



Universidad de Valladolid



ESCUELA DE INGENIERÍAS  
INDUSTRIALES

UNIVERSIDAD DE VALLADOLID

ESCUELA DE INGENIERIAS INDUSTRIALES

GRADO EN INGENIERÍA EN ORGANIZACIÓN INDUSTRIAL

**APLICACIÓN DE BUSINESS  
INTELLIGENCE EN UNA PEQUEÑA  
EMPRESA MEDIANTE EL USO DE  
POWER BI**

Autor:

Rivera Resina, Fernando Javier

Tutor:

López Paredes, Adolfo  
Departamento de Organización de  
Empresas y Comercialización e  
Investigación de Mercados.

Valladolid, septiembre, 2018.





*“El conocimiento es poder”*

*Sir Francis Bacon*





# RESUMEN

El propósito de este documento es mostrar como una pequeña empresa puede obtener información de los datos que generan sus actividades utilizando herramientas de ofimática sin necesidad de tener altos conocimientos de informática o de programación.

Para ello llevo a cabo un ejercicio práctico gracias a la colaboración con una imprenta de Valladolid. En este caso práctico se indicarán las diferentes tareas necesarias para el tratamiento de datos, así como distintas formas de gestión de dichos datos utilizando en todos los casos el software Power BI.

En el documento también presento el mundo del Business Intelligence haciendo referencia a sus componentes principales y a la repercusión que tiene la utilización de este tipo de herramientas en una empresa.

## **Palabras clave:**

Business Intelligence, pequeñas empresas, PyMEs, datos, Power BI.





# ABSTRACT

The purpose of this document is to show how a small business can obtain information from the data generated by its activities using office automation tools without the need for a high level of computer or programming knowledge.

To do this I carry out a practical exercise thanks to the collaboration with a printing house in Valladolid. In this case study, the different tasks required for data processing will be indicated, as well as different ways of managing this data using Power BI software in all cases.

In the document I also present the world of Business Intelligence by referring to its main components and to the repercussion that the use of this type of tools has on a company.

## **Keywords:**

Business Intelligence, small business, SME, Dates, Power BI.





# ÍNDICE

RESUMEN .....	5
Palabras clave:.....	5
ABSTRACT.....	7
Keywords:.....	7
ÍNDICE DE ACRONIMOS.....	13
ÍNDICE DE ILUSTRACIONES.....	15
1. INTRODUCCIÓN.....	21
1.1. Contexto y justificación .....	21
1.2. Objetivo .....	21
1.3. Metodología .....	22
1.4. Organización del documento.....	22
2. BUSINESS INTELLIGENCE .....	23
2.1. Nacimiento y evolución del Business Intelligence .....	23
2.2. Qué es el BI.....	24
2.3. Datos, información y conocimiento.....	25
2.3.1. Datos.....	25
2.3.2. Información .....	26
2.3.3. Conocimiento .....	26
2.4. Componentes del BI.....	27
2.4.1. Fuentes de datos .....	27
2.4.2. ETL .....	28
2.4.3. Datawarehouse.....	28
2.4.4. Datamart .....	29
2.4.5. Dashboard .....	31
2.4.6. Cuadro integral de mando.....	32
2.4.7. Minería de datos.....	33
2.4.8. Reporting.....	33
2.5. Metodología para el diseño e implantación de un sistema BI.....	34
2.5.1. Aspectos Fundamentales para la Implantación de un Sistema BI.....	36
2.5.2. Herramientas necesarias para la implantación completa de BI .....	36
2.5.3. Fase 1 - Determinación de Requerimientos .....	37

2.5.4. Fase 2 - Estrategia de Proyecto .....	39
2.5.5. Fase 3 - Selección de la Tecnología .....	40
2.5.6. Fase 4 - Determinación del Modelo de Información .....	40
2.5.7. Fase 5 - Modelado .....	40
2.5.8. Fase 6 - Procesamientos de Datos.....	43
2.5.9. Fase 7 - Análisis y Producción.....	43
2.5.10. Fase 8 - Difusión:.....	43
3. BI EN LAS EMPRESAS .....	45
3.1. Indicadores de rendimiento .....	45
3.2. BI en Pymes.....	46
3.2.2 Las PyMEs y su proceso de toma de decisiones .....	50
3.3. Encuesta .....	51
3.4. Software para aplicar BI en una empresa .....	55
3.5. ROI con la implantación de herramientas de BI .....	64
3.6. 10 beneficios BI para una empresa .....	67
4. LA EMPRESA Y LOS DATOS .....	71
4.1. La empresa.....	71
4.1.1. Reuniones con la empresa.....	72
4.1.2. División de la empresa .....	73
4.1.3. Como trabaja la empresa.....	74
4.1.4. Programas que utiliza la empresa .....	75
4.2. Los datos .....	75
4.2.1. Datos facilitados por la empresa.....	76
4.2.2. Limpieza de datos.....	80
4.2.3. Obtención de más datos con valor .....	87
4.2.4 Diseño de estructuras de las bases de datos.....	90
5. IMPLANTACIÓN DE BI EN LA EMPRESA.....	95
5.1. Programas utilizados .....	96
5.2. Consideraciones previas a tener en cuenta.....	97
5.3. Implantación en la planta 0 .....	97
5.3.1. Mediante la utilización de Access y Power BI. ....	97
5.3.2. Mediante MySQL Workbench y Power Bi .....	100
5.4. Implantación en planta 1.....	106
5.4.1. Implantación cargando los datos directamente en Bi.....	106
5.5. Utilización del Power Bi.....	107



5.5.1. Presentación del panel general .....	107
5.5.2. Creación de gráficas .....	108
5.5.3. Resultados del uso del Bi.....	108
5.5.4. Recomendaciones para la empresa.....	110
6. CONCLUSIONES Y EXTENSIONES FUTURAS.....	111
7. BIBLIOGRAFÍA .....	113
8. ANEXO: MANUAL.....	119
8.1. Instalación de programas necesarios.....	119
8.2. Implantación utilizando solo el Power Bi.....	119
8.2.1. Creación tablas con Excel .....	119
8.2.2. Cargar datos a Power Bi.....	120
8.2.3. Creación de relaciones .....	120
8.3. Implantación utilizando Access y Power Bi.....	120
8.3.1. Creación base de datos y carga de datos .....	120
8.3.2. Carga base de datos a Power Bi.....	124
8.4. Implantación utilizando MySQL y Power Bi.....	124
8.4.1. Creación de las tablas .....	124
8.5. Utilización de Power Bi.....	129





# ÍNDICE DE ACRONIMOS

- BI: Business Intelligence.
- OLTP: OnLine Transaction Processing.
- OLAP: OnLine Analytical Procesising.
- ETL: Extract, Transform and Load.
- PyME: Pequeña y Mediana Empresa.
- BBDD: Bases de Datos.
- KPI: Key Performance Indicator.
- BSC: Balance ScoreCard.
- CMI: Cuadro de Mando Integral.
- ROI: Retorno de la inversión.
- DW: Datawarehouse.





# ÍNDICE DE ILUSTRACIONES

**Ilustración 2.1. Diagrama de relación entre datos, información y conocimiento.** Figura tomada de [https://www.sinnexus.com/business\\_intelligence/](https://www.sinnexus.com/business_intelligence/).....21

**Ilustración 2.2. Diagrama de componentes de BI.** Ilustración de creación propia basada en <http://www.diazepamresource.com/real-time-data-warehouse-B9573/fantastic-real-time-data-warehouse-j23-on-wonderful-home-design-ideas-with-real-time-data-warehouse/>.....23

**Ilustración 2.3. Fuentes de datos – información – herramientas de análisis**  
Ilustración tomada de [http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529\\_esa.pdf](http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf).....25

**Ilustración 2.4. OLAP frente OLTP** Ilustración de creación propia basada de una encontrada en <https://jorsodom.wordpress.com/2013/09/19/cuales-la-diferencia-entre-una-base-de-datos-oltp-y-olap/>.....27

**Ilustración 2.5. Dashboard Power Bi.** Ilustración tomada de <https://denglishbi.wordpress.com/2018/02/20/how-to-build-a-sql-server-dashboard-with-power-bi/>.....28

**Ilustración 2.6. Diagrama explicativo de los de CMI,** Ilustración tomada de: [https://www.sinnexus.com/business\\_intelligence/cuadro\\_mando\\_integral.aspx](https://www.sinnexus.com/business_intelligence/cuadro_mando_integral.aspx).....28

**Ilustración 2.7. Elemento tecnología BI.** Ilustración de creación propia inspirada en la web: <https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/> .....32

**Ilustración 2.8. Determinación de los requerimientos del BI:** Ilustración tomada de: <https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/>..33

**Ilustración 2.9. Modelo estrella.** Fuente: BALLARD, Chuck; HERREMAN, Dirk; SCHAU, Don; BELL, Rhonda; KIM, Eunsang y VALENCIC, Ann. Data Modeling Techniques for Data Warehousing. IBM’s International Technical Support Organization. p. 47. Febrero, 1998. ....37

<b>Ilustración 2.10. Modelo de copo de nieve. .</b> Fuente: BALLARD, Chuck; HERREMAN, Dirk; SCHAU, Don; BELL, Rhonda; KIM, Eunsang y VALENCIC, Ann. Data Modeling Techniques for Data Warehousing. IBM's International Technical Support Organization. p. 47. Febrero, 1998. ....	38
<b>Ilustración 2.11. Aspectos Fundamentales para la Implantación de un Sistema de BI:</b> Ilustración tomada de: <a href="https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/">https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/</a> .....	40
<b>Ilustración 3.1. Prioridades tecnológicas de las PyMEs frente a las grandes empresas</b> Ilustración tomada de: DRESNER. Small and Mid-Sized Enterprise Business Intelligence Market Study. p. 21. Dresner Advisory Services LLC, 2014. ....	44
<b>Ilustración 3.2. Resultados a la pregunta sobre el almacenamiento de datos.</b> Resultados encuesta <a href="https://docs.google.com/forms/d/1GoXLiMhhUzabHqvDWZPI1gjMXxYjBZZxkmsre4E2JY/edit">https://docs.google.com/forms/d/1GoXLiMhhUzabHqvDWZPI1gjMXxYjBZZxkmsre4E2JY/edit</a> .....	50
<b>Ilustración 3.3. Logo IBM Cognos.</b> <a href="https://www.ibm.com/es-s/marketplace/business-intelligence">https://www.ibm.com/es-s/marketplace/business-intelligence</a> .....	51
<b>Ilustración 3.4 Logo Tableau</b> <a href="https://www.tableau.com/es-es">https://www.tableau.com/es-es</a> .....	52
<b>Ilustración 3.5. Logo Pentaho.</b> <a href="http://www.stratebi.com/pentaho">http://www.stratebi.com/pentaho</a> .....	52
<b>Ilustración 3.6. Logo Tibco.</b> <a href="https://spotfire.tibco.com/es/solutions/tibco-spotfire-para-servicios-financieros">https://spotfire.tibco.com/es/solutions/tibco-spotfire-para-servicios-financieros</a> .....	53
<b>Ilustración 3.7. Logo Ab Initio.</b> <a href="https://www.abinitio.com/en/">https://www.abinitio.com/en/</a> .....	53
<b>Ilustración 3.8. Azure Analysis Services.</b> <a href="https://docs.microsoft.com/es-es/azure/analysis-services/analysis-services-overview">https://docs.microsoft.com/es-es/azure/analysis-services/analysis-services-overview</a> .....	54
<b>Ilustración 3.9. Logo Oracle</b> <a href="https://www.oracle.com/es/index.html">https://www.oracle.com/es/index.html</a> ...54	
<b>Ilustración 3.10. Logo SAP Business Objects.</b> <a href="https://gravitar.biz/business-objects/">https://gravitar.biz/business-objects/</a> .....	54
<b>Ilustración 3.11. Logo Microstrategy.</b> <a href="https://churriwifi.wordpress.com/indice-de-publicaciones/">https://churriwifi.wordpress.com/indice-de-publicaciones/</a> .....	55
<b>Ilustración 3.12. Logo Microsoft Dynamics</b> <a href="https://dynamics.microsoft.com/es-es/">https://dynamics.microsoft.com/es-es/</a> .....	55
<b>Ilustración 3.13. Logo Ultimus</b> <a href="http://www.ultimus.com/es/bpmsoftwaredmodules/process-administrator">http://www.ultimus.com/es/bpmsoftwaredmodules/process-administrator</a> .....	55



**Ilustración 3.14. Logo SharePoint** <https://support.office.com/es-es/article/%C2%BFqu%C3%A9-es-sharepoint-97b915e6-651b-43b2-827d-fb25777f446f> .....56

**Ilustración 3.15. Logo QlikView.** <http://prod.qlik.com/us/> .....56

**Ilustración 3.16. Logo JeyRepors** <https://www.quonext.com/software-gestion/business-intelligence/jet-reports> .....57

**Ilustración 3.17. Logo Sisense.** <https://www.capterra.es/software/86955/sisense>..... 57

**Ilustración 3.18. Logo Clear Analytics.** <https://reviews.financesonline.com/p/clearanalytics/>.....58

**Ilustración 3.19. Logo Google Analytics.** <https://marketingplatform.google.com/about/analytics/>.....58

**Ilustración 3.20. Logo GoodData.** <https://www.gooddata.com/industry/insurance>.....58

**Ilustración 3.21. Logo Eclipse.** <http://www.vogella.com/tutorials/EclipseBIRT/article.html>..... 59

**Ilustración 3.22. Logo Power BI** <https://powerbi.microsoft.com/es-es/>.....59

**Ilustración 3.23. Beneficios tangibles e intangibles de un BI** ilustración tomada de: “Guía para la implementación de una solución de inteligencia de negocios para pequeñas y medianas empresas” de Daniel Felipe Flórez salgado y Laura Margarita del Mar Tejada Cardona.....65

**Ilustración 4.1. Ilustración explicativa de la forma de trabajo de la empresa.** Ilustración de creación propia que explica como genera las facturas, los presupuestos y los albaranes.....71

**Ilustración 4.2. Logo de Sage – Factura plus**.....72

**Ilustración 4.3. Tabla datos planta 0** ..... 74

**Ilustración 4.4. Tabla datos albaranes de la planta 1 facilitado por la empresa.** .....75

**Ilustración 4.5. Tabla datos clientes de la planta 1**.....75

**Ilustración 4.6. Tabla datos facturas de la planta**.....76

**Ilustración 4.7. Tabla datos presupuestos de la planta 1**..... 77

**Ilustración 4.8. Gráfica de Google de la asistencia de clientes a la tienda**...86

Ilustración 4.9. Tabla base datos planta 0. ....	89
Ilustración 4.10. Tabla base datos planta 1.....	91
Ilustración 5.1. Logo MySQL.....	94
Ilustración 5.2. Logo XAMPP.....	94
Ilustración 5.3. Logo Power Bi.....	94
Ilustración 5.4. Logo Excel. ....	94
Ilustración 5.5. Logo Acces.....	95
Ilustración 5.6. Logo Acces.....	97
Ilustración 5.7. Panel de control de XAMPP.....	101
Ilustración 5.8. Vista general gestor MySQL.....	101
Ilustración 5.9. Tablas creadas con el gestor MySQL.....	102
Ilustración 5.10. Creación conexión a BD con MySQL Workbench.....	103
Ilustración 5.11. Panel para la creación de conexiones entre tablas con MySQL Workbench. ....	103
Ilustración 5.12. Código que se genera para la creación de conexiones entre tablas .....	104
Ilustración 5.13. Captura de pantalla de la carga de datos con MySQL Workbench.....	105
Ilustración 5.14. Captura Panel general del Power Bi.....	107
Ilustración 5.15. Panel gráficos Power Bi.....	108
Ilustración 5.16. Captura Dashboard planta 0.....	109
Ilustración 5.17. Dashboard planta 1.....	109
Ilustración 8.1. Captura carga de datos en Power Bi.....	118
Ilustración 8.2. Captura carga de datos en Access.....	119
Ilustración 8.3. Captura selección tipo de importación.....	120
Ilustración 8.4. Captura asistente importación de datos.....	120
Ilustración 8.5. Captura creación de relaciones entre tablas en Access....	121
Ilustración 8.6. Captura panel de creación de relaciones. ....	121



Ilustración 8.7. Captura carga de datos en Power Bi. ....	122
Ilustración 8.8. Captura panel de control de XAMPP.....	123
Ilustración 8.9. Captura creación base de datos en MySQL.....	123
Ilustración 8.10. Captura selección tipo de dato de un campo.....	124
Ilustración 8.21. Captura carga de datos en Power Bi.....	124
Ilustración 8.32. Captura panel de control MySQL Workbench.....	125
Ilustración 8.43. Captura panel creación relaciones en MySQL.....	125
Ilustración 8.54. Captura tablas creadas en MySQL.....	126
Ilustración 8.65. Captura selección importación de datos en MySQL Workbench.....	126
Ilustración 8.76. Captura donde sacar información puerto y servidor.....	127
Ilustración 8.87. Captura panel Power Bi.....	127





# 1. INTRODUCCIÓN

## 1.1. Contexto y justificación

Durante el 2007 se generaron más datos que todos los años anteriores juntos. Los datos contienen una gran información que las grandes empresas utilizan para ofrecer mejores servicios y productos a los clientes. Algunas empresas como Netflix incluso crean algunos sus productos basándose únicamente en los informes reportados a raíz de los datos que obtienen de los clientes o de otras empresas.

Para crear esos informes, son necesarias el uso de unas herramientas que gestionan y analizan los datos, esas herramientas las conocemos como herramientas de business intelligence o herramientas de inteligencia de los necios en español.

Hasta hace pocos años estas herramientas solo estaban al alcance de las grandes empresas que podían acceder a los datos y podían permitirse contratar analistas que tratasen los datos, actualmente, con internet, las redes sociales y los softwares de los que disponen las empresas, las PyMEs también pueden acceder a esos recursos y obtener la información necesaria para mejorar su negocio.

Hay que destacar la importancia de las PyMEs en España, representan más del 99% de las empresas del país. La gran competencia que tiene entre ellas y la amenaza continua de las grandes empresas ha hecho que se vean obligadas a mejorar sus productos y servicios.

La obtención correcta de información y el buen uso de esta, puede marcar la diferencia entre una buena gestión de un negocio y una mala gestión. Por eso es necesario que los dueños y gestores de las PyMEs sepan utilizar estas herramientas.

## 1.2. Objetivo

El objetivo de este trabajo Fin de Grado (TFG) es demostrar cómo sin conocimientos previos de BI y con unos conocimientos informáticos limitados a la ofimática, una pequeña empresa puede obtener información valiosa de sus datos.

Para demostrar esto, he llevado a cabo un caso práctico real estando en contacto con una empresa que me facilito los datos reales que genera el negocio.

### 1.3. Metodología

Para realizar este trabajo fue necesario adquirir unos conocimientos sobre la utilización de softwares que he usado, estos conocimientos los he adquirido realizando el curso online de aplicación práctica de Big Data a Pymes de BPMSat.

He utilizado 5 programas distintos que se pueden descargar gratuitamente o están incluidos en el paquete de Microsoft Office.

### 1.4. Organización del documento

Resumen y palabras clave: Breve resumen del Trabajo Fin de Grado. A continuación, se definen las palabras clave del TFG.

Capítulo 1 – Introducción: En este capítulo se sitúa el contexto y la justificación del proyecto, los objetivos propuestos y las metodologías utilizadas para confeccionarlo.

Capítulo 2 – Bussines Intelligence: En este capítulo se presenta que es el BI, sus principales componentes y las fases de implantación.

Capítulo 3 – BI en las empresas: Se explica cómo afecta el BI a las empresas y se presentan algunos de los softwares de los que puede disponer una empresa para hacer uso de sus datos.

Capítulo 4 – La empresa y los datos: Se presenta la empresa de la cual he realizado el caso práctico y datos de los que se dispone para realizar el TFG.

Capítulo 5 – Implantación del BI en la empresa: Se explican tres formas distintas de aplicar las herramientas de BI utilizando distintos softwares, pero como denominador común, el uso de Power Bi para generar los informes visuales.

Capítulo 6 – Conclusiones: Se realiza un juicio sobre la facilidad y la viabilidad que tiene las empresas para llevar a cabo lo que se propone en el TFG.

Capítulo 7 – Bibliografía: Lista de la bibliografía consultado como apoyo para la realización del TFG.

Capítulo 8 – Anexos: Información adicional al TFG.



## 2. BUSINESS INTELLIGENCE

### 2.1. Nacimiento y evolución del Business Intelligence

El conocimiento es una mezcla de experiencia, valores, información que sirve como marco para la incorporación de nuevas experiencias e información. Se origina y aplica en la mente de los conocedores.

La primera vez que se usa el termino Business Intelligence fue en 1958, *“Capacidad de aprender las interrelaciones de los datos presentados de tal modo que se pueda orientar la acción hacia el objetivo deseado”* – Hans Peter Luhn, 1958-. Esta definición apareció en su artículo *“A Business Intelligence System”*. [1].

Más adelante, en 1989, el analista de Gartner Group, Howard Dresner, describió el BI como *“Los conceptos y métodos para mejorar la toma de decisiones empresariales mediante el uso de sistemas basados en hechos de apoyo”*. En su artículo *“Lack of trust is one of the greatest barriers to success in business intelligence programs”* [2].

Hasta llegar al BI como lo conocemos ahora, fueron importantes unos hitos que a lo largo de los años han definido el BI. Un breve resumen en función de las décadas sería:

- **En los 60 empiezan a almacenarse datos en sistemas de archivos.** A finales de esta década aparece el concepto de bases de datos de la mano de Edgar Frank Cood. *“A Relational Model of Data for Large Shared Data Banks”* [3].
- **En los 70 se desarrollan las primeras BBDD** y con ello una serie de aplicaciones empresariales para la manipulación de estos datos.
- **En la década de los 80 comienza a escucharse el termino Datawarehouse.** Bill Inmon en su libro *“What is a Data Warehouse?”* [4], lo define como *“Un almacén de datos que está orientado a un tema, integrado, variable en el tiempo y no volátil de recogida de datos en apoyo de procesos de toma de decisiones de la dirección.”* Por su parte Ralph Kimball lo define como *“Un almacén de datos que extrae, limpia,*

*conforma y entrega una fuente de datos dimensional para la consulta y el análisis*” en su libro “*The Data Warehouse Toolkit*” [5]. En esta década también aparecen los primeros sistemas de reporting que más tarde explicaremos. Aunque se almacenaban datos en las BBDD, aún era difícil la explotación de estos puestos que no se disponía de aplicaciones que facilitasen la tarea.

Es a finales de esta década cuando empieza a popularizarse el termino BI.

- **En los 90, aunque caras, comienzan las primeras aplicaciones de BI.** No será hasta finales de esa década, cuando aparezca el concepto de datamining y las BBDD multidimensionales.
- En la década pasada, empresas como SAP, IBM, Oracle o Microsoft copan en el mercado con sus plataformas para el BI.
- En esta década, al igual que en años pasados, se han ido mejorando las aplicaciones alimentándose de las tecnologías anteriores.

## 2.2. Qué es el BI

De acuerdo con Revelli, el BI “es el proceso de recolección, tratamiento y difusión de la información que tiene un objetivo: la reducción de incertidumbre en el proceso de toma de decisiones estratégicas” [6]. Por otra parte, Zhan Cui describe el BI como “una forma y método de mejorar el rendimiento del negocio proporcionando asistencia de gran alcance para los ejecutivos encargados de la toma de decisiones y brindarles información útil a la mano” [7].

Una definición más acorde con los objetivos de este proyecto fue propuesta por Robert Stackowiak en la que se define el BI como “el proceso de tomar datos, analizarlos y presentarlo en un conjunto de reportes que condensan la esencia de esta información en la base de las acciones del negocio, lo que permite a la gerencia tomar decisiones fundamentales sobre el negocio” [8].

Después de haber leído las definiciones anteriores y otras muchas me quedo con que **el BI es la habilidad para transformar los datos en información, dicha información en conocimiento de forma que se pueda optimizar el proceso de toma de decisiones en los negocios.**



**Ilustración 2.1. Diagrama de relación entre datos, información y conocimiento.** Figura tomada de [https://www.sinnexus.com/business\\_intelligence/](https://www.sinnexus.com/business_intelligence/)

El BI no es el simple almacenamiento de datos o el análisis de estos, sino que esta tecnología lo que pretende es aglutinar esa gran cantidad de datos que genera la empresa y correlacionarlos de manera que sean útiles para la empresa no solamente obteniendo respuesta de que ha ocurrido o que está ocurriendo, sino que se puedan generar modelos predictivos de futuros eventos.

### 2.3. Datos, información y conocimiento

La forma más sencilla de diferenciar estos términos puede ser que los datos están localizados en el mundo y el conocimiento está localizado en los distintos agentes como máquinas, personas... La información sería el nexo de comunicación entre los dos términos.

Esta definición y los conceptos que se muestran a continuación se basan en definiciones del libro *“El poder del conocimiento”* de Dayana Figarella y Andrés Zamora [9] que a su vez las toman de las definiciones que Thomas Davenport y Laurence Prusak incluyen en su libro *“Working Knowledge: How Organizations Manage what They Know”* [10].

#### 2.3.1. Datos

Los datos son la mínima unidad y corresponden con los elementos primarios de información que por sí solos son irrelevantes en la toma de decisiones.

El nombre de una persona, su código postal son datos que sin un propósito o utilidad no sirven para ayudar a la toma de decisiones.

Los datos pueden provenir de fuentes internas y externas a la organización, pudiendo ser objetivos o subjetivos, cualitativos o cuantitativos...

### 2.3.2. Información

La información es un conjunto de datos procesado con significado (relevancia, propósito y contexto) por lo que son útiles para la toma de decisiones. Los datos se transforman en información añadiéndoles valor:

- **Contextualizando:** Saber el porqué y para qué se genera para qué se generaron
- **Calculando:** los datos pueden procesarse matemática y estadísticamente.
- **Condensado:** Los datos se pueden resumir de forma más concisa.
- **Corrigiendo:** Se han eliminado errores e inconsistencias de los datos.
- **Categorizando:** se conocen las unidades de medida que ayudan a interpretarlos.

Por lo tanto, **la información es la comunicación de conocimientos.**

*“En 2007, por primera vez en la historia, se generó más información en un año de lo que se había producido en los cinco mil años anteriores - el período transcurrido desde la invención de la escritura.” [11]*

### 2.3.3. Conocimiento

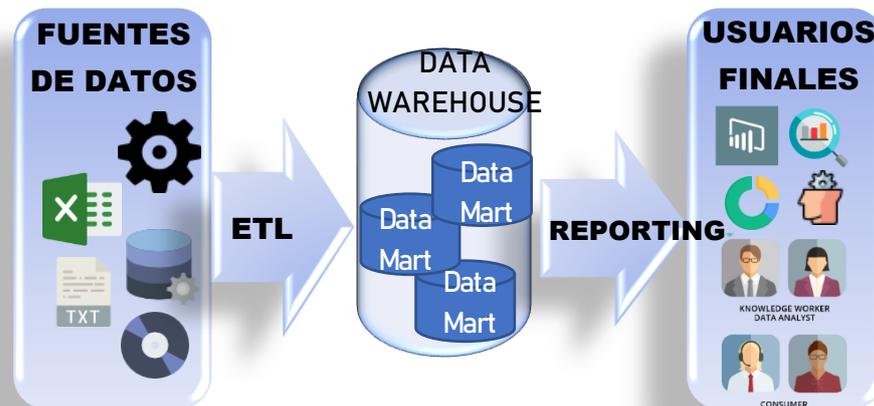
El conocimiento es una mezcla de experiencia, valores, información que sirve como marco para la incorporación de nuevas experiencias e información. Se origina y aplica en la mente de los conocedores.

El conocimiento proviene de la información, así como la información proviene de los datos. Para que la información se convierta en conocimiento es necesario realizar:

- Comparación con otros elementos.
- Predicción de consecuencias.
- Búsqueda de conexiones.
- Conversación con otros portadores de conocimiento.

## 2.4. Componentes del BI

En este apartado explicare los distintos elementos que componen un sistema BI, elementos que están representados en la siguiente figura.



**Ilustración 2.2. Diagrama de componentes de BI.** Ilustración de creación propia basada en <http://www.diazepamresource.com/real-time-data-warehouse-B9573/fantastic-real-time-data-warehouse-j23-on-wonderful-home-design-ideas-with-real-time-data-warehouse/>

### 2.4.1. Fuentes de datos

Los datos que utilizemos pueden ser generados por la propia empresa o de origen externo. Pero la distinción más importante que debemos hacer es entre datos estructurados y datos no estructurados:

- **Datos estructurados:** Es la información que se suele encontrar en la mayoría de las bases de datos. Son archivos de tipo texto que se suelen mostrar en filas y columnas con títulos. Son datos que pueden ser ordenados y procesados fácilmente por todas las herramientas de minería de datos. Lo podríamos ver como si fuese un archivador perfectamente organizado donde todo está identificado, etiquetado y es de fácil acceso.
- **Datos no estructurados:** Los datos no estructurados, generalmente son datos binarios que **no tienen estructura interna identificable**. Es un conglomerado masivo y desorganizado de varios objetos que no tienen valor hasta que

se identifican y almacenan de manera organizada. Podrían ser datos generados en las redes sociales, foros, e-mails, presentaciones Power Point o documentos Word, mientras que datos no-texto podrían ser ficheros de imágenes jpeg, ficheros de audio mp3 o ficheros de video tipo flash.

- **Datos semiestructurados:** Los datos semiestructurados serían aquellos datos que no residen de bases de datos relacionales, pero presentan una organización interna que facilita su tratamiento, tales como documentos XML y datos almacenados en bases de datos NoSQL.

## 2.4.2. ETL

Debemos entender cuál es el proceso de construcción del datawarehouse denominado ETL:

- **Extracción (E):** obtención de información de las distintas fuentes internas y externas.
- **Transformación (T):** filtrado, limpieza, depuración, homogeneización y agrupación de la información.
- **Carga (L):** Organización y actualización de los datos en la BD.

## 2.4.3. Datawarehouse

Un datawarehouse es una base de datos que se caracteriza por integrar y depurar información de varias fuentes para luego poder analizar esta en función de distintos parámetros.

