

# SISTEMA DE ANALISIS DE ECONOMÍA Y DOCENCIA DE LA ETSISI

## **RESUMEN**

En el contexto actual, la gestión y análisis de datos se ha convertido en una herramienta crucial para la toma de decisiones informadas y estratégicas. Las universidades públicas, enfrentan el desafío también de optimizar sus recursos y mejorar así la calidad de la docencia.

Este Trabajo de Fin de Grado se centra en la creación de un sistema de herramientas capaz de automatizar y facilitar el análisis económico y de profesorado al que se enfrenta frecuentemente la Escuela Técnica Superior de Ingenieros y Sistemas Informáticos.

Para ello, el sistema proporcionará un proceso ETL que extraerá de un documento Excel la información que la escuela necesite para sus procesos de negocios.

Una vez hecho esto, se creará un Data Warehouse que nos servirá de base de datos analítica donde se guardarán los datos procesados para facilitar el análisis de estos.

Por último, se creará un cuadro de mando que se podrá proporcionar a los usuarios finales que les servirá de gran utilidad para facilitar sus tareas de toma de decisiones económicas y de docentes.

## **ABSTRACT**

In the current context, data management and analysis have become crucial tools for making informed and strategic decisions. Public universities also face the challenge of optimizing their resources and thus improving the quality of teaching.

This Final Degree Project focuses on the creation of a system of tools capable of automating and facilitating the economic and faculty analysis that the Higher Technical School of Engineers and Computer Systems frequently faces.

To achieve this, the system will provide an ETL process that will extract from an Excel document the information the school needs for its business processes.

Once this is done, a Data Warehouse will be created, which will serve as an analytical database where the processed data will be stored to facilitate its analysis.

Finally, a dashboard will be created that can be provided to end users, which will be of great use in facilitating their economic and teaching decision-making tasks.

# Índice

<b>RESUMEN</b> .....	2
<b>ABSTRACT</b> .....	3
<b>INTRODUCCIÓN</b> .....	5
<b>Contexto</b> .....	5
<b>Objetivo</b> .....	7
<b>Estructura del documento</b> .....	9
<b>PASOS PREVIOS</b> .....	10
<b>Conceptos</b> .....	10
<b>Herramientas</b> .....	11
<b>Pentaho Data Integration:</b> .....	11
<b>Mysql:</b> .....	15
<b>Power Bi:</b> .....	18
<b>Planificación</b> .....	19
<b>PROYECTO</b> .....	20
<b>Análisis</b> .....	20
<b>Diseño</b> .....	23
<b>Implementación</b> .....	34
<b>RESULTADOS</b> .....	56
<b>CONCLUSIONES</b> .....	59
<b>BIBLIOGRAFÍA</b> .....	59
<b>ARCHIVOS</b> .....	59

# **INTRODUCCIÓN**

## **Contexto**

Sin lugar a duda, saber la importancia que los datos tienen hoy en día y su forma de utilizarlos a nuestro favor es algo esencial en la mentalidad de un gran número de empresas en la actualidad.

A raíz del auge de ciertas tecnologías como el Business Intelligence y el uso de las IAs, muchas empresas han optado por mejorar su relación con el apartado analítico para así poder sacar el máximo rendimiento a todos los datos que de una forma u otra han obtenido a lo largo del tiempo.

Según un estudio de la Universidad ISEP, las empresas que optan por utilizar tecnologías Bi obtienen un 5% más de retorno de inversión, aumentan un 53% su eficiencia operativa y mejoran un 48% su toma de decisiones respecto de aquellas empresas que no la utilizan.

Posiblemente uno de los factores que han producido el aumento del Business Intelligence ha sido el gran aumento de datos a lo largo de los últimos años, debido al continuo uso de los dispositivos tecnológicos. Los seres humanos pasamos alrededor de 7 horas navegando en internet y, normalmente, compartiendo los datos por los que se pasa, así como sus búsquedas y por ello sus gustos.

No solo empresas privadas hacen uso del BI, instituciones públicas como con la que nos encontramos en este Trabajo de Fin de Grado también sacan un gran partido a los datos que les rodean para así poder mejorar su rendimiento general.

En este caso hablamos de la Escuela Técnica Superior de Sistemas Informáticos, campus perteneciente a la Universidad Politécnica de Madrid, la cual maneja una gran cantidad de datos internos relacionados con todo tipo de aspectos propios de una universidad. En este trabajo se utilizaron sus datos económicos y de profesorado para poder realizar una herramienta que les permita poder analizar de una forma más sencilla y duradera los datos que extraigan a lo largo de los años.

## Objetivo

El objetivo principal de este Trabajo de Fin de Grado es diseñar y desarrollar un cuadro de mando para el análisis de datos económicos y de profesorado de la Escuela Técnica Superior de Ingenieros y Sistemas Informáticos de la Universidad Politécnica de Madrid. Este cuadro de mando tiene como propósito proporcionar una herramienta intuitiva y eficaz que permita a la institución optimizar la gestión de sus recursos, mejorar la calidad de la docencia y facilitar la toma de decisiones informadas.

La necesidad de contar con un sistema de análisis y visualización de datos surge de la creciente complejidad en la gestión de las instituciones educativas, donde la eficiencia y la efectividad en el uso de los recursos son cruciales. En este contexto, la Escuela Técnica Superior de Ingenieros y Sistemas Informáticos enfrenta desafíos significativos en la administración de sus recursos económicos y en la gestión de su profesorado. Un cuadro de mando bien diseñado permitirá a los administradores y directivos tener una visión integral y detallada de los datos relevantes, facilitando así la identificación de tendencias, patrones y áreas de mejora.

El cuadro de mando se desarrollará utilizando Power BI debido a sus capacidades avanzadas de visualización de datos y su facilidad de uso. Power BI permitirá integrar diversas fuentes de datos, proporcionando una plataforma única para el análisis y la toma de decisiones. A través de este proyecto, se buscará diseñar visualizaciones claras y significativas que permitan a los usuarios comprender fácilmente la información presentada. Estas visualizaciones incluirán gráficos, tablas y otros elementos interactivos que faciliten la explotación y el análisis de los datos.

La creación del cuadro de mando implicará el desarrollo de un proceso ETL (Extracción, Transformación y Carga) que permitirá la recolección y preparación de los datos necesarios. Este proceso garantizará que los datos sean limpiados, transformados y cargados de manera eficiente en un Data Warehouse, donde serán almacenados para su posterior análisis. El Data Warehouse servirá como una base de datos analítica que proporcionará un acceso rápido y confiable a los datos procesados.

La implementación del cuadro de mando también considerará la usabilidad y la accesibilidad para los usuarios finales. El objetivo es que los usuarios finales puedan aprovechar al máximo las capacidades del cuadro de mando para mejorar la toma de decisiones y la gestión de los recursos en la escuela.

En resumen, este Trabajo de Fin de Grado no solo tiene como objetivo mejorar la eficiencia operativa de la Escuela Técnica Superior de Ingenieros y Sistemas Informáticos, sino también fortalecer su capacidad para tomar decisiones basadas en datos. Al proporcionar una herramienta avanzada de análisis y visualización de datos, se espera contribuir significativamente a la misión educativa y administrativa de la institución, promoviendo una gestión más informada y estratégica.



## **Estructura del documento**

Para su mejor comprensión, estructuraré este documento de una forma lineal de proyecto.

Empezaremos con los aspectos generales del proyecto donde se describirán una serie de conceptos necesarios para su claro entendimiento, las herramientas que han sido utilizadas para producir el resultado esperado y la planificación que se siguió a la hora de desarrollarlo.

Una vez vistos los aspectos generales y necesarios del proyecto, nos internaremos en el desarrollo de este. Por ello el documento continuará con el análisis, diseño e implementación del proyecto.

Una vez terminada su explicación se relatarán los resultados obtenidos, así como las conclusiones que se pueden sacar del desarrollo del TFG.

Como final del documento se encuentra las referencias usadas para el proyecto y una explicación de los diferentes archivos generados durante el desarrollo de este.

## **PASOS PREVIOS**

### **Conceptos**

- **Business Intelligence:** Se define como el conjunto de herramientas y buenas prácticas que utilizan las empresas para realizar un análisis eficaz de los datos que poseen y que así sirvan de ayuda para la toma de decisiones de la empresa.
- **ETL:** En inglés, son las siglas de “Extract, Transform and Load”. Se trata de un conjunto de procesos que se encargan de recoger los datos necesarios de la fuente de origen, realizar una serie de operaciones/transformaciones de esos datos para adaptarlos y facilitar así su análisis y, por último, los cargan en el sistema de almacenamiento de destino del que después se puede realizar el análisis directo o, en nuestro caso, montar un panel de control que nos facilitará su visionado y manejo.
- **Data Warehouse:** Es el nombre que se le dan a las bases de datos encargadas de almacenar los datos que van a ser usados para la toma de decisiones de la empresa. Normalmente, son las bases de datos de destino de los procesos ETL, anteriormente comentados. En este caso, esta base de datos está conformada por dos tipos de tablas: dimensión y hechos. Las tablas de dimensión están formada por los datos que proporcionan información característica del proceso de negocio que se está analizando, mientras que las tablas de hechos contienen los valores que nos proporcionan la información analítica.

