

UNIVERSIDAD DE OVIEDO

Facultad de Ciencias

TRABAJO FIN DE MÁSTER

MÁSTER EN ANÁLISIS DE DATOS PARA LA INTELIGENCIA DE NEGOCIOS

CUADRO DE MANDO PARA LA GESTIÓN DE RECURSOS DE EVENTOS

AUTOR: Aníbal Torre Melero

Oviedo, Julio de 2019

Índice

Resumen
1 Introducción
2 Contextualización del proyecto5
3 Contexto empresarial9
4 Advertencia sobre el proyecto 11
5 Proyecto
6 Cálculo de los KPI 14
6.1 % Días de ocupación 14
6.2 Margen de eventos 19
6.3 Usuarios 21
6.4 % Absentismo de espectadores 22
6.5 Personal contratado23
6.6 Consumo
6.7 % Recogida selectiva
6.8 Productos Valor Añadido 32
6.9 % Reservas Canceladas 33
6.10 Consumo per Cápita 35
7 Diseño
8 Conclusiones
9 Bibliografía

Índice de figuras y tablas

Figura 5.1: Llave sintética	. 12
Figura 5.2: Modelo de datos copo de nieve	. 14
Figura 6.1.1: Gráfico de % de días ocupados	. 19
Figura 6.2.1: Gráfico margen total	. 21
Figura 6.3.1: Gráfico número de usuarios	. 22
Figura 6.4.1: Gráfico % de absentismo	. 23
Figura 6.5.1: Gráfico personal contratado	. 25
Figura 6.5.2: Diseño de hoja con botones	. 26
Figura 6.6.1: Gráfico consumo	28
Figura 6.7.1: Gráfico % de recogida selectiva	. 30
Figura 6.7.2: Gráfico de reparto de la recogida selectiva	. 31
Figura 6.8.1: Gráfico beneficio y volumen productos de valor añadido	. 33
Tabla 6.8.2: Productos de valor añadido	. 33
Figura 6.9.1: Gráfico % de reservas canceladas	34
Figura 6.10.1: Gráfico reparto consumo per càpita	36
Figura 7.1: Logo de BSM	38
Figura 7.2: Logo del ayuntamiento de Barcelona	39
Figura 7.3: Filtro de año	. 39
Figura 7.4: Filtro de mes	39
Figura 7.5: Diseño hoja cuadro de mando	40

Resumen

Vivimos en una revolución tecnológica y eso se traduce en muchos cambios de manera acelerada. Las empresas no son ajenas a estos cambios y por tanto en numerosas ocasiones se ven sobrepasadas por la enorme cantidad de datos que generan día tras día. El Big Data lo ha cambiado todo y por ello cada vez más se requieren más profesionales con conocimientos para poder extraer toda la información de esos datos y a su vez la variedad de herramientas y lenguajes de programación para ello va en aumento.

En este trabajo nos centramos en analizar la relevancia del Business Intelligence hoy en día, así como explicando sus distintos procesos. Además, exponemos un proyecto real de visualización de datos explicando el tratamiento de los datos para poder operar con ellos y la obtención de los diferentes KPI's solicitados. Asimismo, se muestra cómo se realizó el diseño, así como las conclusiones a las que se ha llegado en este proyecto.

Abstract

We are living in a technological revolution, which means numerous rapid changes. Companies are not immune to this revolution and on many occasions, therefore, they are overwhelmed by the huge amount of data that they generate day after day. Big Data has changed everything and for this reason companies need more professionals who can extract all the information from these data. At the same time the variety of tools and programming languages for this is increasing.

In this project we focus on analysing the relevance of Business Intelligence nowadays as well as explaining all its different phases. Additionally, we present a real data visualization project, explaining the data processing so as to be able to work with them, and the obtaining of the different KPI's requested. Furthermore, we show how the design was made and the conclusions which have been reached in this project.

3

1.- Introducción

Las empresas llevan manejando datos desde siempre, solo que en estos últimos años gracias a la revolución tecnológica e internet se han incrementado exponencialmente año tras año. Por este motivo, muchas de ellas, se ven incapaces de hacer frente a tal volumen de datos y no saben cómo pueden sacarles provecho.

Algunas de ellas optan por hacer análisis con sus ficheros de Excel, pero eso no es muy eficaz ya que al manejar tantísimos ficheros a la vez es imposible ser consciente de todo lo que rodea a la empresa. Por este motivo surgió el Business Intelligence, que no es un concepto nuevo pero que su auge se produce en estos tiempos, para poder resolver estos problemas de volúmenes enormes de datos y poder extraerles todo el conocimiento.

El propósito de este trabajo es exponer la relevancia de los procesos de Business Intelligence en la gestión de una empresa para así, mediante el uso de la información, mejorar todas las áreas de esta, desde los ingresos y gastos hasta el control de los empleados.

Además, se mostrará un proyecto de visualización de datos; explicando todo el proceso de tratamiento de los datos, así como el cálculo de los diferentes KPI's solicitados por la empresa contratante.

Por último, se presentará el diseño realizado. Aunque pueda parecer algo sin mucha trascendencia, cada vez más las empresas inciden mucho más en ellos ya que se desea que sea intuitivo a la par que agradable para el usuario cuando navegue explorando los resultados solicitados.

4

2.-Contextualización del proyecto

Hoy en día se habla mucho del Business Intelligence, pero mucha gente no sabe con exactitud qué es. El Business Intelligence abarca desde la extracción de datos, la limpieza de estos para poder operar con ellos, su análisis y la creación de reportes que sean muy visuales y fáciles de interpretar.

Según Gartner lo podríamos definir siguiente manera: "El Business Intelligence (BI) es un término general que incluye las aplicaciones, la infraestructura y las herramientas, y las mejores prácticas que permiten el acceso y el análisis de la información para mejorar y optimizar las decisiones y el rendimiento." (Gartner, 2019)

Dentro del Business Intelligence debemos diferenciar dos tipos de procesos:

- ETL (Extract, Transform and Load) es la parte más crucial de cualquier proyecto. Consume la mayor parte del tiempo que se dedica a los proyectos y hay que ser muy cuidadosos ya que si tratamos los datos mal, todo el proyecto estará defectuoso. Las fases del proceso ETL son (Cano, 2007):
 - Extracción: esta fase se basa en la extracción de aquellos datos de los bases de datos operativas que sean necesarios y prepararlos para el resto de las fases del proceso ETL. Mediante el data *staging* (almacenes de datos intermedios) almacenamos datos durante la limpieza de estos. Una vez tenemos acumulados datos de diversas fuentes los cargaremos en el *datawarehouse*.
 - Limpieza: una de las fases más importantes del proceso. En ella deberemos depurar los datos, corregirlos, estandarizarlos, relacionarlos y consolidarlos. Una vez lo hayamos hecho tendremos los datos sin duplicados, sin valores nulos, en la misma medida, ... Con todo esto ya podríamos pasar a la siguiente fase.
 - 3. Transformación: la transformación siempre dependerá de los requerimientos del proyecto. En esta fase se incluyen cambios

de formato, sustitución de códigos, valores derivados y agregados.

- 4. Integración: es la fase donde cargaremos los datos en el datawarehouse. Antes de cargarlos deberemos comprobar que todos los resultados estén bien, ya que si no llevará a los usuarios a tomar decisiones erróneas.
- 5. Actualización: la periodicidad con la que haremos cargas al datawarehouse dependerá de cada proyecto.

Actualmente las ETL con más relevancia en el mundo empresarial son SSIS (Microsoft), PowerCenter, Oracle Data Integrator (Oracle), Pentaho Data Integrator y Business Objects Data Integrator (SAP).

- Herramientas de visualización: tras realizar el proceso ETL podemos proceder a la visualización de los datos. Puede parecer una tarea de escasa importancia, pero nos es de gran ayuda para sacar conclusiones y poder tomar decisiones con fundamento. Para ello debemos tener en cuenta estas consideraciones:
 - Todos los datos deben ser visualizados. Esto no es correcto y además es peligroso, ya que saturar de información no necesaria al usuario puede llevarle a no tomar buenas decisiones.
 - 2. Fiarse en exceso de la visualización. Es cierto que nos es muy útil para ver patrones y tendencias, pero no por ello puede sustituir al pensamiento crítico y analítico. Además, las representaciones gráficas son muy fáciles de manipular puesto que, por ejemplo, a la hora de hacer una comparativa entre dos gráficos pueden estar en distinta escala y por tanto no ser fiable esa información.
 - Hacer cálculos sencillos para comprobar que los datos están correctamente. De esta manera nos evitaremos resultados erróneos al calcular los indicadores.

Una de las grandes virtudes de estas herramientas es la posibilidad de realizar una visualización interactiva. Este concepto abarca desde realizar una

selección especifica de un dato concreto o del conjunto que el usuario necesite, realizar un filtrado de los datos para así disminuir la cantidad de estos y poder sacar conclusiones con mayor facilidad, enlazar información para relacionarla en diferentes gráficos, así como poder organizarla de la manera que el usuario estime oportuno para así fomentar su comprensión (Wang, Wang, y Alexander, 2015).

Actualmente las herramientas de visualización más usadas a nivel mundial son Power BI, QlikView y Tableau.

