



UNAH

UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS

DEPARTAMENTO DE
INGENIERÍA EN SISTEMAS

Bases de Datos II
Proyecto

Catedrático: Emilson Omar Acosta Girón



UNAH

UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS

IS-601, Base de Datos II Primer Periodo 2021

Informe del Proyecto

Elaborado por:

Cuenta	Nombre
20171033802	Edgar Josué Benedetto Godoy

Catedrático: Emilson Omar Acosta Girón



Índice

Índice	2
Introducción	3
Marco teórico	4
Data Warehouse y Data Mart	4
Paradigmas del Data Warehouse	4
Características de un Data Mart	5
Tablas de dimensiones, tabla de hechos y métricas	5
Cubos OLAP y Proceso ETL - Extracción, Transformación y Carga de datos	5
Transformación digital de las empresas en Honduras	6
Productividad en el área de mantenimiento de empresas con plantas industriales en Honduras	7
Capturas de pantalla del proceso de desarrollo del ETL	9
Capturas de pantalla con los pasos para crear el cubo OLAP	23
Conclusiones	37
Recomendaciones	38
Preguntas del negocio utilizadas	35
Explicación de métricas utilizadas	35
Bibliografía	39



Introducción

El desarrollo de un Data Warehouse junto a los Data Mart de las áreas específicas de una empresa es un paso adelante a la centralización de los datos y, por lo tanto, es un avance hacia la optimización de los procesos que se ejecutan dentro de una empresa. Un **Data Warehouse** es un almacén de grandes cantidades de información confiable, oportuna y concisa para realizar análisis puntuales y así mejorar la toma de decisiones dentro de la empresa. Un **Data Mart** cumple la función de orientar la información hacia un aspecto concreto del negocio, de esta forma se convierte en una vista del Data Warehouse.

En la actualidad las empresas punteras en tecnología utilizan el concepto de Data Warehouse para mantener optimizadas las tareas que se realizan en las distintas áreas de la empresa, de igual forma, hay varias empresas que, por su tamaño, les lleva más tiempo y dinero una migración a un sistema de información centralizada.

La **inteligencia de negocios** es una disciplina orientada a la optimización de las tareas ejecutadas por una empresa, esto involucra el análisis de datos de esta. En la actualidad ya existen empresas en Honduras que ofrecen servicios de Inteligencia de negocios, en su mayoría las más confiables son las internacionales, dado que es una estrategia aun en auge y Honduras es un país caracterizado por la corrupción desde el puesto más bajo hasta los altos cargos, algunas empresas que proveen de estos servicios son Go Consultores, BI Dss Technology e IT4 Honduras entre otros.

El Data Warehouse tiene la característica de proveer información que responde a preguntas de negocio específicas las cuales proveen un desglose del **estado actual de la empresa**, por lo tanto, para empresas que no tienen estos conceptos implementados el primer uso de estos es crucial para iniciar con la implementación de servicios de Inteligencia de negocios y automatización de proceso. Muchas empresas grandes en Honduras no han apostado aún por implementar servicios de inteligencia de negocios y automatización de procesos, teniendo aún entre sus filas personal encargado de hacer tareas repetitivas y con altas probabilidades de fallas o errores. Empresas como Cadeca o Zigma alimentos tienen hasta hace muy poco tiempo entre sus filas personal de logística, con lo cual se puede asumir que el proceso de migración a los procesos automatizados viene empezando.

En este caso, el Data Warehouse desarrollado está orientado a una empresa encargada de la producción y manufactura de productos congelados, que posea una planta de producción lo suficientemente grande como para que aún en 2021 no esté implementado el concepto de Inteligencia de negocios como ser Cadeca o Zigma Alimentos. Para este primer acercamiento el área que de la cual se desarrolló un Data Mart es la de mantenimiento de la maquinaria de producción. El desarrollo parte de una base de datos **OLTP** (Online Transaction Processing) debidamente normalizada, para de ella obtener la base de datos **OLAP** (Online Analytical Processing) utilizando el modelo estrella, luego el desarrollo del **ETL** (Extract, Transform and Load) y por último la generación de reportes haciendo uso de los conceptos de Drill-Up y Drill-Down, los cuales se definen como la presentación de información de lo más específico a lo más general y viceversa respectivamente.



Marco teórico

Data Warehouse y Data Mart

Según Nettleton D. (2003), en su libro Análisis de datos comerciales, un Data Warehouse es un compacto, sólido y muy completo almacén de datos concebido de una muy variada de fuente de información. Los datos en un Data Warehouse se distinguen de los de un Data Mart en los entornos operacionales, dado que los primeros suelen ser **datos resumidos** (informes, agregaciones, etc.) mientras que los últimos son **datos de uso diario** (transacciones de cuentas, registros de clientes, etc.).

Según Curto J. & Conesa J. (2010), en su libro "Introducción al Business Intelligence", consideran que un Data Mart: "es un subconjunto de los datos del Data Warehouse cuyo objetivo es responder a un terminado análisis, función o necesidad, con una población de usuarios específica. Al igual que en un data warehouse, los datos están estructurados en **modelos de estrella o copo de nieve**, y un Data Mart puede ser dependiente o independiente de un data warehouse." Nos dice que una posible utilización sería para el Data mining o marketing. El Data Mart está más orientado para un grupo de trabajo o área determinada en de la organización.

El Data Warehouse puede estar compuesto de datos provenientes de distintas fuentes, es decir, distintos gestores de bases de datos como MySQL, MariaDB, SQLite, PostgreSQL, SQL Server, etc. Existen dos paradigmas en la construcción de los Data Warehouse que se orientan en dar la misma solución, pero variando en el concepto de ejecución.

Paradigmas del Data Warehouse

1. **Paradigma Bill Inmon:** Entre sus características principales están, la no volatilidad de los datos, es decir no podrá ser eliminada ni modificada, estos datos deben ser variantes en el tiempo, debe mantener los datos alineados de manera que los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí y por último los datos tienen que ser constantes entre todos los sistemas gestores de bases de datos. El enfoque Inmon también se refiere como **Top-down** porque la información ha de estar a los máximos niveles de detalle. El más utilizado para este tipo de sistema es por ejemplo el **modelo Entidad Relación**.
2. **Paradigma Ralph Kimball:** (Echandi, Porras, & Sandi, 2015). Ralph Kimball es considerado el inventor del Modelo Dimensional y pionero en Data Warehouse e Inteligencia de Negocios. En este paradigma hay **tablas de dimensiones** y **tablas de hechos** también llamadas estas últimas con **facts** tables interconectadas por el **bus structure**. Este tipo de enfocar se conoce como **Bottom-up**, es la unión de diferentes Data Marts que conforma un Data Warehouse de manera estructurada de una forma que se usa un Data Mart como el primer elemento para el análisis, y luego se va incrementando algunos otros que tengan y puedan compartir las dimensiones o inclusive nuevas dimensiones.



Características de un Data Mart

Considerados vistas de los Data Warehouse, enfocados en un área en particular del negocio, tienen características únicas estructuralmente y optimizadas para análisis de datos como:

1. Reducido volumen de datos.
2. Incremento en rapidez de la consulta.
3. Validación rápida de la información.
4. Fácil archivamiento histórico de los datos.

Tablas de dimensiones, tabla de hechos y métricas

Conesa, J., Curto, J. (2011). Introducción al Business Intelligence. Se Puede llamar hechos a los indicadores de negocio. Así pues, por ejemplos las ventas, el stock, las compras, entre otros. Es de suma importancia que estas tablas de hecho posean un alto nivel de detalle de lo que se almacenará, además de incluir las claves de las dimensiones que dan exactamente el nivel de detalle, también sus indicadores. Conesa y Curto definen diferentes tipos de tablas de hecho:

- **Transaction Fact Table:** Permite analizar los datos con mayor detalle.
- **Factless Fact Tables/Coverage Table:** Tablas que carecen de medidas o métricas, adquieren sentido ya que representan el hecho que va a suceder. Normalmente se suele adicionar contadores a estas tablas.
- **Periodic Snapshot Fact Table:** Tablas que se usan para recopilar información cada cierto tiempo, utilizando el mismo concepto de los **Jobs**.
- **Accumulating Snapshot Fact Table:** Muestran el ciclo de vida, de una actividad o un proceso y son multidimensionales.

Las tablas de dimensiones son aquellas que representa la información a través de la cual se desean medir los hechos, también conocida como una colección de miembros, unidades o individuos del mismo tipo. Ejemplo: tiempo, regiones, productos, modelos, etc.

Las Medidas o métricas son atributos numéricos de un hecho que representan el comportamiento del negocio relativo a una dimensión.

Cubos OLAP y Proceso ETL - Extracción, Transformación y Carga de datos

Josep Curto 2007, en su libro Introducción al Business Intelligence. Enseña que el proceso de Extracción, limpieza y carga es una manera de integrar los datos, extraer los datos de la misma organización de origen, transformarlos acorde a lo que necesitamos en de organización, para luego cargar estos datos en los entornos de destino. Los **orígenes** son casi siempre **Bases de datos**, así como también archivos de **hojas de cálculo y texto**. Herramientas de extracción, transformación y carga sirven para mover datos entres orígenes y destinos adicionalmente documentan la transformación de estos datos. Existen funciones avanzadas como:

- Orígenes de datos adicionales: DB, XML, servicios de web entre otros.
- Mejor transformación de datos, usando el SQL y scripts en distintos lenguajes
- Inclusión con productos de gestión de calidad de datos.



Los cubos OLAP proporcionan una presentación multidimensional de los datos de un almacén mediante la creación de cubos que organizan y resumen los datos para mejorar la eficiencia de las consultas analíticas. El diseño de la estructura del almacén de datos puede afectar a la facilidad con la que se podrán diseñar y construir estos cubos, permiten visualizar la información desde distintas perspectivas, y de acuerdo con la perspectiva que se tenga del cubo la información se puede interpretar de una u otra forma.

Transformación digital de las empresas en Honduras

La transformación digital para empresas pequeñas y multinacionales en Honduras ha evidenciado una carencia en el avance tecnológico por reportar tiempos prolongados a la hora de implementar cambios estructurales incluyendo soluciones tecnológicas automatizables. La transformación digital para las Pymes es un proceso que puede llevar a la correcta ejecución de sus operaciones eliminando los procesos ilícitos e incrementando la productividad, incluir tecnologías de la información en los procesos de un negocio es importante, ayudando a la competitividad en el mercado ampliando las fronteras en el mundo digital. No quiere decir que a una empresa que siembra vegetales despidan a sus jornaleros para implementar Drones. Existen métodos para medir el nivel de tecnología que se debe implementar en un negocio para ajustarse a las necesidades y el presupuesto que se desee aportar.

Los constantes casos de corrupción, los manejos de instituciones privadas y públicas con procesos de negocio ilícitos o sospechosos son factores claves a la hora de analizar el tiempo que les toma a las empresas empezar a adaptarse a los cambios tecnológicos, ejemplos claros de esto se encuentran en la banca donde transnacionales como BAC Credomatic es un competidor tecnológicamente superior a Banco Atlántida y otros competidores como Ficohsa son los que se encuentran justo detrás implementando planes de negocio eficientes, ampliando el uso de las tecnologías de la información en cada una de sus áreas, permitiéndoles ser competidores reales y presentado un panorama alentador para aquellas empresas que siguen sus pasos a la hora de posicionarse tecnológicamente en su mercado correspondiente.