-

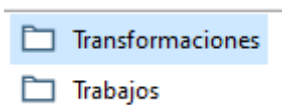
## Herramientas

### Pentaho Data Integration:

























Se trata de una herramienta de código abierto perteneciente a la familia de Pentaho. Esta herramienta permite al usuario la creación de procesos ETL.













Su interfaz es bastante amigable, ya que solamente se basa en dos tipos de procesos, transformaciones y Jobs.



Estos primeros son los que se encargarán de realizar todo el proceso ETL, haciendo uso de su gran número de herramientas que permiten realizar gran cantidad de operaciones, además de las herramientas que permiten realizar la extracción y la carga de los datos de forma sencilla y amable.

- 
- >  Entrada
  - >  Salida
  - >  Streaming
  - >  Transformar
  - >  Utility
  - >  Flow
  - >  Scripting
  - >  Pentaho Server
  - >  Búsqueda
  - >  Uniones
  - >  Almacén de Datos
  - >  Validation
  - >  Statistics
  - >  Big Data
  - >  Cryptography
  - >  Trabajo
  - >  Mapeado
  - >  Bulk loading
  - >  Embebido
  - >  Experimental
  - >  Obsoleto
  - >  History

Por otro lado, los Jobs son los procesos que permiten orquestar varias transformaciones.

- 
- >  General
  - >  Mail
  - >  File management
  - >  Conditions
  - >  Scripting
  - >  Bulk loading
  - >  XML
  - >  Utility
  - >  Repository
  - >  File transfer
  - >  File encryption
  - >  Deprecated

En este caso existen otras herramientas de creación de procesos ETL como puede ser Talent Open Studio o Microsoft SQL Server Integration Services, entre otras. En el caso de mi proyecto he optado por la alternativa de Pentaho debido sus ventajas:

- **Gratuidad:** Su uso no supone un gasto en licencias o mantenimiento de recursos como algunas otras competidoras. Si que existe una edición de pago, pero en este caso la parte gratuita sirve con creces para la realización de este proyecto.
- **Gran comunidad:** Gracias posiblemente a la existencia de la versión gratuita y gracias a que se trata de una herramienta de código libre, se ha creado una gran comunidad alrededor de Pentaho lo que origina un gran abanico de blogs y páginas donde consultar ciertos problemas que uno se puede encontrar al utilizar la herramienta.
- **Conocimiento previo:** En este caso, esta ventaja es propia de mi ya que gracias a experiencias previas en las que he podido participar, partía ya con cierto conocimiento de la herramienta que me ha hecho balancearme sobre utilizarla como herramienta de procesos ETL.

## MySQL:



Para la gestión de los datos analíticos he usado como base de datos de destino una base de datos MySQL.

MySQL es un sistema de gestión de bases de datos relacionales de código abierto desarrollado por Oracle. Esta herramienta nos va a permitir administrar la base de datos de destino de los procesos ETL, en esencia, lugar en el que se van a almacenar los datos una vez procesados por Pentaho Data Integration.

Esta herramienta funciona con un modelo de base de datos relacional, esto significa que los datos se organizan en tablas definidas con una estructura. Existen otras bases de datos que no son relacionales, pero en el caso de realizar una base de datos analítica como la que se debe utilizar en este proyecto

Como su propio nombre indica, utiliza el lenguaje SQL para tratar los datos.

En este caso también existen otras opciones como SQL Server o SQLite, pero al igual que con Pentaho, decidí optar por la opción de MySQL debido a sus características:

- **Gratuidad:** Al igual que con Pentaho, se trata de un software de código abierto haciéndolo bastante accesible.
- **Compatibilidad:** Una gran ventaja a la hora de tomar la decisión de elegir MySQL como gestor de bases de datos es claramente su compatibilidad con Pentaho el cual volcará los datos transformados sin problema en la base de datos.
- **Escalabilidad:** Puede escalar para adaptarse a las necesidades de diferentes aplicaciones. Idóneo en el caso de que se quiera perfeccionar el proyecto y ampliarlo a las distintas áreas que posee la escuela.
- **Conocimiento:** Claramente MySQL se estudia a lo largo de toda la carrera que he estudiado (Ingeniería del Software) es por ello por lo que poseo un conocimiento previo de la herramienta que me ha ayudado bastante a tomar la decisión.
- **Facilidad:** A parte de este gestor, a lo largo de la carrera he tenido la oportunidad de probar otros de ellos, y no he visto ninguno que sea tan sencillo a la hora de realizar casi todas las operaciones necesarias en una base de datos. Además, la instalación de MySQL también es de lo más sencilla necesitando únicamente el instalador de la herramienta el cual te creará el servidor, el prompt y el propio gestor.



La segunda opción que se tuvo en mente a la hora de elegir una base de datos de destino para los datos analíticos fue SQL Server ya que es bastante parecida a MySQL, pero al final opté por la sencillez que nos proporciona esta última.

**[TO-DO] MENCIONAR LAS VERSIONES QUE HE  
UTILIZADO**

## Power Bi:



Con todo ello llegamos a la herramienta sobre la que recae el final del proyecto. PowerBi se trata de un conjunto de servicios que permiten al usuario analizar diferentes orígenes de datos y montar en la herramienta un panel de mandos dinámico con el que se puede realizar una visual analítica sobre esos datos anteriormente mencionados.

Esta herramienta consta de una aplicación de escritorio sobre la que se genera el panel de mandos como si de un PowerPoint se tratase. Consta también de un servicio de software en el que puedes publicar los cuadros de mando que crees para así poder compartirlos con otras personas sin que ellos puedan modificar la aplicación. Por último, también tiene una aplicación móvil con la que puedes modificar y visualizar los paneles de mando de tu cuenta.

En este caso la opción fue bastante sencilla ya que PowerBi es de las mejores herramientas para la creación de cuadros de mando y también poseía un conocimiento previo de la herramienta aunque en este caso en menor medida lo que me ha permitido mejorar con ella a lo largo del proyecto.

## **Planificación**

Lo primero que se tiene que hacer a la hora de afrontar un proyecto es realizar una planificación de lo que se debe de hacer, en que orden y de que forma. Así, podremos tener una visual del alcance que tiene el proyecto, poder hacer una estimación de la duración que tendrá e incluso observar de una forma clara los pasos a seguir para que todo vaya acorde al plan.

Es por ello, que antes de comenzar a trabajar directamente sobre el proyecto, realicé un estudio para tener claro todo lo que se iba a llevar a cabo en el TFG, hablar con mi tutor para aclarar bien el objetivo de este y poder, como se ha mencionado en el apartado anterior, tomar una decisión de las herramientas que se han utilizado.

Como punto de comienzo encontraríamos la primera reunión que tuvimos mi tutor y yo en el que él, me enseñó el tema del TFG, el cual me pareció bastante interesante ya que abarca un tema que, personalmente, me parece muy atractivo, además que podría generar una herramienta que aporte gran valor a los futuros análisis que pueda realizar la escuela en sus datos.

Tras ello, analicé las posibles herramientas que se podrían utilizar en el proyecto y elegí las anteriormente mencionadas.

Una vez hecho todo lo anterior, era momento de comenzar el análisis de los documentos que mi tutor me proporcionó tras esa primera reunión.

## PROYECTO

### Análisis

A la hora de afrontar un proyecto con una metodología tradicional como con el que nos enfrentamos en este caso, lo primero que se debe realizar (tras la planificación y pasos previos) sería un análisis del objetivo que tiene el proyecto, así como de los datos que tenemos para poder realizarlo.

### Origen de los datos

Los datos que se van a utilizar en este TFG provienen de los documentos generados por la propia escuela proporcionados por mi tutor.

Primeramente, la mayoría de los datos son extraídos de forma manual de un PDF que contiene una gran cantidad de datos de los años comprendidos entre 2009 y 2013. Sobre ellos se recogieron los datos económicos y de docencia.

