

**FACULTAD DE INGENIERÍA**

**DEPARTAMENTO DE INGENIERÍA EN SISTEMAS**

**BASES DE DATOS 2**

**LA INTELIGENCIA ARTIFICIAL EN LA LOGISTICA 4.0**

**Elaborado por:**

|  |  |
| --- | --- |
| **Jorge Isaac Coello Berrios** | **20161003898** |
| **Erick Celin Sánchez Sosa** | **20171003793** |
| **Luis Daniel Mejía Valladares** | **20181006117** |

**Catedrático:**

**Emilson Omar Acosta**

**FECHA DE ENTREGA:**

**Tegucigalpa M. D. C, 12 de agosto 2022**

**Índice**

[● Introducción 3](#_Toc111047348)

[● Objetivo General 4](#_Toc111047349)

[● Objetivo Específico 4](#_Toc111047350)

[● Marco teórico 5](#_Toc111047351)

[● Bases de datos utilizadas 12](#_Toc111047352)

[● Preguntas del negocio 16](#_Toc111047353)

[● Métrica utilizada 17](#_Toc111047354)

[● Conclusiones 25](#_Toc111047355)

[● Recomendaciones 26](#_Toc111047356)

[● Bibliografía 27](#_Toc111047357)

# Introducción

Las bases de datos dentro de una empresa son de vital importancia, debido que es en esta en donde se encuentra el registro de la información necesaria para el desarrollo completo y correcto de una empresa. Para el desarrollo del siguiente proyecto, nos basamos en la empresa “Mini Super Río Abajo” (la base de datos no fue proporcionada; pero se tuvo colaboración por parte de los gerentes en cómo funciona su sistema y la manera en que se desenvuelven en relación al mercado y competencias con distintas empresas). Además, se hace uso de Tableau y PENTAHO para el desarrollo del ETL y algunos reportes de acuerdo a la métrica que se asigna por parte de la empresa.

# Objetivo General

Realizar reportes, llenado de Base de Datos OLAP en relación a la PYME “Mini Super Río Abajo”.

# Objetivo Específico

1. Realizar reportes de la base de datos haciendo uso de Tableau.
2. Dada la base de datos (OLTP), hacer el llenado de la base de datos OLAP y el ETL haciendo el uso de PENTAHO.
3. Conocer el funcionamiento y desempeño del Mini Super Río Abajo.
4. Explicación sobre la métrica que se desea implementar con respecto a la base de datos generada.

# Marco teórico



Historia

Todo empezó a través de la pandemia. El edificio que se utiliza para este Mini super, estaba desocupado, y al tener varios empleados (debido que ellos tienen unos buses de la ruta Río Abajo – Centro, Río Abajo – Mercado; una empresa de Aire acondicionado llamada “Aire Eléctrico”, y un restaurante “La Pérgola”); y con el Covid-19, que todo se detuvo, no se tenía un ingreso, y al tener empleados buenos y con bastante trayectoria, iba a ser una injusticia despedirlos, en especial en el área del transporte (al no ir los buses a trabajar, los choferes no tienen ingresos). Por tanto, se optó a poner una venta de abarrotería para que los mismos empleados de transporte trabajaran y ellos hacerse de sus ingresos; para tenerlos generando sus propios ingresos y también generar de manera personal.

Rol

Es una S de R. L (Sociedad de Responsabilidades Limitadas), porque son 2 los socios que tiene, cuyos nombres son Mario Valeriano y Arely Durón (Gerente General).

Misión

Ser una fuente de empleo, en primer lugar, para nuestros trabajadores (al momento de la pandemia) y para la comunidad.

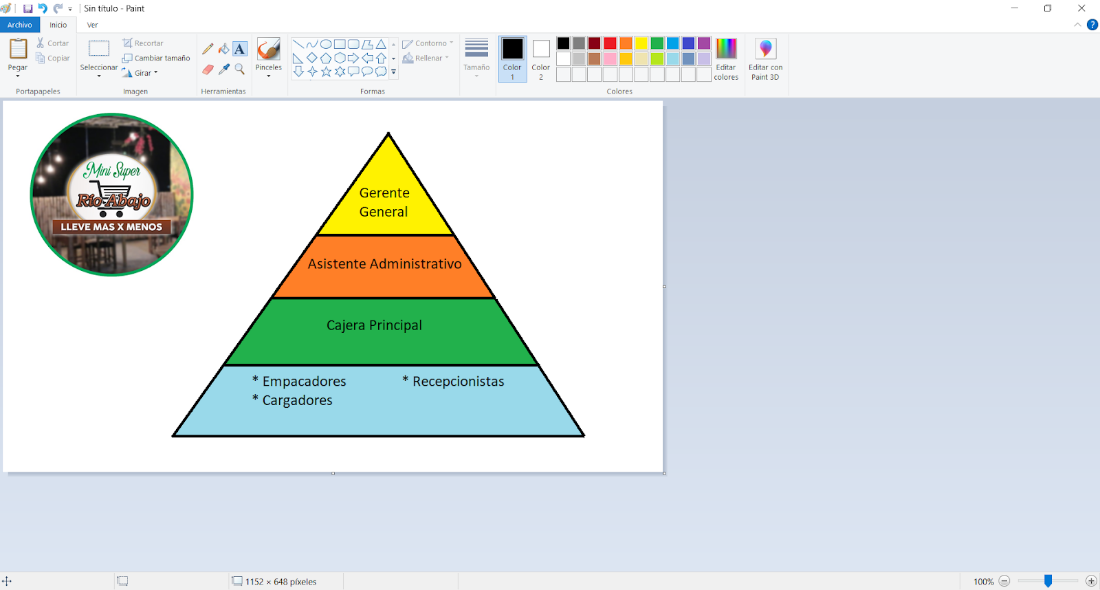
Visión

Que las personas de la comunidad y personas externas que nos visitan, puedan comprar a precios cómodos, sin tener la necesidad de estresarse de salir de la comunidad, teniendo lo básico y un poco más, a precios competitivos.

