

E.T.S. de Ingeniería Industrial,
Informática y de Telecomunicación

Sistema de Business Intelligence para la toma de decisiones de Cruz Roja en la gestión de refugiados



Grado en Ingeniería Informática

Trabajo Fin de Grado

Autor: Euken Rodrigo Marín

Tutor: José Javier Astrain Escola

Pamplona, 27/05/2016

Agradecimientos

En primer lugar, quisiera expresar mi agradecimiento a la Universidad Pública de Navarra y en especial a la Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación, y al Departamento de Ingeniería Matemática e Informática, donde he recibido durante estos años, no sólo una formación académica, sino también una formación humana, como persona.

A mi tutor de Trabajo Fin de Grado, José Javier Astrain Escola. Gracias por atender mis inquietudes la hora de realizar un TFG real como el que he podido realizar. También gracias por su tiempo, consejos, correcciones y ánimos que me han permitido llevar a cabo satisfactoriamente este trabajo.

También agradecer a Raúl Ecay Torres la oportunidad de poder realizar este TFG, donde tanto he podido aprender y comenzar a ver lo que realmente es la exigencia del mundo laboral. Gracias también por toda la ayuda recibida durante estos meses, que sin duda, me ha ayudado a completar el trabajo correctamente.

A Cruz Roja, por la gran labor social y humanitaria que llevan a cabo por todo el mundo. También por la oportunidad de permitirme realizar el trabajo, espero y deseo que sea muy útil y que ayude en todo lo posible en esta grave situación.

A mi familia, en especial a mis padres y hermana, siendo siempre un apoyo para mí.

A mis compañeros y amigos de promoción, por acompañarme durante estos años.

Resumen

Diseño y desarrollo de un sistema de inteligencia de negocio (*Business Intelligence, BI*) para la organización y planificación de recursos en un entorno de incertidumbre como es el de la llegada de refugiados a una organización de ayuda e intervención, en este caso Cruz Roja [1]. La herramienta BI se incluye en un Sistema de Información basado en la integración de tecnologías sencillas, robustas y de mínimo coste que garanticen el correcto funcionamiento del sistema.

Se trata de analizar la gestión por procesos de la atención a los refugiados, con información integrada y en tiempo real, para su dirección, coordinación y posible reorganización. Se pretende compilar de forma ágil y convenientemente anonimizada la información disponible (muchas veces incompleta) y suministrar las herramientas informáticas necesarias para generar informes automatizados que den soporte a la toma de decisiones.

Se incluye también la planificación y organización de los sistemas y tecnologías de la información a emplear, revisando su impacto organizativo y de información, que permitan la adecuada implantación y explotación del sistema BI.

En concreto, Cruz Roja realiza actualmente un control basado en la recogida de formularios registrados en una Base de Datos. Para ello se ha desarrollado un sistema BI que extrae la información de la Base de Datos y permite mostrar unos indicadores claros e intuitivos, necesarios para una correcta toma de decisiones en la gestión de provisiones para los refugiados.

Abstract

Design and development of a Business Intelligence (BI) tool for the organization and planning of resources in an environment of uncertainty as it is the arrival of refugees to an organization of help and intervention.

Based on a Red Cross request, this project consists on the design and development of a Business intelligence tool devoted to improve analysis and decision making process of humanitarian crisis.

Prior the request the Red Cross implemented a form management and data collection tool (Open Data Kit) which feeds data to this system. The BI tool developed takes part of an information system, which is based on the integration of simple, robust and minimum cost technologies.

The tool is devoted to coordinate, evaluate, and may be re-organize, humanitarian operations according to the analysis of the anonymous information concerning the refugees assisted at a set of health posts. The tool captures in real-time the healthcare information, by way of the forms accomplished by doctors, compiles the available (and sometimes incomplete) information, and provides IT tools that support the decision making process and generate automatic reports.

The information collected from the health forms is stored on an existing database of the Red Cross. The developed BI system extracts from this database the required information, processes it, and provides clear and intuitive indicators (dashboards), useful for a right decision-making on the management of provisions for the refugees.

The resulting tool is currently used by Red Cross on the European migration crisis.

Lista de palabras clave

- **Inteligencia de Negocios o BI** (del inglés *business intelligence*), al conjunto de estrategias y aspectos relevantes enfocados a la administración y creación de conocimiento sobre el medio, a través del análisis de los datos existentes en una organización o empresa. El BI se ha convertido en la actualidad en una de las técnicas de análisis de datos con la que las empresas pueden generar conocimiento y tomar decisiones estratégicas.
- **Proceso ETL:** (*Extract, Transform and Load*) “Extraer, transformar y cargar” es el proceso que permite mover datos desde múltiples fuentes, reformatearlos, limpiarlos, y cargarlos en otra base de datos, o *datawarehouse* para analizar, o en otro sistema operacional para apoyar un proceso de negocio.
- **Cubo OLAP:** Son las siglas en inglés de *OnLine Analytical Processing*, procesamiento analítico en línea. Conceptualmente es una base de datos multidimensional en la cual el almacenamiento físico de los datos se realiza también en un vector multidimensional. Es decir, tendremos una tabla de hechos, en donde aparecerá la información sobre la que queramos medir en cualquiera de sus dimensiones y enlazadas a esta, por medio de clave foránea aparecerán las dimensiones de medida del cubo, que pueden ser tantas como queramos.
- **Datawarehouse (DW):** Es un repositorio estructurado de múltiples sistemas y fuentes de datos, a nivel de empresa, orientados hacia áreas de negocio, que contiene datos históricos y que está preparado para facilitar la toma de decisiones.
- **Cuadro de Mando Integral (CMI) o Dashboard:** Herramienta de **gestión empresarial** muy útil para medir la evolución de la actividad de una organización o empresa y sus resultados, desde un punto de vista estratégico y con una perspectiva general. Suelen ser empleados por Gerentes y altos cargos ya que contribuyen de forma eficaz en la visión empresarial, a medio y largo plazo.
- **Open Source:** Característica del software que hace referencia a las libertades de las que disponen los usuarios del mismo para usarlo, distribuirlo, modificarlo y distribuir versiones modificadas del producto, para lo cual exige proporciona acceso al código fuente del mismo. El software Open Source suele estar construido bajo un entorno colaborativo de programadores que hacen sus aportaciones que comparten con la comunidad.
- **Big Data:** Concepto que hace referencia al almacenamiento de grandes cantidades de datos y a los procedimientos usados para encontrar patrones repetitivos dentro de esos datos.

Índice de Contenidos

| | |
|--|----|
| 1. INTRODUCCIÓN | 11 |
| 1.1. Contexto y justificación del trabajo | 11 |
| 1.2. Objetivos del trabajo | 12 |
| 1.3. Estudio del arte | 12 |
| 1.4. Propuesta | 14 |
| 1.5. Enfoque y metodologías | 15 |
| 1.6. Tecnologías | 16 |
| 1.7. Contenidos de la memoria | 17 |
| 2. ITERACIÓN 1-INSTALACIÓN Y CONFIGURACIÓN DE PENTAHO 6.0 | 19 |
| 2.1. Instalación de JAVA 8 | 19 |
| 2.2. Configuración de las variables de entorno | 21 |
| 2.3. Instalación de Pentaho 6.0 | 22 |
| 2.4. Instalación librería Java | 24 |
| 2.5. Instalación de plugins necesarios | 24 |
| 3. ITERACIÓN 2-PROCESO ETL | 26 |
| 3.1. Extracción | 27 |
| 3.1.1 Entrada de tabla | 28 |
| 3.2. Transformación | 30 |
| 3.3. Carga | 31 |
| 3.3.1. Salida Tabla | 32 |
| 3.4. Job | 33 |
| 4. ITERACIÓN 3- CREACIÓN DEL CUBO OLAP | 35 |
| 4.1. Conexión al Datawarehouse | 35 |
| 4.2. Nuevo esquema y añadir cubo | 36 |
| 4.3. Crear dimensiones | 37 |
| 4.4. Dimension Usage | 39 |
| 4.5. Añadir medidas | 40 |
| 4.6. Añadir Data Source y Publicar esquema | 41 |
| 5. ITERACIÓN 4-CREACIÓN DE DASHBOARD | 44 |
| 5.1. Construcción del Dashboard | 46 |
| 5.1.1. Creación de Dashboard | 46 |
| 5.1.2. Visualización del Dashboard | 51 |

| | |
|--|----|
| 6. DASHBOARDS | 52 |
| 6.1. Greece Control Panel | 52 |
| 6.2. Greece Morbidity..... | 53 |
| 6.3. Health Report..... | 54 |
| 6.4. Medicine Consumption | 56 |
| 7. CONCLUSIONES | 58 |
| 8. REFERENCIAS | 59 |
| 9. ANEXO | 61 |
| Anexo 1- Carta de Cruz Roja solicitando la realización del proyecto..... | 61 |

Listado de Ilustraciones

| | |
|---|----|
| Ilustración 1: Muestra del tablero Trello con cada iteración | 15 |
| Ilustración 2 : Componentes del Sistema BI | 18 |
| Ilustración 3: 1-Instalación Java 8 | 19 |
| Ilustración 4: 2-Instalación Java 8 | 19 |
| Ilustración 5: 3-Instalación Java 8 | 20 |
| Ilustración 6: 4-Instalación Java 8 | 20 |
| Ilustración 7: Comando para modificar fichero de variables de entorno | 21 |
| Ilustración 8: Modificar variables de entorno | 21 |
| Ilustración 9: Comprobación de variables | 22 |
| Ilustración 10: Descarga de Pentaho BIServer 6.0 | 22 |
| Ilustración 11: Unzip Pentaho BIServer | 23 |
| Ilustración 12: Iniciar Pentaho | 23 |
| Ilustración 13: Pantalla de inicio Pentaho | 24 |
| Ilustración 14: Marketplace | 25 |
| Ilustración 15: Saiku Analytics | 25 |
| Ilustración 16: Saiku Analytics 3.8.8 Pentaho 6 | 25 |
| Ilustración 17: Iniciar Spoon | 26 |
| Ilustración 18: Pantalla inicial Spoon | 27 |
| Ilustración 19: Fichero, Nuevo, Transformación | 27 |
| Ilustración 20: Conexión datos entrada | 28 |
| Ilustración 21: Entrada tabla | 28 |
| Ilustración 22: Entrada tabla 2 | 29 |
| Ilustración 23: Entrada tabla 3 | 29 |
| Ilustración 24: Transformación Renombrar | 30 |
| Ilustración 25: Transformación Replace | 31 |
| Ilustración 26: Transformación Range | 31 |
| Ilustración 27: Conexión datos salida | 32 |
| Ilustración 28: Salida tabla | 32 |
| Ilustración 29: Ejecutar transformación | 33 |
| Ilustración 30: Vista general JOB | 34 |
| Ilustración 31: Programación JOB | 34 |
| Ilustración 32: Iniciar Schema Workbench | 35 |
| Ilustración 33: Conexión al DW | 35 |
| Ilustración 34: File, New, Schema | 36 |
| Ilustración 35: Añadir cubo | 36 |
| Ilustración 36: Añadir tabla al cubo | 36 |
| Ilustración 37: Se debe añadir dimensiones | 37 |
| Ilustración 38: Añadir dimensión | 37 |
| Ilustración 39: Añadir nombre dimensión, nivel y clave primaria | 38 |
| Ilustración 40: Añadir jerarquías | 38 |
| Ilustración 41: Dimensiones tiempo y localización | 39 |
| Ilustración 42: Añadir dimension usage | 39 |
| Ilustración 43: Dimension Usage | 40 |

| | |
|--|----|
| Ilustración 44: Añadir medidas | 40 |
| Ilustración 45: Medidas | 40 |
| Ilustración 46: Vista general del cubo | 41 |
| Ilustración 47: Manage Data Sources | 41 |
| Ilustración 48: Add Connection | 42 |
| Ilustración 49: Credenciales DW | 42 |
| Ilustración 50: Test data Source | 42 |
| Ilustración 51: Publicar esquema | 43 |
| Ilustración 52: Abrir Saiku Analytics | 44 |
| Ilustración 53: Seleccionar cubo | 45 |
| Ilustración 54: Salida de información | 45 |
| Ilustración 55: Abrir CDE Dashboard | 46 |
| Ilustración 56: Estructura Layout | 47 |
| Ilustración 57: Query MDX | 47 |
| Ilustración 58: Datasource, Query, JNDI, Mondrian Schema | 48 |
| Ilustración 59: Query, Parameters | 48 |
| Ilustración 60: Query | 49 |
| Ilustración 61: Parameters | 49 |
| Ilustración 62: Generic Components | 50 |
| Ilustración 63: Select Components | 50 |
| Ilustración 64: Charts Components | 50 |
| Ilustración 65: Listeners y Parameters | 51 |
| Ilustración 66: Visualización Dashboard | 51 |
| Ilustración 67: Greece Control Panel Dashboard | 52 |
| Ilustración 68: Greece Morbidity Dashboard | 53 |
| Ilustración 69: Health Report Dashboard | 55 |
| Ilustración 70: Medicine Consumption Dashboard | 57 |

1. INTRODUCCIÓN

1.1. Contexto y justificación del trabajo

La realización de este trabajo surge de la necesidad planteada por Cruz Roja en el ámbito sanitario de emergencia: disponer de un sistema de inteligencia de negocio (*Business Intelligence, BI*, conjunto de estrategias y aspectos relevantes enfocados a la administración y creación de conocimiento, a través del análisis de los datos existentes en una organización o empresa), que permita la organización y planificación de recursos en un entorno de incertidumbre como es el de llegada de refugiados a Grecia, véase Anexo 1.

Cruz Roja realiza actualmente un control basado en la recogida de formularios informatizados registrados en una Base de Datos. Este sistema de formularios cumple con las funcionalidades básicas para dar soporte al control de recursos, pero según pasa el tiempo y los datos van creciendo, se hace cada vez más difícil esta organización.

Los formularios son recogidos mediante ODK [2] (Open Data Kit), conjunto de herramientas que permiten recopilar datos a través de dispositivos móviles y enviarlos a un servidor online aunque no se disponga de conexión a Internet o de acceso a una red móvil en el momento de recopilar los datos. Esta variedad de formularios son utilizados por la facilidad de tomar datos a las personas que son atendidas en cada uno de los puestos que hay repartidos en el país griego.

Para facilitar la cumplimentación de estos formularios, la mayoría de campos a introducir son desplegados con información generalizada y estandarizada. Aunque todos siguen un formato estándar, cada formulario, según el tipo de información recogida, rellena filas de distintas tablas de la Base de Datos, y cada una de ellas tiene distintas particularidades, tal y como se indicará en la sección Iteración 2.

Cruz Roja tiene la necesidad de realizar análisis de los datos obtenidos a través de estos formularios, con el que se pretende dar respuesta a lo siguiente:

- Análisis de los pacientes/diagnósticos:
 - ¿Qué rangos de edades?
 - ¿Qué géneros?
 - ¿Dónde se dan?

- Análisis farmacológicos:
 - ¿Qué medicamentos se dispensan?
 - ¿Qué cantidades?

En vista de esta situación y con la idea de seguir utilizando la obtención de datos a través de los formularios debido a su facilidad de movilidad y utilización, el análisis de la información pasó por el planteamiento de crear un nuevo Datawarehouse (DW, almacén de datos estructurado de múltiples sistemas y fuentes de datos orientados hacia áreas de negocio, que contiene datos históricos y que está preparado para facilitar la toma de

decisiones.). Esto supone el diseño y desarrollo de un sistema de inteligencia de negocio (*Business Intelligence, BI*) que permite la organización y planificación de recursos utilizando tecnologías sencillas, robustas y de mínimo coste que garanticen el correcto funcionamiento del sistema. El sistema está compuesto por lo siguiente:

- Un proceso de extracción, transformación y carga (ETL), que permite mover datos desde múltiples fuentes, reformatearlos, limpiarlos, y cargarlos en otra base de datos o DW, para analizarlos, o en otro sistema operacional para apoyar un proceso de negocio.
- Un DW, creado sobre la Base de Datos de los formularios.
- Un conjunto de Dashboard y un cuadro de mando con la herramienta BI para explotar el DW y medir la evolución de la actividad de una organización o empresa y sus resultados, desde un punto de vista estratégico y con una perspectiva general para la toma de decisiones.

1.2. Objetivos del trabajo

La adquisición de datos se realiza en condiciones de campaña, que distan mucho de ser las adecuadas, por lo que esto añade complejidad al sistema y produce la aparición de determinados errores en el proceso y errores en la obtención de la información deseada. Ante los inconvenientes anteriormente indicados sobre el tipo de análisis que se hace de dicha información, o la carga de trabajo innecesaria que se le somete a la Base de Datos, se plantean los siguientes objetivos:

El objetivo principal de este trabajo es:

- Diseñar y desarrollar un sistema de inteligencia de negocio, que permita un análisis correcto de los datos recogidos de forma más sencilla e intuitiva.

Entre los objetivos específicos del trabajo están:

- Evaluar y seleccionar una solución de mínimo coste, primando las opciones Open Source.
- Buscar posibilidades existentes de integración con orígenes o destinos de datos.
- Obtener una serie de decisiones basadas en resultados obtenidos.

1.3. Estudio del arte

La Inteligencia de Negocio, o Business Intelligence (BI), comprende todo el conjunto de estrategias y técnicas relevantes, enfocadas a la creación y clasificación de conocimiento a través de datos e informaciones de una organización o empresa.

Este conocimiento se emplea para facilitar la toma de decisiones de forma correcta e intuitiva, apoyándose en la posible anticipación de acontecimientos futuros, para generar estrategias con el objetivo de mejorar beneficios, competitividad o simplemente prevenir la ausencia de recursos.

Mediante las herramientas de ETL, se extraen datos de distintas fuentes, se tratan y se preparan, para luego cargarlos en el DW.

El ciclo de vida de los sistemas BI dependerá del éxito de las organizaciones o empresas en la toma de sus decisiones. Si estas obtienen beneficios o son realmente más competitivas, el software del BI seguirá siendo utilizado por mucho tiempo, ahora que cada vez las empresas poseen más datos; en caso contrario, serán sustituidos por otros sistemas que aporten resultados más precisos.

En el mercado hay gran variedad de herramientas de Business Intelligence. A continuación se enumeran algunas de ellas:

Software de pago:

- Microsoft SQL Server BI [3].
 - Analysis Services (SSIS).
 - Integration Services (SSAS).
 - Reporting Services (SSRS).
 - Compatibilidad con productos de Microsoft.
- Oracle BI [4].
 - Herramientas de BI integral e informar para inspirar a mejores decisiones en toda la organización.
 - Cuadros de mando interactivos que facilitan la presentación de informes y visualización de datos simple.
- Qlik [5].
- SAP Business Intelligence Platform [6].
 - Plataforma única e integrada que reúne la presentación de informes para proporcionar una visión detallada de la organización.
 - Fácil visualización que hace entender los datos.
- Pentaho Enterprise Edition [7].
- Tableau Business Intelligence [8].
 - Tableau Server o en la nube Tableau Online.
 - En un solo clic, se reciben respuestas en segundos de informes.
- SAP Lumira [9].

Open Source (Software Libre):

- Pentaho Community Edition [10].
 - Desarrollada en Java
 - La más popular
- Palo [11].
- SpagoBI [12].

1.4. Propuesta

Para la realización del proyecto se ha optado por la herramienta **Pentaho Community**. Se trata de una herramienta **Open Source**, que cumple con el objetivo de mínimo coste, ya que es posible trabajar con ella sin ningún tipo de licencia de pago. Además de eso, ofrece una funcionalidad suficientemente potente como para realizar todo lo necesario para este TFG. El único inconveniente que se ha encontrado a la herramienta, ha sido con la Base de Datos que está incorporada por defecto. A la hora de realizar consultas sobre rangos de fechas se dan fallos de query.

Dentro de las herramientas de software libre disponibles en el mercado **Pentaho Community**, es la más extendida y utilizada, lo que aporta una ventaja a la hora de documentarse, aprender y solucionar problemas sobre la misma.

Pentaho es una herramienta de Business Intelligence desarrollada en **Java** [13], y ejecutada en un servidor **Tomcat** [14], lo cual la hace muy accesible ya que es posible utilizarla desde cualquier dispositivo y desde cualquier sistema operativo. Motivo realmente importante para Cruz Roja que dispone de diferentes personas que utilizan distintos sistemas operativos y dispositivos en diversos lugares del mundo. Esto permite que todos puedan acceder indistintamente a Pentaho y realizar las labores pertinentes de una manera adecuada y sin ningún tipo de problemas.

La herramienta BI que se va a montar para la realización del TFG, será **Pentaho 6.0**. Es un conjunto de programas que posee una arquitectura e infraestructura para construir aplicaciones de inteligencia de negocio. Posee un enfoque moderno, simplificado, e interactivo que permite, a partir de un conjunto de información, acceder, descubrir y mezclar cualquier tipo de datos, independientemente de su tamaño y de donde se extraiga, ya sea una base de datos, una hoja de cálculo o un fichero de texto.

Pentaho 6.0 posee una amplia gama de herramientas de análisis cada vez más avanzadas. A partir de informes básicos a modelos de predicción, los usuarios pueden ayudarse a sí mismos para analizar y visualizar los datos necesarios. Además, esta versión añade nuevas funcionalidades respecto a versiones anteriores que parecen interesantes, como pueden ser:

Servicios de datos: Capacidad para exponer cualquier punto de una transformación como una tabla virtual que se puede acceder desde el exterior a través de una interfaz JDBC, y este uso puede ser no sólo externo sino también interno. Esta característica

viene con enormes mejoras en el concepto, no sólo a nivel de interfaz de usuario, sino también sobre las capacidades de almacenamiento en caché globales.

Linaje de PDI: Las organizaciones no sólo necesitan saber qué ETL y transformaciones tienen, necesitan saber exactamente lo que está sucediendo en cualquier, ¿Qué archivos, qué parámetros, cuándo, por qué?, y así sucesivamente. Se puede utilizar PDI para capturar toda esa información en tiempo de ejecución.

1.5. Enfoque y metodologías

Para la realización de este proyecto se ha empleado una metodología ágil interactiva afrontando y diferenciando en distintas iteraciones el proyecto de manera independiente a la vez que se avanzaba de una manera secuencial. Antes de comenzar cada iteración, mediante la herramienta **Trello** [15], se realizaba un primer sprint con las tareas a abordar para esa parte. Según se iba avanzando, estas podían ser modificadas e incluso añadir nuevas tareas que se necesitaran.

