

3. IMPLEMENTAR UN PROCESO ETL

En este tema vamos a ver qué pasos debemos seguir para implementar los procesos ETL con éxito.

Durante la ejecución de un proyecto así, debemos llevar a cabo el análisis en el que se van a definir los orígenes de datos, el destino, la programación de las cargas y también podremos detectar los posibles problemas de acceso a los datos.

Tras realizar la etapa de análisis, se llevará a cabo la etapa de desarrollo en la que vamos a implementar la solución requerida. En esta etapa de desarrollo es en la que se va a usar la herramienta open source Pentaho Data Integration o Kettle y nos centraremos en tres aplicaciones que tiene la suite para la programación y el diseño de cargas. Estas aplicaciones son:

- **Spoon:** entorno gráfico que nos permite la creación de trabajos y transformaciones.
- **Pan:** nos permite ejecutar transformaciones.
- **Kitchen:** nos permite ejecutar trabajos.

Y por último vamos a ver cómo evaluar el rendimiento de las cargas. Esta etapa es importante llevarla a cabo debido a que la ejecución de los procesos puede perder rendimiento conforme pasa el tiempo a causa de que la volumetría de los datos iniciales puede cambiar o porque en el enfoque que se había realizado al comienzo, ahora ya no es factible aplicarlo.

3.1 ANÁLISIS DEL PROCESO DE CARGA

Cuando estamos creando un proceso con una herramienta ETL, las primeras acciones que debemos llevar a cabo son las siguiente:

- Analizar el origen de los datos a cargar.
- Ver qué salida de datos tenemos que generar.

EL PROCESO DE ETL

- Diseñar las operaciones que necesitamos para lograr que los datos que deben cargarse se conviertan en los datos que espera el negocio.
- Establecer la periodicidad de las cargas.
- Estudiar el volumen de los datos.
- Definir si la carga es incremental o total.
- Definir la materialización de los kips year to date,previous year.

La implementación de los procesos será más sencilla cuanto mejor sea el análisis que habremos llevado a cabo. Veamos distintas fases de análisis:

- **Análisis de los orígenes de datos a cargar:** en esta fase tenemos que detectar qué datos son los que se desean cargar, en qué sistema se encuentran y ver si tenemos acceso a esos datos o no. Nos podemos encontrar con las siguientes situaciones:
 - Los datos se encuentran en un Excel, pero la herramienta ETL que utilizamos para cargar no admite este formato.
 - Necesitamos acceso a datos de un servicio web, pero el servidor, en el que está instalado la ETL, no tiene salida a internet.
 - Los datos que se cargan en el Data Warehouse son de un Excel, pero el usuario que los mantiene no puede acceder a las ubicaciones en las que el servidor ETL carga los datos.
 - Los datos que se deben cargar se encuentran en un CRM y no disponemos de conector para acceder.
 - Los códigos en los distintos sistemas orígenes no se codifican del mismo modo. En un sistema origen el código se encuentra guardado en mayúsculas, en otro en minúsculas, con espacios en blanco al final, con espacios en blanco a la derecha, etc. Si suceden estos casos, hay que añadir la lógica en la ETL para convertir los códigos a códigos idénticos.
- **Análisis de los datos destino:** aquí hay que tener en cuenta si los datos los cargamos en una base de datos relacional para la creación de un Data Warehouse, si los datos los cargamos en un entorno transaccional, si generamos un fichero en una ubicación de FTP o en una ubicación de nuestra

red. Si vamos a crear el Data Warehouse (DWH) se van a necesitar credenciales de la base de datos para la creación del modelo de datos que se van a crear. Si utilizamos la ETL como una herramienta de integración contra un transaccional, vamos a necesitar saber las credenciales de esta base de datos, además de la estructura de las tablas que tenemos que cargar, si es FTP sus credenciales, etc. Dependiendo del destino de los datos vamos a hablar con unos u otros responsables de la organización.

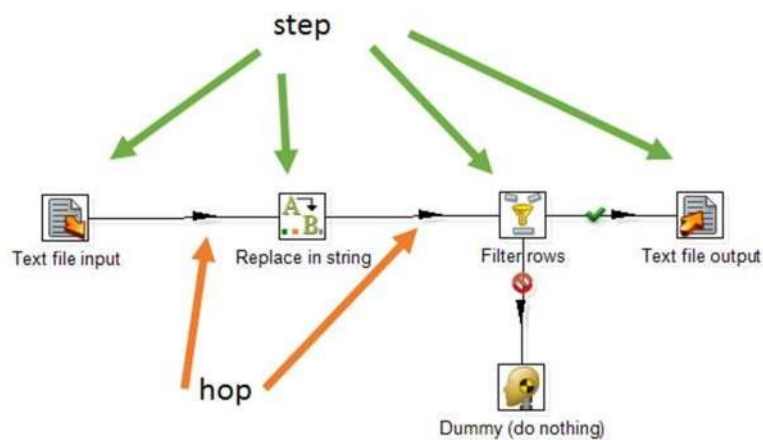
- **Diseño de las operaciones necesarias para la conversión de datos:** según los datos origen y los datos destino, se van a usar unas transformaciones u otras existentes en la herramienta, orígenes, salidas, agrupaciones, ordenaciones, conversiones de cadenas, fórmulas, uniones, etc.
- **Definición de la periodicidad de las cargas:** hay que definir cada cuánto tiempo se van a ejecutar los procesos definidos; si hay un evento que las lanza, si son ad-hoc, si se ejecutan el primer día de mes, si se hacen diariamente, si se lanzan cada martes, el tercer día del mes en época de verano, etc. Si queremos explotar los datos con el detalle de cada mes y el mes actual no se quiere llevar a cabo el análisis, podemos programar las cargas para el primer día de cada mes.
- **Estudiar el volumen de los datos:** según las especificaciones de nuestros servidores y el volumen de los datos, vamos a tener que crear tablas de agregados para que las herramientas de explotación de los datos tengan un buen rendimiento. Puede ocurrir, por ejemplo, que nuestra empresa realice multitud de líneas de facturas a diario y que estos datos se tengan que mostrar con un detalle diario y que además existan unos cuadros de mandos que muestren información mensual. Así, posiblemente deberemos llevar a cabo una tabla de ventas a nivel diario y otra a nivel mensual. Llevar a cabo estas agregaciones tiene como ventaja que las herramientas de explotación de datos van a tener un mejor rendimiento, pero con el inconveniente de que el modelado de los datos será más complejo y habrá más trabajo en la implementación de ETL.

- **Definir cargas totales frente a incrementales:** se llama carga total a aquella que carga todos los registros de las tablas orígenes en la tabla destino y carga incremental es aquella que, únicamente, carga aquellos registros que se han modificado o insertado desde la última carga realizada. Para implementar la segunda opción se necesita que los orígenes tengan los registros auditados, dicho de otro modo, necesitamos saber en qué fecha se ha modificado el registro. Si no disponemos de ese campo de auditoría, tendremos que definir otro campo de fecha y definir algún rango de fechas para cargar. Si, por ejemplo, las facturas no tienen fecha de modificación se podrían cargar aquellas facturas que se han creado en los últimos tres meses, entendiendo así que una factura anterior a tres meses no va a sufrir modificación en el sistema origen.
- **Definir la materialización de los KPIs, year to date, TAM, previous year:** a menudo en la toma de requisitos del proyecto, se llega a la conclusión que el usuario final necesita tener calculados los KPIs de YTD, TAM o previous year. Para ofrecer al usuario esta funcionalidad, podemos implementar en el proceso de carga u ofrecer esta funcionalidad con la herramienta que utilicemos para modelar los cuadros de mandos y los informes. La implementación de estos KPIs en la base de datos va a hacer que el proceso sea algo más complejo, que la validación de los datos tenga aun coste más alto, que la base de datos tenga más volumen de datos y que la venta de tiempo para recargar los datos sea más amplia.