Competencia

Al principio si se tuvieron ciertas dificultades, por ejemplo, el precio de los huevos estaba a L. 30.00 más el cartón, y por esa diferencia la gente decidió ir a comprar al negocio. Y se quiso tener un 20% de utilidad, manteniendo los mismos márgenes. Y no se ha visto a los demás como competencia, porque por la forma de trabajo es: Si el cliente quiere pan, debe llevar la bolsa completa, y en los alrededores se vende menudeado. Y se quiso esa opción para no afectar a las personas que tienen sus negocios que venden al por menor.

Organigrama



Pentaho Data Integration

Pentaho Data Integration, solución tecnológica conocida como Kettel, es una de las herramientas ETL opensource más potentes y versátiles a la hora de diseñar los procesos de integración a la medida de las necesidades de la empresa, fundamentalmente con el objetivo de construir y explotar su Data Warehouse (DW), entre otras utilidades. Uno de sus aspectos más ventajosos consiste en su capacidad de adaptación, lo que permite desarrollar fácilmente los procesos de integración de datos en el DW, centro de la arquitectura de los sistemas de información, entendidos aquellos como el proceso que consiste en su recolección, cruce, enriquecimiento, consolidación y carga.

Como herramienta ETL, Kettel es visual y consigue una mayor productividad del equipo de desarrollo y mantenimiento, permitiendo conectarse a diversas fuentes y arquitecturas para finamente cargar los datos en un repositorio, si bien también puede utilizarse para otras muchas funciones, como sincronizar, enmascarar, migrar datos entre diferentes aplicaciones o, por ejemplo, para la actualización de usuarios a sistemas paralelos. Según Gartner, Pentaho es una herramienta líder de ETL, que cumple las características más importantes que ha de incluir un software de este tipo, incluyéndola por primera vez en su informe del año 2009 como software open source de probada eficacia en proyectos de tratamiento de datos que necesiten la construcción de procesos ETL.

En concreto, la consultora subrayó la importancia de que un software ETL idóneo incluya una serie de características, como las siguientes: conectividad, capacidad de entrega de datos, de metadatos y modelados de datos, de diseño y entorno de desarrollo, de gestión de datos de administración, además de capacidades SOA y un cierto grado de compactación, consistencia e interoperabilidad. Por otro lado, todas ellas comunes a las herramientas ETL punteras más completas, ya sea Cognos, Microsoft, Informática ETL o, como alternativa opensource, también Pentaho ETL.

*Ventajas de ETL Pentaho:*

* ***Gratuidad opensource:*** su uso no supone un gasto en licencias de software, lo que puede suponer un punto a favor, sobre todo para empresas que empiezan.
* ***Edición de pago:*** también existe una edición de pago (Pentaho BI Suite Enterprise Edition) que sí incluye servicio técnico y aporta soluciones más completas que el software libre.
* ***Comunidad amplia:*** intensa activida a nivel de pruebas beta como en foros, muy útiles para la consulta de posibles problemas, si bien sacarle todo el partido implica un mínimo de especialización.
* ***Tecnologías innovadoras y fácil uso:*** tanto en la versión de pago como en la gratuita, se trabaja con las últimas tecnologías (creado bajo las plataformas de desarrollo de Java y MySQL), sin descuidar una instalación y configuración sencillas.
* ***Múltiples programas:*** ser modular y disponer de aplicaciones para realizar muy distintas funciones, entre ellas la integración de datos, es una de sus grandes ventajas, posibilitando tanto un uso conjunto como independiente en diferentes plataformas.

Tableau

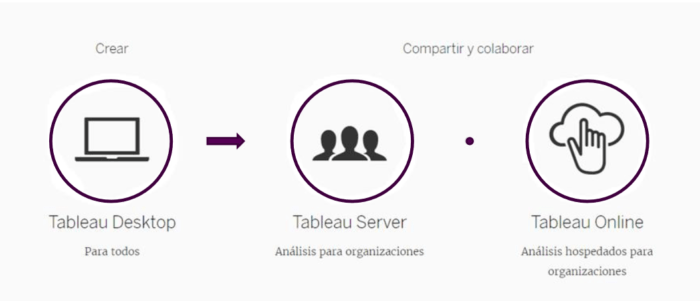
Tableau Software es una solución completa de Business Intelligence, líder en el mercado y que permite a las personas encargadas de tomar decisiones disponer de toda la información necesaria en tiempo real al alcance de sus manos.

Es una herramienta de visualización de datos interactiva, es decir, el usuario tiene la posibilidad de interactuar con los datos: comparar, filtrar, conectar unas variables con otras, etc. Además, los informes y dashboards que se pueden crear con la herramienta son muy visuales lo que facilita la comprensión rápida de los datos.

Tableau, convierte a las empresas en organizaciones “data-driven” haciendo que los datos se conviertan en información útil. Esto se traduce en que las empresas pueden:

* Saber qué está pasando y poder explicar cada situación utilizando los datos.
* Elaborar predicciones sobre qué puede suceder.
* Poner los datos en común con todos los miembros de la organización.
* Recopilar en un único repositorio información dispersa en diferentes sistemas.