DEPARTAMENTO	CATEGORIA	2006	2007	2008	2009	2010	2011	2012	2013
SIN DEPARTAMENTO	PROF. ASOCIADO TIPO 1	2	1	1	1				
SIN DEPARTAMENTO	PROF. ASOCIADO TIPO 3	1	1	1	1	1	1		
SIN DEPARTAMENTO	TITULAR UNIVERSIDAD INT.		1	1	1	1	1	2	2
SIN DEPARTAMENTO	I.D. PROF. ASOCIADO				1	1	1	1	1
ING DE ORG. ADMON DE EMPRESAS Y ESTAD. (ETSI Industriales)	TITULAR E.U.	4	4	4	4	4	4	4	4
ING DE ORG. ADMON DE EMPRESAS Y ESTAD. (ETSI Industriales)	I.D. PROF. ASOCIADO				2				
LINGÜÍSTICA APLICADA A LA CIENCIA Y A LA TECNOLOGÍA	TITULAR E.U.	4	4	4	4	4	4	4	4

GASTOS				INGRESOS	
EJERCICIO 2006					
CÓD. CENTRO	CAPÍTULO	CONCEPTO	IMPORTE	CONCEPTO	IMPORTE
1861	1	GASTOS DE PERSONAL	6.790.709,82	NOMINATIVA	7.257.182,77
1861	2	GASTOS CORRIENTES EN BIENES Y SERVICIOS	448.037,93	MATRÍCULA	1.918.101,20
1861	4	TRANSFERENCIAS Y SUBVENCIONES CORRIENTES	0,00	CÁNONES OTT	37.055,76
1861	6	INVERSIONES REALES	1.046.094,94	OTROS INGRESOS	2.833,56
182105	6	GASTOS OTT	58.750,96		
TOTAL			8.343.593,65	TOTAL	9.215.173,29
EJERCICIO 2007					
CÓD. CENTRO	CAPÍTULO	CONCEPTO	IMPORTE	CONCEPTO	IMPORTE
1861	1	GASTOS DE PERSONAL	7.206.383,86	NOMINATIVA	7.525.997,40
1861	2	GASTOS CORRIENTES EN BIENES Y SERVICIOS	852.957,57	MATRÍCULA	1.598.486,42
1861	4	TRANSFERENCIAS Y SUBVENCIONES CORRIENTES	0,00	CÁNONES OTT	50.359,15
1861	6	INVERSIONES REALES	962.659,33	OTROS INGRESOS	6.775,69
182105	6	GASTOS OTT	70.436,84		
TOTAL			9.092.437,60	TOTAL	9.181.618,66

DIEGO CRUZ CUERVO  
SISTEMA DE ANALISIS DE ECONOMÍA Y DOCENCIA DE LA ETSISI

GASTOS				INGRESOS			
Artículo	Concepto	*	Importe (€)	Artic	Concepto	*	Importe (€)
Cap. 2					Asignación UPM	c	519.777.83
20	Arrendamientos y cánones	c n	10.500,00 5.000,00	32	Cánones	n	75.000
21	Reparación, mantenimiento y conservación	c n	20.000,00 10.000,00	39	Indemnizaciones (sinistros, robos)	n	20.000
22	221.00. Suministro de electricidad	c n	80.000,00 12.000,00	55	Cesiones de locales	n	78.400
	221.01. Suministro de agua	c n	9.000,00 4.000,00				
	221.02. Suministro de gas	c n	50.000,00 8.000,00				
	226.01. Atenciones protocolarias	c n	5.000,00 4.000,00				
	Resto Artículo 22 (Material Suministros y Otros)	c n	115.500,00 28.000,00				
23	Indemnizaciones por razón del Servicio	c n	2.000,00 2.000,00				
Cap. 4	481. Subvenciones, premios	c n	100.00 400,00				
Cap. 6	Inversiones	c n	227.677.83 100.000,00				
TOTAL		c n t	519.777,83 173.400,00 693.177.83				

Por otra parte, se completaron los años entre 2014 y 2023 para el apartado económico con los PDF económicos que se generaron esos años. Así, se pudo realizar un mejor análisis de este apartado de datos.

## Estructura de los datos

Como se ha comentado en el apartado anterior, los datos son extraídos de documentos PDF de forma manual, algo que claramente es muy ineficiente para el futuro plan que pueda tener este proyecto.

Observando los datos generé una plantilla Excel, tanto del apartado económico como del de docencia, en el que los encargados de realizar el análisis de los datos puedan incorporar allí directamente los datos y así facilitar luego los procesos que van a actuar encima de ellos.

DIEGO CRUZ CUERVO  
SISTEMA DE ANALISIS DE ECONOMÍA Y DOCENCIA DE LA ETSISI

Id_economico	Tipo	Articulo	Concepto	Consolidado	Importe	Anio
1	Gasto	20	Arrendamientos y cánones	c	10500	2022
2	Gasto	20	Arrendamientos y cánones	n	5000	2022
3	Gasto	21	Reparación, mantenimiento y conservación	c	20000	2022
4	Gasto	21	Reparación, mantenimiento y conservación	n	10000	2022
5	Gasto	22	221.00. Suministro de electricidad	c	80000	2022
6	Gasto	22	221.00. Suministro de electricidad	n	12000	2022
7	Gasto	22	221.01. Suministro de agua	c	9000	2022
8	Gasto	22	221.01. Suministro de agua	n	4000	2022
9	Gasto	22	221.02. Suministro de gas	c	50000	2022

departamento	categoria	anio	cantidad
Sin Departamento	Profesor Asociado Tipo 1	2006	1,25
Sin Departamento	Profesor Asociado Tipo 3	2006	1
Sin Departamento	Titular Universidad Int.	2006	0
Sin Departamento	L.D. Profesor Asociado	2006	0
Sin Departamento	Profesor Asociado Tipo 1	2007	0,75
Sin Departamento	Profesor Asociado Tipo 3	2007	1
Sin Departamento	Titular Universidad Int.	2007	1
Sin Departamento	L.D. Profesor Asociado	2007	0
Sin Departamento	Profesor Asociado Tipo 1	2008	0,75

## **Diseño**

Una vez hecho el análisis de los datos que me proporcionó mi tutor para poder realizar el proyecto y sabiendo como están estructurados, comenzó una fase de diseño de la base de datos de destino.

Recuerdo que estos datos, que en este punto encontramos en unos archivos Excel estructurados, van a sufrir una serie de transformaciones y van a ser depositados en una base de datos especial llamada Data Warehouse.

Por eso en esta fase del proyecto, estructuré esta base de datos.

## **Data Warehouse**

Sabiendo el concepto de Data Warehouse y analizando la estructura que había generado en la sección anterior, se generó el Data Warehouse correspondiente:

**NOTA:** Se utilizarán dos prefijos para referirnos a los dos tipos de tablas que encontramos en un Data Warehouse.

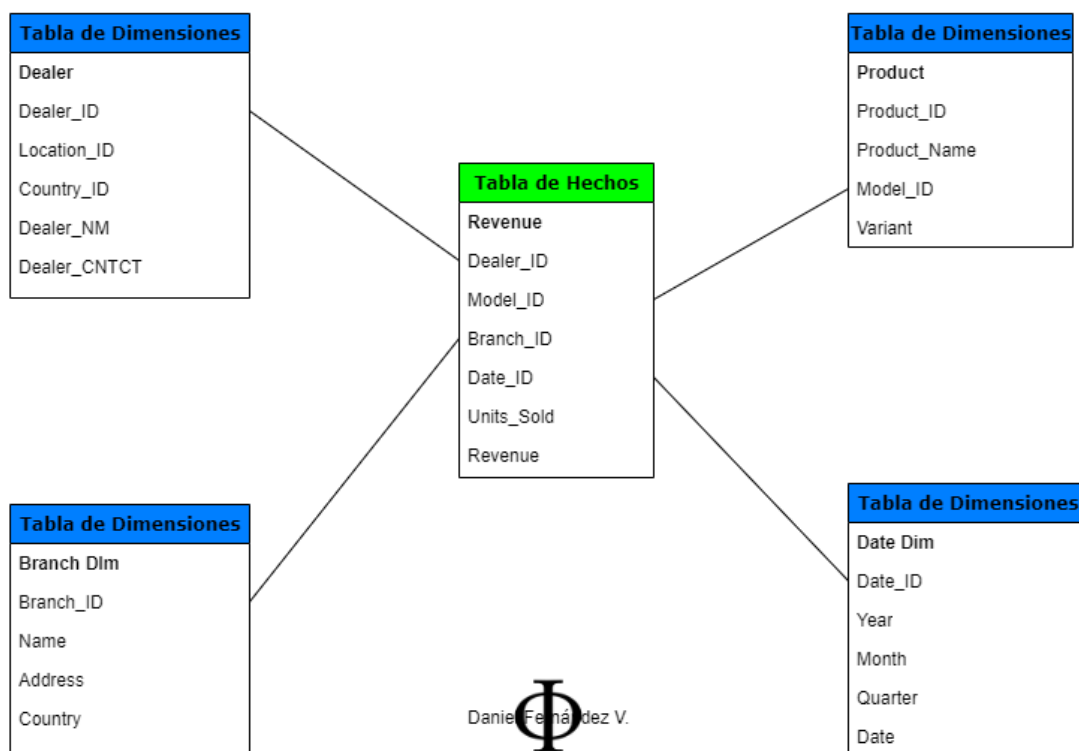
**Dim – Dimensión**

**Fact – Hechos**

Cabe mencionar también que este tipo de bases de datos se pueden estructurar de diferentes formas dependiendo del objetivo final que se quiera lograr con ella:

**-Estrella:** Este esquema está compuesto por una tabla de hechos, que será la tabla principal de la tabla, la cual contará con los datos necesarios para realizar cualquier tipo de cálculo de medidas para realizar el análisis y también tendrá una referencia a cada tabla de dimensión que haya en el modelo, las cuales tendrán información adicional que nos dará el contexto de la situación. Las tablas de dimensión no se conectan a nada que no sea la tabla de hechos.

