

HCM - UNIT4 Business Analytics

Xavier Martin Morales

Resumen—Hoy en día las empresas contienen gran cantidad de información sobre sus empleados. Esta información se ha de poder tratar y analizar de forma eficiente. De esta necesidad surge el concepto Business Intelligence, dónde se quiere proporcionar a la empresa o negocio herramientas para facilitar y mejorar el análisis de estos datos por parte de expertos sin necesidad que éstos utilicen más tiempo del estrictamente necesario. En éste proyecto se especifica el proceso seguido para realizar una aplicación de Business Intelligence, enfocada a la gestión de recursos humanos, dentro de la empresa UNIT4 adaptándose a su entorno de datos y ERP Ekon.

Palabras Clave— Inteligencia de Negocio (BI), Gestión de Recursos Humanos, ETL, OLAP, Data WareHouse, Data Mart, Business Analytic (BA), Reportes

Abstract—Nowadays companies have a huge amount of information about their employees. This information has to be processed and analyzed efficiently. From this necessity arises the concept of Business Intelligence, which aims to provide to the company or business tools to make the data analysis by experts better and easier, without these professionals to spend a lot of time to obtain it. In this project the process followed to make a Business Intelligence application is described, focused on human capital management (HCM), in the company UNIT4 using their data and ERP Ekon

IndexTerms— Business Intelligence (BI), Human Capital Management (HCM), ETL, OLAP, Data WareHouse, Data Mart, Business Analytic (BA), Reporting

1 INTRODUCCIÓN

Business Intelligence (BI)[1][2]proporciona a las empresas un conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa.

Las razones que justifican el uso de BI por parte de una empresa son:

- Proporciona una manera rápida y efectiva de recopilar, abstraer, presentar, formatear y distribuir la información de sus fuentes de datos corporativos.
- Permite a profesionales de la empresa visualizar y analizar datos precisos sobre las actividades fundamentales del negocio para mejorar la toma de decisiones estratégica.

En la Tabla I se muestra algunas de las ventajas que proporciona la gestión con BI sobre la gestión tradicional en las empresas.

TABLA I
COMPARATIVA DE GESTIÓN DE DATOS PARA EL ANÁLISIS

Gestión Tradicional	Gestión con BI
Múltiple captura de datos	Captura única de datos
Visión fragmentada	Visión única
Ciclos de planificación largos	Ciclos de planificación cortos
Falta de control	Control Centralizado

Este proyecto se realiza en un ámbito de prácticas universitarias (UAB) en empresa (UNIT4). UNIT4 es una compañía global de software de gestión de empresas que crea, ofrece y soporta en todo el mundo software y servicios adaptables, con el fin de ayudar a las organizaciones dinámicas a gestionar de manera eficiente sus negocios. La empresa dispone de un entorno ERP propio que comercializa (Ekon) para la gestión de datos. Específicamente, para el ámbito de gestión de recursos humanos, se trabaja con los datos proporcionados por Ekon Laboro.

Actualmente el análisis de los datos en Ekon Laboro se realiza mediante la exportación de éstos a hojas Excel. Se quiere entonces desarrollar una herramienta que permita integrar y realizar análisis sobre los datos que se proporcionan en Ekon Laboro. Se ha utilizado la herramienta UNIT4 Business Analytics, propia de la empresa, para el desarrollo de la estructura de datos y interfaz de usuario. Para el proceso de obtención y carga de datos ya existentes en el entorno Ekon se ha utilizado la versión gratuita de la herramienta Talend OS. V5.4.1.

En este artículo se hace primero una visión general sobre el estado del arte actual y conceptos básicos de BI. Se especifican los objetivos y requerimientos del proyecto así como los pasos realizados y planificación para cumplirlos. Finalmente se muestran los resultados obtenidos y comprobaciones sobre éstos, las conclusiones y una sección para posibles mejoras y acciones futuras.

- E-mail de contacto: xevimartin@gmail.com
- Mención realizada: Enginyeria del Software.
- Trabajo tutorizado por: Josep Lladós Canet (departament de Ciències de la Comunicació)
- Curso 2013/14

2. ESTADO DEL ARTE

El término Business Intelligence aparece por primera vez en un artículo publicado en 1958 por Hans Peter Luhn, un investigador de IBM en aquella época. Aun así no fue hasta mediado de los 90s que BI no empezó a emerger dentro de paquetes de software que ofrecían paquetes de gestión de rendimiento, planificación, reportes, análisis, etc.

En la figura 1 se muestran los diferentes productos más destacados en 2014. El "Magic Quadrant" de Gartner clasifica los proveedores de tecnología según dos criterios (ejes del gráfico):

La habilidad de ejecución (*Ability to execute*), contemplaría entre otros las funcionalidades y/o sector del mercado que cubre en la actualidad.

La Visión de futuro (*Completeness of vision*), comprensión del mercado y hacia donde evoluciona éste.

Divide entonces, según los criterios anteriores, en cuatro agrupaciones:

1) *"Leaders"*: Cubren bien las necesidades del sector y están bien posicionados para las necesidades futuras.

2) *"Visionaries"*: Entienden hacia dónde evoluciona el mercado pero aún no cubren las necesidades actuales del sector.

3) *"Niche Players"*: Se centran satisfactoriamente en un pequeño sector del mercado, o no tienen un foco claro y no son innovadores

4) *"Challengers"*: Cubren bien las necesidades actuales del sector o dominan ampliamente el mercado, pero no demuestran un entendimiento sobre hacia dónde evoluciona éste.



Fig. 1Gartner[3] 2014 MagicQuadrant mercado BI

Dentro de estos hay cinco grandes empresas (SAP, ORACLE, IBM, SAS Institute, Microsoft) que lideran el mercado (con cerca de un 70% de cuota de mercado [4] en 2013).

BI es un producto/concepto en expansión. Aún hoy en día son muchas las empresas que no utilizan los datos que disponen para generar conocimiento y mejorar sus servicios. Ya no es suficiente con tener la información, sino poder interpretarla y analizarla. Tanto medianas como grandes empresas buscan cada vez más soluciones BI que les permitan obtener una rentabilidad a los datos que tienen para poder usarlos como conocimiento.

3. CONCEPTOS BUSINESS INTELLIGENCE

3.1 Arquitectura BI

El concepto de BI incluye tanto aplicaciones, infraestructura, herramientas y estrategias que permiten realizar un análisis de la información.

La figura 2 muestra el esquema básico que sigue BI



Fig2 Esquema workflow BI

3.2 OLAP (On-Line Analytical Processing)

Arquitectura de almacen de datos diseñada para mantener datos extraídos de sistemas de transacciones, operacionales y fuentes externas. Este conjunto de datos está orientado hacia un determinado ámbito, integrado, no volátil y variable en el tiempo que ayuda a la toma de decisiones en la entidad que se utiliza.