*Versiones de Tableau*



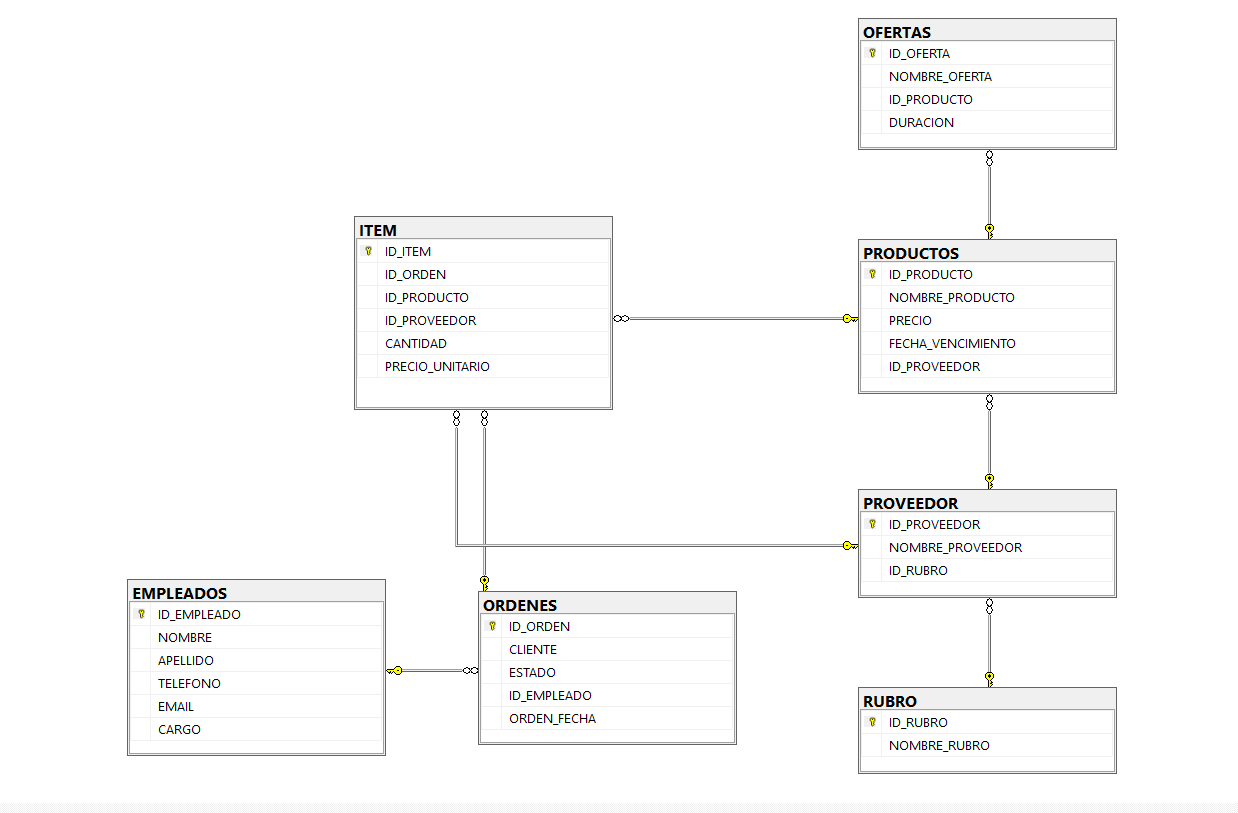
1. ***Tableau Desktop:*** Permite a los usuarios la realización de análisis eficaces de forma autónoma y rápida. Es suficiente con arrastras y soltar para cambiar los parámetros del análisis. Es una herramienta intuitiva que permite visualizar los cambios en el análisis a medida que se van realizando y que es capaz de combinar varios análisis en un único informe, donde además se pueden insertar páginas web o documentos para una mejor explicación y comprensión de las cifras. Tableau Desktop elimina el paso de creación de “universos”, “cubos” o “datawarehouse”, ya que su conexión es directa con los orígenes de datos y no requiere programación. Estos facilita la creación de Data Mashups combinando de forma eficaz datos de diferentes sistemas.
2. ***Tableau Server:*** permite compartir datos e información con los miembros de la organización que se desee para maximizar su impacto, creando un entorno colaborativo. Los usuarios podrán consumir los informes e interactuar con ellos: realizar filtros, actualizar a tiempo real, profundizar datos etc. desde su navegador. Permite difundir el valor de los datos en toda la organización, dando a los usuarios la libertad de explorar los datos en un entorno de confianza, sin limitarlos a preguntas predefinidas, asistentes o tipos de gráficos inflexibles.
3. ***Tableau Online:*** Tableau puede implementarse on-premise en servidores de la organización o en la nube pública usando Tableau Online sin perder en ningún caso el control del servidor por parte de la empresa.

*Ventajas de Tableau:*

* ***Eficacia.*** Permite realizar análisis eficaces para responder a preguntas de negocio complejas rápidamente de forma que las empresas puedan hacer crear su negocio gracias a la mejora en la toma de decisiones.
* Gran capacidad de análisis visual que facilita la comprensión de los datos.
* ***Intuitivo.*** Adopción por parte de los usuarios muy rápida.
* ***Precio.*** Es una solución muy asequible con un precio está por debajo de la media del mercado.
* Plataforma segura, escalable y confiable.