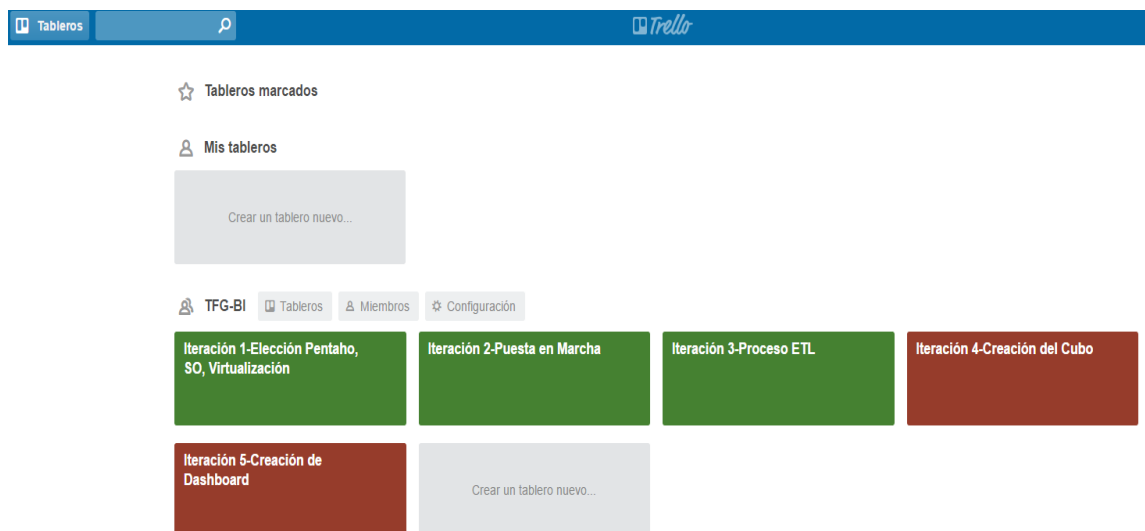


Ilustración 1: Muestra del tablero Trello con cada iteración

La principal manera de aprendizaje ha sido documentarse todo lo posible sobre las distintas partes del proyecto, y una vez se estaba lo suficientemente informado se seguía una estrategia de prueba/error para alcanzar una solución. La principal fuente de información ha sido la Documentación de Pentaho Community, en la que se proporcionan muchos ejemplos útiles para solucionar problemas que surgían durante la realización de cada iteración. Otras fuentes de consulta han sido diversos foros y páginas encontradas por la red a través de búsquedas por palabras clave de los problemas observados.

1.6. Tecnologías

Para la realización del trabajo se ha utilizado la plataforma Proxmox Entorno Virtual [16], o Proxmox VE, que es un entorno de virtualización de servidores de código abierto. Es una distribución Linux basada en Debian con un kernel modificado RHEL y permite desplegar y gestionar máquinas virtuales y contenedores.

Todo el sistema se ha realizado en una máquina de Linux 14.04 LTS ubicada en un contenedor de Proxmox con memoria RAM de 8GB, tamaño de disco 20GB y 2 CPU, antes de migrarlo a la máquina definitiva de producción. El contenedor tiene programadas por defecto copias de seguridad automáticas todos los días (guarda la última semana) y los sábados una semanal, de las que se guardan hasta 4 copias.

La Base de Datos donde se almacenan los formularios recogidos por los voluntarios de Cruz Roja es MySQL [17]. Esta BD nutrirá al proceso ETL.

La primera opción que se plantea para ubicar el DW, es la base de Datos HSQLDB [18]. Se considera ubicarlo ahí porque es la BD que Pentaho trae por defecto, ya instalada y configurada. Al estar integrada con el BI Server, la carga de datos al DW y utilización de estos para crear los Dashboards, puede ser mucho más rápida, que si el DW se ubicase en otra BD externa.

En la fase de desarrollo se modifica la BD donde se ubica el DW, debido a un problema detectado a la hora de realizar consultas sobre rango de fechas. Cuando se seleccionan rangos de fechas grandes a la primera, la BD falla y da error en la consulta. Sin embargo con MySQL esto no pasa, lo que se decide cambiar el DW a esta BD.

A continuación se citan las principales tecnologías que se usan o se requieren durante la realización del trabajo:

- ***Java Development Kit 8:** Es una plataforma de programación para desarrollar y ejecutar software de aplicaciones en lenguaje de programación Java con arquitectura de n niveles distribuidos, basándose ampliamente en componentes de software modulares.
- ***Tomcat 8:** Tomcat es un servidor web que al estar basado en Java, puede ser utilizado en cualquier sistema operativo que lo soporte. Es mantenido y desarrollado por miembros de la Apache Software Foundation y voluntarios independientes. Tomcat se utilizará para el despliegue del servidor BI de pentaho.
- ***Pentaho Business Analytics Platform:** La plataforma Pentaho BI proporciona una arquitectura e infraestructura para construir aplicaciones de inteligencia de negocio. Posee un enfoque moderno, simplificado, e interactivo que permite a los usuarios de negocio acceder, descubrir y mezclar cualquier tipo de datos, independientemente de su tamaño. Con una amplia gama de herramientas de análisis cada vez más avanzadas, desde informes básicos a modelos de predicción, los usuarios pueden ayudarse a sí mismos para analizar y visualizar los datos a través de múltiples medidas y dimensiones.

- ***Data Integration:** Herramienta que ofrece la extracción de gran alcance, transformación y capacidades de carga (ETL). Se puede utilizar esta aplicación independientemente visualizando las transformaciones de diseño y puestos de trabajo que extraen los datos existentes y que estén disponibles para facilitar la presentación de informes y análisis.
- **Report Designer:** El Diseñador de informes es una herramienta gráfica que genera informes de los datos obtenidos a través del motor de integración de datos sin la necesidad de ninguna tabla de etapas intermedias. Puede exportar sus informes en varios formatos, incluyendo PDF, Excel, HTML, XML, CSV.
- ***Marketplace:** Centro de aplicaciones de terceros y propias de Pentaho que pueden aportar una funcionalidad extra.
- **Aggregation Designer:** Es una herramienta para la creación y manejo de tablas agregadas.
- ***Schema Workbench:** Con esta herramienta se realizan los cubos OLAP.
- **Metadata Editor:** Herramienta que permite crear un modelo derivado de un DW con definiciones orientadas al negocio.
- **Hadoop Shims:** Herramienta para Big Data, almacenamiento de grandes cantidades de datos y procedimientos usados para encontrar patrones repetitivos dentro de esos datos, de Pentaho.

*Herramientas que se van a utilizar para la realización del proyecto.

1.7. Contenidos de la memoria

En el capítulo Iteración 1 “Instalación y configuración de Pentaho 6.0” se describen los pasos que se deben seguir para una correcta instalación y configuración, para una adecuada utilización posterior. Para ello es muy importante conocer todos los requisitos y características que posee la versión 6.0 de Pentaho.

El capítulo Iteración 2 “Proceso ETL” hace una descripción del proceso que se debe realizar para obtener las tablas de hechos y dimensiones necesarias para la obtención de un correcto DW con una topología en estrella o copo de nieve.

En la Iteración 3 “Creación del cubo OLAP” se describen las tareas que se deben seguir para una correcta formación de los cubos OLAP. Cada uno de ellos, con su tabla de hechos y dimensiones necesarias para obtener todo tipo de información necesaria para la toma de decisiones.

La Iteración 4 “Creación de Dashboard” se realiza una descripción de cómo se crean y se forman las pizarras o Cuadro de Mandos que se utilizaran para tomar decisiones. Es

una iteración muy importante, ya que es el resultado final de todo el Sistema de Información. Por lo que los Dashboard deben ser claros e intuitivos.

En definitiva, lo que se pretende realizar en este trabajo corresponde a las tareas o procesos que se muestran en el esquema de la Ilustración 2.

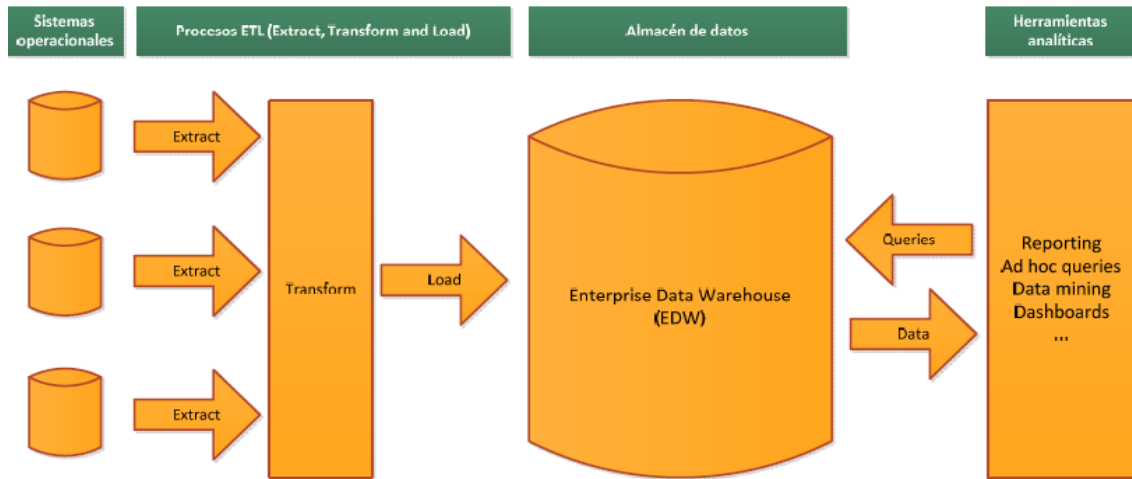


Ilustración 2 : Componentes del Sistema BI

2. ITERACIÓN 1-INSTALACIÓN Y CONFIGURACIÓN DE PENTAHO 6.0

Para lograr una mejor interacción con el contenedor de Proxmox, conviene instalar Putty [19], un cliente SSH y Telnet con el que podemos conectarnos a servidores remotos iniciando una sesión en ellos, que permite ejecutar comandos a través de una consola.

También se puede realizar esta conexión de una manera más gráfica y no por medio de consola. Para ello es necesario tener instalada una máquina virtual con una imagen de un sistema operativo ya creada y configurada para usarse, para este proyecto Ubuntu. Introduciendo en la terminal de esta nueva máquina el comando “ssh -X nombrecontenedor@dirIP” se obtiene un entorno gráfico del contenedor con el que se puede interactuar más gráfica e intuitivamente. El comando ssh -X establece conexiones tipo telnet de forma segura, además de la posibilidad del entorno gráfico anteriormente comentado.

Una vez realizada la primera toma de contacto con el servidor y definida una de las dos formas de interactuar, se comenzará con la instalación del sistema Pentaho Business Analytics Platform.

2.1. Instalación de JAVA 8

Todos los componentes que conforman Pentaho están desarrollados bajo la plataforma Java. Por lo que, en primer lugar, se pasa a descargar la última versión de Java compatible con Pentaho 6.0. En este caso Java 8.

1. “sudo apt-get install software-properties-common python-software-properties”

```
root@inf-srv-tfg1:~# sudo apt-get install software-properties-common python-software-properties
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following extra packages will be installed:
```

Ilustración 3: 1-Instalación Java 8

2. “sudo add-apt-repository ppa:webupd8team/java”

```
root@inf-srv-tfg1:~# sudo add-apt-repository ppa:webupd8team/java
Oracle Java (JDK) Installer (automatically downloads and installs Oracle JDK7 /
JDK8 / JDK9). There are no actual Java files in this PPA.

More info (and Ubuntu installation instructions):
- for Oracle Java 7: http://www.webupd8.org/2012/01/install-oracle-java-jdk-7-in-ubuntu-via.html
- for Oracle Java 8: http://www.webupd8.org/2012/09/install-oracle-java-8-in-ubuntu-via-ppa.html

Debian installation instructions:
- Oracle Java 7: http://www.webupd8.org/2012/06/how-to-install-oracle-java-7-in-debian.html
- Oracle Java 8: http://www.webupd8.org/2014/03/how-to-install-oracle-java-8-in-debian.html

Important!!! For now, you should continue to use Java 8 because Oracle Java 9 is available as an early access release (it should be released in 2016)! You should only use Oracle Java 9 if you explicitly need it, because it may contain bugs and it might not include the latest security patches! Also, some Java options were removed in JDK9, so you may encounter issues with various Java apps. More information and installation instructions (Ubuntu / Linux Mint / Debian): http://www.webupd8.org/2015/02/install-oracle-java-9-in-ubuntu-linux.html
More info: https://launchpad.net/~webupd8team/+archive/ubuntu/java
```

Ilustración 4: 2-Instalación Java 8

3. “sudo apt-get update”

```

root@inf-srv-tfg1: ~
root@inf-srv-tfg1:~# sudo apt-get update
Get:1 http://security.ubuntu.com trusty-security InRelease [64.4 kB]
Ign http://archive.ubuntu.com trusty InRelease
Ign http://archive.canonical.com trusty InRelease
Get:2 http://ppa.launchpad.net trusty InRelease [15.5 kB]
Hit http://archive.canonical.com trusty Release.gpg
Get:3 http://security.ubuntu.com trusty-security/main amd64 Packages [416 kB]
Hit http://archive.canonical.com trusty Release
Get:4 http://archive.ubuntu.com trusty-updates InRelease [64.4 kB]
Get:5 http://ppa.launchpad.net trusty/main amd64 Packages [3359 B]
Hit http://archive.canonical.com trusty/partner amd64 Packages
Get:6 http://ppa.launchpad.net trusty/main i386 Packages [3359 B]
Hit http://archive.canonical.com trusty/partner i386 Packages
Get:7 http://ppa.launchpad.net trusty/main Translation-en [1556 B]
Hit http://archive.canonical.com trusty/partner Translation-en
Hit http://archive.ubuntu.com trusty Release.gpg
Get:8 http://archive.ubuntu.com trusty-updates/main amd64 Packages [696 kB]
Get:9 http://security.ubuntu.com trusty-security/restricted amd64 Packages [13.0
kB]
Get:10 http://security.ubuntu.com trusty-security/universe amd64 Packages [123 k
B]
Get:11 http://archive.ubuntu.com trusty-updates/restricted amd64 Packages [16.0
kB]
Get:12 http://archive.ubuntu.com trusty-updates/universe amd64 Packages [336 kB]
  
```

Ilustración 5: 3-Instalación Java 8

4. “sudo apt-get install oracle-java8-installer”

```

root@inf-srv-tfg1: ~
root@inf-srv-tfg1:~# sudo apt-get install oracle-java8-installer
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following extra packages will be installed:
  gsfonts gsfonts-x11 java-common libfontenc1 libxfont1 x11-common
  xfonts-encodings xfonts-utils
Suggested packages:
  default-jre equivs binfmt-support visualvm ttf-baekmuk ttf-unfonts
  ttf-unfonts-core ttf-kochi-gothic ttf-sazanami-gothic ttf-kochi-mincho
  ttf-sazanami-mincho ttf-arphic-uming firefox firefox-2 iceweasel
  mozilla-firefox iceape-browser mozilla-browser epiphany-gecko
  epiphany-webkit epiphany-browser galeon midbrowser moblin-web-browser
  xulrunner xulrunner-1.9 kongqueror chromium-browser midori google-chrome
The following NEW packages will be installed:
  gsfonts gsfonts-x11 java-common libfontenc1 libxfont1 oracle-java8-installer
  
```

```

root@inf-srv-tfg1: ~
Package configuration
#####

#####Configuring oracle-java8-installer#####
x In order to install this package, you must accept the license terms, x
x the "Oracle Binary Code License Agreement for the Java SE Platform x
x Products and JavaFX ". Not accepting will cancel the installation. x
x x
x Do you accept the Oracle Binary Code license terms? x
x x
x x
#####
x < Yes > < No >
#####
  
```

Ilustración 6: 4-Instalación Java 8

Introduciendo en la terminal el comando “java -version” obtenemos la versión de Java que se acaba de instalar.

2.2. Configuración de las variables de entorno

Se deben configurar dos variables de entorno para que el sistema funcione correctamente. Para ello se debe modificar el fichero `.bashrc` como se puede ver en la siguiente imagen.

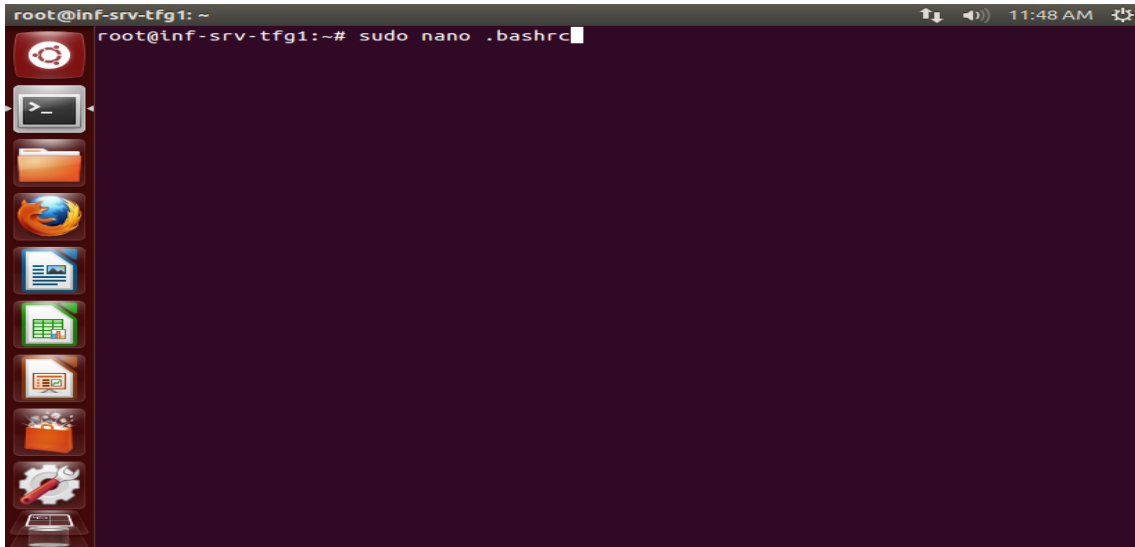


Ilustración 7: Comando para modificar fichero de variables de entorno

Una vez dentro del archivo, al final del mismo, se añaden las 4 últimas filas que se ven en la siguiente imagen. Se le llama `PENTAHO_JAVA_HOME` para diferenciar de futuras aplicaciones que también necesiten de JAVA.

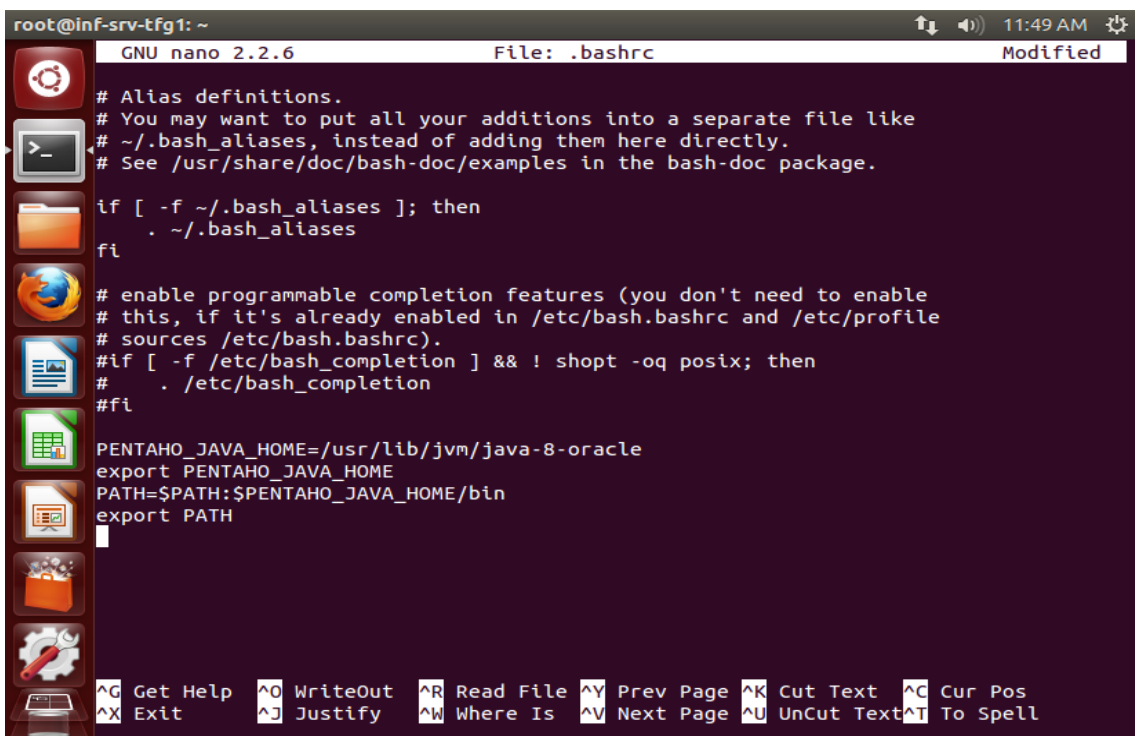
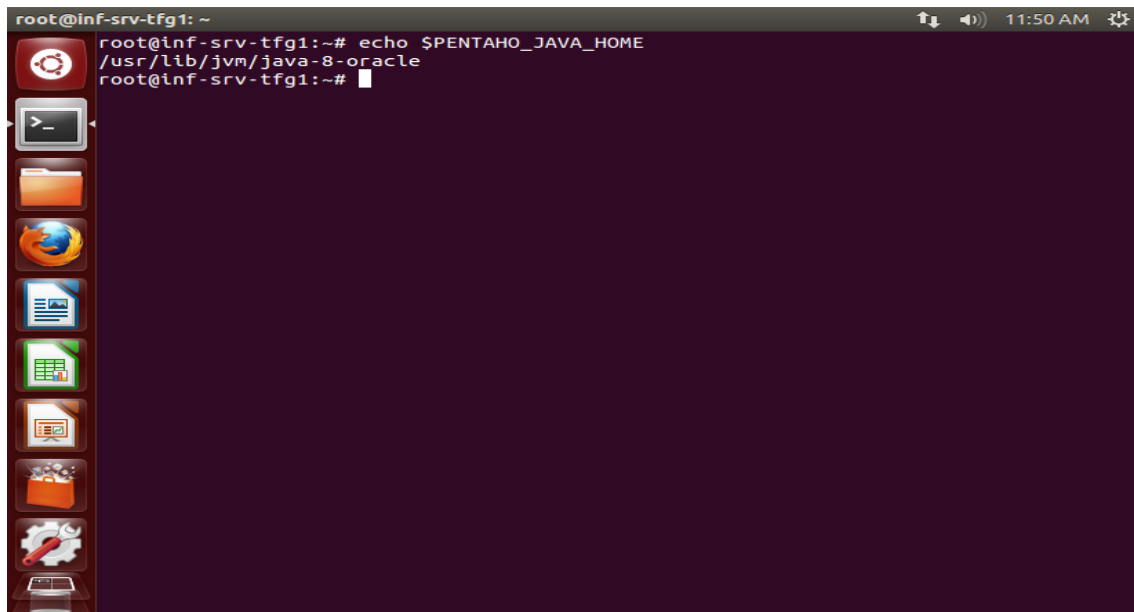


Ilustración 8: Modificar variables de entorno

Se comprueba que la variable PENTAHO_JAVA_HOME apunta a donde se quiere para el buen funcionamiento del sistema.