La herramienta QlikView es una herramienta usada para Business Intelligence desarrollada por la compañía sueca QlikTech en 1993. Antiguamente todo lo relacionada con Business Intelligence se encargaba a los departamentos de IT, desde la extracción de datos, su limpieza, su tratamiento y el crear los diversos reportes o cuadros de mando; al tener que encargarse de todo solían entregar reportes estáticos, pero con el paso de los años la mayoría de las empresas se han ido dando cuenta de que esto no satisface a sus clientes (García y Harmsen, 2013). Por este motivo QlikView permite una navegación por los datos muy dinámica ya que cualquier usuario sin necesidad de conocer la herramienta puede utilizar los reporte e ir usando los diferentes filtros que necesite lo que permite poder ser plenamente conscientes de la situación de la empresa y así poder tomar decisiones mucho más beneficiosas para ellas. Además, se diferencia de otras herramientas de Business Intelligence en que da total libertad al usuario a la hora de realizar cálculos o diseñar sus reportes.

QlikView se basa en modelos de bases de datos relacionales. Aparecieron en la década de 1980 y su principal virtud es que ayudan a eliminar la redundancia que haya en los datos; esto significa en una simplificación y una mejora de los sistemas transaccionales (García y Harmsen, 2013).

Para el usuario una base de datos relacionales es como un conjunto de tablas. Cada fila o tupla es un registro, mientras que cada columna o atributos son las variables. Las diversas tablas están representadas por tuplas y cada una de ellas tiene varios atributos. Deben tener, al menos, dos características:

7

- Es recomendable que toda la información que vayamos a usar esté alojada en el mismo sitio para que no nos lleve a confusiones o redundancias de información.
- Al tratarse de información relacional, va a haber relaciones entre las diferentes tablas a lo largo de la base de datos (Cano, 2007).

Comúnmente se construye un modelo dimensional, en el que creamos una tabla de hechos que será la principal y tendrá relacionadas diversas tablas de dimensión que aportaran información, métricas y profundidad a la información de la tabla de hechos. Los dos esquemas más comunes son el de estrella y el copo de nieve. El de estrella se caracteriza porque todas las tablas de dimensión están exclusivamente ligadas a la tabla de hechos, mientras que el esquema de copo de nieve tiene varias tablas de dimensión ligadas a la tabla de hechos y su vez existen más tablas que complementan con información y métricas adicionales a las tablas de hechos (García y Harmsen, 2013).

En este proyecto al ser la empresa la que impone el software a utilizar, así como los indicadores a hallar, solo se pudo tener libertad en la elección de las diversas representaciones de los indicadores y en el diseño.

3.-Contexto empresarial

GFI es una consultora francesa de IT con presencia en 21 países que ofrece desde servicios de consultoría clásica hasta edición de software. Cuenta con cerca de 20.000 empleados y su cifra de negocio se encuentra alrededor de los 1.400 millones de euros.

Dentro de ella se encuentra el departamento de Business Intelligence donde se proporciona soluciones en materia de obtención, limpieza y tratamiento de datos para maximizar su utilidad. Se realizan proyectos de carácter público como para empresas privadas adaptándose tanto a cuando el cliente desea algo especifico como cuando quiere que se le diseñe un proyecto desde cero ya que no sabe cómo encarar el problema.

En este trabajo se va a exponer la realización de un proyecto encargado por la empresa BSM (Barcelona de Serveis Municipals) a la consultora GFI, donde requirió un proyecto en la herramienta QlikView para poder visualizar distintos indicadores, tanto los resultados como su representación gráfica, así como diversos filtros para navegar por la información según sus necesidades.

La empresa BSM es la encargada de organizar todos los eventos que se realizan en la Anella Olímpica de Barcelona. Esto comprende el Palau Sant Jordi, la Esplanda, el Estadi y el Sant Jordi Club. Los archivos con los datos nos fueron entregados en diversos ficheros Excel y de texto. Cada archivo trata un tema en concreto, pero todos están relacionados. Intentar tratar esta información en cada archivo por separado es una tarea muy complicada ya que no vamos a ser conscientes de todo al tratarlo por separado; de ahí su interés en que se les hiciese un archivo de QlikView en el que se pueda visualizar toda la información al mismo tiempo con sus relaciones y todos los filtros que necesiten.

Los indicadores son de todo tipo; desde el cálculo de ingresos y gastos a la recogida de residuos como los días de ocupación de las diversas instalaciones; todos ellos con el objetivo de aumentar ingresos, reducir costes y realizar espectáculos que gusten a la mayoría de los barceloneses.

Barcelona de Serveis Municipals es una empresa municipal que depende del ayuntamiento de Barcelona y que se encarga de diversos servicios en la

9

ciudad como los aparcamientos, la grúa municipal, el bicing (bicicletas públicas de alquiler), la gestión del zoo y el Park Güell, así como el parque de atracciones del Tibidabo, los cementerios de Barcelona y las instalaciones de la Anella Olímpica.

4.-Advertencia sobre el proyecto

Debido a que en el proyecto existen muchos datos sensibles, la empresa me obliga a cifrar los datos relativos a los nombres de los eventos, los nombres de los promotores y los tipos de actos. Para poder cifrar esta información se usó la función *Hash128* para que así esta se mantenga protegida. La expresión completa es: Hash128(EVENT_NAME). También debo advertir a los usuarios de que los datos son relativos a abril de 2018, por lo que muchos eventos que ya están reservados para fechas futuras aparecen en los archivos, pero en algunos indicadores, como el de la recogida selectiva de residuos, aparecen en blanco ya que no se puede saber la cantidad a priori. Además, todas las cifras están en formato anglosajón por lo que el separador de miles es la coma y el de decimales es el punto.

5.-Proyecto

El objetivo de este proyecto es desarrollar un cuadro de mando que permita visualizar y consultar la información de manera fácil y sencilla para que sea un soporte a la hora de tomar decisiones en la empresa.

Los objetivos de este proyecto son:

- Carga de los datos en QlikView.
- Generar un cuadro de mando en QlikView en el que el usuario pueda interactuar, visualizar y consultar la información relacionada con los eventos.
- Suministrar diversos filtros al usuario que le permitan navegar de una manera sencilla y que además pueda hacer filtrados de los datos para su estudio.

Una vez recibidos los archivos de datos por parte de la empresa en los formatos csv y txt procederemos a su carga y tratamiento para realizar el modelo de datos.

Comenzamos por la pestaña Main. Viene predeterminada con QlikView y en ella se determina la separación de decimales, de miles, el formato de fecha, formato de moneda, así como los meses y los días de la semana.

La primera de las tablas que vamos a cargar es la Eventos. Es la tabla de hechos o, dicho de otra manera, es la tabla principal y de la que dependen las

demás. La siguiente que cargaremos es la tabla de Grupos que ira ligada a la tabla de Eventos por el campo GROUP_ID.

Si nos fijamos detenidamente vemos que hay los mismos campos en ambas tablas, entonces ¿cómo el programa sabe que solo queremos ligar las tablas por el campo GROUP_ID y no por otro/s? Deberemos cambiar los nombres de todos los campos que coinciden en ambas tablas con la función *as* para que así no haga múltiples uniones entre las tablas. Esto es algo que se debe evitar a toda costa ya que puede conllevar a problemas con los datos. En el caso de que la hubiésemos cargado sin darnos cuenta de cambiar los nombres de los demás campos; es muy fácil de identificar ya que crea una nueva tabla que se llama **llave sintética**. El programa entiende que se quiere ligar ambas tablas, pero al haber varios campos posibles crea esta tabla donde aglutina todos los campos posibles. En la siguiente imagen podemos ver la unión entre ambas tablas mediante una llave sintética.



Figura 5.3: Llave sintética

La siguiente tabla que vamos a cargar es la de Recursos de Eventos que estará ligada también a la tabla de hechos Eventos. En este caso el campo para ligar las tablas será EVENT_ID. Una pregunta que puede surgir es, ¿cómo elegir el campo para unir dos tablas? El criterio debe ser el que dé menos lugar a error, por eso siempre que se pueda se debe usar números identificatorios o ID como en este caso. Por ejemplo, el campo EVENT_NAME sería muy peligroso ya que si no están exactamente igual en ambas tablas nos conducirá a inconsistencias

en los resultados, ya que reconoce signos de puntuación y diferencia mayúsculas y minúsculas.

A continuación, cargamos la tabla Subeventos que también irá ligada a la tabla Eventos por el campo EVENT_ID. Es muy común que dos tablas se unan a otra por el mismo campo siempre y cuando sea esta la mejor opción posible.

Sucesivamente, cargamos la tabla Recursos de subeventos que irá ligada a la tabla cargada anteriormente mediante el campo SUBEVENT_ID, ya que las actividades que se realizan disponen de datos de evento y subevento que necesitaremos más adelante para el cálculo de los KPI's.

Adicionalmente, cargaremos las tablas Limpieza y Personal que irán ligadas a la tabla Recursos de eventos por el campo RESOURCE_GROUP_ID, puesto que complementan a esta con información muy relevante que necesitaremos más adelante.