# Bases de datos utilizadas

Base de Datos OLTP



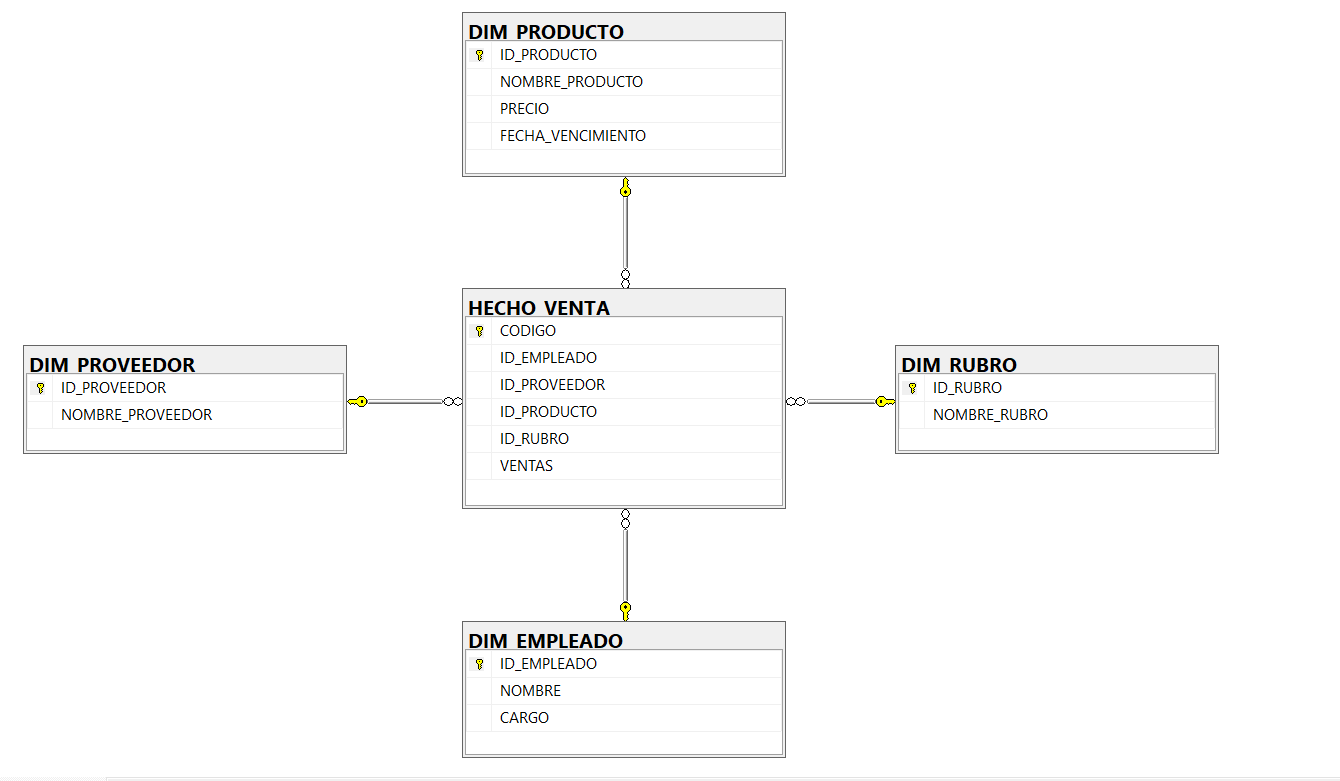
Esta es la base de datos histórica de donde se llenará la base de datos OLAP y se desarrollará el llenado de esta y la respuesta de la métrica dispuesta por la empresa.

Sus tablas son:

* *Empleados*
  + *Id\_Empleado, Nombre, Apellido, Teléfono, Email, Cargo.*
* *Ordenes* 
  + *Id\_Orden, Cliente, Estado, Id\_Empleado, Orden\_Fecha*
* *Item*
  + *Id\_Item, Id\_Orden, Id\_Producto,Id\_Proveedor, Cantidad, Precio\_Unitario*
* *Proveedor* 
  + *Id\_Proveedor, Nombre\_Proveedor, Id\_Rubro*
* *Rubro*
  + *Id\_Rubro, Nombre\_Rubro*
* *Productos*
  + *Id\_Producto, Nombre\_Producto, Precio, Fecha\_Vencimiento, Id\_Proveedor*
* *Ofertas*
  + *Id\_Oferta, Nombre\_Oferta, Id\_Producto, Duración*

Al no ser proporcionada la base de datos del Mini Super Río Abajo, entonces el llenado de esta base de datos fue con la herramienta Mockaroo <https://www.mockaroo.com/>

Base de Datos OLAP



Esta es la base de datos con la que realizará el ETL y los reportes con las herramientas PENTAHO y Tableau.

Sus tablas son:

* Dim\_Producto
  + *Id\_Producto, Nombre\_Producto, Precio, Fecha\_Vencimiento*
* Dim\_Proveedor
  + *Id\_Proveedor, Nombre\_Proveedor*
* Dim\_Empleado
  + Id\_Empleado, Nombre, Cargo
* Dim\_Rubro
  + Id\_Rubro, Nombre\_Rubro
* Hecho\_Venta
  + Código, Id\_Empleado, Id\_Proveedor. Id\_Producto, Id\_Rubro, Ventas

# Preguntas del negocio

Estas fueron las preguntas que se tomaron en cuenta para poder realizar la Base de Datos OLAP:

* De la Dimensión Empleados se quiso saber:
  + ¿Cuál es su código de empleado?
  + ¿Cuál es su nombre?
  + ¿Qué cargo tiene?
* De la Dimensión Proveedor se quiso saber:
  + ¿Cuál es el código del proveedor?
  + ¿Cuál es el nombre del proveedor?
* De la Dimensión Producto se quiso saber:
  + ¿Cuál es el código del producto?
  + ¿Cuál es el nombre del producto?
  + ¿Cuál es el precio del producto?
  + ¿Cuál es la fecha de vencimiento del producto?
* De la Dimensión Rubro se quiso saber:
  + ¿Cuál es el código del rubro?
  + ¿Cuál es el nombre del rubro?
* ¿Cuál es la venta por cliente en base a la fecha?