Debemos comentar que estas tareas comentadas anteriormente hay que llevarlas a cabo independientemente de la herramienta ETL que finalmente utilicemos para crear procesos de carga. Cada ETL tiene sus puntos fuertes y sus debilidades, y es trabajo nuestro identificar cuál es la que mejor se adapta a nuestras necesidades en cada momento.

3.2 TRANSFORMACIONES

Como ya hemos comentado anteriormente, una transformación es un conjunto de pasos o steps que pueden o no conectarse entre sí, dependiendo de su relación mediante lo que llamamos saltos, que procesan registros a partir de un origen de datos y cuya salida son uno o más registros de datos.



**Representación gráfica de steps y hops. Fuente:*

<http://integrandodatos.blogspot.com.es/2013/06/steps-transformations-y-jobs.html>

Cada salto o hop nos permite llevar a cabo un proceso de entrada, transformación o salida de registros. Pentaho Data Integration se puede conectar prácticamente a cualquier origen de datos.

EL PROCESO DE ETL



**Representación gráfica de los orígenes de datos más utilizados. Fuente:*

<http://integrandodatos.blogspot.com.es/2013/06/steps-transformations-y-jobs.html>

A continuación, ofrecemos un ejemplo sobre cómo crear una transformación mediante Pentaho.

Autor: M. Moreno, 2015.

Enlace: <https://about.sofia2.com/2015/02/26/como-generar-una-transformacion-en-kettle-de-una-coleccion-de-mongodb/>

Kettle, ahora conocido como PDI (Pentaho Data Integration) tiene una interfaz intuitiva y gráfica, de arrastrar y soltar. Por su facilidad de manejo para usuarios menos técnicos y una arquitectura probada, escalable y basada en estándares, es cada vez más una gran opción para las empresas de ETL tradicional que buscan herramientas de integración de datos.

A continuación vamos a crear una transformación a partir de una colección de datos de Mongo DB. Para ello vamos a utilizar:

MongoDB: Necesitaremos que esté arrancada la base de datos y saber la dirección y el puerto.

EL PROCESO DE ETL

Kettle: Una versión para usuario conocida como Spoon desde donde descargaremos la información de MongoDB.

1. Insertamos una nueva transformación y arrastramos del menú lateral de Diseño la herramienta MongoDB Input.

Pinchando dos veces en ella accedemos a las opciones de MongoDB. Necesitaremos tener arrancada la base de datos y conocer la colección, el puerto y la dirección.

En nuestro ejemplo hemos incluido una query en formato json para la extracción de datos que está parametrizada por Provincia. Este parámetro lo utilizaremos en el informe para filtrar la información.



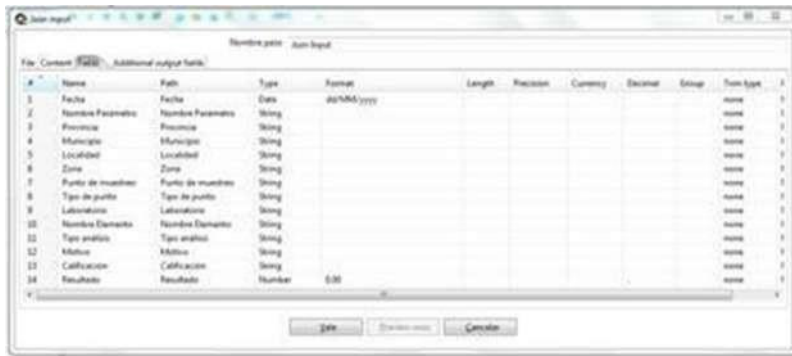
2. Definimos el parámetro de la transformación en 'Editar'-'>'Configuración' e indicamos el valor por defecto:

transformación Parameters Archivado Fechas Dependencias Miscelaneos Monitoring				
Parameters:				
#	Parameter	Default Value	Description	
1	Prov	Araba		

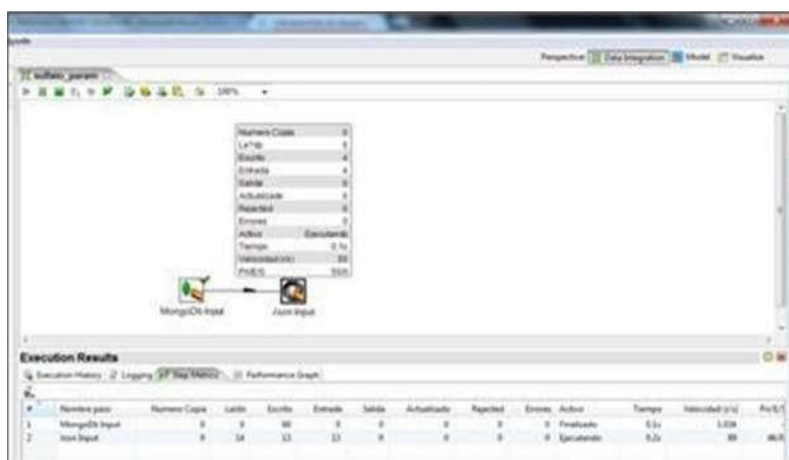
3. Editamos la información la herramienta Json Input.

Informamos los campos que queremos recuperar para realizar el informe indicando el formato y el nombre.

EL PROCESO DE ETL



Unimos los dos iconos y podemos previsualizar el contenido.

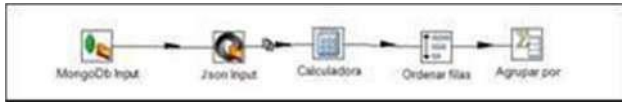


Con esta transformación sería suficiente para generar el informe. No obstante, realizaremos tres pasos más para conseguir más información de valor.

4. Utilizamos la calculadora para definir los campos Año, Mes y Día, a partir de la Fecha.

5. Realizamos el promedio del campo Resultado con la herramienta agrupar por, que para utilizarla deberemos haber ordenado previamente por la clave.





Ejecutamos la transformación y la guardamos.

Hemos realizado un breve incursión por algunas de las herramientas de Kettle, pero podríamos realizar cálculos más complejos, ya que ofrece muchísimas posibilidades.

Algunos enlaces para obtener más información sobre lo que aquí se explica son:

[Spoon User Guide: Transformation Steps](#)

[Pentaho:Create a Parameterized Report with MongoDB](#)

Además de este ejemplo, adjuntamos el siguiente enlace con un ejemplo de transformación con Pentaho Data Integration:

Enlace: <https://www.youtube.com/watch?v=Az0M07E33Jo>

3.3 TRABAJOS

Como también habíamos comentado antes, un trabajo es el conjunto de trabajos, transformaciones y pasos que se ejecutan de forma secuencial. El Trabajo o Job maneja secuencias de tareas, no registros. Con Pentaho se puede programar un Job de dos modos diferentes: por medio del entorno de desarrollo gráfico (Spoon) o con Kitchen.

En relación a la creación de trabajos, en el apartado de anexos hemos dispuesto algunos enlaces con ejemplos prácticos.

3.4 EVALUACIÓN

Implementado el proceso de carga y habiendo comprobado que el resultado es correcto, debemos evaluar el rendimiento de las cargas. ¿Cuánto tiempo tarda la ejecución de toda la carga?, ¿Qué paso es el más lento?, ¿Qué mejoras hay que llevar a cabo?

Como normalmente no se dispone de un tiempo ilimitado para llevar a cabo la ejecución del proceso, es necesario saber cuánto tiempo tarda en hacerse la carga. Es muy normal que la ejecución de las cargas se lleve a cabo durante la madrugada porque es un periodo de tiempo en que los servidores no tienen mucha actividad y los datos no varían. Si es así, las cargas que hemos implementado deberían ejecutarse en unas seis o siete horas.

Un proceso que en un inicio tarda un tiempo aceptable, tras unos años puede ser inaceptable, por ello, esta evaluación debe perdurar en el tiempo. Si no se ha modificado nada en el diseño, ¿a qué se debe este cambio de rendimiento? Probablemente se deba al volumen de datos que movemos en el proceso implementado, que ha variado del “scope” inicial al actual.