Por último, cargaremos una tabla adicional que no estará en ningún archivo a diferencia de las demás, sino que la crearemos escribiendo todo su código en el script. Esta tabla que llamamos Calendario nos será de una gran utilidad puesto que nos brinda la posibilidad de añadir el año, mes, día de la semana, ... En este caso no disponemos de ningún número identificatorio por lo que deberemos ligar esta tabla a la tabla de hechos Eventos por el campo MoveOutEnd.

Con todo esto ya tenemos nuestro modelo. En este caso se trata de un modelo en copo de nieve ya que no todas las tablas están ligadas a la tabla de hechos, sino que algunas de ellas se implementan en otras en busca de que no haya una redundancia de datos (García y Harmsen, 2013). En este caso al ser un modelo bastante simple y con poco volumen de datos el esquema en copo de nieve es el más adecuado; pero en proyectos con un gran volumen de datos no es recomendable porque necesita crear más tablas de dimensiones y más relaciones por lo que influirá mucho en el rendimiento. En la siguiente imagen vemos nuestro modelo final de datos:

13



Figura 5.4: Modelo de datos copo de nieve

6.- Cálculo de los KPI

Los KPI (Key Perfomance Indicator) son una serie de indicadores que se utilizan para sintetizar información para evaluar la eficacia de las acciones que se toman en la empresa y que sea relevante a la hora de tomar decisiones relevantes dentro de esta. Son muy utilizados ya que ofrecen una medición más sencilla, objetiva y fiable. Los KPI deben ser objetivos (no se pueden basar en nuestro juicios), específicos (debemos tener claro qué queremos medir), periódicos (es recomendable actualizar los KPI semanalmente o mensualmente para ir viendo su evolución), cuantificables (debemos poder medir nuestro objetivo para que tenga utilidad), realistas (debemos marcarnos objetivos que sean plausibles de conseguir), coherentes (deben ser acordes a los objetivos de la empresa) y relevantes (deben ser útiles para los objetivos de la empresa) (Parmenter, D. (2015)).

Los KPI solicitados por la empresa son los siguientes:

6.1.- % Días de ocupación: Este indicador permitirá analizar la ocupación de los eventos de la Anella Olímpica con el objetivo de maximizar el uso de los diferentes espacios que comercializan y emprender las acciones

comerciales correspondientes. La fórmula para su cálculo es: *Días ocupados / Días totales*

Para poder calcular este KPI debemos tratar los datos previamente. Primero deberemos crear un campo que se llame Duración. Hay muchos eventos que duran varios días ya que se cuenta desde el día que llegan para montarlo todo hasta el día que se van. Por eso este campo será la resta del día que se van MoveOutEnd menos el día de llegada MoveIn. La primera dificultad para su cálculo es que estos campos no vienen en formato de fecha, por lo tanto, debemos modificarlos. Usando la función Date lo cambiamos a formato fecha. Esta función consta de dos argumentos; en el primero de ellos irá el campo que queremos convertir a fecha y en el segundo el formato de fecha que le queremos dar. En nuestro caso será día, mes y año en ese orden. Esos campos además de las fechas también tienen la hora, lo que distorsionará nuestros cálculos ya que tendrán decimales. Para evitar esto redondearemos hacia abajo la fecha de llegada y hacia arriba la fecha de salida para así tener los días completos. Con todo esto la función completa será así: Date(Floor(Movein), 'DD-MM-YYYY') as Moveln para la fecha de llegada y Date(Floor(MoveOutEnd), 'DD-MM-YYYY') as MoveOutEnd para la fecha de salida. En QlikView siempre que se hace alguna modificación de los datos en el script se debe añadir as para ponerle un nombre; en este caso les pondremos el mismo que ya tenían.

Ahora que ya tenemos ambos campos en formato fecha podemos crear el campo Duración. Los eventos que duran solo un día tienen el campo Moveln en blanco, por lo que necesitaremos usar una función If para poder calcularlo. La expresión es la siguiente: lf(Len(Moveln)<1, 1, fabs(Left(MoveOutEnd,10) – Left(Moveln,10))+1) as Duracion. La función If es un condicional y consta de tres argumentos: el primero es la condición a cumplirse; en este caso es que, si la longitud del campo Moveln es menor que uno, o sea que esté en blanco. El segundo argumento es lo que debe hacer cuando se cumpla lo descrito anteriormente, en este caso que coloque un uno. El tercer argumento es lo que debe hacer en caso contrario, cuando el campo sea mayor o igual que uno.

Como se expuso anteriormente, ambos campos además de fecha tienen hora; para evitarnos problemas con ello usamos la función Left que sirve para escoger caracteres de un determinado campo. Consta de dos argumentos: el primero es el campo al que le quieres aplicar la función y el segundo el número de caracteres que quieres que coja, en este caso 10, ya que las fechas cuentan con dos caracteres para el día, otros dos para el mes, cuatro para el año y dos guiones para separar entre ellos.

Además, al hacer esta resta le debemos sumar uno, porque por ejemplo vamos a suponer un evento que llega el día 1 y se va el día 3; estuvo ocupando el recinto 3 días, pero al hacer la resta 3 - 1 = 2; por lo que debemos sumarle un uno para que nos salga bien el resultado. También le pondremos la función Fabs para darle valor absoluto a los datos por si algunas fechas estuviesen puestas al revés y nos diese negativo. Con todo esto ya tenemos creado el campo Duración.

Para calcular este KPI necesitaremos crear otro campo nuevo que será el tipo de ocupación de cada evento. El campo TIPO_OCUPACION viene en los archivos facilitados por la empresa, pero debemos añadirlo a la tabla Events nosotros mismos. Para ello deberemos mapear el campo TIPO_OCUPACION que se encuentra en otro archivo. Mapear sirve para reemplazar un valor llave por uno de equivalencia.

Previamente usamos la función **MAPPING** para indicar que queremos crear una tabla de mapeo. Este tipo de tablas tienen las siguientes propiedades:

- Solo puede tener dos columnas; la de equivalencia y la que el mapeo debe regresar para el valor de equivalencia.
- Es una tabla temporal. Una vez finalizada la recarga del script de QlikView se eliminará del modelo.

Una vez tengamos la tabla de mapeo hecha debemos usar la función **Applymap** para poder insertarla en la tabla Events. Esta función consta de tres parámetros:

- En el primero, ponemos el nombre de la tabla de mapeo. En nuestro caso MAP_Estado.
- En el segundo, el campo que se desea buscar en la tabla de mapeo. Para nosotros es el campo STATUS.
- En el tercero, pondremos lo que queremos que haga en el caso de que no se encuentre una correspondencia en la tabla de mapeo.

Con todo esto la función de la tabla de mapeo será: MAP_Estado: Mapping LOAD Capitalize(Status), Tipo_ocupacion FROM Ficheros\Salida\Estado.csv (txt, codepage is 28591, embedded labels, delimiter is ';', msq)

Y la de applymap será: applymap ('MAP_Estado', Capitalize (STATUS), 'Bloqueig') as TIPO_OCUPACION.

Una vez que ya tenemos el mapping hecho ya podemos calcular nuestros KPI. Para ello usaremos un set analysis. El set analysis o análisis de conjuntos en español es una función de QlikView nos permite hacer cálculos que serían imposibles de otra manera. Nos permite restringir, predefinir y extender el conjunto de datos en que los objetos basan sus cálculos. Es de gran utilidad que solo afecte a la expresión donde está siendo utilizado y no al conjunto completo ya que nos permite insertar en nuestro documento diversos objetos con distintos filtros a la vez. Las principales ventajas del set analysis son:

- Te permite comparar resultados de varios periodos de tiempo a la vez sin tener que cambiar el estado de selección.
- A la hora de realizar un cálculo puedes excluir o restringir ciertos valores de campo.
- Para crear acumulados hasta cierta fecha.
- Para ignorar selecciones en algún campo que no sean relevantes para nuestros cálculos.
- Para ignorar todas las selecciones.
- También se puede usar como reemplazo de la función if, ya que esta consume altos recursos de hardware.

Para poder calcular estos KPI's debemos filtrarlo en función del recinto donde se celebre el espectáculo. Podríamos hacerlo con un cuadro de lista, pero filtraríamos todo el archivo. Para solo filtrar en el objeto de hoja debemos usar set analysis.

Tiene una sintaxis un tanto compleja que vamos a ver a continuación. Lo primero de todo es la función que queremos utilizar (Sum para sumar, Count para

contar, ...) Luego, los valores que deseamos incluir o excluir; en este caso los haremos de inclusión. La siguiente parte es indicar los campos que queremos incluir en nuestro calculo. En nuestro caso será el recinto donde se celebra el espectáculo (Palau Sant Jordi, Esplanda, Estadi o Sant Jordi Club) así como el estado del reciento (se considera que está ocupado si está reservado, bloqueado o tiene una pre-reserva).

Por último, debemos indicar el campo que queremos que calcule, en este caso es Duración. Además, podemos añadirle la función Num en la que su primer argumento es lo que queramos que convierta en numérico y el segundo es el formato, en nuestro caso será #,##0 para que la coma sea el separador de miles. La función final será así: Num(Sum({<VENUE = {'Esplanda'}, TIPO_OCUPACION={'Bloqueig', 'Reserva', 'Pre-reserva'}>} Duracion),'#,##0'). Para calcular el resto solo habría que cambiar el recinto.