# Métrica utilizada

Para el caso de la métrica en donde se intenta obtener el total de ventas por fecha de cada empleado se hace el uso de la siguiente consulta SQL:

SELECT CONVERT (DATE, ORDEN\_FECHA) FECHA, EMPLEADOS.ID\_EMPLEADO, PROVEEDOR.ID\_PROVEEDOR, PRODUCTOS.ID\_PRODUCTO,

RUBRO.ID\_RUBRO, SUM (ITEM.CANTIDAD \* ITEM.PRECIO\_UNITARIO) TOTAL FROM ORDENES

INNER JOIN EMPLEADOS ON ORDENES.ID\_EMPLEADO = EMPLEADOS.ID\_EMPLEADO

INNER JOIN ITEM ON ORDENES.ID\_ORDEN = ITEM.ID\_ORDEN

INNER JOIN PROVEEDOR ON ITEM.ID\_PROVEEDOR = PROVEEDOR.ID\_PROVEEDOR

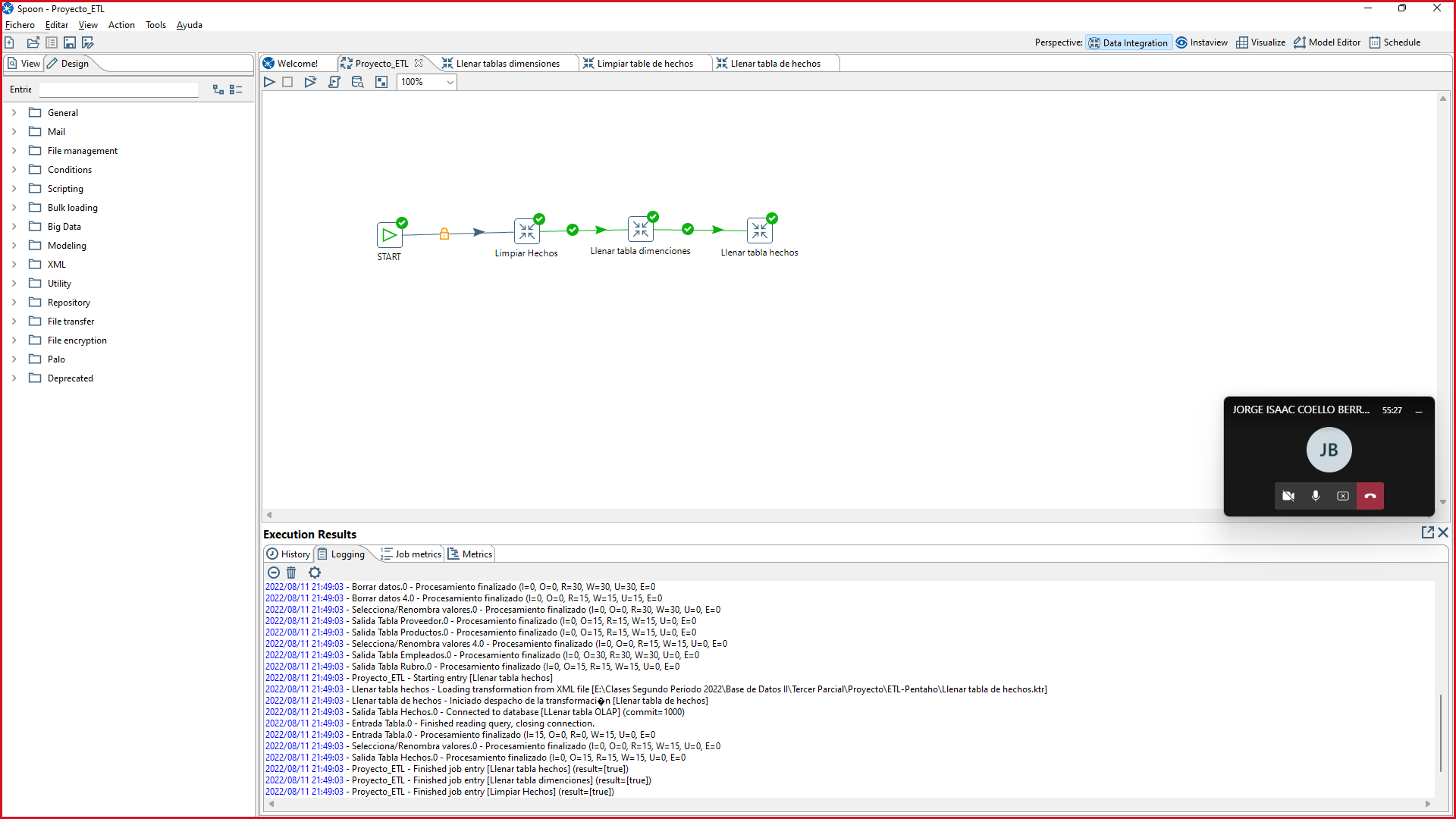