Este esquema es el óptimo para herramientas de BI, además de ser el más sencillo y óptimo para el disco.

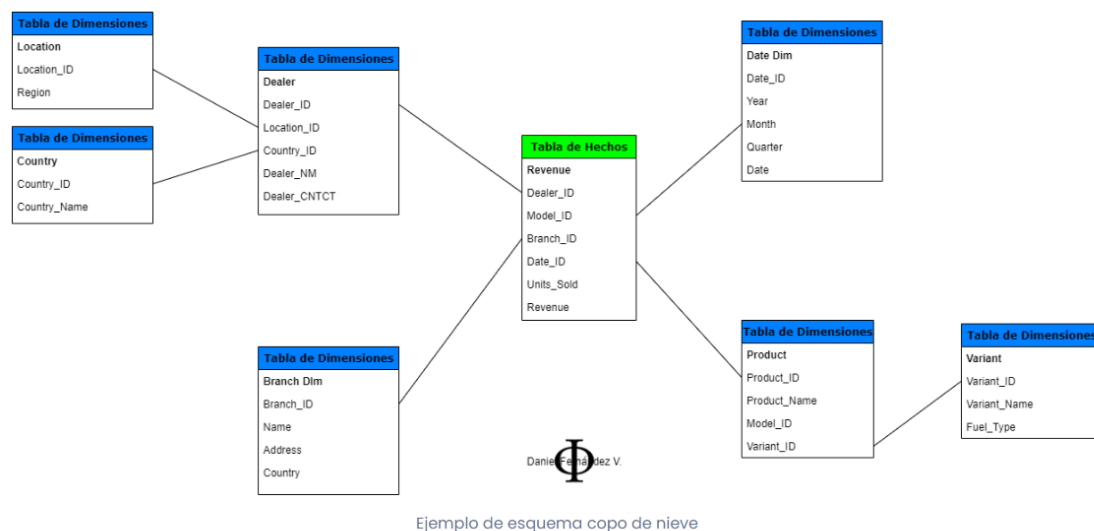


Ejemplo de esquema de estrella (Star)

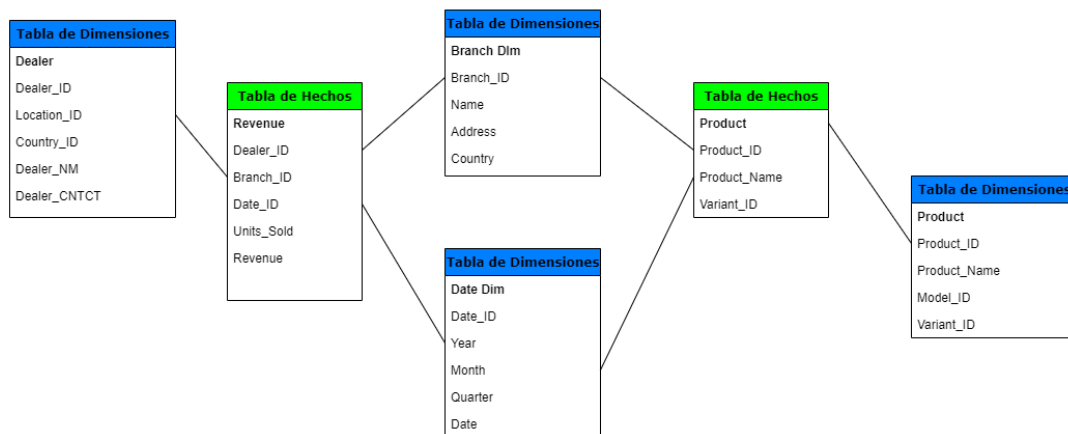


**-Copo de nieve:** Este esquema se parece bastante al esquema anterior, con la única diferencia que existen en él dimensiones que contienen referencias a otras dimensiones las cuales no se conectan directamente con la tabla de hechos.

Los datos en este esquema se encuentran normalizados, es por ello por lo que ocupa menos espacio en disco y se reduce así el tiempo de consulta. La única desventaja es que el trabajo de mantenimiento es mayor.



**-Constelación:** En ella hay dos tablas de hechos que comparten alguna tabla de dimensión. Esto nos permite en la herramienta BI poder filtrar por esa dimensión y que automáticamente ambas tablas estén filtradas. Tiene un mayor espacio en disco y puede llegar a dificultar las consultas y por ende el rendimiento.



Daniel Fernández V.

Ejemplo de esquema Galaxy

Este esquema es el que se terminó siguiendo ya que encontraremos una dimensión común entre las dos tablas de hechos que representan los sectores en los que se divide el cuadro de mando. En un principio iban a ser dos esquemas de estrella totalmente independiente uno del otro.

## Apartado Económico

### Dim\_concepto\_economico:

Esta tabla contiene los datos de los distintos conceptos/artículos en los que la escuela ha gastado o ingresado dinero. En caso de que en algún año posterior se realice un gasto o ingreso sobre un concepto nuevo, solamente será necesario incorporar toda su información necesaria en esta tabla y luego referenciarlo en la tabla de hechos.

Los atributos de esta tabla son los siguientes:

### Table: **dim\_concepto\_economico**

#### Columns:

<b>concepto_sk</b>	bigint PK
Concepto	tinytext
version	int
date_from	datetime
date_to	datetime

### **Dim\_tipo\_economico;**

Esta tabla contiene los dos tipos de datos que puede haber en la tabla económica: Gasto o Ingreso.

Los atributos de esta tabla son los siguientes:

### **Table: dim\_tipo\_economico**

#### **Columns:**

<u>tipo_sk</u>	bigint PK
Tipo	tinytext
version	int
date_from	datetime
date_to	datetime

### **Fact\_economico:**

Esta tabla posee los datos analíticos necesarios para posteriormente realizar cualquier estudio que se quiera realizar. Aparte de estos datos analíticos, también posee referencias a las tablas de dimensión que nos proporcionarán información extra sobre los datos para poder agruparlos o incluso filtrarlos.

Los atributos de esta tabla son los siguientes:

### **Table: fact\_economico**

#### **Columns:**

Id_economico	int
Articulo	int
Consolidado	tinytext
Importe	double
Anio	int
tipo_sk	bigint
concepto_sk	bigint

## Apartado Profesorado

### Dim\_categoria\_profesorado:

Esta tabla contiene los diferentes tipos de profesores que ha tenido la escuela a lo largo de los años 2009 y 2013. A través de esta podremos en la tabla de hechos hacer un análisis de cuantos profesores de un tipo hubo o incluso poder agruparlos por uno de estos tipos.

Los atributos de esta tabla son los siguientes:

### Table: `dim_categoria_profesorado`

#### Columns:

<code>categoria_sk</code>	<code>bigint PK</code>
<code>categoria</code>	<code>tinytext</code>
<code>version</code>	<code>int</code>
<code>date_from</code>	<code>datetime</code>
<code>date_to</code>	<code>datetime</code>

### **Dim\_departamento\_profesorado:**

Esta tabla contiene los diferentes departamentos que tiene la escuela, que nos servirán para posteriormente en la tabla de hechos poder analizar el profesorado respecto a cada departamento haciendo referencia desde esa tabla a esta.

Los atributos de esta tabla son los siguiente:

#### **Table: dim\_departamento\_profesorado**

##### **Columns:**

<b>departamento_sk</b>	bigint PK
departamento	tinytext
version	int
date_from	datetime
date_to	datetime

### **Fact\_profesorado:**

Esta tabla es la tabla principal de la sección de profesorado. En ella al igual que en la tabla de hechos de economía, encontramos los datos analíticos de la sección. Además, también tiene las referencias a las distintas tablas de dimensión para poder filtrar y agrupar por ellas.

Los atributos de esta tabla son los siguientes:

### **Table: fact\_profesorado**

#### **Columns:**

departamento_sk	bigint
categoria_sk	bigint
anio	double
cantidad	double



A parte de las dimensiones anteriormente mencionadas, en el modelo final de PowerBi encontramos una última dimensión común a ambas tablas de hechos: Dim\_ejercicio.

Esta dimensión servirá para poder filtrar ambas tablas de hechos por un año en concreto y así poder ver y analizar los datos de una forma más sencilla. Esta tabla, únicamente cuenta con una columna de sencilla. Esta tabla, únicamente cuenta con una columna de ejercicio en el que tiene los años entre 2000 y 2050 para poder incorporar sin problema los datos de esos años.

En caso de continuar con el proyecto, y en caso de que se quiera incorporar la información de algún año posterior a 2050, tendremos que añadir en esta tabla esos años.

## **Implementación**

Una vez puesto todo lo anterior en contexto comienza la parte de implementación, que como su propio nombre indica se trata de la parte de desarrollo del proyecto en la que se realiza todo lo anteriormente mencionado.

Para ello el orden de realización de las diferentes partes será el orden lógico que se tiene que seguir para poder obtener los resultados queridos. Es decir, nuestro objetivo es realizar un panel de mando sobre unos datos que nos llega en formato PDF.