- Si pasamos de vender en nuestra empresa 10.000 líneas de facturas al día a 1.000.000 lo más seguro es que el proceso de carga aumente.
- Si el proceso que hemos implementado no se ejecuta dentro del tiempo establecido, deberemos ampliar esta ventana de tiempo o disminuir el volumen de datos que debemos tratar o rediseñar el proceso de carga, ampliar la dimensión de los servicios, etc.

En Spoon podemos ver el resultado de las ejecuciones en el panel de *execution results*. En la pestaña *logging* vamos a observar a qué hora se ha iniciado la ejecución y a qué hora finaliza la misma. En la pestaña de *Step Metrics* podemos consultar el tiempo que sucede en cada paso, además de las estadísticas de registros que entran en cada paso, registros rechazados, registros que salen, etc.

Como ya hemos visto anteriormente, podemos configurar las transformaciones ya existentes en los Jobs creados para enviar el resultado de sus ejecuciones a un fichero de texto.

También podemos llevar a cabo la configuración de la herramienta para tener un repositorio con las trazas de las transformaciones que se han ejecutado.

Debemos comentar que la configuración de este archivado puede provocar que la ejecución de los procesos se lleve a cabo de forma algo más lenta.

3.5 CREACIÓN DE DASHBOARDS

En el apartado anterior hemos hablado sobre la creación de un modelo de datos y la estructuración de los datos en tablas mediante procesos ETL. Así, en este último apartado vamos a ver de qué manera coger estos datos estructurados, relacionarlos entre sí y visualizarlos de manera gráfica con el fin de poder sacar toda la información necesaria para nuestros análisis.

Para ello nos centraremos en el estudio de los dashboards, también llamados cuadros de mando. En líneas generales se trata de una herramienta de visualización de datos que permite mostrar, mediante la representación gráfica de los diferentes KPI's o indicadores, el estado de un negocio. Además de visualizar el estado de nuestro negocio, se trata de una herramienta muy útil en la toma de decisiones, ya que nos enseña de manera gráfica y sencilla aquello que va según lo previsto y aquello que no, así como las tendencias que están a la alza y a la baja. Ayudándonos ello a redirigir nuestros objetivos y nuestro negocio hacia el logro de mejores beneficios.

A modo de resumen, la creación de un dashboard nos reportará los siguientes beneficios:

- Se enfoca en la toma de decisiones, lo que implica que se optimicen todas las estrategias de la empresa. No obstante, cabe remarcar el hecho de que el dashboard no sirve para definir acciones correctivas, sino preventivas.
- Permite llevar a cabo un seguimiento adecuado y recurrente del grado de cumplimiento de los objetivos establecidos, tanto a nivel global de toda la empresa como a nivel concreto de un departamento o proyecto específico.
- Refleja los principales indicadores inmersos en el negocio.
- Actúa como guía para identificar el origen de un dato específico, tanto si es positivo como negativo. Además de ello, permite la detección adecuada de los datos para facilitar la toma de decisiones.

EL PROCESO DE ETL

- Analiza en detalle aquellas causas por las que se generaron dichos datos.
- Permite contemplar todas las causas posibles que han generado algún problema.

Como veremos, en la actualidad disponemos de un gran número de herramientas que nos lo generan automáticamente. No obstante, en la creación de los dashboards, debemos tener en cuenta los siguientes elementos:

- El primer elemento a considerar son las KPI's o indicadores que se tratarán en el dashboard. Aquí es muy importante que incluyamos solo las adecuadas, de modo que será necesario realizar una búsqueda exhaustiva y profunda de aquellas palabras clave que añaden valor adicional a nuestro negocio, ya que serán estas las que nos ayudarán a posicionarnos y a darnos información relevante sobre la empresa y el proyecto.
- Al tratarse de una herramienta visual, que nos aporta información realmente valiosa para nuestro negocio, es muy importante que la información y los datos estén bien estructurados, de modo que su interpretación será rápida y sencilla.
- Ligado con el elemento anterior, el dashboard debe ser comprensible. Eso es, debe facilitarnos la contextualización de los datos para su posterior comparación e interpretación, de manera que nos permita hacer valoraciones útiles e interesantes con el fin de mantener o redirigir la dirección de nuestro negocio o proyecto.
- Por supuesto, debe ser personalizado, de modo que tratará los objetivos específicos establecidos para cada empresa.

Además de estos cuatro elementos esenciales, resulta interesante señalar algunas características más que debemos tener en cuenta a la hora de elaborar un dashboard. Estas son:

- No debemos abusar en cuanto al número de indicadores que incluyamos, y es que deberemos incluir solo aquellos que sean realmente necesarios. Ello es muy importante ya que, como hemos dicho, se trata de una herramienta de visualización que nos ayudará a entender de manera rápida y fácil los

datos de que disponemos, por lo que deberemos evitar incluir indicadores que realmente no tengan importancia para lo que queremos analizar.

- Además, estos indicadores o KPI deben ser accionables para el negocio, es decir, deben permitirnos llevar a cabo acciones estratégicas para mejorar.
- Es muy importante que los datos e informaciones sean compartidos en la totalidad de la empresa, por lo que su visualización debe ser lo más clara y fácil de interpretar posible. Así, un dashboard debe ser breve, hablar un lenguaje común y ofrecer una representación gráfica adecuada para los datos que representa y lo suficientemente visual para que su estudio resulte atractivo.
- Además de los indicadores perfectamente visualizados, el dashboard debe ir acompañado por un análisis en el que explique qué ha ocurrido, las recomendaciones para poner solución o mejorar la situación y el impacto que ello puede tener sobre el negocio.

Como hemos comentado, las empresas disponen de toda una serie de herramientas para crear sus propios dashboards. Básicamente, de lo que se trata es de introducir los datos objeto de estudio, que serán convertidos en información gráfica con el fin de facilitar su interpretación.

A continuación, adjuntamos un ejemplo sobre cómo crear un dashboard con la herramienta Pentaho.

El ejemplo ha sido extraído de la web de Jortilles a partir del siguiente enlace:

<http://www.jortilles.com/crea-cuadro-mando-pentaho-cde/>

1. Introducción

Un Dashboard o cuadro de mando, es una aplicación que se utiliza para presentar el contenido de una serie de indicadores que muestran el comportamiento de los datos.

En este tutorial realizaremos un Dashboard con Pentaho, utilizando la herramienta que nos da Pentaho para hacer cuadros de mando, es decir CDE.

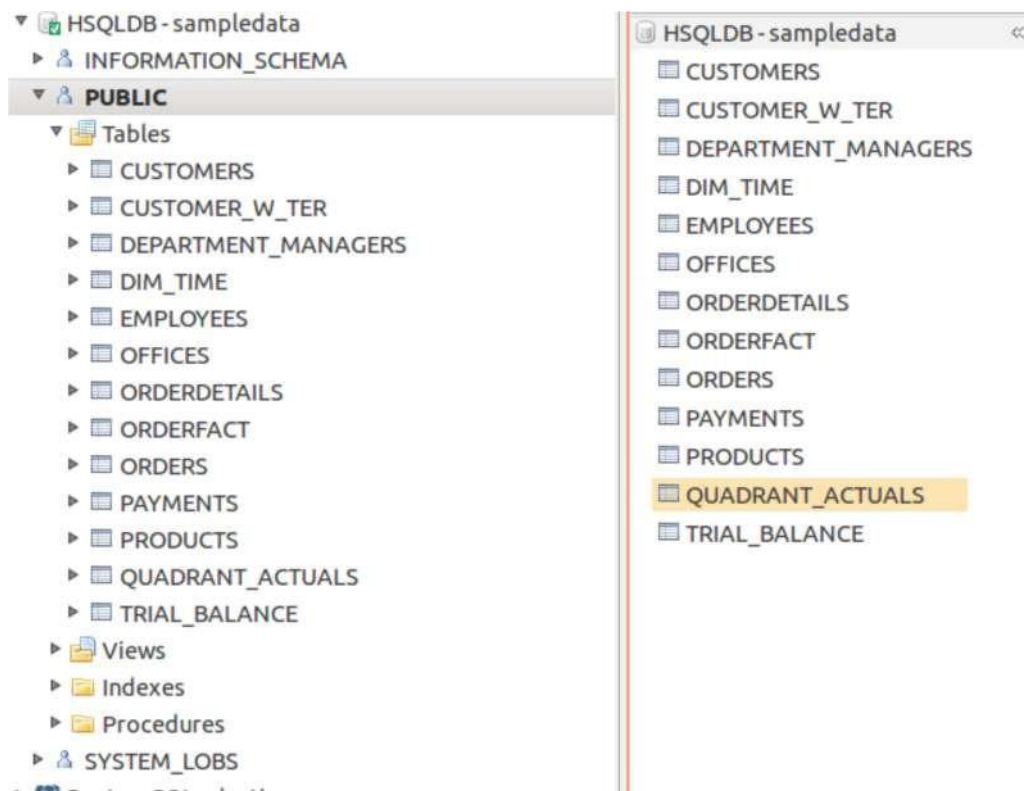
<http://www.webdetails.pt/ctools/cde/>

2. Requerimientos

Para realizar este Dashboard necesitamos instalar Pentaho Bi Server , disponible para descargar desde la web de SourceForge