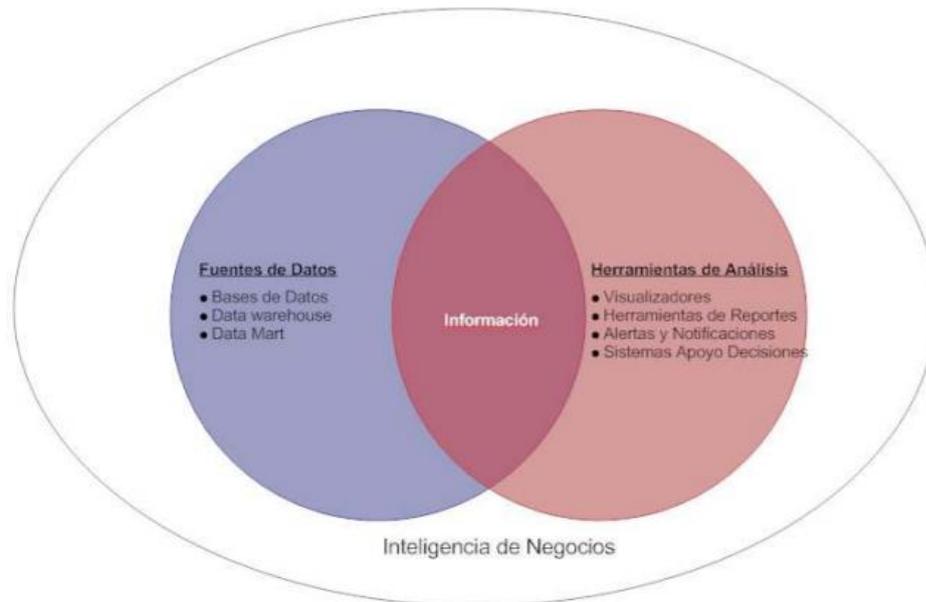
Crear un datawarehouse supone desde el punto de vista técnico el primer paso para implantar BI.

La ventaja de estas bases de datos se debe a la forma de las estructuras en las que se almacena la información, cubos relacionales, tablas en estrella o copo de nieve que más adelante explicaremos.

Anteriormente ya he hecho referencia al termino que acuño Bill Inmon, sin embargo, es más complejo que un simple almacén, Bill Inmon expuso que un datawarehouse se caracteriza por ser:

- **Temático:** Los datos se organizan por temas para facilitar su acceso y entendimiento por los usuarios finales.
- **Histórico:** El datawarehouse debe tener la información cargada con los distintos valores que toma la variable tiempo para así poder hacer comparaciones del resto de datos en función de este y hacer análisis de tendencias.

- **Integrado:** los datos almacenados deben integrarse en una estructura consistente, la información suele estructurarse también en distintos niveles de detalle para adecuarse a las necesidades del cliente.
- **No volátil:** la información debe estar almacenada para ser leída, no modificada. La introducción de nueva información debe ser una actualización sobre lo que ya exista.



**Ilustración 2.3. Fuentes de datos – información – herramientas de análisis** Ilustración tomada de [http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529\\_esa.pdf](http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf)

#### 2.4.4. Datamart

Un datamart es una base de datos especializada en el almacenamiento de datos de una determinada área de negocios. Se caracteriza por tener una estructura óptima de los datos para analizar la información al detalle desde distintas perspectivas. Los datamart pueden ser alimentados con los datos del datawarehouse.

Para crear el datamart, es preciso encontrar la estructura óptima para el análisis de su información. Esta estructura puede estar montada sobre una base de datos OLTP o sobre una OLAP.

Los datamarts tienen las siguientes ventajas:

- Poco volumen de datos.

- Mayor rapidez de consulta.
- Consultas SQL y/o MDX sencillas.
- Validación directa de la información.
- Facilidad para hacer histórico de datos.

### OLAP y OLTP

Las características de una OLTP son las siguientes:

- Información desgranada.
- Nula variabilidad y rigidez del esquema.
- Información reciente.
- Procesamiento de transiciones, realiza un dato por primera vez.
- Se pueden realizar informes básicos.
- La información está resumida al detalle
- La información se actualiza en tiempo real.
- El volumen de datos de entrada suele ser muy alto.

Características de OLAP:

- Trabaja con una agregación de datos.
- Presenta datos resumidos, sin detalles.
- Análisis de grandes cantidades de datos para así encontrar más patrones.
- Los datos deben estar depurados.
- La limpieza de los datos nos permite tener datos más fiables

Las empresas que tiene un almacén de datos, suele disponer de más de una base de datos, en función de los datos que almacenen en cada base de datos, podrán elegir un sistema u otro.

Quizás la manera más visual y sencilla de comprender las diferencias entre ambas sea mediante una imagen en la que diferenciar donde no coinciden ambas dos, OLTP frente a OLAP.

	OLTP	OLAP
Usuario típico	Operacional	Analista
Registros accedidos	Decenas	Millones
Unidad de procesamiento	Transacción simple	Agregación
Redundancia de datos	Causa ineficiencia	Clave para procesamiento eficiente
Función	Almacenamiento/consulta	Análisis
Modelo de datos	Relacional	Multidimensional

**Ilustración 2.4. OLAP frente OLTP** Ilustración de creación propia basada de una encontrada en <https://jorsodom.wordpress.com/2013/09/19/cual-es-la-diferencia-entre-una-base-de-datos-oltp-y-olap/>

### 2.4.5. Dashboard

**Es una representación gráfica de los principales indicadores (KPI)** que intervienen en la consecución de los objetivos de negocio, y que está orientada a la toma de decisiones para optimizar la estrategia de la empresa.

Debe representar las KPIs necesarias, no por tener más indicadores será mejor la información, hay que saber hacer limpieza. Además, se deben presentar las KPIs de forma que sean relevantes, con un contexto. También, es importante una visualización que sea fácil de interpretar, que nos permita tomar decisiones de forma correcta y rápida.

Un dashboard es una captura de los principales indicadores de nuestro negocio, no un lugar donde estudiar un determinado tema, cruzar variables, etc. etc.

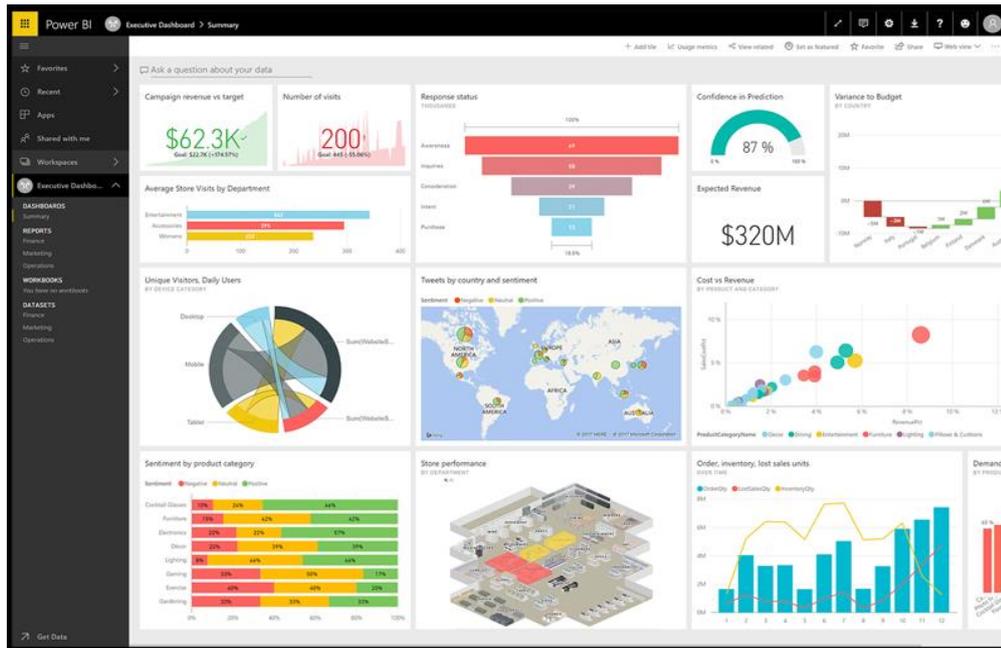


Ilustración 2.5. Dashboard Power BI. Ilustración tomada de <https://denglishbi.wordpress.com/2018/02/<20/how-to-build-a-sql-server-dashboard-with-power-bi/>

## 2.4.6. Cuadro integral de mando

EL CMI, también conocido como **Balanced Scorecard (BSC)**, es un sistema de planificación que es utilizado en los negocios y la industria para establecer y monitorizar los objetivos de la empresa y de sus diferentes áreas o unidades, es decir se usa para la gestión del rendimiento empresarial. Fue creado por dos profesores de la Universidad de Harvard en 1990, Robert S. Kaplan y David P. Norton. [12]



Ilustración 2.6. Diagrama explicativo de los de CMI, Ilustración tomada de: [https://www.sinnexus.com/business\\_intelligence/cuadro\\_mando\\_integral.aspx](https://www.sinnexus.com/business_intelligence/cuadro_mando_integral.aspx)

Se originó como una medida que añade las medidas de desempeño no financieras estratégicas para las métricas financieras tradicionales para dar a los gerentes una visión más amplia del desempeño organizacional.

Un **cuadro de mando integral** requiere de un formato prescriptivo, por eso no pueden faltarle los siguientes elementos, que ayudan a asegurar que el BSC está intrínsecamente ligado a las necesidades estratégicas críticas de la organización:

- **Perspectivas:** planteadas como agrupaciones de áreas estratégicas a un alto nivel. Perspectivas financieras, del cliente, interna y de crecimiento.
- **Metas:** expresando la visión que hay detrás del plan estratégico.
- **Métricas:** que provienen de los indicadores clave de rendimiento definidos.

#### 2.4.7. Minería de datos

El datamining o minería de datos, es una tecnología que nos permite trabajar con grandes cantidades de bases de datos de manera automática o semiautomática, con el objetivo de encontrar patrones, tendencias o reglas que nos ayuden a comprender el comportamiento de los datos. Para conseguirlo, hace uso de la estadística y de algoritmos que se aproximan a lo que conocemos como IA. Luis Carlos Molina *“Data Mining: Torturando a los datos hasta que confiesen”* [13]

#### 2.4.8. Reporting

Es el proceso mediante el cual, se hace informes utilizando la información del Data Warehouse con el objetivo de que el usuario final pueda ver la información. La generación de informes requiere partir de la reflexión (operativa o estratégica) para saber qué información se necesita, quien la necesita y con qué periodicidad.

## 2.5. Metodología para el diseño e implantación de un sistema BI

La implantación de un sistema BI en una empresa siempre vendrá condicionada por su actividad, por la cantidad de datos que genere y por el tamaño de esta.

El nivel tecnológico que hay en los sistemas de BI, permite que se pueda adaptar a las necesidades de una empresa, el problema puede surgir porque los gestores de las empresas no estén preparados para controlar estas tecnologías, lo cierto es que una que la herramienta BI no tiene necesidad de ser especialmente compleja.

Las herramientas de BI son capaces de dar soporte en estas tecnologías:

**Captación de Información:** incluye todos los procesos orientados a integrar información, calidad de datos, a capturar información no existente previamente, etc. Nos referimos a todas las tecnologías relativas a ETL y a los diferentes sistemas de captura de información utilizable por sistemas BI

**Manejo de la Información:** incluye todos los sistemas relacionados con Bases de Datos, Modelos Multidimensionales, sistemas de agregación de información, etc.

**Visualización y Distribución:** Toda la información deberemos visualizarla para poder hacer algo con ella. Para conseguirlo están las soluciones de reporting, visualización, broadcasting, etc....

**Análisis de la Información:** es el grupo funcional que dio origen a la denominación de Business Intelligence..., es en muchos casos irrelevante en los proyectos de BI, siendo sustituida la capacidad de análisis, por multitud de “fotos fijas” que presentan información.

**Gestión de las decisiones adoptadas:** Una vez adoptada una decisión queda ponerla en marcha y gestionar su ejecución. Esta es por tanto la funcionalidad que cierra el ciclo y que nos permite tener una tecnología capaz de gestionar todo el proceso. La parte final de todo proceso de análisis es la acción, pero si dejamos la gestión de las acciones fuera de la tecnología BI, difícilmente vamos a poder cerrar el círculo y realimentar nuestro sistema

La ilustración de a continuación muestra los diferentes elementos que conforman el mundo de la tecnología de BI y la gestión del proceso de toma de decisiones:



**Ilustración 2.7. Elementos tecnología BI.** Ilustración de creación propia inspirada en la web: <https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/>

Estas son las ramas necesarias si queremos tener el control de nuestra información para poder gestionar la toma de decisiones.

**Un proyecto de implantación de BI se enfrenta a varios problemas, pero el primero de ellos es determinar por qué se quiere poner en marcha. Si no conocemos la problemática a la que nos enfrentamos, no podremos determinar el alcance del proyecto.**

Cuando abordamos un proyecto de BI es necesario determinar su alcance, esto no tan sencillo como parece y constituye la primera etapa que expondré en este capítulo: La Determinación de Requerimientos

### 2.5.1. Aspectos Fundamentales para la Implantación de un Sistema BI

- Es necesario conocer las necesidades de la empresa y como evolucionaran en el futuro.
- Hay que consultar a los usuarios finales de la organización, ellos son los que darán uso las herramientas por lo cual debemos contar con su opinión. La gestión de las herramientas será suya y quien mejor que ellos para determinar las necesidades.
- El tipo de tecnología que se implante (dentro de la tecnología BI), dependerá de la empresa, sus necesidades, el tipo de empresa y su modelo de gestión.
- Hay que planificar todo el proyecto y gestionarlo.

### 2.5.2. Herramientas necesarias para la implantación completa de BI

Podemos distribuir las soluciones de BI en tres categorías:

- **Herramientas para la gestión de datos** (del inglés, *data management tools*): Permiten depurar y estandarizar datos de distinta procedencia.
- **Las aplicaciones para descubrir nuevos datos** (del inglés, *data discovery applications*): Permiten la recopilación y evaluación de información (*data mining* o [minería de datos](#)), y analizar con técnicas de análisis predictivo para realizar proyecciones de futuro.
- **Herramientas de reporting.**: Una vez recopilada y tratada toda esa información preexistente o nueva, ayudan a las empresas a visualizarla de manera gráfica e intuitiva.

### 2.5.3. Fase 1 - Determinación de Requerimientos

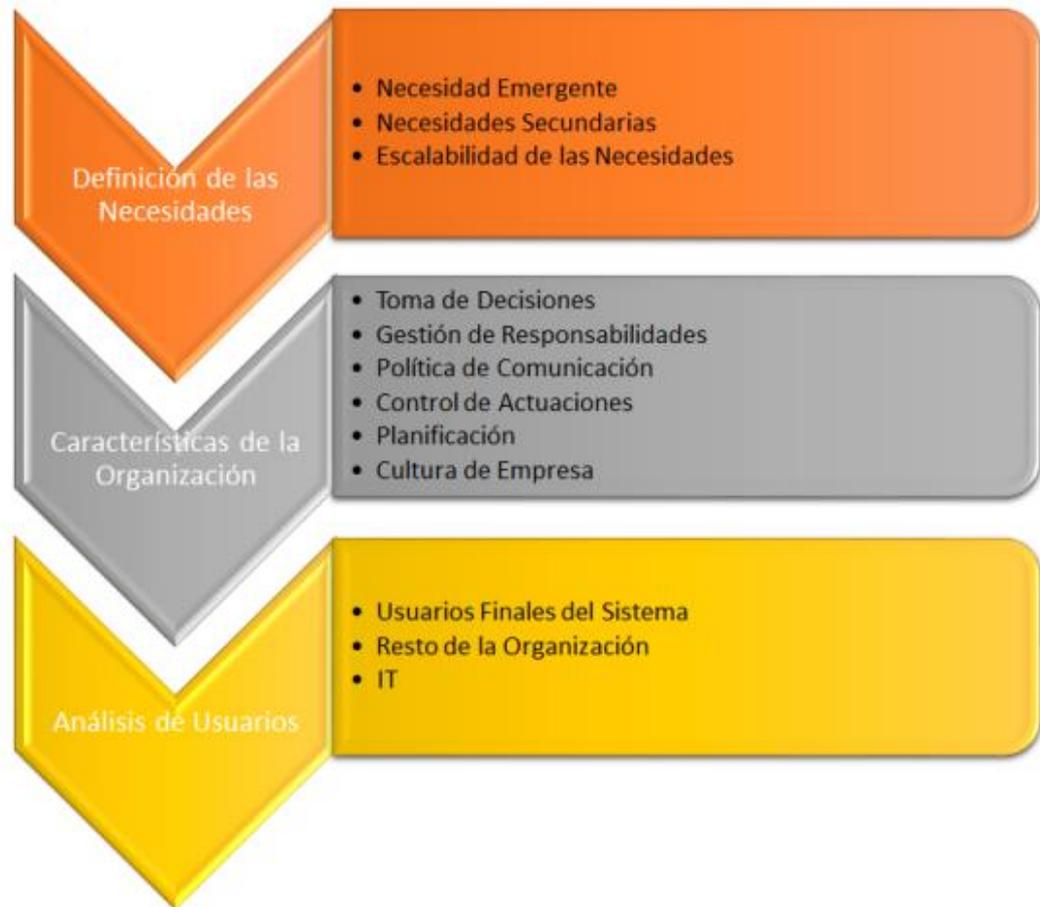
Tiene tres etapas diferenciadas pero interrelacionadas:

#### 1.1. Determinación de Necesidades.

El objetivo de esta etapa es definir qué necesidades tenemos cuando nos planteamos la realización de un proyecto de BI. Podemos estructurar esta etapa en tres actividades:

- Determinación de la Necesidad Emergente - La primera pregunta que debemos plantearnos es qué necesidad tenemos dentro de la empresa y, sobre todo, para qué queremos un sistema de BI.  
Tenemos un problema con la información que manejamos.  
Tenemos un problema de accesibilidad a la información.  
Tenemos un problema para gestionar con dicha información.
- Analizar además de la necesidad emergente o principal, el resto de necesidades que podemos tener. Uno de los errores más frecuentes es hacer o modificar un proyecto de BI en función de una primera y/o principal necesidad detectada si atender a otras necesidades secundarias.
- Establecer el alcance de estas necesidades: ¿se trata de un único Departamento o Área afectado?; ¿podrían estar afectados por esta necesidad otras áreas?; ¿el alcance es general de toda la Compañía, aunque haya surgido de un área concreta?

Todo proyecto de BI debe tener esas dos capacidades: la transversalidad (es decir, poder servir para cualquier otra área de la empresa) y la escalabilidad (es decir, poder ser válida para cubrir las necesidades globales)



**Ilustración 2.8. Determinación de los requerimientos del BI:** Ilustración tomada de: <https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/>

## 1.2. Características de la Organización y su importancia en las Necesidades

Como una organización toma decisiones, depende de la organización interna de esta. Debemos analizar los siete factores para tenerlos en cuenta:

- Quien toma las decisiones.
- Como las toma.
- Hay decisiones automáticas o siempre interviene una persona.

Los requerimientos funcionales del sistema BI deberán estar en coordinados con la estructura de la Organización;

Nuestro sistema debe ser acorde a la actividad de la empresa presente como la futura a medio plazo.

- Gestión de Responsabilidades. ¿Tenemos un modelo jerárquico o descentralizado de responsabilidades? A la hora de ver cómo vamos a gestionar las decisiones es fundamental analizar cómo funcionan las responsabilidades en la Organización
- Política de Comunicación. Cómo fluya la información en nuestra Organización va a condicionar la posibilidad de explotar un sistema de BI:
- Sistemas de Control de Actuaciones- ¿Cómo se gestionan los proyectos a desarrollar? ¿Quién asigna las responsabilidades? ¿Cómo se persigue su cumplimiento?
- Planificación- La existencia o inexistencia de una cultura de programación y planificación de las actuaciones a realizar.
- Cultura de Empresa. Al final, lo que estamos analizando en este apartado son todos los aspectos que tienen que ver con la cultura de empresa que tiene nuestra organización.

### 1.3. Definición de las características de los usuarios.

Debemos tener en cuenta **cómo son los usuarios finales**. ¿Están acostumbrados a tomar decisiones? ¿Analizan la información suministrada o se limitan a ejecutar a partir del análisis realizado por otros? ¿Que usan hoy día para acceder a la información? ¿Cómo analizan los datos?

### 2.5.4. Fase 2 - Estrategia de Proyecto

Ahora es el momento de pensar cómo vamos a realizar el proyecto para dar servicio y solución a las necesidades que hemos planteado en la fase anterior.

Debemos plantearnos:

- Las necesidades a corto y medio plazo.
- Debemos entender el nivel de importancia del proyecto, como la a impactar en la empresa.
- Como en cualquier proyecto tenemos que ver cuáles son nuestros puntos fuertes, y cuáles nuestras desventajas.
- Debemos intentar prever las dificultades que vamos a encontrarnos y como vamos a subsanarlas.

- Cuál es la responsabilidad de cada persona en el proyecto.
- Cómo afrontaremos cada fase del proyecto
- Establezcamos los niveles de éxito y fracaso aceptables para cada aspecto incluido, y las alternativas si no se consiguen.
- Establezcamos la política de comunicación que vamos a seguir.
- Por último, cuál será nuestra estrategia.

Si valoramos estos puntos antes de empezar con el proyecto, nos facilitará mucho el trabajo puesto que estaremos preparados para los problemas y podremos resolverlos o incluso evitar que ocurran.

### **2.5.5. Fase 3 - Selección de la Tecnología**

Debemos saber que softwares y hardwares utilizaremos en el proyecto. Esto hay que hacerlo cuanto antes porque condicionará lo que se vamos a poder hacer y cómo vamos a poder hacerlo.

Debemos analizar qué tecnología cubre funcionalmente nuestras necesidades y qué tecnología podemos soportar desde nuestra organización. El cruce de ambas debería darnos la mejor alternativa para soportar el proyecto.

### **2.5.6. Fase 4 - Determinación del Modelo de Información**

Esta fase es primordial para saber que necesitamos de nuestro sistema de información, por ejemplo:

- Que variables vamos a controlar.
- Dimensiones que queremos gestionar (Lo importante de cada variable)

Esto será importante para que cuando implantemos nuestro proyecto, cubra las necesidades de gestión de la empresa.

### **2.5.7. Fase 5 - Modelado**

Cuando tenemos la información necesaria gracias a las preguntas que nos hemos realizado antes, es fundamental estudiar los datos que tenemos y realizar el proceso de modelado.

En el proceso de modelado recogemos los datos que nos facilita la empresa, los examinamos y creamos una estructura utilizable para el

análisis. Dichas estructuras son bastante similares para casi todos los proyectos de BI.

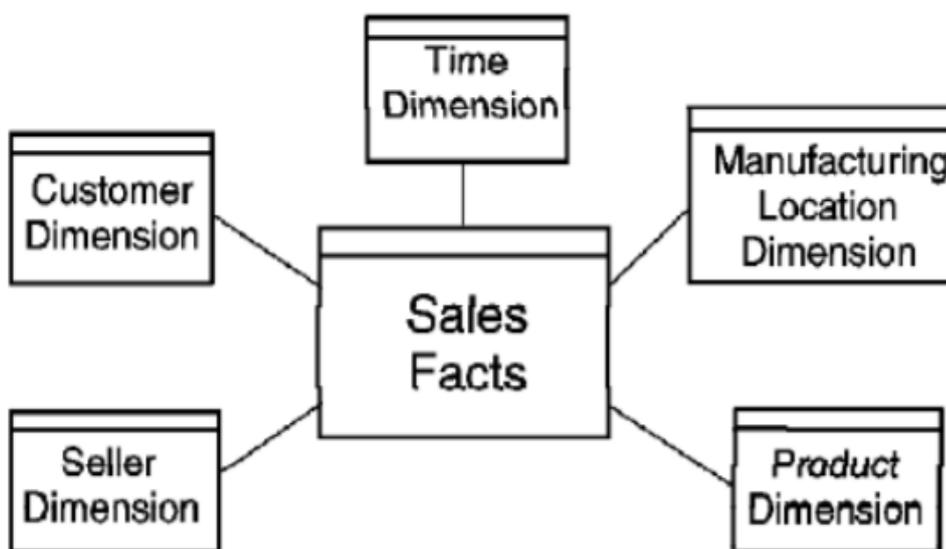
Existen dos tipos de modelos básicos:

- **Modelo Estrella:** Es la estructura más sencilla. Hay una tabla de hechos central que contiene los datos para el análisis, rodeada de las tablas de dimensión.

Lo que caracteriza a la arquitectura de estrella es que sólo ha una tabla de dimensiones para cada dimensión. Esto quiere decir que la única tabla que tiene relación con otra es la de hechos, lo que significa que toda la información relacionada con una dimensión debe estar en una sola tabla.

Las tablas de dimensiones deben tener una clave primaria, la clave principal de la tabla de hechos estará compuesta por las claves principales de las tablas de dimensiones.

En las tablas de dimensiones se encuentran los atributos de los datos cuantificables que se encuentran en la tabla de hechos.



**Ilustración 2.9. Modelo estrella.** Fuente: BALLARD, Chuck; HERREMAN, Dirk; SCHAU, Don; BELL, Rhonda; KIM, Eunsang y VALENCIC, Ann. Data Modeling Techniques for Data Warehousing. IBM's International Technical Support Organization. p. 47. Febrero, 1998

Pros y contras:

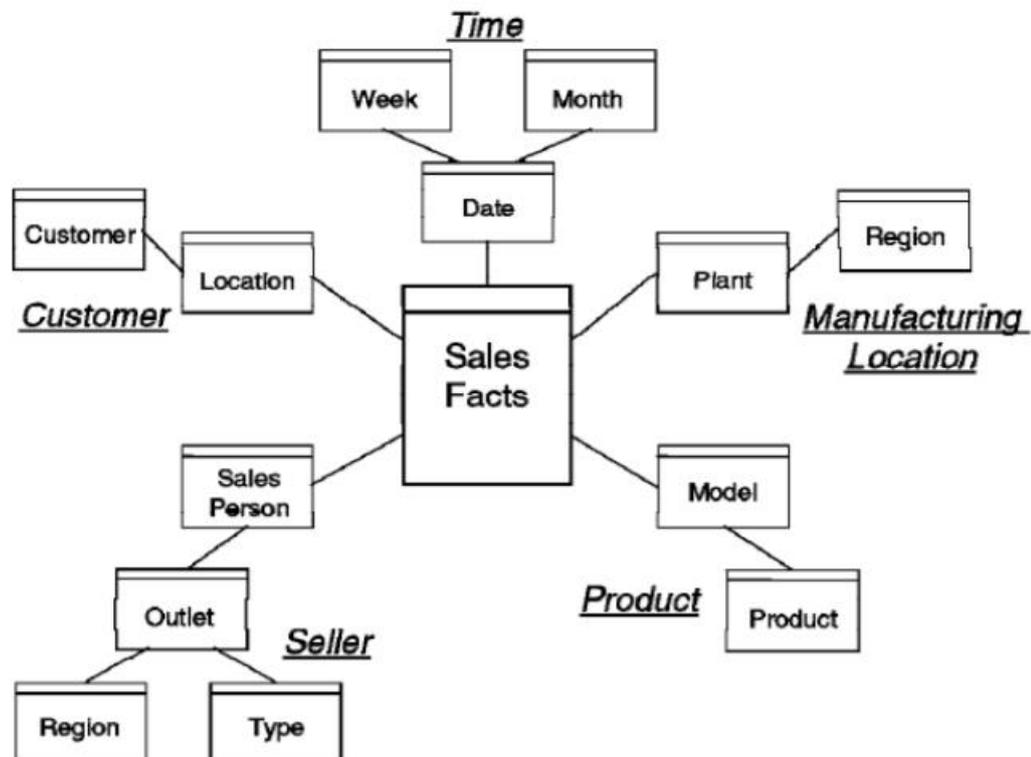
- Al ser un esquema simple, **es rápido**. Permite acceder tanto a datos agregados como de detalle.
- El diseño de esquemas en estrella permite implementar la funcionalidad de una base de datos multidimensional utilizando una clásica base de datos relacional.

- Debido a que las condiciones solo relacionan la tabla de hechos con las dimensiones, las consultas a la base de datos no son complicadas.
- Son más simples de manejar que los modelos copo de nieve.

**Modelo Copo de Nieve:** Este modelo se diferencia principalmente del modelo de estrella en que no solo la tabla de hechos tiene relación con otras tablas, sino que hay otras tablas que se relacionan con las dimensiones sin tener relación con los hechos.

Para conseguir un esquema en copo de nieve se ha de tomar un esquema en estrella y conservar la tabla de hechos, centrándose únicamente en el modelado de las tablas de dimensión, que ahora se dividen en subtablas.

El modelo fue concebido para facilitar el mantenimiento de las dimensiones, sin embargo, esto hace que se vinculen más tablas, haciendo la extracción de datos más difícil así como vuelve compleja la tarea de mantener el modelo.



**Ilustración 2.10. Modelo de copo de nieve.** Fuente: BALLARD, Chuck; HERREMAN, Dirk; SCHAU, Don; BELL, Rhonda; KIM, Eunsang y VALENCIC, Ann. Data Modeling Techniques for Data Warehousing. IBM's International Technical Support Organization. p. 47. Febrero, 1998

Pros y contras:

- En los esquemas de copo de nieve **se evita la redundancia** de datos y con ello se ahorra espacio.
- Se puede usar un esquema de copo de nieve en un Data Warehouse, aunque estos sean grandes y complejos, pero nunca en sistemas donde el tiempo de respuesta sea un factor crítico para los usuarios.

### 2.5.8. Fase 6 – Procesamientos de Datos

El siguiente paso es dar formato a los datos hasta llegar al modelo que hemos elegido en la fase anterior. Podemos hacerlo apoyándonos en la creación de nuevas bases de datos o agregando datos a otras ya existentes.

Esta fase es lo que hemos presentado anteriormente como proceso ETL. Es muy importante trabajar con cuidado y prestar atención a los datos con los que se trabajan pues tienen que tener coherencia a la entrada con la salida.

### 2.5.9. Fase 7 - Análisis y Producción

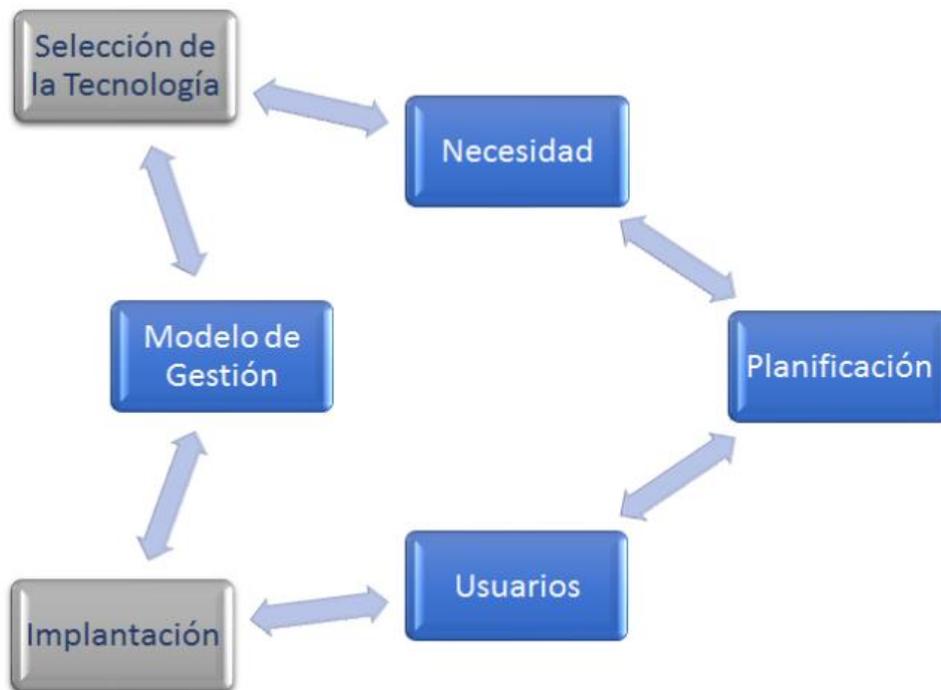
Hasta ahora lo que hemos hecho ha sido cargar y gestionar los datos. En esta fase es en la que se le da valor a los datos creando conocimiento e “inteligencia”. Esta fase la llamamos reporting, y consiste en recoger la información, transformarla y crear unos informes en función de los requisitos del cliente.