Con todo esto ya tenemos calculados nuestros KPI sobre el número de días de ocupación. Además, el cliente nos pidió una gráfica del porcentaje de días que está ocupado cada recinto por mes. Para ello deberemos insertar un gráfico en nuestra hoja. Luego escogeremos que tipo de grafico queremos; en este caso usaremos uno de líneas. A continuación, le añadiremos la dimensión Mes ya que es la que se mostrará; el año lo tendrá que filtrar el usuario. Ahora le agregaremos las expresiones de los cuatro recintos para que así se muestre en el gráfico. El cálculo será los días de ocupación de cada recinto dividido de los días que tiene cada mes. Por lo tanto, tenemos que calcular los días que tiene cada mes ya que no lo tenemos.

En esta ocasión volveremos a hacer un **set analysis**, pero de exclusión en lugar de inclusión como anteriormente. Lo que haremos es contar los días que sean distintos y excluiremos los campos de VENUE y TIPO_OCUPACION ya que si no afectarán y solo nos indicará los días que hay algún evento y no todos los días que tiene cada mes.

La expresión completa quedará de la siguiente manera: Num (sum ({<VENUE = {'Recinto1'}, TIPO_OCUPACION= {'Bloqueig', 'Reserva', 'Prereserva'}>} Duracion),'#,##0') / Count (DISTINCT {\$< VENUE, TIPO_OCUPACION =>} [Día])

Ahora solo nos queda añadir las expresiones con los cuatro recintos y ya tendríamos el gráfico hecho. Para terminar, cambiaremos el símbolo por puntos,

disminuimos el tamaño de línea, aumentamos el de los símbolos, ponemos la leyenda en horizontal, le indicamos que no muestre el título y le damos 100% de transparencia para así poder colocarlo en cualquier parte de la hoja.



Figura 6.1.1: Gráfico de % de días ocupados

6.2.- Margen de Eventos: Este indicador permitirá calcular el margen de beneficio de cada uno de los eventos que se han llevado a cabo con el objetivo de analizar su rendimiento y aumentar el margen de beneficio para próximos eventos. La fórmula para su cálculo es: *Ingresos de evento – Gastos de evento*, donde Ingresos de evento es (*Cesión de uso + montaje y desmontaje - rápel - descuentos*) + (*Ingresos de cesión de salas - descuentos*) + *Ingresos de ventas comerciales + Ingresos de servicios + Restauración + Tiendas*; mientras que Gastos de eventos está compuesto por *Gastos de acciones comerciales + Gastos de servicios*.

Este KPI lo podemos calcular directamente sin necesidad de transformar ningún dato. Para calcularlo debemos sumar los ingresos de los eventos y los de los subeventos y restar los costes de los eventos y los de los subeventos. Para ello usaremos la siguiente expresión: Num(Round(Sum(TOTAL_INCOME) + Sum(TOTAL_INCOME_SUB)),'#,##0'). Nótese que la función Sum aparece dos veces y no se simplifica escribiendo Num(Round(Sum(TOTAL_INCOME + TOTAL_INCOME_SUB)),'#,##0'); esto es porque la segunda manera sería incorrecta. Ambos campos (TOTAL_INCOME y TOTAL_INCOME_SUB) no son un número, sino una sucesión de ellos; por lo tanto, necesitamos hacer el sumatorio de cada uno por separado para que sí sean dos números y poder sumarlos. Para los costes la formula será Num (Round (Sum (TOTAL_COST) + Sum (TOTAL_COST _SUB)),'#,##0'). El margen total se calcula restando los ingresos totales menos los costes totales, por lo que su expresión será Num (Round (Sum (TOTAL_INCOME) + Sum (TOTAL_INCOME_SUB) - Sum(TOTAL COST) - Sum (TOTAL COST SUB)),'#,##0')

Además, calcularemos el margen por día reservado que será el margen total dividido de todos los días que alguno de los recintos está reservado, con lo que nuestra expresión será Num(Round(Sum(TOTAL_INCOME) + Sum (TOTAL_INCOME_SUB) – Sum (TOTAL_COST) – Sum(TOTAL_COST_SUB)) / Sum({<TIPO_OCUPACION={'Bloqueig', 'Reserva', 'Prereserva'}>}Duración), '#,# #0').

El cliente también quiere ver como se reparten los ingresos y gastos de los eventos y los ingresos y gastos de los subeventos. Por lo tanto, las expresiones serán Round(Sum(TOTAL_INCOME)) para los ingresos de los eventos, Round(Sum(TOTAL_INCOME_SUB)) para los ingresos de los subeventos, Round(Sum (TOTAL_COST) para los costes de los eventos y Sum(TOTAL_COST_SUB)) para los costes de los subeventos.

A la hora de representar este KPI, opté por utilizar un gráfico de indicador al que le quité la escala, la aguja de indicador, una amplitud de 180 grados, que su área interior sea del 50 por ciento y lo pinté de rojo para representar los gastos. Como este tipo de indicadores no permite que se les añada varios valores lo dejaremos así y lo mandaremos a una capa negativa. A continuación, añadiremos otro gráfico similar, pero en las expresiones le añadimos la formula del peso de los ingresos respecto del total, con lo que la expresión será Round (Sum (TOTAL INCOME) + Sum (TOTAL INCOME SUB)) / (Round (Sum (TOTAL INCOME) + Sum(TOTAL INCOME SUB)) + Sum(TOTAL COST) + Sum(TOTAL COST SUB)), en el indicador que rellene hasta el valor obtenido, le damos la misma amplitud y área interior del anterior, le quitamos la escala y lo pintamos de gris para representar los ingresos. Lo ponemos encima del gráfico anterior, entonces cada vez que filtramos hace el efecto óptico de que ambos gráficos varían en función del filtro que se aplique, pero solo cambia el de los ingresos, el de los costes siempre se queda igual. Para la representación de los costes/ingresos de los eventos y subeventos hice los mismos gráficos con sus respectivas expresiones. La representación quedará de la siguiente manera:



Figura 6.2.1: Gráfico margen total

6.3.- Usuarios: Este indicador permitirá cuantificar el número de usuarios que acceden a las instalaciones por cada evento.

Un usuario, por definición, es una persona que entra a las instalaciones, el cual podría ser un asistente o un visitante. A continuación, se definen estos dos conceptos:

- Asistente: persona que entra a la instalación con tiquet, independientemente del importe (gratuito o de pago).
- Visitante: persona que entra a la instalación sin tiquet por actos corporativos o ventas globales.

Por lo tanto, la fórmula para calcular este indicador es: Asistentes + Visitantes.

Este KPI también se puede calcular directamente como el anterior. En el caso de los espectadores, asistentes y usuarios ya nos vienen en los datos. Por lo tanto, solo habrá que usar la función Sum y le añadimos Num para darle el formato necesario. Las expresiones serán las siguientes: Num (Sum (NumberOfSpectators), '#,##0') para el número de espectadores, Num (Sum (Attendance), '#,##0') para el número de asistentes y Num (Sum (NumberOfVisitors), '#,##0') para el número de visitantes.

En el caso del número de usuarios deberemos calcularlo. Se trata de la suma del número de visitantes y el número de asistentes por lo que la expresión será Num (Sum (NumbreOfVisitors) + Sum (Attendance), '#,##0').

En cuanto a su representación gráfica, la forma más intuitiva de verlo es con un gráfico de barras. Lo insertamos, le añadimos la dimensión EVENT_NAME para que se vea el nombre de cada evento al lado de cada barra, le añadimos la expresión de usuarios descrita anteriormente, lo orientamos horizontalmente y le activamos la barra de herramientas del eje X para cuando el número de elementos sea superior a 10. Hacemos esto porque el número de eventos es elevado y si se ven todos a la vez es más difícil de interpretar. Por último, lo ordenaremos de mayor a menor para mejorar aún más su comprensión. El gráfico resultante será el siguiente:



Figura 6.3.1: Gráfico número de usuarios

6.4.- % Absentismo de espectadores: Este indicador permitirá cuantificar el porcentaje de espectadores que no llegaron a acceder a las instalaciones para cada evento. Con la finalidad de comprender que representa este indicador es necesario tener en consideración las siguientes definiciones:

- Asistente: persona que entra a la instalación con tiquet, independientemente del importe (gratuito o de pago).
- Espectador: persona con tiquet, que podría entrar o no a la instalación, independientemente del importe del tiquet (gratuito o de pago).

La fórmula para el cálculo de este indicador es: 1 - [(Volumen Asistentes) / (Volumen Espectadores)].

Para obtener este ratio deberemos dividir el número de asistentes entre el número de entradas, lo redondeamos para que tenga cuatro decimales y luego lo multiplicamos por 100. Por último, le añadimos el símbolo de porcentaje. La expresión final será esta: Round(Sum(Attendance) / Sum(NumberOfSpectators), 0.0001)*100 & '%'.