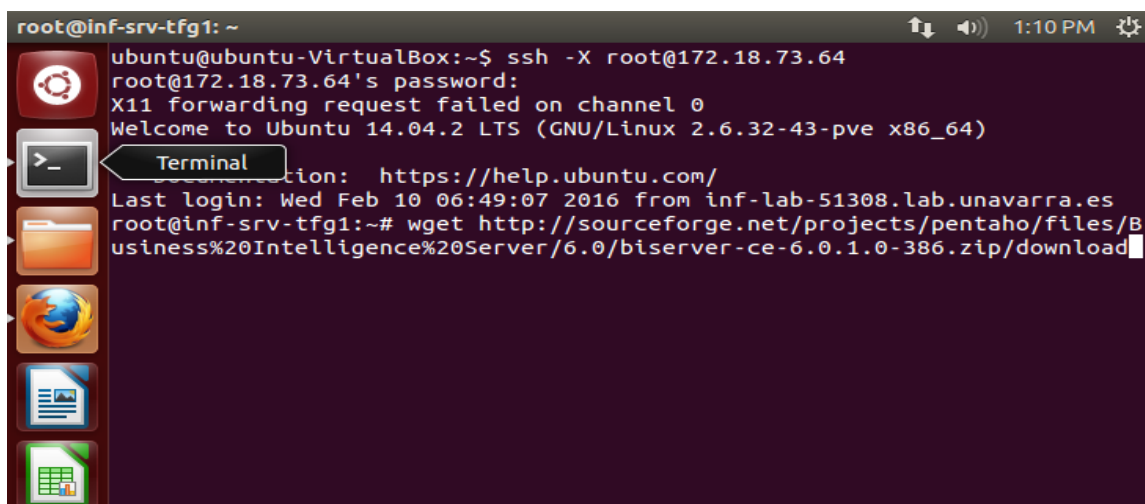
```
root@inf-srv-tfg1: ~  
root@inf-srv-tfg1:~# echo $PENTAHO_JAVA_HOME  
/usr/lib/jvm/java-8-oracle  
root@inf-srv-tfg1:~#
```

Ilustración 9: Comprobación de variables

2.3. Instalación de Pentaho 6.0

Después de realizar estos cambios de variables de entorno ya está todo preparado para descargar e instalar Pentaho 6.0.


Se descarga BI server 6.0 de la página web community oficial introduciendo los siguientes comandos. wget que permite realizar descargar de una forma muy sencilla a través de consola:



```
root@inf-srv-tfg1: ~  
ubuntu@ubuntu-VirtualBox:~$ ssh -X root@172.18.73.64  
root@172.18.73.64's password:  
X11 forwarding request failed on channel 0  
Welcome to Ubuntu 14.04.2 LTS (GNU/Linux 2.6.32-43-pve x86_64)  
Last login: Wed Feb 10 06:49:07 2016 from inf-lab-51308.lab.unavarra.es  
root@inf-srv-tfg1:~# wget http://sourceforge.net/projects/pentaho/files/B  
usiness%20Intelligence%20Server/6.0/biserver-ce-6.0.1.0-386.zip/download
```

Ilustración 10: Descarga de Pentaho BIServer 6.0

Antes de descomprimir el archivo anteriormente descargado, es necesario crear una carpeta, la que se llamará “pentaho”, en la raíz /home/ de nuestro sistema. Para crear la carpeta simplemente añadimos por línea de comando “mkdir pentaho”. Una vez creada la carpeta, se descomprime en ella para tener ordenados y localizados todos los componentes necesarios durante la realización del proyecto.

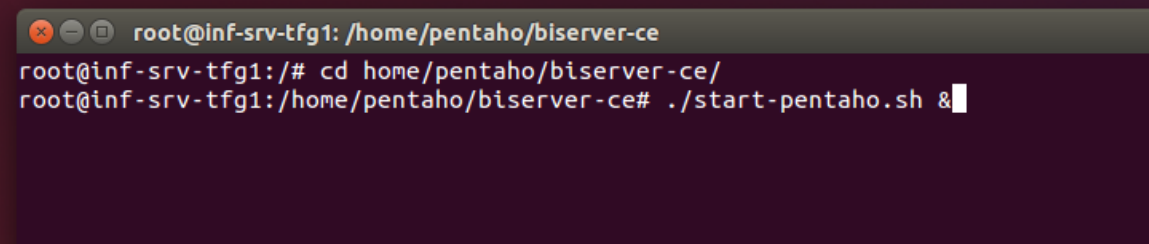


```
root@inf-srv-tfg1: ~
ubuntu@ubuntu-VirtualBox:~$ ssh -X root@172.18.73.64
root@172.18.73.64's password:
X11 forwarding request failed on channel 0
Welcome to Ubuntu 14.04.2 LTS (GNU/Linux 2.6.32-43-pve x86_64)

 * Documentation:  https://help.ubuntu.com/
Last login: Wed Feb 10 06:49:07 2016 from inf-lab-51308.lab.unavarra.es
root@inf-srv-tfg1:~# unzip download
```

Ilustración 11: Unzip Pentaho BIServer

Una vez finalizado el proceso de descompresión, tendremos todo preparado para poder iniciar Pentaho BI. Para ello solo se debe ejecutar el archivo que se encuentra dentro de la carpeta “/home/pentaho/biserver-ce/”, llamado “start-pentaho.sh”, con el que iniciaremos el servidor Tomcat sobre el que se aloja Pentaho y accedemos desde el navegador introduciendo la url “inf-srv-tfg1.lab.unavarra.es:8080”.



```
root@inf-srv-tfg1: /home/pentaho/biserver-ce
root@inf-srv-tfg1:/# cd home/pentaho/biserver-ce/
root@inf-srv-tfg1:/home/pentaho/biserver-ce# ./start-pentaho.sh &
```

Ilustración 12: Iniciar Pentaho

Para este trabajo se utiliza el rol de administrador que trae por defecto Pentaho, al cual se accede introduciendo los siguientes datos de login, usuario: “admin”, contraseña: “password”. Cabe destacar que Pentaho permite gestionar los roles de usuario, y así poder dar permisos específicos a distintos usuarios dependiendo de la utilización o gestión de Pentaho.



Ilustración 13: Pantalla de inicio Pentaho

Para el resto de componentes no es necesario configurar nada a nivel de variables locales. Simplemente descargamos el archivo comprimido desde la página web de Pentaho Community y lo descomprimimos en la carpeta “PENTAHO” que tenemos creada en la raíz “home/pentaho/”, como se ha realizado anteriormente. Dentro de esa carpeta descomprimida encontraremos un archivo con el nombre de la aplicación en cuestión y ejecutándolo directamente se iniciará dicha aplicación.

Para este proyecto es necesario contar, además de con Pentaho BI Server que ya tenemos satisfactoriamente instalado y configurado, con:

- Data Integration
- Schema Workbench
- Report Designer

2.4. Instalación librería Java

La base de datos, MySQL ya que es la base de datos donde se almacenan los formularios que Cruz Roja recoge, es necesario tenerla conectada para su posterior utilización, en este caso al Data Integration, para cargar los datos del proceso ETL. Para ello se debe añadir la librería java <http://dev.mysql.com/downloads/connector/j/> en la siguiente ruta: /home/pentaho/data-integration/lib.

2.5. Instalación de plugins necesarios

En esta versión de Pentaho, no es necesario instalar el Marketplace ya que viene instalado por defecto.

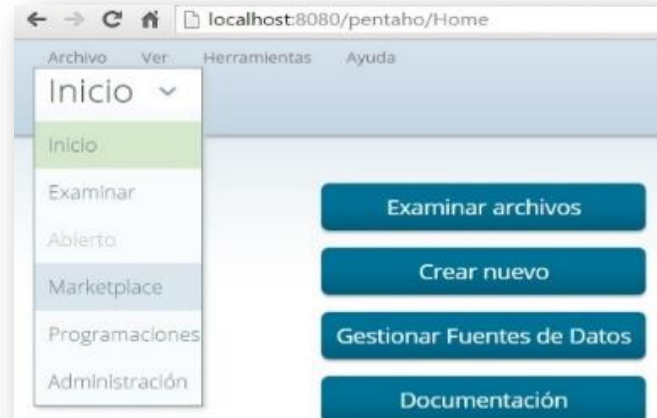


Ilustración 14: Marketplace

Por lo que se va directamente a instalar la versión 3.8.8 de Saiku Analytics. Este plugin va a permitir hacer consultas de una manera mucho más intuitiva y gráfica. Se instalará esta versión puesto que es la versión más reciente para Pentaho 6.0.

Buscamos el plugin Saiku Analytics como se ve en la siguiente ilustración.

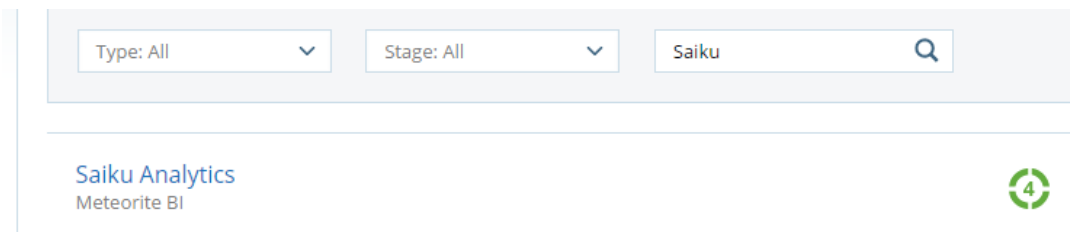


Ilustración 15: Saiku Analytics

Y a continuación se instala.

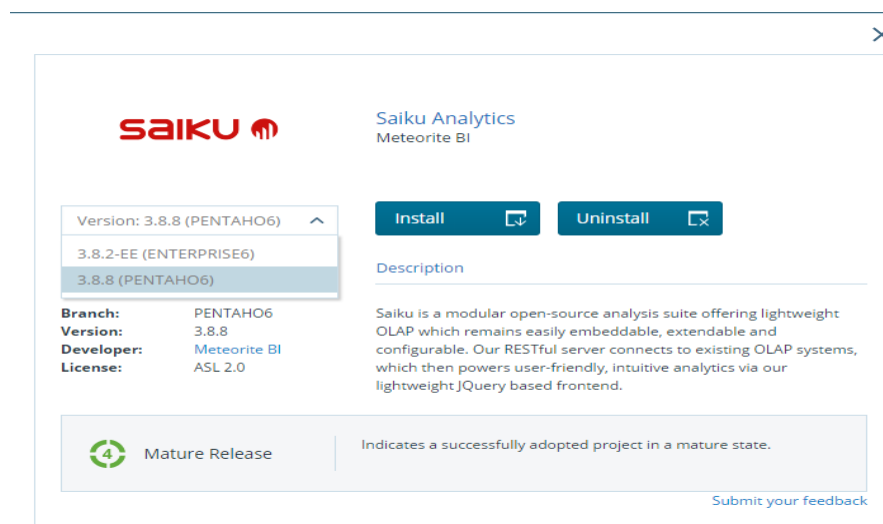


Ilustración 16: Saiku Analytics 3.8.8 Pentaho 6

3. ITERACIÓN 2-PROCESO ETL

Una vez instalado y configurado todo lo anterior ya se dispone de lo necesario para poder utilizar Pentaho y sus componentes. Cuando se encuentra en este punto, toca ponerse a desarrollar el sistema BI, empezando por el proceso ETL.

Para la realización de este TFG se dispone de una Base de Datos con todos los datos de los formularios necesarios para cumplir este proceso y poblar nuestro DW. Cabe destacar que la extracción se puede realizar de distintas fuentes de datos,

Ya se ha explicado cómo trabaja Cruz Roja sobre el terreno, y cómo obtienen los datos en formularios, mediante desplegados con palabras, frases o abreviaturas que se almacenan en la Base de Datos y posteriormente poblar el DW. Cabe destacar que la extracción de datos se puede realizar de diferentes fuentes, ya sean archivos Excel, archivos de texto o como en el caso de este trabajo, diferentes bases de datos, lo que supone realizar un proceso de unificación.

Para realizar el proceso ETL se debe tener instalada la herramienta Data Integration, que posee, entre otras cosas, un entorno gráfico llamado “spoon”. Aquí se realizarán todas las operaciones necesarias para definir correctamente nuestro Datawarehouse, de aquí en adelante DW. Para ello se ejecuta el archivo “spoon.sh” que se encuentra en la carpeta “/home/pentaho/data-integration/” como se aprecia en la Ilustración 17.

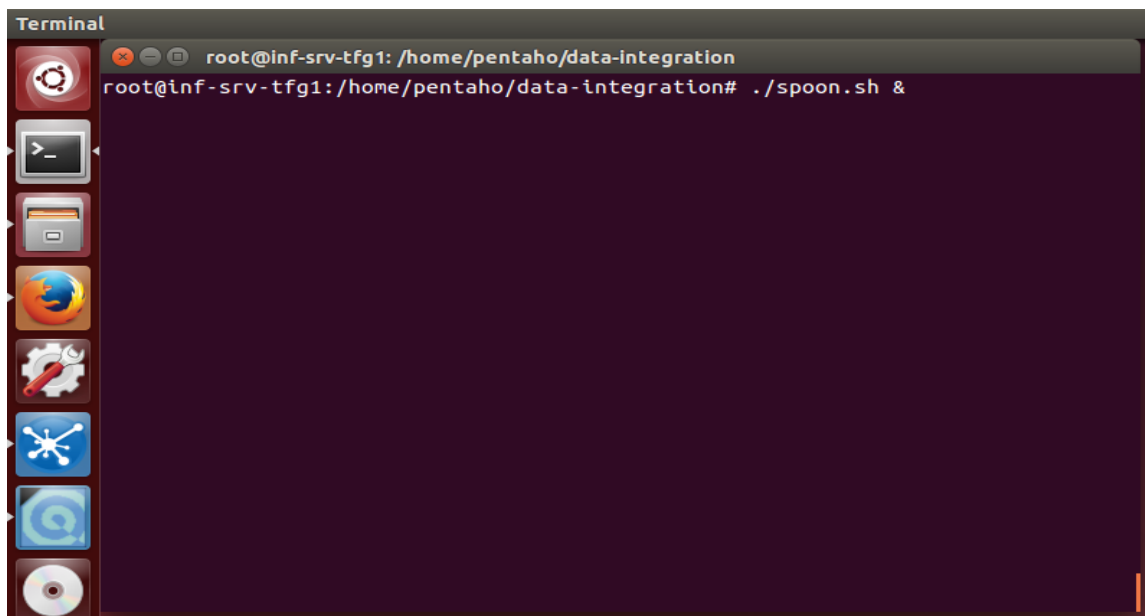


Ilustración 17: Iniciar Spoon

Una vez iniciada la aplicación nos encontraremos, tal y como se aprecia en la siguiente ilustración, con una pantalla inicial en la que aparece una presentación y una serie de documentación sobre la aplicación.

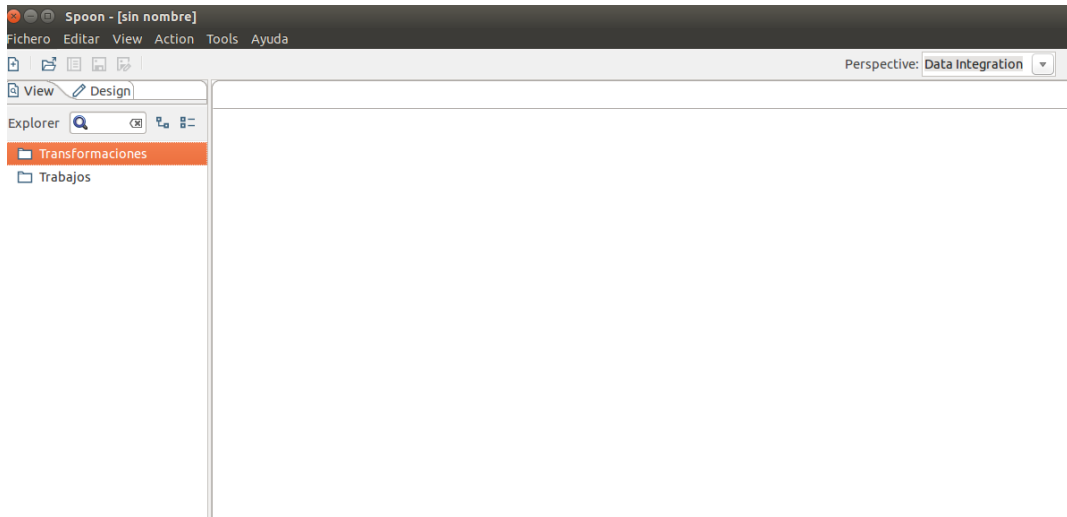


Ilustración 18: Pantalla inicial Spoon

3.1. Extracción

Para iniciar el proceso se debe crear un nuevo archivo “.ktf” que tiene de nombre “Transformación” como se indica en la siguiente ilustración.

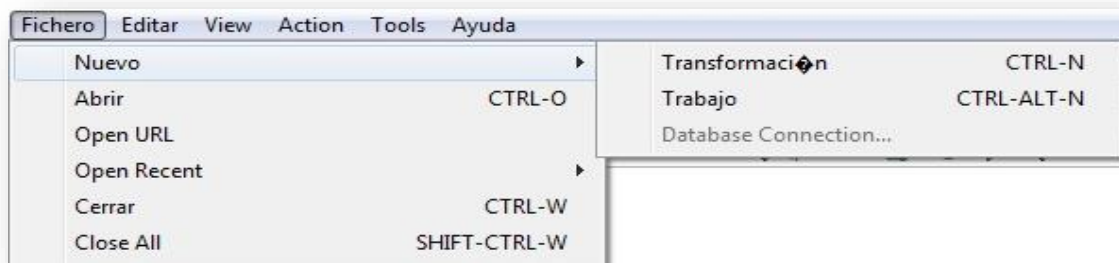


Ilustración 19: Fichero, Nuevo, Transformación

Como ya se ha explicado anteriormente, se pueden extraer los datos de diversas fuentes y en nuestro caso se hará desde una base de datos MySQL. Para realizar dicha conexión hay que seguir los siguientes pasos:

1. Doble click en entrada de datos.
2. Elegir el tipo de BD, en este caso MySQL.
3. Rellenar los campos como se ve en la Ilustración 20.
4. Probamos que la conexión es correcta.

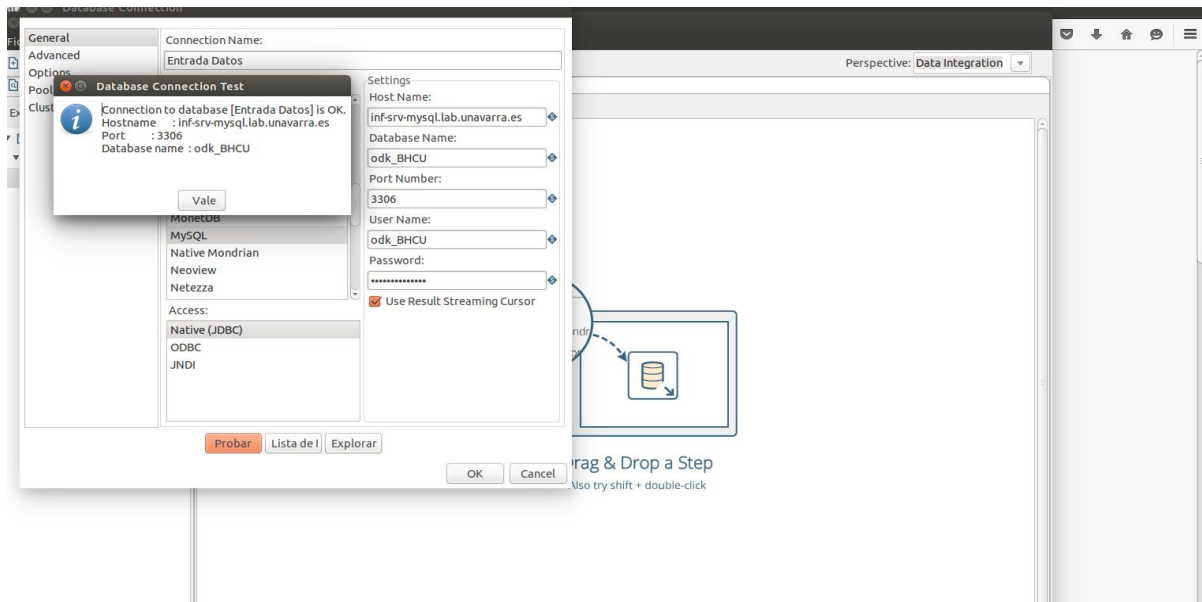


Ilustración 20: Conexión datos entrada

3.1.1 Entrada de tabla

En primer lugar, se debe dar entrada a los datos mediante la opción “Entrada Tabla”, donde elegiremos la conexión a la BD anteriormente creada como fuente de ingreso de los datos necesarios.

La opción “Entrada Tabla” se encuentra en el menú de la izquierda, en la carpeta “Entrada”, donde arrastraremos la opción requerida a la ventana de trabajo en la parte central y más amplia de spoon.

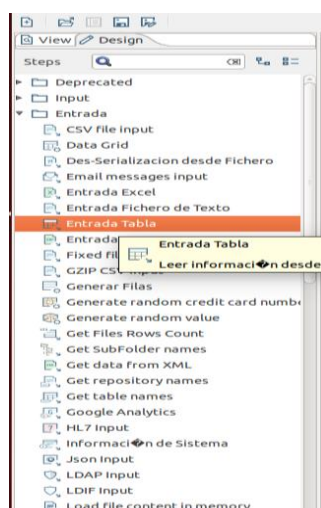


Ilustración 21: Entrada tabla

Al hacer doble click sobre el icono, se abrirá una pantalla donde se elegirá la conexión de entrada de datos anteriormente configurada. Se podrá realizar una consulta SQL que nos permitirá obtener las filas y columnas que consideremos oportunas para poblar

nuestro DW. En la Ilustración 22 se puede ver un ejemplo de cómo extraer los datos de una tabla de la BD de extracción.

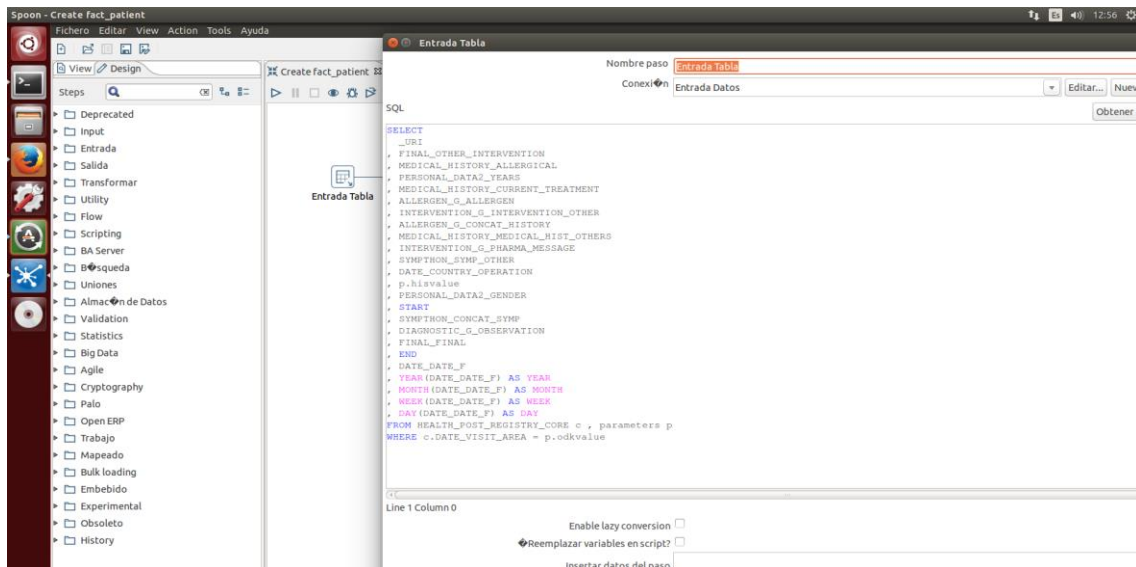


Ilustración 22: Entrada tabla 2

Para una mayor claridad y simplicidad del trabajo se aconseja crear tantas transformaciones como tablas vaya a tener el DW. Así podremos acceder al archivo “.ktl” que queremos modificar sin tener que estar pendientes de cambiar otras tablas que no queremos.