3.2 Cubo

Un Cubo es una estructura multidimensional formada por:

1) **Tabla de Hechos (FactTable)**: Tabla central que representa los datos numéricos en el contexto que se analizó. Contiene los datos a mostrar y analizar. Contiene los registros asociados a la actividad que se va a realizar

2) Una o más **medidas (Measures)**: Las medidas son las magnitudes o cantidades almacenadas en la BBDD y que son objetos de análisis complejo o inspección.

3) Una o más **dimensiones (Dimensions)**: Representa una entidad de negocio. Dan contexto a los datos numéricos de la *FactTable*. Describen quien, qué, cuando dónde, cómo, etc. Presentan los datos organizados en jerarquías.

En la figura 3 se muestra una representación simple de la estructura de un cubo, almacenando las ventas por regiones, según períodos de tiempo. Un cubo no está limitado a 3 dimensiones, sino que puede tener todas aquellas que se necesiten definir.

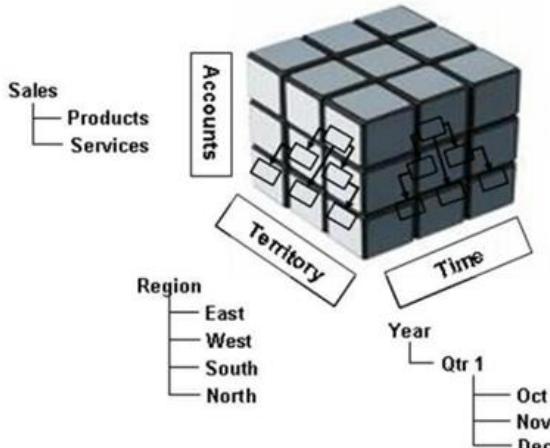


Fig 3 Representación de estructura de un cubo OLAP[5].

3.3 Data Warehouse y Data Mart

DATA WAREHOUSE: Arquitectura de almacenado de datos diseñada para mantener datos extraídos de sistemas de transacciones, operacionales y fuentes externas. Este conjunto de datos está orientado hacia un determinado ámbito, no volátil y variable en el tiempo que ayuda a la toma de decisiones en la entidad que se utiliza.

DATA MART: Subconjunto de datos con el propósito de ayudar a un área específica dentro del negocio para que se puedan tomar mejores decisiones.

3.4 Extract Transform Load (ETL)

Proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos, limpiarlos y cargarlos a otra base de datos o sistema operacional para apoyar un proceso de negocio.

4. BI APPLICADO A LA GESTIÓN DE RRHH

Originalmente la mayoría de Informes y conocimiento que se ofrecía, y en el cual están centrados muchos de los productos que ofrece el mercado de BI, era sobre información económica y financiera de la empresa. Pero cada vez son más los sectores en los cuales se aplica BI para extraer el conocimiento de los datos que se disponen, como es el caso de la gestión de RRHH. Se ha hecho un pequeño estudio sobre las más comunes y principales funcionalidades que ofrecen diferentes productos (microStrategy, IBM:Cognos,SaS, etc.) BI en el sector de RRHH.

Los puntos en común en varias aplicaciones de gestión de RRHH(HR management) son:

- 1) Detección, análisis y/o predicción del "attritionrate" (desgaste de plantilla) (bajas, movimientos de plantilla).
- 2) Talent Management.
- 3) Entrenamiento (LMS Learning Management System).
- 4) Reclutamiento (ATS, Application Tracking System).
- 5) Performance (por sectores, departamentos, etc.).
- 6) Succes Plan, ("Historial de trabajador").
- 7) Absentismo.
- 8) Costes personal.

Como vemos el objetivo de estos productos es ofrecer el máximo conocimiento posible a partir de los datos que puede contener cualquier empresa sobre sus trabajadores.

5. OBJETIVOS

Para el proyecto realizado y el cual sirve como base de este artículo, se quiere definir todo el proceso necesario para conseguir integrar y automatizar una herramienta de análisis de datos de RRHH a partir del origen de datos de la suite Ekon.Laboro.

En la figura 4 se muestra un esquema básico sobre el proceso que se seguirá para realizar el proyecto y las herramientas que se utilizarán para cada uno de los pasos que se siguen.



Fig. 4 Esquema del proceso realizado para la generación de los Informes.

El proyecto por lo tanto ha de:

- 1) Extraer los datos (desde diferentes fuentes), preferiblemente de forma diaria y automática.
- 2) Transformar datos para utilizarlos en cubos OLAP.
- 3) Cargar los datos con el software UNIT4 Business Analytics para poder generar/diseñar los diferentes informes.

En la Tabla II se especifican los objetivos marcados para el proyecto y la prioridad de éstos.

TABLA II
OBJETIVOS Y PRIORIDAD DEL PROYECTO

Objetivo	Prioridad
OBJ1. Desarrollar una APP de UNIT4 Business Analytics para el análisis de datos de RRHH	Crítico
OBJ2. Mostrar información sobre las características de la plantilla	Prioritario
OBJ3. Mostrar información sobre el tipo de contrataciones	Prioritario
OBJ4. Mostrar información sobre los costes de la plantilla	Prioritario
OBJ5. Mostrar información sobre el absentismo	Prioritario
OBJ6. Implementar proceso ETL que permita crear los cubos a partir de los datos disponibles en la empresa	Prioritario
OBJ7. Automatizar el proceso ETL con Jenkins	Secundario
OBJ8. Diseño de la aplicación intuitiva para los usuarios	Prioritario
OBJ9. Aprender metodologías de trabajo en el desarrollo de software	Secundario
OBJ10. Aprender herramientas/funcionamiento necesarios para la implementación de BI	Secundario

6. REQUERIMIENTOS

Teniendo en cuenta los datos disponibles, los objetivos que se proponen y el tipo de información que empresas del sector suelen proporcionar se ha decidido que la aplicación disponga de las siguientes funcionalidades:

1) Muestra información sobre Plantilla:

- Perfil de la plantilla por sexo
- Perfil de la plantilla por edad
- Perfil de la plantilla por población
- Perfil de la plantilla por antigüedad
- Plantilla por categorías
- Plantilla por puesto de trabajo
- Evolución de la plantilla (Alta/Baja)

2) Muestra información sobre Contratación:

- Modelo de contratación
- Grupos de cotización
- Por convenio
- Por Régimen de la Seguridad Social
- Bajas por Tipo de baja
- Contratación por Nivel de formación
- Contratación por Nivel Salarial

3) Muestra información sobre costes:

- Coste por categoría
- Coste por puesto de trabajo
- Coste por trabajador
- Costes general