Es necesario una buena comunicación con el cliente para definir los informes que quiere y certificar que el proyecto va cumpliendo los objetivos marcados.

### 2.5.10. Fase 8 - Difusión:

En esta fase de difusión, se entrega el producto final, es decir, los informes, a los diversos clientes que lo requieren. Esto básicamente implica el uso de herramientas BI para la publicación de dichos informes o la posibilidad de tener herramientas de fácil uso para que los mismos usuarios tengan la capacidad de revisar los datos de manera rápida y sencilla.

Toda esta información la he obtenido de la Tesis de Master de la UPC “Diseño e implantación de un sistema de Business Intelligence para la gestión analítica de la facturación y los KPIs relevantes para el negocio en la empresa” [11] y de la web “decisionesytecnología ” [14]



**Ilustración 2.11. Aspectos Fundamentales para la Implantación de un Sistema de BI:**  
Ilustración tomada de: <https://decisionesytecnologia.wordpress.com/2013/02/07/metodologia-para-el-diseno-e-implantacion-de-un-sistema-de-bi/>

## 3. BI EN LAS EMPRESAS

En la actualidad, las operaciones de cualquier empresa generan una gran cantidad de datos que generalmente suelen ser almacenadas en bases de datos o de forma más sencilla en hojas de cálculo de Excel.

El fin de cualquier empresa es incrementar sus ganancias, para obtener este fin es muy importante mantener y sacar rentabilidad a los clientes que tiene, pero también es muy importante adquirir nuevos clientes. El BI permitirá a la empresa sacar una ventaja competitiva importante para conseguir los objetivos anteriormente mencionados.

La estrategia de BI debe abordar el cómo contribuirá al logro de las metas del negocio, cómo permitirá tomar mejores decisiones para incrementar el rendimiento de la organización y cómo se incorporará en los diferentes procesos de administración. Una estrategia de BI exitosa debe estar alineada con los objetivos empresariales, realizar un mejor uso de la información y permitir la integración de BI en las actividades estratégicas, tácticas y operacionales.

Los objetivos, las decisiones y los procesos empresariales deben promover la estrategia de BI para asegurar la adopción exitosa de esta tecnología. Conducir la estrategia de BI teniendo en cuenta el contexto empresarial asegura la alineación de la visión organizacional, las metas y los procesos con las iniciativas de BI.

Se debe determinar y cuantificar los beneficios que el negocio desea obtener y qué necesidades se están afrontando. Algunos ejemplos de los objetivos que pueden ser abordados por la estrategia de BI incluyen: cómo incrementar la lealtad del cliente, aumentar el número de clientes, elevar la participación en el mercado, promocionar efectivamente los productos, analizar los segmentos de mercado, entre otros. Estos son los objetivos finales que se deben considerar en la definición de la estrategia de BI más no las herramientas o tecnologías a implementar [15].

### 3.1. Indicadores de rendimiento

Como es evidente, el BI usa KPIs o indicadores claves del rendimiento para poder analizar y valorar el estado actual de la empresa y poder establecer un plan de actuación.

Estos indicadores pueden ofrecer información valiosa para la toma de decisiones por lo que es importante asegurarse de que los procesos críticos están siendo monitoreados [16].

Los KPIs (por sus siglas en inglés, Key Performance Indicators) son un método para determinar el éxito del negocio utilizando métricas objetivas y directrices cuantificables. Pueden ser utilizados para evaluar el estado global de la organización, comparar el plan de negocio con el desempeño actual o hacer seguimiento al progreso de una actividad en particular, meta, departamento o producto.

Aunque haya ciertos KPIs universales para todo el mundo, cada organización puede establecer los que ella crea más importantes para evaluar su actividad. Estos pueden analizar aspectos estratégicos, de procesos, económicos... en definitiva, pueden valorar si la empresa esta cumpliendo con sus objetivos marcados.

Los KPIs en sí mismo no son metas ni objetivos, si no, que son una forma de medir el correcto funcionamiento de la organización para la consecución de sus objetivos. Proporcionan a la empresa información fundamental sobre el rendimiento del negocio, especialmente en periodos muy cortos. Sin el uso de estos indicadores tomaría mucho tiempo y esfuerzo realizar la recopilación y procesamiento de los datos para obtener información sobre el estado general de la empresa; en esta medida se presentarían demoras en la identificación y tratamiento de los problemas que se puedan presentar.

Los KPIs incluyen información estadística como tendencias de ventas, margen de rentabilidad, grado de satisfacción del cliente, desempeño de los departamentos, inventario en tiempo real o cualquier otro factor que se considere crítico para el éxito de la empresa. Los KPIs son parte integral de una solución de BI puesto que contribuyen a una correcta implementación de esta tecnología en las organizaciones.

### 3.2. BI en Pymes

En España las PyMEs representan el 99% de las empresas del país. El 75% de las pymes pertenece al sector servicios; el 9,6% a la construcción; el 8,3% a la industria y el 7% al sector agrario. Si atendemos a la distribución del empleo, el grueso se lo reparten los servicios (75,3% empleo) y la industria (14,7%). Datos obtenidos del informe “Retrato de PYME - enero 2018” [17] realizado por el gobierno.

Las PyMEs se caracterizan por tomar decisiones con información incompleta y basadas en predicciones; sin embargo, la incorporación de datos de calidad podría mejorar significativamente sus niveles de innovación y competitividad.

Desde hace años han existido en el mercado herramientas de BI, las cuales resultan de especial interés, pues facilitan la recopilación y el análisis de los datos, presentándolos de forma tal que agregan valor al proceso de toma de decisiones. Tales herramientas estaban destinadas a la gran empresa, por lo que los costos y tiempos de implementación solían ser altos. Sin embargo, los avances tecnológicos experimentados durante los últimos años han permitido que las PyMEs tengan a su disposición tecnologías de BI -a precios accesibles- que faciliten la obtención de datos en tiempo y forma. A pesar de la existencia de la tecnología, no son muchas las PyMEs a nivel mundial que implementan BI, siendo las principales excusas el tiempo, el dinero y la complejidad.

**El procesamiento de la información se ha convertido gradualmente en un factor fundamental para generar ventajas competitivas** en el mercado actual. Con los avances en las tecnologías de la información, la existencia de una gran variedad de productos y mayores demandas por parte de los clientes, las empresas están obligadas a operar en entornos altamente complejos y dinámicos. Las organizaciones que sobreviven y tienen éxito en estas condiciones de mercado necesitan tomar decisiones de manera oportuna, eficaz y apropiada [18].

Con el objetivo de incrementar la eficiencia, muchas organizaciones han implementado sistemas de información en las operaciones de sus negocios para recolectar, combinar, acceder y analizar grandes cantidades de datos. Algunas de estas herramientas de análisis son proporcionadas por las tecnologías de BI, que conducen no solo a la mejora del rendimiento general de la empresa, sino que también ayudan en la predicción de tendencias del mercado mediante el análisis de datos históricos.

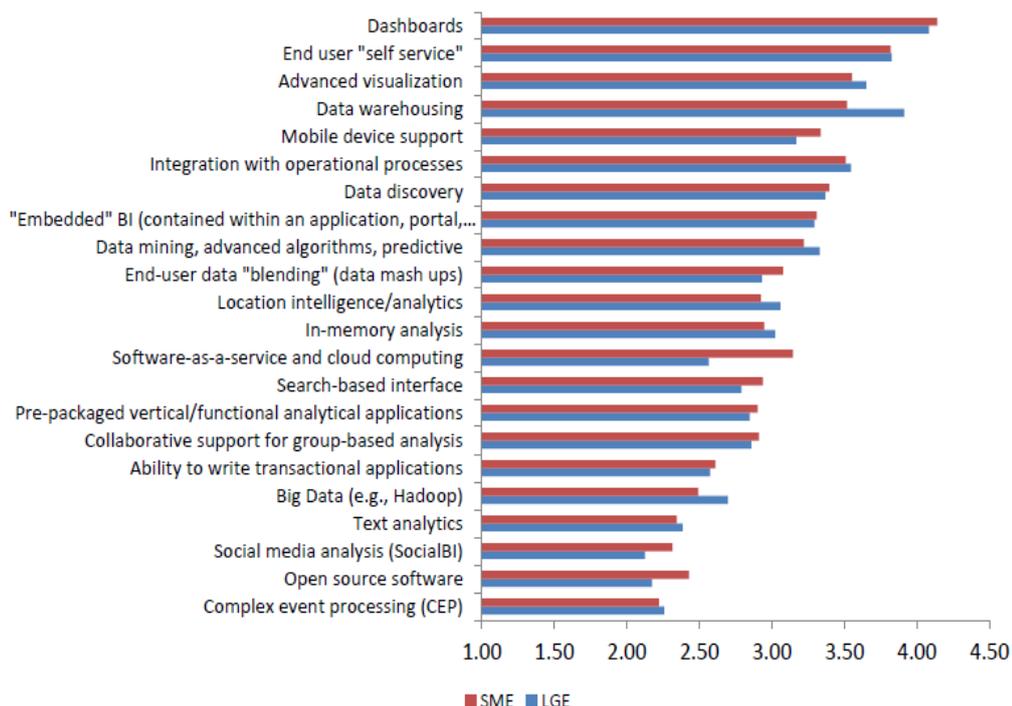
A pesar de que las tecnologías de BI normalmente se consideran reservadas para las grandes empresas, **la demanda actual de BI no se limita al tamaño de las organizaciones** [19]. Durante un tiempo considerable, las grandes compañías se han dado cuenta de la importancia de los datos existentes en sus sistemas de información y han realizado grandes inversiones en sistemas que les proporcionen capacidad de análisis y predicción. La oferta de soluciones de BI para las PYME es muy reducida, ya que los clientes objetivo de los proveedores de herramientas de BI siguen siendo las grandes empresas, debido a que son este tipo de organizaciones quienes demuestran una mayor receptividad para acoger proyectos de esta

naturaleza y también debido a su capacidad financiera [20]. Solo recientemente algunas PYMEs comenzaron a reunir los requisitos para adquirir este tipo de soluciones. De hecho, tanto las pequeñas como las medianas empresas observan la necesidad de incorporar este tipo de herramientas en la misma medida que las grandes compañías.

En un entorno empresarial en constante cambio, los requisitos a nivel de tecnologías de la información, los desafíos del mercado y las presiones empresariales que enfrentan las PYMEs no son muy diferentes de los que afrontan las grandes empresas. La combinación de datos históricos con la tecnología de BI permite a las pequeñas y medianas empresas tomar decisiones estratégicas pertinentes a través de técnicas de exploración de grandes volúmenes de datos, potenciando el análisis de patrones desconocidos para generar una ventaja importante en el entendimiento de los procesos del negocio.

La gestión de grandes cantidades de datos y su transformación en información útil representa un gran desafío para las PYMEs. Generalmente las organizaciones requieren información para el análisis de situaciones desfavorables; este análisis es un enfoque reactivo a una situación negativa o adversa. En una perspectiva de gestión proactiva, si los resultados son positivos, las organizaciones deben continuar desarrollando acciones para aprovechar y optimizar los beneficios obtenidos. Una empresa proactiva se enfoca en identificar nuevas oportunidades y prepararse para anticipar posibles fallos. Por otra parte, las organizaciones reactivas desgastan toda su energía y todos los recursos en revertir los resultados negativos, a expensas de la mejora de sus habilidades y la exploración de oportunidades. **Las compañías deben entonces decidir si adoptar un enfoque proactivo o reactivo ante el comportamiento del negocio y del entorno.**

La explotación de las tecnologías de BI es importante para el desarrollo de las pequeñas y medianas empresas. En una investigación realizada por Gibson y Arnott en 2003, se concluyó que si las grandes organizaciones continúan aprovechando las últimas tecnologías de soporte a la decisión, y las pequeñas empresas siguen sin correr el riesgo de adoptar inteligencia empresarial, la brecha de poder solo se mantendrá en aumento. Si este desequilibrio persiste, a las pequeñas organizaciones les resultará cada vez más difícil competir en la economía moderna [21]. Sin embargo, existen muchas consideraciones y riesgos que las organizaciones deben tener en cuenta para evaluar la adopción de soluciones de BI.



**Ilustración 3.1. Prioridades tecnológicas de las PyMEs frente a las grandes empresas.**  
 Ilustración tomada de: DRESNER. Small and Mid-Sized Enterprise Business Intelligence Market Study. p. 21. Dresner Advisory Services LLC, 2014.

### ¿Por qué no aprovechan las PYMEs el uso del BI?

La falta de recursos es uno de los principales factores que se aborda en el estudio de las PYMEs. Algunos de estos recursos incluyen las finanzas, la tecnología, los conocimientos y los recursos humanos. Debido a su presupuesto financiero restringido y el poco número de trabajadores, la mayoría de los empleados deben realizar múltiples tareas y no se especializan en un área en particular [22]. Por otra parte, la mano de obra no cualificada sumada a la falta de especialización técnica resulta en que los gerentes de las PYMEs sean conservadores al adoptar innovaciones de tecnología. Esta puede ser la razón por la que la adopción de nuevas tecnologías en las PYMEs es un proceso lento que se realiza con especial precaución.

Por ejemplo, Fuller-Love encontró que los propietarios o gerentes acceden a la adopción de BI solo cuando perciben la promesa de éxito, ya que no quieren correr riesgos [23]. Este es especialmente el caso de BI, debido a que los costos de mantenimiento y de ejecución son muy altos, así como la probabilidad de fracaso en la implementación en comparación con otras tecnologías.

Tal como plantea Suleymanov [24] (2011), BI no era accesible para las PyMEs por las siguientes razones:

- Los costos de implementación de BI eran muy altos; tanto que las PyMEs.
- Los productos instalados resultaban poco flexibles, aun cuando era conocido que las necesidades de negocio cambian constantemente.
- Los requerimientos de infraestructura eran demasiado altos, por lo que la empresa debía contar con un alto presupuesto para invertir en BI.
- El tiempo de implementación era muy alto, generalmente superior a los 6 meses.
- Existían muy pocas soluciones disponibles para las PyMEs.

Cayón[25] (2015) afirma que los avances en Cloud Computing durante los últimos años están acelerando la adopción de BI en las PyMEs, incluyendo la apertura de la posibilidad de implementar BI.

### 3.2.2 Las PyMEs y su proceso de toma de decisiones

Rojo[26] afirma que conducir exitosamente una PyME no es una tarea fácil. Además de a obtener beneficios, las PyMEs se enfrentan a una fuerte competencia en la que la sobrevivencia depende en gran medida de las decisiones tomadas.

Por otra parte, Elegant J. [27] expone que, independientemente del tamaño de las empresas, el contar con información crítica del negocio y con reportes integrados es indispensable para el éxito en el mercado.

EL BI no solamente ofrece beneficios a las grandes empresas. Por el contrario, las PyMEs representan el sector que puede apreciar de forma más tangible los beneficios que BI puede brindarle[28].

En cuanto a la gestión de las PyMEs, Valda [29] afirma que los empresarios que las lideran son rígidos, autocráticos e incuestionables en el poder, a la vez que sobrevaloran la experiencia, repitiendo modelos exitosos previos sin dar relevancia a los cambios de contexto que ocurren en las empresas.

Así mismo, su conocimiento y comprensión del entorno constituye un recurso clave para mantener la ventaja competitiva, por lo que experimentan procesos de tomas de decisiones incrementales, iterativos y no lineales, basados en información incompleta y guiados por corazonadas.

De esta manera, la PyME se enfrenta a tomar decisiones tanto operativas como estratégicas con información escasa e incompleta, dando mayor relevancia a experiencias previas de sus líderes.

Por otra parte, la evolución de la tecnología ha facilitado que las herramientas de BI, de gran valor para la generación de datos y reportes de calidad, se encuentren a disposición no sólo de la gran empresa, sino – también- de la PyME, cuyo volumen de datos y capacidad económica es mucho menor. Sin embargo, resulta interesante destacar que, aún con el abaratamiento de los costos, la implementación de tales soluciones no es utilizada masivamente por las PyMEs. *“Investigación sobre la aplicación de business intelligence en la gestión de las PyMEs”* [30]

### 3.3. Encuesta

Si bien los beneficios del BI a las PyMEs son amplios, es necesario que cada una evalúe si realmente lo necesita y si el retorno de la inversión es atractivo. Para ello, Cruz [31] propone cuatro preguntas sencillas que todo empresario PyME debe responderse antes de iniciar un proceso de implementación:

- ¿En base a qué, se toman las decisiones en la empresa?: Hoy en día son muchas las PyMEs que utilizan Microsoft Excel como herramienta para almacenar datos y generar reportes; es decir, que los gerentes ya se encuentran utilizando datos para tomar decisiones.
- ¿Con qué intensidad utilizan las redes sociales?: Con la masificación de las redes sociales, son muchas las PyMEs que utilizan alguna de ellas para darse a conocer y atraer nuevos clientes. Los datos que se generan en las redes suelen tener gran potencial y ser muy valiosos para tomar decisiones de marketing, por lo cual vale la pena analizarlos mediante herramientas de BI.
- ¿Cuáles son los sistemas con los que cuenta la PyME para registrar las operaciones?: Cualquiera que sea el sistema,

siempre que exista, se cuenta con datos para ser explotados y analizados al máximo. Si la PyME cuenta con dichos datos, puede pensar en utilizar BI para transformarlos y relacionarlos para tomar mejores decisiones que las que tomaría con la simple intuición.

- ¿Cuánto costaría implementar BI?: La respuesta a esta pregunta es bastante complicada. El ejecutivo PyME podría comenzar a plantearse algunas cuestiones como prioridades, tecnología requerida, costo de desarrollo, área de negocio con mayor necesidad, coordinador, entre otros.

Continuando con la propuesta de Cruz respecto a las preguntas que debe hacerse un empresario con una PyME, he decidido realizar una encuesta a distintas PyMEs con la intención de conocer la situación del uso del BI. Además de las preguntas que propone, he añadido algunas de tal forma que las preguntas del formulario han sido las siguientes:

- ¿Número aproximado de trabajadores?
- ¿En que se apoya para la toma de decisiones?
  - Datos recopilados anteriormente.
  - En base a experiencias previas.
  - Intuición.
- ¿Qué uso le da a las redes sociales de la empresa?
  - Como herramienta de marketing.
  - Para estar en contacto con clientes.
  - Ambas.
  - Prácticamente no las uso.
- ¿Qué herramientas utiliza para el almacenamiento de datos?
  - Excel.
  - Base de datos.
  - Herramientas googles.
  - Otros.
- ¿Sabe lo que es el Business Intelligence?
  - Sí.
  - No.
  - He oído hablar pero no sé qué es.
- ¿Considera que utiliza una herramienta de BI en su empresa?
  - Sí.
  - No.
- ¿Ha considerado implementar herramientas de BI en su negocio?

- Sí.
  - Sí, pero ahora no puedo permitírmelo.
  - No.
- ¿Si ha implementado herramientas de BI, tuvo algún problema al hacerlo? (Puede marcar varios)
  - No uso herramientas de Bi.
  - No tuve ningún problema al implementar.
  - Hubo problemas en cuanto a los plazos de implementación.
  - Ha habido problemas con el uso del BI una vez implementado.
  - Los reportes obtenidos inicialmente no fueron los esperados.
  - Fue difícil alinear las necesidades de la empresa con lo que ofrecía la herramienta BI.
  - Hubo problemas con los costes finales de implantación.
  - Otros.
- ¿Dentro de la empresa quien genera los reportes? (Puede seleccionar varios)
  - El gerente.
  - Área de tecnología.
  - Otros.
- ¿Percibió mejoras con la implementación de BI?
  - No he implementado.
  - Sí, más o menos lo esperado.
  - Sí, pero menos de lo esperado.
  - Sí, más de lo esperado.
  - No.

Se envió un formulario con distintas preguntas abiertas y cerradas y aunque la muestra de la encuesta no es representativa, nos sirve para hacernos una idea del uso del BI en las PyMEs españolas.

Con un total de 30 respuestas, estos fueron los resultados más interesantes:

- 5 de las empresas encuestas (16,7%) dicen no saber que es el BI, incluso 4 de ellas (13,3%) no han oído hablar de BI. El resto de las empresas (70%) sí que saben que es el BI.
- En la pregunta abierta sobre en que se apoyan para la toma de decisiones, 18 de las empresas (60%) se apoyan en los datos recopilados anteriormente. 28 (93,3%) se basan

también en las experiencias previas, solo 11 empresas (36,7%) toman sus decisiones apoyándose en la intuición.

- En cuanto al uso de las redes sociales, 11 (36,7%) de las empresas no hacen uso de las redes sociales. Una de las empresas (3,3%) solo hace uso de las redes para estar en contacto con clientes mientras que 9 empresas (30%) hacen uso de las redes como herramienta de marketing. 9 de los encuestados (30%) hacen uso de las redes con estas dos finalidades, tanto marketing como estar en contacto con los clientes.
- En cuanto a las herramientas con las que trabajan para el almacenamiento y tratamiento de datos:

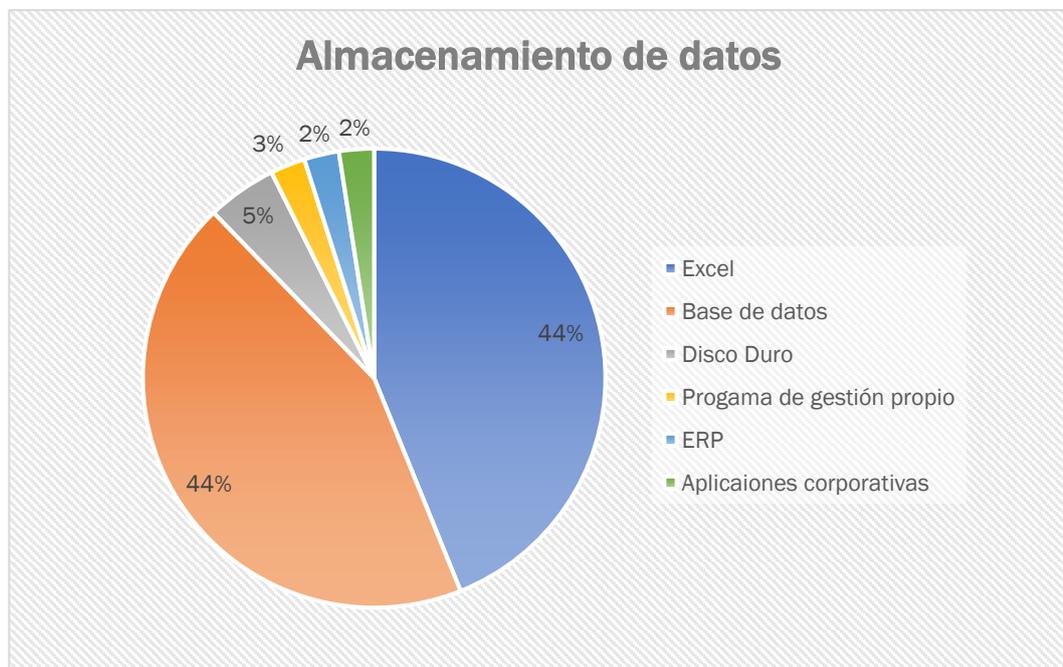


Ilustración 3.2. Resultados a la pregunta sobre el almacenamiento de datos.

Pongo los mismos términos usados por los encuestados

- Sobre la existencia del BI en la empresa, solo 12 empresas (40%) afirman disponer de ellas.
- Destaca que 13 de las empresas (43,3%) no se plantean implementar herramientas de BI, 3 (10%) de las empresas si querrían hacerlo pero creen que no pueden permitírselo en este momento y 14 (46,7%) de las empresas encuestadas quiere implementar de cero o continuar implementando herramientas de BI.
- El 15,4% de las empresas que tienen herramientas de BI en su empresa no han tenido problemas. Por lo contrario, el resto

han tenido problemas con los plazos de implantación, los reportes no eran los esperados, problemas con los costes finales e incluso problemas a la hora de alinear las necesidades de la empresa con las herramientas.

- En 12 de las empresas los reportes los generan el gerente (40%), en 8 (26,7%) los genera el área tecnológica, mientras que en el resto de empresas (33,3%), son otros quienes se encargan.
- El 37,5% de las empresas que han implementado el BI obtuvieron los resultados esperados. El mismo porcentaje considera que obtuvieron menos mejoras de lo esperado; el 25% restante cree haber obtenido más mejoras de lo que esperaban inicialmente.

Me parece importante destacar que algunas empresas que realizaron el cuestionario y no son PyMEs pues tiene más de 3000 trabajadores, desconocen que es BI y por lo tanto no hacen uso de herramientas.

Gracias al cuestionario, podemos ver como las PyMEs españolas desconocen que es el BI y por lo tanto no obtienen una ventaja competitiva gracias a este.

### 3.4. Software para aplicar BI en una empresa

En este capítulo expondré brevemente las principales herramientas que tenemos para la aplicación de BI y explicaré porqué he decido trabajar con Power BI.



Ilustración 3.3. Logo IBM Cognos.

Con **Cognos Analytics** los profesionales de las empresas podrán implementar inteligencia de negocios a través de su organización para mejorar el rendimiento y obtener una ventaja competitiva. Es una herramienta para profesionales, ideal para las grandes empresas, no para pymes. [32]

**Tableau** ofrece la simplicidad y autoservicio demandado por los usuarios junto con el gobierno y la automatización que espera. Ofrece un enfoque guiado a analytics por empoderar a los usuarios a resolver problemas individualmente o en grupo de trabajo por cuenta propia, mientras que proporciona una solución de autoservicio gobernado que puede fácilmente ampliarse cuando el negocio necesite crecer.



Ilustración 3.4 Logo Tableau

**Tableau**, es una herramienta de BI que permite analizar, visualizar y compartir grandes volúmenes de información en forma rápida, flexible y amigable.

Tableau está orientado a que personas de todos los ámbitos de negocio puedan manejar información fácilmente y analizarla en forma atractiva. El usuario tan solo tiene que arrastrar los campos de su interés para lograr el cruce de información y obtener una atractiva visualización. Con esto se evitan engorrosos procesos de carga de datos y se elimina la necesidad de programar. [33]



Ilustración 3.5. Logo Pentaho.

**Pentaho** es una herramienta de BI desarrollada bajo la filosofía del software libre para la gestión y toma de decisiones empresariales. Es una plataforma compuesta de diferentes programas que satisfacen los requisitos de BI. Ofreciendo soluciones para la gestión y análisis de la información, incluyendo el análisis multidimensional **OLAP**, presentación de informes, minería de datos y creación de cuadros de mando para el usuario.

Existen 3 productos con diferentes enfoques:

*·Pentaho Report Designer:*

Editor basado en eclipse con prestaciones profesionales y de calidad y con capacidad de personalización de informes a las necesidades de negocio destinado a desarrolladores.

*Pentaho Report Design Wizard:*

Herramienta de diseño de informes, que facilita el trabajo y permite a los usuarios obtener resultados de forma inmediata. Está destinada a usuarios con menos conocimientos técnicos.

A través de pasos sencillos permite conectarse a todo tipo de bases relacionales, integrar el resultado dentro del portal pentaho. [34]



Ilustración 3.6. Logo Tibco.

**TIBCO Spotfire** es una herramienta de BI bastante versátil y con una gran comunidad online. Su principal diferencial se encuentra

en el análisis de los procesos de ventas y la experiencia de cliente.

Con Spotfire Web Player se pueden explorar datos mediante visualizaciones interactivas en forma de gráficos de barras, tablas cruzadas, gráficos de dispersión, etc., así como flujos de trabajo guiados que lo ayudarán a dar respuesta a los eventos que afecten a su negocio. Spotfire Web Player permite filtrar los datos de manera interactiva y ofrece respuestas al instante. Con Spotfire Web Player se pueden encontrar tendencias y patrones, y obtener información detallada sobre datos específicos al mismo tiempo.

Spotfire Web Player es una herramienta eficaz para compartir y distribuir información dentro y fuera de una organización. Se trata de una aplicación basada en navegador web, de modo que no necesita instalación y puede usarse dentro de un portal o sitio web. [35]



Ilustración 3.7. Logo Ab Initio.

**Ab Initio** comprende una amplia variedad de soluciones relacionadas con el procesamiento de datos en paralelo, bajo la que podemos entender

información estructurada y no estructurada tanto por lotes como analizando eventos complejos. Su tecnología trabaja mediante la interfaz gráfica de usuario. [36]

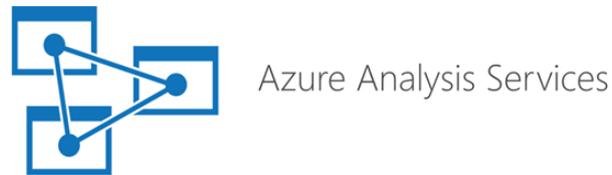


Ilustración 3.8. Azure Analysis Services.

**Azure Analysis Services** es una herramienta de Microsoft de uso exclusivo en la nube. Este motor de análisis permite visualizar los datos con distintas herramientas de datos (por ejemplo, Power BI), caracterizándose por su gran velocidad y disponibilidad así como por su capacidad para conectar datos operativos en tiempo real con DirectQuery. [37]



Ilustración 3.9. Logo Oracle

**ORACLE**, una de las empresas más importantes del sector también posee su propia herramienta de analítica, aunque quizás sería más apropiado decir que posee una solución de BI específica para cada caso de uso y que integra muchas de estas funcionalidades en sus principales soluciones de negocio. Como es obvio, especialmente recomendado para análisis de datos en bases de datos SQL. [38]



Ilustración 3.10. Logo SAP Business Objects.