INNER JOIN PRODUCTOS ON ITEM.ID\_PRODUCTO = PRODUCTOS.ID\_PRODUCTO

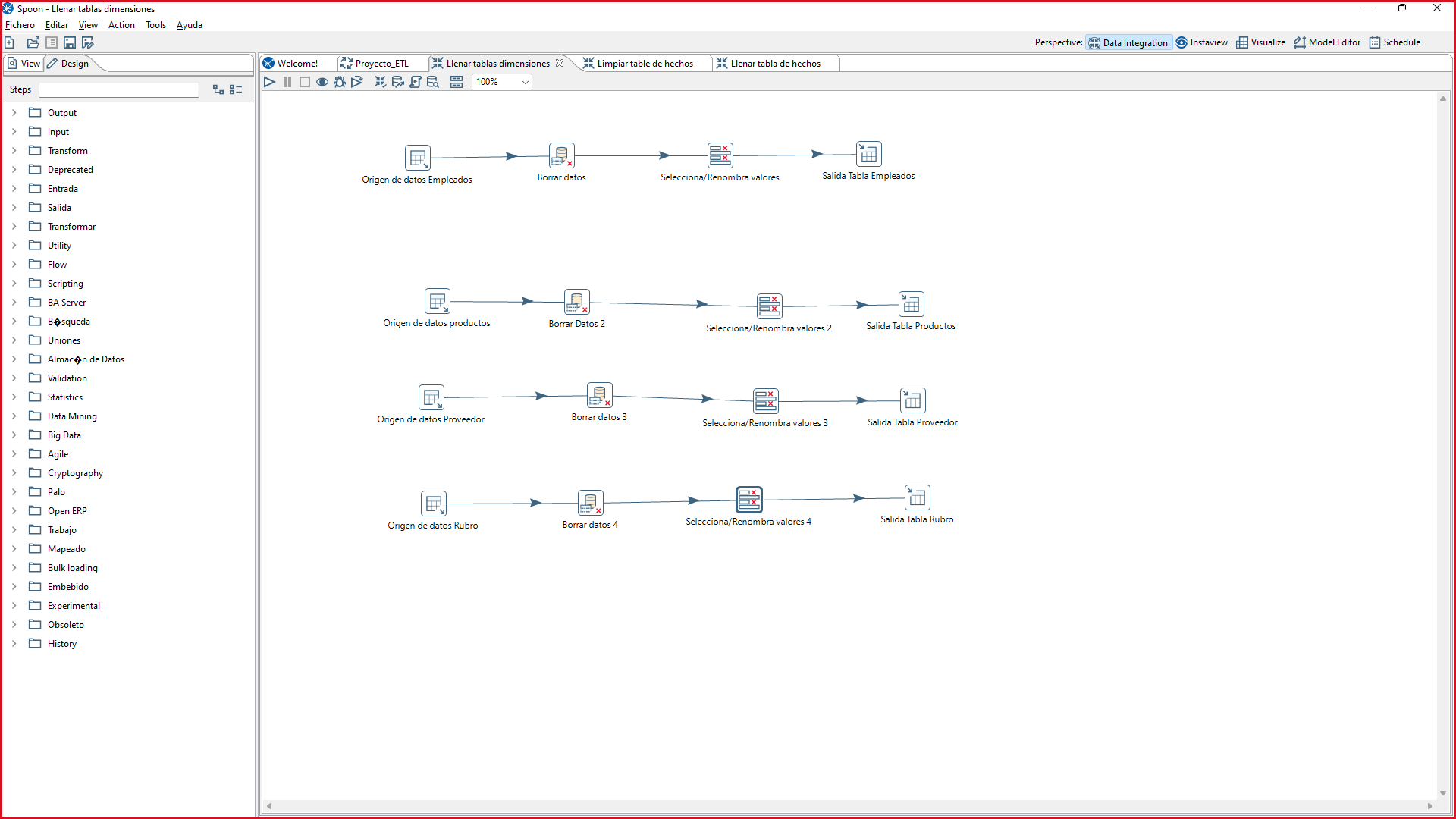
INNER JOIN RUBRO ON PROVEEDOR.ID\_RUBRO = RUBRO.ID\_RUBRO

GROUP BY ORDEN\_FECHA, EMPLEADOS.ID\_EMPLEADO, PROVEEDOR.ID\_PROVEEDOR, PRODUCTOS.ID\_PRODUCTO, RUBRO.ID\_RUBRO

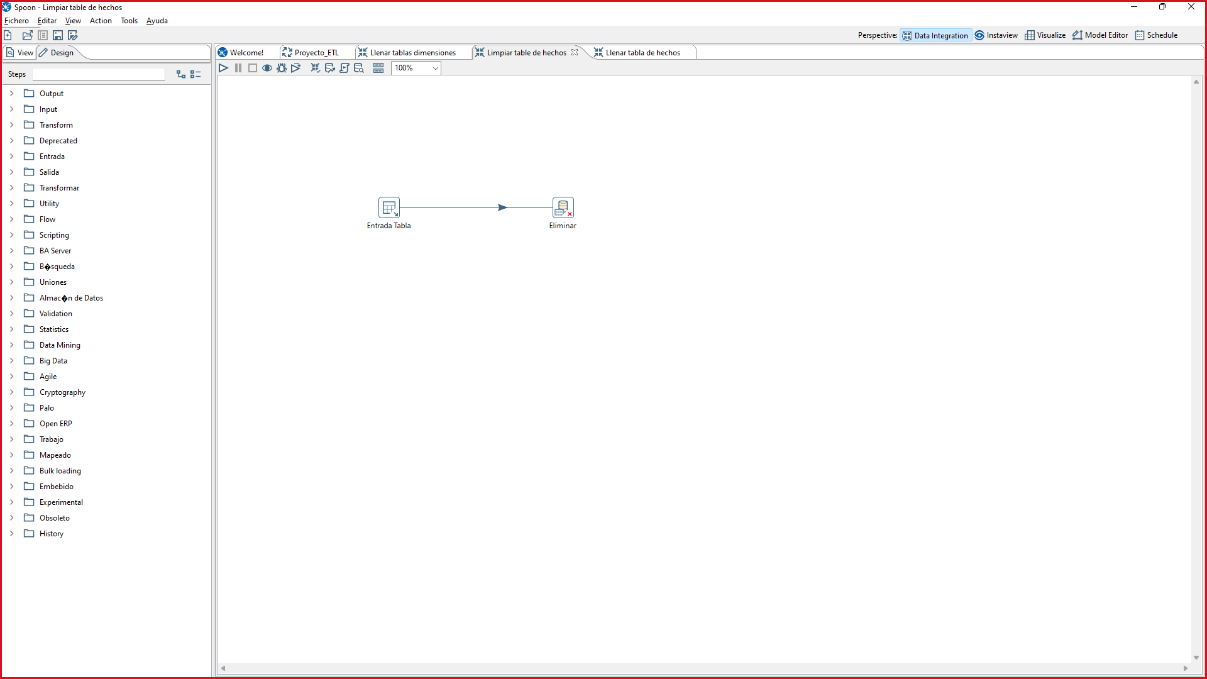
Por tanto, se hace la selección de todos los campos en la tabla Hechos\_Venta, donde el código es la fecha en la que se realizó la venta. Y se irá agrupando de acuerdo a cada empleado, proveedor, producto y rubro; para que se nos haga la suma de todas las ventas que realizó cada empleado.