- 4) Muestra información sobre Absentismo:
- Evolución del Absentismo por Motivo de Ausencia
 - Comparativa del absentismo (por años)
 - Top en absentismo por Motivo de Ausencia
 - Absentismo por puesto de trabajo
 - Absentismo por categoría

Se ha utilizado Talend OS para el proceso ETL, Jenkins para automatizar este proceso y UNIT4 Business Analytics para la generación de vistas e informes

7. METODOLOGÍA

7.1 Planificación

La empresa UNIT4 trabaja con la metodología ágil de desarrollo SCRUM[6]. SCRUM se basa en un proceso iterativo e incremental. Aunque este se aplica por equipos, se han seguido la mayoría de conceptos y metodologías de trabajo aplicadas para SCRUM. Estos básicamente han sido:

1) Definición del ProductBacklog: Inventario de funcionalidades y mejoras que se han de incorporar. Conjunto de historias valoradas por puntos (en este caso equivalen a 2 días de trabajo/ persona)

2) Requerimientos atomizados (Historias): Subdivididas en tareas, pueden ser de análisis, desarrollo, pruebas. Suelen ser pequeñas (1-12h)

3) Planificación mediante Sprints: En nuestro caso serán de 5 sprints de 4 semanas aproximadamente. Cada Sprint consiste de un conjunto de historias

4) Reuniones: DailyScrums, Sizing (puntuar las historias), Pre-Planning y Planning (Historias por Sprint y valoración de tareas en horas) y Sprint Reviews (al final de cada Sprint)

El proyecto se ha realizado entre los meses de febrero y junio de 2014. En la Fig. 5 se muestra el diagrama de Gantt con los intervalos de tiempo de cada Sprint. Cada Sprint tendrá aproximadamente (depende los días hábiles) 10 puntos, los puntos representan 2 días de trabajo. Se muestra de manera general lo que se ha realizado en cada Sprint:

Sprint 0 (9.5 puntos)

- Búsqueda de información BI & BA 3.5p
- Instalaciones y tutoriales diferentes herramientas 4p
- Reuniones (Daily, Planning, Revision) 1p
- Búsqueda funcionalidades herramientas HCM 0.5p
- Documentación (Informe Inicial) 0.5p

Sprint 1 (10 puntos):

- Ánalisis: Productbacklog 4p
- Definición e implementación Prototipo 4.5p
- Reuniones (Daily, Planning, Revision) 1p
- Documentación (Informe progreso I) 0.5p

Sprint2 (8 puntos)

- Implementación funcionalidades 2p
- Definición e implementación ETL 4p
- Reuniones (Daily, Planning, Revision) 1p
- Documentación (manuales) 1p

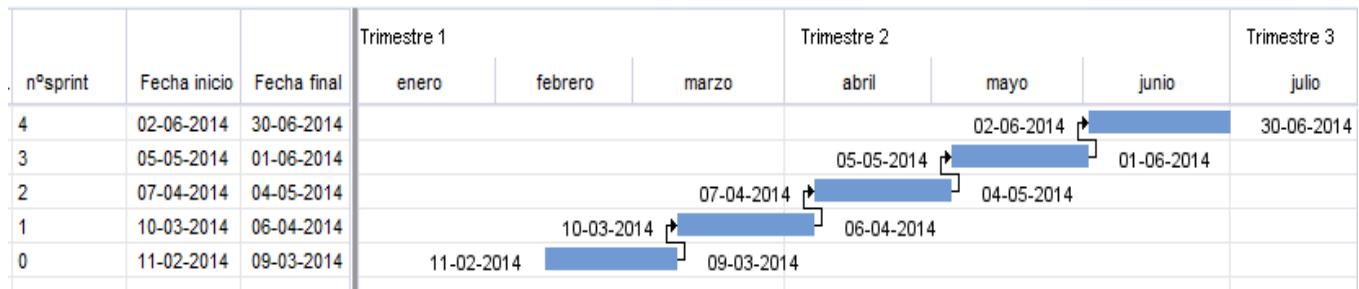


Fig. 5 Gantt con los Sprints a realizar. Cada Sprint tiene una duración de 4 semanas hábiles comprendidas desde el 11-02-2014 al 30-06-2014

Sprint3 (10puntos)

Diseño de la interficie con U4BA 6p
 Ejecución tests Unitarios 2p
 Reuniones (Daily, Planning, Revision) 1p
 Documentación (Informe de progreso II) 1p

Sprint4 (10puntos)

Mejoras de interfície gráfica (2p)
 Cierre (2p)
 Presentaciones (2p)
 Documentación (Manuales, Artículo) 4p

7.2 Análisis de los datos de origen

Los datos se obtienen de diferentes tablas existentes en las bases de datos de la suite ekon.Laboro. La suite Ekon.Laboro contiene más de 1000 tablas y se han utilizado más de 30 tablas para recopilar la información que se requería. En el Apéndice 1 se muestra el esquema (simplificado a las tablas y campos que se usan) de entidad-relación de las tablas de origen de los datos.

Se trabaja sobre una tabla central (*no_tra_laborales*) que contiene (mediante relaciones a otras tablas o de forma directa) información sobre los contratos realizados. A partir de esta información se define la estructura de las diferentes dimensiones que se utilizan.

7.3 Análisis y diseño del cubo

7.3.1 Tablas de dimensiones

Las dimensiones siguen una estructura de árbol, definida en una misma tabla, con el siguiente formato: (*id*, *name*, *parent_id*). Las dimensiones deben cumplir:

1) *id* es único para cada registro.

2) *parent_id* hace referencia a un *id* existente en la propia dimensión (o ser null en caso que sea un elemento raíz).

Las dimensiones nos permitirán filtrar y agrupar todos los elementos de las Tablas de Hechos. Pongamos el ejemplo de cómo se ha realizado la dimensión jerarquía, esta se usa para filtrar con un menú desplegable el contenido que se mostrará en los informes.

Ejemplo Jerarquía

Tenemos un conjunto de tablas relacionadas que nos proporcionan la siguiente información:

Empresa -- *Cuentas de cotización* -- *Centros de Trabajo* -- *Contratos Trabajadores*

Para definir la dimensión jerarquía creamos un primer nivel raíz con valor ("+","Todos",null). En un segundo nivel se introducen las empresas (*id_empresa*, *nombreEmpresa*, "+"). En el siguiente nivel se definen las cuentas de cotización de cada empresa (*id_cuenta*, *nombreCuenta*, *id_empresa*). Lo mismo entonces para un cuarto nivel para Centros de Trabajo. Una vez definida una dimensión hacemos referencia a ésta introduciendo el *id* del Centro de Trabajo (el nodo hoja) en la Tabla de Hechos. Esto nos permite que, en caso de seleccionar en el menú desplegable un nodo de un nivel, se muestren los hijos de éste. En el caso anterior, si seleccionamos una cuenta de la seguridad social, los centros de trabajo asociada a esta, y así sucesivamente.