Las evaluaciones comunes que se hacen para determinar el nivel de madurez de un negocio para asumir los retos que involucra la transformación tecnológica digital se pueden resumir en 3 vectores principales.

- Modelo organizacional
- Gestión de la innovación
- Capacidades digitales

Los planes de transformación digital de un negocio deben atender a las necesidades de este, limitándose inicialmente a entender ciertos factores clave como:

- Implementar tecnología no por sí sola no quiere decir que se esté transformando a la empresa por completo.
- Cómo la tecnología ayuda a hacer más ágil y agregar valor al cliente final.



- Cómo mejorar el proceso de comprar para transformar servicios y productos en un bien fácil de obtener en línea.
- Si no se dispone el precio de los productos o servicios en la web de forma automatizada con actualizaciones en tiempo real no se debe pretender iniciar con IoT (Internet of Things).

Las empresas han evidenciado una serie de pasos que se pueden seguir para generar un traslado de lo cotidianamente hecho en papel a empezar a automatizar procesos de forma orgánica, permitiendo a sus participantes ser parte de un proyecto que consta de etapas caracterizadas en desarrollar área por área el uso de las tecnologías de la información en la empresa. Los pasos por seguir son:

1. Analizar.
2. Crear un cambio cultural.
3. Aplicar tecnología en los procesos internos.
4. Aplicar tecnología en los procesos relacionados con la atención al cliente.
5. Realizar mediciones estadísticas usando la información a partir de los datos recopilados.
6. Mejorar planes a partir de mediciones realizadas.

Como organización siempre se dispone de datos, parte de lo que se buscan es convertir datos en información, información en estrategia y estrategia en resultados tangibles que permitan a la organización seguir creciendo. La transformación digital para empresas PYMES es una catapulta que permite poder tener un mejor alcance, ser más ágil y poder competir con los competidores multinacionales.

Productividad en el área de mantenimiento de empresas con plantas industriales en Honduras

Beltrán J. (1998), en su obra Indicadores de Gestión, dice que básicamente que la productividad es la relación que hay entre producción y RRHH. También ampliando el significado podemos decir que la productividad cuantitativamente hablando es la relación entre la cantidad producida y la cantidad de recursos empleados en la producción.

Por otra parte, Niebel B. (2007) en su libro Motion and time study, refiere que el aumento de recursos administrados a RRHH por hora de trabajo o tiempo gastado da como resultado el aumento de la productividad. Estos RRHH son la base para el incremento de la productividad ya que se considera como el recurso más relevante de cualquier organización. También nos dice que el capital es un recurso muy importante para el desarrollo industrial y que la ayuda del factor tecnológico hace que incremente, es importante saber que estos recursos, como el capital y la tecnología, pueden llegar a ser desperdiciados por las personas que no estén dispuestas a utilizarlas adecuadamente. La gestión del personal se convierte en el eje principal del manejo de tiempos productivos en las operaciones en plantas productoras. Algunos indicadores de



productividad que pueden aplicarse al área de mantenimiento de máquinas de producción de plantas industriales que posean cámaras frías son:

- 1- Utilización de Máquinas: Numero de maquinas en funcionamiento / Total de maquinas
- 2- Problemas disponibles: Total de horas de problemas mecánicos y eléctricos / Total de horas registradas.

En los sistemas de producción también hay indicadores de productividad que pueden ayudar al desarrollo o adquisición de herramientas informáticas que permitan optimizar la toma de decisiones incluyendo conceptos como el Data Warehouse, para poder centralizar los datos que se obtienen y de esta forma acelerar la toma de decisiones cruciales que aumenten la productividad en el área de mantenimiento y de esta forma secuencialmente aumentar la productividad en el área de producción permitiendo su funcionamiento prolongado.

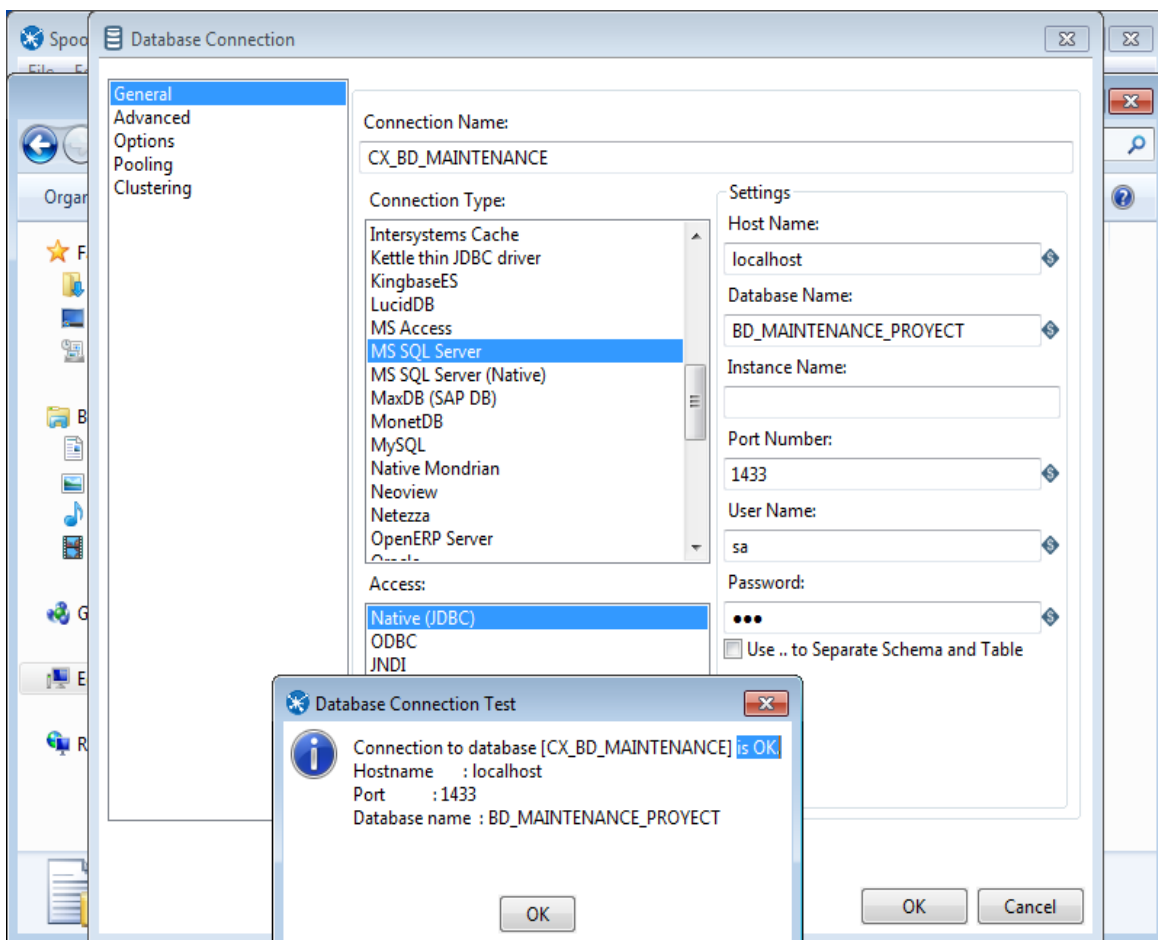
Salazar B. (2016) en su publicación Indicadores De Los Sistemas De Producción. Se enseña que tener indicadores de gestión en una aplicación computacional es sumamente relevante para la ejecución de procesos productivos. La productividad es igual que decir que un sistema de producción es eficiente, o cociente entre el resultado del sistema productivo (Ventas, productos) y la cantidad de recursos utilizados; En el interior de un sistema de producción hay índices de productividad, tantos como existan recursos, ya que estos pueden operar como indicadores de gestión tradicionales.

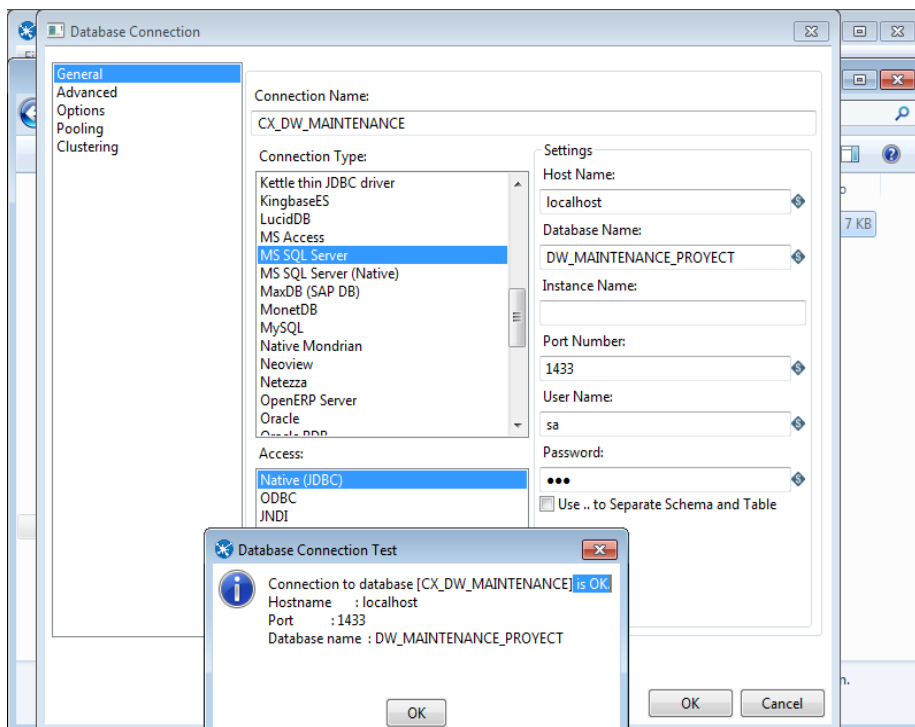
Jiménez Á. & Delgado E. & Gaona G. (2001) en su obra Modelo de productividad de David Sumanth plantea un modelo de productividad, con el objeto de dar a las empresas pequeñas y medianas una aplicación que les deje establecer como se encuentra en la actualidad esta productividad, mostrando aquellos insumos que son palpables o medibles en la organización son los más relevantes y debemos optimizar las estrategias de incremento de la productividad. Se sustenta basándose en la teórica del Modelo de Productividad Total Operativo (MPT) el cual fue creado por David J. Sumanth y se basa en cuatro principales frases que: Medición, Evaluación, Planeación y Mejoramiento, que básicamente son Ciclo de Mejoramiento de la Productividad.

1. **Medición:** Elaborar una manera de cuantificar, que se base en los inputs y outputs que muestran el valor de la producción y los insumos utilizados indicándonos la variación, que también pueden ser medibles directamente.
2. **Evaluación:** Esta es una faceta que se encuentra entre medición y la planeación, desarrollando maneras metodológicas para analizar la productividad en dos tiempos diferentes presupuestados para luego compararlos con tiempos reales.
3. **Planeación:** Fase analítica que comienza desde un punto futuro en el cual se establecen objetivos, medidos en metas, desarrollándose líneas de acción para llegar a estos objetivos con muy definidas direcciones de acción.
4. **Mejoramiento:** Iniciadas por la organización y con base en la Planeación, buscando que sea funcional, se fija en la eficiencia y la eficacia del trabajador, implementado estrategias para hacerlo viable.