# ETL Pentaho

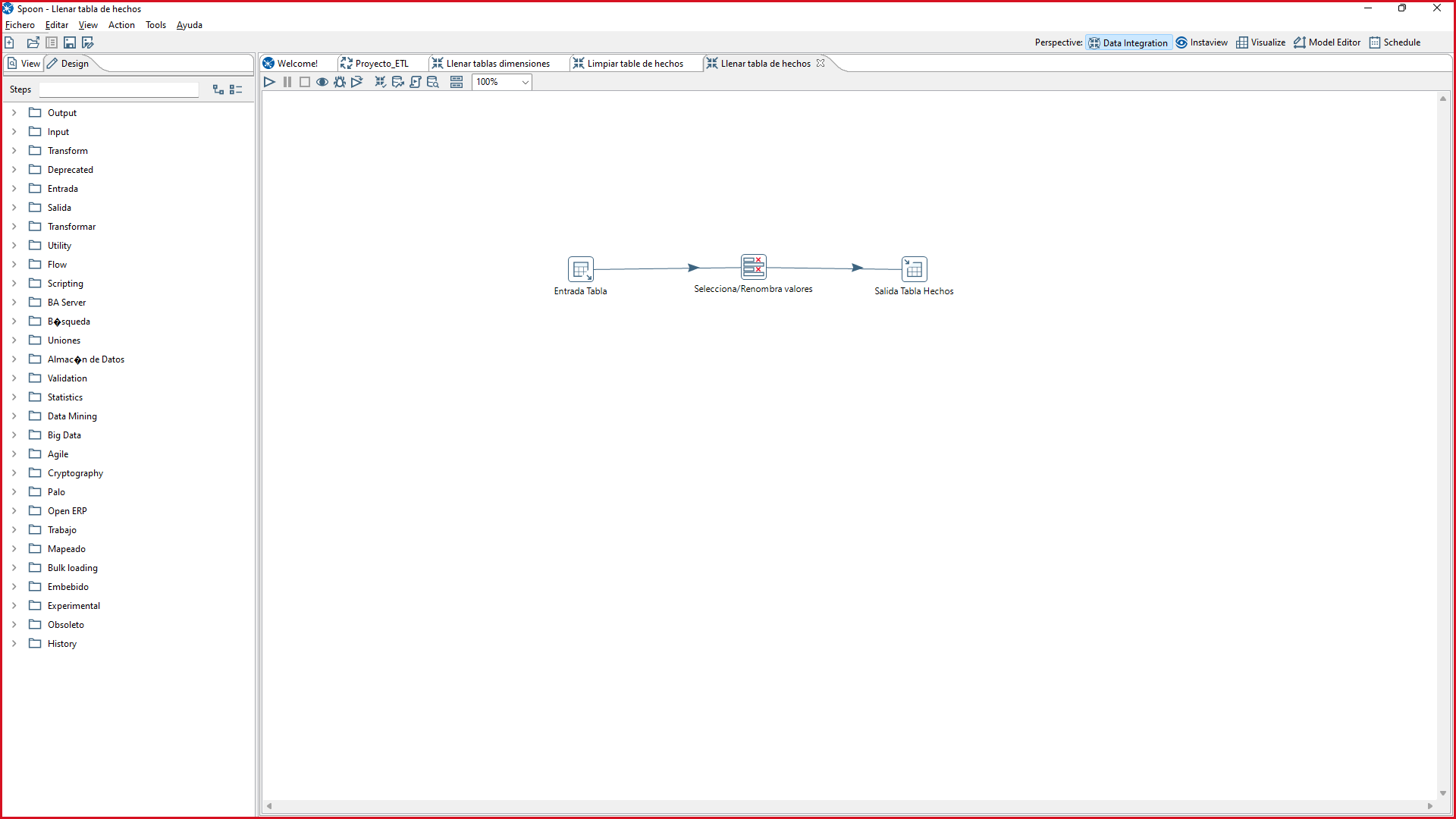
* ETL final
* Llenado Tabla Dimensiones



* Limpiar tabla hechos

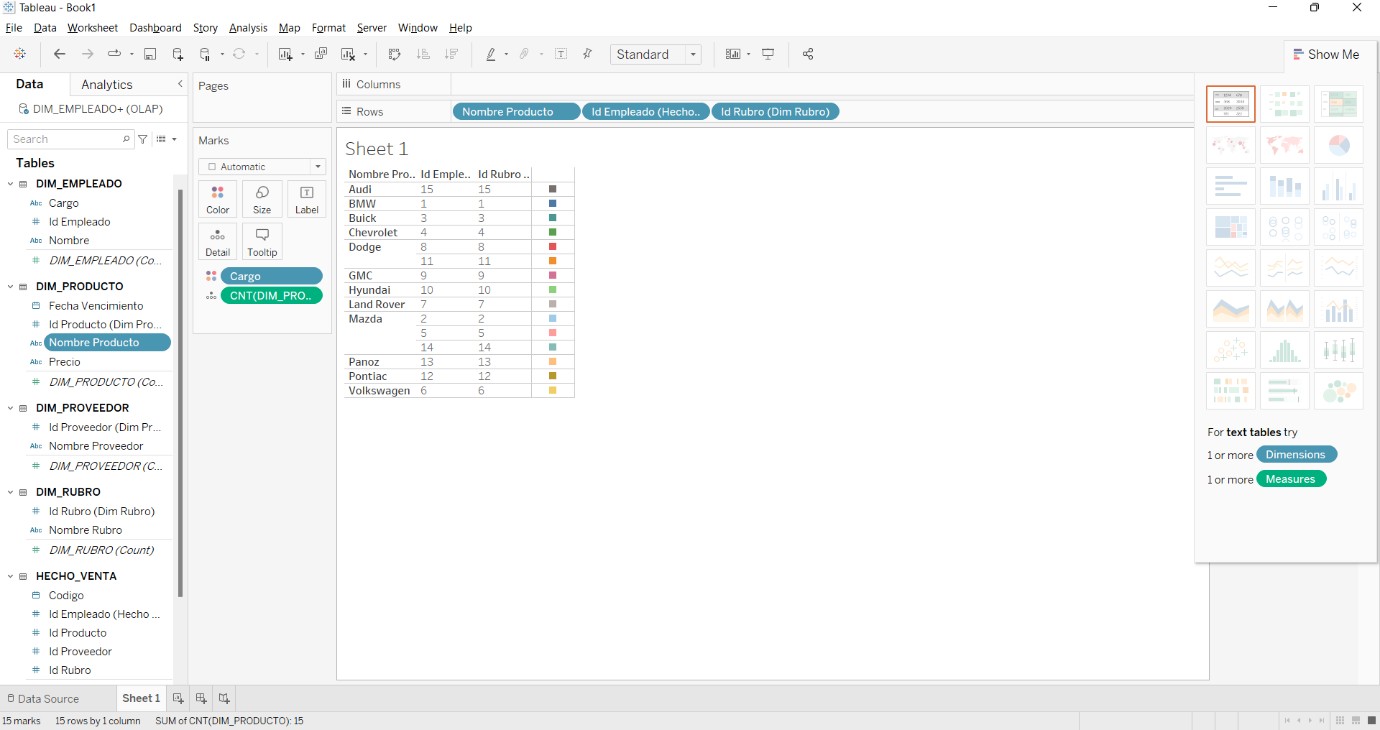


* Limpiar tablas de dimensión



# Reportes Tableau

Reporte con un color según el modelo del vehículo



REPORTE del empelado con mayores ganancias Gráfico

Descripción generada automáticamente

Verificar a que RUBRO pertenece.

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

Precio con el ID\_PRODUCTOImagen que contiene Aplicación

Descripción generada automáticamente

Año que se vence por Producto

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Recuento de VENRTAS:Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

PRECIO de cada producto.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

PRECIO de primeros 6Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**PRECIO de los PRODUCTOS según el RUBRO.**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Fecha de vencimiento según el RUBRO.Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

# Conclusiones

* El uso de ETL facilita en gran manera la administración de datos en esquemas los cuales dependen del uso constante de transferencia de datos y esto ayuda en el ámbito transaccional de las empresas.
* Podemos decir que los ETL son proceso que organizan los flujos datos entre distintos sistemas en una organización y aporta los métodos y herramientas necesarias para movilizar los datos desde múltiples fuentes a un almacén de datos, para transformarlos, limpiarlos y cargarlos en otra base de datos.
* El modelo de estrella y copo de nieve en las bases de datos sirve para tener un mejor control de la información, de acuerdo con las preguntas que quiere responder la empresa.
* Las bases de datos relacionales nos ayudan a tener un buen control de cómo va la relación entre cada tabla. Por tanto, al haber realizado bien este paso... Llegaremos a lo más parecido a la realidad.