SAP **Business Objects** permite convertir datos -principalmente financieros y procedentes del ERP de la firma- en información útil, distribuyéndola a los ejecutivos que la necesiten. Además, permite completar esos datos con otras fuentes externas y consumir los resultados desde cualquier dispositivo. Un completo panel de control de la empresa quizás menos sencillo que el de sus rivales, pero más completo a la hora de examinar procesos de negocio. [39]



Ilustración 3.11. Logo Microstrategy.

**Microstrategy** quizás sea el proveedor más antiguo de BI, sector en la que lleva operando desde 1989. Creada en origen para analizar y dar forma a datos relacionales, sus últimas versiones ya permiten crear cuadros de mando personalizados en los que atender tendencias y previsiones con información alojada en Hadoop y otras BBDD y motores específicos del Big Data y la información no estructurada. [40]



Ilustración 3.12. Logo Microsoft Dynamics

**Microsoft Dynamics** destaca por ser una herramienta que además de ser una solución de BI, permite gestionar otros ámbitos de la empresa como la gestión financiera, contabilidad, cadena de suministro y fabricación, inventario, pedidos, proveedores, ventas servicio posventa, creación automática de informes con los datos generados. [41]

**Ultimus** proporciona soluciones y servicios de tecnología basada en bpm para automatizar y optimizar los procesos de negocio para resolver sus retos de negocio - de forma rápida y eficiente.



Ilustración 3.13. Logo Ultimus

Tecnología BPM: La Ultimus Business Process Management Suite es una plataforma de software completa en toda la empresa que le da un control total sobre su negocio.

Soluciones Empresariales: Ya sea que esté buscando una solución por la industria, la disciplina núcleo, o un desafío empresarial, Ultimus puede ofrecer las mejores soluciones de su clase para adaptarse a las necesidades de su empresa.

Servicios de Ciclos de Vida Maximizar el rendimiento de su empresa con integraciones propias, entrenamiento de software, soporte global, asesoría en sitio, y mantenimiento. [42]



Ilustración 3.14. Logo SharePoint

**Microsoft SharePoint** es una plataforma de colaboración empresarial, formada por productos y elementos de software que incluye, entre una selección cada vez mayor de componentes, funciones de colaboración, basado en el navegador web, módulos de administración de procesos, módulos de búsqueda y una plataforma de administración de documentos (gestión documental)

SharePoint puede utilizarse para sitios de webhost que acceda a espacios de trabajo compartidos, almacenes de información y documentos, así como para alojar aplicaciones definidas como los wikis y blogs. Todos los usuarios pueden manipular los controles propietarios llamados "web parts" o interactuar con piezas de contenido, como listas y bibliotecas de documentos. [43]



Ilustración 3.15. Logo QlikView.

**QlikView** es una plataforma de business discovery que ofrece un BI de autoservicio para todo tipo de usuarios de una empresa u organización. Con QlikView puede analizar datos y utilizar sus descubrimientos para apoyar la toma de decisiones. QlikView le permite formular y responder sus propias preguntas y seguir sus propias rutas hacia el conocimiento. QlikView le permite a usted y a sus colegas llegar a las decisiones de forma colaborativa. [44]

Jet Reports es una avanzada herramienta de reporting y BI sobre Microsoft Dynamics NAV que permite navegar con facilidad por los grandes volúmenes de información que una organización genera, obteniendo los datos que se desean y llegando al mismo origen gracias a sus potentes capacidades de desglose (*drill-down*).



Ilustración 3.16. Logo JeyRepors

Con Jet Reports se analizan con rapidez y precisión la evolución del negocio, apoyado en los informes y cubos OLAP, con gráficos y tablas fáciles de interpretar.

Además, el software de reporting y Business Intelligence Jet Reports se adapta a las necesidades de las empresas en estos ámbitos, permitiendo que los usuarios puedan fácilmente crear desde listados básicos hasta complejos informes de BI. [45]



Ilustración 3.17. Logo Sisense.

Sisense hace que sea fácil revelar instantáneamente información comercial de datos complejos: cualquier fuente de datos, cualquier tamaño. Se centra en la redefinición de todos los aspectos del análisis empresarial para facilitar que cualquier persona descubra información comercial; está diseñado para ayudar a los usuarios empresariales a tomar datos complejos y crear instantáneamente información que impulse a toda su empresa. [46]



Ilustración 3.18. Logo Clear Analytics.

Clear Analytics proporciona soluciones de software que requieren menos recursos humanos en la recopilación y ensamblaje de datos y completa la tarea en un tiempo inhumanamente imposible. Ayuda a predecir los riesgos y analizar el crecimiento con anticipación,

permitiéndoles tomar mejores decisiones de negocio. Incluso los empleados sin experiencia podrán producir informes y análisis de alto nivel en Excel, ayudando así a mejorar el negocio. [47]



Ilustración 3.19. Logo Google Analytics.

**Google Analytics** permite hacer un seguimiento de sitios web, blogs y redes sociales. Además, pone a tu disposición informes predeterminados y personalizables. Tal y como se indica en su entrada en la Wikipedia, Analytics “ofrece información agrupada del tráfico que llega a los sitios web según la audiencia, la adquisición, el comportamiento y las conversiones que se llevan a cabo en el sitio web”.

De hecho trabaja sobre elementos tan variados e importantes como :el número de visitas, la duración de las mismas, las fuentes de tráfico o las páginas visitadas

Al ser una herramienta de Google es la que mejor compatibilidad tiene con el resto de herramientas de la compañía, lo que tratándose de Google supone una gran ventaja. [48]



Ilustración 3.20. Logo GoodData.

**Good Data** es una organización que se basa en un esquema empresarial de cooperativa, en la que cualquier persona puede convertirse en socia. No es un herramienta gratuita, pero permite la realización de un gran número de gráficos y es bastante intuitiva. [49]



Ilustración 3.21. Logo Eclipse.

Eclipse Birt Project es un proyecto de software de código abierto que proporciona capacidades de creación de informes y de inteligencia de negocio para clientes pesados (fat clients) y aplicaciones web, especialmente aquellas basadas en Java y Java EE. BIRT es un proyecto de software de alto nivel dentro de la Eclipse Foundation (fundación Eclipse), un consorcio de proveedores de la industria del software sin ánimo de lucro y una comunidad de código abierto.

El objetivo del proyecto es cubrir un amplio rango de necesidades de creación de informes dentro de una aplicación típica,<sup>1</sup> abarcando desde informes operacionales o de empresa hasta procesamiento multi-dimensional analítico en línea (OLAP). Inicialmente, el proyecto se ha enfocado en capacidades que permitan a los desarrolladores de aplicaciones diseñar e integrar fácilmente informes dentro de aplicaciones.

BIRT tiene 2 componentes principales: un diseñador de informes visuales dentro de Eclipse IDE para crear informes BIRT, y un componente de rutina para generar informes que pueden ser puestos en uso en cualquier entorno Java. [50]



Ilustración 3.22. Logo Power BI.

**Microsoft Power BI** es la solución destinada a la inteligencia empresarial, que permite unir diferentes fuentes de datos (más de 65), modelizar y analizar datos para después, presentarlos a través de paneles e informes; que puedan ser consultarlos de una manera muy fácil, atractiva e intuitiva.

Power BI puede unificar todos los datos de su organización y externos, se encuentren a en la nube o localmente, podemos utilizar

información de un Excel propio, de Dynamics CRM, Dynamics NAV, SAP, MailChimp, Salesfoce, Facebook, Google Analytics, SQL Server... o de datos externos. Por ejemplo, datos del Instituto Nacional de Estadística, del Banco Mundial.... Y un largo etcétera. [51]

Con Power BI los usuarios sin necesidad de grandes conocimientos, podrán crear paneles, gráficos e informes de todo tipo.

Es tan sencillo porque con los datos preparados (modelización), los usuarios solo tendrán que elegir y arrastrar qué tipo de gráficos desean que aparezcan en el panel o informe.

Con Power BI podemos visualizar todo aquello que nos interese, por ejemplo:

- Como van nuestros **flujos de caja**.
- Número de **oportunidades de negocio**.
- Número de **ofertas**,
- **Pedidos** y sus **facturas**.
- Conocer en un **mapa** qué comunidades autónomas, regiones y países **donde tengo a mis clientes**, compran más mis productos y servicios.
- Como van las campañas de **marketing**.
- Logro de los **objetivos**...

Además, dispone de una **aplicación para móviles**.

### **3.5. ROI con la implantación de herramientas de BI**

El tema del ROI en Business Intelligence siempre ha sido un tema complejo y debatido, nunca ha quedado muy claro si es efectivo medirlo. Laura Madsen,[52], experta en este campo, escribió el artículo “*Now is the Time: An ROI for Business Intelligence*” [53] en el que me baso en este apartado:

## Cómo medir el valor de las iniciativas de BI

En una gran empresa siempre se están evaluando los costes frente a los beneficios de una decisión empresarial. Por este motivo es importante conocer los costes que conlleva cualquier iniciativa.

En el caso del BI, es bastante complicado medir esto, por eso que la gran mayoría de empresas no se hayan aventurado a evaluar dichos costes.

Actualmente, con las reducciones de costes en todas las áreas, se ha hecho inevitable valorar el coste y la rentabilidad de todas las iniciativas antes de afrontarlas.

La métrica más usada para poder medir estas iniciativas es el Retorno de la inversión (ROI).

Hay dos razones principales por las que se calcula el ROI:

- Para incorporarlo dentro de un plan de proyecto que se va a iniciar y necesita cuantificar los fondos a destinar.
- Para medir el resultado económico de un proyecto que acaba de concluir.

En el primer caso, el ROI servirá para conocer qué resultados o ventajas me puede proporcionar el sistema por anticipado, incluso justificar la no realización de un proyecto.

En el segundo caso, nos servirá para poder comparar con el ROI inicial previsto y analizar las desviaciones.

Para poder calcular el ROI de forma adecuada tendremos que identificar por separado todos los costes en que hemos incurrido, así como los beneficios podremos obtener cuando la iniciativa esté completada o en producción.

### Identificación de los costes:

En primer lugar, hay que identificar los tipos de costes. Los principales tipos de costes son de hardware, software y trabajo.

**Hardware** se refiere a toda la infraestructura física que se necesita para implementar un DW o Business Intelligence. Servidores, puestos cliente, redes, discos, copadoras, dispositivos móviles, dispositivos externos, etc...

Los costes de **software** incluyen todos los gastos por licencias que se tienen que pagar a los fabricantes. Se incluyen las herramientas ETL, la base

de datos, las herramientas de análisis y reporting, balance scorecard, data mining... términos que hemos visto con anterioridad.

Los costes **laborales** incluyen como en cualquier otro proyecto el trabajo tanto de los propios miembros de la empresa como de posibles subcontratas.

También podemos clasificar los costes como iniciales o repetitivos:

Los **costes iniciales** son aquellos que solo se realizan en una ocasión y no se repiten (licencias de software, hardware, costes laborales especificados para la implementación y formación).

Los **costes repetitivos** son aquellos en los que se sigue incurriendo, una vez que el sistema BI o DW ya está en funcionamiento (costes de administración del sistema, soporte, nuevos desarrollos, resolución de problemas, actualizaciones de software, etc...)

Toda esta información previa sobre costes se la podemos pedir a los propios vendedores, a las empresas de integración o consultoría.

#### **Identificar los beneficios:**

Los beneficios son bastante más difíciles de calcular que los costes. Es en esta parte donde se pueden cometer errores y es más difícil interpretar los resultados del ROI. Para simplificar este paso separamos los beneficios obtenidos en incrementos de ingresos y en reducción de costes.

**Incrementos de ingresos.** Es muy difícil calcular que parte del incremento de ingresos se debe a una mejor utilización de la información, Sin embargo será factible achacar las nuevas oportunidades o detección de hábitos del consumidor con un mejor análisis, pero cuantificarlo también será difícil. Con todo y con eso, debemos intentar cuantificar esas ventajas competitivas generadas gracias a la mejora en la toma de decisiones.

Estos son algunos de los ejemplos que pueden encuadrarse en este grupo:

- Incremento de la cifra de negocio.
- Identificar nuevas oportunidades de negocio.
- Mejorar la satisfacción y el servicio a los clientes.
- Mejorar el tiempo respuesta al mercado.
- Acelerar el desarrollo de nuevos productos.
- Facilitar la toma de decisiones.
- Aumentar la habilidad para hacer frente a los competidores.

**Ahorro de Costes.** Este tipo de beneficios son más fáciles de calcular puesto que podemos comparar los costes del sistema actual con los antiguos sistemas de reporting, análisis y almacenamiento.

Además, es mucho más fácil cuantificar y dar números para estos beneficios.

Estos son algunos ejemplos:

- Reducción de los costes operativos.
- Automatización de procesos manuales.
- Mejora de los procesos operativos.
- Mejora de la agilidad de la organización.
- Aumenta la diseminación de la información.
- Mejora el análisis.
- Mejora la toma de decisiones.

### 3.6. 10 beneficios BI para una empresa

EL objetivo de los sistemas de información es ayudar a tomar decisiones. Según Saima Solución [53], empresa líder en implementación de herramientas BI, estos son los 10 principales beneficios que obtenemos gracias a su uso.

#### 1. Nos olvidamos de las especulaciones

Nos olvidamos de la toma de decisiones mediante la intuición o esperando suerte, gracias al BI tomamos decisiones fundamentadas en datos e información, las decisiones podrán ser o no correctas, pero eso ya dependerá de los conocimientos de quien las tome.

#### 2. Nos da respuestas más rápidas

Otro gran incentivo para implementar una herramienta de BI es que puedes conseguir respuestas inmediatas a las preguntas de tu negocio. Puedes tomar decisiones sin tener que revisar un montón de páginas de informes y perder el tiempo.

#### 3. Puedes conseguir importantes métricas sobre tu negocio

Con los programas de BI, cualquier miembro de tu organización puede tener acceso a los datos de tu compañía, por ejemplo a las métricas de tu negocio, los cuadros de mando y otros informes importantes. Tan sólo hace falta una conexión a tus datos y puedes acceder a todo el funcionamiento de tu organización.

#### 4. Conseguir información valiosa sobre el comportamiento de tus clientes

Uno de los mayores beneficios de tener una plataforma BI en tu organización es tener la posibilidad de investigar lo que el mercado está adquiriendo. Se te da la oportunidad de convertir esa información en algo que te produzca beneficios.

#### 5. Identificar las oportunidades de venta cruzada y up-selling

Las plataformas BI te permiten utilizar los datos de clientes para mejorar tus modelos de venta tradicionales, así como para ayudar a tu personal de ventas a implementar técnicas de cross-selling y up-selling en los puntos de venta más adecuados.

#### 6. Permite optimizar las operaciones

Cuando se implementa un sistema de BI, podemos empezar a conseguir *insights* detallados sobre el rendimiento de nuestras operaciones. Esto nos permite conocer las áreas en las que necesitamos mejorar o modificar algo a fin de optimizar y racionalizar nuestras operaciones.

#### 7. Desarrollar la eficiencia

Hay muchas organizaciones que pierden una gran cantidad de tiempo buscando información desde sus diferentes fuentes de datos en los momentos en los que necesitan obtener una comprensión más profunda de su empresa. Cuando conseguimos agrupar esa información, aún necesitamos cruzar y convertir dichos datos, comunicar las opiniones sobre los resultados y ver si son lo suficiente precisos.

Con los sistemas de BI, todos los datos que necesitamos vienen desde una única fuente, podemos visualizarlos en un panel y crear rápidamente un informe. Esto ahorra mucho tiempo y energías, además de eliminar las ineficiencias de nuestros procesos de trabajo.

#### 8. Encuentra cuáles son tus costes de producción reales

Los programas BI pueden darte unas estadísticas más elaboradas sobre los costes de producción de tu compañía, así como de tu habilidad de hacer cambios en esa producción para conseguir una rentabilidad mayor.

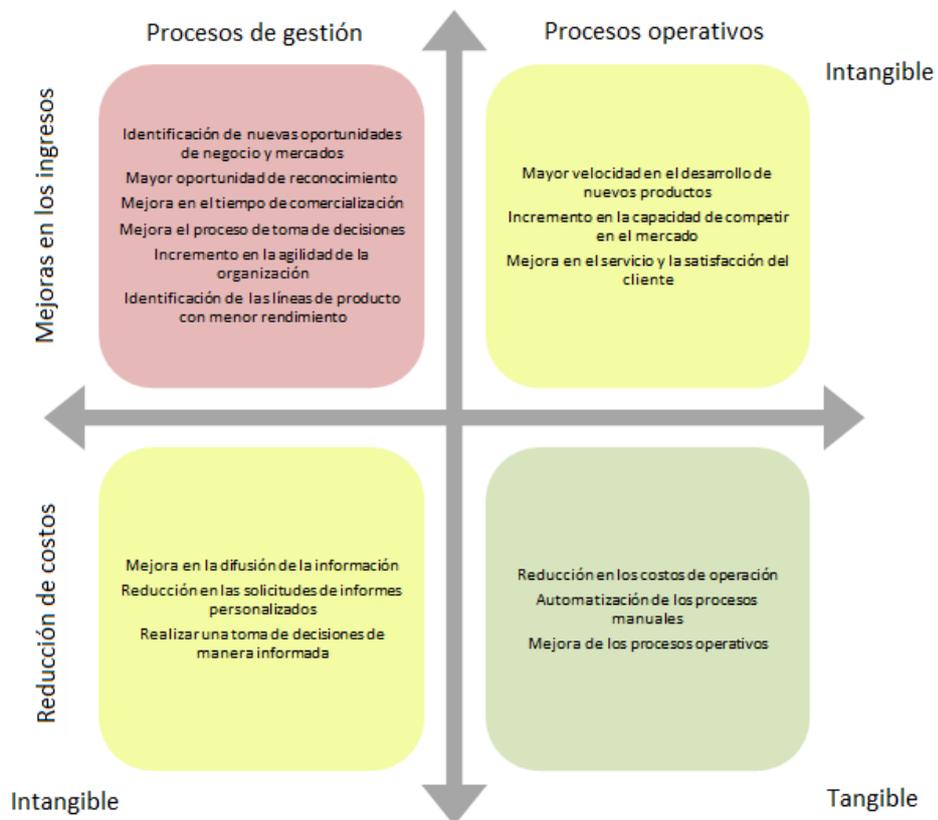
#### 9. Crear un mejor inventario

Los programas de BI nos ayudan también a llevar a cabo mejor el inventario. De esta forma, tus clientes pueden recibir lo que necesitan en el momento

en el que ellos quieren y tu empresa no tiene la carga adicional de tener que almacenar el exceso de inventario.

### 10. Entender mejor el pasado, presente y futuro de tu empresa

Uno de los puntos fuertes de los sistemas de Inteligencia de Negocio es darte la oportunidad de conocer el estado de tu negocio en un período concreto de tiempo, por ejemplo, la forma en que se vendieron las existencias de un producto determinado. Estos datos se pueden utilizar para tomar ventajas sobre tu competencia.



**Ilustración 3.23. Beneficios tangibles e intangibles de un BI** Ilustración tomada de: “Guía para la implementación de una solución de inteligencia de negocios para pequeñas y medianas empresas” de Daniel Felipe Flórez salgado y Laura Margarita del Mar Tejada Cardona.



## 4. LA EMPRESA Y LOS DATOS

En este cuarto capítulo del trabajo presentare a la empresa y las tareas que ha sido necesarias para la limpieza de datos y la creación de la base de datos.

Este capítulo y el siguiente están totalmente condicionados por los datos que la empresa me facilitó, dependiendo de la cantidad y la calidad de estos datos se podrán elaborar distintos informes de los que podremos obtener más o menos información.

### 4.1. La empresa

La empresa en la que se centra la parte práctica de este TFG, es una empresa del sector servicios, concretamente una imprenta/copistería.

La empresa podemos distinguir dos áreas claramente diferenciadas, un área tipo “B2C” y otra “B2B”. Más adelante explicaremos esta división de la empresa y que supone tanto para la empresa como para el TFG.

No es una imprenta muy grande por lo que algunos pedidos necesitan ser subcontratarlos a otras imprentas más grandes.

Datos de interés de la empresa:

- Fundada en 1995.
- Horarios:
  - Lunes a viernes de 8.30 a 14.00 – 16.00 a 20.30
  - Sábados 10.30 a 14.00
- El pasado año 2017 facturo por valor de XXXX.
- Cuenta con 4 trabajadores a lo largo del año.
- Situada en el barrio de Santa Clara de Valladolid.
- Está dividida en dos partes que explicaremos más detenidamente.
- La imprenta oferta X productos personalizados.
- Los clientes son particulares, empresas e instituciones públicas como universidades o ayuntamientos.
- Tiene clientes en varias comunidades autónomas, no solo en Castilla y León.

#### 4.1.1. Reuniones con la empresa

Para una correcta implantación de un sistema BI es necesario conocer la empresa, como funciona, que productos ofrece, a que clientes y sobre todo que espera obtener con la implementación de un sistema BI.

A lo largo del proyecto ha habido un total de cinco reuniones en la empresa además de distintas llamadas para aclarar algunas cosas habladas durante las reuniones.

La primera reunión con la empresa fue de la mano de mi tutor de TFG Adolfo. En ella conocí al propietario y le presentamos la idea del proyecto. Por su parte, el propietario nos explicó cómo trabaja la empresa, como está dividida y los problemas que tenía para poder sacar rendimiento a los datos que genera.

La siguiente reunión la tuve con el dueño y encargado de lo que llamamos “planta 1”. En esa reunión se me explicaron los procedimientos, los distintos productos que produce la empresa y los clientes con quien trabaja. Además, en esta segunda reunión se me facilitaron los datos para poder empezar a idear como iría el proyecto.

La tercera reunión tuvo lugar con la encargada de lo que llamamos “planta 0”. Esta zona es una copistería, me explico cómo trabajan y que datos genera la máquina donde cobra. Me facilito todos los datos que generan en esa zona para que también pudiese trabajar con ellos.

En la cuarta reunión, tras haber trabajado con los datos que me facilitaron, detecte varias incongruencias como duplicados de clientes debido a nombres cambiados, información muy pobre en la planta 0 para la gran cantidad de ventas y otras cosas bastante mejorables. Por este motivo concertamos otra cita para explicarme cómo funcionan los programas.

En la última reunión, además de explicarme con trabajan con los programas y cómo funcionan estos, llegamos a unas conclusiones de lo que deberían aportar los programas y como podría ayudar esto a la forma de trabajar. El dueño me explico del coste que tiene el mantenimiento anual de esos programas y los motivos por los que no se habían actualizado durante los últimos años.

Además de estas reuniones, hubo llamadas y contactos telefónicos para aclarar pequeñas dudas sobre el funcionamiento de la empresa, de cómo se elaboran los presupuestos y otros pequeños detalles que me ayudasen a entender el conjunto de la empresa para poder entender mejor los datos.

#### 4.1.2. División de la empresa

En el punto anterior he comentado que la empresa está claramente diferenciada en dos áreas. Podemos distinguir una parte de la empresa que se dedica a la venta de productos al por menos, productos elaborados por ellos como fotocopias o encuadernaciones, y productos elaborados por otras empresas como bolígrafos, celo, grapadoras... El otro área se dedica a la venta de productos personalizados a empresas o instituciones públicas aunque también venden a particulares.

Esta división no es solo de negocio, si no que dentro del local de la empresa el área B2B está en la primera planta, mientras que el área B2C está en la planta baja.

B2B se refiere a empresas que venden a otras empresas y B2C se refiere a empresas que venden a particular.

Para poder comprar algo al por menor, es necesario acudir a la tienda donde se atiende en un mostrador. En cambio, para ventas al por mayor o servicios personalizados, además de poder hacerlos acudiendo a la tienda, se pueden realizar los pedidos por teléfono o por email.

Otra diferencia entre las áreas son los trabajadores, cada área tiene sus trabajadores y por lo general, salvo necesidad, no se trabajan en la otra área.

También se diferencian en los productos que venden, el B2B vende una variedad más reducida de productos, pero de un valor más alto al ser elaborados y personalizados para cada cliente, en cambio en la otra área se venden un gran número de productos distintos, pero de un valor mucho menor.

Posiblemente la diferencia más importante para lo que nos atañe para este TFG, es que cada área usa unos programas distintos para presupuestar y facturar. Esto supone que cada área genere los datos en un formato, con distinta información incluso que haya duplicidad de clientes con nombres cambiados debido que comprenden tanto al por mayor como al por menor.

Además, el área B2B genera mucha menos cantidad de datos, pero con más información, hay menos ventas, pero las que hay se tiene registrado el producto, la cantidad, cliente... En cambio, en la sección de la planta baja se realizan un gran número de ventas al día, productos que pueden valer menos de 50 céntimos y por ese motivo no se registra tanta información de cada transacción.

Debido a estos motivos, tome la decisión de realizar por separado dos proyectos de implantación de BI, uno para la zona que llamaré “planta 1” y otro para “planta 0”. Estos nombres se deben a la distribución física de las áreas dentro del local.

### 4.1.3. Como trabaja la empresa

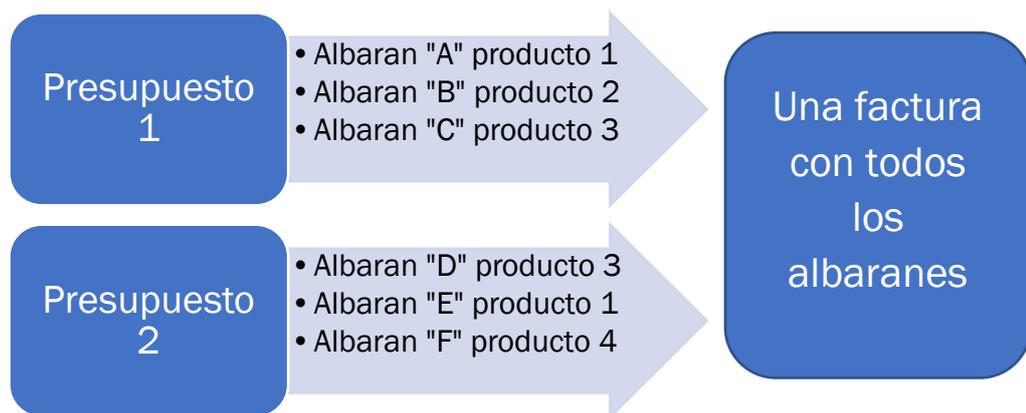
#### Planta 0

En la planta 0 la forma de trabajar es sencilla, el trabajador marca su código, marca el precio total de la venta y se genera una línea de datos con esa información junto con la fecha.

#### Planta 1

En la planta 1 el funcionamiento es más complejo, y es fundamental comprender bien el funcionamiento para entender cómo y porque se generan los datos de una determinada forma.

Cuando un cliente pide un presupuesto, si el cliente es un cliente fidelizado o con vistas a serlo, se apunta su nombre al generar el presupuesto. Una vez que se genera el presupuesto y el cliente lo aprueba, se genera un albarán para cada tipo de producto que va en dicho presupuesto. Al cabo de mes, se genera una factura que incluye todos los albaranes de ese mes de ese cliente. Además, si un cliente hace el mismo pedido más de una vez no se genera presupuesto nuevo, si no que se trabaja con el anterior.



**Ilustración 4.1. Ilustración explicativa de la forma de trabajo de la empresa.** Ilustración de creación propia que explica como genera las facturas, los presupuestos y los albaranes.

#### 4.1.4. Programas que utiliza la empresa



Ilustración 4.2. Logo de Sage – Factura plus.

En la planta 0 utilizan el software **Facturaplus** para las cosas que necesitan una factura y no un ticket. Por ejemplo, para venta de grandes cantidades de un producto a una empresa. Facturaplus es un programa elaborado por Sage y es uno de los mejores programas de facturación, diseñado para facilitarte el control de los procesos comerciales de tu negocio: desde la emisión de documentos de compra / venta hasta el control de las situaciones de stock, la gestión del reaprovisionamiento o la emisión de remesas bancarias, entre otras operativas.

Además, este software de facturación, gracias a sus listados, informes y gráficos te permite conocer en todo momento el estado de tu negocio, y su histórico, ayudándote a tomar las mejores decisiones y a satisfacer las necesidades de tus clientes.

Esta información es la que podemos encontrar en su web, pero en el negocio en el que se centra el TFG, no aprovechan todas las herramientas que nos facilita Facturaplus.

Para los cobros en caja y la gestión del TPV, usan un programa personalizado para la empresa creado por **AMG**, en el tienen acceso fácil a los diferentes productos que venden pero tiene un problema importante para sacar rendimiento a los datos que más adelante veremos.

### 4.2. Los datos

Cualquier proyecto de implantación de un sistema BI estará condicionado por los datos de los que dispone la empresa. Es evidente que cuantos más datos disponga la empresa más información podremos obtener

de ellos y como consecuencia una mayor ventaja competitiva frente a la competencia. Pero no es solo una cuestión de cantidad de datos, si no que también es muy importante la calidad de estos.

En esta parte práctica en la que presentaremos las dos áreas de la empresa por separado, veremos cómo en una hay una mayor cantidad de datos al haber muchas más ventas, pero en el otro área, veremos que aunque haya menos ventas y con ellos se generen menos filas de datos, estos datos son más completos, con información de producto, cliente, fechas de presupuesto y facturación de tal modo que podremos sacar más información.

Los datos de particulares y empresas no serán facilitados en este documento, si no que cada cliente tendrá un número.

#### **4.2.1. Datos facilitados por la empresa**

En ambas áreas, la empresa facilito unos datos estructurados pero los cuales eran necesario limpiar para poder trabajar con ellos.

Hay datos externos que solo he utilizado en una de las áreas pudiendo haberlos utilizado en las dos. He decidido hacerlo así para ver que no es necesario hacer siempre lo mismo, si no que podemos aportar distintos datos a los que ya tenemos en función del proyecto, que aunque no completemos correctamente nuestros datos por desconocimiento no quiere decir que nuestro sistema este mal hecho, simplemente nos dará menos información de la que podríamos sacar.

##### **Datos planta 0**

En esta área se nos facilitaron varios archivos en formato Excel. Los datos correspondían al año 2017 puesto que no disponían de años anteriores. Al ser el área de la copistería hubo una gran cantidad de ventas durante el año 2017. Más de 20.000 ventas a lo largo del año que van desde 0,05€ hasta 850€.

Información que contenían los datos:

Cada línea contiene la siguiente información

- Documento: Id de la venta
- Agente: Trabajador que hizo la venta
- Cliente: En todos los casos es “cliente al contado”
- Fecha: Día en el que se hizo la venta.
- Base: Importe venta sin IVA
- Cuota: Importe venta con IVA

- Importe: Importe venta + IVA

Documento	Fecha	Ciente	Agente	Base	Cuota	Importe
114612	03/01/2017	Ciente de Contado	1	157,73	33,12	190,85
114613	03/01/2017	Ciente de Contado	2	16,40	3,45	19,85
114614	03/01/2017	Ciente de Contado	2	2,46	0,52	2,98
114615	03/01/2017	Ciente de Contado	2	2,89	0,61	3,50
114616	03/01/2017	Ciente de Contado	3	1,12	0,23	1,35
114617	03/01/2017	Ciente de Contado	3	0,83	0,17	1,00
114618	03/01/2017	Ciente de Contado	1	2,93	0,62	3,55
114619	03/01/2017	Ciente de Contado	1	6,17	1,30	7,47

Ilustración 4.3. Tabla datos planta 0 .