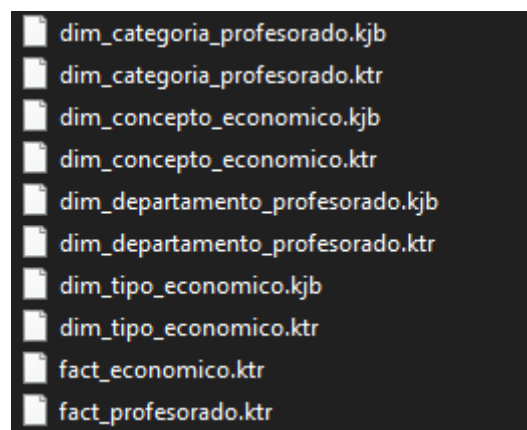
Estos datos fueron pasados a Excel para así poder incorporarlos en Pentaho y que nos resultara más fácil su manejo.

Una vez hecho eso pasarán por los procesos de transformación para adaptar los datos a nuestro modelo y para poder acomodarlos se depositarán en una base de datos propia.

Con ello en la base de datos se terminará por el desarrollo del cuadro de mando con una última herramienta.

## ETL

Como se acaba de comentar, tras la actualización de los datos a formato Excel, se realizaron 6 Transformaciones y 3 Jobs en la herramienta de Pentaho para así poder acomodar los datos y conformar el Data Warehouse.



Comenzaremos por la categoría de docencia y dentro de ella con las transformaciones de las tablas de dimensión.

Hay que destacar que lo más habitual es comenzar por las tablas de dimensión ya que la tabla de hechos está formada por referencias a las tablas de dimensión con las que tiene relación, es por ello por lo que, para poder coger esas referencias, se tienen que tener las tablas ya creadas y formadas. Dicho de otro modo y como veréis posteriormente, se necesitan las tablas de dimensión para crear la tabla de hechos.

## Dim\_categoria\_profesorado:



Esta transformación consta de 4 pasos:

- **Tab\_profesorado:** Se trata del paso de obtención de los datos del Excel para poder introducirlos en la herramienta. Aquí lo más importante es coger bien los datos de la hoja en la que se encuentran del Excel. En mi caso adicionalmente parametricé la ruta desde donde se coge el Excel para que en caso de que se cambie la ruta solamente sea necesario acceder al archivo de la parametrización y ahí cambiarlo, sin tener que tocar nada de la transformación.

#	Fichero/Directorio
1	<code>\${PATH_PRINCIPAL}\${EXCEL_PROFESORADO}.xlsx</code>

- **Seleccionar categoría:** En este caso, por el Excel, nos vienen todos los datos del Excel, algo que no necesitamos puesto que solo nos interesa el apartado de categoría, es por ello que este paso se encarga de seleccionar únicamente la columna de categoría que nos viene del Excel.

Selecciona/Renombra valores

Nombre paso:

Selecciona & Modifica | Eliminar | Meta-información

Campos:

#	Nombre campo	Renombrar a	Longitud	Precisión
1	categoria			

Obtener campos a seleccionar

Edit Mapping

Include unspecified fields, ordered by name ☐

Help Vale Cancelar

- **Filas únicas:** Tras el paso anterior, nos hemos quedado únicamente con la columna de categoría, pero como os podéis imaginar existen valores repetidos dentro de los datos que nos quedan. Es por ello por lo que este paso nos elimina este inconveniente quedándose una única referencia por categoría, por lo que ya tenemos las categorías que van a conformar la dimensión.

#	categoria
1	Profesor Asociado Tipo 1
2	Profesor Asociado Tipo 3
3	Titular Universidad Int.
4	L.D. Profesor Asociado
5	Profesor Asociado Tipo 1
6	Profesor Asociado Tipo 3
7	Titular Universidad Int.
8	L.D. Profesor Asociado
9	Profesor Asociado Tipo 1
10	Profesor Asociado Tipo 3
11	Titular Universidad Int.
12	L.D. Profesor Asociado

- **Dim\_categoria\_profesorado:** Y, con todo esto, llegamos al último paso de la transformación. Este último paso, como os podéis imaginar, se trata de cargar la tabla de dimensión.

Para ello, Pentaho trae consigo un paso que nos permite crearla y mantenerla actualizada en caso de que se quiera realizar más cargas. En ella puedes poner el nombre que quieras a la tabla, seleccionar los campos que tendrá y también puedes realizar un control de versiones con su rango de fechas y, lo más importante, puedes crear una clave que se vaya incrementando de forma automática para cada valor de la tabla.

De esta manera, podrás luego desde la tabla de hechos a la tabla de dimensión.

The screenshot shows the 'Actualizar la dimensión' (Update Dimension) step configuration in Pentaho Data Integration. The window is titled 'Esquema/actualización en dimensión'. The 'Nombre paso' (Step Name) is 'dim\_categoria\_profesorado'. The 'Actualizar la dimensión?' (Update dimension?) checkbox is checked. The 'Conexión' (Connection) is 'TFG'. The 'Esquema destino' (Destination Schema) is empty. The 'Tabla destino' (Destination Table) is 'dim\_categoria\_profesorado'. The 'Número de transacción' (Transaction Number) is '100'. The 'Enable the cache?' checkbox is checked. The 'Pre-load the cache?' checkbox is unchecked. The 'Cache size in rows (0 = cache all)' is '5000'. Below this, there is a 'Campos' (Fields) tab. Under 'Campos clave (para buscar fila en dimensión):' (Key Fields (to search for row in dimension)), there is a table with two columns: 'Campo de Dimensión' (Dimension Field) and 'Campo en flujo' (Field in flow). The first row has 'categoria' in both columns. Below the table, the 'Campo de clave técnica' (Technical Key Field) is 'categoria\_sk'. The 'Nuevo nombre' (New name) field is empty. Under 'Creación de clave técnica' (Technical Key Creation), the 'Utilizar máximo tabla + 1' (Use maximum table + 1) radio button is selected. Below this, the 'Campo de Versión' (Version Field) is 'version'. The 'Campo Fecha flujo' (Flow Date Field) is empty. The 'Campo inicio rango fecha' (Start Date Range Field) is 'date\_from'. The 'Año m. n.' (Month. Year) is '1900'. The 'Use an alternative start date?' checkbox is unchecked. The 'Campo final rango fecha' (End Date Range Field) is 'date\_to'. The 'Año m. n.' is '2199'. At the bottom, there are buttons: 'Vale' (OK), 'Cancelar' (Cancel), 'Obtener Campos' (Get Fields), and 'SQL'.

Cosa importante también que mencionar es la conexión a la base de datos ya que en este último paso es necesario ya tener creada la base de datos, aunque no las tablas ya que estas

DIEGO CRUZ CUERVO  
SISTEMA DE ANALISIS DE ECONOMÍA Y DOCENCIA DE LA ETSISI

también se pueden crear a medida pulsando en el botón inferior que pone SQL.

Con esto ya no solo se explica la dimensión de categoría sino que las otras 4 dimensiones que tiene el Data Warehouse también se realizan de la misma forma, añadiendo alguna de ellas un paso de ordenación que es opcional y cambiando claramente el paso de la selección o el paso del Excel en el que se recogen los datos económicos en vez de los de docencia.

	categoria_sk	categoria	version	date_from	date_to
►	0	NULL	1	NULL	NULL
	1	Profesor Asociado Tipo 1	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	2	Profesor Asociado Tipo 3	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	3	Titular Universidad Int.	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	4	L.D. Profesor Asociado	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	5	Titular E.U.	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	6	L.D. Ayudante	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	7	Titular Universidad	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	8	Catedratico universidad	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	9	Catedratico E.U.	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	10	Profesor Asociado Tipo 2	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	11	L.D. Profesor Colaborador	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	12	Titular E.U. Interino	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	13	L.D. Profesor Ayudante...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	14	L.D. Profesor Contratad...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	15	L.D. Profesor Contratad...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
✱	NULL	NULL	NULL	NULL	NULL

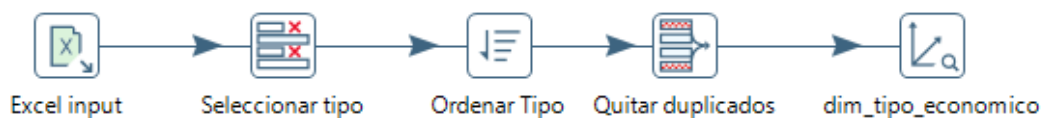


## Dim\_departamento\_profesorado:



	departamento_sk	departamento	version	date_from	date_to
▶	0	NULL	1	NULL	NULL
	1	Sin Departamento	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	2	Ingenieria de Organización, Administracion de E...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	3	Linguistica Aplicada a la Ciencia y a la Tecnologia	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	4	Tecnologia de la Edificacion	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	5	Informatica Aplicada	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	6	Lenguajes, Proyectos y Sistemas Informaticos	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	7	Matematica Aplicada (E.U. Informatica)	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	8	Organizacion y Estructura de la Informacion (Tit...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	9	Sistemas Inteligentes Aplicados	1	1900-01-01 00:00:00	2200-01-01 00:00:00
*	NULL	NULL	NULL	NULL	NULL