7.3.2 Tablas de Hechos

Contienen la información que se desea mostrar, las medidas (measures) y los filtros o dimensiones que definen estos valores.

Para la información que se necesita mostrar se ha dividido en tres tablas de hechos:

1) **Fact_Plantilla**: contiene la información necesaria para los informes de Evolución de Plantilla y Contratación.

2) **Fact_Table_Absentismo**: contiene la información necesaria para los informes de Absentismo.

3) **Fact_Table_Costes_RRHH**: contiene la información necesaria para los informes de Costes de RRHH.

Para simplificar tanto el proceso ETL, como la definición de las vistas (en las que se definen las dimensiones por las que se filtra la información), se ha dividido en diferentes tablas de hechos.

Cada una de estas tablas contiene tantas columnas como dimensiones a las que esta referenciada y las medidas correspondientes en otras columnas (en este caso cada tabla de hechos solo usa actualmente una medida)

En el Apéndice A2 se muestran los esquemas de las Tablas de Hechos.

7.3.3 Medidas

Las medidas son las unidades que se quieren mostrar. Estas se definen en las Tablas de Hechos. En este caso se han definido tres tipos de medidas:

1) Cantidad: En Fact_Plantilla, indica el número de personas, como la información corresponde con un contrato a una persona este valor es siempre 1.

2) Dias_Ausente: En Fact_Table_Absentismo, indica el número de días por més que el trabajador ha estado de baja.

3) Coste: En Fact_Table_Costes_RRHH, puede contener el coste del Salario Bruto por mes o el coste que le supone a la empresa (Salario Bruto más coste SS).

7.4 Definición del proceso ETL

La estructura que se utiliza se divide en varios paquetes principales:

DataSources, donde se define la extracción de datos de las fuentes de origen

Execute_data_cubes_scripts, donde se define la ejecución de los scripts para la carga de datos a cubos (contenido de las tablas int_)

Load_context, donde se carga el archivo de configuración y define las bases de datos origen y destino y la localización de los scripts

Months_Between_Dates, devuelve los meses entre dos fechas, para poder generar historiales de plantilla previos a la implantación por meses

Process_cube, donde se define el procesado del cubo (con plugins de U4BA para Talend)

Save_data_int_tables, donde se define el proceso de transformación de los datos extraídos y se guarda en las tablas int_TableName (usadas por U4BA para los cubos)

Test, donde se definen los test unitarios para comprobar la integridad de los datos obtenidos al procesarlos

El Propósito principal de este proceso es:

1) Recoger los datos

2) Transformarlos para la inserción en tablas int_ (generadas por U4BA para insertar los datos)

2.1) Des normalización de las tablas

2.2) Toda tabla int_ debe cumplir (para que se cargue correctamente al procesar el cubo):

Formato de tres campos (gen_id, name, parent)

Gen_id único

Parent debe existir en la propia tabla (o ser null)

2.3) Generación de la estructura/árbol a partir de la tabla_int

3) Cargar datos a cubo

4) Cargar datos (desde tablas int_):

7.4.1 Ejemplo ETL

Como ejemplo se muestra el procedimiento seguido para el caso Población.

Tres tablas contienen la información necesaria para crear la dimensión dim_trabajador_poblacion. La figura 6 muestra el esquema de las tablas necesarias para el ejemplo y la Tabla III los diferentes niveles de la dimensión, con la relación a las tablas origen y su valor:

El primer paso que se realiza es la extracción de datos de la fuente origen. Para ello cargamos los campos que necesitamos de las tres tablas.

El segundo paso es llenar la dimensión. Para cada nivel de la dimensión se transforman los datos origen a la

estructura correspondiente de dimensión (id, name, parent_id). Dentro de este proceso además se comprueba, mediante SQL, que las id de la dimensión sean únicas y que parent_id esté relacionado con algún id. En este caso los valores de la dimensión población (dim_trabajador_poblacion) serán:

- 1) Nivel 0: Todos
- 2) Nivel 1: Países
- 3) Nivel 2: Provincias
- 4) Nivel 3: Poblaciones

El tercer paso es llenar el valor en la Tabla de Hechos. En el caso de dimensión Población ésta se utiliza para la tabla de Hechos Fact_Plantilla. Para llenar entonces el campo de dim_trabajador_poblacion de Fact_Plantilla se debe, por cada registro de no_tra_laborales (contratos laborales), obtener los valores de xpais_id, xprovincia_id, x población_id de la tabla que contiene información sobre las personas (físicas o jurídicas) data_xc_personas. El valor será entonces el id del Nivel 3 de la dimensión

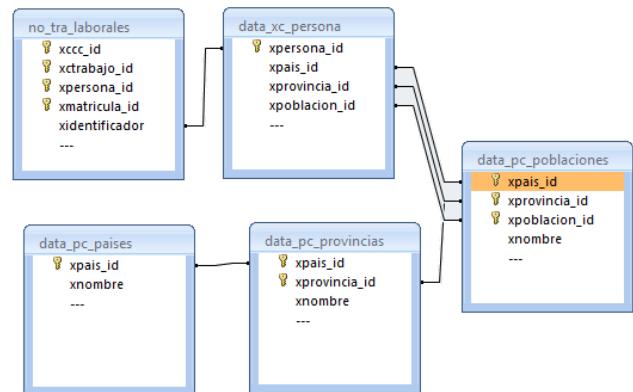


Fig. 6 Diagrama (parcial) tablas necesarias para la creación de la dimensión Población y la referencia de ésta en la Tabla de Hechos Fact_Plantilla

TABLA III
DIMENSIÓN POBLACIÓN. TABLAS Y CAMPOS ORIGEN. VALORES DE LA DIMENSIÓN

Nivel	Origen	Id	Name	Parent_id
Nivel 0	SQL Insert	""	"Todos"	NULL
Nivel 1	Data_pc_paises	"PAIS_+"+xpais_id	Xnombre	""
Nivel 2	Data_pc_provincias	"PAIS_+"+xpais_id+"PROV_+"+xprovincia_id	Xnombre	"PAIS_+"+xpais_id
Nivel 3	Data_pc_poblaciones	"PAIS_+"+xpais_id+"PROV_+"+xprovincia_id+"POBL_+"+xpoblacion_id	Xnombre	"PAIS_+"+xpais_id+"PROV_+"+xprovincia_id

7.5 Definición de los Informes

Todas las vistas e interfaz gráfica están realizadas con la herramienta UNIT4 Business Analytics. La herramienta se utiliza tanto en la definición del esquema de los cubos como en las diferentes funcionalidades a mostrar.