<https://sourceforge.net/projects/pentaho/files/Business%20Intelligence%20Server/6.0/>

La base de datos que utilizaremos será la que viene de prueba, “SampleData”, al configurar Pentaho BI Server.



Estas son las tablas que tenemos disponibles para realizar nuestro Dashboard.

En nuestro Dashboard, la información que queremos presentar contendrá:

- Un selector con los nombres de los managers.
- Un gráfico de barras que mostrará la varianza acumulada por cada departamento.

3. Caso de uso: Creación de un Dashboard

Una vez descargado el programa, descomprimos, y ejecutamos en el terminal

```
./start-pentaho.sh
```

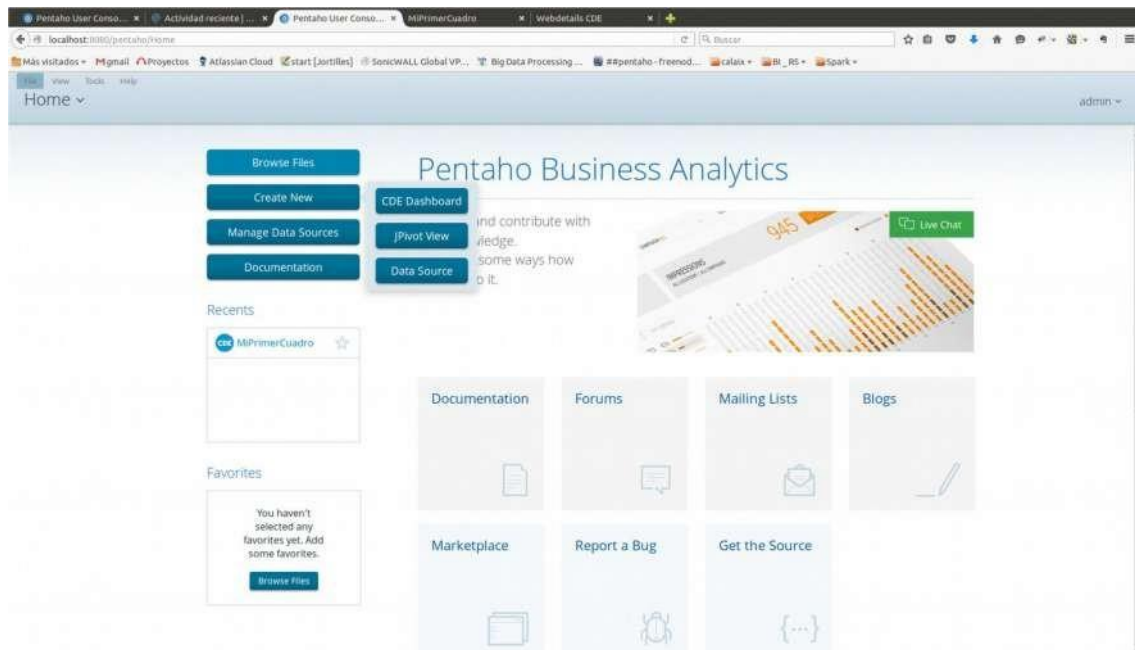


Abrimos el navegador con la ruta localhost:8080, y accedemos con el usuario admin



Seleccionamos Create New → CDE Dashboard

EL PROCESO DE ETL

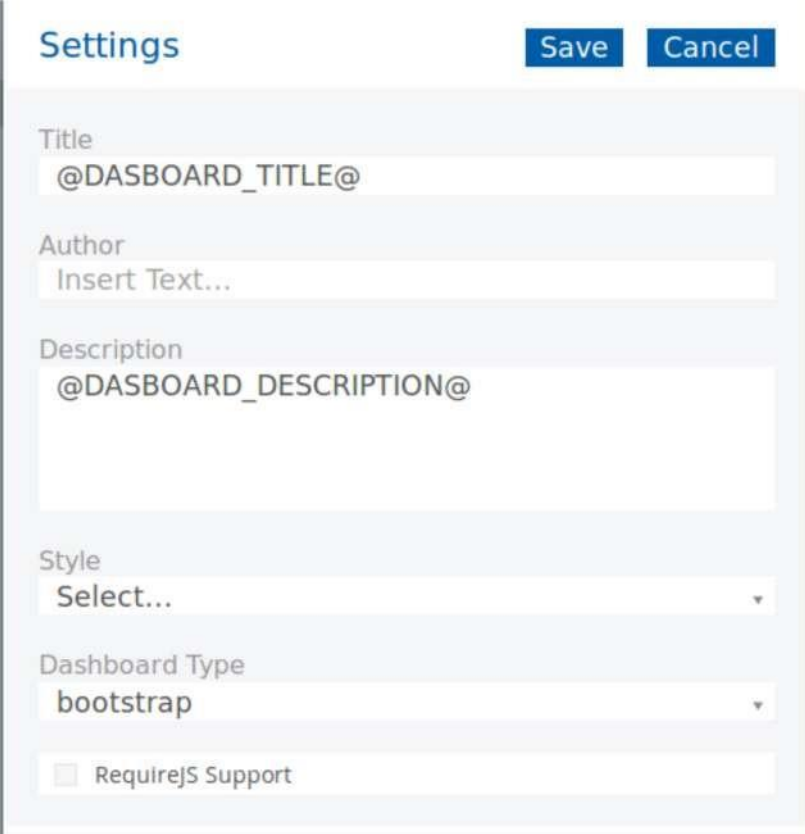


En el Dashboard tenemos tres vistas



- Layout Panel, es decir la presentación gráfica de los datos
- Components Panel, los objetos que mostrarán los datos, como gráficos, tablas, etc.
- DataSource Panel, son las consultas a los orígenes de datos

Primero guardaremos el documento con el nombre deseado, en este caso elegiremos como base el framework de Bootstrap:



The image shows a 'Settings' dialog box with a light blue header and two buttons: 'Save' and 'Cancel'. The settings are organized into sections with light blue backgrounds. The 'Title' section has a text input field containing '@DASHBOARD_TITLE@'. The 'Author' section has a text input field containing 'Insert Text...'. The 'Description' section has a text area containing '@DASHBOARD_DESCRIPTION@'. The 'Style' section has a dropdown menu with 'Select...' and a downward arrow. The 'Dashboard Type' section has a dropdown menu with 'bootstrap' and a downward arrow. The 'RequireJS Support' section has a checkbox that is currently unchecked.

Settings

Save Cancel

Title
@DASHBOARD_TITLE@

Author
Insert Text...

Description
@DASHBOARD_DESCRIPTION@

Style
Select...