## Dim\_tipo\_economico:



	tipo_sk	Tipo	version	date_from	date_to
▶	0	NULL	1	NULL	NULL
	1	Gasto	1	1900-01-01 00:00:00	2200-01-01 00:00:00
	2	Ingreso	1	1900-01-01 00:00:00	2200-01-01 00:00:00
*	NULL	NULL	NULL	NULL	NULL

## Dim\_concepto\_economico:

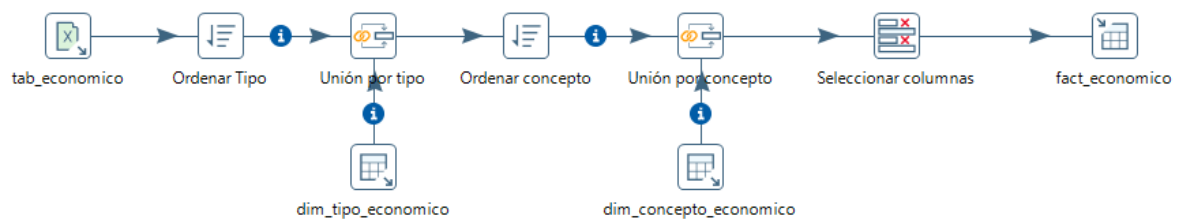


	concepto_sk	Concepto	version	date_from	date_to
▶ 0	NULL		1	NULL	NULL
1		221.00. Suministro de electricidad	1	1900-01-01 00:00:00	2200-01-01 00:00:00
2		221.01. Suministro de agua	1	1900-01-01 00:00:00	2200-01-01 00:00:00
3		221.02. Suministro de gas	1	1900-01-01 00:00:00	2200-01-01 00:00:00
4		226.01. Atenciones protocolarias	1	1900-01-01 00:00:00	2200-01-01 00:00:00
5		480. Becas	1	1900-01-01 00:00:00	2200-01-01 00:00:00
6		481. Subvenciones, premios	1	1900-01-01 00:00:00	2200-01-01 00:00:00
7		Arrendamientos y cánones	1	1900-01-01 00:00:00	2200-01-01 00:00:00
8		Asignación UPM	1	1900-01-01 00:00:00	2200-01-01 00:00:00
9		Cesiones de locales	1	1900-01-01 00:00:00	2200-01-01 00:00:00
10		Cánones	1	1900-01-01 00:00:00	2200-01-01 00:00:00
11		Cánones OTT	1	1900-01-01 00:00:00	2200-01-01 00:00:00
12		Gastos corrientes en bienes y se...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
13		Gastos de personal	1	1900-01-01 00:00:00	2200-01-01 00:00:00
14		Gastos OTT	1	1900-01-01 00:00:00	2200-01-01 00:00:00
15		Indemnizaciones (sinistros, robos)	1	1900-01-01 00:00:00	2200-01-01 00:00:00
16		Indemnizaciones por razón del S...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
17		Inversiones	1	1900-01-01 00:00:00	2200-01-01 00:00:00
18		Inversiones reales	1	1900-01-01 00:00:00	2200-01-01 00:00:00
19		Matrícula	1	1900-01-01 00:00:00	2200-01-01 00:00:00
20		Nominativa	1	1900-01-01 00:00:00	2200-01-01 00:00:00
21		Otros ingresos	1	1900-01-01 00:00:00	2200-01-01 00:00:00
22		Reparación, mantenimiento y co...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
23		Resto articulo 22 (Material, Sumi...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
24		Transferencias y subvenciones c...	1	1900-01-01 00:00:00	2200-01-01 00:00:00
*	NULL	NULL	NULL	NULL	NULL

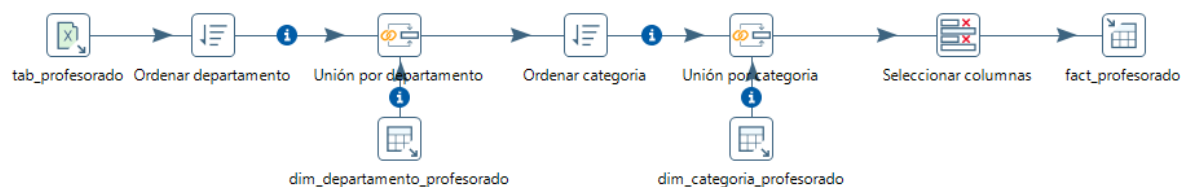
Una vez creadas las tablas de dimensión como se ha comentado anteriormente se realizaron las tablas de hechos que contienen referencias a las tablas de dimensión con las que están relacionadas, así como los datos que van a servir para realizar los distintos cálculos para hacer el análisis querido.

Aquí encontramos 2 tablas de hechos que en su construcción se parecen bastante:

### Fact\_economico:



### Fact\_profesorado:



El flujo de esta transformación es sencillo:

Comienza con el paso de extracción de datos de los excels en los que se encuentran la información para cada una de las categorías. Como se ha comentado anteriormente se encuentran parametrizadas para así poder modificarlo más fácilmente en caso de querer cambiar la zona en la que se quiera depositar.

#	Fichero/Directorio
1	<code>\${PATH_PRINCIPAL}\${EXCEL_PROFESORADO}.xlsx</code>

#	Fichero/Directorio
1	<code>\${PATH_PRINCIPAL}\${EXCEL_ECONOMICO}.xlsx</code>

Una vez tenemos los datos en el flujo, vienen los pasos necesarios para poder hacer referencias desde la tabla de hechos a las tablas de dimensión con las que están relacionadas cada una.

Para ello hay que tener en cuenta que existe un campo común entre la tabla de hechos y las tablas de dimensión. Por ese campo se realizará una unión entre las tablas teniendo como resultado, que en cada fila existirá tanto la descripción de la dimensión como su referencia:

[illegible]

Execution Results									
<div> <div>Logging</div> <div>Execution History</div> <div>Step Metrics</div> <div>Performance Graph</div> <div>Metrics</div> <div>Preview data</div> </div>									
<div> <div> <div> <div></div> <div> <div> <div> <div></div> <div></div> <div></div> </div> </div> <div> <div> <div></div> <div></div> <div></div> </div> </div> <div> <div> <div></div> <div></div> <div></div> </div> </div> <div> <div> <div></div> <div></div> <div></div> </div> </div> </div> </div> <div> <div> <div></div> <div></div> <div></div> </div> </div> <div> <div> <div></div> <div></div> <div></div> </div> </div> </div> </div>									
#	Id_economico	Tipo	Articulo	Concepto	Consolidado	Importe	Anio	tipo_sk	Tipo_1
1	1	Gasto	20	Arrendamientos y cánones	c	10500,0	2022	1	Gasto
2	2	Gasto	20	Arrendamientos y cánones	n	5000,0	2022	1	Gasto
3	3	Gasto	21	Reparación, mantenimiento y conservación	c	20000,0	2022	1	Gasto
4	4	Gasto	21	Reparación, mantenimiento y conservación	n	10000,0	2022	1	Gasto
5	5	Gasto	22	221.00. Suministro de electricidad	c	80000,0	2022	1	Gasto
6	6	Gasto	22	221.00. Suministro de electricidad	n	12000,0	2022	1	Gasto
7	7	Gasto	22	221.01. Suministro de agua	c	9000,0	2022	1	Gasto
8	8	Gasto	22	221.01. Suministro de agua	n	4000,0	2022	1	Gasto
9	9	Gasto	22	221.02. Suministro de gas	c	50000,0	2022	1	Gasto
10	10	Gasto	22	221.02. Suministro de gas	n	8000,0	2022	1	Gasto
11	11	Gasto	22	226.01. Atenciones protocolarias	c	5000,0	2022	1	Gasto
12	12	Gasto	22	226.01. Atenciones protocolarias	n	4000,0	2022	1	Gasto
13	13	Gasto	22	Resto articulo 22 (Material, Suministros y Otros)	c	115500,0	2022	1	Gasto

(Las capturas anteriores son un ejemplo con la dimensión del tipo económico, esto se repite para las demás dimensiones).

Como podéis observar, se nos quedan ciertos datos los cuales son repetidos o innecesarios. Estamos consiguiendo la referencia a la tabla de dimensión para no sobrecargar la tabla de hechos con datos los cuales son simplemente de contexto.

Es por ello por lo que el siguiente paso a realizar es una selección de campos de la tabla para poder eliminar las descripciones que están guardadas ya en las tablas de dimensión.

#	Nombre campo	#	Nombre campo
1	Id_economico	1	departamento_sk
2	tipo_sk	2	categoria_sk
3	Articulo	3	anio
4	concepto_sk	4	cantidad
5	Consolidado		
6	Importe		
7	Anio		

Con ello ya tendríamos formadas las tablas de hechos asique lo último que nos quedaría para completar la tabla de hechos es el paso de carga en el Data Warehouse.