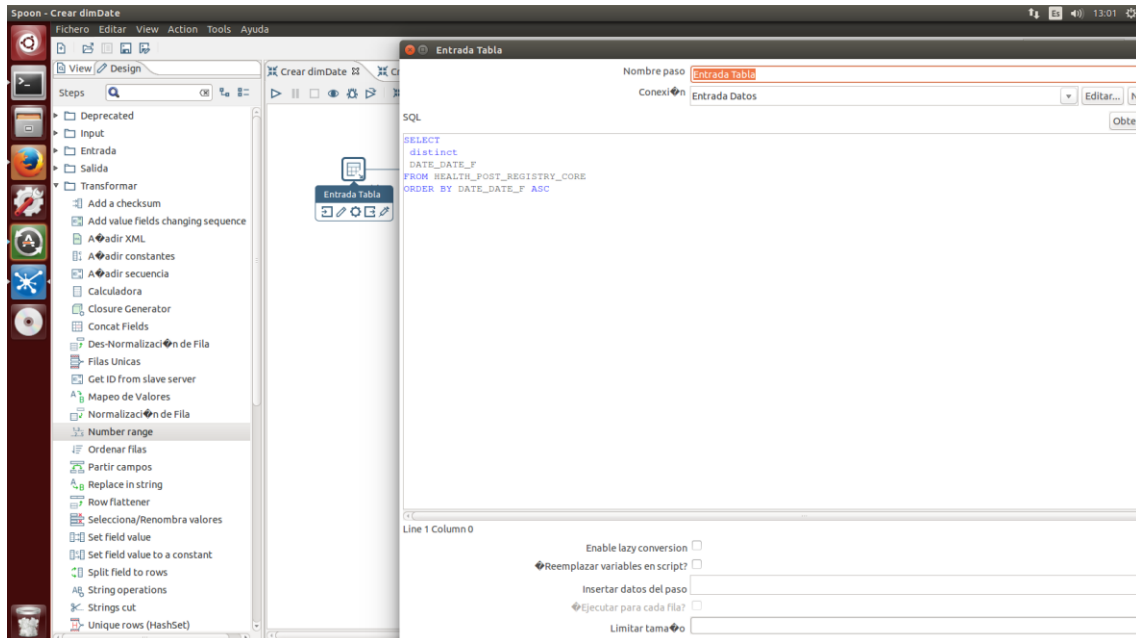


Ilustración 23: Entrada tabla 3

3.2. Transformación

La transformación es la segunda tarea que se realiza dentro del proceso ETL. Este es un paso muy importante, ya que se deben unificar formatos, limpiar y tratar los datos de las diferentes fuentes de obtención de los mismos. En este paso se pueden realizar transformaciones como seleccionar ciertas columnas de las tablas cargadas, traducción de códigos, división de una columna en varias, u obtener valores calculados entre dos columnas para facilitar las operaciones en el futuro.

Estos son unos pocos ejemplos de lo que se puede hacer en la transformación, en la que destacaría unificar formatos, ya que el DW debe ser integrado, la información no se modifica ni se elimina, una vez almacenado un dato, se convierte en información solo de lectura, y se mantiene para futuras consultas. Y no volátil, el DW puede contener datos de varios sistemas operacionales, y estos deben ser consistentes.

Siguiendo con el proyecto y con la extracción de datos que ya se había realizado, se continúa con la segunda fase del proceso.

En primer lugar se pueden renombrar las columnas para un correcto entendimiento de que información contiene cada una en el DW, se aconseja seguir una notación común para diferenciar las tablas de hechos, tabla central de un esquema en estrella o en copo de nieve y contiene los valores de las medidas de negocio o dicho de otra forma los indicadores de negocio, de las dimensiones, tablas que contienen atributos (o campos) que se utilizan para restringir y agrupar los datos almacenados en una tabla de hechos cuando se realizan consultas sobre dicho datos, en la Ilustración 24 se puede observar dicha tarea.

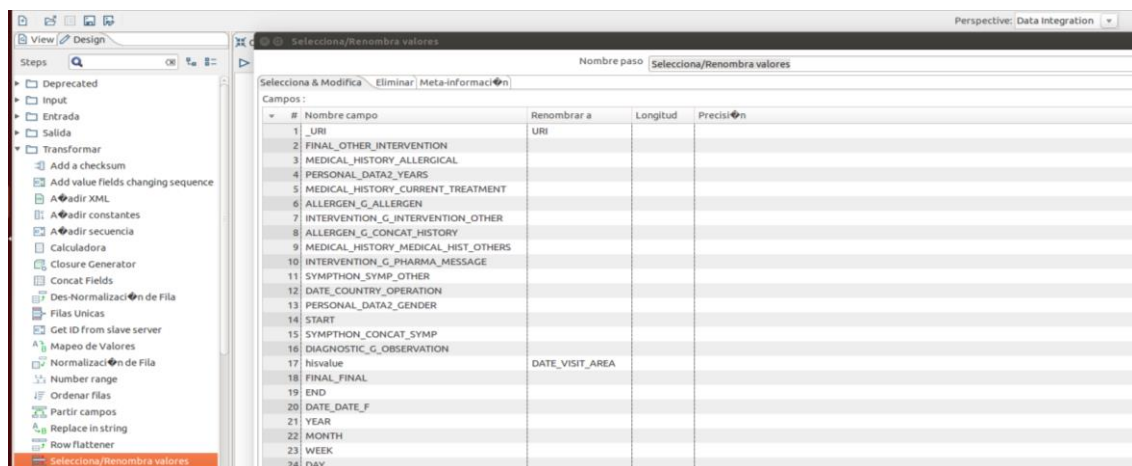


Ilustración 24: Transformación Renombrar

Otra tarea muy importante a la hora de unificar formatos es la de “replace”, donde se pueden reemplazar cadenas de texto, en la BD los datos de algunos campos, como por ejemplo FINAL_FINAL, vienen con “_” en vez de “ ”, por lo que se utilizará esta tarea para solucionar y unificar esto.

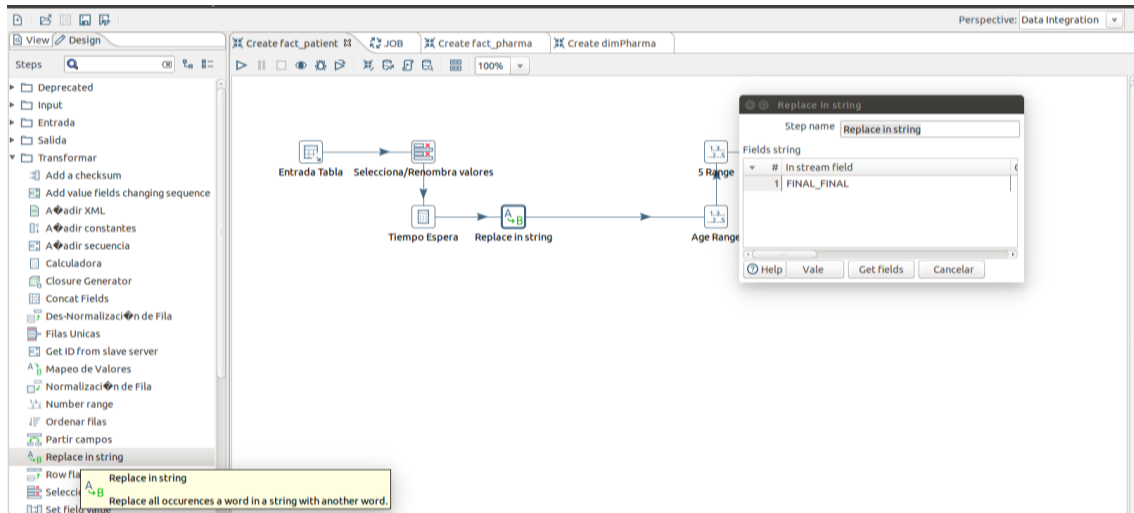


Ilustración 25: Transformación Replace

Para terminar con este proceso, se pueden crear nuevas columnas con datos calculados que posteriormente son utilizados para medir ciertas cosas. En nuestro ejemplo se utilizarán tareas como “calculadora”, para calcular el tiempo medio de consulta y la tarea “range”, que creará nuevas columnas con los rangos de edad calculados de una columna que ya se posee.

En la siguiente ilustración se puede observar un ejemplo de una columna calculada con la tarea “range”.

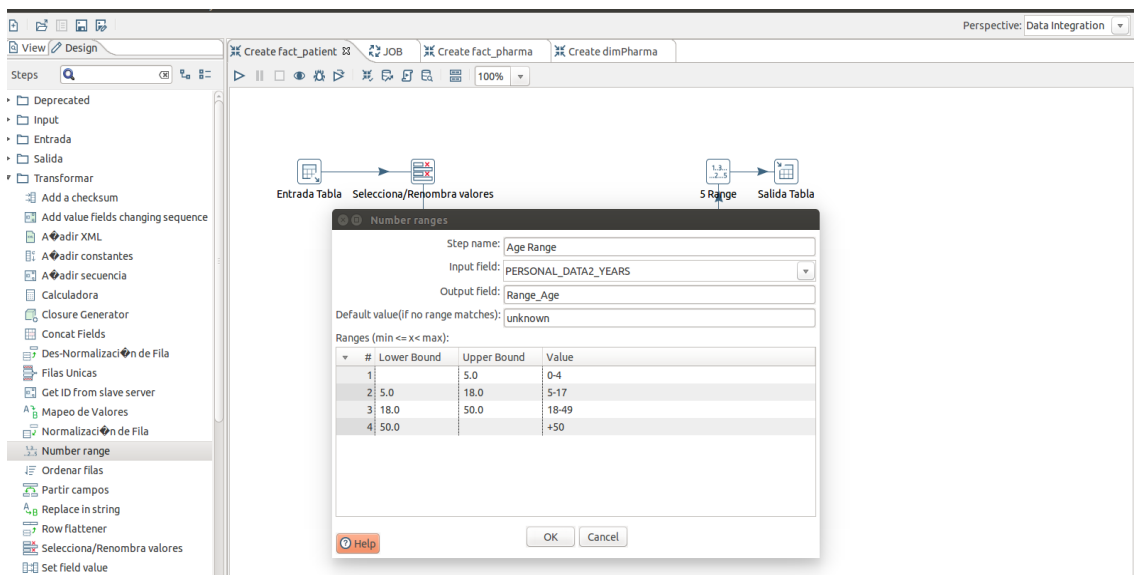


Ilustración 26: Transformación Range

3.3. Carga

La tercera y última fase del proceso ETL consiste en la carga de los datos en el DW. Es un proceso muy parecido al de extracción, pero a la inversa. Como se ha realizado anteriormente, hay que establecer la conexión con el DW siguiendo prácticamente los mismos pasos:

1. Doble click en Salida de Datos.
2. Elegir el tipo de BD, en este caso MySQL.
3. Rellenar los campos como se ve en la siguiente ilustración.
4. Probamos que la conexión es correcta.

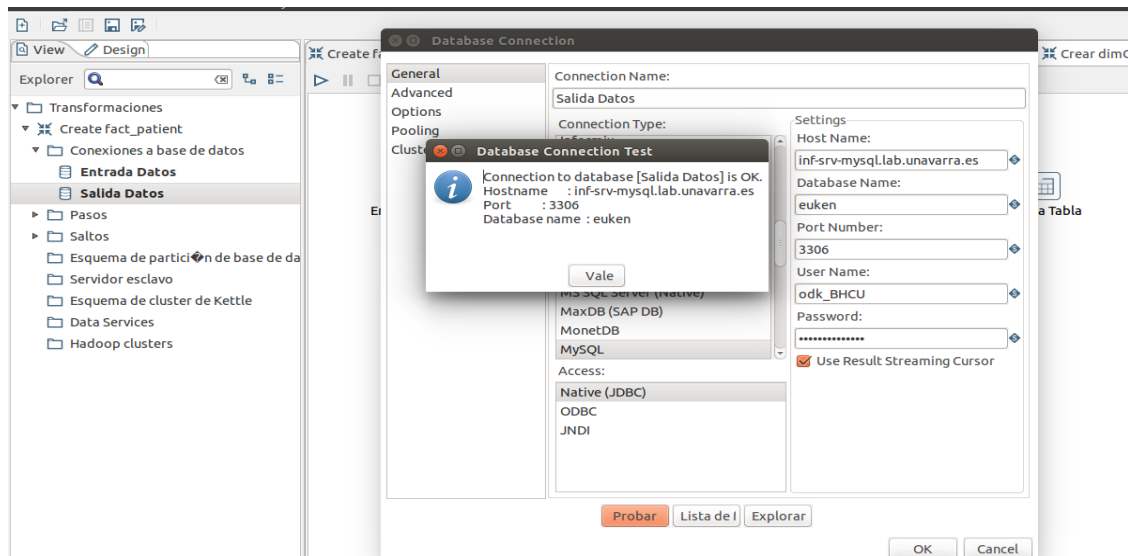


Ilustración 27: Conexión datos salida

3.3.1. Salida Tabla

Para realizar la carga de datos en el DW, se les debe dar salida mediante la tarea “Salida Tabla”, donde elegiremos la conexión a la BD anteriormente creada como carga de los datos necesarios en el DW.

La opción “Salida Tabla” se encuentra en el menú de la izquierda en la carpeta “Salida”, donde arrastraremos la opción requerida a la ventana de trabajo como se puede apreciar en la Ilustración 28.

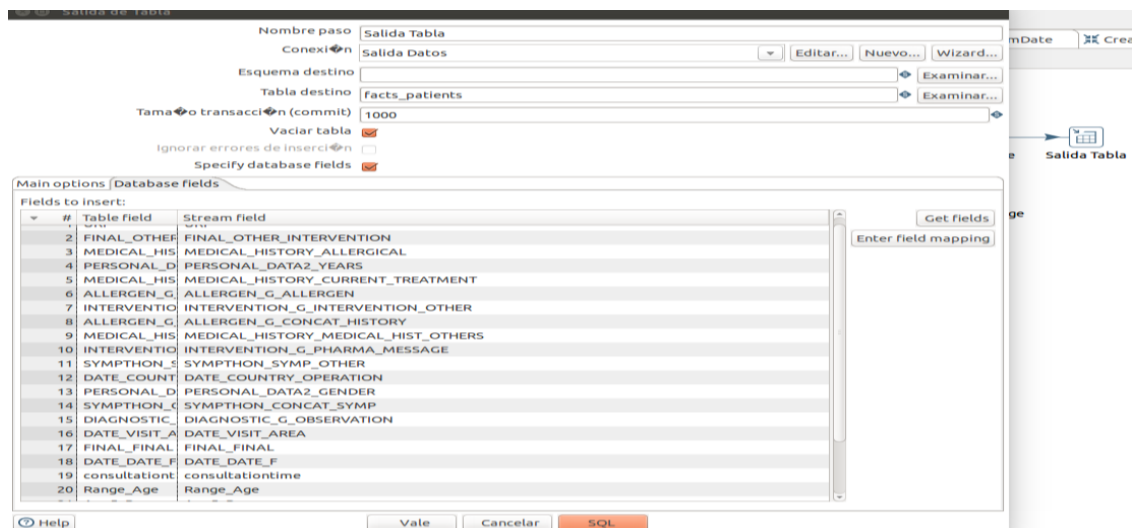


Ilustración 28: Salida tabla

Se puede observar que en conexión hemos seleccionado Salida Datos y hemos nombrado la tabla que se almacenará en el DW. También seleccionamos vaciar tabla y especificamos los campos de la tabla, en la pestaña Database fields, pulsando Get fields se añadirán automáticamente todos los campos que habíamos extraído y transformado.

Pulsando sobre el botón “SQL” se ejecuta una consulta que creará la tabla en el DW. Después de realizar todo esto, el DW tendrá las tablas creadas con todos los campos seleccionados, pero los datos todavía no habrán sido cargados.

Una vez que se tienen todas la tareas, tanto las de extracción, transformación y carga, habrá que ejecutar el fichero de transformación. En la Ilustración 29 se puede observar.

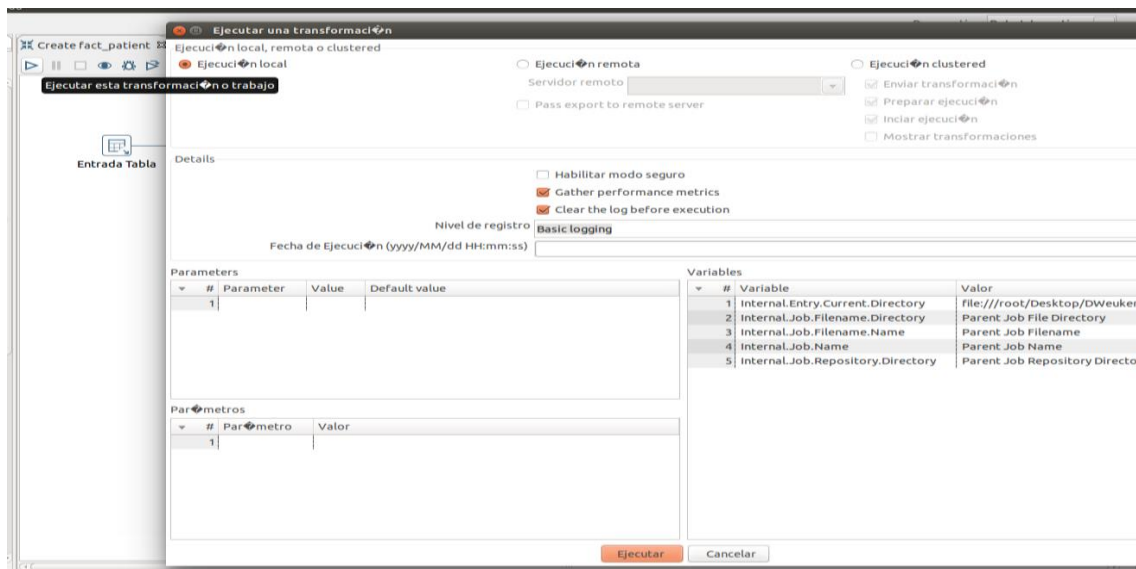


Ilustración 29: Ejecutar transformación

3.4. Job

Una vez realizado todo el proceso ETL anteriormente definido, spoon da la posibilidad de crear archivos “.kjb” en los que se pueden realizar todos los procesos de extracción, transformación y carga con una única ejecución.

Se arrastra del menú de la izquierda la tarea “start” y tantas tareas TRANSFORMACIÓN como procesos ETL se hayan realizado para poblar el DW. En las opciones de la tarea TRANSFORMACIÓN se añade la ruta donde se encuentra cada una de ellas. Cada TRANSFORMACIÓN debe estar unida a una tarea “sucess”. Se unen a diferentes por si falla una TRANSFORMACIÓN, no afecte a los demás procesos ETL, y así localizar rápidamente donde se ha producido el error. En la siguiente ilustración se puede ver:

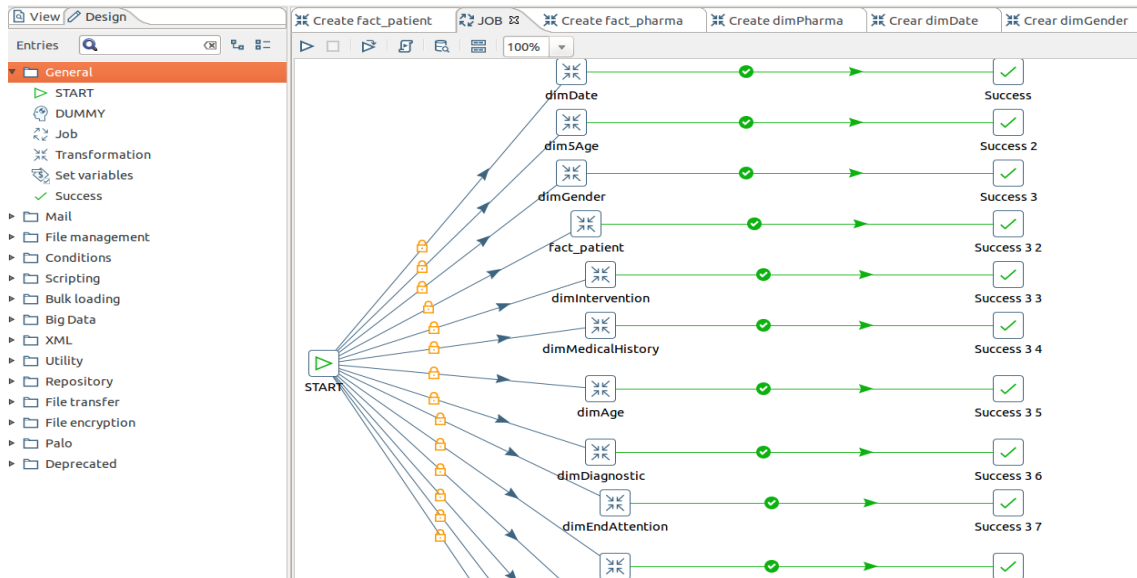


Ilustración 30: Vista general JOB

JOB se ejecutará de la misma manera que se ha hecho cada TRANSFORMACIÓN. También cabe la posibilidad de programar ejecuciones de JOB pulsando sobre la tarea “start”.

The 'Job Scheduling' dialog box has the following settings: Repeat is unchecked; Type is set to 'Daily'; Interval in seconds is 0; Interval in minutes is 60; Time of day is 12:00; Day of week is Monday; and Day of month is 1. At the bottom, there are buttons for 'Help', 'Vale', and 'Cancelar'.

Ilustración 31: Programación JOB

4. ITERACIÓN 3- CREACIÓN DEL CUBO OLAP

Después de realizar todo lo explicado durante la segunda iteración, y asignar correctamente tanto las claves primarias y extranjeras desde la BD donde se aloja el DW, se dispone de un modelo en estrella necesario para continuar con la siguiente iteración, en la que se construirán diferentes cubos OLAP de manera correcta.

Schema workbench es la herramienta gráfica para la creación de cubos en Pentaho. Para acceder a ella es necesario tener instalada dicha herramienta. Para abrirla se ejecuta el archivo “workbench.sh” que se encuentra en nuestro caso en la carpeta “/home/pentaho/schema-workbench/”, como se aprecia en la Ilustración 32.

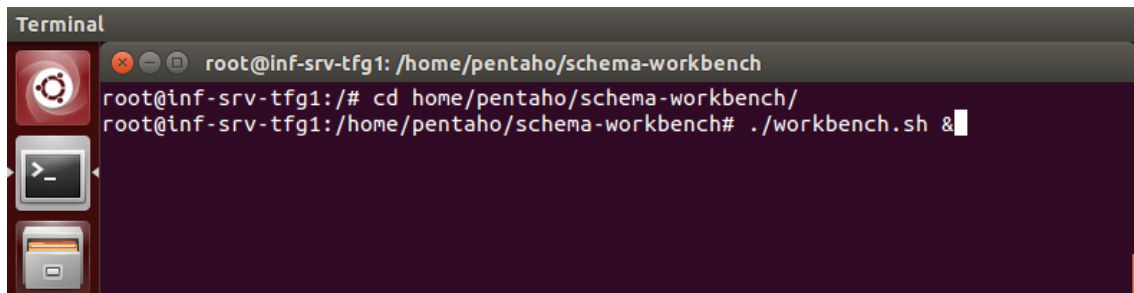


Ilustración 32: Iniciar Schema Workbench

Antes de empezar a trabajar con *Schema Workbench*, es necesario que la base de datos esté conectada, en este caso a *Schema Workbench*, para cargar los datos del DW. Para ello, se debe copiar el archivo JAR, que se puede encontrar descargándolo desde <http://dev.mysql.com/downloads/connector/j/>, y copiarlo en la siguiente ruta: “/home/pentaho/schema-workbench/drivers”.