La primera vez que vi los datos pensé que con esos datos poco podría hacer, únicamente cruce el importe de las ventas con el agente y con la fecha. De esta forma obtuve cuanto vendía cada trabajador y en qué días se vendió más.

La solución a esta falta de datos es completar los que nos facilita la empresa con otros datos que podemos obtener y/o “crear”. Para ello generaremos otros archivos con datos que poder cruzar con los que ya teníamos. De este apartado de obtención y generación de datos nuevos hablaremos más adelante.

### Datos planta 1

De esta planta, la empresa pudo facilitarme los datos de cuatro años. Estos datos eran mucho más completos y más complejos de tratar pero tenían muchos errores que hacían que la información obtenida de ellos fuese errónea.

La empresa facilito 4 archivos formato “.txt”:

- Albaranes
- Clientes
- Facturas
- Presupuestos

**Albaranes:** El archivo albaranes contenía muchos campos pero gran cantidad de ellos no servían para nada puesto que eran redundantes, no aportaban valor o no eran posibles cruzarlos con otros. Los campos eran:

- N°: Numeración empezando por el 1
- Número pedido líneas: otra numeración

- N° Completo presupuesto: Número que identifica a que presupuesto corresponde el albara
- Fecha: La fecha del albarán
- Nombre cliente
- Cantidad líneas: Cantidad de producto
- Producto pedido: Producto del albarán
- Estado: Recogido/No recogido
- Impreso: Si/No
- Facturado Si/no
- Cancelado Pedido Líneas: Si/No

Albaranes: Bloc de notas

No.	Numero Pedido	Líneas	N° completo Presupuesto	Pedido	Líneas	Fecha	Nombre corporativo	Destino	Cantidad	Líneas	Producto	Pedido	Líneas	Referencia	Estado	Impreso	Factu
1	2008030001	2/3	10/03/2008	1	20000	Sobres	calculo	Recogido	Sí	Sí	Sí						
2	2008030001	2/3	20/03/2008	1	0	Sobres	calculo	Recogido	No	Sí	Sí						
3	2008030002	4/1	20/03/2008	1	1000			Recogido	No	Sí	Sí						
4	2008030003	10	27/03/2008	2	GIRO, S.L.	1000			Recogido	Sí	Sí	Sí					
5	2008040002	65	13/04/2008		10			Recogido	No	Sí	Sí						
6	2008040003	101/1	17/04/2008		200			Recogido	Sí	Sí	Sí						
7	2008030001	2/3	25/05/2008	PACO	0	Sobres	calculo	No listo	No	Sí	Sí						
8	2008040001	54	25/05/2008	2	.	65000		Recogido	No	Sí	Sí	Sí					
9	2008060001	143	02/06/2008		S.L.	400		Recogido	No	Sí	No	Sí	No				
10	2008030002	4/1	02/06/2008		0		No listo		No	Sí	Sí						
11	2008050004	114/1	02/06/2008		1000			Recogido	No	Sí	No						
12	2008050003	122	02/06/2008		1000			Recogido	No	Sí	No						
13	2008050002	131/1	02/06/2008		10			Recogido	No	Sí	No						
14	2008050001	127	02/06/2008		2000			Recogido	No	Sí	No	Sí	No				
15	1	124/3	02/06/2008		L50000			Recogido	No	Sí	No						
16	2009030001	23/1	27/03/2009		250			Recogido	No	Sí	No						
17	2009040003	159/5	02/04/2009		3700			Recogido	No	Sí	No	Sí	No				
18	2009040002	157/3	02/04/2009		4000			Recogido	No	Sí	No						
19	2009040001	159/1	02/04/2009		3700			Recogido	No	Sí	No						
20	2009040005	172	08/04/2009		10			Recogido	Sí	Sí	No						
21	2009040007	182	14/04/2009		12000			Recogido	Sí	Sí	No						

Ilustración 4.4. Tabla datos albaranes de la planta 1.

**Cientes:**

- Nombre: Nombre empresa
- Código postal
- Población
- Teléfono

Nombre	Código postal	Población	Teléfono
CO	47003	Valladolid	
1 A 1	47014	Valladolid	9
1 A 2	47012	Poligono de San Cristobal	
1RD 3	47172	Villarmentero de Esgueva	6
A P	47170	Renedo de Esgueva	9
A.M.P.A.		47011 Valladolid	
	Z	47011 Valladolid	9
DE	A	47340 Sardón de Duero	
	S.L.	26004 Logroño	
A		47195 Arroyo de la Encomienda	98

Ilustración 4.5. Tabla datos clientes de la planta 1.

### Facturas:

- Número: todo "0"
- Fecha
- Número Pedido Segmentos: Número
- Nª Completo presupuesto pedido segmentos: Nº presupuesto
- Nombre cliente
- Referencia
- Segmentos
- Subtotal
- Total IVA
- IVA incluido

Numero	Fecha	Numero Pedido Segmentos N°	Segmentos	Nombre Cliente	Referencia Segmentos	Subtotal (€)	Total IVA (€)	IVA incl. (€)
0	04/05/2009	2009040023	258	Folleto Publicitario	133,00	21,00	154,00	
0	04/05/2009	2009050010	200		665,00	106,00	771,00	
0	04/05/2009	2009040021	242	ES Sala Factory 16 05 09			616,00	616,00
0	04/05/2009	2009040020	241	Sala Factori 16 05 09	616,00		616,00	
0	04/05/2009	2009040010	194		72,00	12,00	84,00	
0	04/05/2009	2009040005	172		69,00	11,00	80,00	
0	04/05/2009	2009030001	237		200,00	32,00	232,00	
0	04/05/2009	2008050003	122		112,00	18,00	130,00	
0	04/05/2009	2008030002	4/1		137,00		993,00	
0	04/05/2009	2008060001	143			137,00		
0	04/05/2009	2008050001	127		530,63	84,90	615,53	
0	04/05/2009	1 124/3	PUB		119,92	869,41		
0	30/04/2009	2009040007	182		673,00	100,00	773,00	
0	30/04/2009	2009040006	181	Publicitario	603,00	96,00	699,00	
0	04/05/2009	2009040022	246		64,00	466,00		
0	04/05/2009	2009040009	192	Control transporte	402,00	64,00	466,00	
0	04/05/2009	2009040019	235	oja Expandida	87,00	14,00	101,00	
0	04/05/2009	2009040024	258	Cartel Publicitario	387,00	62,00	449,00	
0	04/05/2009	2009040026	260	Folleto Publicitario	387,00	62,00	449,00	
0	04/05/2009	2009040011	210	Facturas	0,00		0,00	
0	04/05/2009	2009040013	173		59,00	428,00		
0	04/05/2009	2009040015	203	titación	369,00	59,00	428,00	
0	04/05/2009	2009040016	203	ograma	369,00	59,00	428,00	
0	04/05/2009	2009040014	211	arios Recibos	98,00	16,00	114,00	
0	01/06/2009	2009050021	309	multitectura Ambiental"	597,50	95,60	693,10	
0	01/06/2009	2009050012	272	legas Hmos. Villar	597,50	95,60	693,10	
0	01/06/2009	2009050013	271	DE CASTILLA	597,50	95,60	693,10	
0	01/06/2009	2009050004	274		0,00	0,00		
0	01/06/2009	2009050001	268		0,00	0,00		
0	01/06/2009	2009050002	269		0,00	0,00		
0	01/06/2009	2009050003	270		0,00	0,00		
0	02/06/2009	2009050024	310	MENTERO Carteles "Fiesta Mancomunidad"	0,00		0,00	
0	02/06/2009	2009050039	365	itario	217,00	34,72	251,72	
0	26/06/2009	2009060038	496	Sobre Americano	89,35	14,30	103,65	
0	26/06/2009	2009060019	467	BAUTISTA Triplico	235,80	37,73	273,53	
0	26/06/2009	2009060001	407	Talonnario Facturas	169,50	27,12	196,62	

Ilustración 4.6. Tabla datos facturas de la planta.

### Presupuestos:

- N° Completo: N° Presupuesto
- Fecha
- Producto
- Referencia
- Páginas
- Nombre cliente
- Teléfono
- Contacto
- Cantidad
- Venta Total
- Coeficiente: Coeficiente que pone el trabajador
- Último pedido: Supuesto último pedido de ese mismo cliente

Nº completo	Fecha	Producto	Referencia	Páginas	Nombre Cliente	Teléfono	Contacto	ClienteMóvil	Contacto Cliente	Cantidad	Venta Total (€)	Coefficiente	ultimo pedido (fecha)	Cliente
2/3	10/03/2008	Sobres calculo	20			11557,86		0,00%	20/03/2008					
4/1	20/03/2008		4			804,92	0,00%	20/03/2008						
10	27/03/2008		2			79 97	1800	222,40	28,00%	10/06/2009				
23/1	29/03/2008		2			79 97	250	199,95	28,00%	10/06/2009				
54	07/04/2008		4			79 97	65000	2485,09	10,00%	10/06/2009				
65	13/04/2008		4			79 97	10	68,40	25,00%	10/06/2009				
101/1	16/04/2008		4			200	194,90	13,00%	27/09/2016					
114/1	04/05/2008		6			79 97	1800	251,35	28,00%	10/06/2009				
122	04/05/2008		2			79 97	1000	111,99	28,00%	10/06/2009				
124/3	15/05/2008		2			50000	749,49	14,00%	20/02/2018					
127	18/05/2008		6			609 40	15 03	2000	530,63	15,00%	27/12/2009			
131/1	24/05/2008		2			79 97	10	186,05	28,00%	10/06/2009				
143	02/06/2008		2			137,20	25,00%	02/06/2008						
156/2	27/03/2009		6			4000	1000	249,59	28,00%	05/08/2017				
157/3	01/04/2009		32			4000	3112,27	15,00%	24/07/2009					
159/1	01/04/2009		32			3700	2932,22	15,00%	24/07/2009					
159/4	01/04/2009		32			3700	2827,71	15,00%	24/07/2009					
159/5	01/04/2009		32			3700	2827,71	15,00%	24/07/2009					
164	02/04/2009		32			4000	2956,38	10,00%						
165/1	02/04/2009		4			30000	1616,53	12,00%						
165/2	02/04/2009		4			40000	1984,34	12,00%						
168	02/04/2009		2			5000	163,72	28,00%						
170	02/04/2009		2			5000	199,10	28,00%						
171	03/04/2009		48			677 84	35 31	5000	4849,97	10,00%	10/12/2012			
172	03/04/2009	Talones - Juegos	2			VIALES	4041	983 25	38 49	10	68,03	28,00%	05/08/2017	
173	06/04/2009		2			5000	289,44	13,00%	27/09/2016					
174	06/04/2009		2			10000	2942,40	10,00%						
175/1	06/04/2009		2			4000	532,59	12,00%	23/09/2010					
175/2	06/04/2009		2			36000	1350,43	12,00%	23/09/2010					
176/1	08/04/2009		6			5000	422,87	15,00%	27/09/2016					
176/2	08/04/2009		6			10000	636,16	15,00%	27/09/2016					
176/3	08/04/2009		6			20000	996,36	15,00%	27/09/2016					
177/3	08/04/2009		2			677 84	35 31	400	173,62	17,00%	10/12/2012			
177/4	08/04/2009		2			677 84	35 31	500	182,05	17,00%	10/12/2012			
177/5	08/04/2009		2			677 84	35 31	100	75,00	-40,96%	10/12/2012			
177/6	08/04/2009		2			677 84	35 31	200	116,00	-6,20%	10/12/2012			
178	08/04/2009		2			677 84	35 31	100	75,00	-40,96%	10/12/2012			
179	08/04/2009		2			4000	190,50	28,00%						
180	08/04/2009		4			10000	1265,51	17,00%	08/04/2009					
181	08/04/2009		2			10000	602,21	12,00%	12/11/2012					
182	08/04/2009		4			12000	672,62	20,00%	08/04/2009					
183	09/04/2009		48			677 84	35 31	5000	2854,16	12,00%	10/12/2012			

Ilustración 4.7. Tabla datos presupuestos de la planta 1.

Como se puede observar, aunque haya menos cantidad de ventas, los datos son más completos. Esto también da lugar a que la posibilidad de que haya defectos en los datos sea mayor.

En estos archivos vemos que en algunos casos los datos son incompletos, incongruentes, no aportan valor y están desordenados. Por este motivo es necesario una labor de limpieza de datos de la cual hablaremos más adelante.

#### 4.2.2. Limpieza de datos

Para poder trabajar con los datos, es necesario tratarlos. El proceso por el cual corregimos y eliminamos los datos de una base de datos o una tabla se conoce como limpieza de datos. Es la parte más tediosa para la implantación de un sistema BI, pero además de servirnos para implementar las herramientas de BI en nuestro negocio, nos sirve para saber qué datos tienen importancia y por lo tanto deben generar nuestro negocio con los programas que usa.

El diseño de la estructura de nuestra base de datos y la limpieza de datos son dos procesos que se retroalimentan el uno al otro, puesto que a medida que vemos que datos tenemos, podemos diseñar las tablas y también a la inversa, el diseño de la estructura nos orienta en el proceso de limpieza. En mi opinión, para realizar bien estos dos procesos, es necesario conocer bien el funcionamiento de la empresa, lo que nos permitirá

discriminar mejor entre los datos de valor y los que no y nos permitirá crear una base de datos más fiel al funcionamiento de la empresa.

**Para la implementación de un sistema BI es necesario comprender los datos que tenemos**, y por eso en este apartado explicare los datos que los que disponemos y como he tratado los diferentes archivos que me facilito la empresa, que problemas tenían y como los he solucionado.

## Limpieza de datos planta 0

### Documento

Documento es el ID Venta que viene generado por la máquina. Para una mayor comodidad de trabajo decidí cambiarlo y poner que empezase por 1, pero podía haberlos mantenido.

### Agente

Agente hace referencia al nombre y apellidos que de la persona que realizo la venta, decidí asignar a cada trabajador un ID para que en la tabla de hechos apareciese ese ID.

Cree una **nueva tabla empleado**, la que luego crearía una nueva tabla de dimensión con ese ID\_empleado y los datos de cada empleado.

### Ciente

En todos los casos es “Cliente al contado”. Puesto que no aporta ningún valor decidí eliminar esta columna.

### Fecha

Este campo indica el día en el que se hizo la venta, pero no dice nada respecto a la hora. Por ese motivo fue necesario crear la columna hora de la forma que explicaré más adelante.

### Base

Importe venta sin IVA, decidí quedarme con este importe frente al importe sin el IVA incluido.

### Cuota

IVA de las ventas, al disponer de “base” y de “importe”, no aporta gran información

## **Importe**

Es el importe final de la venta con el IVA incluido.

## **Limpeza de datos planta 1**

### **Albaranes**

#### **Nº:**

ID identificativo del albarán que servirá como foreign key.

#### **Número pedido líneas:**

Número que no aporta ninguna información al no ser un valor en sí mismo, ni estar relacionado con otra columna. Por este motivo decidí eliminarla.

#### **Nº Completo presupuesto:**

Número que identifica a que presupuesto corresponde el albarán. Tiene una gran importancia porque nos servirá para relacionar el albarán, con las facturas y los presupuestos.

#### **Fecha:**

Fecha en la que se genera el albarán. Para nosotros será la más importante puesto que nos indica cuando se formalizo el pedido y podremos sacar tiempos desde que se hace un presupuesto hasta que se aprueba.

## **Nombre cliente**

Nombre de la empresa cliente. Como veremos más adelante supone un gran problema y su limpieza fue parte fundamental del proyecto. Lo explicaré mejor más adelante.

#### **Producto pedido:**

Producto que se ha pedido en ese albarán. Como expliqué anteriormente, cada albarán lleva referenciado un solo producto

#### **Cantidad líneas:**

Cantidad de producto que se ha pedido.



### **Estado:**

Si el producto fue recogido o no, salvo un par de excepciones siempre es “Si”, y ese par de excepciones eran causadas por errores humanos. Puesto que no aporta ningún valor decidí eliminarlo

### **Impreso:**

Si el producto fue impreso o no, salvo un par de excepciones siempre es “Si”, y ese par de excepciones eran causadas por errores humanos. Puesto que no aporta ningún valor decidí eliminarlo

### **Facturado**

Siempre es si, por ese motivo no aporta ninguna información y lo eliminé.

### **Cancelado Pedido Líneas:**

Si el producto fue cancelado o no, salvo un par de excepciones siempre es “Si”, y ese par de excepciones eran causadas por errores humanos. Puesto que no aporta ningún valor decidí eliminarlo.

### **Clientes**

#### **Nombre:**

Nombre de la empresa cliente, como he dicho, más adelante explico detalladamente como trabaje con este campo.

#### **Código postal**

CP de la empresa.

#### **Población**

Población de la empresa. En relación a la población decidí añadir dos campos más, si pertenecía a la provincia de Valladolid y si pertenecía a Valladolid capital.

#### **Teléfono**

Teléfono de contacto de la empresa.

### Facturas

#### **Número:**

Una numeración que al ser siempre 0 no aporta nada y borré.

#### **Fecha**

Fecha en la que se generó la factura, normalmente a final de mes

#### **Número Pedido Segmentos:**

Número que no aporta ninguna información al no ser un valor en sí mismo, ni estar relacionado con otra columna. Por este motivo decidí eliminarla.

#### **Nº Completo presupuesto pedido segmentos;**

Número que identifica a que presupuesto corresponde la factura. Tiene una gran importancia porque nos servirá para relacionar la actura, con el albarán y los presupuestos

#### **Nombre cliente:**

Nombre del cliente a quien se le factura, como digo hubo que tratar de forma especial ese campo y lo explico más adelante.

#### **Referencia**

Descripción que da hace del producto el responsable de la planta. Puesto que cada una es de una forma, es imposible compara y analizar con otros datos y decidí eliminar el campo.

#### **Segmentos**

Hace referencia a uno de los productos que incluye el presupuesto, para ser exacto, al primero que aparece en este. Al no aparecer todos, no es un dato real y decido eliminar el campo para utilizar el campo "producto" de albaranes.

#### **Subtotal**

Valor de la factura sin IVA. Decidi trabajar con este campo.

#### **Total IVA**

IVA de la factura.



## **IVA incluido**

Valor factura con IVA.

## **Presupuestos**

### **Nº Completo:**

Número de presupuesto, seguramente el campo más importante de todas las tablas y que explicare su tratamiento en el siguiente punto.

### **Fecha**

La fecha en la que se genera el presupuesto, es decir, la fecha en la que pasamos la oferta al cliente. Podemos analizar el tiempo con las otras fechas, pero debido a como se gestionan los presupuesto y albaranes, puede ser poco real.

### **Producto**

Producto correspondiente al presupuesto, el problema es que como un presupuesto es incluye varios productos, solo sale referenciado uno. Por ese motivo es mejor trabajar con producto de albarán.

### **Referencia**

Descripción sobre los productos que están presupuestados. Todos son distintos entre sí y por eso es imposible analizarlos individualmente y decidí eliminarlos.

### **Páginas**

Cantidad de productos totales. Al no estar indicado la cantidad de cada producto.

### **Nombre cliente**

Nombre que da el responsable de planta al cliente en el momento de hacer el presupuesto

### **Teléfono – Contacto**

Contacto del cliente que al ya tenerlo en la tabla clientes lo borre para no tener datos duplicados

### **Cantidad**

Similar a páginas.

### **Venta Total**

Valor total del presupuesto

### **Coeficiente:**

Coeficiente que pone el responsable a cada producto de forma subjetiva. Como me explico el cliente, no es ningún dato significativo y por eso lo elimine

### **Último pedido:**

Supuesto último pedido de ese mismo cliente, tendría valor en el caso de estar bien, pero el programa genera mal este campo y por eso lo eliminé.

## **Operaciones especiales con algunos campos**

### **Nº presupuesto**

El número ha sido el dato más importante porque es la única forma de relacionar facturas, albaranes y presupuestos. Cada albarán está relacionado con un presupuesto y cada presupuesto está relacionado con una factura.

No ha sido necesario modificar nada ni hacer limpieza en este campo, el problema es que un mismo presupuesto puede volver a utilizarse en el caso que un mismo cliente haga el mismo pedido. Esto supone que hay albaranes referenciados con presupuestos de hace varios años, lo que nos falsea el cálculo del tiempo entre que se genera.

### **Clientes planta 1**

El valor cliente aparece en la tabla facturas, albaranes y presupuestos. Como comenté anteriormente, el primer paso cuando un cliente quiere comprar algo es generarle un presupuesto, al generar ese presupuesto el responsable de la tienda pone un nombre de cliente, el nombre que él quiere, El problema surge porque al generar el albarán y la factura, el nombre que aparece ya es el nombre real de la empresa.

Otro problema que surge es que algunas empresas estaban guardadas dos veces con distinto nombre aunque muy parecido, con una



letra de menos, con y sin “S.L.” ... Para solucionar todo esto y que los datos de las ventas estuviesen realmente bien relacionados, en primer lugar cree una tabla con todos los clientes en la que fue necesario asignar un ID a cada empresa cliente evitando que haya empresas repetidas. Después de tener el ID asignado a cada cliente, se asignó ese ID a cada una de las ventas en función del cliente en las tres tablas. Esto lo hice con el comando “buscarv”, pero como no siempre tenía el mismo nombre, luego fue necesario ir uno por uno comprobando que estaba bien asignado o asignando si no lo estuviese.

Por último, el campo cliente no siempre estaba rellenado, apoyándome en la columna que hacía referencia al número de presupuesto, pude sacar a que cliente eran esas ventas.

## Producto

El producto sale reflejado en varias tablas de los datos que me facilito la empresa. Había dos columnas que hacían referencia al producto. Una de ellas hacía referencia al tipo de producto, y otra era una referencia redactada por el responsable sobre lo que era. Este último campo no nos sirve puesto que en cada producto añade o cambia un detalle que lo hace único. El otro campo en que se habla del tipo del producto que se hace, es el que nos servirá para saber qué productos se consumen más.

Para organizar mejor este campo, cree una tabla “productos” con los distintos productos que ofrece la empresa y asigne una referencia a cada producto.

### 4.2.3. Obtención de más datos con valor

Con estos datos que puede facilitarnos la empresa, no se puede sacar mucha información que aporten un valor añadido a la empresa, por ese motivo he intentado sacar más información de otras fuentes de información.

#### Análisis de variables.

Para intentar conseguir información más valiosa, es necesario analizar que variables pueden afectar al negocio.

Debido a la naturaleza del negocio de cada planta, no afectarán de la misma forma unas variables u otras a las distintas plantas.

Posibles variables que podrían afectar al negocio:

- Fecha.
  - Festividades / fechas destacadas del calendario.
  - Mes.
  - Principio o final de semana.
  - ...
- Hora.
- Temperatura.
- Lluvia.
- Lugar de residencia del cliente/empresa.
- Empleado que atiende.
- Producto que vende.

Después de plantear que variables pueden afectar a nuestro negocio, debemos decidir cuales si afectan e intentar hacernos con esa información.

Por ejemplo, para comprar algo en la planta 0 es necesario acudir a la tienda, esto no ocurre en la planta 1. Por lo tanto, las condiciones climáticas pueden afectar notablemente a las ventas de la planta 0 y no tanto a la planta 1.

Si está lloviendo o hace mucho calor, es probable que nadie vaya a comprar un bolígrafo a la tienda, pero si es una compra realmente importante de más valor o que necesite y la tienda cierre al día siguiente, es probable que sí que salga a compra por lo que el clima o la fecha si afectan.

**Para la planta 0 decidí que las variables** que podía conseguir fiablemente y que si afectan son las siguientes:

- Hora.
- Época del año.
- Temperatura.
- Lluvias.

Otras variables que pueden influir de las que no dispongo y puedo conseguir son la edad de los clientes, el sexo, donde viven, el producto que compran... Para tener estos datos, daré unas recomendaciones a la empresa en otro apartado.

**Para la panta 1 los datos eran mucho más completos** de los que de por sí ya podíamos sacar bastante información. Una separación que hice de los clientes fue si eran una entidad pública o privada. Otra información que podría ser de valor es el sector al que se dedica la empresa o el número de trabajadores.

Para analizar estas dos variables, he buscado la información de las 10 empresas a las que más y a las 10 que menos factura la imprenta para ver si tiene relación el número de trabajadores o el sector con que compren más o menos.

En la planta 1 la hora o el clima no tienen tanta importancia dado que los presupuestos se pueden hacer telefónicamente u online sin necesidad de acudir a la tienda.

### Obtención de datos de las nuevas variables.

Después de analizar que variables pueden afectar al negocio, es necesario buscar los datos que nos permitan relacionar esas nuevas variables con los datos que ya teníamos antes.

### Hora

Una variable muy importante es conocer a qué **hora** se hace la venta en la planta 0. Puesto que no dispongo de ese dato, opte por apoyarme en las herramientas de Google que registran cuando va la gente a la tienda. Gracias a esas estadísticas, pude sacar porcentualmente cuánta gente va a la tienda en cada franja horaria en función del día, y así asignar porcentualmente una hora venta a cada transacción según el día.

Por ejemplo, los lunes según las estadísticas de Google de 9 a 10 han 5 personas, de 10 a 11 hay 9 y así con el resto de horas, saque un porcentaje en gracias esos datos y fui asignando una hora de venta una en una para cada día del año. Ha sido un trabajo muy tedioso puesto que cada día tiene un número distinto de ventas y es importante hacerlo siguiendo el orden de la venta de cada día para no solo saber el número de ventas en función de la hora, si no también saber si el valor de las ventas varía dependiendo de la hora.

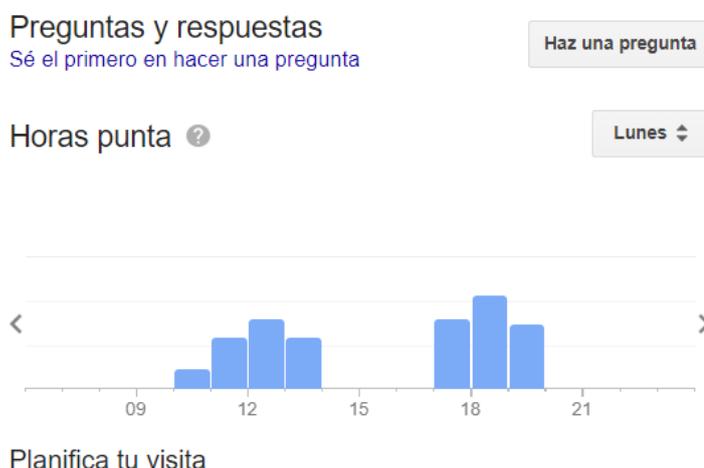


Ilustración 4.8. Gráfica de Google de la asistencia de clientes a la tienda.

## Época año

Otra variable importante es analizar la **época del año**, para hacer esto fue más sencillo puesto que dividí el año en varias épocas exámenes, en función de las vacaciones y de si es época “normal” de trabajo.

## Empleado

Para analizar mejor el **empleado**, he creado una tabla añadiendo en la que salen los distintos empleados con un código de identificación que registrará las ventas, nombre, apellidos, sexo, edad, teléfono y email. De esta forma podemos analizar las ventas también en función del sexo y la edad del vendedor.

## Clima

Los datos del **clima** se pueden obtener en AEMET descargando un Excel para cada día del año que contiene los datos de todas las estaciones meteorológicas. Después de descarga Excel a Excel y quedarme con la información de la estación de Valladolid, descubrí la web [www.datosclima.es](http://www.datosclima.es) que por 1,5€ pude descargarme todo el año en un solo Excel y solo tuve que quedarme con Valladolid.

## Búsqueda de datos no estructurados

También busque datos no estructurados en la web, pero debido a que la empresa no hace uso de las redes sociales no encontré nada de valor que pudiese aportar información sobre la empresa.

## 4.2.4 Diseño de estructuras de las bases de datos

Una vez realizada la limpieza de datos, es el momento de diseñar nuestra base de datos. Para ello debemos tener muy claro que datos contiene cada tabla y cómo podemos y queremos relacionarlas. También hemos tenido que plantear junto con el cliente que información queremos sacar de nuestros datos.

## Planta 0

En este caso el diseño de la estructura es más sencilla. Después de la limpieza de datos y la obtención de otros nuevos, tenemos las siguientes tablas sus respectivos campos.

- Hora:
  - Id\_hora.
  - Franja.
  - Turno.
- Fecha:
  - Id\_fecha.
  - Día\_semana.
  - Día\_mes.
  - Mes.
  - Año.
  - Epoca\_año.
- Clima:
  - Fecha.
  - T<sup>a</sup>\_Max.
  - Lluvia\_mañana.
  - Luvia\_tarde.
  - Lluvia\_día.
- Empleado:
  - Id\_empleado.
  - Nombre.
  - Apellido.
  - Edad.
  - Sexo.
  - Teléfono.
  - Email.
- Ventas:
  - Id\_Venta.
  - Fecha.
  - Hora.
  - Empleado.
  - Importe.