Dashboard Type
bootstrap

☐ RequireJS Support

Lo guardamos en la carpeta pública, con el nombre elegido y continuamos:

Save as

Ok

Cancel

Choose Format

☒ Dashboard

☐ Widget

Choose Folder

public

Steel Wheels

etc

home

File Name*

MiPrimerCuadro

Title

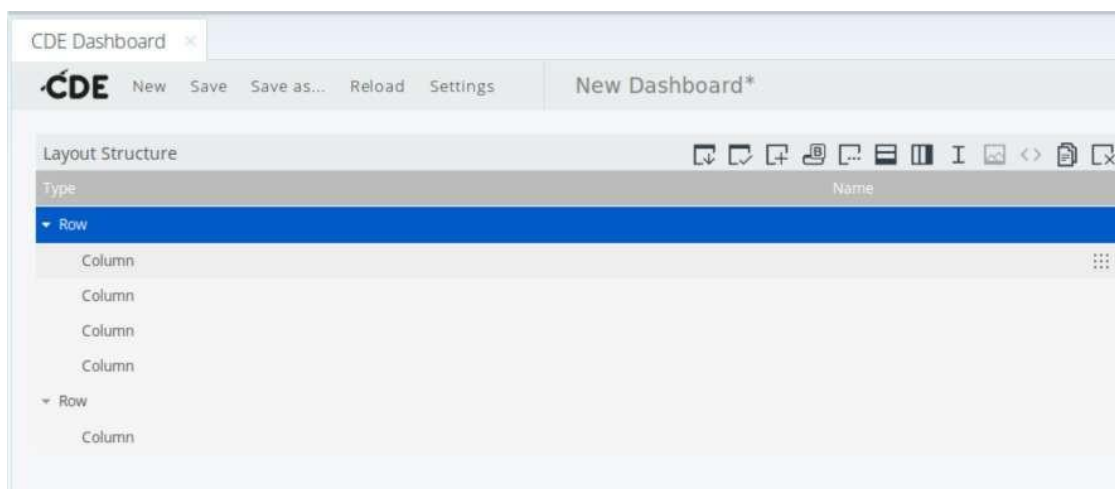
Insert Text...

Description

Insert Text...

➤ Layout Panel

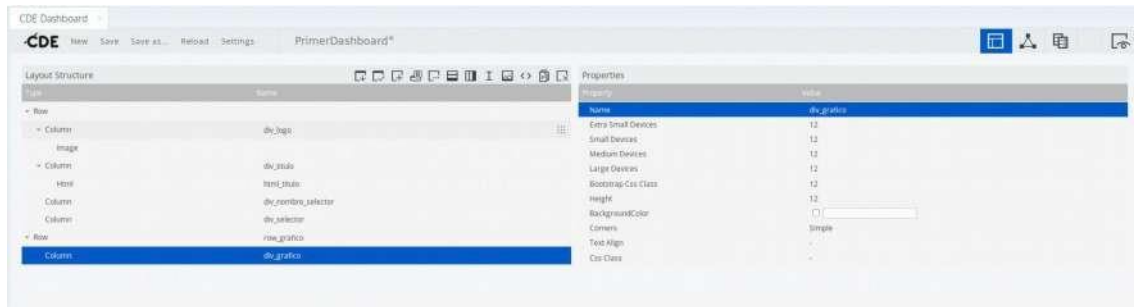
Empezamos a crear la estructura física donde irán los objetos. En esta capa podemos añadir los estilos css, código javascript y html



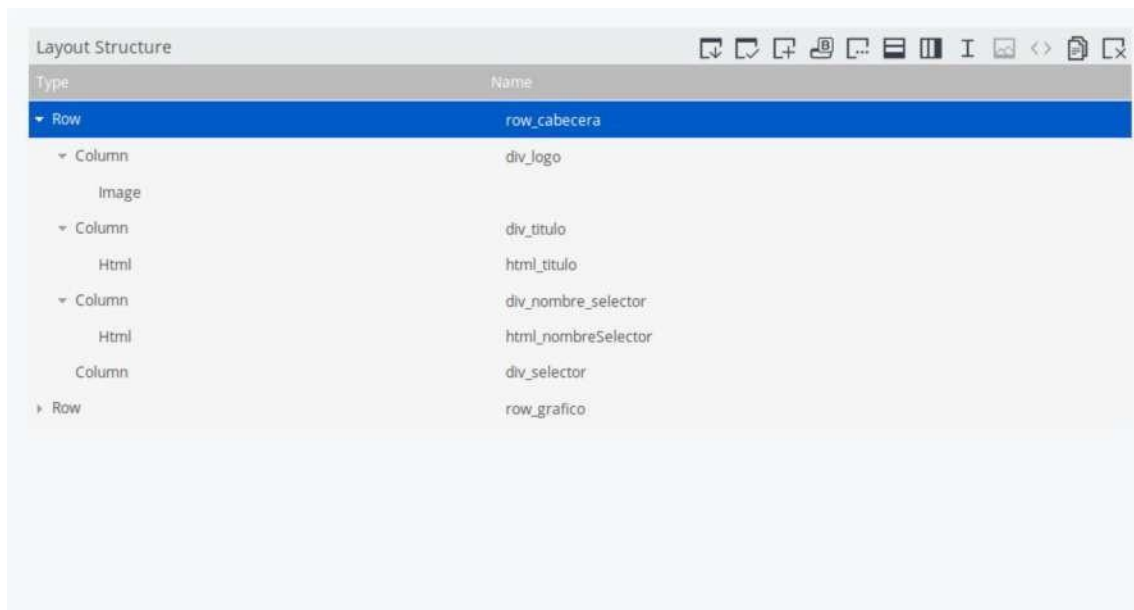
EL PROCESO DE ETL

Como estamos trabajando con Bootstrap, se basa en modelo de cajas (filas y columnas) la suma total de las columnas tiene que ser 12.

Definimos las filas que tendrá nuestro diseño, y dentro de cada fila las columnas que necesitaremos.



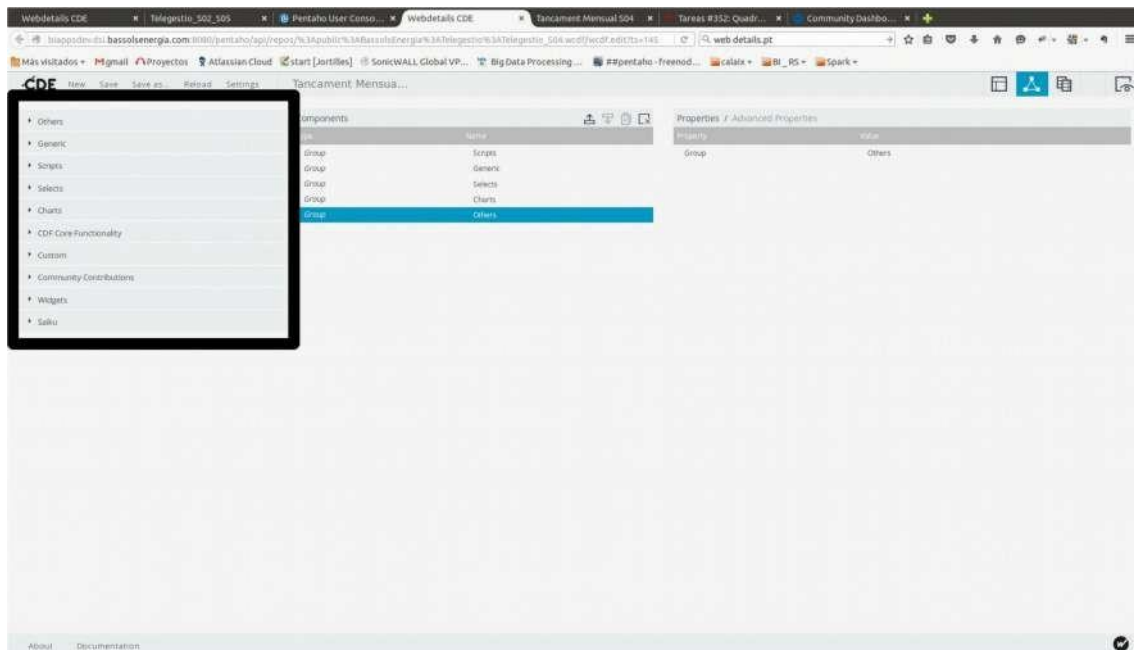
La Primera fila corresponde a la cabecera (con todos sus elementos) y la segunda corresponde al gráfico.



