Business Economic Performance in Strategic Research: A Comparison of Approaches". *Academy of Management Review.* Vol.11, No.4, pp.801-814.

Venkatraman, N. y V. Ramanujam (1987). "Measurement of business economic performance: an examination of method convergence". *Journal of Management*. Vol.13, No.1, pp.109-122.

Venkatraman, N. (1989). "Strategic orientation of business enterprises: the construct, dimensionality and measurement". *Management Science.* Vol.35, No.8, pp.942-962.

Vernette, E. (1995). "Eficacia de los instrumentos de estudio: evaluación de las escalas de medida". *Investigación y Marketing. Vol.*48, julio, pp.49-66.

Villalon Martínez, Nery María (2006). *Recomendaciones para la implementación de inteligencia de negocios basada en un data warehouse*. Tesis. Instituto Tecnológico y de Estudios superiores de Monterrey.

Wernerfelt, Birger (1984). "A resource-based view of the firm". *Strategic Management Journal*. Vol.5, No.2, pp.171-80.

West, J. (1997). Managing performance in the white spaces. In G. Rummler (Ed.), *Performance improvement pathfinders: Models for organizational learning systems* (pp. 108-123). Silver Spring, MD: International Society for Performance Improvement.

Whitten, J., Bentley, L. Barlow, V. (1991). *System Analysis and Design Methods*.

Williams, R. (1994). The Non-Designer's Design Book

Womack, J. y Jones, D. (1996). *Lean Thinking.* Simon & Schuster

Wood y Fildes (1976). *Forecasting for Business. Methods and Applications.* London: Long-man.

Wurman, R. S. (1997). Information Architects. Graphis Inc.

Yin, R.K. (1989). *Case study research: Design and methods (*Revised Edition). London: Sage

Zeeuw, E. (1995). "La función del EFQM". *Anuario de Calidad 1995.* Gaceta de los Negocios, pp.66-68.

Zeleni, Milan (2004). *Knowledge-Information Circulation Through the Enterprise: Forward to the roots of knowledge management*. En Yong Shi, Weixuan Xu, Zhengxin Chen (2004) *Data mining and knowledge management*. Springer.

Zerilli, A. (1990). *Fundamentos de organización y dirección general*. Bilbao: Deusto.

Zhang, Z. (2000). "Developing a Model of Quality Management Methods and Evaluating their Effects on Business Performance". *Total Quality Management*. Vol.11, No.1, pp.129-137.

Ziegler, Mike (2003). *Process Owner. http://www.isixsigma.com/dictionary/Process_Owner-436.htm*. 2/8/2009

Zwicky, F. (1947). "Morphology and nomenclature of jet engines". *Aeronautical Engineering Review.* Junio.

Anexo