Con estas tablas, la estructura de la base de datos es muy clara. La tabla de hechos será la tabla ventas y el resto serán tablas de dimensión. Las relaciones entre las tablas son evidentes:

- Hora (Ventas) – Id\_hora (Hora)
- Fecha (ventas – Id\_fecha (fecha)
- Empleado (ventas) – Id\_empleado(empleado)
- Id\_fecha (fecha) – fecha (clima)

La tabla nos quedaría así:

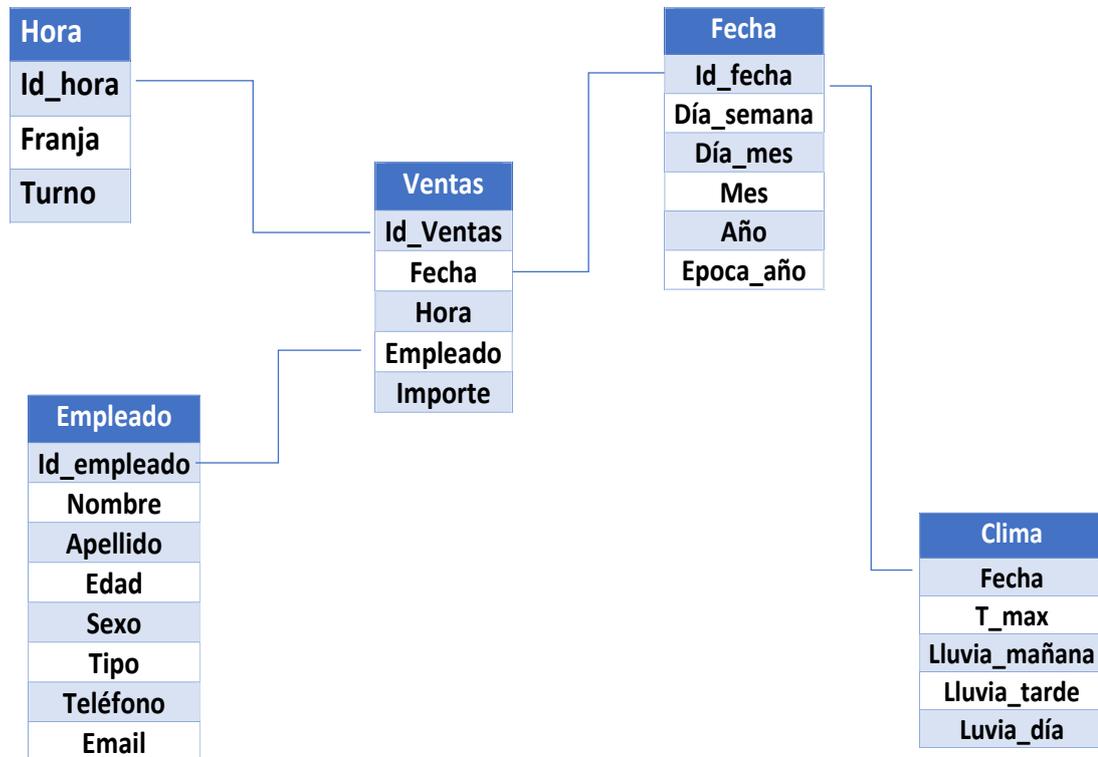


Ilustración 4.9. Tabla base datos planta 0.

Con esta disposición de las tablas, podemos relacionar el importe de la venta con todos los campos referentes al empleado como pueden ser la edad o el sexo, con la franja horaria en la que se hizo la venta y también con la tabla fecha y clima, pudiendo saber establecer una conexión por ejemplo entre las ventas y la temperatura o el día de la semana.

### Planta 1

En esta planta fue mucho más complicado establecer las relaciones entre las tablas debido a los datos de los que disponíamos. Después de la limpieza de datos me quede las siguientes tablas con sus respectivos campos:

- Presupuestos:
  - N° completo.
  - Fecha.
  - Venta total.



- Facturas:
  - Id\_Factura.
  - N\_Factura.
  - Fecha.
  - Subtotal.
  - Total\_IVA.
  - IVA\_incl.
- Clientes:
  - Id\_Cliente.
  - Nombre.
  - CP.
  - Población
  - Ciudad\_Valladolid.
  - Provincia\_Valladolid.
  - Teléfono.
- Albaranes
  - Id\_Albarán.
  - Id\_Fecha.
  - Id\_Presupuesto.
  - Id\_Cliente.
  - Id\_Producto.
  - Cantidad\_Albarán.
  - Id\_Factura.
- Calendario
  - Id\_Fecha.
  - Día\_semana.
  - Día\_mes
  - Mes.
  - Año.

Mi primera intención fue que la tabla de hechos fuese la tabla presupuestos debido a que el número de presupuesto es lo único que relaciona la tabla albaranes, con presupuestos y con facturas, pero debido a como se generan los datos al generar cuando encargar un pedido a la empresa, creí que la mejor tabla para ser la tabla de hechos, sería la tabla albaranes.

Esta decisión atendía a que albaranes es la unidad mínima que se genera cuando el cliente acepta el presupuesto. Se genera un albarán para cada producto, con sus cantidades y con el precio de dicho producto.

Una vez realizada la implantación con esa estructura, me percaté que no funcionaba correctamente y volví a mi idea inicial de realizarlo en torno a

presupuesto. Aunque no podía relacionar tantos campos como quería, lo que sí hacía, lo hacía correctamente. Finalmente, las relaciones que quedaron fueron las siguientes:

- N° Completo (presupuesto) – N° presupuesto (Factura).
- N° Completo (presupuesto) – N° presupuesto (Albarán).
- Fecha (Presupuesto) – Id\_Fecha (Fecha).
- Fecha (Factura) – Id\_Fecha (Fecha).
- Fecha (Albarán) – Id\_Fecha (Fecha).
- Id\_Cliente (Presupuesto) – Id\_Cliente (Cliente).
- Id\_Cliente (Albaran) – Id\_Cliente (Cliente).
- Id\_Producto (Albaran) – Id\_Producto (Producto).

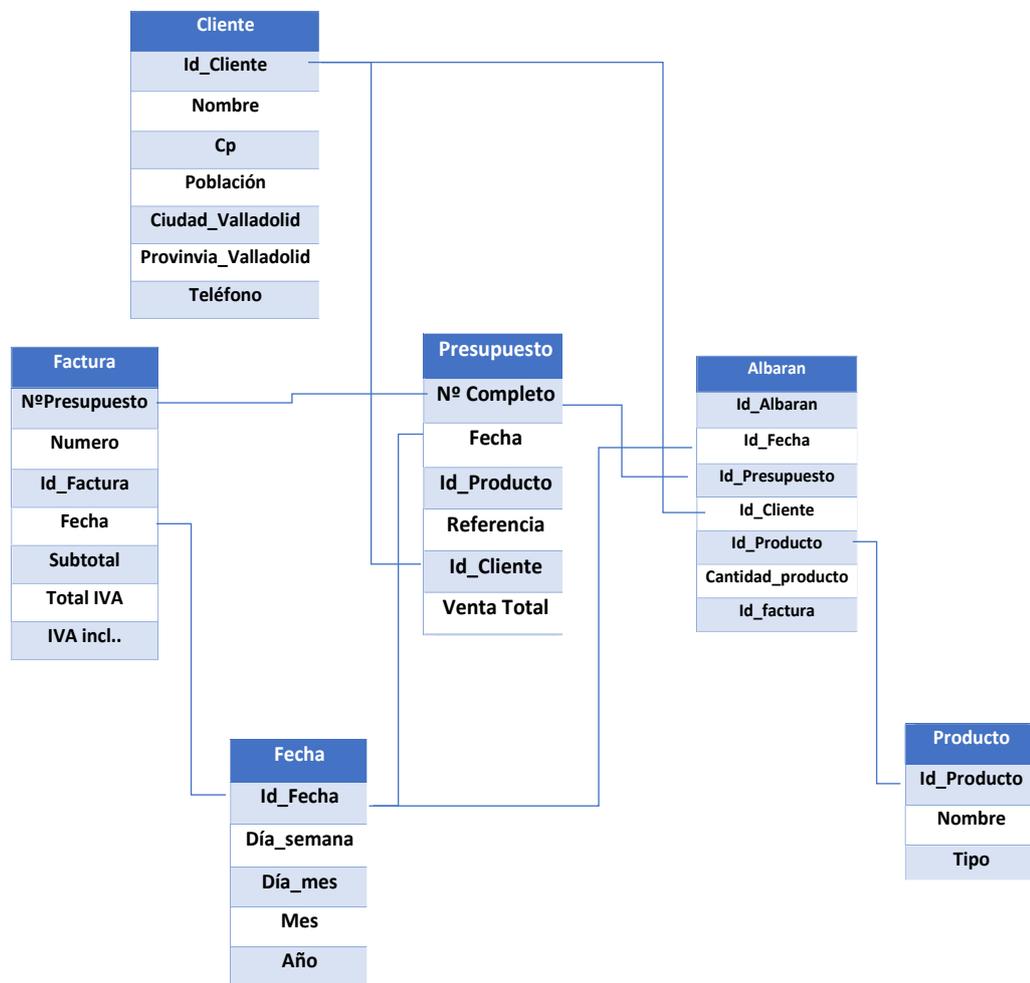


Ilustración 4.10. Tabla base datos planta 1.



## 5. IMPLANTACIÓN DE BI EN LA EMPRESA

Después de explicar el funcionamiento de la empresa, los datos que aporta, como los hemos tratado y los nuevos datos que he obtenido de otras fuentes, es el momento de explicar cómo podemos implantar las herramientas de BI en la empresa.

**Voy a explicar cómo realizar la implantación del BI de 3 formas distintas**, utilizando distintitos programas, pero en todos los casos será común el uso del Power BI y del Excel. En todos los casos explicaremos sus ventajas y desventajas.

La primera de las formas será la más sencilla, después del correcto tratamiento de los datos, cargar correctamente las tablas al Power BI, crear las relaciones entre tablas y ya estaría disponible para trabajar con ello.

La segunda forma es algo más complicado, primero crearemos una base de datos en Access, cargaremos los datos, crearemos las relaciones entre las tablas y después desde Power BI cargaremos dicha tabla y podremos trabajar con ella.

La última forma es la más complicada de todas, pero nos permitirá tener nuestros datos seguros sin riesgo de perderlos al estar cargados en la web. Con el gesto de bases de datos de MySQL, XAMPP, crearé una base de datos la cual editare con MySQL Workbench. Después cargaré esta base de datos al Power BI y ya podré trabajar con ella.

Debido a la naturaleza de la empresa, será necesario realizar dos proyectos de implantación por separado, uno para cada planta.

La implantación en la planta 0 la haré mediante la forma 2 y 3, la de la planta 1 la haré mediante la forma 1.

## 5.1. Programas utilizados



Ilustración 5.1. Logo MySQL

**MySQL Workbench** es una herramienta visual de diseño de bases de datos que integra desarrollo de software, Administración de bases de datos, diseño de bases de datos, gestión y mantenimiento para el sistema de base de datos MySQL. [54]



Ilustración 5.2. Logo XAMPP

**XAMPP** es un paquete de software libre, que consiste principalmente en el sistema de gestión de bases de datos MySQL, el servidor web Apache y los intérpretes para lenguajes de script PHP y Perl. [55]



Ilustración 5.3. Logo Power Bi.

**Power Bi** programa del que ya hemos hablado anteriormente. En nuestro proyecto lo hemos utilizado para la visualización de información mediante la elaboración de dashboard.



Ilustración 5.4. Logo Excel

**Microsoft Excel** es una aplicación de hojas de cálculo que forma parte de la suite de oficina Microsoft Office. Es una aplicación utilizada en tareas financieras y contables, con fórmulas, gráficos y un lenguaje de programación. [56]



Ilustración 5.5. Logo Access

Microsoft Access es un sistema de gestión de bases de datos incluido en el paquete ofimático denominado Microsoft Office. Access es un gestor de datos que recopila información relativa a un asunto o propósito particular, como el seguimiento de pedidos de clientes o el mantenimiento de una colección de música, etc. [57]

## 5.2. Consideraciones previas a tener en cuenta

Antes de ponernos con la implantación, es importante saber algunas cosas que pueden crearnos problemas:

- Para la carga de datos directamente en Power Bi, será mejor generar una tabla con cada uno de nuestros archivos Excel con datos y cargar dicha tabla. No se exactamente porque, pero esto facilita el trabajo al programa a la hora de cargar los datos. De no hacerlo, si hay un gran número de datos, da problemas al cargarlo o los carga pero no les muestra por pantalla.
- Debemos asignar una clave foránea o foreign key a todas las tablas
- La clave foránea de una tabla no puede tener valor null.
- La clave foránea no puede repetirse en una tabla, debe ser única.
- Un campo de una tabla que está relacionado con un campo de otra tabla deberá ser del mismo tipo de campo.
- Una clave foránea no podrá ser de tipo texto, lo pondremos como cadena de caracteres (Varchar).

## 5.3. Implantación en la planta 0

La implantación de herramientas BI en la planta 0 la haremos de dos formas. La primera creando la base de datos con Acces, y la segunda creando y gestionando la base de datos en MySQL.

### 5.3.1. Mediante la utilización de Access y Power BI.

En primer lugar, debemos familiarizarnos con el programa. Si vista principal es muy similar a la de Word o Excel con una barra superior de herramientas. A la izquierda nos aparecerán las tablas que tenemos cargadas, por defecto nos aparecerá una para cargar datos pero que desaparecerá en cuanto carguemos nuestros datos y la cerremos.

Debemos crear la base de datos y cargar nuestros datos en ella, esto lo podemos ir realizando más o menos al mismo tiempo, puesto que cuando carguemos nuestros datos desde un archivo Excel, se irán generando las tablas para nuestra base de datos.

El proceso se divide en tres partes, carga de datos, creación de relaciones entre las tablas y cargar la base de datos al power BI. Todo esto está explicado y detallado con imágenes en el anexo, aquí explicaré lo que he tenido que hacer yo para nuestro caso en concreto.

### **Carga de datos**

Paso 1 – Selección de archivos:

En este paso debemos seleccionar que queremos cargar desde un archivo Excel nuestros datos.

Datos externos > Excel > Importar el origen de datos en una nueva tabla de la base de datos actual.

Examinamos y seleccionamos nuestros datos.

Paso 2 – Importación de hojas de cálculos:

En este paso es importante realizar bien unos pequeños detalles para que después no nos de problemas. Es importante indicar que la primera fila contiene encabezados de columnas, si no los confundirá como si de otro dato se tratase. En la siguiente deberemos indicar el tipo de dato de cada columna y si está indexado o no. Un dato que está relacionado con otro está indexado, pero podemos seleccionar si está duplicado o no. Cuando sea la clave foránea de una tabla no puede ser duplicado, pero si no es la clave foránea si. Por ejemplo, Id\_fecha en la tabla venta está duplicado porque hay varias ventas un mismo día, pero en la tabla fecha Id\_Fecha es la clave foránea, solo hay una línea para ese día y por eso no puede estar duplicado. En nuestro caso los datos serán de tipo:

- Id\_Venta: Entero.
- Fecha: Fecha con hora.
- Hora: Fecha con hora.
- Trabajador: Entero.
- Importe: doble.
- Temperatura: texto corto.
- Lluvias: texto corto.
- Franja horaria: Entero.
- Turno: texto corto.

- Edad: texto corto.
- Sexo: texto corto.
- Nombre: texto corto.
- Apellido: texto corto.
- Email: texto corto.
- Teléfono: Doble

### Paso 3 – Elegir clave primaria

Debemos elegir la clave primaria, si hemos hecho lo anterior bien y no hay campos en blanco en nuestra clave primaria no debería haber problema.

Deberemos realizar estos pasos con todas nuestras tablas, una vez hecho, nos aparecerán dichas tablas en la parte izquierda de la pantalla.

### Creación de relaciones entre las tablas

#### Paso 1 – Selección de las tablas que llevarán relación

Herramienta de base de datos > Relaciones > Se abrirá pa pestaña diseño y agregamos las tablas.

#### Paso 2 – Relacionar las tablas

Pestaña diseño > Modificar relaciones > Crear una nueva > Seleccionamos los campos que queremos relacionar.

En caso de haber hecho bien estos pasos, deberíamos tener la tabla creada correctamente y en nuestro caso nos queda una estructura así:

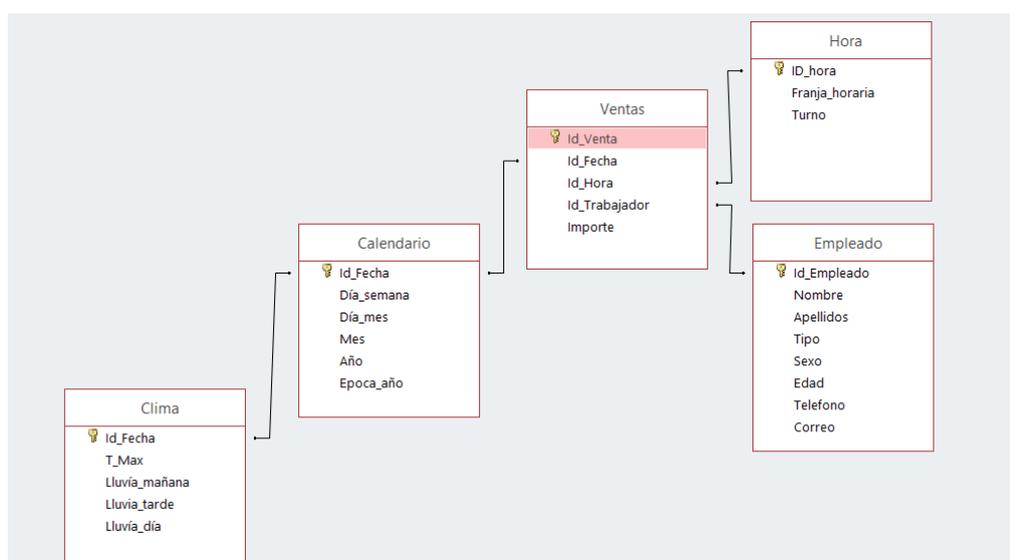


Ilustración 5.6. Estructura base de datos en Acces.

### **Cargar la base de datos al Power Bi**

En este momento debemos abrir el Power Bi, al igual que otras herramientas de Microsoft tiene una interfaz parecida a Word o Excel.

Para cargar la base de datos, deberemos seguir estos pasos seleccionar “Get Data” > Access Data Base > Seleccionamos la tabla y la cargamos.

En ese momento ya tendremos todo disponible para trabajar con Power Bi para sacar nuestros gráficos y nuestro dashboard. Como realizar esto lo explico en otro punto más adelante.

### **5.3.2. Mediante MySQL Workbench y Power Bi**

Este método es el más complejo de los tres, pero será la forma más segura de conservar nuestra base de datos al tenerla en línea. Solo se podrá acceder con la IP del que haya creado la tabla u otras IP a las que este quiera autorizar.

Deberemos descargar e instalar XAMPP y el MySQL Workbench. En el anexo dejaré los enlaces donde poder realizar la descarga e indicaré como realizar la instalación. Además todo el proceso explicado aquí, lo estará más detalladamente en el anexo.

Deberemos crear las tablas, crear las relaciones entre ellas y cargar los datos. Para hacer todo esto iremos alternando entre un programa y otro realizando distintos pasos con cada programa.

#### **Creación de las tablas**

##### **Paso 1 – Ejecutar el XAMPP**

Deberemos ejecutar el XAMPP e inicializar Apache y MySQL. Esto lo realizaremos pinchando en “start” en los dos casos. A continuación abriremos MySQL pinchando en “admin” y se nos abrirá un gestor de MySQL en el navegador web.

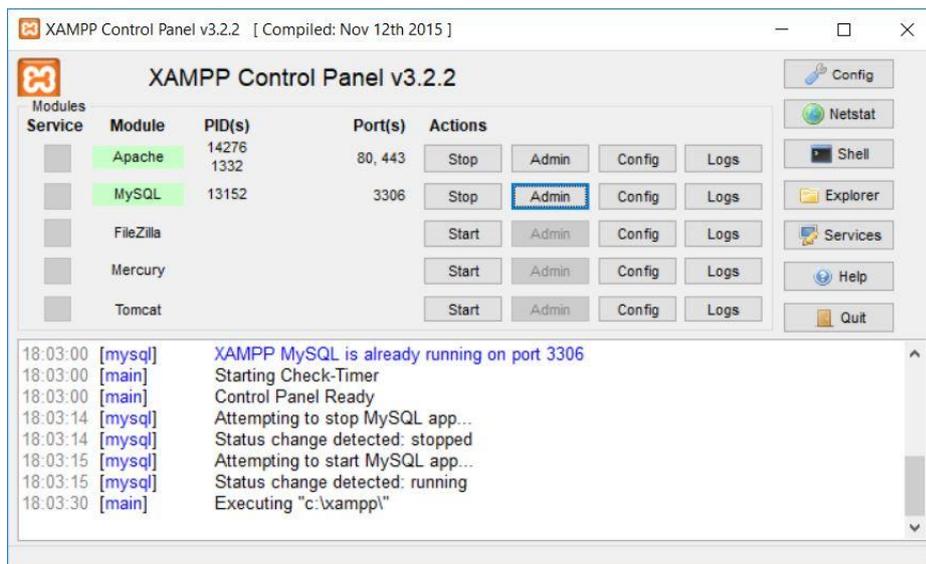


Ilustración 5.7. Panel de control de XAMPP.

### Paso 2 – Creación de la base de datos

En el navegador se nos habrá abierto un gestor de MySQL en el cual podremos navegar por las bases de datos que tenemos creadas y sus tablas gracias a un menú que tenemos a la izquierda.

En ese mismo margen izquierdo tendremos que pinchar en “nueva” y nos pedirá un nombre para la base de datos y la crearemos.

### Paso 3 – Creación de las tablas

En el mismo margen izquierdo por el que ponemos navegar por las distintas bases de datos que tengamos, podemos también navegar por las tablas de estas y crear tablas. Seleccionaremos nueva tabla, nos pedirá nombre y deberemos definir los campos de dicha tabla.

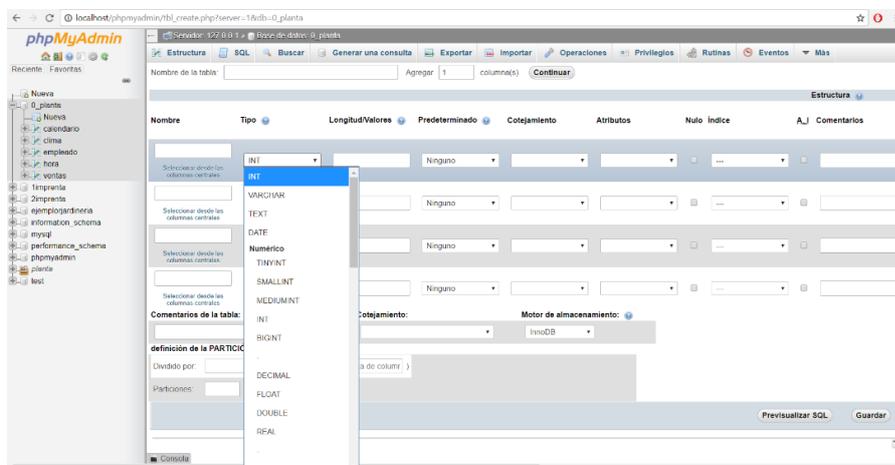


Ilustración 5.8. Vista general gestor MySQL.

Como podemos ver en la imagen, estos campos pueden ser de varios tipos, los que más utilizaremos y para que los utilizaremos son los siguientes:

- Varchar: Es para una cadena de caracteres. Este tipo campo nos servirá para campos que incluyan texto y números, además de ser el que mejor funciona si hay guiones o barras.
- Text: Este campo será para tipo texto.
- Date: Para fechas (aunque usaremos varchar para estas)
- Int: Para números enteros.
- Double: Para número decimales.

En nuestro caso definiremos los campos de la siguiente forma:

- Tipo varchar: fecha, hora, venta.
- Tipo text: día semana, mes, época año, temperatura máxima, lluvia mañana, lluvia tarde, lluvia día, nombre, apellidos, tipo, sexo, email, edad, turno.
- Tipo int: día mes, año, empleado, teléfono, franja horaria.
- Tipo doublé: importe.

De tal forma que las tablas quedaban así:

Nombre	Tipo
Id_fecha	varchar(10)
Dia_semana	text
Día_mes	int(11)
Mes	text
Año	int(11)
Epo_año	text

Nombre	Tipo
Id_fecha	varchar(10)
T_Max	text
Lluvia_mañana	text
Lluvia_tarde	text
Lluvia_día	text

Nombre	Tipo
Id_empleado	int(11)
Nombre	text
Apellido	text
Tipo	text
Sexo	text
Telefono	int(11)
Email	text
Edad	text

Nombre	Tipo
Id_venta	varchar(10)
Id_fecha	varchar(10)
Id_Hora	varchar(10)
Id_Trabajador	int(11)
Importe	double

Nombre	Tipo
Id_hora	varchar(10)
Franja	int(11)
Turno	text

Ilustración 5.9. Tablas creadas con el gestor MySQL.

## Creación de relaciones entre tablas

Paso 1 – Creación de conexión a la base de datos.

Es en este paso cuando debemos abrir el MySQL Workbench

Deberemos crear una nueva conexión con nuestra base de datos. Esto se hace pulsando el botón “+”, damos un nombre a la conexión y creamos nuestra conexión. Nos saldrán unos avisos que deberemos aceptar. Si todo ha salido correctamente nos saldrá la la conexión a la base de datos.

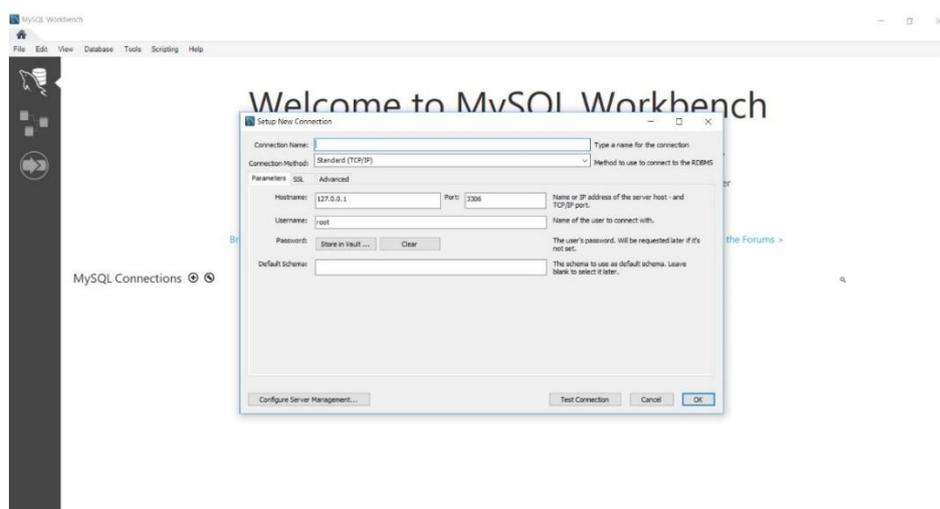


Ilustración 5.10. Creación conexión a base de datos con MySQL Workbench.

Paso 2 – Creación de relaciones entre las tablas.

Ahora nos saldrá la conexión a la base de datos, la cual seleccionaremos y ya podremos trabajar sobre ella.

Una vez abierta la base de datos, deberemos seleccionar la tabla en la que queremos crear las restricciones. Pinchando sobre un símbolo que es una llave inglesa se nos abrirá un panel para trabajar sobre esa tabla.

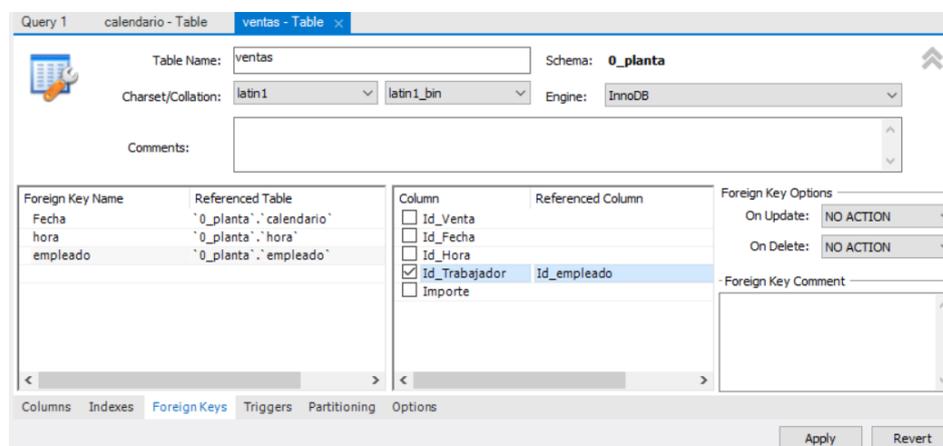


Ilustración 5.11. Panel para la creación de conexiones entre tablas con MySQL Workbench.

Seleccionamos la pestaña foreign keys y marcamos las relaciones que queremos hacer seleccionando la tabla con que queremos relacionar y que campos queremos relacionar. Cuando demos a “Apply” nos saldrá una ventana como esta con el siguiente código. Aceptaremos y la relación estará creada.

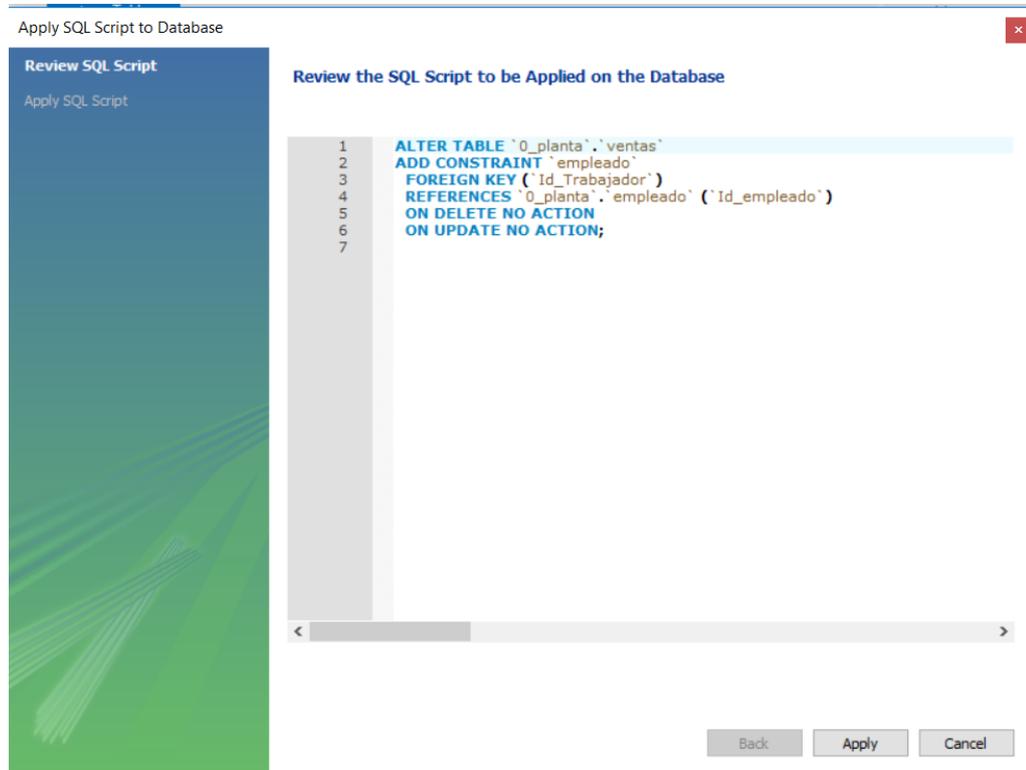


Ilustración 5.12. Código que se genera para la creación de conexiones entre tablas.

Realizamos esto para todas las tablas.

### Carga de datos

Paso 1 – Debemos guardar los archivos Excel como CSV delimitados por comillas.

Paso 2 – Haciendo clic con el botón derecho sobre la tabla a la que queremos importar datos seleccionamos “Table Data Import Wizard”.