➤ Component Panel

Aquí tenemos diferentes objetos para utilizar, organizados por categoría (gráficos, selectores.):

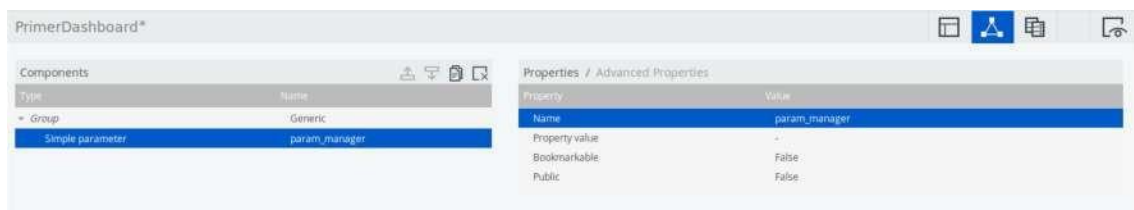
EL PROCESO DE ETL



Una vez definida la estructura de la presentación, vamos a crear los objetos necesarios. Necesitamos:

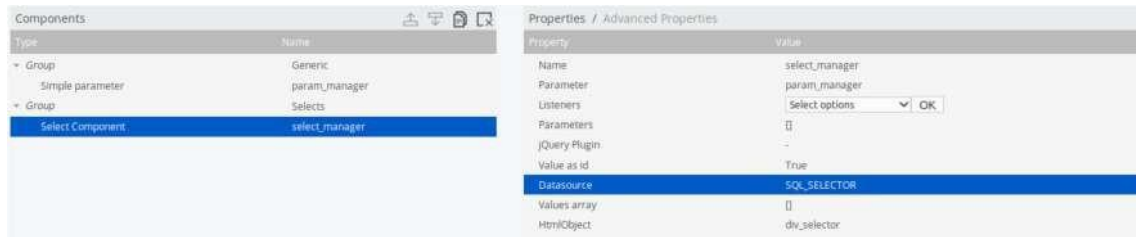
- Un selector, es decir un desplegable que muestre el nombre de los managers para hacer la consulta.
- Gráfico de barras para mostrar el resultado.

Para crear el selector, primero debemos crear una variable (simple parametre) donde guarde el valor asignado:



Después creamos el selector, (select component) es decir el desplegable, que devolverá el valor del parámetro que hemos creado anteriormente:

EL PROCESO DE ETL

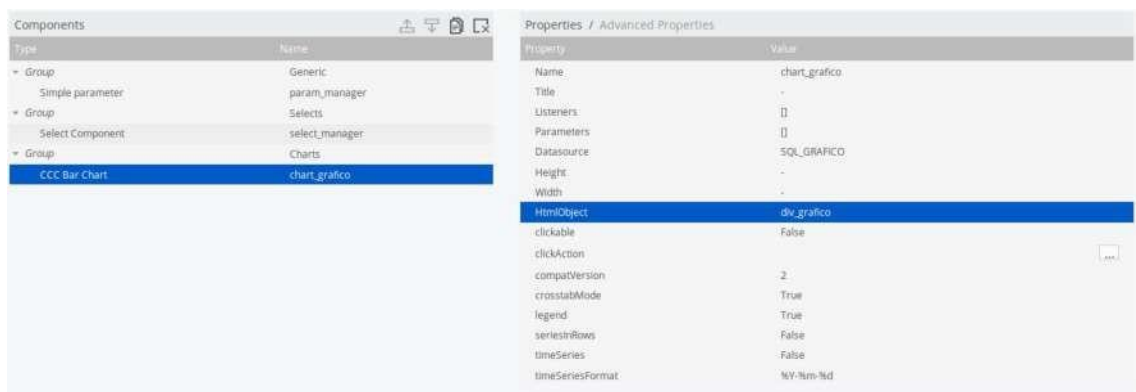


DataSource: Ponemos este valor, que será el nombre de la consulta que crearemos a continuación: SQL_SELECTOR

HtmlObject: Es el nombre del objeto que hemos declarado en el layout, donde irá, en este caso, el selector DIV_SELECTOR

Y el último objeto que necesitamos es el gráfico de barras:

DataSource:



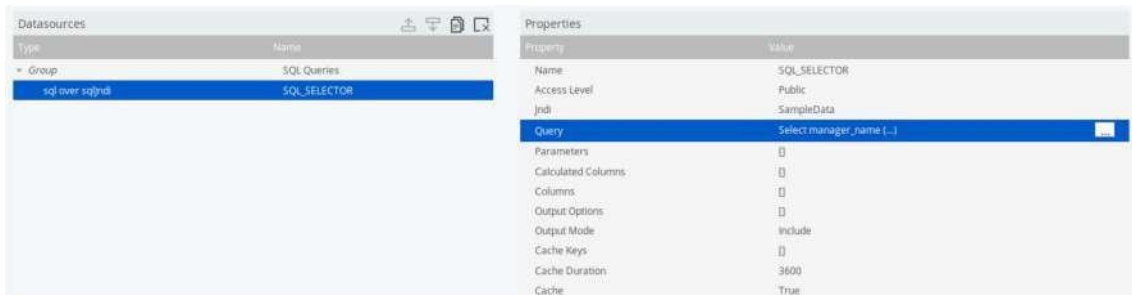
DataSource: Ponemos este valor, que será el nombre de la consulta que crearemos a continuación: SQL_GRAFICO

HtmlObject: Es el nombre del objeto que hemos declarado en el layout, donde irá, en este caso, el selector DIV_GRAFICO

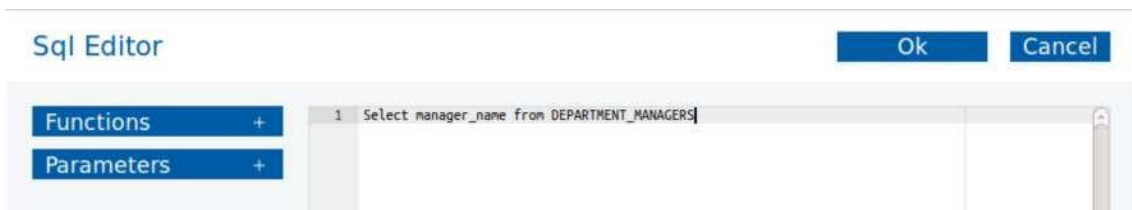
➤ Datasource Panel

- Para realizar la consulta SQL_SELECTOR. Seleccionaremos “sql over sqlIndi”

EL PROCESO DE ETL



Esta consulta devolverá el nombre de los managers.



- Para realizar la consulta SQL_GRAFICO



Si vamos al navegador y actualizamos, ya podemos ver el resultado, con el selector y el gráfico. Aunque aún no se actualizan los datos al cambiar el valor del selector.



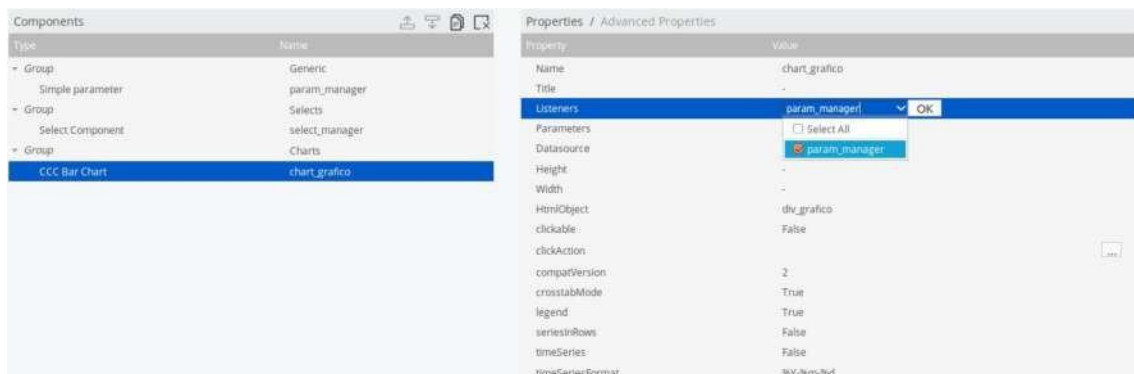
Para actualizar los datos al cambiar la selección, debemos activar los “listeners”, y los parámetros.