La interfaz gráfica se divide, como anteriormente en la definición de los datos, en cuatro grupos de informes agrupados por:

- 1) Evolución Plantilla.
- 2) Contratación.
- 3) Costes.
- 4) Absentismo.

Se define también un botón "prompter" que permite seleccionar y modificar el período en el cual se quiere mostrar los datos, y en caso de querer filtrar por alguna dimensión, un prompter de la dimensión deseada.

La Vista se divide en 3 partes principales:

1) Informe:

A) Muestra una tabla donde se puede ver en detalle (hasta la información del trabajador) la información definida.

B) Muestra una gráfica donde se puede ver la evolución del concepto seleccionada por periodo de tiempo.

C) Muestra una segunda gráfica (en casos donde hay gran cantidad de conceptos se presenta en un solo gráfico) con más información adicional (porcentajes, evolución por años, etc.).

2) Selección de concepto (SubInformes): Permite navegar por los diferentes conceptos definidos en esta Vista.

3) Prompter/s de filtrado: Prompter/s en la parte superior izquierda que permite seleccionar la fecha inicial y período del que se quieren mostrar los datos o un valor específico para una dimensión de filtrado.

Las vistas "detalle por trabajadores" se filtrarán también por la categoría seleccionada en los informes. Estos informes a su vez permitirán seleccionar un trabajador para ver el "Historial del Trabajador".

7.5.1 Vista Informes.

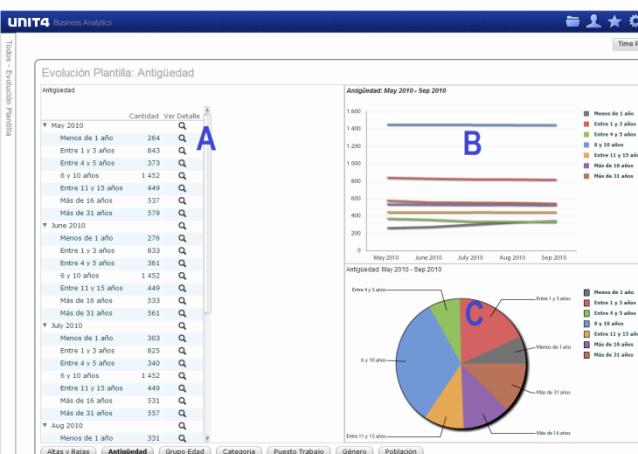


Fig. 7 Vista de un informe generado con UNIT4 Business Intelligence

Para cada elemento que muestra datos en el informe (Tablas, gráficas) se deben definir los valores que se mostrarán en las filas, columnas y las condiciones. En las condiciones se introduce la medida que se quiere mostrar (por ejemplo cantidad) y los valores (fijos o según se seleccione) de las dimensiones que servirán como filtro. En el caso de seleccionar una empresa, por ejemplo, filtrar los resultados solo por esa empresa.

8. RESULTADOS

8.1 Resultados

Durante los cuatro meses que se ha estado desarrollando este proyecto, en la empresa UNIT4, se ha conseguido hacer funcionar en los servidores de dicha empresa una aplicación de BI a partir de los datos que dispone en su CRM Ekon.Laboro.

Más de 40 informes, realizados con la herramienta UNIT4 Business Analytics, ofrecen información en tablas y gráficas de los cuatro temas que se marcó como objetivos (OBJ1 a OBJ5). Dentro de cada uno de estos temas (Evolución de Plantilla, Contratación, Absentismo, Costes RRHH) se accede a diferentes informes, correspondientes a las funcionalidades previamente establecidas, que permiten tener una visión general del concepto que se está analizando. A su vez estos informes pueden servir como filtro para nuevos informes que muestran más en detalle (trabajadores) un aspecto seleccionado y llegar a obtener un "Historial de Trabajador". Finalmente desde un concepto amplio, como podría ser absentismo por puesto de trabajo en una empresa, se consigue ver el detalle de los trabajadores con ausencias en un período de tiempo.(OBJ8)

Aun así la parte principal no es ver el detalle del trabajador, sino poder ofrecer de un modo fácil y rápido una visión general de lo que se necesita.

Se han creado las bases de datos para almacenar los cubos y el proceso ETL, con la herramienta Talend OS, para llenar estos datos de forma automática y diaria (herramienta Jenkins)(OBJ6,OBJ7 y OBJ10). Se ha implementado el proceso ETL de tal forma que permite, mediante la modificación de la información de conexión en un archivo de configuración, conectarse a cualquier base de datos de un cliente que esté usando Ekon.Laboro.

8.2 Pruebas Realizadas

8.2.1 Pruebas Unitarias y de integridad de los datos

Se definen Unitests, creados con la herramienta Talend OS, para los diferentes informes que se muestran, agrupados por los 4 grupos de informes. En cada test se realizan tres pasos:

- 1) Sentencia SQL contra la Base de Datos de Ekon.Laboro (Datos originales)
- 2) Sentencia SQL ("equivalente") contra la Base de Datos que contiene los cubos
- 3) Comparación automatizada de los resultados, informando de test superado en caso de ser los mismos resultados.

8.2.2 Pruebas Configuración

Se han realizado pruebas de ejecución en diferentes dispositivos así como generar una nueva base de datos contra la que almacenar el cubo, teniendo que modificar únicamente las propiedades de conexión del fichero de configuración.

9 CONCLUSIÓN

Las empresas poseen y pueden obtener grandes cantidades de datos que les permiten mejorar y/o optimizar su rentabilidad en los servicios que ofrecen. Convertir estos datos en información útil, de fácil y rápida interpretación para aquellos que la necesitan es el propósito de Business Intelligence.

En este proyecto se ha conseguido definir una automatización del proceso ETL sobre el entorno Ekon (el cual contiene los datos de RRHH) y realizar los diferentes pasos necesarios para poder proporcionar a la gestión de RRHH:

- 1) Un entorno dedicado exclusivamente al análisis
- 2) Más información a partir de los datos existentes
- 3) Acceso más rápido y fácil a la información

Esta solución BI proporciona entonces la información necesaria a los empleados para la toma de decisiones, y podría permitir dar nuevas herramientas haciendo uso de técnicas de análisis predictivo para la generación de nuevos datos.

Desde un punto de vista más personal, realizar este proyecto conjuntamente con las prácticas en la empresa UNIT4 me ha aportado diferentes aptitudes y conocimientos:

El uso de su herramienta de gestión SCRUM me ha permitido conocer más en detalle esta metodología ágil de trabajo.

Estar trabajando en un entorno laboral me ha permitido ver de un modo directo el estilo y métodos de trabajo en una empresa del sector informático.