En cuanto a su representación gráfica haremos el mismo grafico que en el anterior KPI, pero en este caso lo ordenaremos por orden alfabético. En su expresión no pondremos el ratio de asistencia como en el párrafo anterior, sino que pondremos el de absentismo puesto que en el gráfico se ve cada evento por separado facilitando su comprensión. El gráfico será el siguiente:



Figura 6.4.1: Gráfico % de absentismo

6.5.- Personal contratado: Conjunto de indicadores que permitirán tener una visión global del personal contratado para cada evento. En particular, se calcularán los siguientes indicadores:

- Total personas contratadas
- Total horas personal contratado
- Total importe personal contratado

Para contabilizar dichos indicadores se deberá tomar en cuenta tanto el personal externo como el interno.

Como este KPI se divide en tres partes vamos a analizar cada uno por separado. Comenzamos por el total de personas contratadas. Para hallar este valor necesitamos sumar los trabajadores internos de la empresa y los externos. Necesitamos saber el tipo de cada uno para poder calcularlo, y el campo Tipo viene con muchos valores nulos y ausentes. Por lo tanto, tendremos que modificarlo para poder operar con él.

Usaremos la función **Trim** que sirve para quitar los espacios en un campo. Esta función es muy útil para eliminar distorsiones que provoquen nulos. Mediante una orden If, lo que haremos es que, si la longitud del campo Tipo es menor que uno después de aplicarle la función Trim, o lo que es lo mismo, que esté en blanco, pues entonces en esos casos sean sin tipo, ya que no tenemos ninguna manera de saber cuáles de esos son internos y externos; por eso los dejamos sin clasificar y en el caso contrario que aparezca su tipo. La expresión será la siguiente: lf(Len(Trim(Tipo))<1, 'SIN TIPO', Tipo).

Ahora ya podemos empezar a calcular los valores. Usaremos un set analysis donde queremos que sume en el campo Q1 (donde figura el número de trabajadores por cada actividad) aquellos en los que su tipo es Interno. Además, queremos que tenga un formato numérico. La expresión completa es: Num(Sum({<Tipo = {'Interno'}>}Q1),'#,##0').

A continuación hacemos lo mismo cambiando el tipo por Externo de la siguiente manera Num(Sum({<Tipo = {'Externo'}>}Q1),'#,##0'). Para calcular los trabajadores totales solo tenemos que sumar las dos expresiones descritas.

Para representar estos resultados he optado por un gráfico de tarta en el que la dimensión utilizada será el Tipo. En cuanto a las expresiones solo tiene una y es la suma de internos y externos; en el estilo no usamos el que es un círculo completo sino el que tiene el centro en blanco y habilitamos la leyenda. El resultado es el siguiente:



Figura 6.5.1: Gráfico personal contratado

El segundo de ellos, el reparto de horas, lo calcularemos usando set analysis como en el caso anterior. La fórmula que usaremos es la siguiente Num(Floor(Sum({<Tipo = {'Interno'}>} Q2)),'#,##0') para los trabajadores internos y Num(Floor(Sum({<Tipo = {'Externo'}>} Q2)),'#,##0') para los externos. El total es la suma de ambas expresiones. Se habrá dado cuenta de que en estas expresiones aparece la función **Floor**. Sirve para redondear hacia arriba y no tener decimales en el reparto de las horas. Su representación gráfica es exactamente la misma, pero con su fórmula conjunta.

El tercero de ellos, el importe de las horas, también usaremos set analysis para hallarlo. La fórmula para los trabajadores internos es Num(Floor (Sum ({<Tipo = {'Interno'}>} TOTAL_COST)),'#,##0') y para los trabajadores externos es Num(Floor (Sum ({<Tipo = {'Externo'}>} TOTAL_COST)),'#,##0'). El total es la suma de ambas expresiones y su representación gráfica vuelve a ser un gráfico de tarta con su debida expresión.

Al tratarse de tantos resultados y de tres gráficos si lo mostramos todo a la vez en la hoja queda muy saturado y dificulta su comprensión. Por este motivo usaremos botones para cada uno de los indicadores explicados anteriormente. Al pulsar cada botón, el usuario irá interactuando con los diversos indicadores. Son muy útiles ya que hacen muy amena la experiencia del usuario con el cuadro de mando y además facilita mucho la comprensión de los datos ya que en ningún momento se mezclan resultados de dos indicadores diferentes. Para crear estos botones necesitamos crear una variable. En el panel de variable añadimos una nueva, la llamamos vChart y le damos un valor cualquiera; en este caso un 1.

Luego añadimos tres botones. Esto lo hacemos desde insertar objeto en la hoja, objeto de botón. Les ponemos los títulos de los indicadores y le asignamos la acción de establecer variable, le indicamos la variable vChart que creamos antes y a cada botón le asignamos un valor distinto, en este caso le indicaremos 1, 2 y 3 respectivamente. Luego, deberemos entrar en cada cuadro de texto y cada gráfico y en la pestaña diseño le indicamos mostrar condicional. La condición será vChart =1 para los objetos relacionados con el reparto de trabajadores, vChart =2 para los objetos relacionados con el reparto de horas y vChart =3 para los objetos relacionados con el importe de las horas. De esta manera cuando seleccionamos cada botón aparecerán solo los objetos relacionados con él.

Por último, nos interesa que el color de los botones sea distinto cuando está seleccionado que cuando no, para que así el usuario sepa en todo momento que está viendo. Le asignaremos el color gris cuando esté seleccionado y el rojo cuando no. Para ello deberemos buscar la composición de ambos colores en la escala RGB para poder darle un color calculado. La expresión será lf (vChart = 1, RGB(242,242,242), RGB (245,0,0)) para el botón del reparto de trabajadores; para los demás solo habrá que cambiar el valor de vChart. El resultado final se puede ver en la siguiente ilustración:



Figura 6.5.2: Diseño de hoja con botones

6.6.- Consumo: Conjunto de indicadores que permitirán tener una visión global del consumo de los actos con el objetivo de optimizar y analizar los recursos destinados a los eventos.

En concreto se contabilizará el consumo de agua, electricidad y gas por unidad de consumo e importe.

Comencemos por la electricidad; para calcular su coste deberemos usar una vez más set analysis. Debemos ser conscientes que los csv que estamos utilizando han sido escritos a mano por diversas personas, por lo que es posible que haya erratas. Para que el cálculo sea fiable deberemos hacer el set analysis por el ID del recurso y no por su nombre ya que se minimizarán los errores. Una vez sepamos su ID ya podremos empezar a hallar los indicadores. Aplicaremos esta misma técnica tanto para la electricidad, el agua y el gas.

Comenzaremos por la electricidad; para hallar su gasto y la cantidad consumida haremos un cálculo muy parecido. Partiendo de que tenemos eventos y subeventos con sus consumos respectivamente. Por ello, la expresión para calcular el coste de la electricidad es Num (Round (Sum ({\$<RESOURCE_GROUP_ID = {1744}>} TOTAL_COST) + Sum ({\$<RESOURCE_GROUP_ID_2 = {1744}>} TOTAL_COST_SUB)), '#,##0'), donde 1744 es el numero identificatorio de la electricidad, TOTAL_COST son los costes totales de los eventos y TOTAL_COST_SUB son los costes totales de los subeventos.

Para hallar la cantidad consumida es tan sencillo como usar la expresión anterior y cambiar los campos TOTAL_COST y TOTAL_COST_SUB por los campos Q2 y Q2_2 que son las cantidades totales de los eventos y de los subeventos respectivamente con lo que la expresión final es Num (Round (Sum ({ $\CRESOURCE_GROUP_ID = {1744} > } Q2$) + Sum ({ $\CRESOURCE_GROUP_ID = {1744} > } Q2$) + Sum ({ $\CRESOURCE_GROUP_ID = {1744} > } Q2_2$)),'#,##0')

Ahora es el turno del agua; los cálculos serán exactamente los mismos cambiando el ID del recurso. En este caso, el ID del agua es 1745 por lo que las expresiones serán Num (Round (Sum ({\$<RESOURCE_GROUP_ID = {1745}>} TOTAL_COST) + Sum ({\$<RESOURCE_GROUP_ID_2 = {1745}>} TOTAL_

COST_SUB)) ,'#,##0') para el coste total de agua y Num (Round (Sum({ \COST_SUB)) ,'#,##0') para el consumo total de agua. GROUP_ID_2 = {1745}>} Q2_2)) ,'#,##0') para el consumo total de agua.

Por último, en cuanto al gas, los cálculos volverán a ser los mismos que en los indicadores anteriores. Su ID del recurso es el 1746, así que sus expresiones son Num (Round (Sum ({\$<RESOURCE_GROUP_ID = {1746}>} TOTAL_ COST) + Sum ({\$<RESOURCE_GROUP_ID_2 = {1746}>} TOTAL_ COST_SUB)),'#,##0') para el coste total del gas y Num (Round (Sum ({\$<RESOURCE_GROUP_ID = {1746}>} Q2) + Sum ({\$<RESOURCE_ GROUP_ID_2 = {1746}>} Q2_2)),'#,##0') para la cantidad total de gas.

Ahora que ya tenemos todos los cálculos hechos vamos a representarlo gráficamente. Para ello vamos a utilizar un gráfico de líneas y en su estilo se lo cambiamos por uno de áreas apiladas. En las dimensiones le añadimos el año y el mes. Mostramos su leyenda para que sea más sencillo interactuar con el gráfico. En cuanto a las expresiones le añadimos las del coste de electricidad, agua y gas respectivamente. Les añadimos los nombres en la etiqueta y habilitamos que estas expresiones sean la leyenda. A la hora de asignar los colores la electricidad será amarillo, el agua azul y el gas gris.