EL PROCESO DE ETL

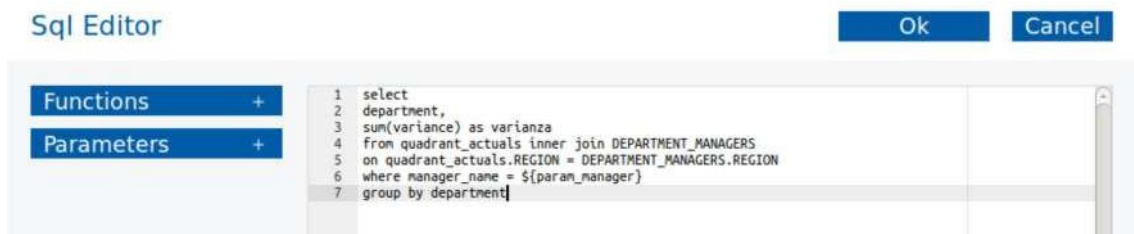
Un listener está relacionado con un parámetro, cuando se produce un cambio de este se desencadena una actualización del componente, en este caso, cuando se produce un cambio de manager, se actualiza el gráfico.

Parameters, es el valor o los valores que se pasan como argumentos para ejecutar la acción. En este caso solo pasamos un parámetro, el manager.

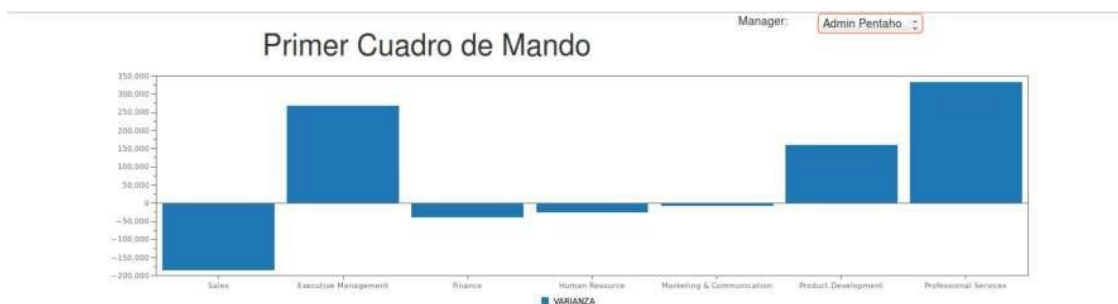
Volvemos a la pestaña de Componentes, y en el gráfico, ponemos en ambos campos “param_manager”



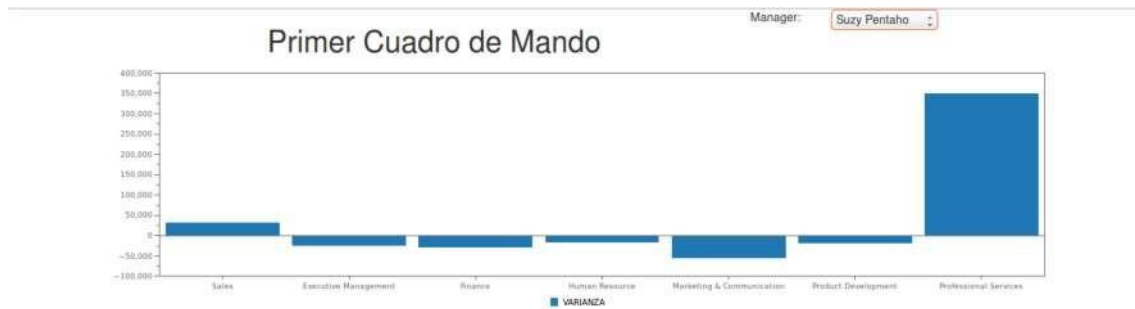
Ahora debemos actualizar el “datasource” SQL_GRAFICO, con el valor del parámetro:



Ahora sí que cambia la tabla en función del nombre del manager:



EL PROCESO DE ETL



Ahora vamos a añadir un panel de bootstrap, para mejorar la presentación del gráfico de barras. Para añadirlo simplemente clicamos en el icono.

Layout Structure			
Type	Name		
▼ Row	row_cabecera		
Column	div_logo		
▶ Column	div_titulo		
▶ Column	div_nombre_selector		
Column	div_selector		
▼ Row	row_grafico		
▼ Column	div_grafico		
▼ Bootstrap Panel			
▼ Panel Header			
Html	html_titulo_grafico		
Panel Body	div_contenedor		

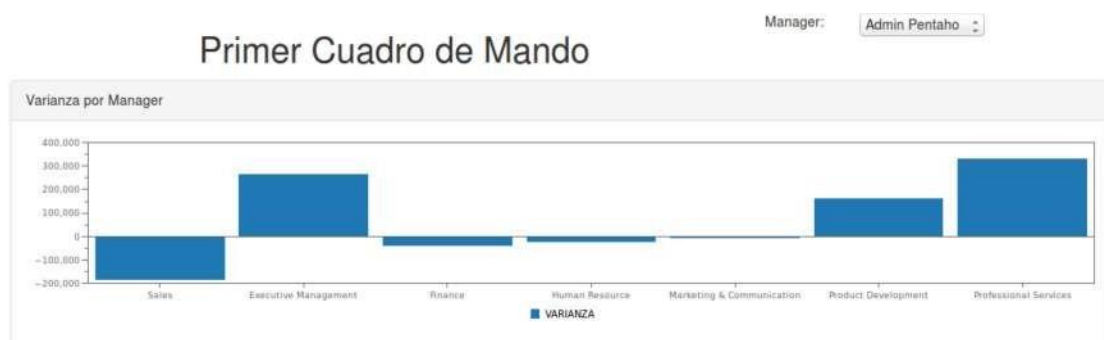
En el Panel Header, declararemos el título del gráfico.

Y el en panel body irá el gráfico de barras. Por ello debemos actualizar el nombre del objeto html que estaba asignando al gráfico.

EL PROCESO DE ETL

Properties / Advanced Properties	
Property	Value
Name	chart_grafico
Title	-
Listeners	[param_manager]
Parameters	[[{"param_manager": "p (...)
Datasource	SQL_GRAFICO
Height	200
Width	-
HtmlObject	div_contenedor
clickable	False
clickAction	...
compatVersion	2
crosstabMode	True
legend	True
seriesInRows	False
timeSeries	False
timeSeriesFormat	%Y-%m-%d

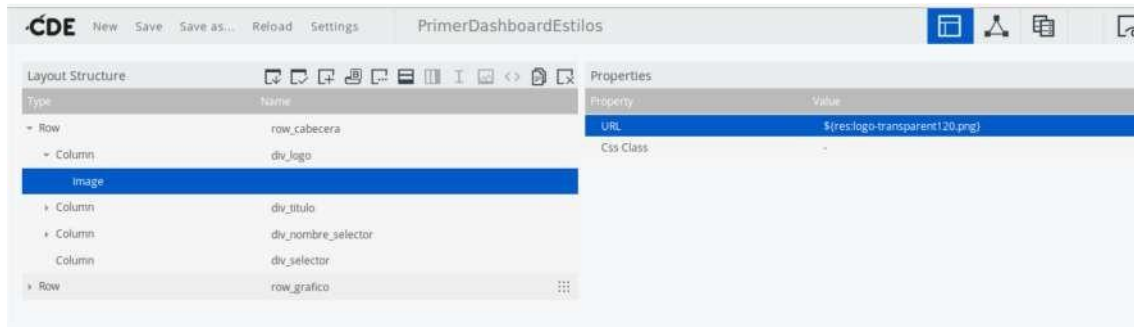
Si volvemos a actualizar el navegador ya podemos ver el panel con el título en el gráfico.



