Modelado de un sistema de Business Intelligence para un sistema de gestión de almacén. Desarrollo de la carga de datos

Proyecto Final de Carrera

Ingeniería Informática

**Autor**: Carlos Expósito Cerro

**Director**: Andrés Boza García

[Fecha]

Resumen

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed nisi turpis, iaculis a pulvinar quis, luctus et lorem. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nullam vitae purus eros, id auctor dolor. Sed et nisl quis nibh fermentum cursus ut at elit. Etiam condimentum porta leo quis tempor. Quisque commodo lobortis aliquet. Etiam tincidunt, libero ut vehicula euismod, justo augue lobortis sem, et facilisis velit lacus tristique dolor.

**Palabras clave:** Business Intelligence, Pentaho, Kettle, Data Warehouse.

# Tabla de contenidos

[Tabla de contenidos 6](#_Toc364701862)

[1 Introducción 7](#_Toc364701863)

[1.1 Presentación 8](#_Toc364701864)

[1.2 Contexto 8](#_Toc364701865)

[1.3 Objetivos 8](#_Toc364701866)

[1.4 Estructura del proyecto 8](#_Toc364701867)

[2 Base Teórica y (en su caso) entorno del problema 9](#_Toc364701868)

[3 Propuesta/Solución/Análisis/Trabajo Realizado (Mi aportación) 11](#_Toc364701869)

[4 Conclusiones 13](#_Toc364701870)

[5 13](#_Toc364701871)

# Introducción

## Presentación

La presentación es lo primero que un lector del proyecto va a leer, por tanto debe ser un texto introductorio, claro, no complejo, que pueda acercar al lector a lo que se va a tratar a continuación.

El proyecto se centra principalmente en los procesos de extracción, carga y transformación o procesos ETL, aunque abarca todos los aspectos de un proceso de Business Intelligence, desde la creación del almacén de datos hasta la creación de informes obtenidos desde ese almacén.

## Contexto

Cuando se trate de un trabajo en empresa, debe presentarse brevemente el contexto en el que se ha desarrollado el proyecto.

En cualquier proyecto de ingeniería se pueden identificar y diferenciar ciertas fases que son comunes en cualquier ámbito en el que trabajemos, tales como la identificación y análisis de requisitos, ---------, por ejemplo. En el presente documento se hablará de algunas de éstas fases sobre las que el lector debe saber y tener en cuenta que se centran en el ámbito de la Ingeniería Informática y proyectos relacionados en este ámbito, por lo que de ahora en adelante cuando se hable de alguna de las fases antes mencionadas de un proyecto, nos referiremos única y exclusivamente a fases dentro de un proyecto de Ingeniería Informática.

## Objetivos

Qué objetivos se propusieron al comienzo del proyecto. Por tanto, los objetivos del proyecto hablan del futuro del proyecto.

## Estructura del proyecto

Debes indicar como está estructurada la memoria del proyecto. Por ejemplo, La memoria está estructurada en dos grandes bloques, el primer bloque trata sobre … e incluye los capítulos 2, 3 y 4. En el capítulo 2 se trata sobre ….En el segundo bloque en el que se estructura la memoria es … e incluye los capítulos … .

# Base Teórica y (en su caso) entorno del problema

## Identificación de requisitos

En la mayoría de proyectos dentro del ámbito de la Ingeniería, y más concretamente dentro de la Ingeniería Informática, una de las primeras fases de todo proyecto es la de *identificar y analizar los requisitos* del cliente y del problema. En este aspecto podemos diferenciar entre requisitos funcionales, que son aquellos que están relacionados con la funcionalidad de la aplicación o proceso a desarrollar (por ejemplo, en el caso del desarrollo de una aplicación software, la funcionalidad que debe tener implementada la aplicación y que por tanto podrán utilizar sus usuarios), y los requisitos no funcionales, que son aquellos requisitos que vienen determinados por factores externos a la funcionalidad del sistema, como por ejemplo requisitos tecnológicos, legales, técnicos, etc., y que también se deben cumplir.

## Bases de Datos

### Bases de datos SQL

### Bases de datos No-SQL

## Data Warehousing o Almacenes de datos

### Qué es un Data Warehouse

*Un almacén de datos es un sistema que extrae, limpia, ajusta y muestra datos de diversas fuentes en un almacén de datos multidimensional, implementando y dando soporte a las consultas y al análisis para la toma de decisiones*

Un almacén de datos o Data Warehouse es por tanto ….

Hoy en día todavía se confunde muy a menudo lo que es un Data Warehouse; podríamos preguntar a diez expertos en la materia y obtener diez respuestas diferentes. La mayor diferencia solemos encontrarla a la hora de describir cuales son los componentes que forman parte de un proyecto de Data Warehouse. Según , los mayores errores conceptuales sobre lo que es un almacén de datos son los siguientes:

1. Un almacén de datos NO es un producto. Contrariamente a lo que muchos proveedores reclaman, no se puede comprar un almacén de datos debido a que un almacén de datos incluye análisis del sistema, manipulación y tratamiento de los datos, migración de datos y un diseño dimensional y acceso a datos. Un único producto no puede abarcar todas las tareas involucradas en la construcción de un almacén de datos.
2. Un Data Warehouse NO es un lenguaje. No se puede aprender a *programar* un almacén de datos como si de un lenguaje de programación se tratara. Un almacén de datos está formado por varios componentes, cada uno de los cuales puede requerir varios lenguajes de programación o de especificación de datos.
3. Un Data Warehouse NO es un proyecto.
4. Un Data Warehouse NO es un modelo de datos. Un modelo de datos por sí solo no se puede considerar un Data Warehouse, ya que éste incluye por definición una serie de procesos ETL. Después de todo, incluso el mejor diseño o modelado de datos sería inútil sin los propios datos.
5. Un Data Warehouse NO es una copia del sistema operacional. Es un error bastante común creer que copiando el sistema operacional a un sistema independiente es crear un almacén de datos. Al igual que un modelo de datos por sí mismo no puede crear un Data Warehouse, tampoco puede crearse únicamente haciendo una copia de los datos en otro sistema, ya que un Data Warehouse implica también la reestructuración de los datos.

### Origen del Data Warehouse

### Diseño de Data Warehouse

### Diseño de dimensiones

### Diseño de hechos

### Herramientas ROLAP

### Herramientas MOLAP

### Cubos OLAP

## Business Intelligence o Inteligencia de negocios

## Data Mining o Minería de datos

## Big Data

## Procesos ETL

### Extracción

### Carga

### Transformación

## Herramientas Business Intelligence

### Herramientas ETL

### Herramientas de generación de Informes

## Herramienta Pentaho Data Integration: Kettle

## Sistema de Gestión de bases de datos PostgreSQL

# Propuesta / Solución / Análisis / Trabajo Realizado (Mi aportación)

# Conclusiones

-Acercamiento a los objetivos planteados

- Conclusión a nivel de funcionalidad (aplicabilidad de la herramienta en el mercado, empresas, decisores, …).

- Conclusiones del trabajo propio (impresiones, aprendizajes, experiencias, …).

- Líneas de trabajo futuro.

ANEXOS

Bibliografía (En orden alfabético del apellido del autor).

blabla