En cuanto al contenido del proyecto, Business Intelligence, me ha permitido poner en práctica conceptos aprendidos durante la carrera así como obtener nuevos conocimientos sobre técnicas y herramientas (como procesos ETL, generación de informes para perfiles no técnicos, etc.)

10 TRABAJO FUTURO

El conjunto de informes que se ha realizado no se deben ver tan solo como un producto cerrado. La estructura implementada y las herramientas disponibles permiten adaptar la solución a diferentes necesidades que pueda tener cualquier cliente. Se pueden añadir, a las agrupaciones de reportes, y sus correspondientes *FactTables*, nuevos reportes y dimensiones, así como adaptar las existentes para mostrar la información en un formato distinto.

El objetivo principal que persigue BI es dar

conocimiento de la forma más directa y sencilla para el usuario, con lo que dependiendo del tipo de datos y empresa/personas pueda variar la forma en la que presentar los datos.

El punto principal en el que se ha basado el proyecto es sobre el "Reporting" y muestra de historiales sobre pasado y presente. Una de las partes, contemplada ya dentro del amplio concepto de BI, es el uso de Reports para informar también sobre la previsión futura y/o aplicar técnicas de análisis de datos para poder mostrar esos análisis como nuevas funcionalidades.[7][8][9]

Para aplicar y generar estos nuevos datos a partir de los existentes se necesitaría de herramientas específicas para éste paso, pero una vez generados los nuevos datos se les aplicaría el mismo procedimiento que para los datos originales. Diferenciado los originales de los predichos, mediante una dimensión, se podría además comparar los valores calculados con los reales.

AGRADECIMIENTOS

Agradezco tanto a la Universitat Autònoma de Barcelona, como a la empresa UNIT4, la oportunidad que me han dado de realizar este proyecto en una empresa. Agradezco también al profesor encargado de tutorizar este proyecto, Josep Lladós, y al responsable de mi proyecto en la empresa, Jordi Soler, por sus consejos y guía durante los cuatro meses de desarrollo del proyecto.

BIBLIOGRAFÍA

- [1] J. Ranjan, «Business Intelligence: Concepts, Components, Techniques and Benefits,» *Journal of Theoretical and Applied Information Technology*, 2009.
- [2] A. A. R. Gomez, «Inteligencia de Negocios: Estado del Arte,» *Scientia et Technia*, 2010
- [3] «Gartner,» [En línea]. Available: <http://www.gartner.com/>. [Último acceso: 24 Junio 2014].
- [4] «Business Intelligence share market,» [En línea]. Available: <http://www.forbes.com/sites/louiscolumbus/2014/04/29/2013-business-intelligence-and-analytics-market-share-update-sap-continues-market-leadership/> [Último acceso: 24 Junio 2014].
- [5] «Business Intelligenceinfo,» [En línea]. Available: www.businessintelligence.info/definiciones/que-es-olap.html [Último acceso: 24 Junio 2014].
- [6] [En línea]. Available: www.proyectosagiles.org. [Último acceso: 24 Junio 2014].
- [7] SaS, «The Current State of Business Analytics: Where do we go from here?» 2012.
- [8] T. Klatt, M. Schlaefke y K. Moeller, «Integrating business analytics into strategic planning for better performance,» *Journal of business strategy*, 2011.
- [9] SaS, «The HR imperative for predictive modeling. Elevate recruitment and retention to strategic status and advanced analytics»

APENDICE

A1. Esquema de la base de datos origen

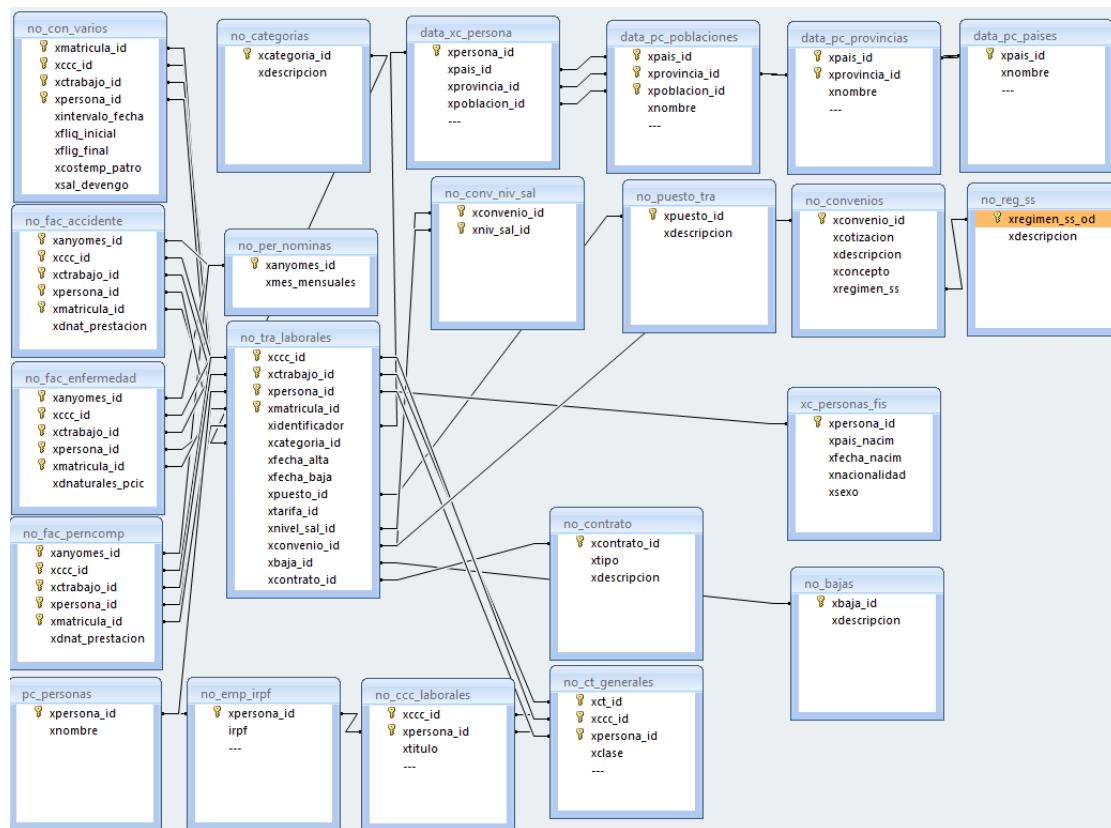
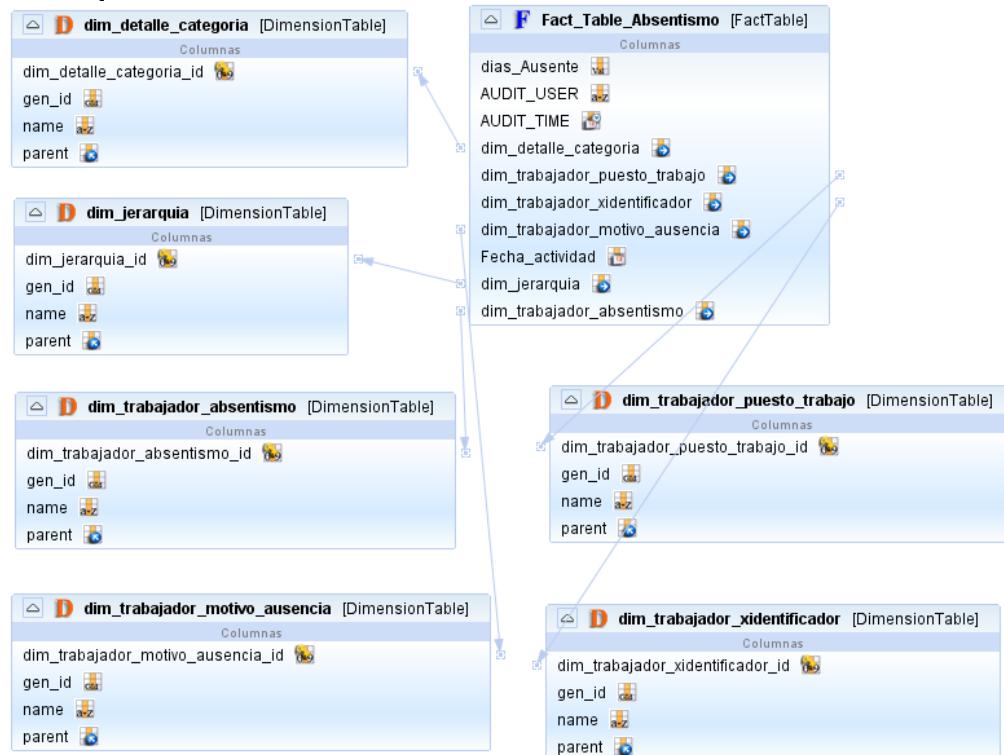