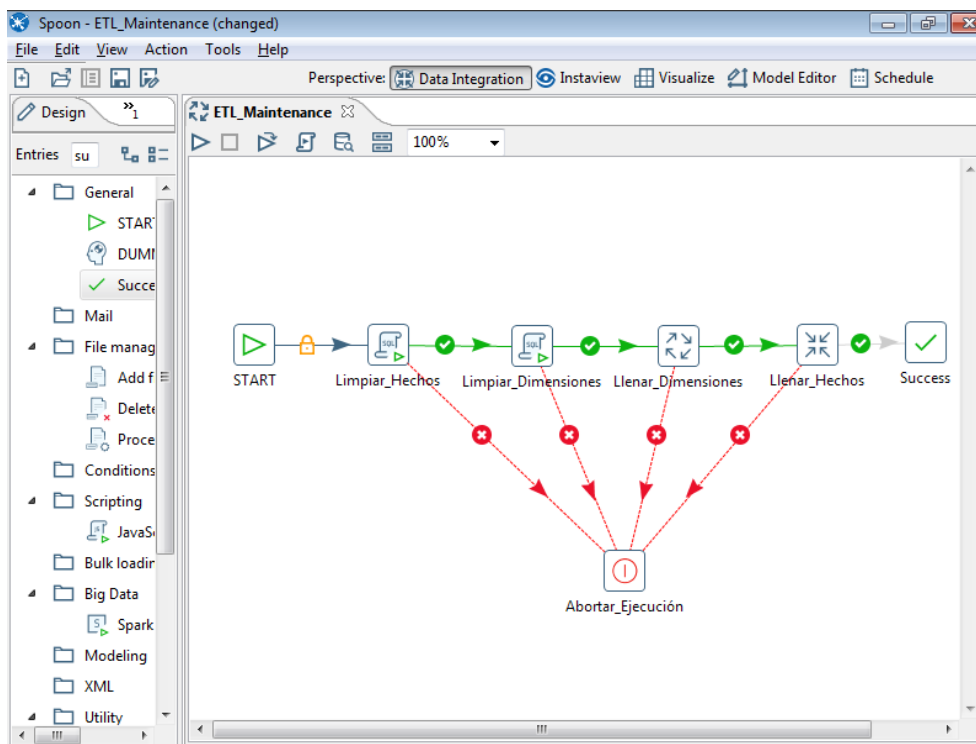
Capturas de pantalla del proceso de desarrollo del ETL

- 1- Crear las conexiones con las bases de datos OLTP y OLAP.





- 2- Creación del flujo de datos principal asignado un nombre descriptivo a cada uno de los procesos del flujo de datos.





- 3- Definir las consultas SQL para limpiar las tablas de dimensiones y la tabla de Hechos

Execute SQL Script ...

Job entry name: Limpiar_Hechos

Connection: CX_DW_MAINTENANCE [Edit...] [New...] [Wizard...]

SQL from file: ☐

SQL filename: [Browse...]

Send SQL as single statement? ☐

Use variable substitution? ☐

SQL Script:

```
DELETE FROM FactMaintenances;
```

Line 1 Column 0

[?] Help [OK] [Cancel]

Execute SQL Script ...

Job entry name: Limpiar_Hechos

Connection: CX_DW_MAINTENANCE [Edit...] [New...] [Wizard...]

SQL from file: ☐

SQL filename: [Browse...]

Send SQL as single statement? ☐

Use variable substitution? ☐

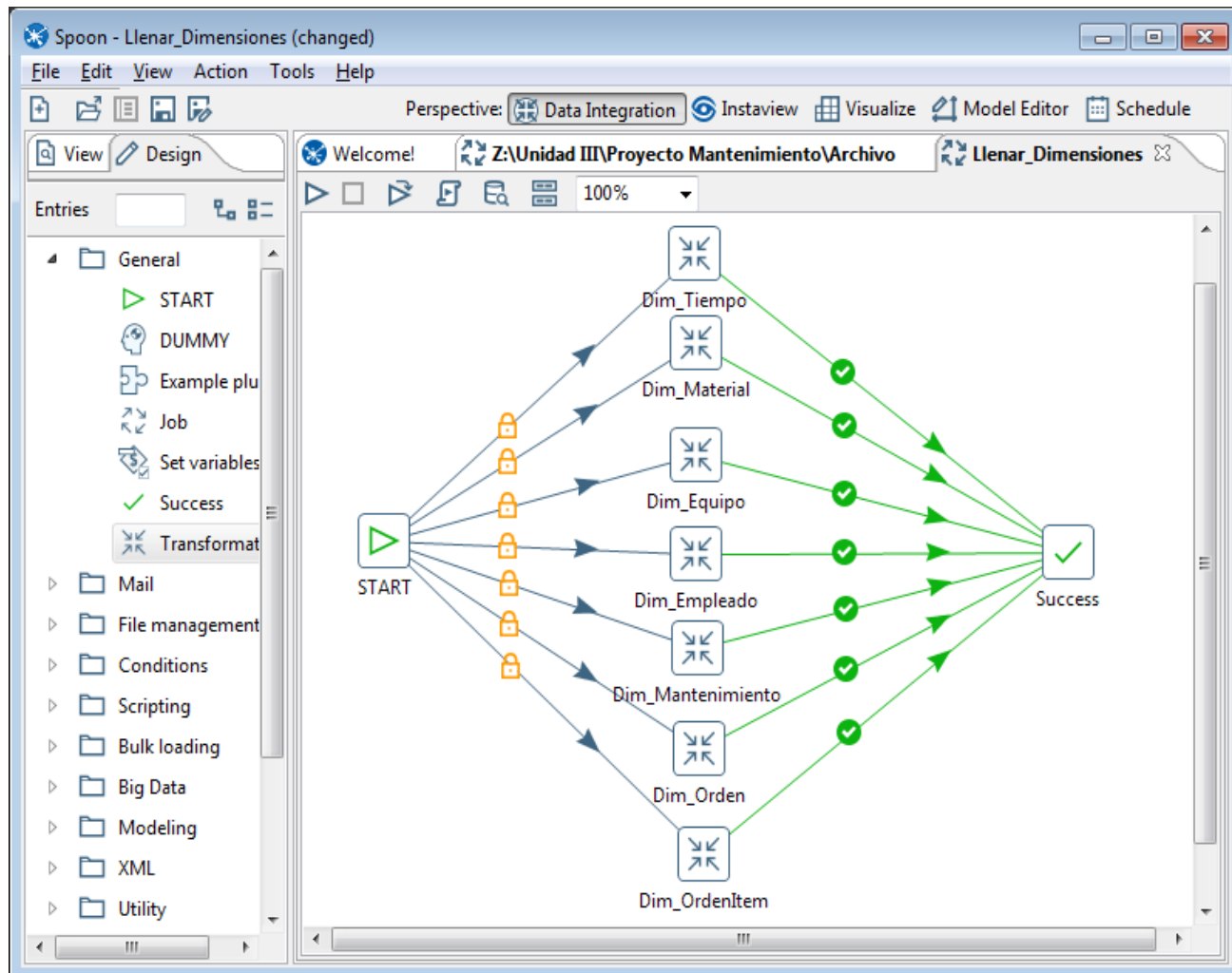
SQL Script:

```
DELETE FROM FactMaintenances;
```

Line 1 Column 0

[?] Help [OK] [Cancel]

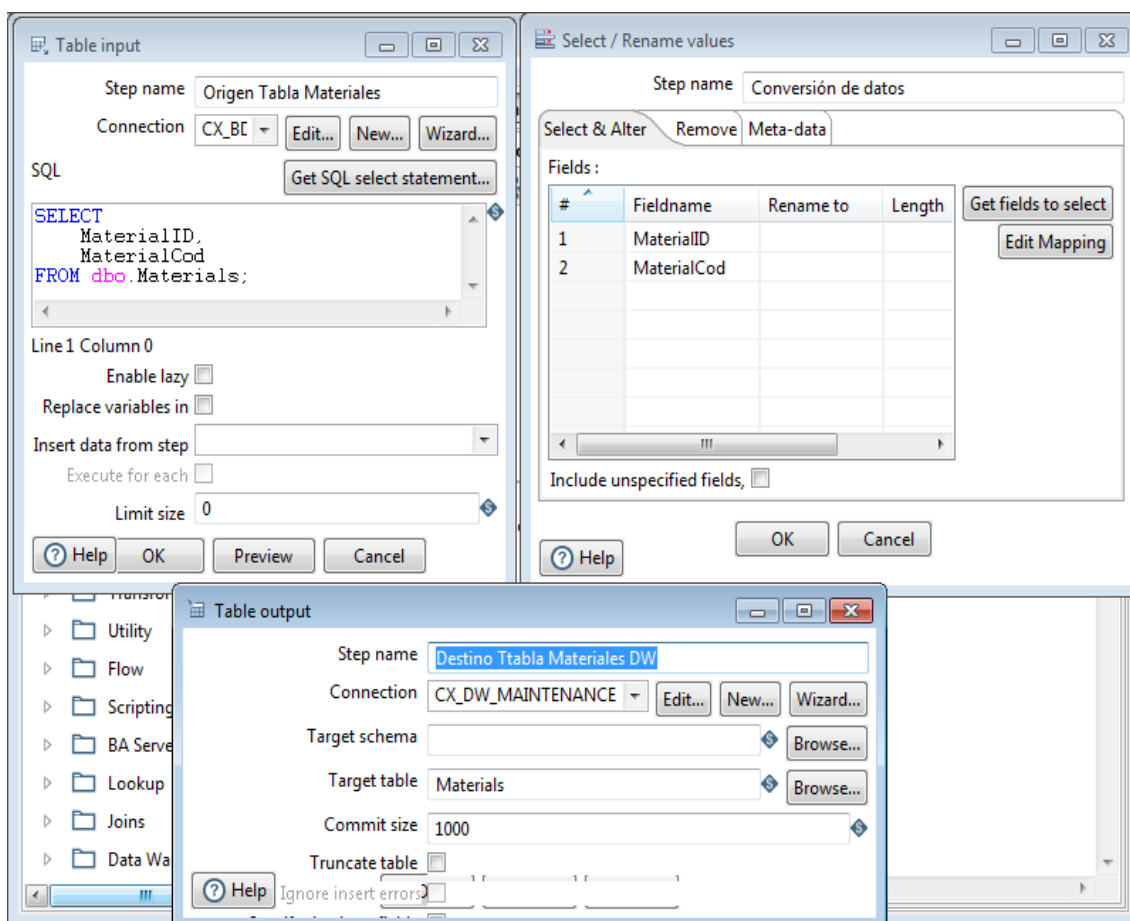
4- Creación del flujo de datos de la carga en las tablas de dimensiones.



5- Definir el origen, la transformación y el destino de los datos para cada una de las dimensiones.

a) Dimensión Material:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)





b) Dimensión Equipo:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

The image displays three screenshots of the SQL Server Enterprise Manager Data Transformation Services (DTS) task wizard, illustrating the steps for creating a dimension table.