4.1. Conexión al Datawarehouse

Una vez iniciada la aplicación, y como sucede durante el proceso ETL, hay que establecer una conexión con nuestro DW, accediendo a “options” y posteriormente a “database connection”. Aparecerá un cuadro de opciones de conexión como el que se muestra en la Ilustración 33.

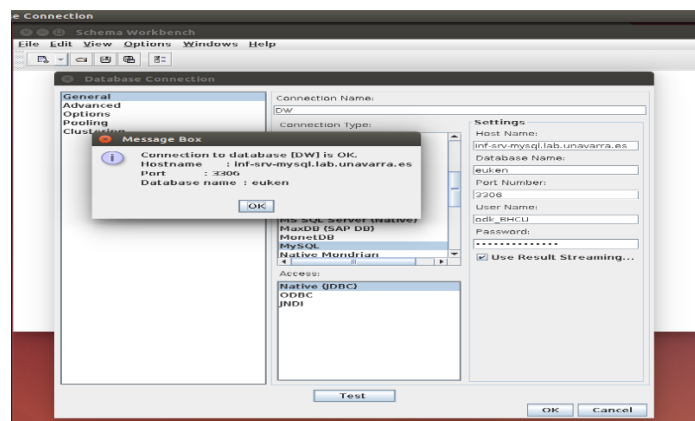


Ilustración 33: Conexión al DW

4.2. Nuevo esquema y añadir cubo

Para iniciar el proceso de creación del cubo se tiene que abrir un nuevo esquema, tal y como se muestra en la Ilustración 34.

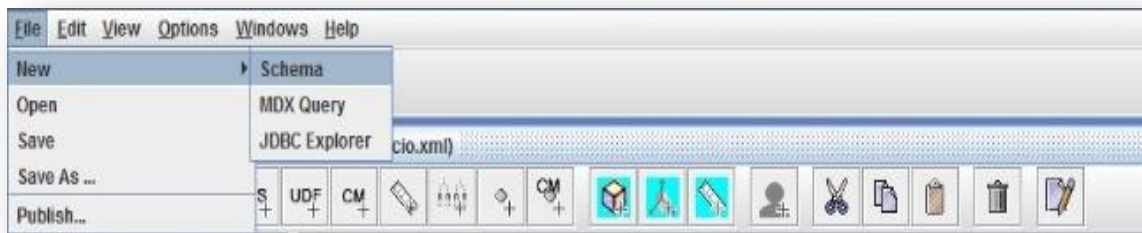


Ilustración 34: File, New, Schema

Una vez se tiene el esquema creado, ya está todo listo para iniciar la creación del cubo OLAP. Se pulsa con el botón derecho sobre Schema y se añade el cubo.

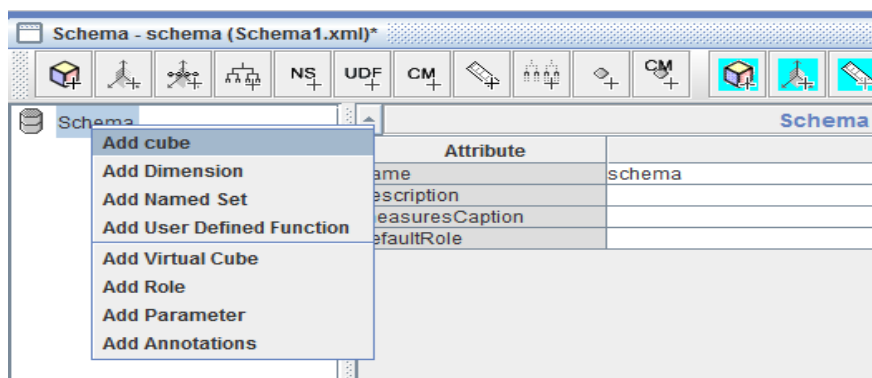


Ilustración 35: Añadir cubo

A continuación se debe añadir la tabla de hechos correspondiente al cubo que se quiere crear, de similar forma a la anterior, se pulsa el botón derecho pinchando en el cubo creado anteriormente y se añade la tabla de hechos, en este caso "patients". En la Ilustración 36 se puede observar el proceso.

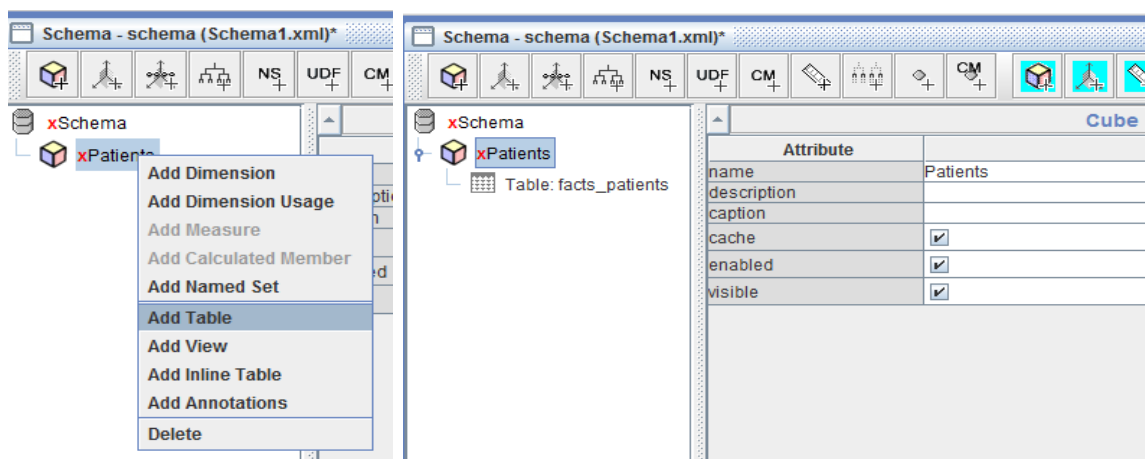


Ilustración 36: Añadir tabla al cubo

Una vez se añade la tabla de hechos al cubo, la propia herramienta, pide que se añadan dimensiones.

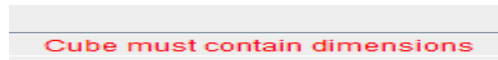


Ilustración 37: Se debe añadir dimensiones

4.3. Crear dimensiones

A la hora de crear o añadir las dimensiones, se decide crearlas en el esquema y no en el cubo como suele ser lo habitual. Esto se realiza para poder reutilizar las mismas dimensiones en varios cubos del mismo esquema.

En primer lugar se selecciona el esquema, y pulsando el botón derecho o seleccionando en la parte superior del área de trabajo, se elige “Add Dimension”. En la siguiente ilustración se puede observar cómo debe realizarse.

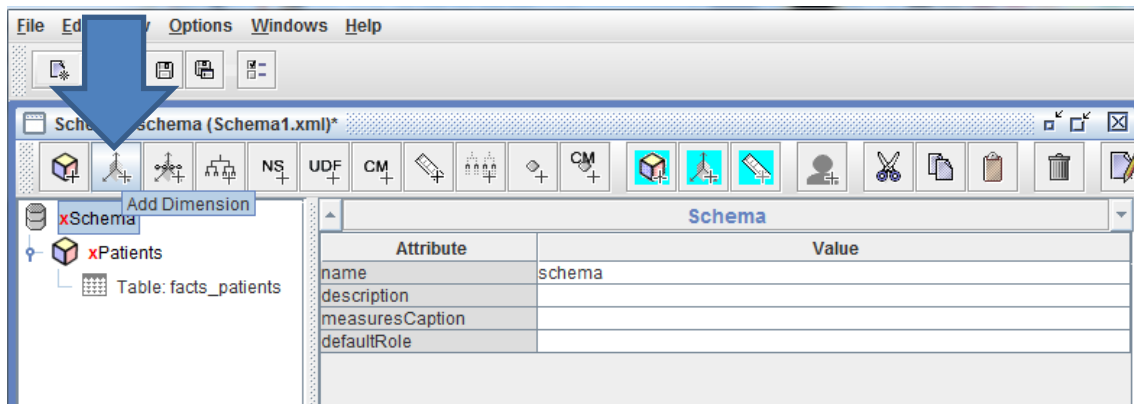


Ilustración 38: Añadir dimensión

Una vez se ha creado la dimensión, se le añade el nombre correspondiente, tanto el nombre de la misma, como el nombre del nivel. También, como en el caso del cubo, se añade la tabla de dimensión correspondiente, en este caso la dimensión tiempo. Esta dimensión es muy importante durante todo el proceso. Se añade la clave primaria de la dimensión, como ya se ha explicado anteriormente, las claves tanto primarias o extranjeras son muy importantes para el buen funcionamiento del cubo y esquema en estrella. En la siguiente ilustración se pueden observar todos los pasos a dar.

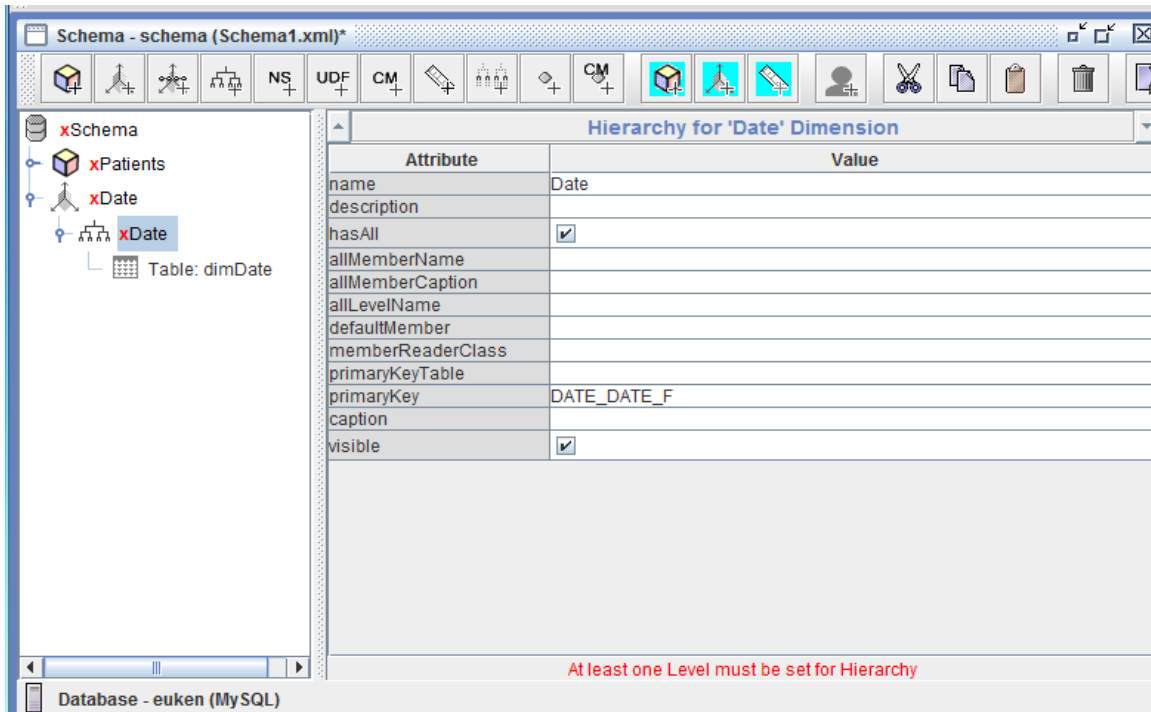


Ilustración 39: Añadir nombre dimensión, nivel y clave primaria

Por último, para terminar de completar una dimensión, la propia herramienta vuelve a solicitar añadir, en este caso jerarquías. Una vez añadidas, pulsando botón derecho sobre el nivel, se da un nombre a ésta y se añaden los datos que se ven en la Ilustración 40.

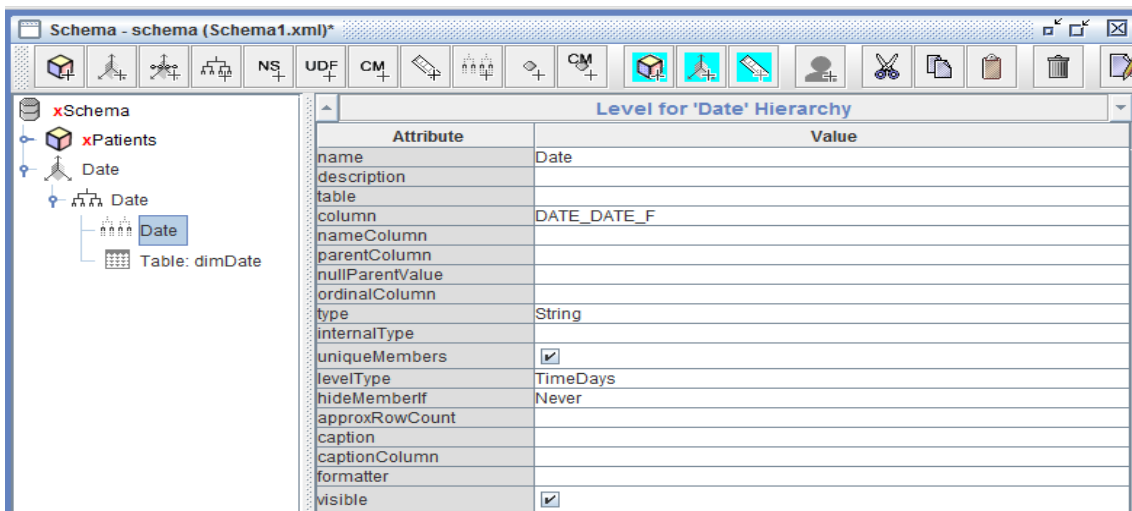


Ilustración 40: Añadir jerarquías

Cabe la posibilidad que la dimensión posea varias jerarquías, para ello se añaden las necesarias, también pulsando el botón derecho sobre el nivel. A continuación se puede ver un ejemplo de la dimensión “Location” con dos jerarquías, “Country” y “Area”.

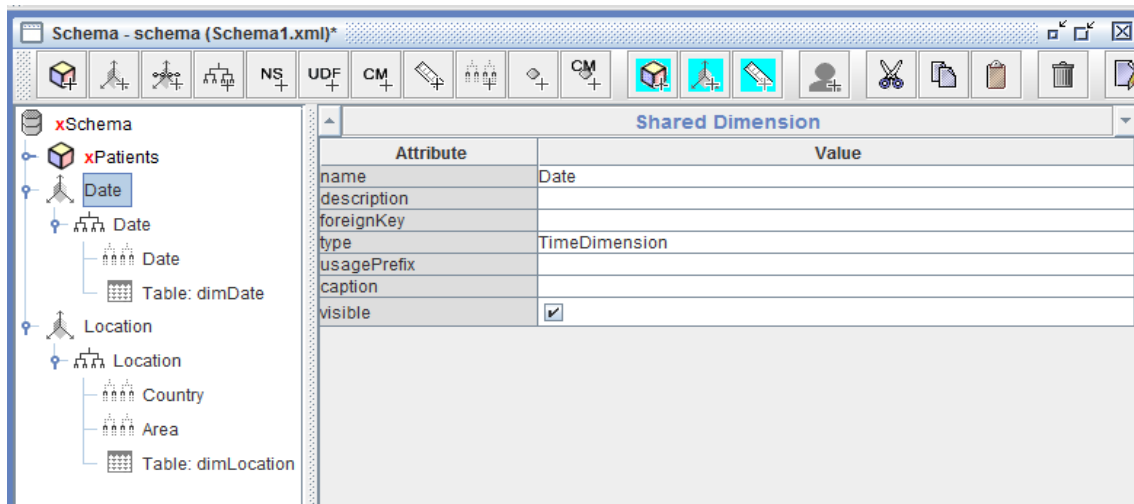


Ilustración 41: Dimensiones tiempo y localización

4.4. Dimension Usage

Una vez creadas todas las dimensiones que se van a utilizar en cualquier cubo del mismo esquema, ya sea “Patients”, “Diagnostic” o “Pharma”, mediante la opción “Add Dimension Usage” las dimensiones creadas en el esquema se pueden añadir a cualquier cubo que vaya a hacer uso de ella. En la siguiente ilustración se puede ver que, una vez seleccionado el cubo, se debe pulsar el botón indicado para añadir la “Dimension Usage”.

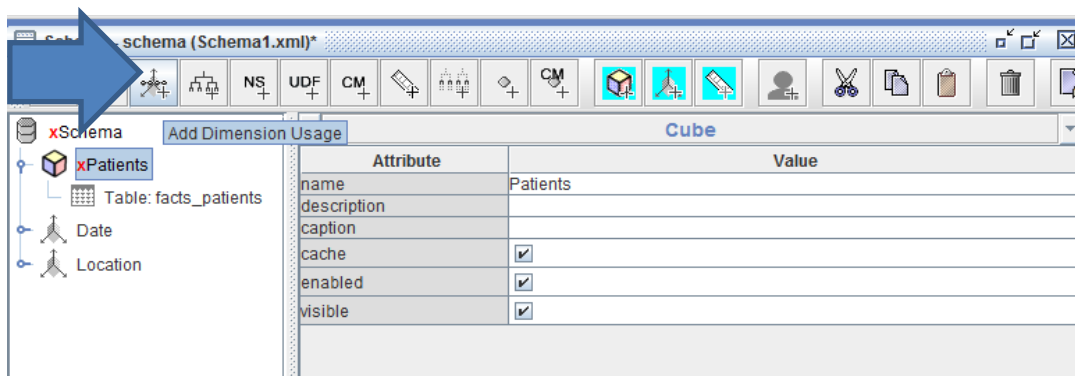


Ilustración 42: Añadir dimension usage

La “Dimension Usage” se añade al cubo; se nombra con el nombre de la dimensión; se añade la dimensión creada anteriormente en el campo *source*; y por último y más importante, se añade la clave extranjera de la tabla de hechos para relacionarlo correctamente. En la siguiente ilustración se puede observar qué campos es necesario rellenar.

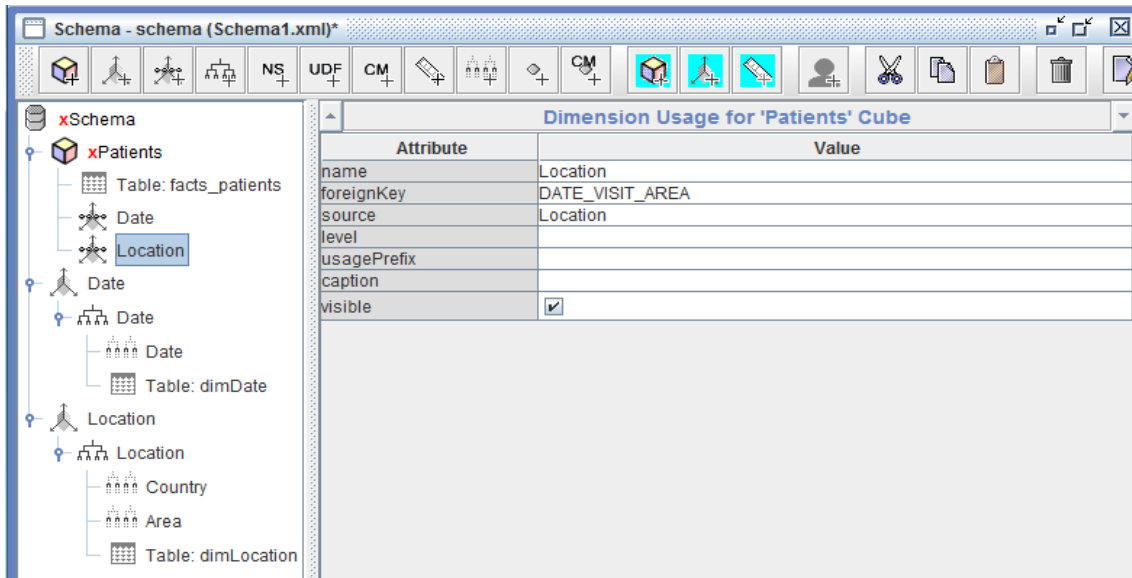


Ilustración 43: Dimension Usage

4.5. Añadir medidas

Por último, la propia herramienta solicita añadir medidas. Como su nombre indica, son los campos de la tabla de hechos que se quieren medir. Añadir una medida es muy sencillo, se debe pulsar el botón derecho sobre el cubo y se elige “Add Measure”.

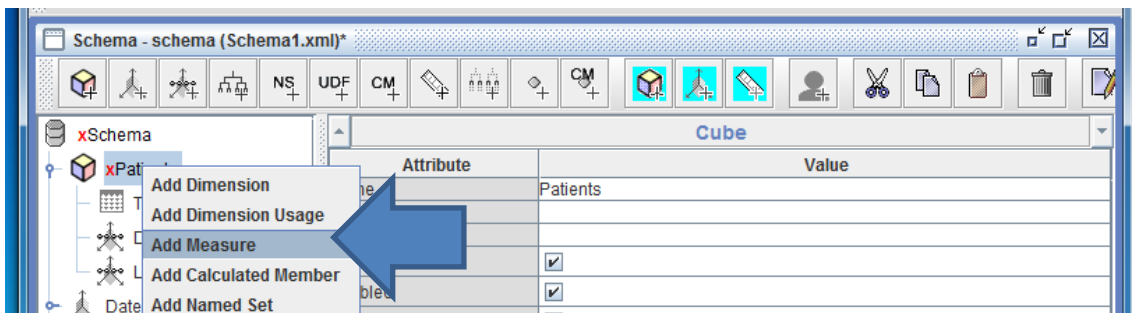


Ilustración 44: Añadir medidas

Una vez añadida la medida, se selecciona la columna a medir y qué se quiere hacer con ella. Entre otras opciones están la de contar, contar distintos, sacar la media, máximo o mínimo entre otros. En la siguiente ilustración se puede observar cómo se ha realizado lo anteriormente explicado.

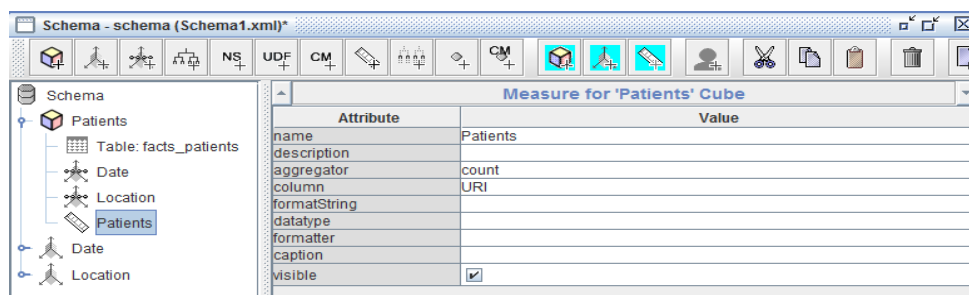


Ilustración 45: Medidas

A continuación se muestra cómo quedaría el esquema definitivo con todos los cubos, dimensiones y medidas creadas.