Paso 3- Seleccionamos la el archivo CSV, dejamos los valores predefinidos y comprobamos que relaciona bien nuestras columnas del CSV con los campos de la tabla. Damos a siguiente hasta el final y ya estarían cargados.

En caso de que nos diga que no podemos leer, debemos el Encoding de utf-8 a latín. El error puede deberse al uso de la “ñ” en algunos datos.

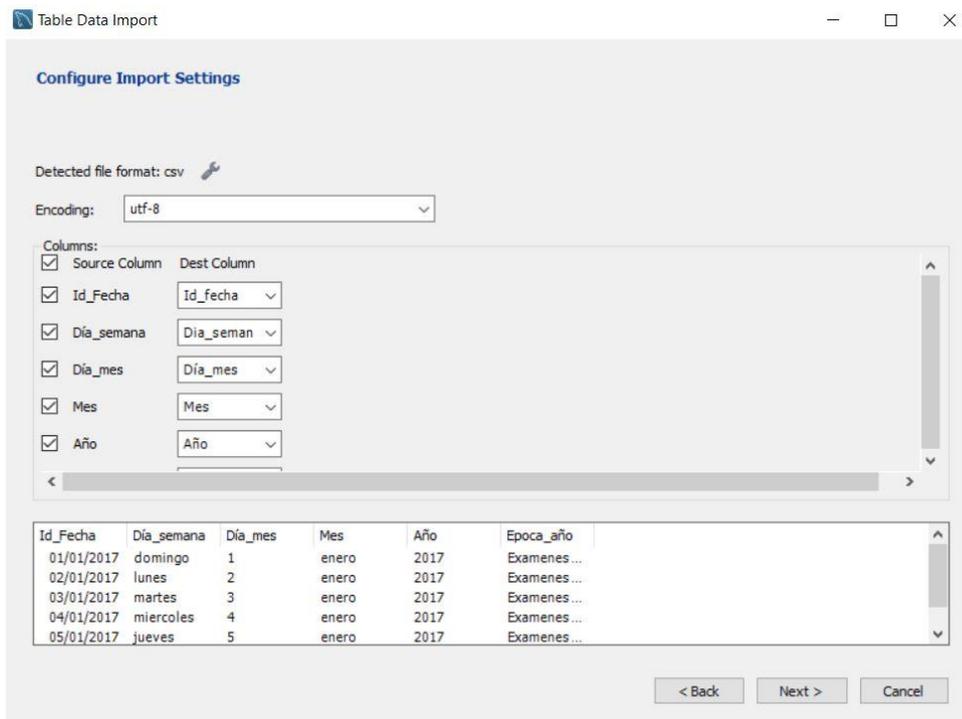


Ilustración 5.13. Captura de pantalla de la carga de datos con MySQL Workbench.

### Cargar base de datos al Power Bi

Para este paso tanto Apache como MySQL deberán seguir inicializados en el panel de control de XAMPP.

Paso 1 - Este paso es relativamente similar a como he indicado antes, lo único que al seleccionar el tipo de base de datos, debemos seleccionar "MySQL Data Base".

Paso 2 - Si es la primera vez que intentamos conectar a una base de datos de MySQL nos aparecerá un error. Debemos pinchar en "learn more", se nos abrirá una web en el navegador de donde podremos descargar un archivo MySQL que deberemos instalar y nos permitirá conectar la base de datos.

Repetimos paso 1 habiendo reiniciado previamente el Power Bi.

Paso 3 - Ahora nos aparecerá un cuadro de dialogo donde deberemos poner el número de servidor que aparece en el panel de control de XAMPP y el nombre de la base de datos.

Paso 4 - Conectamos y seleccionamos las tablas que queremos cargar.

Llegados a este punto ya estaremos preparados para generar nuestros informes visuales con Power BI.

## 5.4. Implantación en planta 1

Para la planta 1 tenía unas ideas teóricas que finalmente no he podido aplicar a la práctica. Debido a como se generan los datos, ha sido imposible establecer ciertas relaciones que si pensé que podría haber establecido. Tras intentarlo de muchas formas y no conseguirlo, decidí que no era un fracaso dentro del proyecto, si no que podría utilizarlo como comparativa con la planta 0 y como ahí sí he podido.

Como he explicado anteriormente, mi idea era utilizar como tabla de hechos la tabla albarán, pero esto suponía ciertos problemas a la hora de referenciar y además la tabla albarán no tiene un campo que sea el valor de albarán para relacionarlo con el dinero. Finalmente volví a la idea inicial de utilizar como tabla de hechos la tabla presupuesto.

La implantación de la tabla 1 lo he hecho de la forma que nos falta.

### 5.4.1. Implantación cargando los datos directamente en Bi.

A priori, esta es la forma más simple de todas, pero pueden surgir otros problemas que de la otra forma no surgen. En este caso tendremos dos pasos, la carga de los datos y crear las relaciones entre tablas.

#### Creación de tablas y carga de datos en Power Bi

Para cargar los datos en Excel en Power Bi es recomendable que estos estén estructurados como una tabla con encabezados dentro del Excel. Esto facilitará que el Power Bi coja y cree correctamente las tablas.

Seleccionamos en Data > File > Excel. Seleccionamos el archivo a subir y dentro de este la tabla.

Realizamos ese paso con todo y ya tendríamos las tablas creadas y los datos cargados.

#### Creación de relaciones entre las tablas.

Para este paso debemos navegar por la barra de la derecha y seleccionar ver la relación entre tablas. Power Bi nos habrá creado algunas automáticamente, puede que este bien o no, revisamos y añadimos otras nuevas.

Para ello hacemos clic en “manage relationships” y añadimos una nueva. Se hace seleccionando las tablas y los campos que queremos.

## 5.5. Utilización del Power Bi

Una vez que tenemos todos los datos y la estructura cargados en Power Bi, la forma de trabajar será similar sea cual sea el método por el cual hemos llegado hasta aquí.

Incluso aunque se traten de distintas bases de datos se trabaja de forma similar. Debemos relacionar que variables queremos analizar en función de que campos.

### 5.5.1. Presentación del panel general

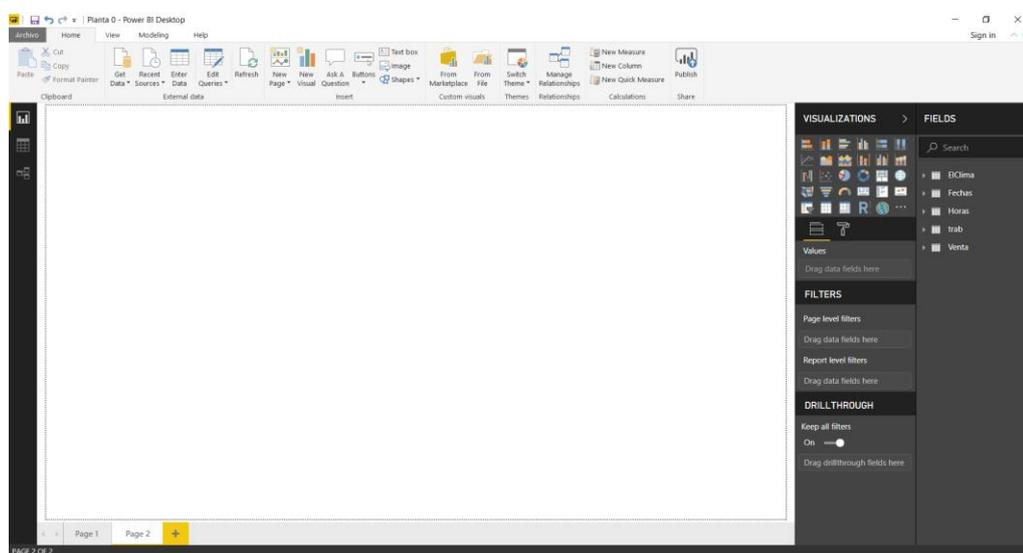


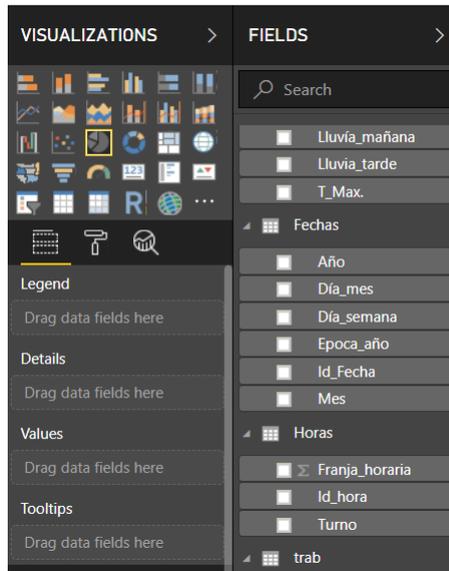
Ilustración 5.14. Panel general del Power Bi.

En la zona de la izquierda podemos elegir entre ver las gráficas, las tablas de datos o las tablas y sus relaciones.

A la derecha será el panel para crear gráficos, seleccionaremos el tipo de gráfica y los valores de campos queremos que nos represente en función de que otros campos.

En la parte superior tenemos el panel de control para guardar, cargar otros datos, modificar tablas y conexiones... Es un panel realmente intuitivo.

### 5.5.2. Creación de gráficas



Debemos seleccionar el tipo de gráfica que queremos y arrastrar al recuadro “values” el campo que queremos analizar. En “Legend” y “detail” arrastramos en función de que queremos analizar. Por ejemplo, arrastraríamos “Valor” de la tabla venta a “values” y “edad” de la tabla trabajados a “details”. Se nos generaría un gráfico que nos relacionaría esto. Para las gráficas de barras podemos seleccionar los ejes y para otros tipos de gráficas otras restricciones, pero eso lo explico en el anexo 1.

Ilustración 5.15. Panel gráficos Power Bi.

### 5.5.3. Resultados del uso del Bi.

Inicialmente pensé que podría sacar alguna relación más, por ejemplo, relacionar cada producto con los ingresos que genera. Pero debido a como genera los datos la empresa es imposible. Como mencioné anteriormente en varias ocasiones, esto está muy limitado por los datos que genera la empresa y lo que es más importante... por como los genera. La planta 0 genera menos campos, pero mejor estructurados lo que nos permite un mejor análisis.

En el dashboard de la planta 0 representamos:

- Una tabla que relaciona la venta de cada trabajador en función del mes.
- Tenemos representado también el volumen de venta en función de:
  - Día del mes con un gráfico de líneas.
  - Edad del empleado mediante un gráfico circular.
  - Día de la semana mediante un gráfico de barras.
  - De la temperatura mediante un gráfico de anillo.
  - Si llovió o no ese día con barras horizontales.
  - De la hora mediante un gráfico de área.
  - Del sexo del empleado mediante un gráfico de anillo.
- Además, tenemos una botonera para separar indicar si queremos que se nos muestre solo gráficos de los turnos de mañana, solo de por los turnos de tarde o ambos casos.

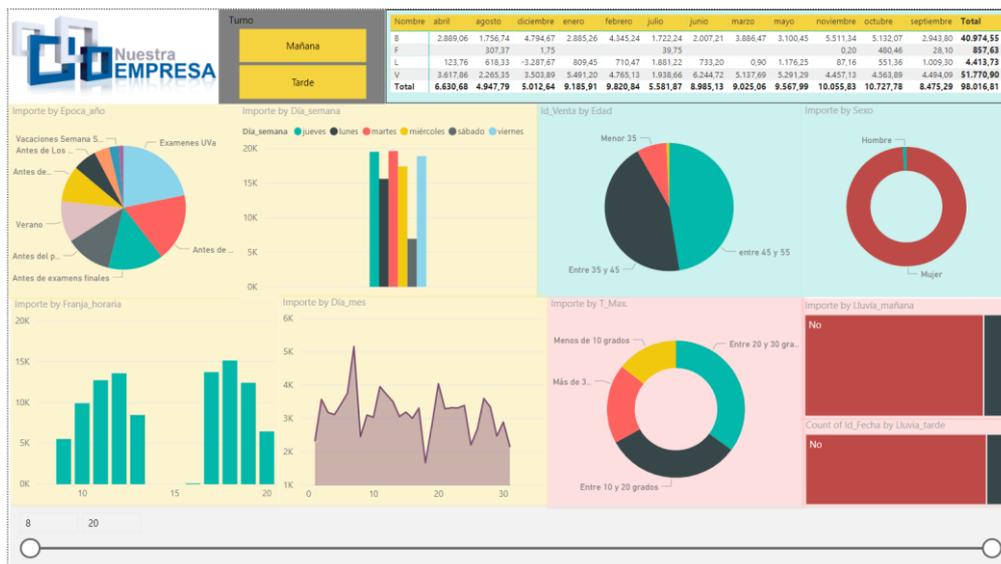


Ilustración 5.16. Dashboard planta 0

En la planta 1 representamos:

- Una tabla que relaciona las ventas de cada cliente en función del mes.
- Una tabla con la cantidad vendida de cada producto.
- Un mapa donde sale representado en mapa las distintas promociones donde se han realizado las ventas.
- La cantidad de productos vendidos en función del mes.
- El total facturado en función:
  - Del año con un gráfico de barras.
  - De los meses del año con un gráfico de anillo.
  - Del mes y del año con un gráfico de barras.



Ilustración 5.17. Dashboard planta 1

## 5.5.4. Recomendaciones para la empresa

### Recomendaciones sobre los datos

Las recomendaciones para la empresa van especialmente dirigidas a que mejoren fácilmente los datos de los que disponen.

En ninguna de las dos plantas he podido relacionar los ingresos en función del producto. En la planta 0 es difícil que al hacer una venta el empleado marque el producto en el TPV dentro de todos los que tienen, pero los productos deberían estar divididos en secciones. De esta forma, hubiese podido analizar los ingresos en función de los productos. En la planta 1, si los albaranes tuviesen el precio, al ser un albarán para cada producto podríamos relacionar los importes de ventas en función del producto. Supongo que para que el precio saliese referenciado en los albaranes, habría que hacer alguna modificación informática en el programa que los genera.

Respecto a los clientes, sería bueno tener más información. En la planta 0, se podría hacer un sistema de fidelización de cliente. Por ejemplo, se podrían hacer carnets de cliente que supusiese algún descuento para el cliente por tenerlo. De esa forma, si el cliente se identifica al hacer una compra, tendríamos registrado sexo, edad, dirección del cliente y otros aspectos que pueden resultar de interés para la empresa.

En la planta 1, se podría añadir más información del cliente, por ejemplo, el sector de la empresa o el número de empleados.

Además, como explique en otro apartado, fue imposible encontrar datos no estructurados en las redes. La empresa debería tener redes sociales donde los clientes puedan valorar los productos y así también poder sacar información de ellos.

### Recomendaciones empresariales

En cuanto a las decisiones empresariales debería tomarlas el gerente de la empresa, pero con las gráficas que se representan en el dashboard podemos sacar conclusiones acerca de las fechas, las horas o los trabajadores que más ingresos generan para la empresa. También podemos saber que productos se consumen más o por ejemplo, si merece la pena mantener abierta la tienda las últimas horas de un día que llueva de diciembre.

## 6. CONCLUSIONES Y EXTENSIONES FUTURAS

Es este capítulo expongo las conclusiones que he extraído tanto del proceso de recolección de información de la parte teórica como de la parte práctica.

El principal objetivo de este documento era demostrar cómo sin conocimientos previos de informática el gerente o dueño de una pequeña empresa podía sacar rendimiento de los datos que genera su empresa pero gracias a la lectura teórica previa al trabajo, he descubierto este mundo el cual, personalmente opino que no se le da la importancia que merece en el grado que he estudiado.

El mundo del Business Intelligence es un mundo en constante desarrollo y que ya ha evolucionado no solo al análisis de los datos generados en el pasado, si no qué al análisis de las situaciones empresariales en el futuro, esta rama se conoce como Business Analytics.

En mi opinión, cualquier gran empresa que no destine una parte de sus recursos a esta tecnología, estará en clara desventaja competitiva con sus competidores lo que suele suponer unos malos resultados para el negocio.

En el caso que nos concierne, el de las pequeñas empresas, puede que en parte sea distinto. Tener esta desventaja competitiva respecto al resto de pequeñas empresas, es una desventaja competitiva que quizá pueda sustituir con otros factores, pero está claro que la empresa que maneje estas herramientas dispondrá de un valor añadido del que no disponen el resto de empresas.

En cuanto a la implementación del sistema de BI en una pequeña empresa, mi primera opción era realizar solo una de las opciones, pero me di cuenta qué la implementación por ese método si podía ser algo complicado implementar para los dueños de algunas empresas.

La tarea de limpieza de datos es una tarea muy tediosa que debe ser solucionada desde el origen, una correcta generación de los datos es fundamental para facilitar la posible implantación.

Por otro lado, como refleja la pequeña encuesta, opino que el mundo del BI es el gran desconocido dentro de las pequeñas empresas, si las

empresas estuviesen más formadas respecto a este tema, estoy convencido que pondrían más empeño en implementar alguna herramienta de BI por simple que sea.

Cualquier empresa puede tener un registro de sus ventas, sus trabajadores, sus proveedores incluso sus clientes, eso es más que suficiente para poder sacar unas conclusiones que mejoren la gestión de la empresa gracias a los datos.

En definitiva, las herramientas de BI son una de las mayores oportunidades de mejora y de marcar la diferencia para las empresas desde la llegada de la informática a estas, las posibilidades de BI son crecientes y varían de acuerdo con la situación de cada empresa y el momento de cada mercado, pero el hecho es que el futuro de análisis de datos ha llegado.

**En cuanto a la línea de trabajo futuro**, la posibilidad de analizar datos no estructurados obtenidos de la web aportaría más información al negocio.

Continuar este proyecto hacia la implementación de una sencilla herramienta de Business Analytics sería interesante, quizá no se podría hacer una herramienta de predicción como utilizan las grandes empresas, pero si una herramienta que cambiando ciertas variables y configurando los gráficos adecuadamente pueda hacer una simulación de una posible situación futura.



## 7. BIBLIOGRAFÍA

- [1] Hans Peter Luhn (1958) “*A Business Intelligence System*”.  
<http://altaplana.com/ibm-luhn58-BusinessIntelligence.pdf>
- [2] Howard Dresner, “*Lack of trust is one of the greatest barriers to success in business intelligence programs*” -  
<http://sandhill.com/article/dresners-point-in-business-intelligence-trust-is-an-elephant-in-the-room/>
- [3] Edgar Frank Cood. “*A Relational Model of Data for Large Shared Data Banks*” <https://dl.acm.org/citation.cfm?id=362685>
- [4] Ralph Kimball (1996) -“*What is a Data Warehouse?*”  
<http://www.essai.rnu.tn/Ebook/Informatique/The%20Data%20Warehouse%20Toolkit,%203rd%20Edition.pdf>
- [5] Ralph Kimball; “*The Data Warehouse Toolkit*”,  
<http://www.essai.rnu.tn/Ebook/Informatique/The%20Data%20Warehouse%20Toolkit,%203rd%20Edition.pdf>
- [6] Carlo Revilli. *Intelligence stratégique sur Internet*. Paris. 1998: Dunod.
- [7] CUI, Zhan; Damiani, Ernesto y Leida, Marcello. "Benefits of ontologies in real time data access." *Digital EcoSystems and Technologies Conference, 2007. DEST'07. Inaugural IEEE-IES. IEEE*, p. 392-397. 2007.
- [8] Robert Stackowiak,; Joseph Rayman y Rick Greenwald. *Oracle Data Warehousing & Business Intelligence*. John Wiley & Sons, Wiley Publishing, Inc, Indianapolis, 2007.
- [9] Dayana Figarella y Andrés Zamora; “*El poder del conocimiento*”  
[http://sisbib.unmsm.edu.pe/m\\_noticias/archivos/libros/2013/dic2013\\_titulo.pdf](http://sisbib.unmsm.edu.pe/m_noticias/archivos/libros/2013/dic2013_titulo.pdf)
- [10] Thomas Davenport y Laurence Prusak ; “*Working Knowledge: How Organizations Manage what They Know*” ; 229099904\_Working\_Knowledge\_How\_Organizations\_Manage\_What\_They\_Know

- [11] Jaap Bloem, Menno van Doorn, Sander Duivestijn Me the media: rise of the conversation society - VINT editions (research institute of Sogeti), 2009, p. 270.
- [12] Robert S. Kaplan y David P. Norton, *“El cuadro de mando integral”*; [http://aulavirtual.iberoamericana.edu.co/recursosel/documentos\\_para-descarga/Cuadro%20de%20Mando%20Integral,%20da%20Edici%C3%B3n%20%20Robert%20S.%20Kaplan%20&%20David%20P.%20Norton.pdf](http://aulavirtual.iberoamericana.edu.co/recursosel/documentos_para-descarga/Cuadro%20de%20Mando%20Integral,%20da%20Edici%C3%B3n%20%20Robert%20S.%20Kaplan%20&%20David%20P.%20Norton.pdf)
- [13] Luis Carlos Molina; Artículo *“Data Mining: Torturando a los datos hasta que confiesen”*; en UOC <http://www.uoc.edu/web/esp/art/uoc/molina1102/molina1102.html>
- [14] Alba Rifà Ribé Libro *“Diseño e implantación de un sistema de Business Intelligence para la gestión analítica de la facturación y los KPIS relevantes para el negocio en la empresa”*
- [15] Kevin Pant, Prashant. Business intelligence (BI). How to build successful BI strategy. Deloitte Touche Tohmatsu Limited. 19 p. 2009.
- [16] Jeff Law. Business Intelligence: Good performance indicators. Sage ERP X3. White Paper.
- [17] Informe del gobierno sobre las PyMEs; *“Retrato de PYME - enero 2018”*
- [18] Alan Habjan, Ane Popovic, A. Achieving Business Process Change With Improved Business Intelligence Systems: A Case Of Slovenian Company. 7th WSEAS International Conference on Applied Computer Science. Venecia, Italia. 2007.
- [19] Pegasus Software. 2008. Using Business Intelligence to navigate in stormy times: an SME's Guide.
- [20] Teresa Guarda; Manuel Santos; Filipe Pinto; Maria Augusto y Carlos Silva. Business Intelligence as a Competitive Advantage for SMEs. International Journal of Trade, Economics and Finance, Vol. 4, No. 4. Agosto, 2013.
- [21] Marcus Gibson y David Arnott. "BI for Small Business: Assessment, Framework & Agenda." 7th Pacific Asia Conference on Information Systems. 2003.



- [22] Mark Kirchmer. Small and Medium Enterprises Also Benefit from MPE. High Performance Through Process Excellence, p. 147-157. 2011.
- [23] Neil Fuller Love. Management development in small firms. International Journal of Management Reviews, 8, p. 175 - 190. 2006
- [24] Olivier Suleymanov. (2011). Business Intelligence Strategy. Business Intelligence for Small and Middle size Enterprises. 2016, de <http://es.slideshare.net/olimjons/business-intelligence-strategy-for-smes>
- [25] Martin Cayón. (2015). Cloud Computing, propulsor en la adopción de TI en las pymes. Recuperado el 20 de Septiembre de 2016, de <http://mundocontact.com/cloudcomputing-propulsor-en-la-adopcion-de-ti-en-las-pymes/>
- [26] Pedro Rojo. “Cómo se benefician las PyMEs de una solución de Business Intelligence.” 2016 de <http://dataiq.com.ar/blog/beneficiosbusiness-intelligence-para-pymes/>
- [27] John Elegant BI. (2009). The Value Proposition of Business Intelligence (BI) for Small and Medium Enterprises <https://www.researchgate.net/file.PostFileLoader.html?id=565f06165dbbbd285c8b4568&assetKey=AS%3A302227481137154%401449068054695>
- [28] Lantares Solutions. Aplicaciones prácticas del análisis de datos para las PYME. Recuperado el 18 de Septiembre de 2016, de <http://www.lantares.com/blog/aplicaciones-practicas-del-analisis-de-datos-para-las-pyme>
- [29] Valda, J. (2010). Principales características del empresario PyME argentino. Recuperado el 16 de Mayo de 2016, de <http://www.grandespymes.com.ar/2010/10/07/principales-caracteristicas-del-empresario-pyme-argentino/>
- [30] Proyecto de final de carrera defendido en la universidad de Palermo, Argentina. “*Investigación sobre la aplicación de business intelligence en la gestión de las PyMEs en Argentina*”- Clarysabel Tovar
- [31] Santiago Cruz. (2014). Soy PyME, ¿Necesito Business Intelligence?. Recuperado el 18 de Septiembre de 2016, de [http://www.yankuik.com/es/pymes\\_y\\_bi/](http://www.yankuik.com/es/pymes_y_bi/)

- [32] Información sobre IBM cognos analytc - <https://www.ibm.com/es-es/marketplace/business-intelligence>
- [33] Tableau - <https://www.tableau.com/es-es>
- [34] Pentaho - <http://www.stratebi.com/pentaho>
- [34] Tibco - <https://spotfire.tibco.com/es/solutions/tibco-spotfire-para-servicios-financieros>
- [36] Ab Initio - <https://www.abinitio.com/en/>
- [37] Azure Analysis Services - <https://docs.microsoft.com/es-es/azure/analysis-services/analysis-services-overview>
- [38] Oracle - [https://es.wikipedia.org/wiki/Oracle\\_Database](https://es.wikipedia.org/wiki/Oracle_Database)
- [39] SAP Business Objects - <https://gravitar.biz/business-objects/>
- [40] Microstrategy - <https://churriwifi.wordpress.com/indice-de-publicaciones/>
- [41] Microsoft Dynamics - <https://dynamics.microsoft.com/es-es/>
- [42] Ultimus  
<http://www.ultimus.com/es/bpmsoftwaremodules/process-administrator>
- [43] Microsoft SharePoint - <https://support.office.com/es-es/article/%C2%BFqu-%C3%A9-es-sharepoint-97b915e6-651b-43b2-827d-fb25777f446f>
- [44] QlikView - <http://prod.qlik.com/us/>
- [45] Jet Reports - <https://www.quonext.com/software-gestion/business-intelligence/jet-reports>
- [46] Sisense - <https://www.capterra.es/software/86955/sisense>
- [47] Clear Analytics - <https://reviews.financesonline.com/p/clear-analytics/>
- [48] Google Analytics - <https://marketingplatform.google.com/about/analytics/>
- [49] Good Data - <https://www.gooddata.com/industry/insurance>



- [50] Eclipse Birt - <http://www.vogella.com/tutorials/EclipseBIRT/article.html>
- [51] Power Bi - <https://powerbi.microsoft.com/es-es/>
- [52] About Laura Madsen - <https://www.linkedin.com/in/lauramadsen>
- [53] Laura Madsen - *"Now is the Time: An ROI for Business Intelligence"*  
- <http://www.b-eye-network.com/view/10222>
- [54] MySQLWorkbench - [www.dev.mysql.com/downloads/workbench/](http://www.dev.mysql.com/downloads/workbench/)
- [55] XAMPP - <https://www.apachefriends.org/es/index.html>
- [56] Excel - <https://support.office.com/es-es/article/informaci%C3%B3n-general-sobre-f%C3%B3rmulas-en-excel-ecfdc708-9162-49e8-b993-c311f47ca173>
- [57] Access - <https://products.office.com/es-es/access>





## 8. ANEXO: MANUAL

En este anexo se incluye un manual más detallado de los pasos descritos en el capítulo 5. Este manual no es solo para el caso del proyecto, si no más general.

### 8.1. Instalación de programas necesarios

Microsoft Access y Microsoft Excel vienen con el paquete Office. Por lo general casi todas las empresas tienen instalado este paquete en sus ordenadores. En el caso de no disponer de estos programas, Apache Open Office nos ofrece programas similares al Access y Excel además del Word y otros de los que viene con el paquete de Microsoft.

XAMPP, MySQL Workbench y Power Bi pueden descargarse de forma gratuita. Power Bi tiene una versión de pago que incluye más herramientas.

Enlaces de descarga:

Apache Open Office: <https://www.openoffice.org/es/descargar/>

XAMPP: <https://www.apachefriends.org/es/index.html>

MySQL WorkBench: <https://www.mysql.com/products/workbench/>

Power Bi: <https://powerbi.microsoft.com/es-es/get-started/>

Extensión P Bi: <https://dev.mysql.com/downloads/connector/net/>

Al realizar la instalación seleccionar siempre los valores predeterminados, ninguno de instaladores incluye otros programas secundarios.

### 8.2. Implantación utilizando solo el Power Bi

Podemos realizar la implantación cargando las tablas directamente en el Power Bi, para ellos es recomendable tener nuestros datos en formato tabla.

#### **8.2.1. Creación tablas con Excel**

Para crear una tabla con Excel hay que seleccionar todos los datos y en la pestaña “inicio” seleccionar “Dar formato a tabla”. En caso de tener encabezados, hay que indicarlo cuando Excel nos lo pregunte, si no se hace, entenderá los encabezados como otra línea de datos.

## 8.2.2. Cargar datos a Power Bi

1 - En la pestaña “home” en “get data” seleccionamos Excel y se

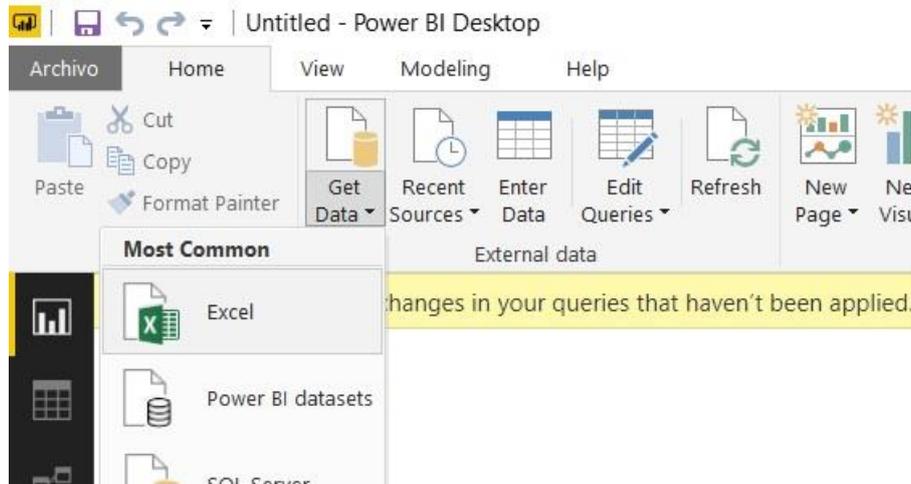


Ilustración 8.9. Captura carga de datos en Power Bi

2 - Se nos abrirá una ventana que nos permitirá elegir entre las distintas pestañas y tablas que incluye el archivo Excel. Seleccionamos la tabla que queremos cargar.

## 8.2.3. Creación de relaciones

En “home”, pinchamos en “Manage Relationships” y se nos abrirá una ventana. En “new” podremos crear las últimas relaciones de las tablas que queramos.