Table input: This window shows the 'Step name' as 'Origen Tabla Equipos'. The 'Connection' is set to 'CX_BD_MAI'. The 'SQL' field contains the following query:

```
SELECT
    EquipmentID,
    EquipmentCod,
    Name
FROM dbo.Equipments
```

The 'Line 1 Column 0' section shows options for 'Enable lazy conversion', 'Replace variables in', 'Insert data from step', 'Execute for each row?', and 'Limit size' (set to 0). Buttons for 'Help', 'OK', 'Preview', and 'Cancel' are visible.

Select / Rename values: This window shows the 'Step name' as 'Conversión de datos'. The 'Fields' table lists the selected fields:

#	Fieldname	Rename to
1	EquipmentID	
2	EquipmentCod	
3	Name	

Buttons for 'Get fields to select', 'Edit Mapping', 'Include unspecified', 'OK', and 'Cancel' are present.

Table output: This window shows the 'Step name' as 'Destino Tabla Equipos DW'. The 'Connection' is set to 'CX_DW_MAINTENANCE'. The 'Target schema' is empty, and the 'Target table' is 'Equipments'. The 'Commit size' is set to 1000. The 'Truncate table' checkbox is checked. Buttons for 'Help', 'OK', 'Cancel', and 'Preview' are visible.

c) Dimensión Empleado:

1. Origen de datos (Input Table)
2. Concatenar Nombre y Apellido (Concat Fields)
3. Conversión de datos (Select Values)
4. Destino de datos (Output Table)

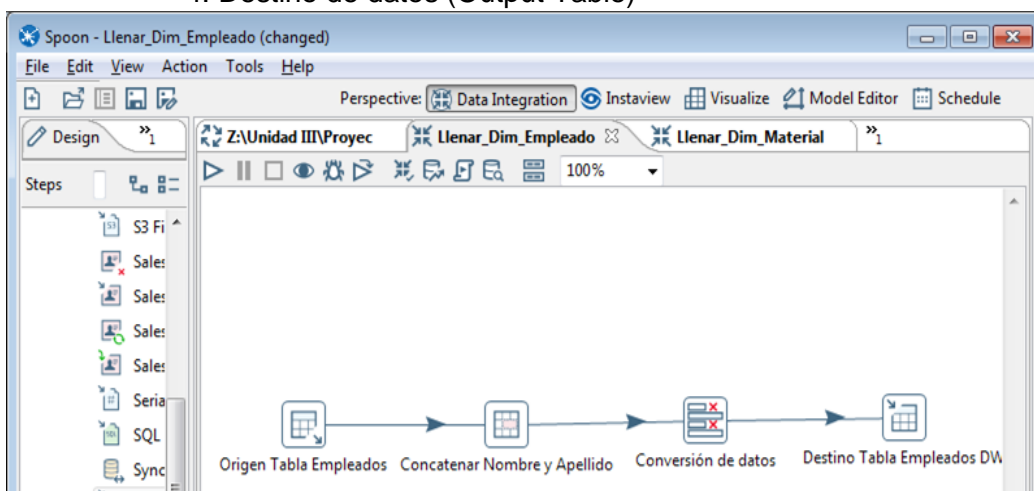


Table input

Step name: Origen Tabla Empleados

Connection: [Edit...](#) [New...](#) [Wizard...](#)

SQL: [Get SQL select statement...](#)

```
SELECT
EmployeeID,
FirstName,
LastName,
Salary
FROM dbo.Employees;
```

Line 1 Column 0

Enable lazy: ☐

Replace: ☐

Insert data:

Execute for: ☐

Limit size: 0

[Help](#) [Preview](#) [Cancel](#)

Concat Fields

Step name: Concatenar Nombre y Apellido

Target Field: WholeName

Length of: 0

Separator: [Insert TAB](#)

Enclosure:

Fields

#	Name	Type
1	FirstName	String
2	LastName	String

[Get Fields](#) [Minimal width](#)

[Help](#) [OK](#) [Cancel](#)

Select / Rename values

Step name: Conversión de datos

Select & Alter [Remove](#) [Meta-data](#)

Fields:

#	Fieldname
1	EmployeeID
2	WholeName
3	Salary

[Get fields to select](#) [Edit Mapping](#)

[Help](#) [OK](#) [Cancel](#)

Table output

Step name: Destino Tabla Empleados DW

Connection: CX_DW_MAINTENANCE [Edit...](#) [New...](#) [Wizard...](#)

Target schema: [Browse...](#)

Target table: Employees [Browse...](#)

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐



d) Dimensión Mantenimiento:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

The image displays three screenshots of the SQL Server Data Tools (SSDT) wizards used for creating a data dimension.

Table input wizard: The 'Step name' is 'Origen Tabla Mantenimientos'. The 'Connection' is 'CX_BD_I'. The 'SQL' field contains the following query:

```
SELECT
MaintenanceID,
MaintenanceType,
WorkedHours
FROM dbo.Maintenances;
```

Select / Rename values wizard: The 'Step name' is 'Conversión de datos'. The 'Fields' table is as follows:

#	Fieldname	Rename to
1	MaintenanceID	
2	MaintenanceType	
3	WorkedHours	

Table output wizard: The 'Step name' is 'Destino Tabla Mantenimientos DW'. The 'Connection' is 'CX_DW_MAINTENANCE'. The 'Target schema' is empty, and the 'Target table' is 'Maintenances'. The 'Commit size' is '1000'. The 'Truncate table' checkbox is unchecked.



e) Dimensión Ordenes:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

The image displays three screenshots of SQL Server Enterprise Manager wizards, arranged in a 2x2 grid with the bottom-right cell empty.

Top Left: Table input wizard

- Step name: Origen Tabla Ordenes
- Connection: CX_BD_MA
- SQL:

```
SELECT
  OrderID,
  OrderCod,
  TransactionType,
  AccountDate
FROM dbo.Orders;
```
- Line 1 Column 0
- Enable lazy conversion: ☐
- Replace variables in: ☐
- Insert data from step:
- Execute for each row?: ☐
- Limit size: 0
- Buttons: Help, OK, Preview, Cancel

Top Right: Select / Rename values wizard

- Step name: Conversión de datos
- Tab: Select & Alter
- Fields table:

#	Fieldname	Rename to
1	OrderID	
2	OrderCod	
3	TransactionType	
4	AccountDate	

- Buttons: Get fields to select, Edit Mapping
- Include unspecified: ☐
- Buttons: Help, OK, Cancel

Bottom Left: Table output wizard

- Step name: Destino Tabla Ordenes DW
- Connection: CX_DW_MAINTENANCE
- Target schema:
- Target table: Orders
- Commit size: 1000
- Truncate table: ☐
- Buttons: Help

f) Dimensión Orden Items:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

Table input

Step name: Origen Tabla Orden Items

Connection: CX_BD_M [Edit...] [New...] [Wizard...]

SQL: [Get SQL select statement...]

```
SELECT
  OrderItemID,
  AmountLemp,
  AmountDoll,
  AmountAdd
FROM dbo.OrderItems;
```

Line 1 Column 0

Enable lazy: ☐

Replace variables in: ☐

Insert data from step:

Execute for each row?: ☐

Limit size: 0

[?] Help [OK] [Preview] [Cancel]

Select / Rename values

Step name: Conversión de datos

Select & Alter [Remove] [Meta-data]

Fields:

#	Fieldname	Rename to	Len
1	OrderItemID		
2	AmountLemp		
3	AmountDoll		
4	AmountAdd		

[Get fields to select] [Edit Mapping]

Include unspecified: ☐

[?] Help [OK] [Cancel]

Table output

Step name: Destino Tabla Orden Items DW

Connection: CX_DW_MAINTENANCE [Edit...] [New...] [Wizard...]

Target schema: [Browse...]

Target table: OrderItems [Browse...]

Commit size: 1000

Truncate table: ☐

[?] Help



g) Dimensión Tiempo:

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

The image displays three screenshots of the SQL Server Data Tools (SSDT) interface, illustrating the configuration of a data flow task for a Time Dimension.

Table input (Step 1: Origen Tabla Tiempo)

Step name: Origen Tabla Tiempo
Connection: CX_BD_MAIN

SQL:

```
SELECT
    CONVERT(DATE, AccountDate) TimeID,
    DATEPART(YEAR, AccountDate) Year,
    DATEPART(MONTH, AccountDate) Month,
    DATEPART(WEEK, AccountDate) Week,
    DATEPART(QUARTER, AccountDate) Trimester,
    DATENAME(WEEKDAY, AccountDate) DayOfWeek
FROM Orders
GROUP BY Orders.AccountDate
```

Line 1 Column 0

Enable lazy conversion ☐
Replace variables in ☐
Insert data from step
Execute for each row? ☐
Limit size 0

Buttons: ? Help, OK, Preview, Cancel

Select / Rename values (Step 2: Conversión de datos)

Step name: Conversión de datos

Fields:

#	Fieldname	Rename to
1	TimeID	
2	Year	
3	Month	
4	Week	
5	Trimester	
6	DayOfWeek	

Buttons: ? Help, OK, Cancel

Table output (Step 3: Destino Tabla Tiempo DW)

Step name: Destino Tabla Tiempo DW
Connection: CX_DW_MAINTENANCE
Target schema:
Target table: Time

Buttons: ? Help, Browse...



6- Transformación encargada de llenar la tabla de Hechos

1. Origen de datos (Input Table)
2. Conversión de datos (Select Values)
3. Destino de datos (Output Table)

Table input

Step name: Origen Tabla Hechos

Connection: CX_BD_MAINTENANCE

SQL

```
SELECT
  CONVERT(DATE, dbo.Orders.AccountDate) TimeID,|
  dbo.Materials.MaterialID,
  dbo.Equipments.EquipmentID,
  dbo.Employees.EmployeeID,
  dbo.Maintenances.MaintenanceID,
  dbo.Orders.OrderID,
  dbo.OrderItems.OrderItemID,
  (SELECT SUM(dbo.OrderItems.AmountLemp) FROM dbo.OrderItems) TotalAmountLemp,
  (SELECT SUM(dbo.OrderItems.AmountDoll) FROM dbo.OrderItems) TotalAmountDoll,
  (SELECT SUM(CASE WHEN dbo.Maintenances.MaintenanceType = 'PREVENTIVO' THEN dbo.Maintenanc
  (SELECT SUM(CASE WHEN dbo.Maintenances.MaintenanceType = 'CORRECTIVO' THEN dbo.Maintenanc
  (SELECT COUNT(CASE WHEN dbo.Maintenances.MaintenanceType = 'PREVENTIVO' THEN dbo.Maintena
  (SELECT COUNT(CASE WHEN dbo.Maintenances.MaintenanceType = 'CORRECTIVO' THEN dbo.Maintena
  (SELECT COUNT(dbo.Materials.MaterialID) FROM dbo.Materials) TotalMaterials,
  (SELECT COUNT(CASE WHEN dbo.Maintenances.MaintenanceType = 'PREVENTIVO' THEN dbo.Maintena
  (SELECT COUNT(CASE WHEN dbo.Maintenances.MaintenanceType = 'CORRECTIVO' THEN dbo.Maintena
  (SELECT SUM(dbo.Employees.Salary) FROM dbo.Employees) TotalSalary
FROM dbo.OrderItems
INNER JOIN dbo.Materials ON dbo.OrderItems.MaterialID= dbo.Materials.MaterialID
INNER JOIN dbo.Orders ON dbo.OrderItems.OrderID = dbo.Orders.OrderID
INNER JOIN dbo.Employees ON dbo.Orders.EmployeeID = dbo.Employees.EmployeeID
```

Line 2 Column 46

Enable lazy conversion ☐

Replace variables in script? ☐

Insert data from step

Execute for each row? ☐

Limit size 0

Help OK Preview Cancel



Select / Rename values

Step name:

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Ref
1	TimeID	
2	MaterialID	
3	EquipmentID	
4	EmployeeID	
5	MaintenanceID	
6	OrderID	
7	OrderItemID	
8	TotalAmountLemp	
9	TotalAmountDoll	
10	TotalHoursPrev	
11	TotalHoursCorr	
12	TotalPrev	
13	TotalCorr	
14	TotalMaterials	
15	TotalEquipmentPrev	
16	TotalEquipmentCorr	
17	TotalSalary	

Get fields to select
Edit Mapping

Include unspecified ☐

Help OK Cancel

Table output

Step name:

Connection: Edit... New... Wizard...