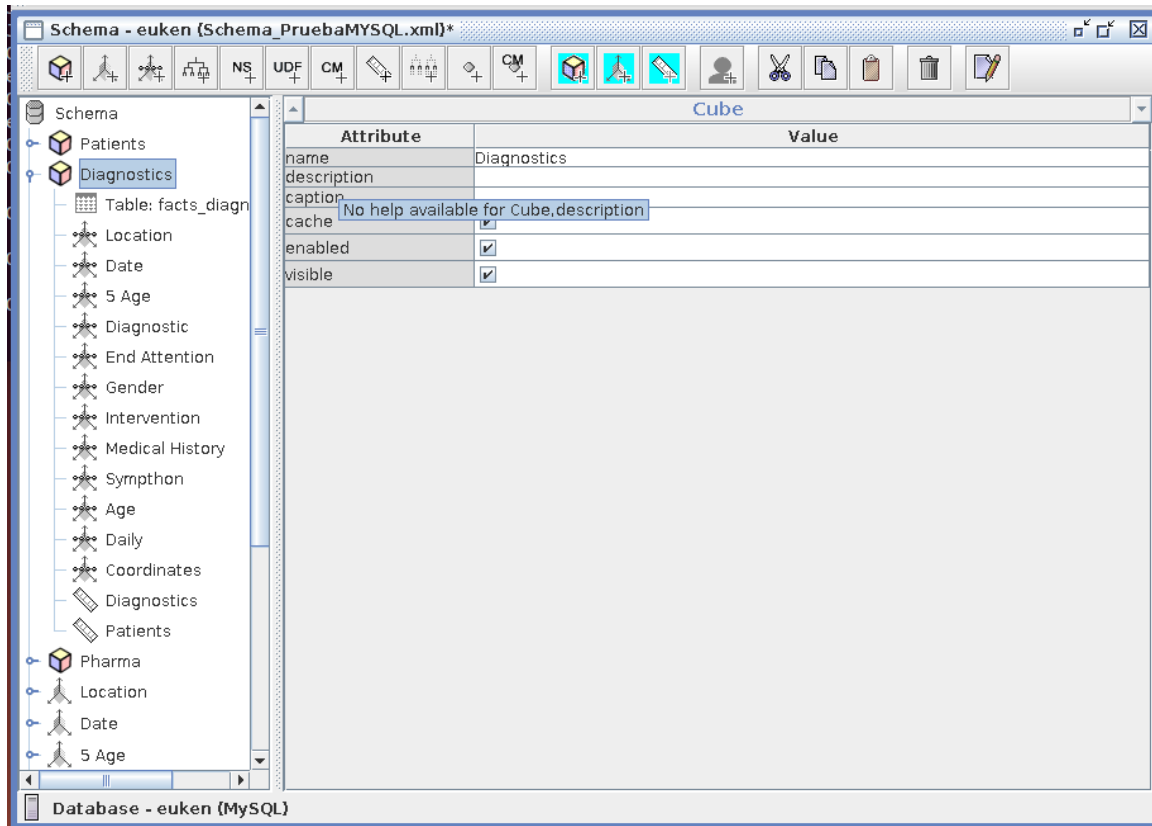


Ilustración 46: Vista general del cubo

4.6. Añadir Data Source y Publicar esquema

Para poder utilizar el esquema y cubo creados, es necesario publicarlo. Para realizar esto se debe tener un Data Source creado en la consola de Pentaho.

Se abre Pentaho como se ha explicado en prácticas anteriores mediante el archivo ejecutable “./start-pentaho.sh” en la ruta “/home/pentaho/biserver-ce”. Una vez abierto, se deben seguir los siguientes pasos:

1. Se pulsa sobre el botón Create New Data Source, tal y como se aprecia en la ilustración.

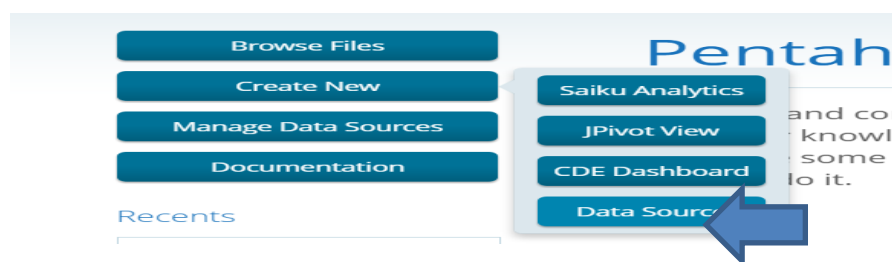


Ilustración 47: Manage Data Sources

- Como se muestra en la ilustración, seleccionamos como *Source Type*: *Database Table(s)*, posteriormente, y como se muestra a continuación, se añade una nueva conexión pulsado sobre “+”.

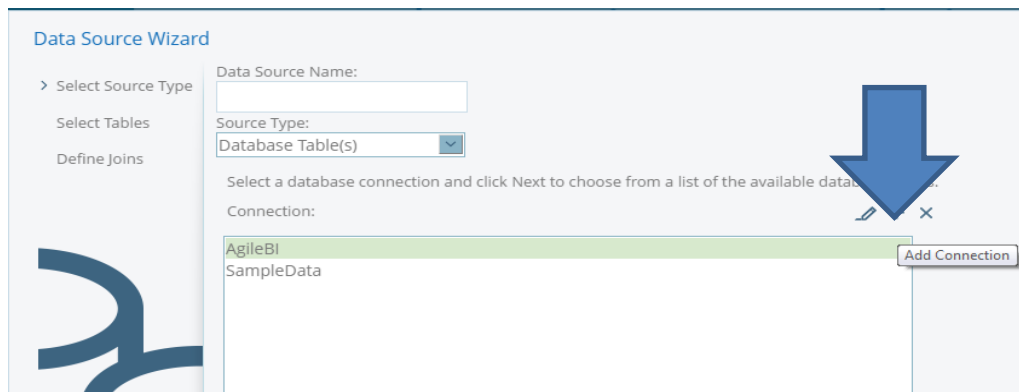


Ilustración 48: Add Connection

- Se rellenan los campos de la ventana de conexión a la base de datos.

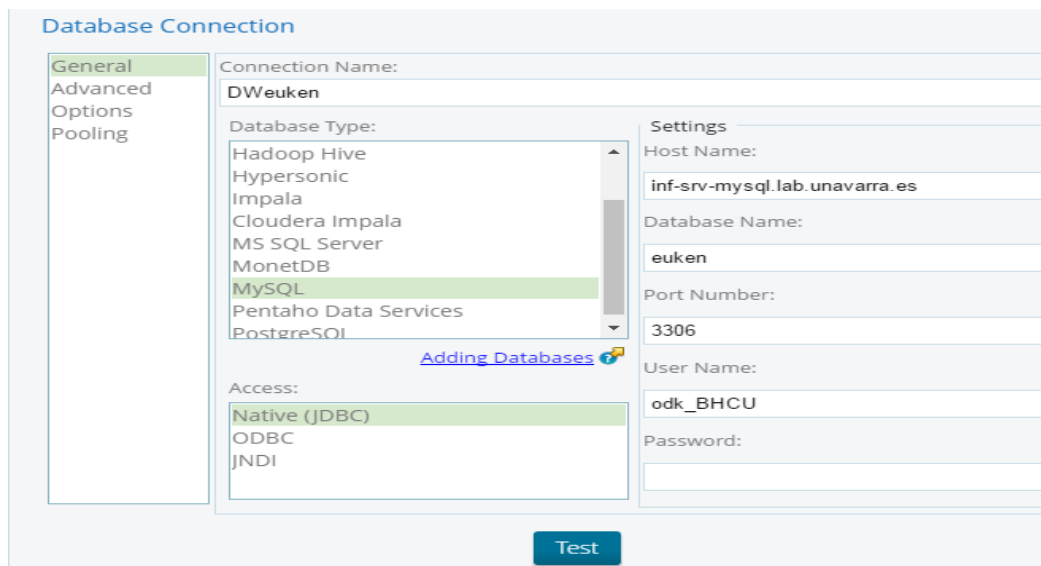


Ilustración 49: Credenciales DW

- Se comprueba la conexión realizando el test y se vuelve a la pantalla inicial. Una vez realizado todo esto ya se tiene la conexión correcta para poder publicar el cubo anteriormente creado.

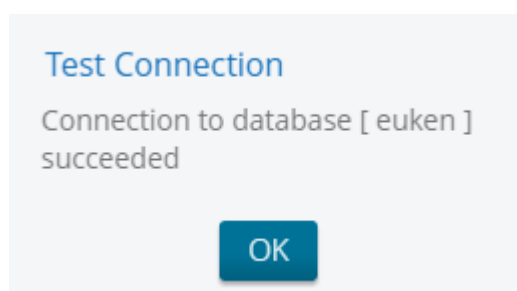


Ilustración 50: Test data Source

Una vez añadido correctamente el Data Source, se publica el esquema tal y como se indica a continuación.

En File, se selecciona Publish, y se rellena la ventana que aparece con la URL donde se tiene instalado Pentaho, las credenciales de acceso y el nombre del Data Source creado anteriormente. En las siguientes imágenes se pueden ver estos pasos.

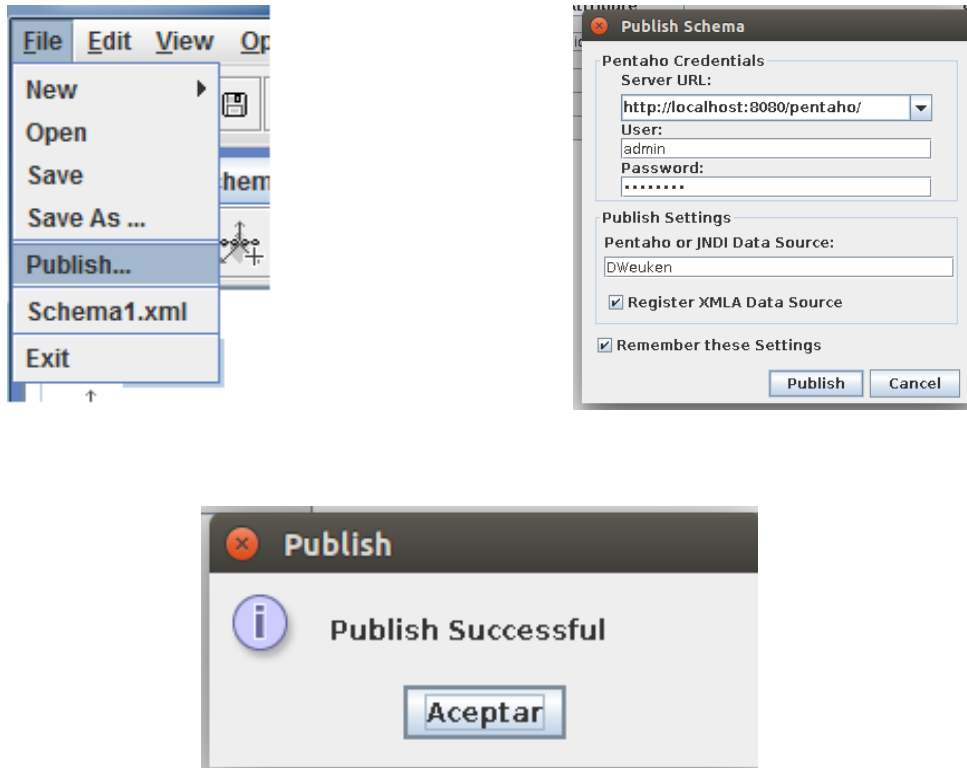


Ilustración 51: Publicar esquema

5. ITERACIÓN 4-CREACIÓN DE DASHBOARD

Para asegurarse de que se ha construido y publicado el cubo OLAP correctamente, este se puede visualizar de una forma muy sencilla y rápida con la herramienta descargada desde el MarketPlace Saiku Analytics.

Para ello se accede a dicha herramienta a través de la página de inicio de Pentaho, pulsando sobre “Crear nuevo” y “Saiku Analytics”.

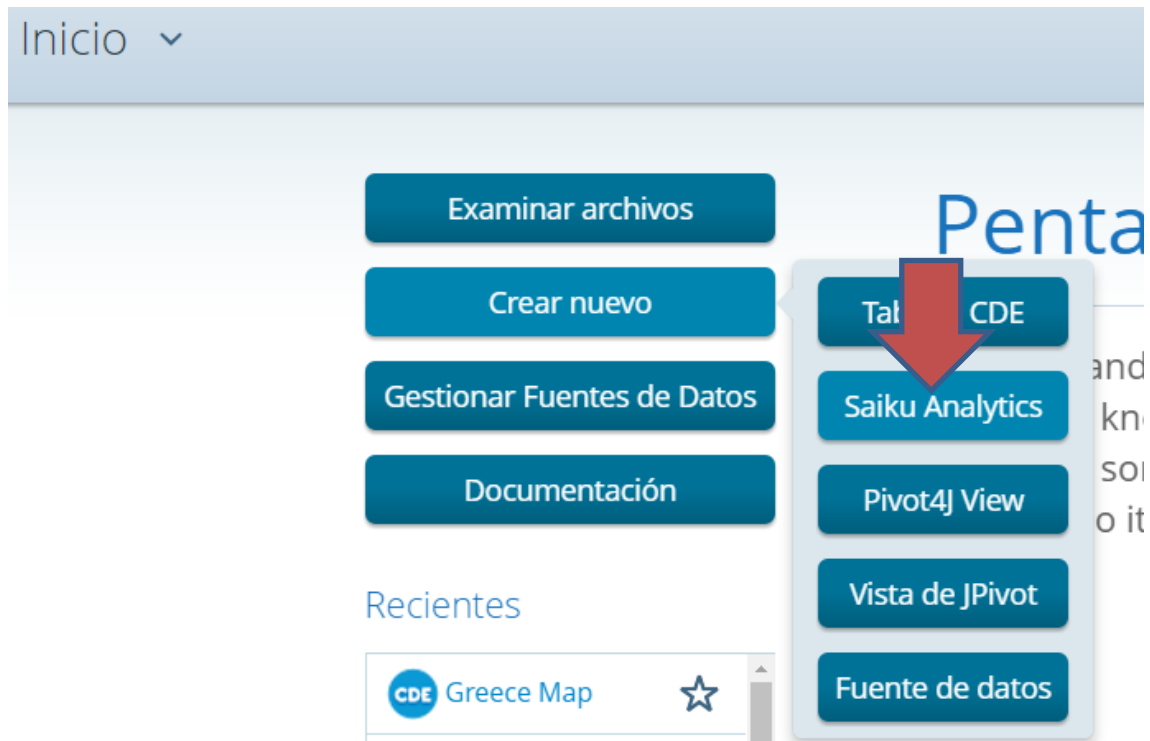


Ilustración 52: Abrir Saiku Analytics

Una vez abierto Saiku, creamos una nueva query y seleccionamos en cualquiera de los cubos que hemos publicado anteriormente. En la siguiente ilustración se puede observar que se selecciona el cubo “Diagnostics”.

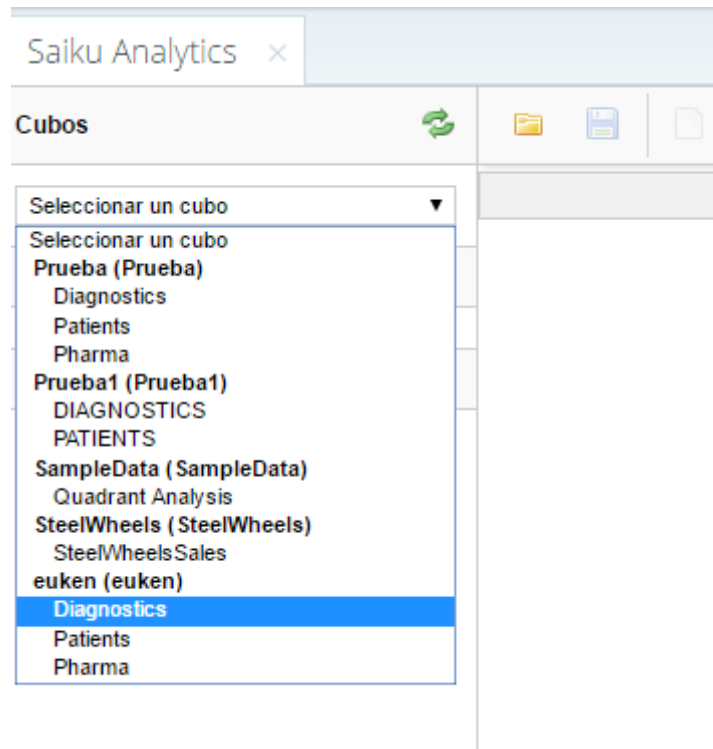


Ilustración 53: Seleccionar cubo

Seleccionado el cubo del que se quieren realizar las consultas de datos, esta herramienta proporciona una interfaz bastante intuitiva. Se pulsan o arrastran, las medidas o dimensiones a donde se quiera para observar los datos que devuelve la consulta.

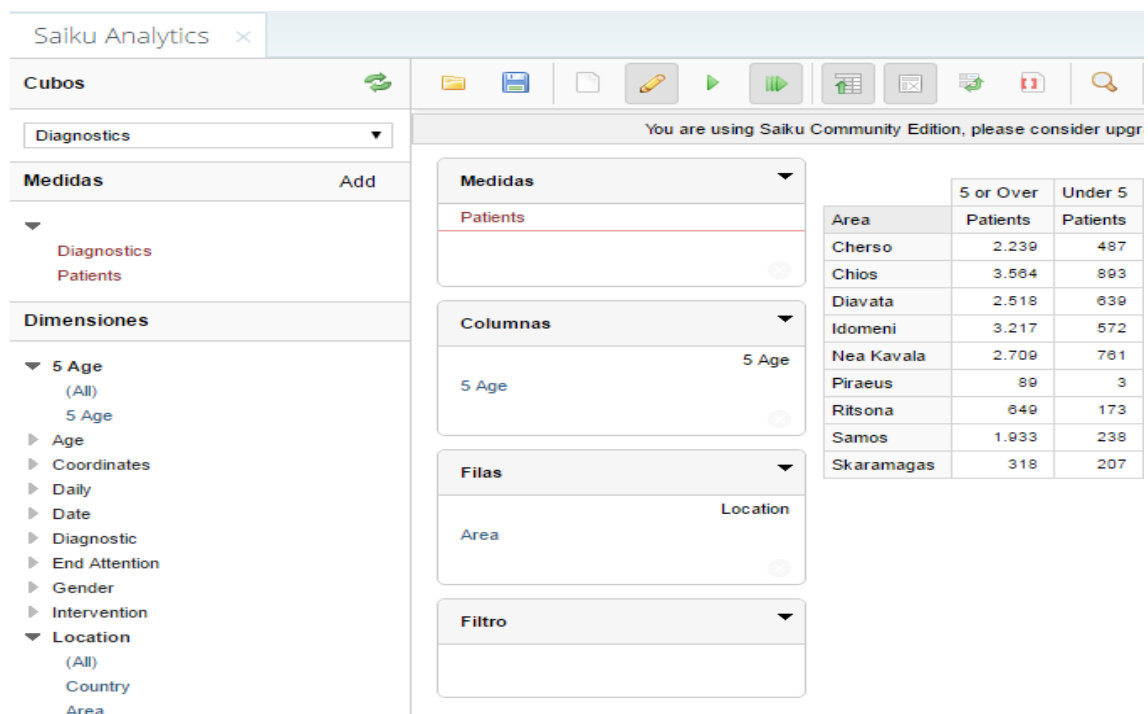


Ilustración 54: Salida de información

La utilización de esta herramienta es útil para la realización de futuros Dashboard, ya que se pueden obtener consultas MDX [20], lenguaje de consulta para bases de datos multidimensionales sobre cubos OLAP, se utiliza en Business Intelligence para generar reportes para la toma de decisiones basados en datos históricos, con la posibilidad de cambiar la estructura o rotación del cubo. Una consulta MDX es muy similar a una consulta SQL, devuelve un conjunto de celdas, que es resultado de tomar un subconjunto del cubo original. Sin embargo, contiene diferentes funciones y al utilizar varias dimensiones puede volverse bastante compleja. MDX utiliza en varias situaciones las jerarquías.

5.1. Construcción del Dashboard

La creación de Dashboards, se realiza a través de la herramienta CDE de Pentaho. Para mostrar y manipular los datos de una manera rápida e intuitiva, el cuadro de mando ha de ser fácil de manejar, puesto que se va a requerir de él en los momentos de toma de decisiones. Este permitirá analizar diferentes datos e informaciones de una manera ágil para tomar una decisión adecuada.

5.1.1. Creación de Dashboard

Una vez se han introducido las credenciales de acceso y aparece la pantalla de inicio de Pentaho, se pulsa sobre “Crea nuevo” y “CDE Dashboard”.



Ilustración 55: Abrir CDE Dashboard

Layout Structure

En el editor que se ve en la siguiente imagen, es donde se crea el Dashboard. Lo primero que se debe realizar es la estructura Layout, la cual indica la disposición visual de los

componentes del cuadro de mando. En la imagen también se puede observar un ejemplo de la estructura del Dashboard.

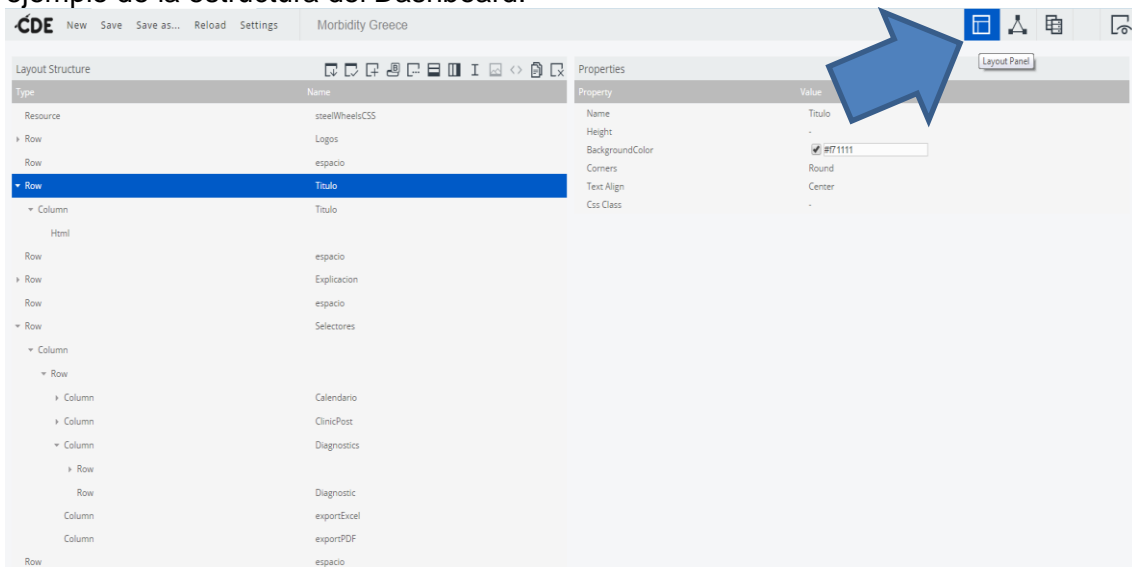


Ilustración 56: Estructura Layout

Datasource

En la pestaña Datasources se añaden las consultas MDX necesarias para la creación de tanto selectores, tablas o gráficos que se vayan a añadir al Dashboard.

Es aquí donde toca jugar con la herramienta Saku Analytics, que se usa para realizar dichas consultas y así poder recoger el código que estas proporcionan. A continuación se puede observar donde hay que pulsar en Saiku para poder copiar la consulta.