Para ello, al igual que en las tablas de dimensión, uniendo el flujo con el paso final y dándole internamente en ese paso al botón de SQL, se nos creará la tabla con la estructura que se necesita para así solo tener que terminar de configurar el paso y poder ejecutarlo.

**Salida de Tabla**

Nombre paso:  fact\_profesorado

Conexión: TFG [v] [Editar...] [Nuevo...] [Wizard...]

Esquema destino:  [Examinar...]

Tabla destino: fact\_profesorado [Examinar...]

Tamaño transacción (commit): 1000

Vaciar tabla: ☒

Ignorar errores de inserción: ☐

Especificar campos de base de datos: ☒

---

Main options Database fields

Fields to insert:

#	Table field	Stream field
1	departame...	departament...
2	categoría_sk	categoría_sk
3	año	año
4	cantidad	cantidad

[Get fields]  
[Enter field mapping]

[?] Help   [Vale]   [Cancelar]   [SQL]

Con ello ejecutado, ya tendríamos el Data Warehouse creado y con los datos ya transferidos desde el archivo Excel.

## Fact\_economico:

	Id_economico	Articulo	Consolidado	Importe	Anio	tipo_sk	concepto_sk
►	5	22	c	80000	2022	1	1
	6	22	n	12000	2022	1	1
	29	22	c	80000	2023	1	1
	30	22	n	12000	2023	1	1
	7	22	c	9000	2022	1	2
	8	22	n	4000	2022	1	2
	31	22	c	9000	2023	1	2
	32	22	n	4000	2023	1	2
	9	22	c	50000	2022	1	3
	10	22	n	8000	2022	1	3
	33	22	c	55000	2023	1	3
	34	22	n	8000	2023	1	3
	11	22	c	5000	2022	1	4
	12	22	n	4000	2022	1	4
	35	22	c	5000	2023	1	4
	36	22	n	4000	2023	1	4
	41	4	c	4000	2023	1	5
	42	4	n	1000	2023	1	5
	17	4	c	100	2022	1	6
	18	4	n	400	2022	1	6
	43	4	c	100	2023	1	6
	44	4	n	400	2023	1	6
	1	20	c	10500	2022	1	7
	2	20	n	5000	2022	1	7
	25	20	c	10500	2023	1	7
	26	20	n	5000	2023	1	7



## Fact\_profesorado:

	departamento_sk	categoria_sk	anio	cantidad
►	5	9	2006	1
	5	9	2007	0
	5	9	2008	0
	5	9	2009	0
	5	9	2010	0
	5	9	2011	0
	5	9	2012	0
	5	9	2013	0
	8	9	2006	6
	8	9	2007	6
	8	9	2008	5
	8	9	2009	5
	8	9	2010	5
	8	9	2011	4
	8	9	2012	4
	8	9	2013	4
	9	9	2006	0
	9	9	2007	1
	9	9	2008	1
	9	9	2009	1
	9	9	2010	1
	9	9	2011	1
	9	9	2012	1
	9	9	2013	1
	5	8	2006	0
	5	8	2007	0

[TO-DO -> HABLAR DE LOS JOBS CUANDO LOS TENGAS]

## BBDD

Con todo lo anterior realizado, los datos son depositados en nuestro Data Warehouse, base de datos creada para almacenar los datos que van a ser usados para el análisis.

Este Data Warehouse en principio contará con las tablas que se han creado en las transformaciones anteriores, pero se podría ir ampliando con las diferentes secciones que se quieran ir añadiendo a la herramienta final.

Ahora que tenemos los datos en el Data Warehouse, podemos validar que todo este acorde a lo planeado mediante una herramienta de administración de bases de datos como es, en este caso, MySQL Workbench.

Con esta aplicación se podrán realizar las consultas para poder observar los datos que han sido incorporados en las tablas en su totalidad:

```
SELECT *  
FROM [insertar nombre tabla]
```

## **Cuadro de mando**

### **Definición**

Un cuadro de mando de Power BI es una herramienta que permite el diseño, la visualización y el análisis de datos clave en tiempo real para la toma de decisiones empresariales. Este cuadro de mando se compone de diversos gráficos, tablas, mapas y otros elementos visuales que enseñan la información más relevante y permiten a los usuarios explorar los datos mediante filtros y segmentaciones.

El objetivo principal de un cuadro de mando es proporcionar una visión clara y concisa de las métricas importantes de una organización, facilitando el monitoreo de los datos y la identificación de tendencias y anomalías.

Power BI, desarrollado por Microsoft, ofrece capacidades avanzadas de conectividad a diversas fuentes de datos, transformación y modelado de datos, y creación de visualizaciones personalizadas, lo que permite a los usuarios construir cuadros de mando altamente efectivos y adaptados a sus necesidades específicas.

## **Diseño de un Cuadro de Mando en Power BI**

El diseño de un cuadro de mando en Power BI implica varios pasos clave que aseguran que la herramienta sea efectiva, intuitiva y visualmente atractiva. A continuación, se detallan los componentes y consideraciones principales para un diseño óptimo:

### **1. Identificación de Objetivos y Usuarios**

#### **Objetivos:**

- Determinar los objetivos específicos del cuadro de mando, tales como monitoreo de las medidas, análisis financiero, control de docentes, etc.
- Definir las preguntas clave que el cuadro de mando debe responder.

#### **Usuarios:**

- Identificar a los usuarios finales para así comprender sus necesidades y preferencias.
- Asegurarse de que el nivel de detalle y la complejidad del cuadro de mando sea adecuado para su audiencia.

### **2. Selección de Datos y Medidas**

#### **Fuentes de Datos:**

- Conectar Power BI al Data Warehouse empresarial.
- Asegurar la seguridad de accesos para prevenir futuros problemas.

### **Medidas:**

- Seleccionar los datos más relevantes para los objetivos definidos por el usuario.
- Definir métricas claras y calculables que reflejen la evolución y progreso de los datos.

## **3. Diseño Visual**

### **Diseño y Navegación:**

- Organizar el cuadro de mando de manera lógica y coherente, con una estructura clara que guíe al usuario.
- Utilizar pestañas y enlaces para permitir una navegación fluida entre diferentes vistas o niveles de detalle.

### **Elementos Visuales:**

- Elegir los tipos de visualización adecuados (gráficos de barras, líneas, mapas, tablas, etc.) según la naturaleza de los datos y la información que se desea resaltar.
- Asegurar que los gráficos sean fáciles de interpretar, evitando la sobrecarga de información y priorizando la claridad.

### **Estilo:**

- Mantener una paleta de colores coherente y utilizar colores con propósito (por ejemplo, verde para la sección económica y rosado para la parte de profesorado).
- Aplicar formatos de número, fechas y texto de manera consistente.

## **4. Segmentación y Detalle**

### **Filtros y Segmentaciones:**

- Incluir opciones de filtrado y segmentación que permitan a los usuarios explorar los datos de diferentes maneras (por tipo de financiación, período de tiempo, partida, etc.).

### **Vista de detalle:**

- Implementar una vista de detalle para que los usuarios puedan profundizar en los datos y analizar la procedencia de las métricas.

## **5. Pruebas y Refinamiento**

### **Feedback:**

- Realizar pruebas para recoger feedback sobre la usabilidad y efectividad del cuadro de mando con diferentes usuarios.
- Ajustar el diseño y funcionalidad en base a las observaciones y sugerencias recibidas.

### **Rendimiento:**

- Optimizar las consultas y modelos de datos para asegurar un rendimiento eficiente y tiempos de carga rápidos.
- Revisar y depurar cualquier problema de rendimiento o errores en los datos visualizados.

## **6. Implementación y Mantenimiento**

### **Publicación:**

- Publicar el cuadro de mando en Power BI Service y configurar las opciones de seguridad y acceso según las necesidades de la organización.

- Asegurar que los usuarios tienen el acceso adecuado y que el cuadro de mando está actualizado regularmente.

**Mantenimiento:**

- Establecer un plan de mantenimiento para revisar y actualizar el cuadro de mando periódicamente, asegurando que los datos y medidas reflejen los datos correctamente.
- Incorporar mejoras y nuevas funcionalidades según sea necesario para mantener la relevancia y utilidad del cuadro de mando.

Un cuadro de mando bien diseñado en Power BI no solo proporciona una visión clara y concisa del desempeño organizacional, sino que también empodera a los usuarios para tomar decisiones informadas y basadas en datos.