Target schema: Browse...

Target table: Browse...

Commit size:

Truncate table ☐

Ignore insert errors ☐

Specify database fields ☐

Main options Database fields

Partition data over tables ☐
Partitioning field:

Partition data per month ☒
Partition data per day ☐

Use batch update for ☒

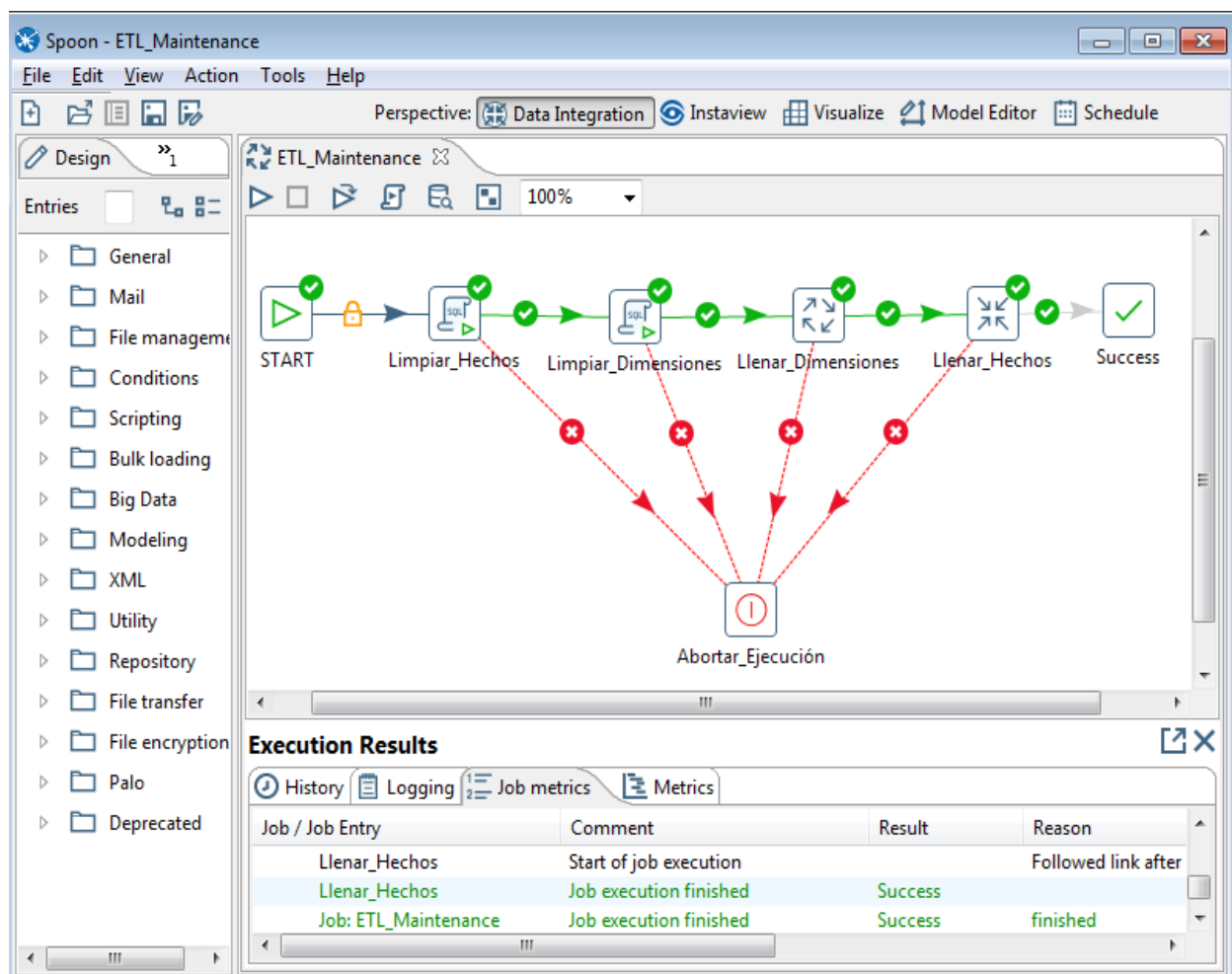
Is the name of the table ☐
Field that contains name:

Store the tablename field ☒

Return auto-generated ☐
Name of auto-generated:

Help OK Cancel SQL

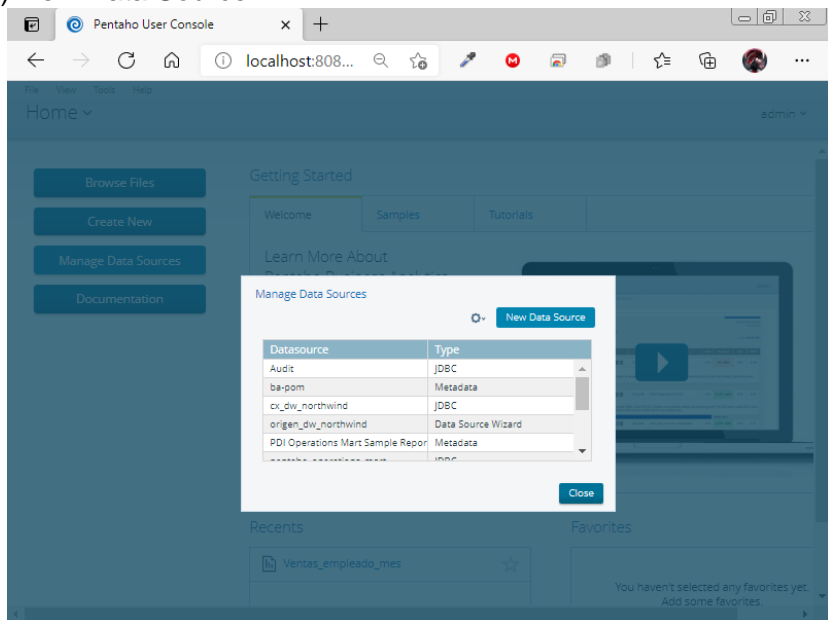
7- Ejecución del proceso ETL



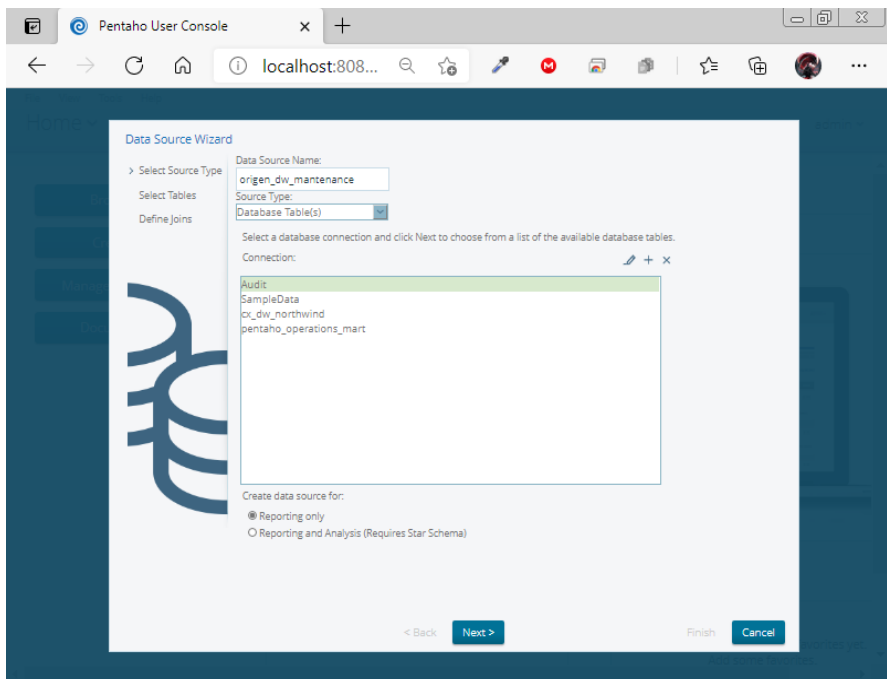
Capturas de pantalla con los pasos para crear el cubo OLAP

1- Crear la conexión con el gestor de bases de datos SQL Server.

- 1) Manage Data Sources
- 2) New Data Source

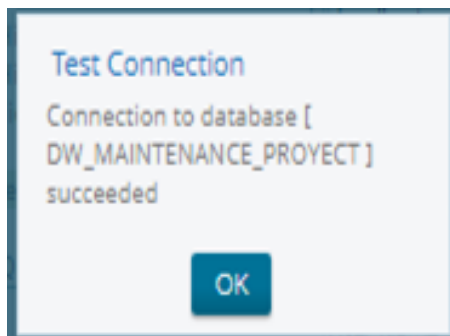
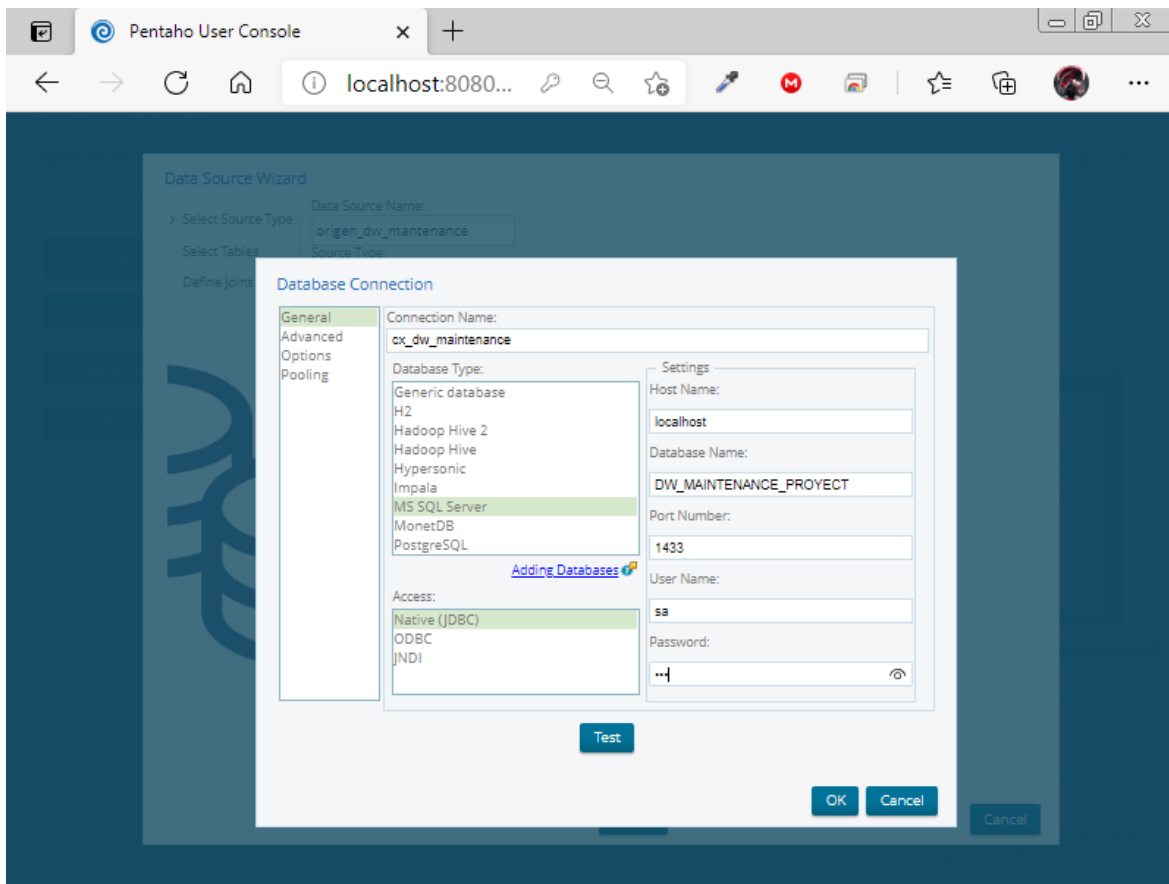