Figura 6.6.1: Gráfico consumo

6.7.- % **recogida selectiva**: Este indicador permitirá contabilizar el porcentaje de residuos de recogida selectiva respecto al total de residuos generados en los actos, cuya finalidad es reducir costes y la cantidad de residuos

generados para próximos eventos. La fórmula para su cálculo es: *Kg residuos recogida selectiva / Kg totales residuos*.

Para poder hallar este indicador necesitamos un campo que se encuentra en otro archivo por lo que tendremos que mapearlo. Hacemos un MAPPING LOAD del campo Volumen y lo unimos por el campo RESOURCE_ GROUP _ID ya que es el campo que comparten ambos archivos. La creación de la tabla de mapeo es la siguiente:

MAP_Limpieza:

Mapping LOAD RESOURCE_GROUP_ID,

Volumen

FROM

Ficheros\Neteja.csv

(txt, utf8, embedded labels, delimiter is ';', msq);

Ahora que ya tenemos la tabla de mapeo creada vamos a insertarla en nuestra tabla Recursos de eventos usando la función Apllymap. La expresión completa es Applymap ('MAP_Limpieza', RESOURCE_GROUP_ID,0) as VOLUMEN_EVENTO. Lo que significa esta expresión es que nos inserta la tabla MAP_Limpieza por el campo RESOURCE_GROUP_ID y en caso contrario que aparezca un 0. Le cambiamos el nombre y llamamos al campo VOLUMEN _ EVENTO. Ahora que ya tenemos el campo creado podemos empezar a hallar los indicadores.

El primero de ellos es la recogida total de residuos y se trata de la suma del campo VOLUMEN_EVENTO por lo que su expresión es Num(Sum (VOLUMEN_EVENTO),'#,##0').

Ahora vamos a calcular otro de los indicadores que es el de la cantidad de kilos de basura que se recogieron selectivamente. Para ello haremos un set analysis de la suma dentro del campo VOLUMEN_EVENTO de los recursos envases (se incluyen de vidrio y plástico), orgánicos y de papel y cartón. Con todo ello la expresión completa para su cálculo es Num (Sum ({\$<RESOURCE_4}

= {'80- Contenidor Envasos', '81- Contenidor Orgànica', '82- Contenidors paper cartrò'}>} [VOLUMEN_EVENTO]) ,'#,##0').

Ahora que ya tenemos calculados estos dos indicadores, hallar el porcentaje de recogida selectiva será tan sencillo como dividir la cantidad recogida selectivamente entre la cantidad total recogida de residuos. La expresión será la siguiente Round(Sum ({\$<RESOURCE_4 = {'80- Contenidor Envasos', '81- Contenidor Orgànica', '82- Contenidors paper cartrò'}>} [VOLUMEN EVENTO]) / sum (VOLUMEN EVENTO),0.0001) *100.

A la hora de representarlo no podemos crear directamente el gráfico puesto que, en el campo Selectiva, que nos indica que contenedores se reciclaron y cuáles no, solo aparece los que se reciclaron y los que no, aparecen en blanco.

Para resolver este problema debemos volver a usar la función Trim para así eliminar estos espacios en blanco y poder sustituirlo por un texto que nos ayude a operar. En este caso pondremos NO para así diferenciar los contenedores que se reciclan y los que no. La expresión completa es lf(Len(Trim(Selectiva))<1, 'NO', Selectiva); lo que esta expresión ejecuta es que cuando la longitud del campo Selectiva es menor que uno después de aplicarle la función Trim, o lo que es lo mismo que esté vacío, pues que aparezca un NO y en caso contrario que aparezca lo que figura en el campo Selectiva. Con todo esto ya podemos insertar nuestro gráfico de tarta que contendrá la expresión de la recogida total de residuos y la dimensión Selectiva. Le asignamos el color azul a la recogida selectiva y el rojo a la cantidad que no se recoge selectivamente.



Figura 6.7.1: Gráfico % de recogida selectiva

Además, vamos a añadir un segundo gráfico para complementar más la información. Vamos a reflejar dentro de la recogida selectiva que peso tiene cada una de las tres categorías. Le añadimos la dimensión RESOURCE y en cuanto a las expresiones solo le añadimos la de la recogida selectiva total, le añadimos la leyenda y le asignamos colores en función del color de los contenedores. El resultado es el siguiente:





Al habilitar la leyenda, los nombres aparecen tal y como vienen en el archivo y en este caso está en catalán. Para poder cambiarlo a español debemos usar la función If en una expresión un tanto complicada. En un apartado anterior se explicó en que se basaban los tres argumentos de la función If; el tercero de ellos es el que se usa para en caso contrario a la condición que se da en el primero de ellos, pero también sirve para encadenar otra función If. Por lo tanto, esta expresión significa que en caso de que el recurso sea los contenedores de envases aparezca envases en el gráfico y en caso contrario si el recurso es el contenedor orgánico que aparezca orgánico en el gráfico y en caso contrario si el recurso es los contenedores de papel y cartón que aparezca papel y cartón. La expresión completa es la siguiente: If (RESOURCE_4= '80- Contenidor Envasos', 'Envases', If (RESOURCE_4= '81- Contenidor Orgànica', 'Orgánico', If (RESOURCE_4 = '82- Contenidors paper cartrò', 'Papel y Cartón', RESOURCE_4))).

6.8.- Productos Valor Añadido: Este indicador permitirá contabilizar los productos de valor añadido que se han vendido al cliente final con la finalidad de obtener más ingresos por acto y mejorar la cuenta de explotación.

El indicador deberá mostrar tanto el número de unidades vendidas de cada producto, así como el importe. La fórmula para su cálculo es: *Corporate for brands* + *Corporate for partners* + *Vips Corporate* + *Fast Pack* + *entradas web Anella* + *Vasos vendidos* + *Acciones Especiales*.

Para calcular estos valores volveremos a usar set analysis. En este caso los productos de valor añadido los componen las acciones especiales, la venta de vasos y los VIP Corporate. Estos últimos tienen un coste, por lo que el beneficio de productos de valor añadido es la suma de los ingresos de esas tres categorías menos los costes de los VIP Corporate. Con todo esto la fórmula de la expresión es la siguiente Num (Round (Sum ({\$<CATEGORY = {'Accions Especials', 'Venda Gots', 'VIP Corporate'}>} [TOTAL_INCOME]) - Sum ({\$<CATEGORY = {'VIP Corporate'}>} [TOTAL_COST])),'#,##0').

En cuanto al volumen de productos de valor añadido la expresión es más sencilla ya que no hay nada que restar. La expresión completa es la siguiente: Num (Round (Sum ({\$<CATEGORY = {'Accions Especials', 'Venda Gots', 'VIP Corporate'}>} [Q2])),'#,##0').

A la hora de insertar en la hoja un gráfico que represente bien estos KPI escogeremos el de barras, le añadimos la dimensión EVENT_NAME para ver cada evento, aunque en este trabajo estén cifrados, le añadimos las expresiones anteriores, le habilitamos la leyenda y le asignamos el color gris para el beneficio y el azul para el volumen. El resultado final es el siguiente:



Figura 6.8.1: Gráfico beneficio y volumen productos de valor añadido

También vamos a añadir una tabla donde se desglose las cantidades de cada categoría para así dar más visibilidad de donde salen los KPI. Para ello insertamos una tabla simple, le añadimos las expresiones de beneficio y volumen descritas anteriormente. Al estar los nombres de las categorías en catalán los cambiamos a español usando la función If como ya hicimos en el apartado anterior. La expresión para modificar las etiquetas es If (CATEGORY= 'Accions Especials', 'Acciones Especiales', If (CATEGORY= 'VIP Corporate', 'VIP Corporativo', If (CATEGORY= 'Venda Gots', 'Venta de vasos', CATEGORY))). Por último, le añadimos un fondo gris para que resalten más los datos. El resultado final es el siguiente:

Categoría	Beneficios V.A. 🛆	Volumen V.A.
Acciones Especiales	12,656	7,460
VIP Corporativo	78,284	76,284
Venta de vasos	131,978	145,253
TOTAL	222,917	228,997

Tabla 6.8.2: Productos de valor añadido

6.9.- % **Reservas Canceladas**: Este indicador permitirá cuantificar el porcentaje de reservas canceladas con respecto al total de reservas, con el objetivo de minimizar el número de reservas canceladas y conocer las razones por las cuales las reservas o pre-reservas fueron canceladas. Este indicador es el único que tiene en cuenta las reservas canceladas, todos los demás indicadores no las consideran. La fórmula para el cálculo de este indicador es: *Volumen reservas canceladas / Volumen total reservas*.

Lo primero de todo debemos hallar el volumen total de eventos y cuántos de ellos fueron cancelados para poder calcular el ratio. Para calcularlo deberemos usar set analysis para contar todos los eventos menos los que están bloqueados ya que esos no se consideran en este ratio. La expresión completa es COUNT({\$<TIPO_OCUPACION = {'Cancelados', 'Pre-reserva', 'Reserva'}>} STATUS).