Fig8. Diagrama Parcial Base de Datos Laboro, adaptado a la información necesaria para el proyecto

A2. Esquema de las tablas de hecho

Fig9. Esquema de la Tabla de Hechos Absentismo (*Fact_Table_Absentismo*)

Aunque no representado en las siguientes, cada dimensión de las tablas de hechos *Fact_Plantilla* y *Fact_Table_Coste_RRHH* está relacionada con la tabla correspondiente a la dimensión

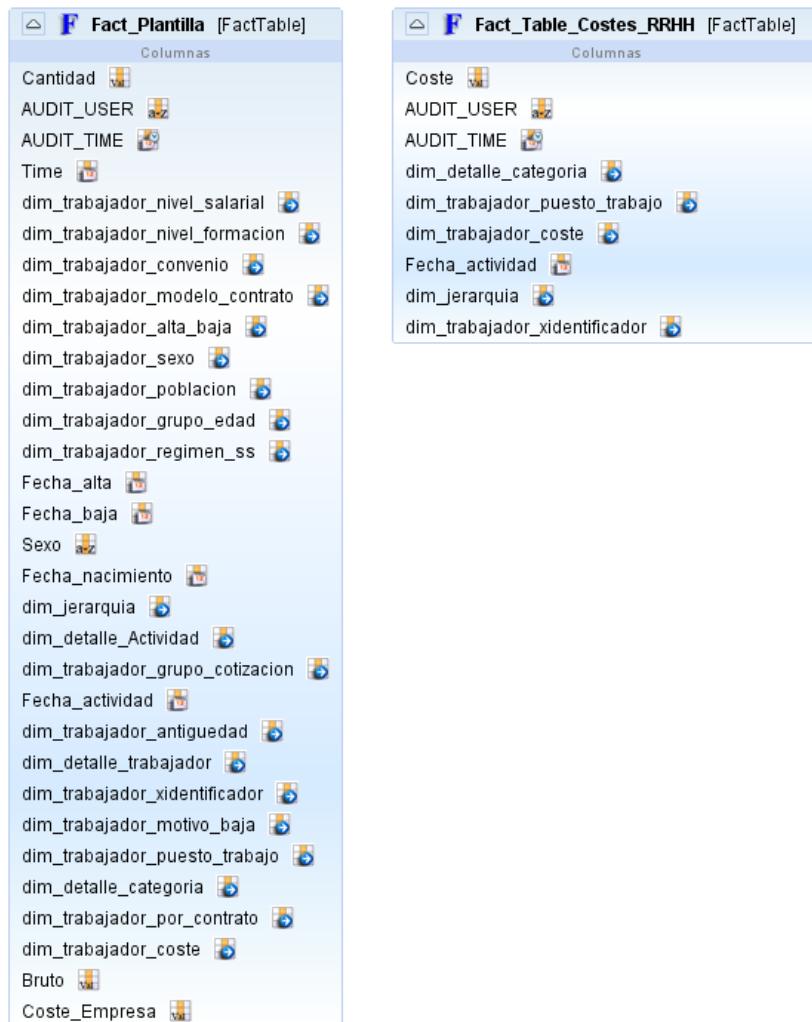


Fig 10. Estructura de las Tabla de Hechos Plantilla y Costes RRHH (*Fact_Plantilla*, *Fact_Table_Absentismo*)