2- Asignar un nombre al origen de datos y seleccionar el tipo del origen de datos Database Tables.

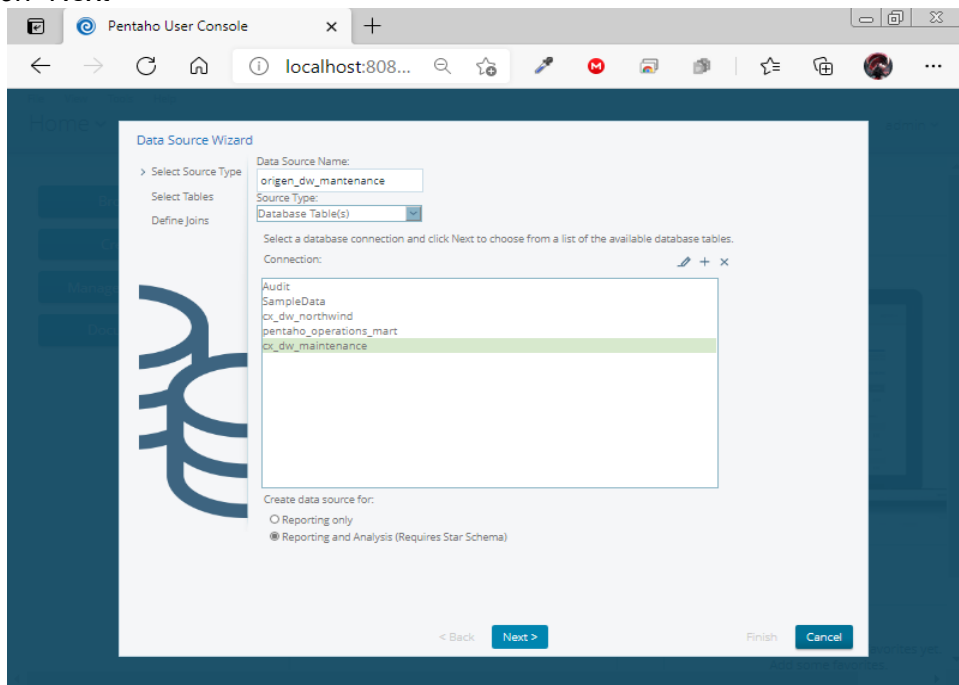


3- Crear y probar la nueva conexión para obtener la información desde el gestor de base de datos SQL Server.

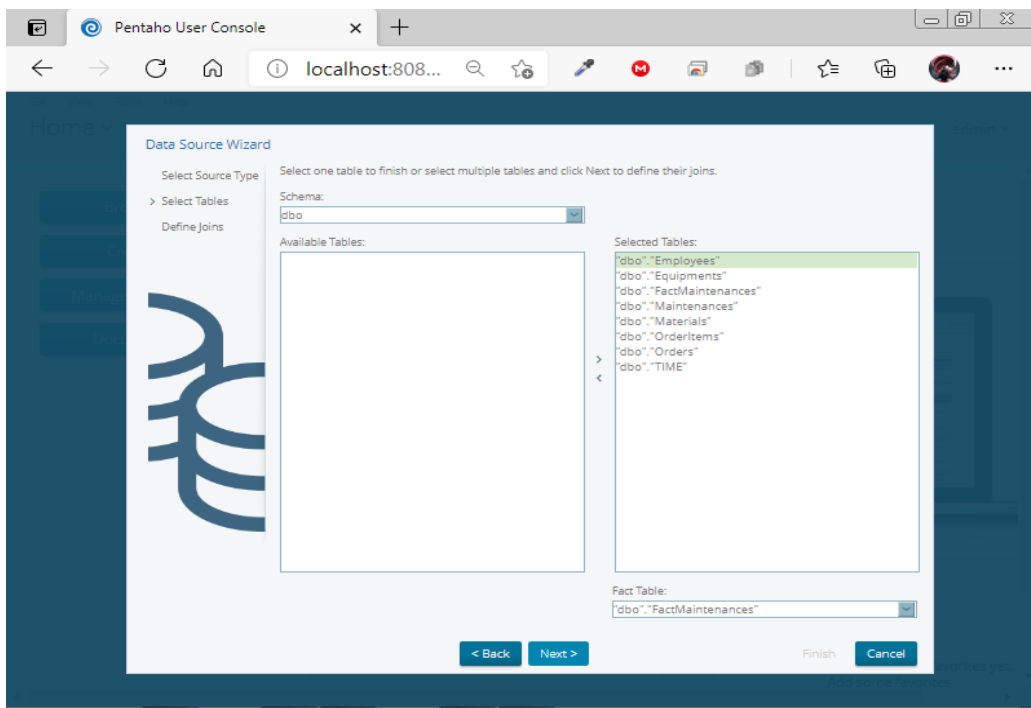
- 1) Clic en el botón “+”
- 2) Colocar los parámetros necesarios



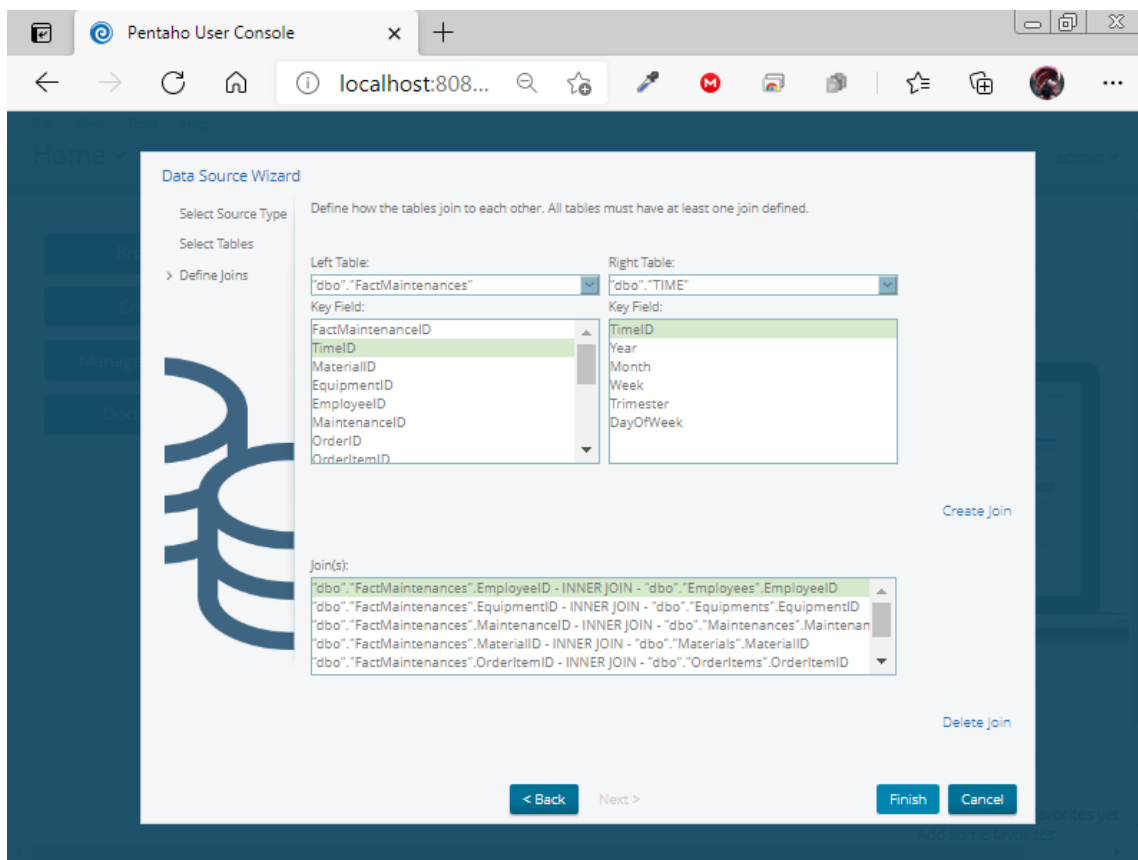
4- Seleccionar la opción de “Reporting and Analysis (Requires Star Shema)” luego clic botón “Next”



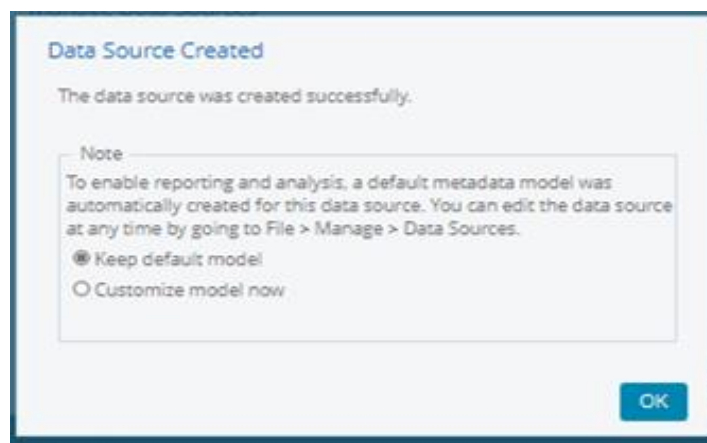
5- Seleccionar las tablas a utilizar, haciendo referencia a la tabla de hechos en la sección de “Fact Table”



- 6- Configurar la relación entre las tablas de dimensiones y la tabla de hechos, seleccionando el campo que relaciona a alguna tabla de dimensión con la de hechos y dando clic en el botón “Create Join”.

