

CIPFP <i>www.fpmislata.com</i>							
<b>Actividad:</b>	<b>Pentaho Data Integration (PDI). Practica Global</b>						
<b>Ciclo:</b>	IABD	<b>Modulo:</b>	BDA	<b>Curso:</b>		<b>Agrupación:</b>	1
<b>Alumno/a:</b>						<b>Grupo:</b>	

## Actividad

Una empresa de venta de componentes de ordenador quiere explotar la información que almacena sobre sus ventas y sus empleados. Y para conseguirlo te ha contratado para que le diseñes un proyecto de Big Data, creando un Data Warehouse para explotar los datos de la empresa.

El Data WareHouse debe responder como minimo a las siguientes consultas:

- Consulta de Ventas por Cliente:
  - Obtener la lista de clientes junto con la cantidad total de compras y el monto total gastado.
- Consulta de Productos Más Vendidos:
  - Identificar los productos más vendidos en términos de cantidad y monto total de ventas.
- Consulta de Rentabilidad de Productos:
  - Analizar la rentabilidad de cada producto en base a la diferencia entre el precio de venta y el costo.
- Consulta de Actividad de Usuarios
  - Obtener un registro de las acciones realizadas por cada usuario en la plataforma.
- Consulta de Ventas del Último Año:
  - Recuperar las ventas que ocurrieron en el último año.
- Productividad de los usuarios:
  - Número de ventas por hora trabajada.

Los orígenes de datos que nos proporciona la empresa son:

- Base de Datos Relacionales (MySQL)
  - Nombre de la BD: **Empresa**
  - Tablas: **Clientes, Ventas, Productos**
- Archivo Excel:
  - Nombre del archivo: **Datos\_Empleados.xlsx**
  - Hoja de Cálculo: **Empleados**
  - Contenido: **Datos Empleados**
- Datos en MongoDB Atlas:
  - Base de datos: **Sistema**
  - Colecciones: **Usuarios, Logs**

CIPFP <a href="http://www.fpmislata.com">www.fpmislata.com</a>							
<b>Actividad:</b>	<b>Pentaho Data Integration (PDI). Practica Global</b>						
<b>Ciclo:</b>	IABD	<b>Modulo:</b>	BDA	<b>Curso:</b>		<b>Agrupación:</b>	1
<b>Alumno/a:</b>						<b>Grupo:</b>	

- Acceso con Compass:
  - [mongodb+srv://LuisGB:Lgb\\_123@cluster0.fymln.mongodb.net/](mongodb+srv://LuisGB:Lgb_123@cluster0.fymln.mongodb.net/)
- Acceso con Shell:
  - mongosh "mongodb+srv://cluster0.fymln.mongodb.net/" --apiVersion 1 --username LuisGB

CIPFP <i>www.fpmislata.com</i>							
<b>Actividad:</b>	<b>Pentaho Data Integration (PDI). Practica Global</b>						
<b>Ciclo:</b>	IABD	<b>Modulo:</b>	BDA	<b>Curso:</b>		<b>Agrupación:</b>	1
<b>Alumno/a:</b>						<b>Grupo:</b>	

El proyecto deber tener una ETL que recoja los datos y cree el Data Warehouse. Los requisitos mínimos del proyecto son:

- Analisis de los datos
- ETL con:
  - Minimo un Job
  - Minimo 2 Transformaciones
  - Debe tener validacioines de los datos
  - Debe tener monitorización del proceso
- Data Warehose creado en una BD en MySQL

#### Entrega de la práctica:

- Los jobs y transformaciones
- Un volcado de la BD (utiliza una solo), para facilitar la corrección
- Un ejemplo de cada tipo de consulta relizao en el Data Warehouse creado.

#### AYUDA:

- Enlace interesante para el trabajo con MongoDB en Pentaho Data Integration
  - <https://www.linkedin.com/pulse/describre-mongodb-y-trabajar-con-sus-datos-en-pentaho-dario-bernabeu/?originalSubdomain=es>