A3. JOB TALEND

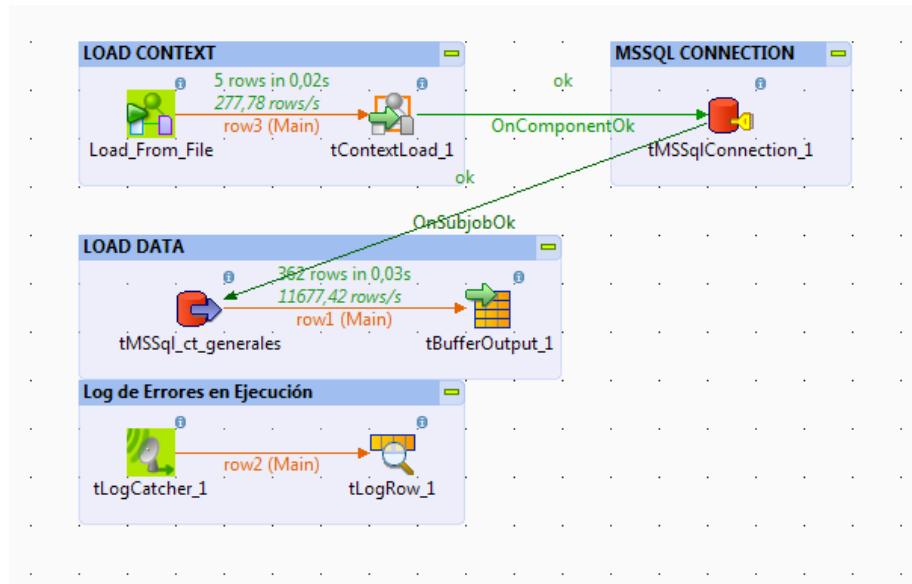


Fig 11. Vista Diseño de un Job Talend

A4. VISTAS INFORMES

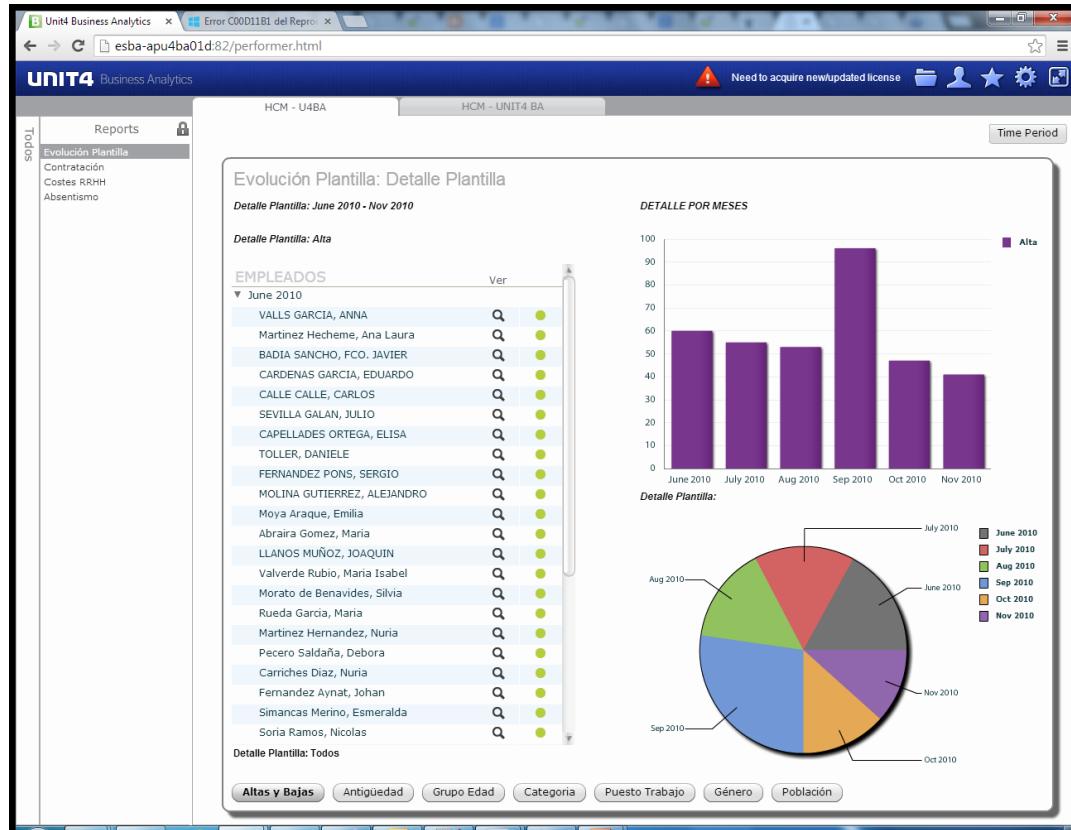


Fig 12. Vista de un Informe de la aplicación HCM-U4BA. En este informe se muestra la evolución de altas en un período de tiempo seleccionado

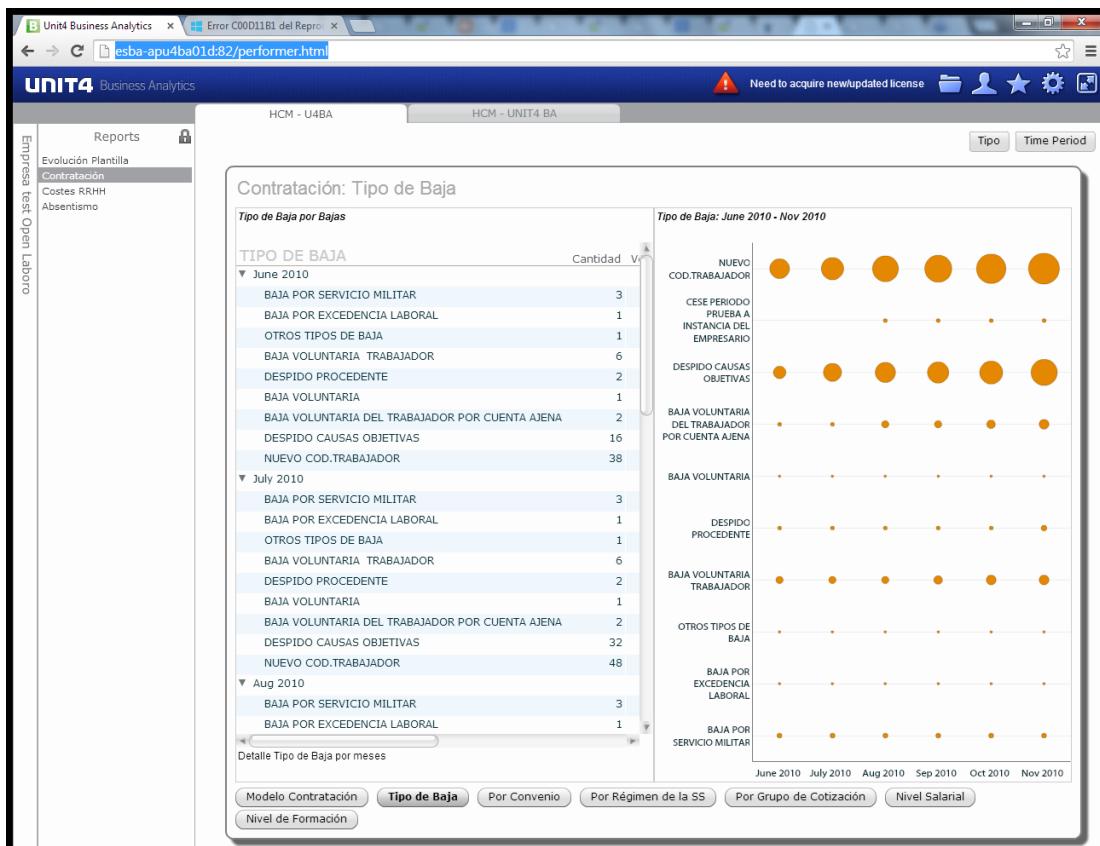


Fig 13. En este informe se muestra la evolución y comparación entre tipos de bajas