* La forma en la que Tableau nos muestra los reportes, nos ayuda a que podamos tener esa información de manera más interactiva, evitando hacer trabajos más tardados y de menor calidad.
* Tableau se trata de un programa muy completo y especialmente diseñado para facilitar el almacenamiento de datos e información en las empresas y organizaciones.
* Hace que sea más fácil explorar y administrar los datos. Asimismo, permite descubrir y compartir información más rápidamente a fin de generar grandes cambios en los negocios y en el mundo

# 

# Recomendaciones

* Al momento de crear un ETL las tablas deben tener ID diferentes o llaves primarias con registros diferentes, ya que al momento de ejecutar el ETL al tener alguna repetida puede generar errores y que el mismo no se ejecute completamente.
* Lo primordial es que nuestros registros lleven un formato acorde a nuestro gestor de base de datos ya sea fechas, decimales o de tipo booleanos. Ya que el ETL al ejecutarse puede tener un error de lectura de datos.
* Tener en cuenta el tipo de datos con el cual se van a crear los registros de las tablas en las bases de datos, porque si se pone un dato diferente al tipo que se escogió dará error.
* Analizar bien cómo es el modelo que se quiere seguir, ya sea estrella o copo de nieve. Y así se sabrá cuál es el más efectivo de acuerdo con las preguntas brindadas por la empresa
* Yo recomiendo este software por que brinda modelos de gobernanza y seguridad de grado empresarial. De esta manera, podrá mantener sus datos al alcance de las personas adecuadas en todo momento, incluso al expandir el análisis en toda la organización. Además, la administración es simple y eficaz.

# Bibliografía

* Ventajas en la integración de datos con la herramienta ETL Pentaho, Power Data, 3 junio 2017, <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/ventajas-en-la-integraci-n-de-datos-con-la-herramienta-etl-pentaho>
* ANALYTICS Tableau Software, Neteris, <https://neteris.com/software/tableau-software-visualizacion-datos/>