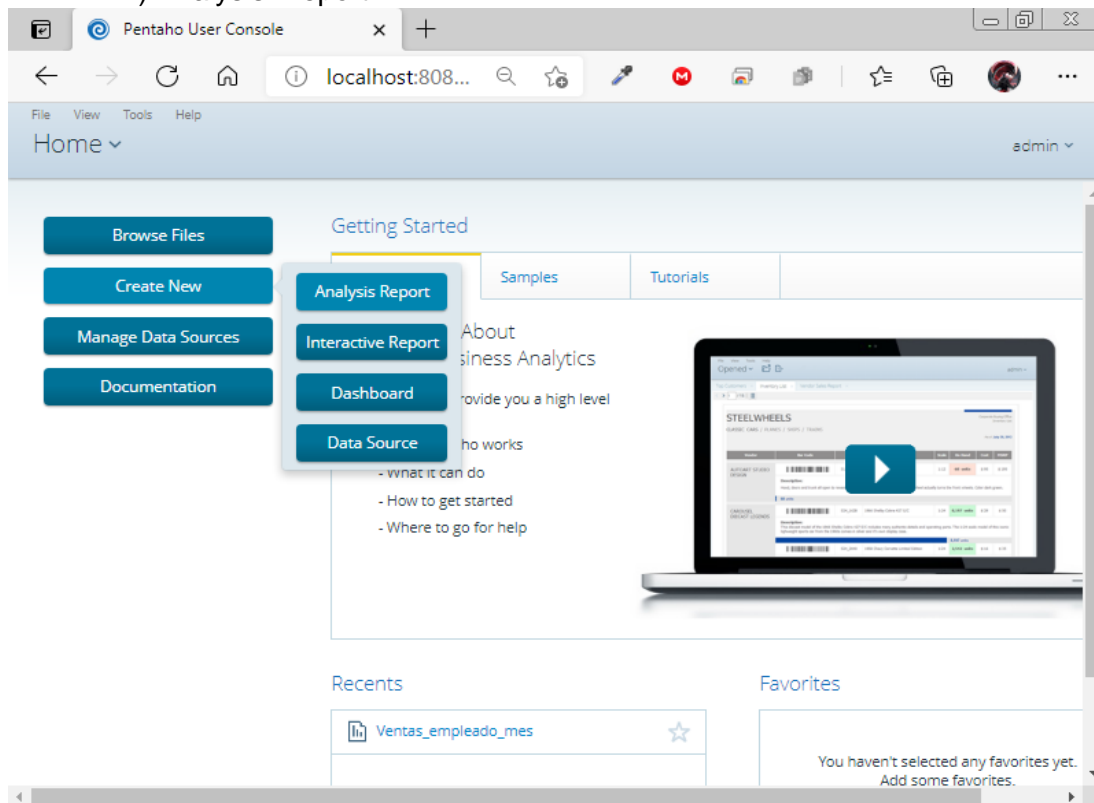


- 7- Mantener el modelo generado por defecto dando

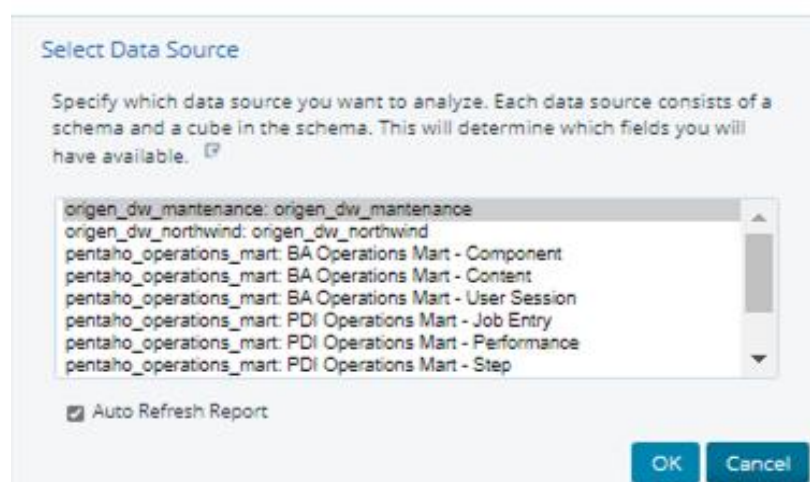


8- Crear un nuevo reporte

- 1) Create New
- 2) Analysis Report

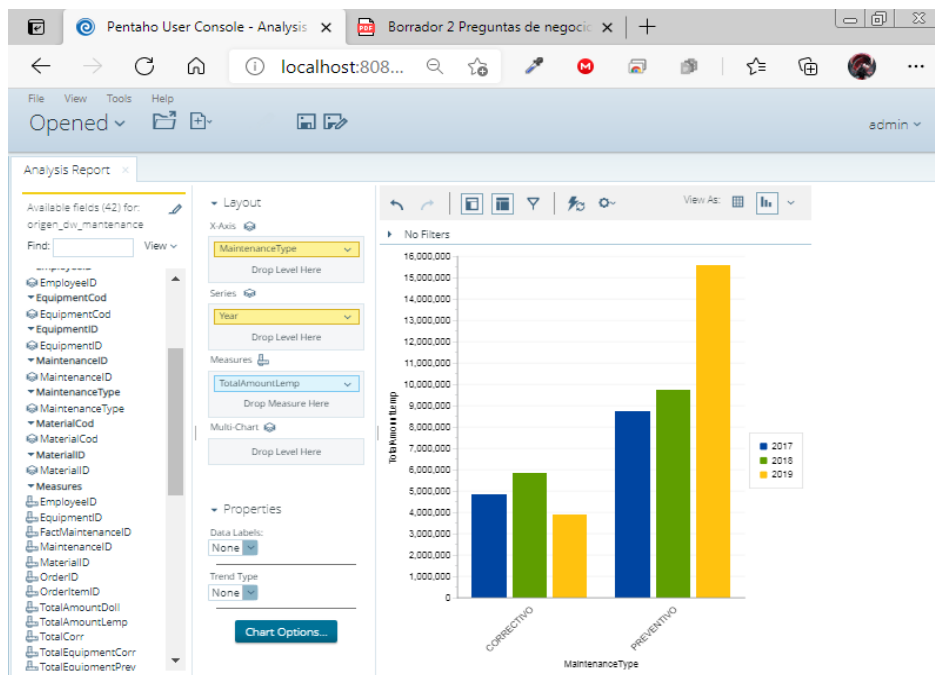


9- Seleccionar el origen de datos correspondiente

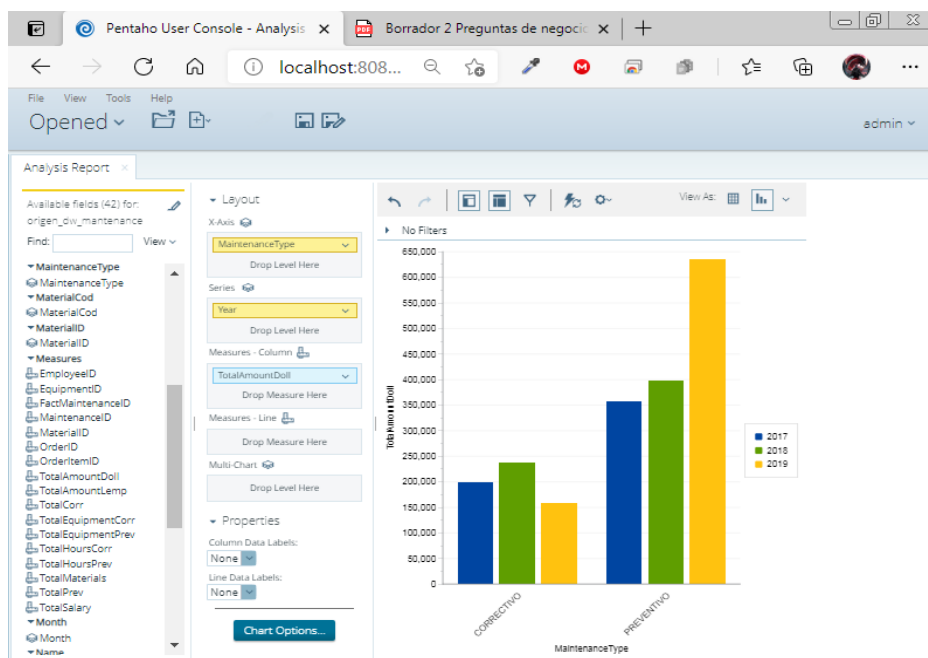


10- Seleccionar las columnas, filas y métricas necesarias para responder las preguntas de negocio

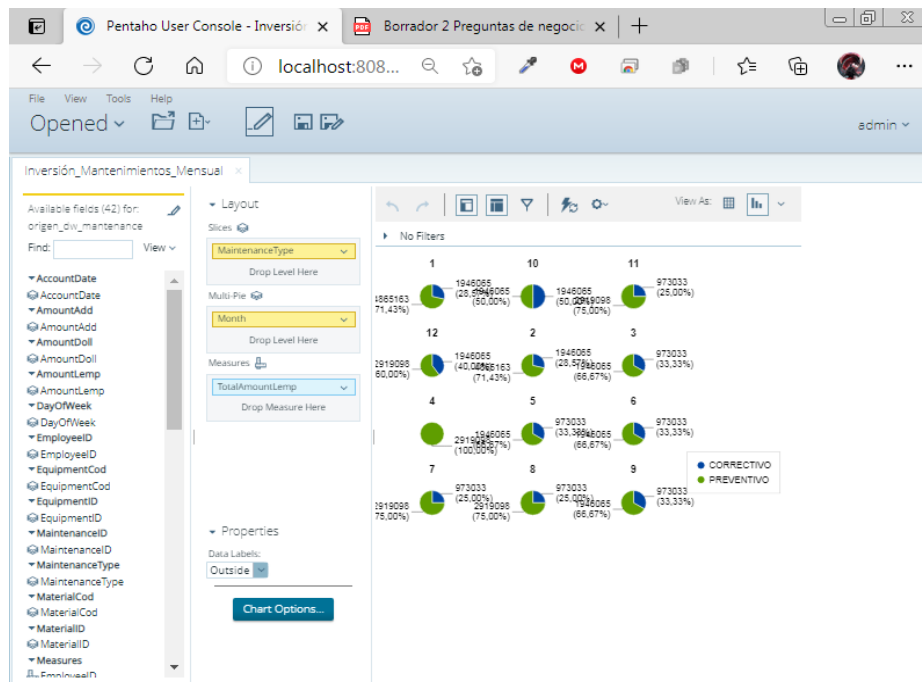
1. Conocer la inversión que se hace en lempiras por órdenes de mantenimiento al año