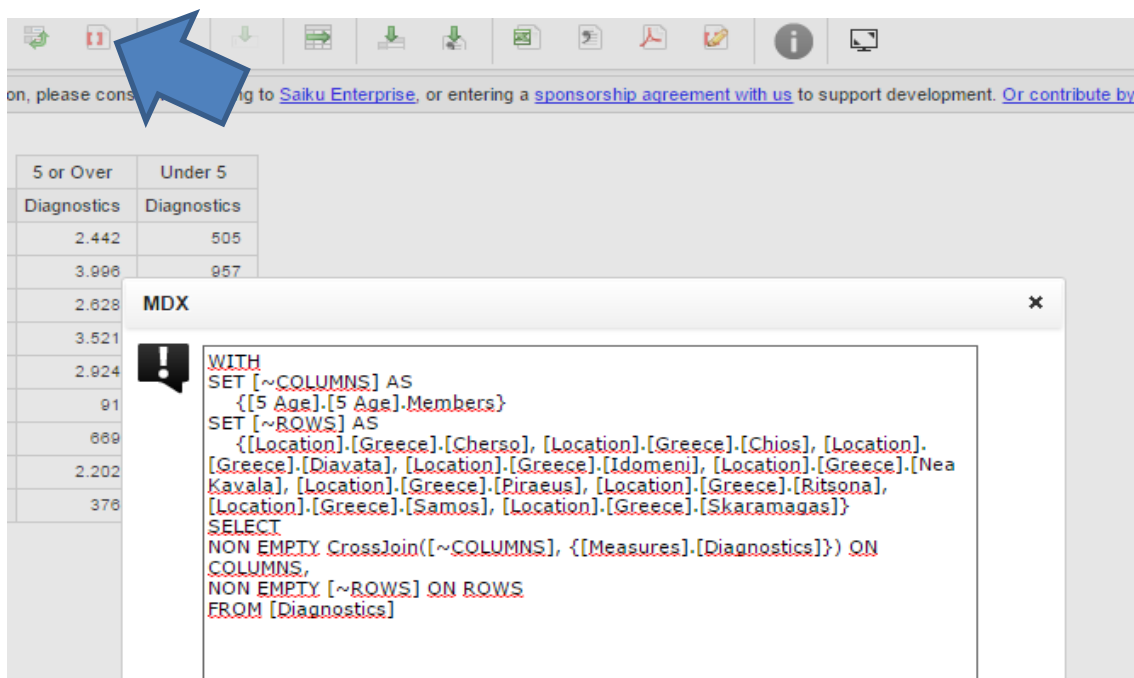


Ilustración 57: Query MDX

Una vez se tiene el código de la consulta, en la barra izquierda de la ventana Datasource, se añaden tantas “MDX Queries-mdx over mondrianjndi” como consultas se quieran añadir. El código se pega en el desplegable Query. Una vez insertado se debe especificar también:

- *El Jndi (java Naming and Directory Interface)*: Permite descubrir y buscar objetos y datos a través de un nombre.
- *Mondrian Schema*: Esquema sobre el que se ha realizado la consulta.

En la siguiente ilustración se puede observar dónde se ubican y qué añadir en los componentes anteriormente nombrados.

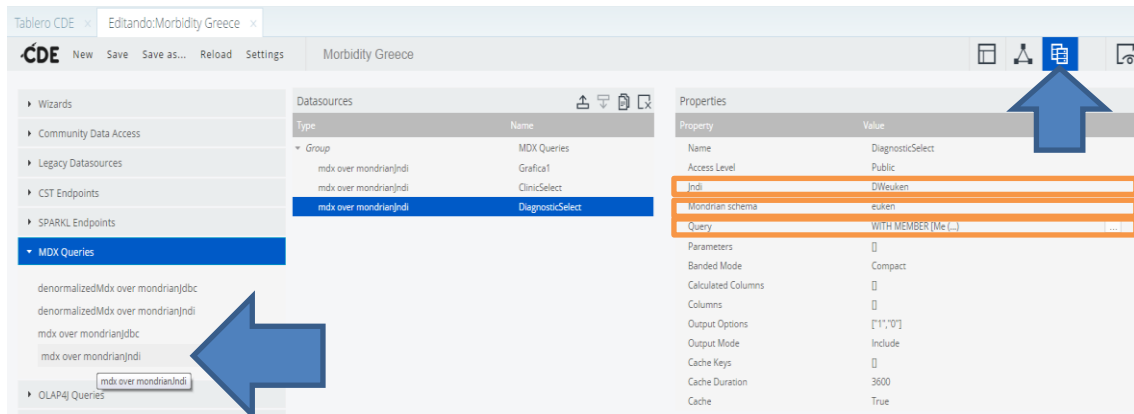


Ilustración 58: Datasource, Query, JNDI, Mondrian Schema

Es posible que las consultas que se realicen se quieran enlazar con distintos selectores. Para ello es necesario crear los parámetros que definirán la consulta MDX. Esto se realiza modificando las casillas que se ven recuadradas a continuación.

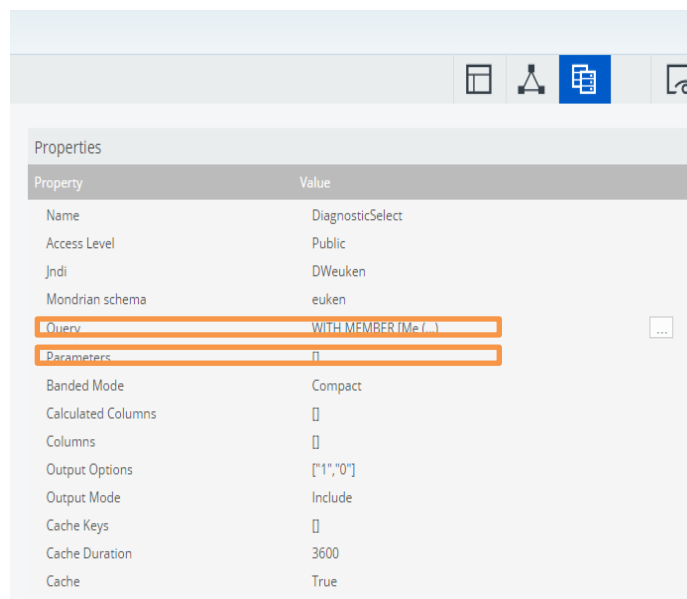


Ilustración 59: Query, Parameters

En primer lugar se modifica la Query, eliminando los parámetros estáticos que se han seleccionado como filtros en la herramienta Saiku Analytics. Estos se modifican por parámetros variables que se declaran en la fila seleccionada "Parameters".

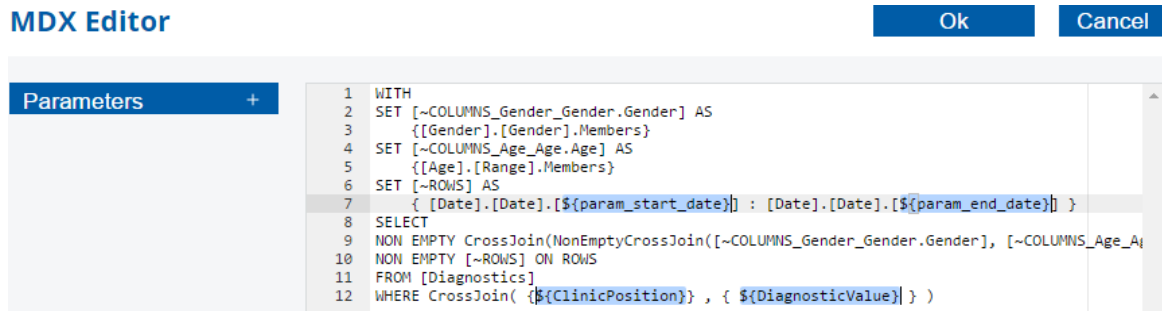


Ilustración 60: Query

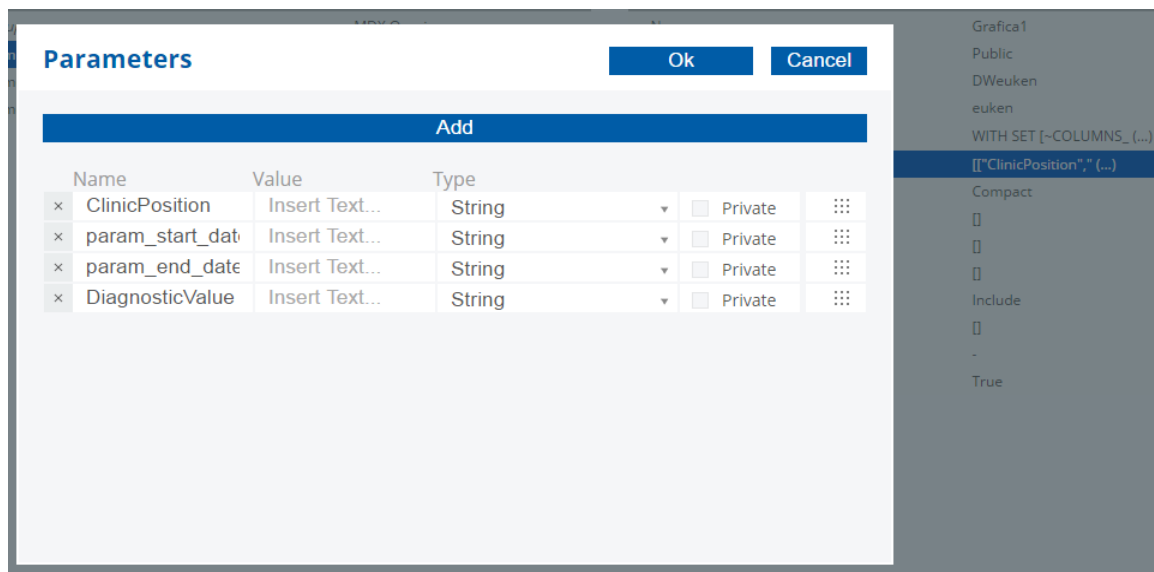


Ilustración 61: Parameters

Components

Una vez realizado correctamente lo anterior, queda añadir los componentes del Dashboard. Se añaden los selectores, gráficos e incluso mapas necesarios para la creación del cuadro de mando intuitivo.

Para la realización de este Dashboard, se ha utilizado un gráfico de barras (CCC Bar Chart), dos selectores comunes y dos selectores de tiempo, para establecer un rango de fecha con inicio y fin.

En primer lugar se debe añadir de la barra que aparece a la izquierda, en la sección "Generic", dos "Simple parameter" y dos "Date parameter". En la Ilustración 62 se muestra cómo se realiza.

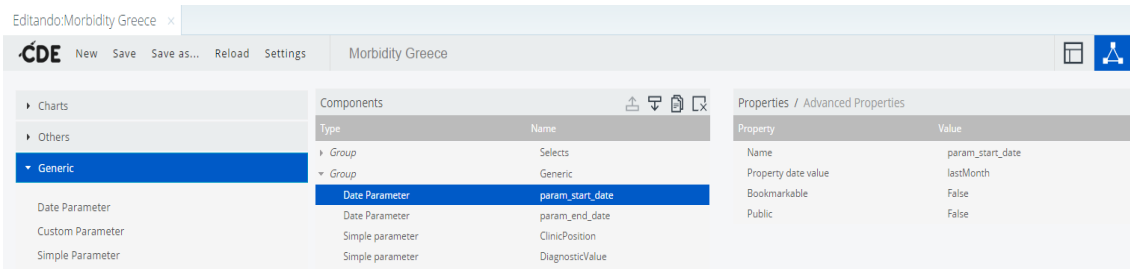


Ilustración 62: Generic Components

A continuación, es necesario añadir los componentes de los selectores de dichos parámetros. Para completar correctamente cada uno de ellos, deben introducirse correctamente los campos *“DataSource”*, lee la consulta anteriormente creada, *“Parameter”* y *“HtmlObject”*, lugar donde se va a ubicar en el Layout.

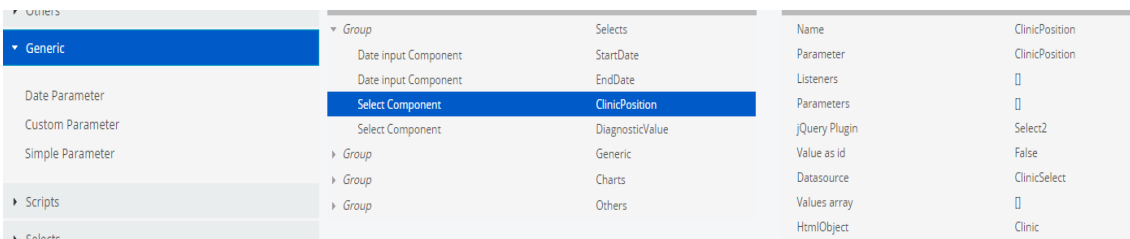


Ilustración 63: Select Components

Una vez se tienen los selectores correctamente definidos, se pasa a introducir el gráfico que se quiera añadir al Dashboard. En este caso se introduce un gráfico de barras, en el que se muestra la morbilidad en Grecia en un rango de fechas, dependiendo de la localización y según los diagnósticos.

Para ello se añade un nuevo CCC Bar Chart en el que realizan modificaciones similares a las anteriores, es decir, completar campos como *“DataSource”*, *“HtmlObject”*, *“Height”*, *“Width”*. Pero también hay que incluir correctamente *“Parameters”* y *“Listeners”*, donde se le especifica al gráfico cuándo debe actualizarse y qué debe cambiar.

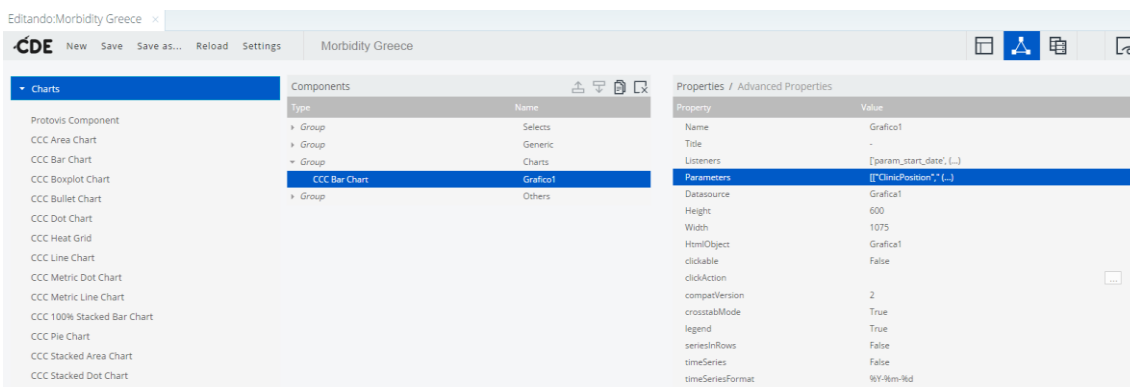


Ilustración 64: Charts Components

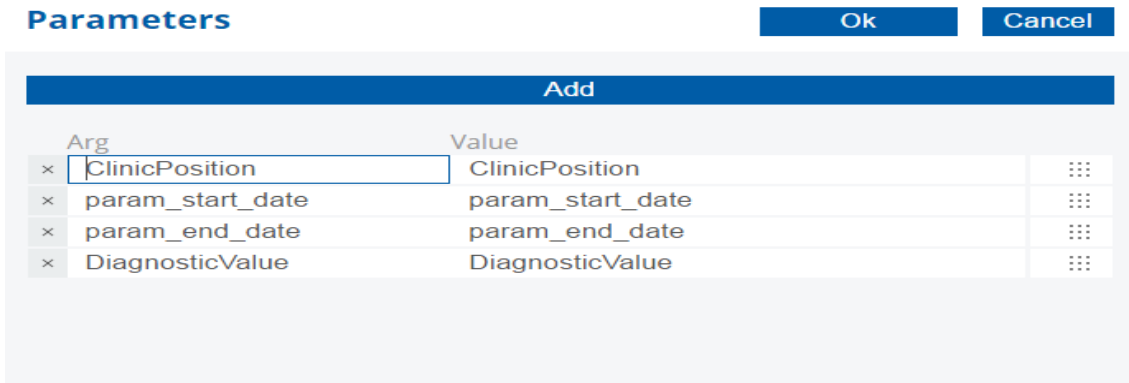



Ilustración 65: Listeners y Parameters

5.1.2. Visualización del Dashboard

Una realizado el Dashboard, es necesario guardarlo en la ruta que se quiera y tenerlo preparado para poder utilizarlo. Mientras se van generando las distintas partes del Dashboard, o una vez finalizado, se puede visualizar utilizando la pestaña  situada en la parte superior izquierda del editor.

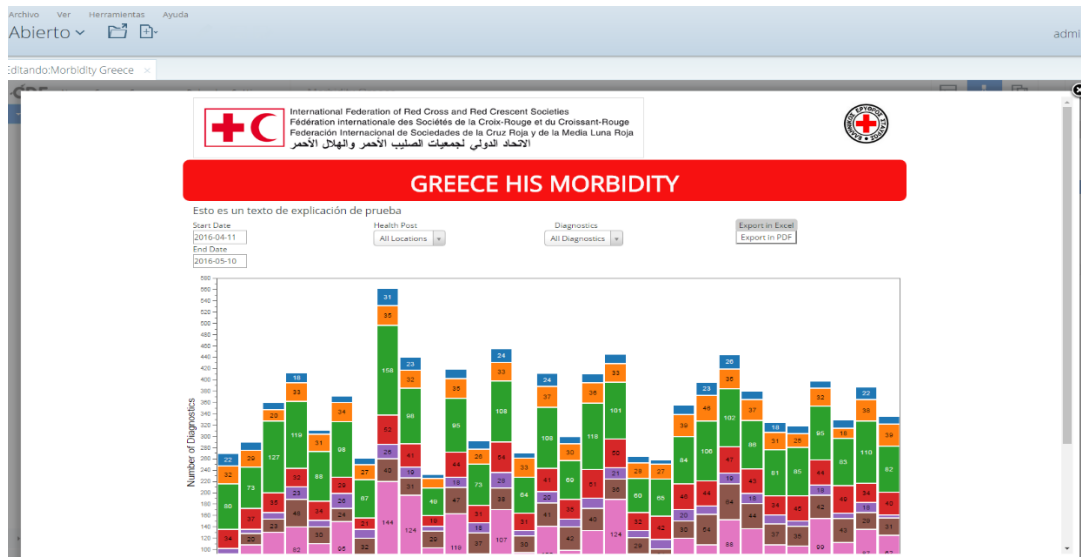


Ilustración 66: Visualización Dashboard

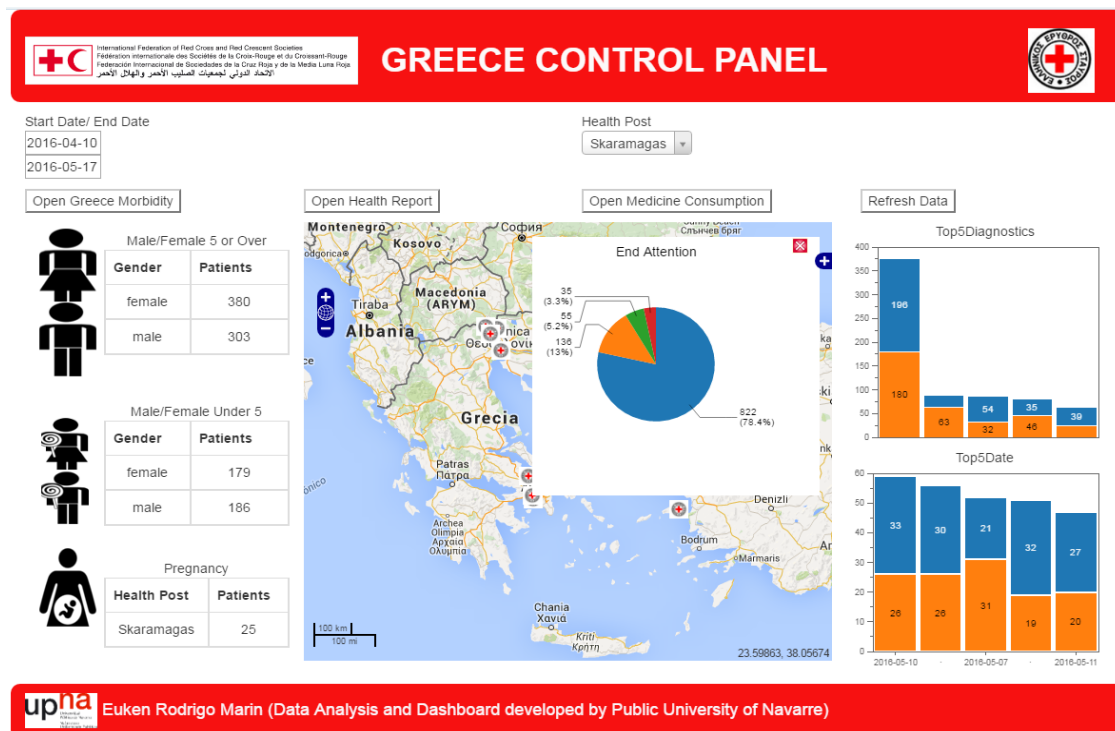
6. Dashboards

En este proyecto se han realizado más Dashboard que el explicado durante la memoria y se realizan todos de una manera similar. Lo que realmente interesa de la construcción de todo el Sistema BI, es la facilidad que se le da al usuario para poder visualizar dichos Dashboards. Es por eso que se ha instalado un plugin, concretamente *Community Startup Tabs*, que permite añadir como pantalla de inicio al hacer Login, diferentes Cuadros de Mandos según el rol de usuario.

Esto permite una comodidad al médico que quiere realizar un análisis de datos que rellenando sus credenciales tiene un fácil acceso.

6.1. Greece Control Panel

El primer Dashboard que se encuentra al iniciar sesión, es una pantalla con información general de la situación de refugiados en Grecia. En él se pueden ver cuántas personas han pasado por los distintos puestos repartidos por todo el país, disgregado en sexo y edad. También hay un control de las embarazadas, de los cinco diagnósticos que más se han dado y de los cinco días con más pacientes. Todo esto se puede filtrar por un rango de fecha y por los distintos puestos clínicos del país. En la parte central hay un mapa que realiza labores tanto de selector, cómo de indicador de las distintas clínicas mostrando datos sobre que se realiza al acabar las visitas. También posee enlaces a los distintos Dashboards que se han realizado, para facilitar la navegación entre ellos. Y por último, se ha añadido un botón que refresca la BD por si hay actualizaciones de última hora que se quieren visualizar. A continuación se puede ver una ilustración.



6.2. Greece Morbidity

Pulsando el botón “*Open Greece Morbidity*”, se abre otro Dashboard, concretamente el explicado durante la memoria, el cual, permite un análisis de la morbilidad en Grecia. Se puede observar un gráfico de barras donde indica por días, el número de diagnósticos desglosado por rango de fechas y sexo. El Dashboard posee también dos selectores que permiten filtrar por puesto clínico y diagnóstico concreto. Todo esto seleccionando un rango de fechas concreto. Se puede observar en la Ilustración 68.