Por último, vamos a añadir nuestro logo en el Dashboard:

Ya teníamos definido el espacio en la capa del layout, en la fila correspondiente a la cabecera, insertamos una imagen.

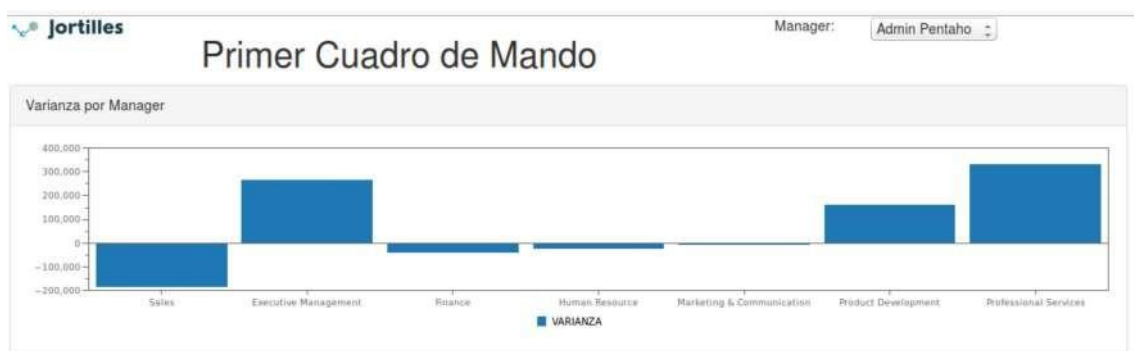
EL PROCESO DE ETL



Y en las propiedades, debemos introducir la url donde hemos puesto la imagen. En nuestro caso la hemos subido en la misma carpeta del dashboard, la carpeta pública.

Como url será: `$(res:nombre_logo.png)`

Volvemos a actualizar el navegador, y ahora sí, ya tenemos nuestro primer Dashboard finalizado:



Además de este ejemplo, a continuación ofrecemos dos ejemplos gráficos de dashboards creados a partir de la herramienta de Google, Google Analytics y con Excel.

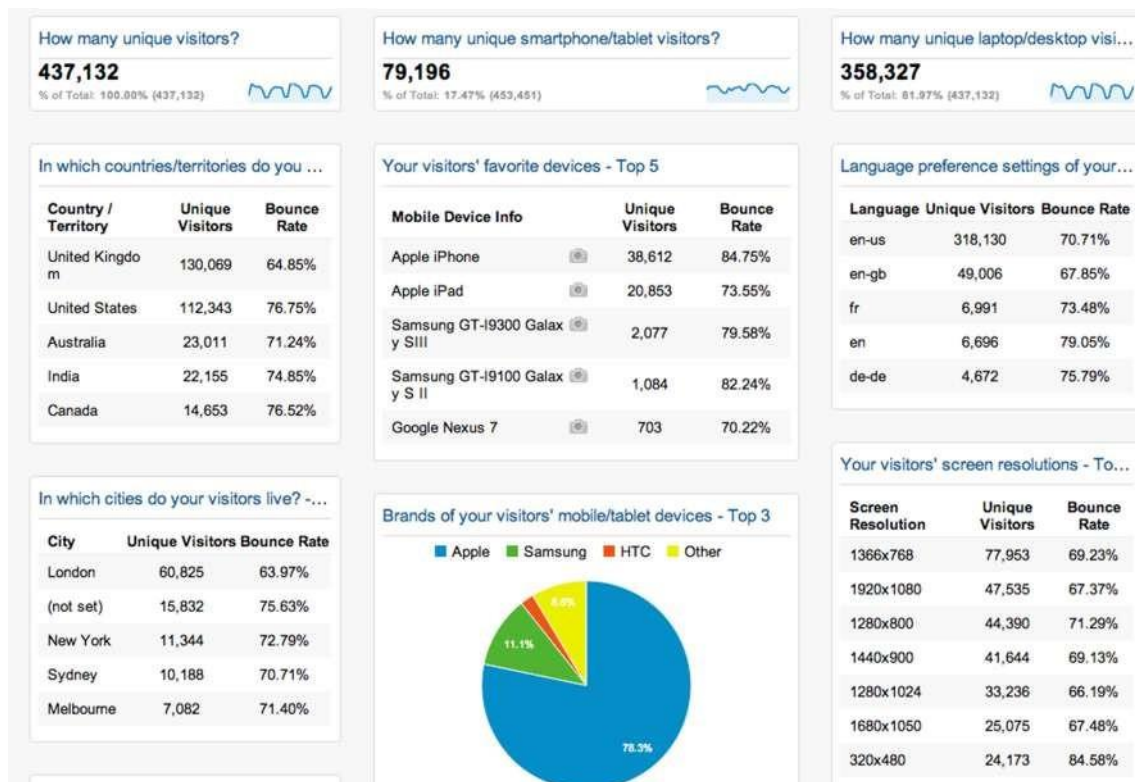
Además, y para complementar la información, en el apartado de los anexos hemos dispuesto algunos enlaces a tutoriales en los que se detalla de manera precisa cómo podemos crear un dashboard utilizando Google Analytics y Excel.

EL PROCESO DE ETL



*Ejemplo de dashboard extraído del blog *temporaExcel*. Enlace:

<http://temporaexcel.blogspot.com.es/2012/11/como-hacer-dashboard-cuadro-de-mando-economico-financiero-2.html>



*Ejemplo de dashboard extraído del blog *Ecoconsultancy*. Enlace: <https://ecoconsultancy.com/blog/62828-10-useful-google-analytics-custom-dashboards>

Comentar que, además de estas herramientas, existen otras herramientas para la creación de un dashboard. Para su consulta, adjuntamos el siguiente enlace, en el que se ofrece un listado de dichas herramientas.

Enlace: <https://www.capterra.com/data-visualization-software/>

Para acabar, como hemos comentado antes, es importante que el dashboard vaya acompañado de un documento en el que se reflejen los resultados e impresiones del análisis.

Aquí resulta interesante hablar sobre los **informes personalizados**.

Tal y como su propio nombre indica, se trata de informes personalizados, es decir, son informes creados en función de las necesidades del momento, en los que obtendremos información y datos sobre el tema que hayamos escogido, de modo que nos ofrecerá solamente aquella información concreta que nosotros o nuestra empresa quiera.

De este modo, una vez disponemos de toda la información referente a nuestra empresa, con los informes personalizados lo que haremos será acotar dicha información a los datos que realmente nos interesan.

Por ejemplo, podemos tener información relacionada con el departamento de ventas de un trimestre o año en concreto. Aquí, disponemos de datos tales como la facturación, el importe del ticket medio, el número de compras diarias, semanales y mensuales, los productos más y menos vendidos, aquellos productos que se han vendido juntos, la procedencia de los compradores, el modo de pago, etc. Como vemos, se trata de una gran cantidad de información, toda ella útil dependiendo de lo que en ese preciso momento nos interese. Así, si lo que queremos es hacer promociones por la compra de dos productos juntos, lo que nos interesará será conocer toda la información relativa a aquellos productos que se han vendido juntos durante el período de tiempo analizado o aquellos productos que los compradores han consultado mientras compraban un producto concreto, por lo que crearemos un

EL PROCESO DE ETL

informe personalizado con dicha información, que nos permita valorar si la acción que queremos llevar a cabo reportará o no beneficios.

Como beneficios que nos ofrecen este tipo de informes destacamos:

- Nos ofrecen de manera unificada información que tenemos repartida en varios informes, presentando dicha información de manera unificada en función de las métricas y parámetros que hayamos seleccionado.
- Nos permite comunicar y reportar esta información de manera mucho más rápida, clara y precisa, tanto a los miembros de la empresa como a los clientes.
- Además de seleccionar el tipo de información que queremos incluir, también nos permiten establecer parámetros en relación al tiempo.

Para la realización de estos informes podemos servirnos de herramientas como las que ofrece Google Analytics.