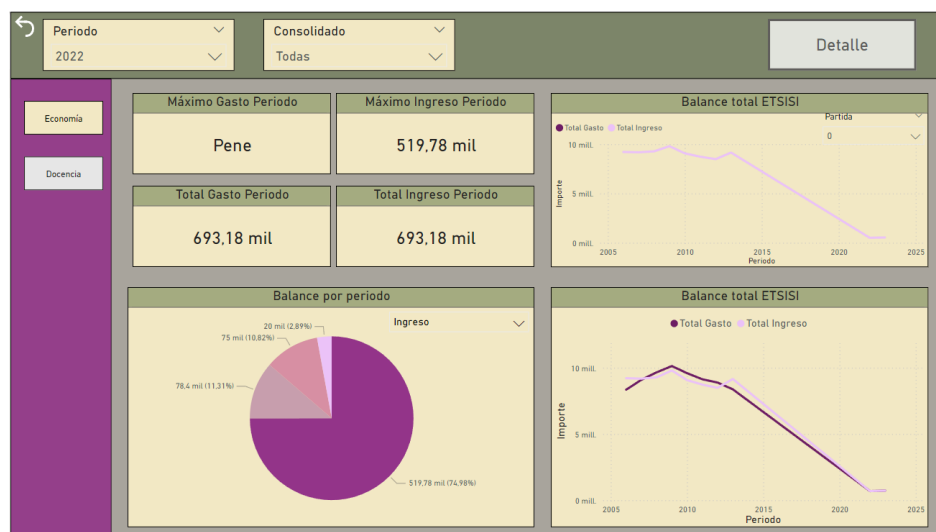
Dicho todo esto, lo último que falta por mencionar es que al final Power BI es una herramienta con gran parte de diseño la cual concede al usuario una gran libertad de decisiones para poder formar un cuadro de mando único.

## RESULTADOS

- **Inicio:** La primera vista es la portada del cuadro de mando. Además, sirve de página principal en el que seleccionar a que categoría de los datos ir en la zona superior.



- **Economía:** Esta vista es la página principal del apartado de economía del cuadro de mando. Contiene una serie de visualizaciones que muestra la variación de gasto e ingreso económico que ha tenido la ETSISI a lo largo de los años.

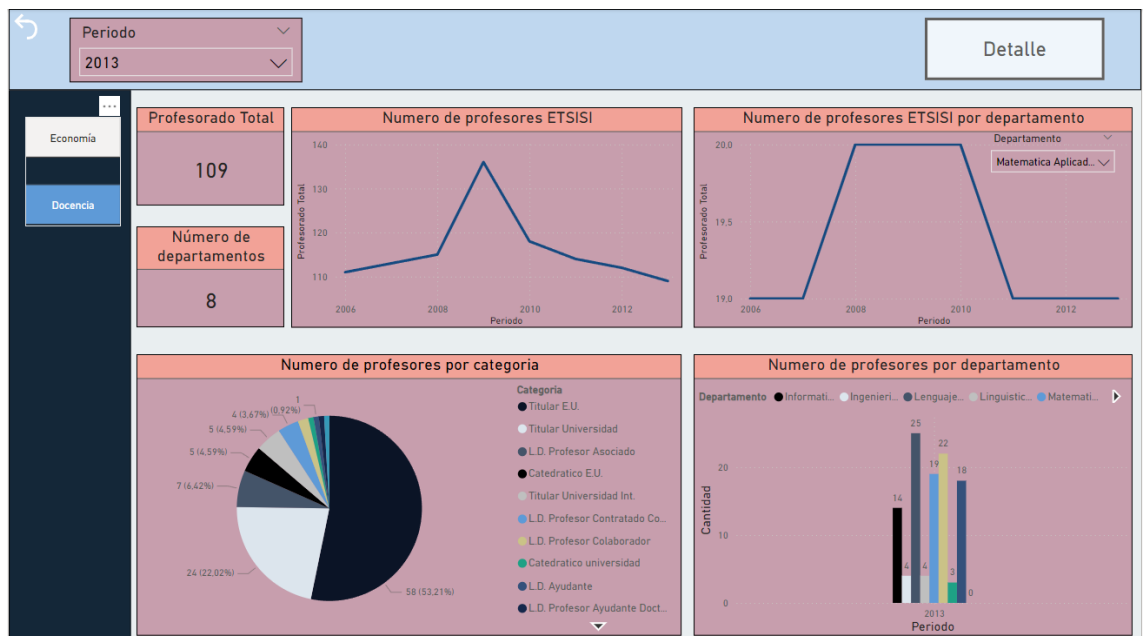




- **Economía – Detalle:** Para poder observar de donde proceden los datos mostrados en la vista anterior, se ha implementado esta vista con la que puedes ir al detalle de la economía que ha tenido la escuela.

Periodo	Concepto	Artículo	Tipo	Consolidado	Importe
2022	221.00. Suministro de electricidad	22	Gasto	c	80.000,00
2022	221.00. Suministro de electricidad	22	Gasto	n	12.000,00
2022	221.01. Suministro de agua	22	Gasto	c	9.000,00
2022	221.01. Suministro de agua	22	Gasto	n	4.000,00
2022	221.02. Suministro de gas	22	Gasto	c	50.000,00
2022	221.02. Suministro de gas	22	Gasto	n	8.000,00
2022	226.01. Atenciones protocolarias	22	Gasto	c	5.000,00
2022	226.01. Atenciones protocolarias	22	Gasto	n	4.000,00
2022	481. Subvenciones, premios	4	Gasto	c	100,00
2022	481. Subvenciones, premios	4	Gasto	n	400,00
2022	Arrendamientos y cánones	20	Gasto	c	10.500,00
2022	Arrendamientos y cánones	20	Gasto	n	5.000,00
2022	Asignación UPM	0	Ingreso	c	519.777,83
2022	Cánones	32	Ingreso	n	75.000,00
2022	Cesiones de locales	55	Ingreso	n	78.400,00
2022	Indemnizaciones (siniestros, robos)	39	Ingreso	n	20.000,00
2022	Indemnizaciones por razón del Servicio	23	Gasto	c	2.000,00
2022	Indemnizaciones por razón del Servicio	23	Gasto	n	2.000,00
2022	Inversiones	6	Gasto	c	227.677,83
2022	Inversiones	6	Gasto	n	100.000,00
2022	Reparación, mantenimiento y conservación	21	Gasto	c	20.000,00
2022	Reparación, mantenimiento y conservación	21	Gasto	n	10.000,00

- **Docencia:** La vista general de la categoría de docencia. En ella se puede observar la variación de profesorado a través del tiempo, así como su variación dentro de un departamento o la distribución de categoría de docencia por año.



DIEGO CRUZ CUERVO  
SISTEMA DE ANALISIS DE ECONOMÍA Y DOCENCIA DE LA ETSISI

- **Docencia – Detalle:** Esta vista similar a la de “Economía – Detalle”, muestra los datos de los que parte la vista anterior para que, en caso de encontrar una anomalía o algo destacado, se pueda ver que datos influyen en ello.

Periodo		General	Departamento	Categoria	Profesorado Total
2013					
Periodo	Departamento	Categoria			Profesorado Total
2013	Informatica Aplicada	Catedratico E.U.			0
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	Catedratico E.U.			4
2013	Sistemas Inteligentes Aplicados	Catedratico E.U.			1
2013	Informatica Aplicada	Catedratico universidad			1
2013	Lenguajes, Proyectos y Sistemas Informaticos	L.D. Ayudante			1
2013	Linguistica Aplicada a la Ciencia y a la Tecnologia	L.D. Ayudante			0
2013	Informatica Aplicada	L.D. Profesor Asociado			0
2013	Ingenieria de Organización, Administracion de Empresas y Estadistica	L.D. Profesor Asociado			0
2013	Lenguajes, Proyectos y Sistemas Informaticos	L.D. Profesor Asociado			3
2013	Linguistica Aplicada a la Ciencia y a la Tecnologia	L.D. Profesor Asociado			0
2013	Matematica Aplicada (E.U. Informatica)	L.D. Profesor Asociado			1
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	L.D. Profesor Asociado			1
2013	Sin Departamento	L.D. Profesor Asociado			1
2013	Sistemas Inteligentes Aplicados	L.D. Profesor Asociado			1
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	L.D. Profesor Ayudante Doctor			1
2013	Lenguajes, Proyectos y Sistemas Informaticos	L.D. Profesor Colaborador			2
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	L.D. Profesor Contratado Coctor			4
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	L.D. Profesor Contratado Coctor O.A.			1
2013	Linguistica Aplicada a la Ciencia y a la Tecnologia	Profesor Asociado Tipo 1			0
2013	Organizacion y Estructura de la Informacion (Titulo Provisional)	Profesor Asociado Tipo 1			0
2013	Sin Departamento	Profesor Asociado Tipo 1			0
2013	Sistemas Inteligentes Aplicados	Profesor Asociado Tipo 1			0
2013	Lenguajes, Proyectos y Sistemas Informaticos	Profesor Asociado Tipo 2			0
2013	Sistemas Inteligentes Aplicados	Profesor Asociado Tipo 2			0

## CONCLUSIONES

## BIBLIOGRAFÍA

Concepto de business intelligence : <https://www.tableau.com/es-es/learn/articles/business-intelligence>

Mysql: [https://www.arsys.es/blog/mysql#Que\\_es\\_MySQL](https://www.arsys.es/blog/mysql#Que_es_MySQL)

PowerBi: <https://learn.microsoft.com/es-es/power-bi/fundamentals/power-bi-overview>

Esquemas Data Warehouse: <https://datamanagement.es/2020/04/03/esquemas-data-warehousing/>

## ARCHIVOS