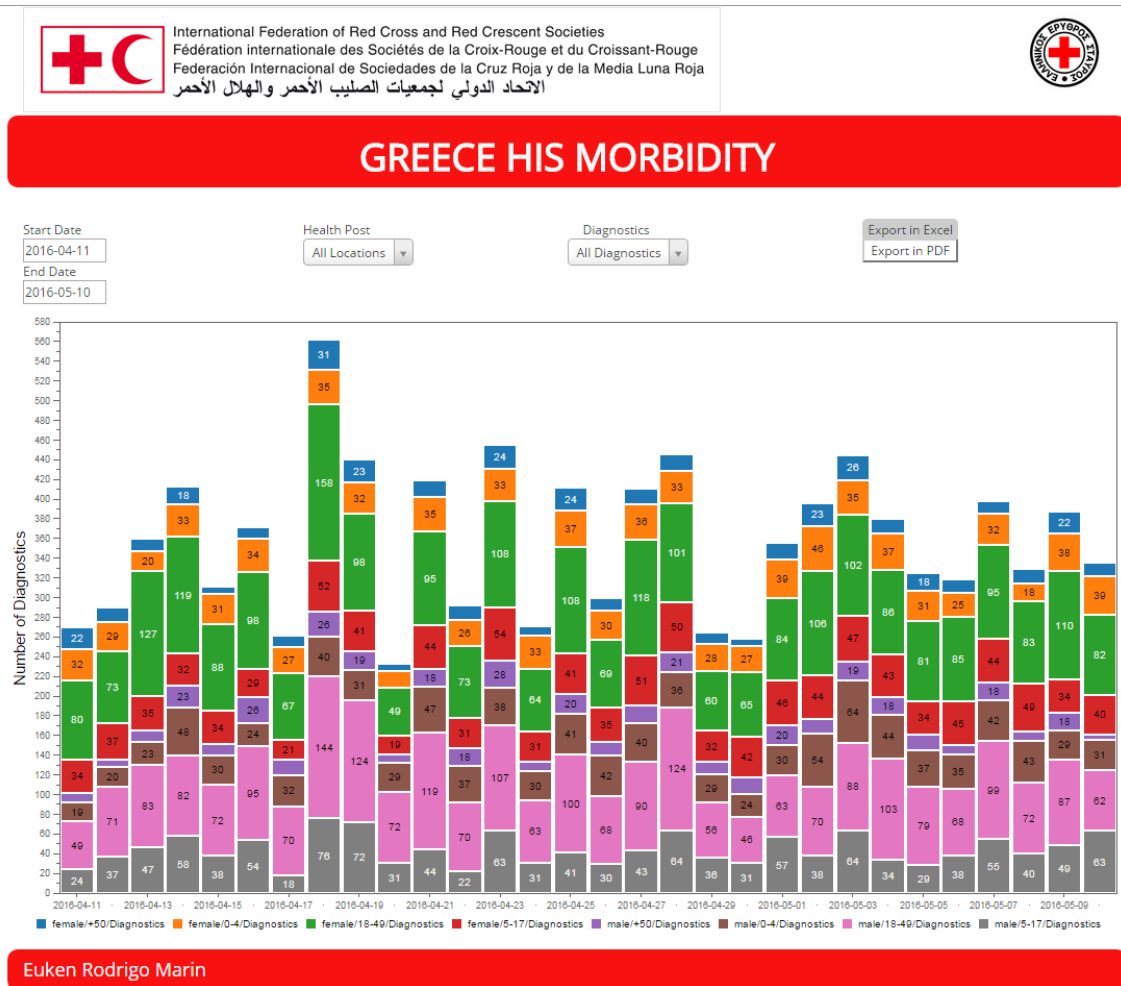
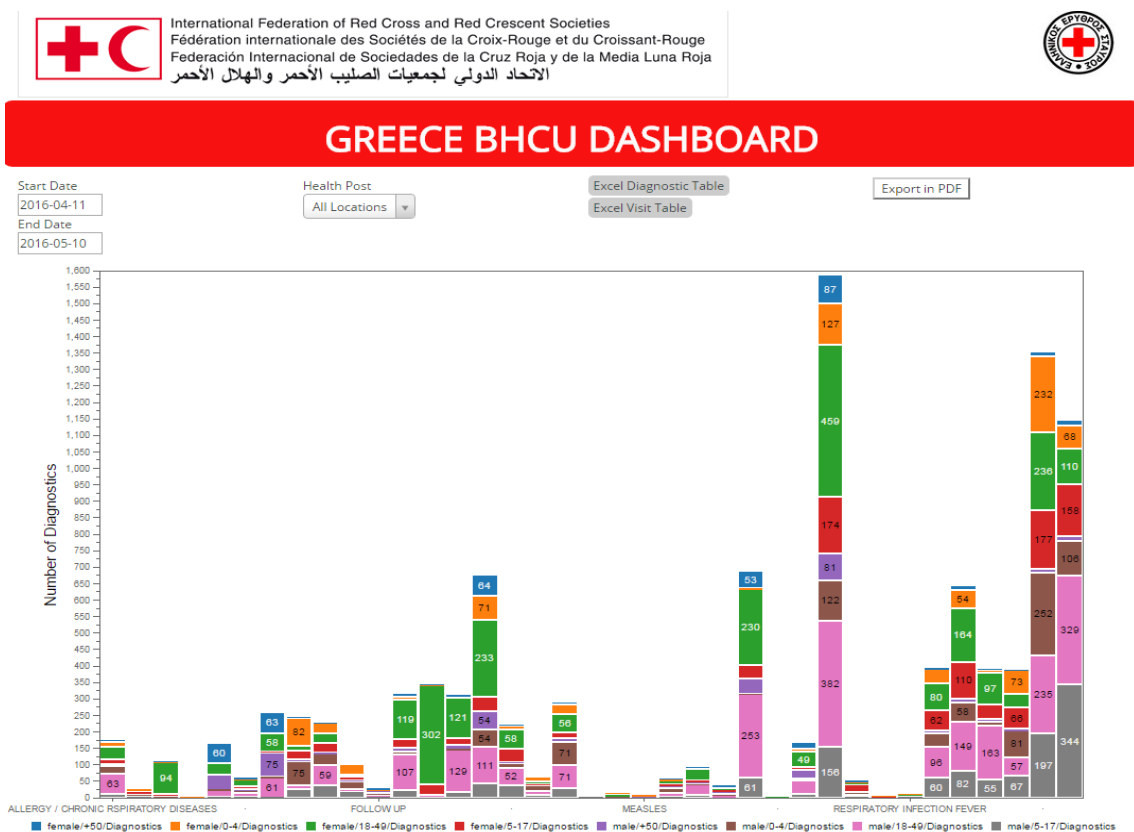


Ilustración 68: Greece Morbidity Dashboard

6.3. Health Report

Volviendo al Dashbord principal y pulsando sobre “Open Health Report”, se abre otro Cuadro de Mando en el que se puede observar en primer lugar, un gráfico de barras que indica el número de diagnósticos por cada diagnóstico disgregado en edad y sexo acompañado de una tabla que permite exportarla a Excel. También hay una tabla que indica el número de pacientes que terminan la visita, son destinados a un hospital, evacuados al hospital o les dan cita para volver en unos días. Esta información también está separada por rangos de edades y sexo. Por último se observa una pequeña tabla con información de los pacientes que han visitado las clínicas, o bien por la mañana, tarde o noche.



Total Numbers of Diagnostics

Show entries Search:

| Diagnostic | female/+50/Diagnostics | female/0-4/Diagnostics | female/18-49/Diagnostics | female/5-17/Diagnostics | male/+50/Diagnostics | male/0-4/Diagnostics | male/18-49/Diagnostics | male/5-17/Diagnostics |
|---|------------------------|------------------------|--------------------------|-------------------------|----------------------|----------------------|------------------------|-----------------------|
| ALLERGY / CHRONIC RESPIRATORY DISEASES | 9 | 14 | 38 | 16 | 5 | 23 | 63 | 11 |
| ALLERGY OTHER RESPIRATORY | | 3 | 5 | 3 | 2 | 6 | 5 | 2 |
| ANTE-NATAL | 2 | 3 | 94 | 11 | | 1 | 1 | |
| BLOODY DIARRHOEA | | 1 | 1 | | | | | |
| CARDIOVASCULAR DISEASE | 60 | | 35 | | 46 | 1 | 19 | 4 |
| DEHYDRATION / HYPOTENSION | 2 | 2 | 21 | 9 | 4 | 6 | 12 | 4 |
| DIABETES | 63 | | 58 | 1 | 75 | 1 | 61 | 1 |
| DIARRHOEA, not bloody | 6 | 82 | 16 | 27 | 4 | 75 | 11 | 26 |
| EYE INFECTIONS | 4 | 31 | 29 | 29 | 3 | 37 | 59 | 39 |
| FEVER OF UNKNOWN ORIGIN | | 32 | 5 | 12 | 1 | 25 | 7 | 19 |
| FOLLOW UP | 1 | 2 | 2 | | 6 | 8 | 6 | 2 |
| GASTRIC PROBLEMS | 11 | 10 | 119 | 27 | 12 | 8 | 107 | 24 |
| GYNAECOLOGICAL AND OBSTETRICAL AFFECTIONS | 4 | 1 | 302 | 31 | | | 9 | |
| HEADACHE | 12 | | 121 | 22 | 13 | 1 | 129 | 17 |

| | | | | | | | | |
|-----------------------|----|----|-----|-----|----|-----|-----|-----|
| WOUNDS, BURNS, TRAUMA | 17 | 68 | 110 | 158 | 14 | 106 | 329 | 344 |
|-----------------------|----|----|-----|-----|----|-----|-----|-----|

Showing 1 to 37 of 37 entries

Previous Next

Medical Visits Attending

Show entries Search:

| End Attention | female/+50/Diagnostics | female/0-4/Diagnostics | female/18-49/Diagnostics | female/5-17/Diagnostics | male/+50/Diagnostics | male/0-4/Diagnostics | male/18-49/Diagnostics | male/5-17/Diagnostics |
|---------------------------|------------------------|------------------------|--------------------------|-------------------------|----------------------|----------------------|------------------------|-----------------------|
| Concluded | 356 | 823 | 2317 | 1038 | 349 | 860 | 2155 | 1145 |
| Evacuated to hospital | 1 | | 6 | 1 | 1 | 1 | 5 | 2 |
| Further visit | 119 | 103 | 301 | 91 | 104 | 182 | 241 | 135 |
| Referal to Hospital | 10 | 15 | 82 | 19 | 21 | 21 | 67 | 26 |
| Referal to other resource | 4 | 6 | 26 | 22 | 3 | 5 | 28 | 19 |

Showing 1 to 5 of 5 entries

Previous Next

Patients Assisted by Shift

Show entries Search:

| Shift | Patients |
|---------|----------|
| Morning | 8093 |
| Evening | 1665 |
| Night | 158 |

Showing 1 to 3 of 3 entries

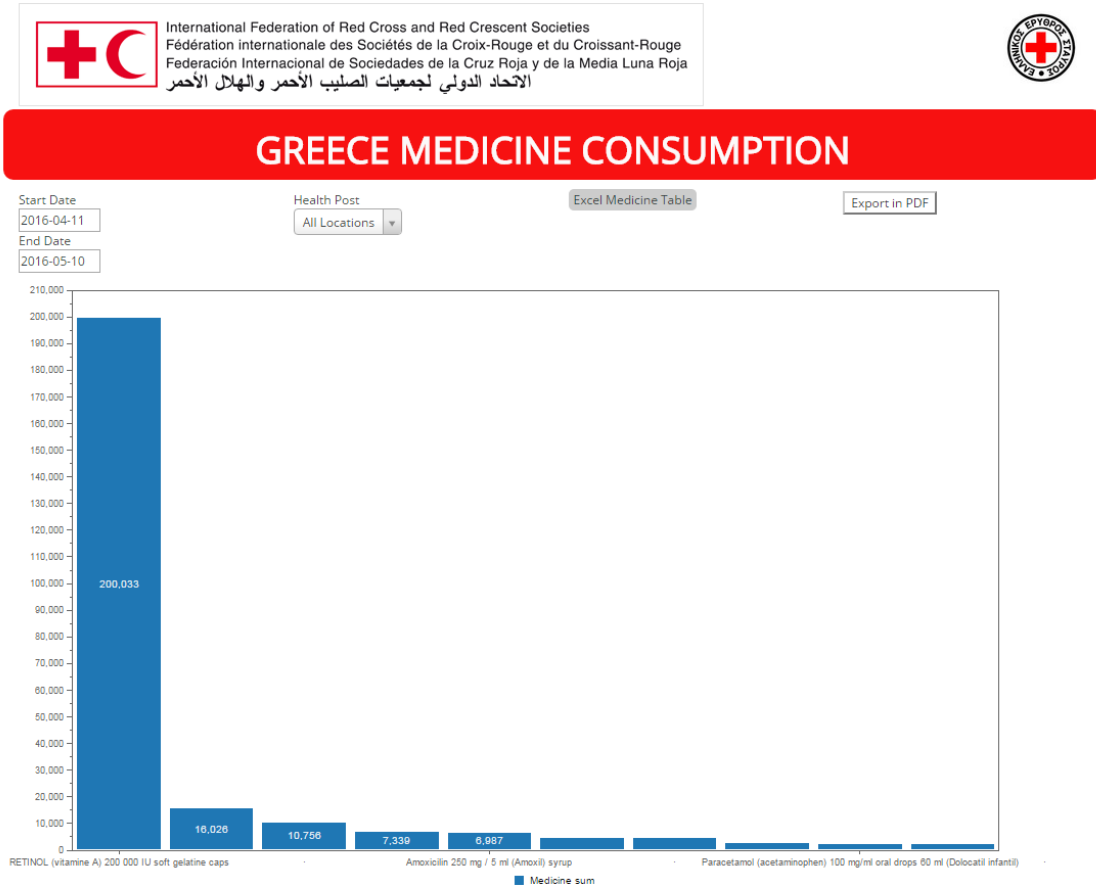
Previous Next

Euken Rodrigo Marin

Ilustración 69: Health Report Dashboard

6.4. Medicine Consumption

En este último Dashboard, el análisis que se realiza, es sobre la medicación que se dispensa en todos los puestos clínicos. En primer lugar se observa un gráfico de barras con los diez medicamentos que más se han dispensado. Junto a este gráfico hay una tabla con todos los medicamentos junto con la cantidad que se han dado a los refugiados. Este gráfico sirve sobre todo para controlar el stock de medicamentos en cada lugar, ya que gracias al selector “Health Post” el Dashboard varía según el puesto seleccionado.



Medicine Consumption

Show entries

Search:

| Pharma.Pharma | Medicine sum |
|---|--------------|
| ACICLOVIR 400 mg tab | 40 |
| AMITRIPTYLINE 25mg tab | 343 |
| AMOXYCILLIN + AC CLAVULANIQUE 400+57/5ml powd 70 ml | 24 |
| AMPICILLIN 2g im/iv powder vial | 21 |
| ASCORBIC ACID (vitamine C) 500 mg tab | 81 |
| ATENOLOL 50 mg tbl | 271 |
| AZITHROMYCIN 500 mg tabs/caps | 4 |
| Acetylsalicylic acid 100 mg tablets | 227 |
| Albendazole 400 mg tablets (Eskazole) | 8 |
| Aluminum oxide + magnesium hydroxide 600/300 mg (Maaloxc) chewable tablets | 123 |

| | |
|--|------|
| Ambroxol Hydrochloride (Mucosolvan) | 23 |
| Ambroxol oral solution 0.3% (MUCOSOLVAN) | 685 |
| Ambroxol tablet 30 mg (Mucosolvan) | 17 |
| Amlodipine 5 mg tablets | 69 |
| Amoxi/Clavulonic 500mg/125mg | 715 |
| Amoxicilin 250 mg / 5 ml (Amoxil) syrup | 6987 |
| Amoxicillin 500 mg | 5024 |
| Artificial Tears Artelac EDO | 55 |
| Azitromycine 500 mg tablets (Zitromax) | 61 |
| BENZYL BENZOATE 25% lotion 200 ml btl | 708 |
| BISACODYL 5 mg tab | 5 |
| BUTYLHYOSCINE (butylscopolamine) 20mg/ml 1 ml amp | 30 |
| Beclometasone 005 mg/dose inhaler 200 doses | 10 |
| Betametasona (Celecrem) | 14 |
| Betametasona + fusidic acid (Betafusin) | 7 |
| Bisoprolol 5 mg film coated tablets | 91 |
| Bromazepan (Lexatin) | 10 |
| Buphenine HCl Zinc (Opino-Gel) cream | 0 |
| CALCIUM GLUCONATE 100mg/ml 10 ml amp small packaging | 30 |
| CEFIXIME 100mg/5 ml suspension 100 ml | 10 |

Showing 1 to 30 of 145 entries

Previous 2 3 4 5 Next

Euken Rodrigo Marin

Ilustración 70: Medicine Consumption Dashboard

7. CONCLUSIONES

Business Intelligence es un área que últimamente ha tenido, y va a seguir teniendo, mucho auge en la mayoría de organizaciones o empresas que manejan una gran cantidad de datos. Esto se debe a las características y valor añadido que este tipo de herramientas son capaces de aportar a las compañías.

Una organización o empresa que dispone de un sistema de información de este tipo, perfectamente integrado y con una buena interpretación de las personas que deben tomar decisiones, adquiere una ventaja competitiva con respecto a las compañías que no lo poseen.

En el caso de este proyecto, al tratarse de un producto Open Source, su implantación no supone un desembolso económico al haber diversas opciones que permiten la implantación de la plataforma, en la que sólo es necesario disponer de hardware.

La realización de este sistema de información supone a Cruz Roja un salto de calidad a la hora del análisis que realizan. Además de una mayor facilidad a la hora de tomar decisiones, este sistema les permite ahorrar tiempo de trabajo intentando limpiar y descifrar la BD de los formularios. Ahora, gracias al proceso ETL se realiza todo mucho más rápido, y también se localizan con mayor facilidad posibles errores.

En el caso concreto de Pentaho 6.0, decir que ha resultado fácil e intuitivo realizar gran parte del proyecto, debido a las herramientas gráficas que posee, como pueden ser el *Spoon* o *Schema Workbench*.

Pese a ser un herramienta Open Source, los Dashboard que se han podido realizar, poseen una gran variedad de gráficos, selectores, tablas..., lo que permite facilitar la toma de decisiones a las personas de Cruz Roja que vayan a utilizar el sistema de información.

Gracias a la realización del proyecto, se ha visto que la BD que Pentaho trae instalada y configurada por defecto, no soporta hacer grandes consultas con la dimensión tiempo, especialmente añadiendo rangos de fecha. Esto es muy importante en el sistema de información realizado, ya que es imprescindible analizar la evolución en el tiempo, tanto de las personas que pasan por los puestos clínicos con distintos diagnósticos, como los medicamentos que se van dispensando.

Se han cumplido todos los objetivos fijados al comienzo del proyecto, atendiendo a peticiones o problemas que surgían según se avanzaba en la realización del mismo, permitiendo a Cruz Roja disponer del sistema de información que debandaban.

8. REFERENCIAS

- [1] <<**Cruz Roja**>> [En línea]. Available: <http://www.cruzroja.es/> . [Mayo 2016].
- [2] <<**ODK**>> [En línea]. Available: <https://opendatakit.org/>. [Mayo 2016].
- [3] <<**Microsoft SQL Server BI**>> [En línea]. Available: <https://www.microsoft.com/es-es/server-cloud/products/sql-server-editions/sql-server-business-intelligence.aspx>. [Mayo 2016].
- [4] <<**Oracle BI**>> [En línea]. Available: <https://www.oracle.com/es/solutions/business-analytics/business-intelligence/index.html>. [Mayo 2016].
- [5] <<**Qlik**>> [En línea]. Available: <http://global.qlik.com/es>. [Mayo 2016].
- [6] <<**SAP Business Intelligence Platform**>> [En línea]. Available: <http://go.sap.com/spain/solution/platform-technology/analytics/business-intelligence-bi.html>. [Mayo 2016].
- [7] <<**Tableau Business Intelligence**>> [En línea]. Available: <http://www.tableau.com/>. [Mayo 2016].
- [8] <<**SAP Lumira**>> [En línea]. Available: <http://saplumira.com/>. [Mayo 2016].
- [9] <<**Pentaho Enterprise Edition**>> [En línea]. Available: <http://www.pentaho.com/>. [Mayo 2016].
- [10] <<**Pentaho Community Edition**>> [En línea]. Available: <http://community.pentaho.com/>. [Mayo 2016].
- [11] <<**Palo**>> [En línea]. Available: <http://sourceforge.net/projects/palo/>. [Mayo 2016].
- [12] <<**SpagoBI**>> [En línea]. Available: <http://www.spagobi.org/>. [Mayo 2016].
- [13] <<**Java**>> [En línea]. Available: <http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>. [Mayo 2016].
- [14] <<**Tomcat**>> [En línea]. Available: <http://tomcat.apache.org/>. [Mayo 2016].
- [15] <<**Trello**>> [En línea]. Available: <https://trello.com/>. [Mayo 2016].
- [16] <<**Proxmox**>> [En línea]. Available: <https://www.proxmox.com/en/>. [Mayo 2016].
- [17] <<**MySQL Connector Java**>> [En línea]. Available: <http://dev.mysql.com/downloads/connector/j/>. [Mayo 2016].
- [18] <<**HSQLDB**>> [En línea]. Available: <http://hsqldb.org/>. [Mayo 2016].

- [19] <<Putty>> [En línea]. Available: <http://www.putty.org/>. [Mayo 2016].
- [20] **George Spofford, Sivakumar Harinath, Christopher Webb, Dylan hai Huang, Francesco Civardi** (2006). MDX Solutions with Microsoft SQL Server Analysis Services 2005 and Hyperion Essbase (2nd edition). Wiley.
- Microsoft** (2012). Multidimensional Expressions (MDX) Reference (SQL Server 2012 Books Online). Microsoft Corporation.

9. ANEXO

Anexo 1- Carta de Cruz Roja solicitando la realización del proyecto

Raúl Ecay Torres Ecay Torres
Dister Management Unit
International Department
Spanish Red Cross
Reina Victoria, 48
Madrid - 28008

18/01/2016

Jose Javier Astrain Escola
Profesor Universitario -- Area de Leguajes y sistemas informáticos
Universidad Pública de Navarra
Campus Arrosadía s/n
Pamplona - 31006

Estimado Profesor Jose Javier Astrain Escola:

Le escribo en respuesta a su ofrecimiento de colaboración por parte de un alumno de Grado de su universidad en relación a la crisis de migrantes solicitantes de asilo en Grecia motivada por los conflictos armados de oriente medio.

En estos momentos estamos en proceso de implantación de un sistema básico de historia clínica y un monitor en línea de salud pública.

Si bien en estos momentos el seguimiento se hace con herramientas básicas, el posible bloqueo de fronteras podría derivar en una situación de asentamientos semi-permanentes y población retenida que requerirán de herramientas de análisis y seguimiento con el fin de monitorizar los indicadores de salud de esta población y la detección temprana de posibles focos epidémicos.

Nuestra necesidad a corto plazo es la explotación de la base de la base datos actual generada por el ecosistema OpenDataKit y proporcionar herramientas de seguimiento y análisis que nos permitir mejorar la atención sanitaria a esta población a la vez que poder anticiparnos a sus necesidades en el escenario de bloqueo de fronteras.

Si entrara dentro de sus posibilidades, le ruego nos contacte a la mayor brevedad.

Le agradecemos su ofrecimiento que si se llegara a materializar nos será de gran ayuda para mejorar la atención sanitaria a la población forzada a migrar por diferentes conflictos armados.

Atentamente,



Raúl Ecay Torres