## 8.3. Implantación utilizando Access y Power Bi.

En este método crearemos la base de datos en Access, cargaremos los datos en este y después cargaremos esta base de datos con los datos incluidos al Power Bi.

### 8.3.1. Creación base de datos y carga de datos

En Access aprovecharemos la carga de los datos para crear las tablas de nuestra base de datos. Esto podemos hacerlo porque al hacer limpieza de datos hemos dejado los archivos Excel como queremos que sean las tablas de nuestra base de datos.

1 - Selección origen de datos.

En la pestaña “datos externos”, en “nuevo origen de datos” seleccionamos desde un archivo Excel.

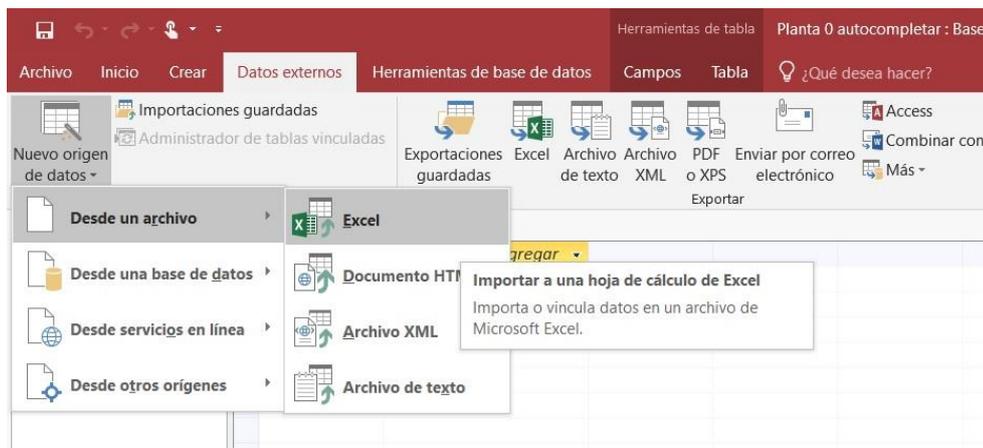


Ilustración 8.2. Captura carga de datos en Access.

Al seleccionar se nos abrirá una ventana que nos preguntará que archivo queremos seleccionar. También nos preguntara si queremos vincular la tabla al origen de los datos, si queremos anexar una copia de los registros de las tablas o si queremos importar el origen de datos en una nueva tabla de la base de datos. Seleccionaremos esta última opción.

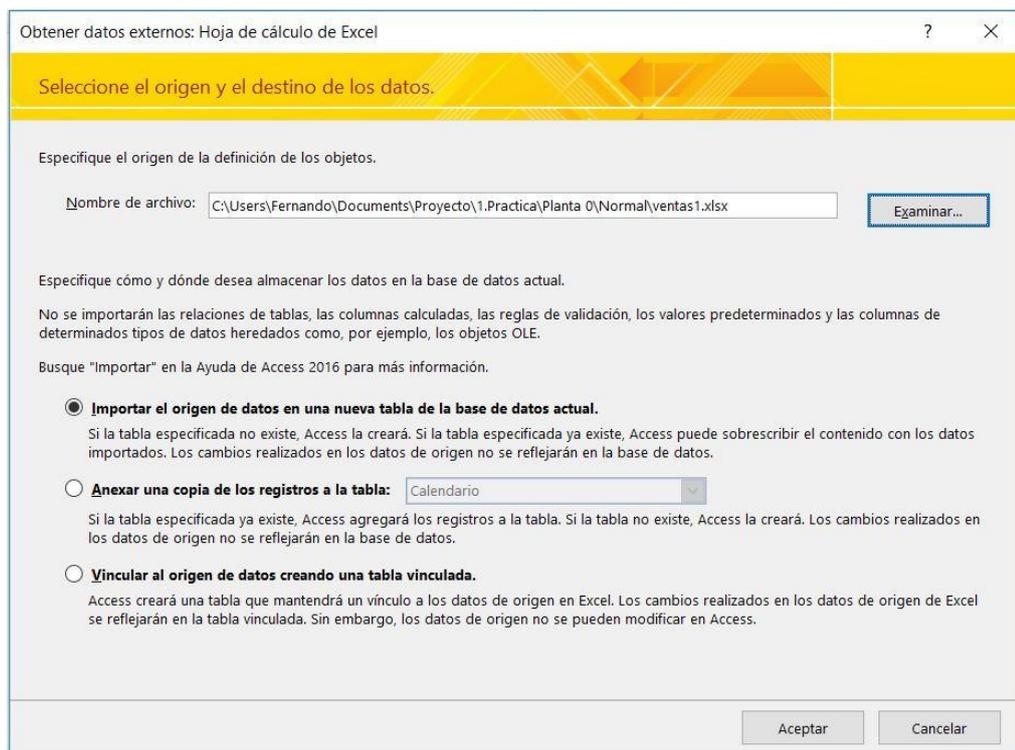


Ilustración 8.3. Captura selección tipo de importación.

## 2 -Importación de los datos

Se nos abrirá una ventana con una vista previa de la tabla, igual que en Excel deberemos indicar si tiene encabezados y clicar en siguiente.

En la siguiente venta tendremos que indicar que tipo de dato es cada columna, si está indexado o no y si está duplicado o no.

El estar indexado funciona como el índice de un libro, nos permite realizar búsquedas más rápidamente. Las claves primarias deberán ir indexadas y ser únicas.

Los tipos de datos generalmente pueden ser enteros, decimales (doble), texto largo, texto corto, fecha y hora y cadena de caracteres.

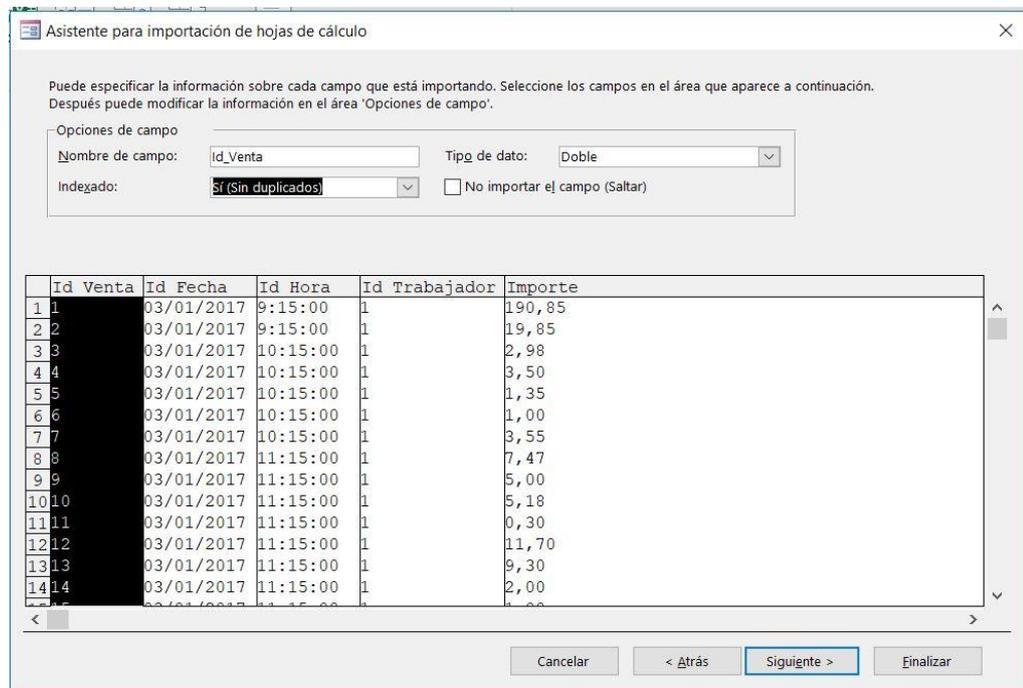


Ilustración 8.4. Captura asistente importación de datos.

En la siguiente ventana deberemos elegir cual será nuestra clave principal de esa tabla. Si hemos seleccionado mal entre las opciones de indexado, nos dará error.

Deberemos repetir este proceso con todas nuestras tablas, al acabar deberán salirnos todas nuestras tablas en la zona de la izquierda.

### 3 – Creación de relaciones entre tablas

En la pestaña “herramientas de datos” seleccionamos “relaciones”. Se nos abrirá una nueva pestaña “diseño”. En esa pestaña diseño pinchamos en “modificar relaciones” y seleccionamos las tablas que queremos cargar para modificar sus relaciones.

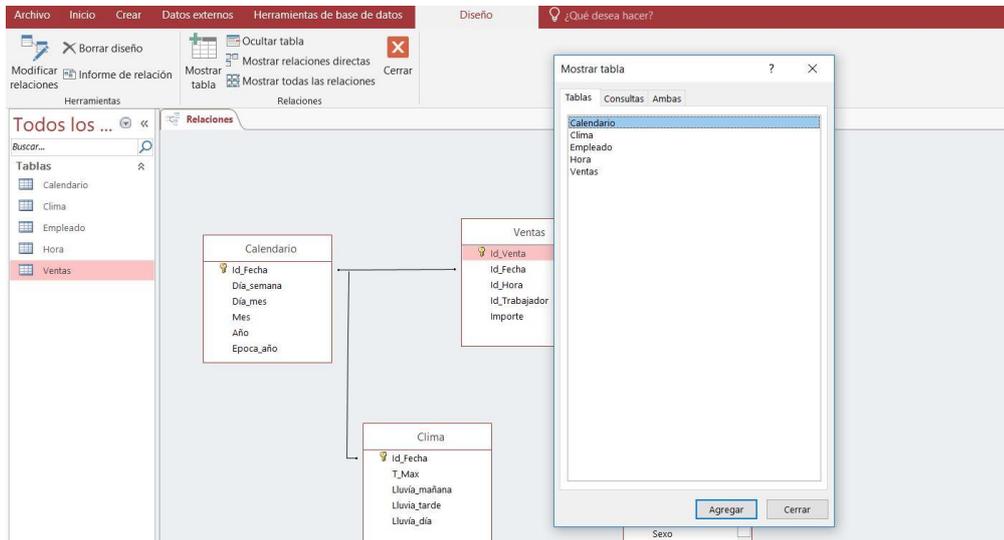


Ilustración 8.5. Captura creación de relaciones entre tablas en access.

Pinchamos en “modificar relaciones” y se nos abrirá una ventana donde podemos crear las relaciones. Seleccionamos las tablas y los campos que queremos relacionar.

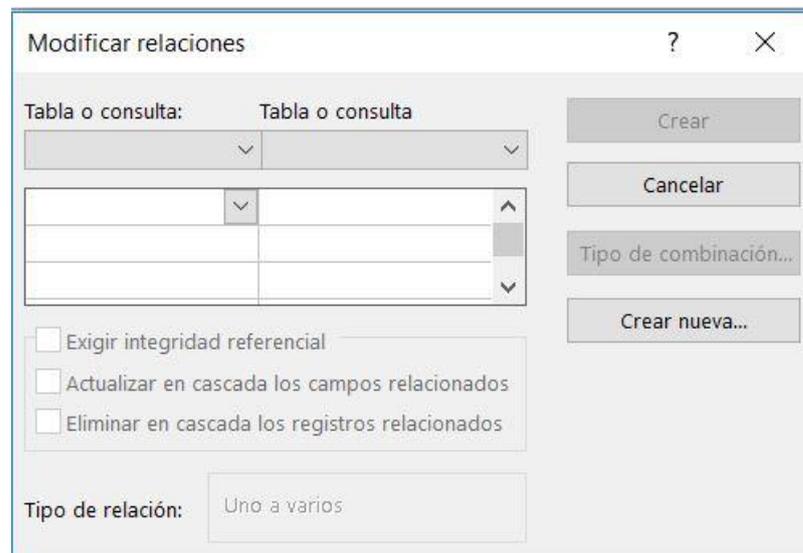


Ilustración 8.6. Captura panel de creación de relaciones.

Hacemos todas las relaciones que queramos y deberían aparecernos todas las conexiones entre las tablas.

### 8.3.2. Carga base de datos a Power Bi.

Este paso es muy similar al de la carga de las tablas Excel a Power Bi. En la pestaña “home”, seleccionamos “get data” y “Access Data Base”. Seleccionamos la base de datos que queremos cargar y dentro de esa base de datos nos deja seleccionar las tablas que queremos cargar. Cargamos todas las tablas y ya tendríamos cargado los datos, con las relaciones entre las tablas.

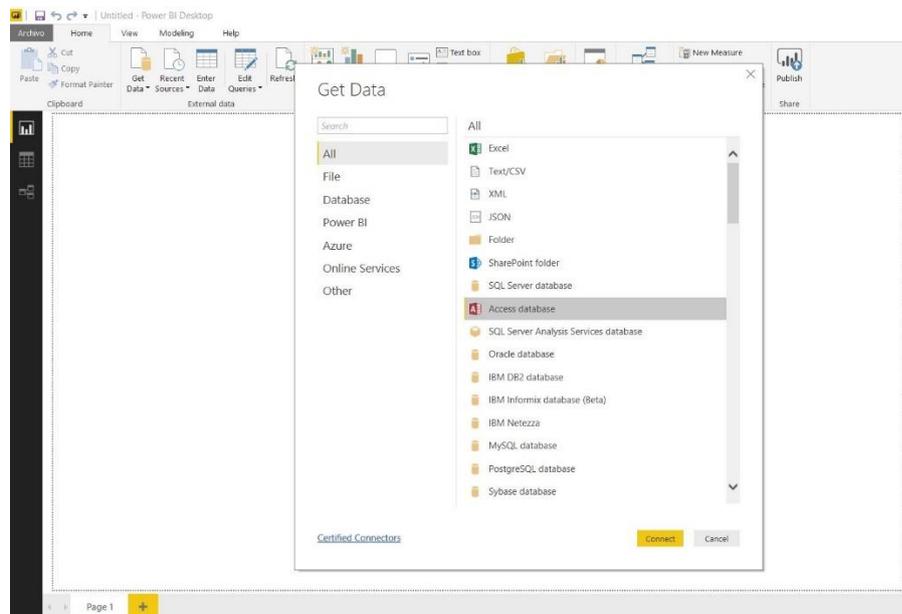


Ilustración 8.7. Captura carga de datos en Power Bi

## 8.4. Implantación utilizando MySQL y Power Bi.

Para crear una base de datos usaremos XAMPP y MySQL Workbench. Primero crearemos las tablas, luego las relaciones y por último cargaremos los datos.

### 8.4.1. Creación de las tablas

1 – Lanzar el MySQL:

En primer lugar tendremos que abrir el XAMPP e inicializar Apache y MySQL pulsando en “start”, a continuación abrir MySQL pinchando en “admin”.

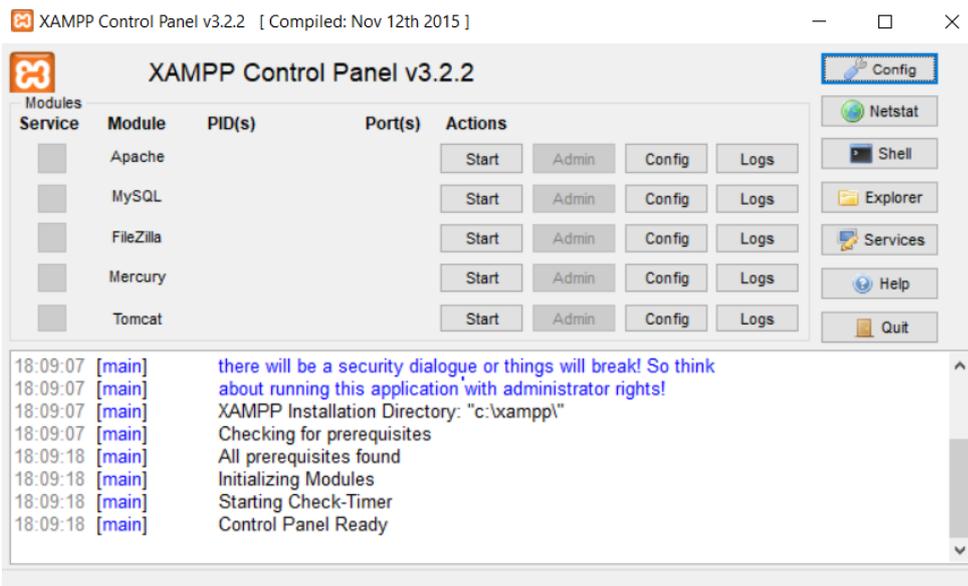


Ilustración 8.7. Captura panel de control de XAMPP.

2 – Crear la base de datos y tablas.

En la siguiente pantalla nos pedirá que demos nombre a la base de datos que queremos crear y entonces ya podremos crear tablas pinchando en “nueva tabla”.

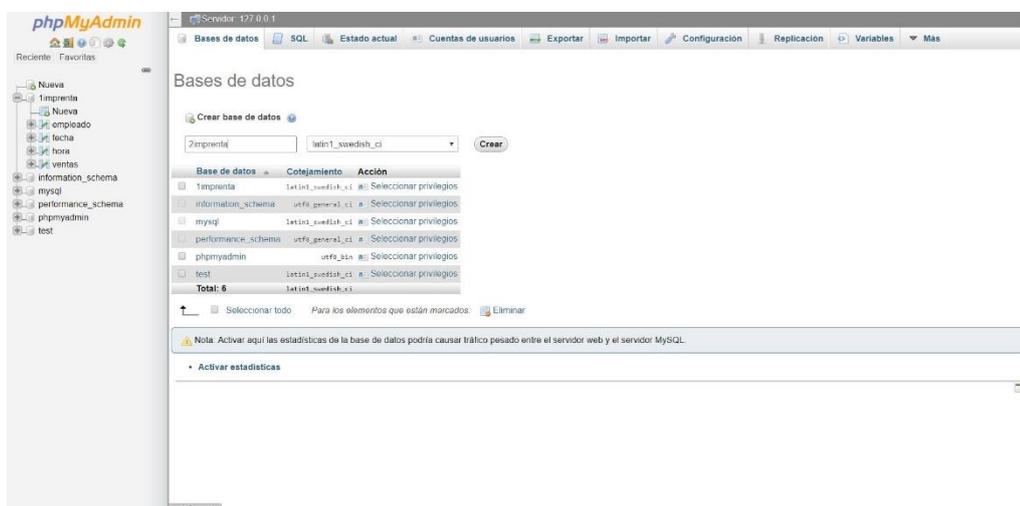


Ilustración 8.8. Captura creación base de datos en MySQL.

Para la creación de tablas deberemos definir el tipo de datos que ira en cada columna, los más habituales son los mismos que los mencionados anteriormente para crear la base de datos con Access.

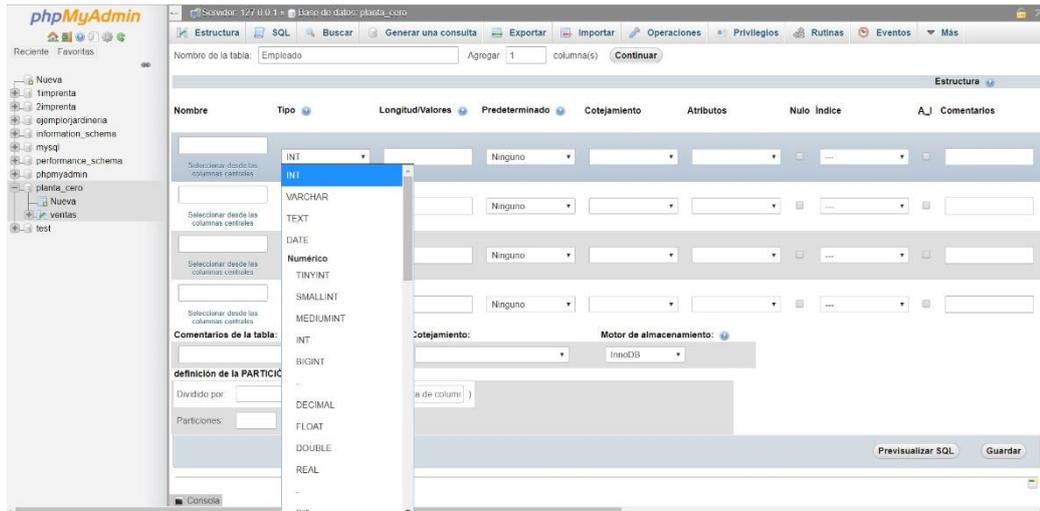


Ilustración 8.9. Captura selección tipo de dato de un campo.

### 3 – Creación de relaciones,

Para la creación de relaciones entre tablas usaremos el MySQL Workbench. Al abrirlo, pulsando en “+”, crearemos una conexión a la base de datos que queramos. Nos saldrán varios avisos que tendremos que aceptar.

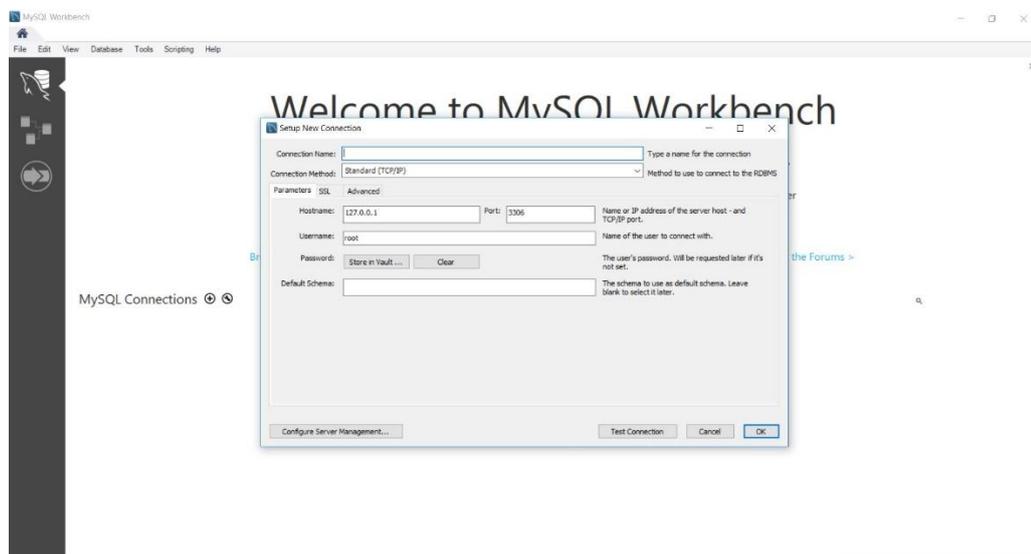


Ilustración 8.100. Captura carga de datos en Power Bi

Creamos la conexión y pinchando sobre un símbolo de una llave inglesa que viene junto al nombre de cada tabla se nos abrirá un ventana para gestionar esa tabla.

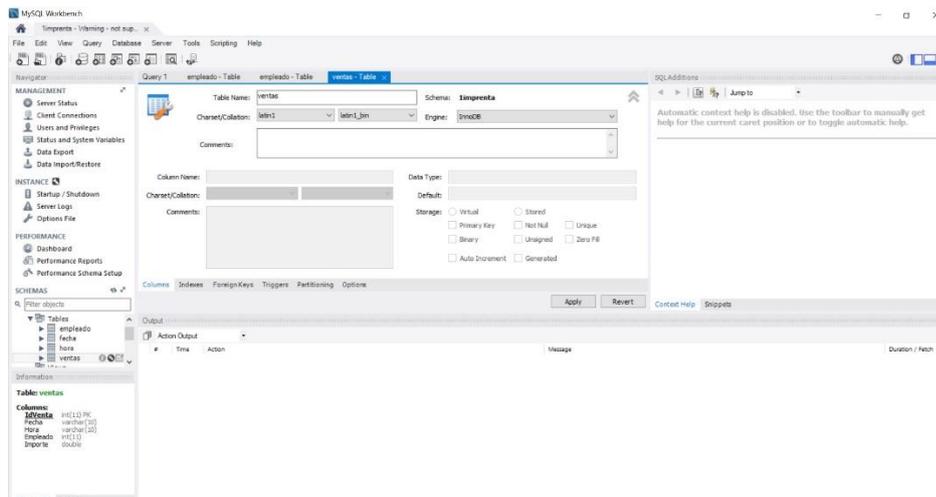


Ilustración 8.111. Captura panel de control MySQL Workbench.

Pinchando en “foreigns Keys” se nos abrirá una pestaña donde podremos crear las relaciones entre tablas. Seleccionamos las tablas que deseamos crear, damos a “Apply” y nos saldrá el código que ejecutará, dando a “Ok” se creará nuestra relación.

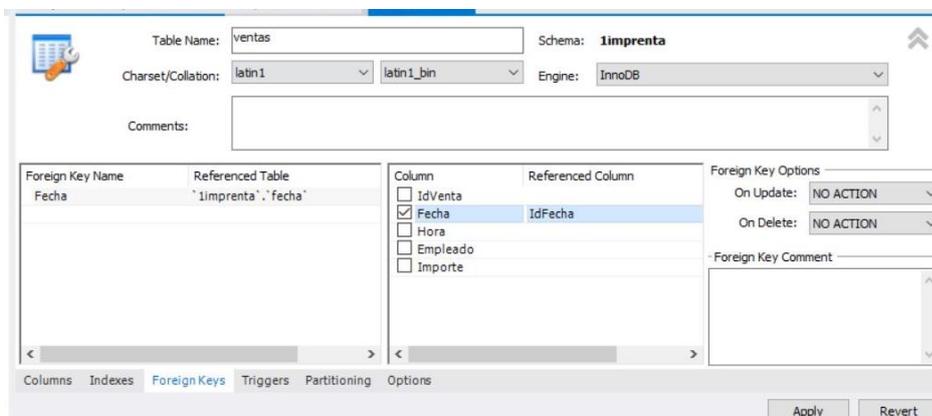


Ilustración 8.122. Captura panel creación relaciones en MySQL Workbench.

Si volvemos al MySQL lanzado por XAMPP y pinchamos en “diseñador”, nos saldrán las relaciones que hemos creado.



Ilustración 8.133. Captura tablas creadas en MySQL.

#### 4 – Carga de datos.

Volvemos a MySQL Workbench, pulsando con el botón derecho sobre la tabla que queremos cargar de datos, seleccionamos “table data import wizard”. Y seleccionamos el archivo que queremos importar.

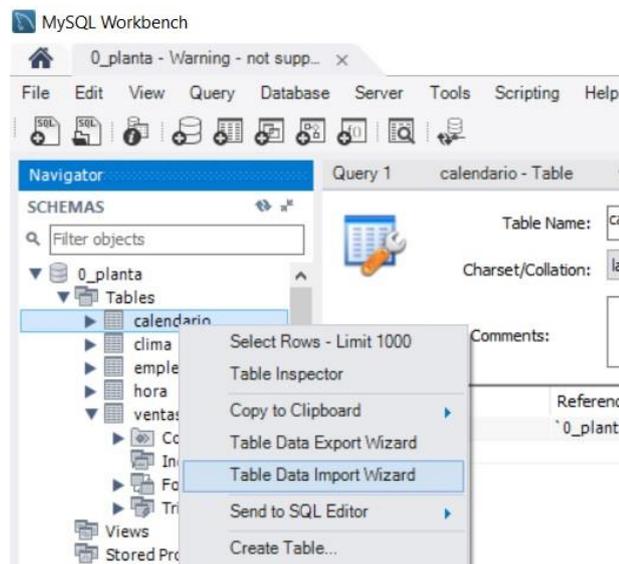


Ilustración 8.144. Captura selección importación de datos en MySQL Workbench.

Es muy importante que ese archivo sea de tipo CVS delimitado por comas, si no, no podrá importarlo.

En las siguientes ventanas tendremos que identificar cada columna del archivo CVS con cada columna de la tabla de datos y ya estarían cargados los datos.

## 5 – Conexión con Power Bi.

Al igual que en los casos anteriores, pincharemos en “get data” y seleccionamos “MySQL Data base”. Al ser la primera vez nos saldrá un aviso de que debemos instalar una extensión para continuar. El enlace de dicha extensión esta al principio del capítulo con el resto de enlaces.

Una vez instalada la extensión, nos pedirá información para poder conectar a la base de datos. En gestos MySQL de XAMPP tenemos el servidor, en el panel de control desde donde lanzamos el MySQL y Apache, tenemos el puerto.

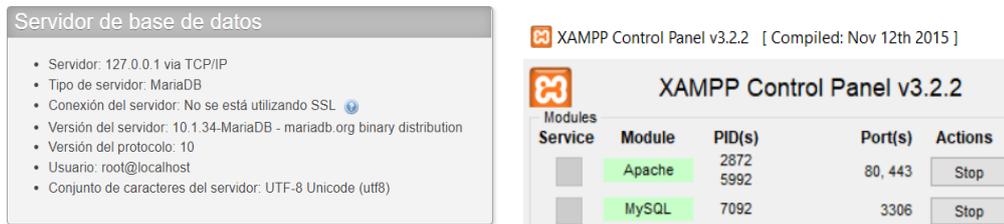


Ilustración 8.155. Captura donde sacar información puerto y servidor.

Con estos datos, el nombre de nuestra base de datos podremos cargar a Power Bi nuestra base de datos.

## 8.5. Utilización de Power Bi

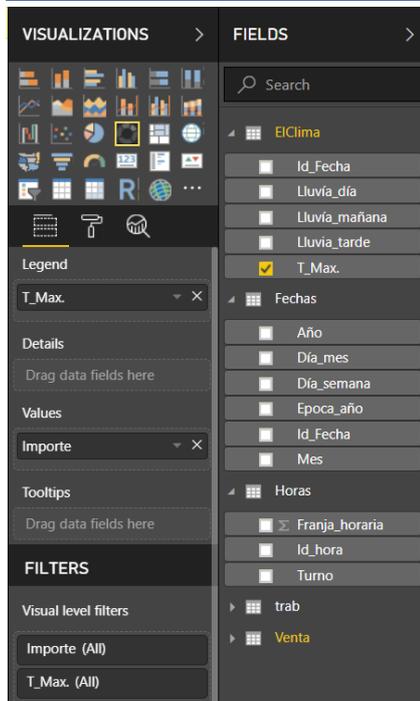


Ilustración 8.166. Captura panel power Bi.

En la parte de la izquierda podemos cambiar lo que queremos ver, podemos elegir entre las gráficas, las relaciones entre tablas y las tablas con sus datos.

En el panel de la zona derecha, el de la foto, podemos crear distintos gráficos. Debemos seleccionar el gráfico que queremos y arrastrar a la casilla “Values” el campo de la tabla que queramos analizar. En las casillas “x axis”, “y axis”, “Details” y “Legend” arrastraremos en función de que queremos analizar. En función de estos campos será mejor un tipo de gráfico y otro. Pinchando sobre el logo que es como un rodillo de pintura, podremos acceder a las herramientas para modificar estéticamente los gráficos.