El cálculo de los eventos cancelados es exactamente igual que el del total pero en el set analysis solo mantenemos los cancelados por lo que la expresión es Count({\$<TIPO_OCUPACION = {'Cancelados'}>} STATUS).

Por lo tanto, el ratio será tan sencillo como dividir los cancelados entre el total, por lo que la expresión es Round(COUNT ({\$<TIPO_OCUPACION = {'Cancelados'}>} STATUS) / COUNT ({\$<TIPO_OCUPACION = {'Cancelados', 'Pre-reserva', 'Reserva'}>} STATUS),0.0001) * 100.

Para representarlo volveremos a usar un gráfico de tarta ya que es el más fácil de comprender. En él añadiremos la expresión del total de eventos, la dimensión TIPO_OCUPACION y le daremos el color rojo a los eventos cancelados, el azul para los pre-reservados y el gris para los reservados. El gráfico resultante es el siguiente:



Figura 6.9.1: Gráfico % de reservas canceladas

6.10.- Consumo Per Cápita: Este indicador permitirá analizar el consumo promedio per cápita de los usuarios que han asistido a eventos en la Anella Olímpica, con el objetivo de maximizar el beneficio en restauración y tiendas.

El consumo per cápita contabilizará el precio de la entrada, el consumo en restauración y en tiendas. Se considerará la entrada como consumo externo y el consumo de restauración y tiendas, como interno. La fórmula para su cálculo es: [(Recaudación Promotor) + (Recaudación Restauración) + (Recaudación Tiendas)]/ Volumen Usuarios.

Antes de comenzar con su cálculo debemos modificar tres campos que son necesarios para hallar estos indicadores. Se trata de los campos PromoterIncome, FBIncome y MerchIncome. Los datos son cantidades de dinero y vienen con la siguiente estructura (1 000€) por lo que deberemos modificarlos para poder empezar a calcular. En principio, usando una función replace para eliminar el espacio y el símbolo del euro debería ser suficiente, pero no afecta al espacio. Esto es debido a que no era un espacio, sino el separador de miles americano, pero como nosotros no lo tenemos vemos un espacio. Usando la tabla ASCII descubrí que se trataba de los caracteres 194 y 160 por lo tanto la expresión completa para tratar estos campos y poder empezar a calcular con ellos es: Num (Replace (Replace (Replace (Replace (PromoterIncome, '€', ''), chr (194), "), chr (160), "), ',', '.'), '#,##0') as PromoterIncome. Esta expresión es bastante compleja por lo que vamos por partes. Usamos Replace (PromoterIncome, '€', ") para que elimine el símbolo de euro que nos afectaba a los cálculos, luego utilizando Replace(Replace, chr (194), "), chr (160), ") eliminamos ambos caracteres que componen el espaciador de miles americano para así ya poder empezar a calcular los indicadores. Por último, usando Num (Replace, ',', '.'), '#,##0') reemplazamos la coma por punto ya que tenemos nuestros datos en formato anglosajón y le damos formato numérico.

Comencemos calculando el consumo per cápita del precio de la entrada. La fórmula será tan sencilla como dividir los ingresos de los promotores entre el número de usuarios y recordemos que el número de usuarios era la suma del número de visitantes y el número de asistentes. Por lo tanto, la expresión será

35

Round(Num (PromoterIncome) / (Sum (NumberOfVisitors) + Sum (Attendance)), 0.01).

Continuamos con el consumo per cápita en restauración. Su divisor será el mismo que en el anterior KPI, pero el numerador serán los ingresos de restauración. Su expresión es Round (Num (FBIncome) / (Sum (NumberOfVisitors) + Sum (Attendance)), 0.01).

El tercero de ellos se trata del consumo per cápita en tiendas. Una vez más su divisor será el mismo y el numerador los ingresos de las tiendas. Su expresión es Round (Num (MerchIncome) / (Sum (NumberOfVisitors) + Sum (Attendance)), 0.01).

Por lo tanto, el consumo per cápita será la suma de los tres KPI anteriores. Su expresión es: Round ((Sum (PromoterIncome) + Sum (FBIncome) + Sum (MerchIncome)) / (Sum (NumberOfVisitors) + Sum (Attendance)), 0.01).

Para representarlo usaremos una vez más un gráfico de tarta. En este caso no le añadiremos ninguna dimensión y en las expresiones le añadiremos las del promotor, tiendas y restauración. Mostramos la leyenda y le asignaremos el color gris al promotor, el rojo a las tiendas y el azul a la restauración.



Figura 6.10.1: Gráfico reparto consumo per cápita

Una vez vistos todos los KPIs y la manera de hallarlos podemos pasar a los apartados de diseño del cuadro de mando y conclusiones.

7.-Diseño

Una vez tenemos todos los KPIs calculados con sus debidas representaciones comenzamos a realizar el diseño del cuadro de mando. Puede parecer que el diseño no es algo importante, pero los cuadros de mando deben ser lo más agradables e intuitivos posibles para los usuarios.

A diferencia de la mayoría de los documentos empresariales que deben ser sobrios e impersonales, en los cuadros de mando ocurre todo lo contrario. Debemos ser conscientes de que el 70% de la información que procesamos es a través de la vista; por este motivo es tan importante los colores que usemos, así como la organización de la información en las diferentes pestañas de nuestro cuadro de mando (Few, S. (2006)). Además, los cuadros de mando suelen ser muy utilizados por los altos cargos de las empresas para consultar información rápida por lo que la organización de la información es crucial.

A la hora de elegir los colores que vamos a utilizar debemos de ser muy cuidadosos; puesto que es muy frecuente pensar que cuanto más colorido mejor y eso es incorrecto ya que la principal cualidad que debe tener un cuadro de mando es que sea de fácil comprensión y si lo saturamos de colores influirá negativamente. Nuestro cerebro procesa los colores y les asigna un significado. Por ejemplo, el color rojo lo asociamos con el peligro por lo que se suele usar para reflejar costes o pérdidas. Colores más fríos (como el gris o el azul) se usan para reflejar cantidades o ingresos. Colores muy fuertes como el amarillo y ciertos tonos del verde no suelen ser aconsejables nunca, pero en ocasiones hay peticiones expresas por parte del cliente que incumplen esta teoría.

En cuanto a la organización de la información debemos de ser cuidadosos y no saturar al usuario. La forma más fácil de interpretar la información es mediante gráficos, pero el uso excesivo de estos va a ser contraproducente. Cuantas menos pestañas tenga un cuadro de mando más rápida será su consulta y por tanto se presupone más eficiente. Debemos tener cuidado con esto, ya que es muy común ver hojas con muchísima y diversa información y es preferible repartirla en diversas pestañas para facilitar su comprensión (Few, 2006).

37

También es crucial la importancia a la hora de elegir los objetos de visualización que utilicemos. Los gráficos de tarta son los más usados, pero cuando vamos a representar un indicador compuesto por muchos elementos deja de ser útil ya que es muy difícil de interpretar y es preferible usar un gráfico de barras. Además, los márgenes se suelen usar para incluir los logos de las empresas y los diversos filtros que solicite el cliente para así dejar la zona central a los indicadores.

Con todas estas pautas a la hora de diseñar un cuadro de mando procedo a la explicación del diseño de mi proyecto.

A la hora de realizar el diseño del cuadro de mando se suele optar por usar los colores de la empresa, en este caso el rojo y el negro, para que así se sientan identificados; puesto que son los colores institucionales del Ayuntamiento de Barcelona. En algunos indicadores es un problema ya que ambos colores se relacionan con cosas malas; en este contexto a pérdidas. Por este motivo he usado los colores gris, azul y beige ya que son muy vistosos e intuitivos para el usuario. El rojo lo he usado para cosa negativas como pérdidas, gastos o reservas canceladas.

En cuanto al fondo le cambiamos el color predeterminado por el blanco, añadí el logo de BSM arriba a la izquierda, ya que es la empresa para la que se hizo este proyecto y también el del ayuntamiento de Barcelona puesto que BSM es una empresa municipal y depende del ayuntamiento. Los logos son los siguientes:

B-SM de Serveis

Barcelona Municipals

Figura 7.1: Logo de BSM



Figura 7.2: Logo del ayuntamiento de Barcelona

También le vamos a añadir una barra gris para que resalten más los logos y además una barra roja ya que si solo dejamos la barra gris queda un poco frio el diseño. La barra roja será exactamente el mismo tono que el del escudo del ayuntamiento de Barcelona. Para poder obtener el tono exacto usaremos el ColorPix.exe, de esta manera podemos saber que ese color tiene 218 de rojo, 26 de verde y 36 de azul en la escala RGB.

Ahora comenzamos a colocar los filtros. En la parte de arriba colocaremos dos cuadros de lista; uno con los años y otro con los meses. Los colocamos de manera horizontal y centrados en la pantalla ya que van a ser los filtros más utilizados. Es muy común usar cuadros de mando para consultar periodos de tiempo y realizar comparaciones; por ese motivo lo colocamos ahí.



En la parte izquierda añadiremos los filtros solicitados por el cliente. En este caso se trata del nombre de los eventos, el estado de estos, los promotores, la instalación donde se realizan y el tipo de acto. En este caso no podemos desplegarlos como hicimos con el año y el mes ya que son muchísimos y no sería operativo. Por este motivo vamos a crear unos desplegables para permitir hacer selecciones múltiples de una manera sencilla y eficaz. Para crear estos

desplegable debemos insertar un gráfico y escogemos que sea una tabla pivotante. Luego, le añadimos la dimensión que queramos que aparezca en el desplegable, le habilitamos el selector desplegable y no añadimos ninguna expresión ya que no queremos calcular nada ahora mismo. Como ya expliqué al comienzo de este trabajo los nombres de los eventos, los promotores y los tipos de actos están cifrados para así no incumplir con la ley de protección de datos.

Para mejorar aún más la experiencia del usuario si seleccionamos algún filtro lo más practico es que se refleje lo que estamos filtrando; para ello deberemos añadir una expresión en su etiqueta: If (GetSelectedCount (EVENT_NAME) <1, 'Nombre Evento', GetFieldSelections (EVENT_NAME, '|', 30)). Esta expresión lo que significa es que cuando 'EVENT_NAME' tengo un valor inferior a 1 (cuando no esté seleccionado el filtro) aparezca Nombre evento y cuando ocurra lo contrario (que sí esté seleccionado) aparezca lo que se está seleccionando, si se selecciona más de un elemento a la vez que haya una barra separándolos y que como máximo puedan aparecer 30 simultáneamente.

De esta manera logramos un diseño sencillo y a la vez intuitivo para que cualquier tipo de usuario sin necesidad de conocer esta herramienta pueda navegar, consultar y filtrar los datos de manera sencilla.

Con todo ello nuestro diseño final de las hojas del cuadro de mando es el siguiente:



Figura 7.5: Diseño hoja cuadro de mando

8.-Conclusiones

Este trabajo refleja la importancia del Business Intelligence y más concretamente de los cuadros de mando en el mundo empresarial actual. Cada día las empresas tienen más datos, pero la mayoría de las veces no saben qué hacer con ellos ni cómo gestionarlos, de ahí la creciente demanda en especialistas de todo tipo relacionado con los datos, como ingenieros, economistas o matemáticos.

A lo largo de estas páginas hemos podido comprobar que las herramientas de visualización, en este caso QlikView, aportan muchísimo valor ya que cualquier empleado puede ver los indicadores (KPI) más importantes de la empresa en cualquier momento, así como aplicar los filtros que le sean oportunos. A la hora de tomar decisiones estratégicas es crucial tener esta información a mano, ya que si solo usamos ficheros de Excel va a ser mucho más difícil ser conscientes de la situación real de la empresa, además de que la carga de datos desde Excel, cuando tenemos una gran cantidad de ellos, genera problemas y demoras.

Las conclusiones que podemos extraer del proyecto son las siguientes:

En el KPI del porcentaje de días ocupación podemos observar que los espectáculos y festivales están de moda, lo que se refleja en la gran cantidad de días que se encuentran ocupadas las instalaciones. Este KPI es muy interesante porque refleja perfectamente los picos de consumo que ocurre en casi todos los sectores. Apenas hay espectáculos en marzo y abril (semana santa), julio y agosto (vacaciones), así como septiembre (la vuelta al colegio, el gasto de los ahorros en las vacaciones, ...). Por el contrario, vemos que los principales picos se producen en mayo y junio (ya hace buen tiempo para los espectáculos y la gente aún no se fue de vacaciones) y en la época próxima a la navidad. También podemos fijarnos en que el Palau Sant Jordi y el Sant Jordi Club son las instalaciones que más eventos reciben, esto es debido a que no son al aire libre y en invierno prácticamente todos los eventos se realizan en estos dos recintos.

- En lo relativo al margen de los eventos podemos apreciar que los espectáculos gozan de buena salud ya que dejan un beneficio cercano a los 4 millones tanto en 2016, como en 2017. Debemos ser conscientes de que estos datos son de abril de 2018, por eso en las cantidades de los eventos que son a posteriori es difícil poder sacar conclusiones.
- En cuanto al número de usuarios también registra unas grandes cifras, casi 900.000 usuarios en 2016 y 2017. Los macroconciertos de artistas internacionales ayudan a estas altas cifras, así como los eventos diurnos del festival Sónar.
- El KPI del absentismo de espectadores refleja unos datos un tanto difíciles de interpretar. Si bien el absentismo es bajo por regla general, los eventos benéficos hacen que se distorsione este indicador. Un ejemplo curioso es un evento que vendió una enorme cantidad de entradas, pero sin embargo tuvo una tasa de absentismo bastante grande que hace distorsionar el resultado final. Esto se debe a que el importe de la entrada se destinaba a una causa benéfica, por lo que mucha gente que la apoyaba y que no reside en Barcelona compró la entrada y luego no asistió.
- En lo relativo a los KPI del personal contratado podemos observar que la gran mayoría de empleados son subcontratados, pero sin embargo en el reparto de horas no hay esa diferencia tan grande. Esto se debe a que los empleados internos se encargan de los trabajos más importantes y por tanto que más horas requieren mientras que los externos se encargan de tareas como seguridad durante los eventos, limpieza o cátering.
- Referente al consumo los resultados son los esperados. El gasto en electricidad supera con creces al del agua y gas siendo su coste de 442.064€ frente a los 49.329€ del agua y los 33.759€ del gas.
- Respecto a la recogida selectiva podemos concluir que aún queda mucho por mejorar ya que apenas la mitad de los residuos que se generan en los espectáculos se reciclan, 50,98% en 2016 y 51,87% en 2017 con el objetivo de ir reduciendo la generación de residuos, así como sus costes.
- En relación con los productos de valor añadido reflejan buenos beneficios, 108.851€ en 2016 y 100.911€ en 2017. Podemos observar que las acciones especiales como pueden ser los meet & greet apenas generan

beneficios, en 2016 no hubo y en 2017 solo 12.056€. Los VIP sí generan unos beneficios más cuantiosos principalmente relacionado con grandes artistas internacionales generando unos beneficios de 56.055€ en 2016 y 20.509€ en 2017. Lo que mayores beneficios genera, dentro de los productos de valor añadido, es la venta de bebidas; 52.795€ en 2016 y 68.347€ en 2017. Es curioso ver como se distribuyen, ya que no hay una correlación en función del número de usuarios, sino en el tipo de espectáculo. Por ejemplo, los conciertos de artistas pop con miles de usuarios no generan abundantes beneficios en la venta de bebidas y sin embargo en los conciertos de rock, aun teniendo muchos menos usuarios, se obtienen beneficios cuantiosos ya que los espectadores consumen muchas más bebidas de media.

- Por lo que concierne al porcentaje de reservas canceladas vemos que los resultados son curiosos. En 2016 solo el 15,61% de las reservas se cancelaron mientras que en 2017 asciende al 71,65%. Esto se debe a que el volumen total de eventos casi se triplicó, por lo que es muy común en este negocio que se hagan reservas con años de antelación y por diversos motivos luego se cancelen sin que se haya llegado siquiera a promocionarlo ni a vender entradas.
- En consideración al consumo per cápita podemos percatarnos de que casi todo está relacionado con el consumo de los promotores mientras que la restauración y las tiendas generan bastante menos en comparación con lo generado por los promotores. Esto se debe a que todo el mundo tiene que comprar entrada para acceder a los eventos, pero no es necesario que consuman ningún alimento o bebida ni que compren algo de merchandising. Además, muchos eventos se realizan por la tarde, sobre todo actividades infantiles y conciertos de artistas orientados a adolescentes, por eso el consumo per cápita de la restauración es tan bajo.

En conclusión, podemos ver como el negocio de los espectáculos en Barcelona está en auge y no para de dar beneficios cuantiosos. Si solo nos centrásemos en los beneficios nos estaríamos equivocando ya que debemos analizar todas las causas que los generan para así dar una explicación de cara a conseguir una maximización de los ingresos y una minimización de los costes consiguiendo atraer los espectáculos que más entradas venden, así como los que más ingresan en productos de valor añadido, restauración y tiendas.

Un análisis como el que se ha hecho sería imposible sin una herramienta de Business Intelligence, ya que usando muchos libros de Excel simultáneamente habría un montón de factores que se nos pasarían por alto, además de ser una herramienta más incómoda en lo que se refiere al uso de los filtros, muchas veces necesarios para comprender las causas de la situación de la empresa.

9.- Bibliografía

Cano, J (2007): "BUSINESS INTELLIGENCE: Competir con Información", Fundación Cultural BANESTO: 100-105, 351-352

Few, S. (2006): "Information Dashboard Design", https://perceptualedge.com: 30, 36-45

García, M. Harmsen, B. (2013): "QlikView 11 for Developers", Packt Publishing: 8, 88-92

Gartner (2019): IT Glossary: https://www.gartner.com/it-glossary/businessintelligence-bi (Consultada el 18 de julio de 2019)

Parmenter, D. (2015): "KEY PERFORMANCE INDICATORS: Developing, Implementing and Using Winning KPIs", John Wiley & Sons: 7-11

Wang, L. Wang, G. y Alexander, C. (2015): "Big Data and Visualization: Methods, Challenges and Technology Progress", Digital Technologies, vol. 1, no. 1: 33-38