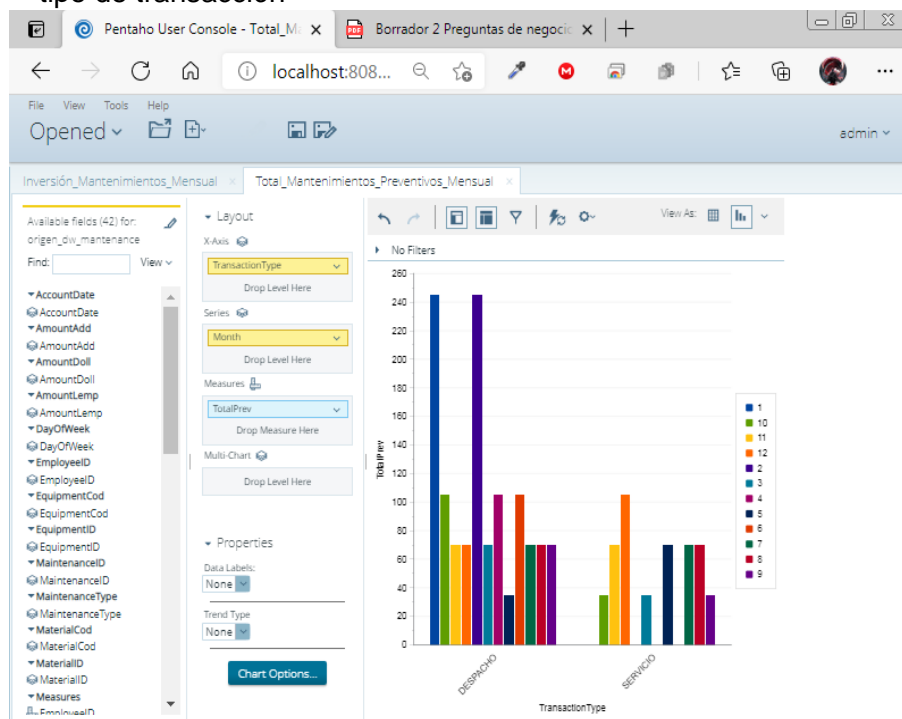
2. Conocer la inversión que se hace en dólares por órdenes de mantenimiento al año



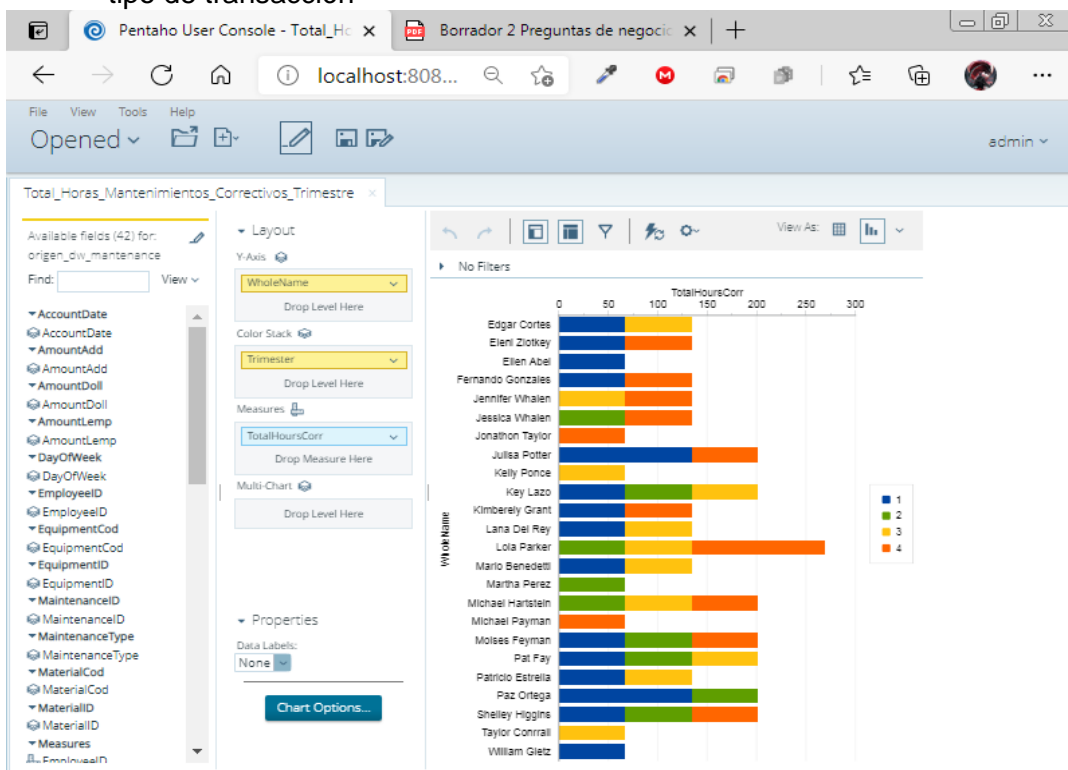
3. Cantidad invertida en mantenimientos preventivo al mes
4. Cantidad invertida en mantenimiento correctivos al mes



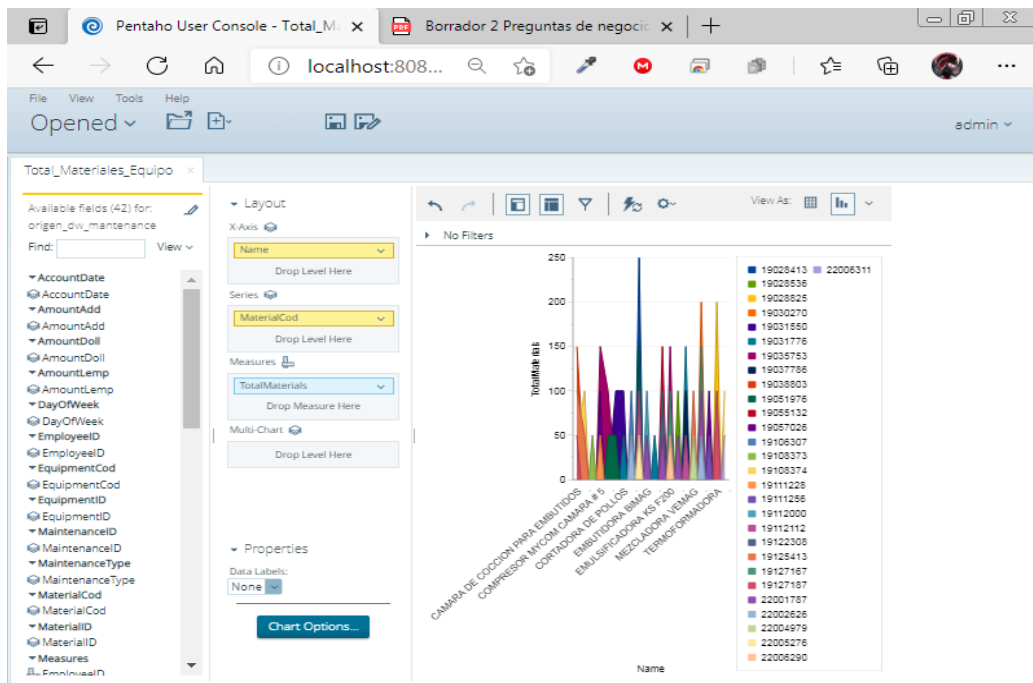
5. Cantidad de ordenes por mantenimientos preventivos mes a mes, a partir del tipo de transacción



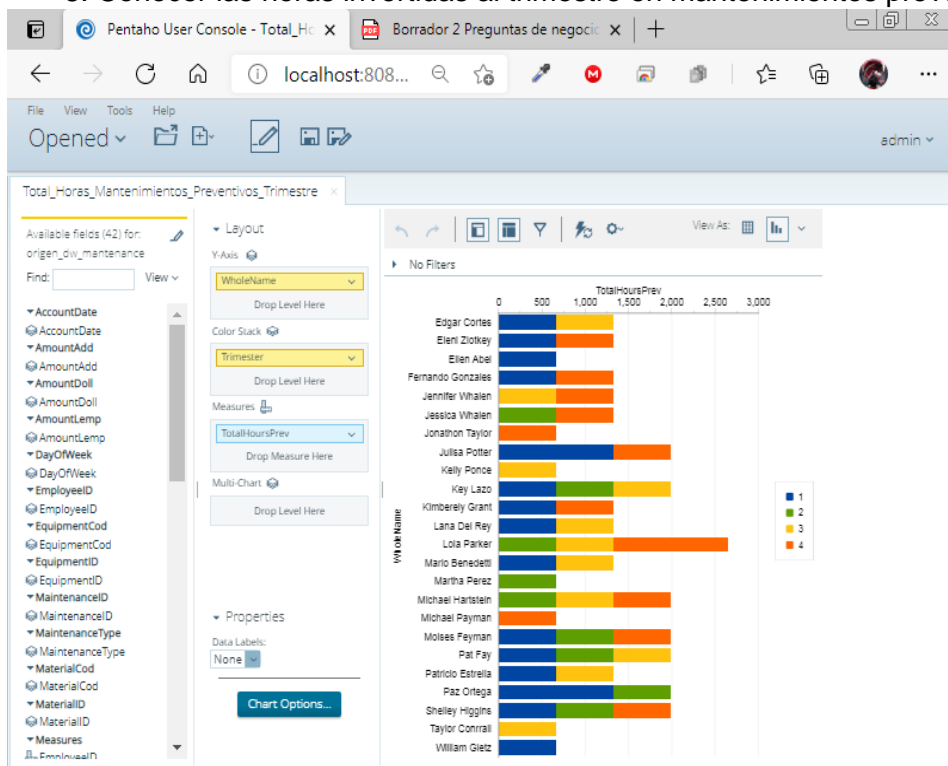
6. Cantidad de ordenes por mantenimientos correctivos mes a mes, a partir del tipo de transacción



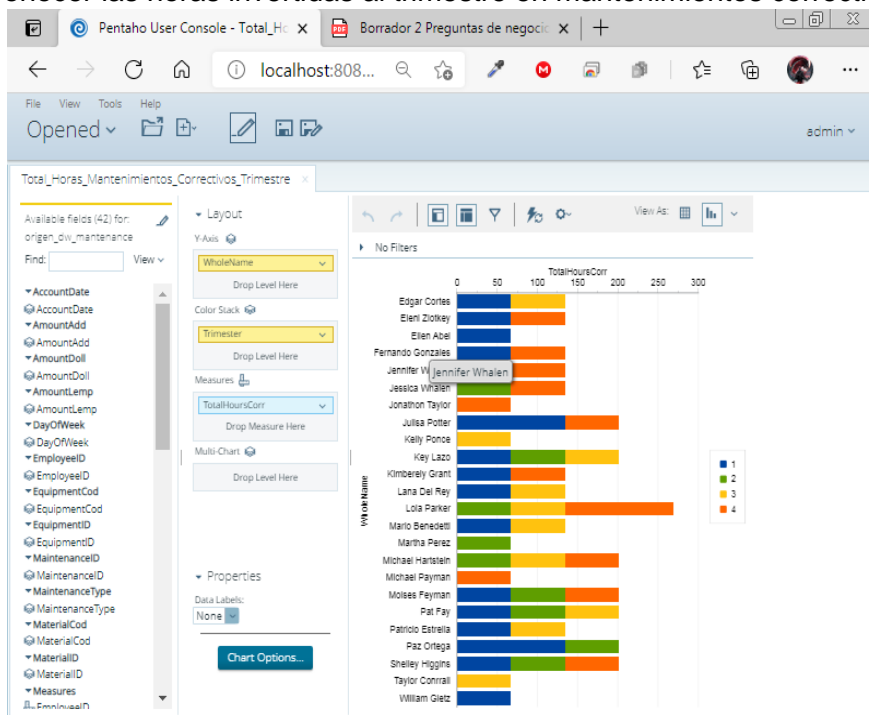
7. Conocer la cantidad de materiales utilizados a partir del equipo



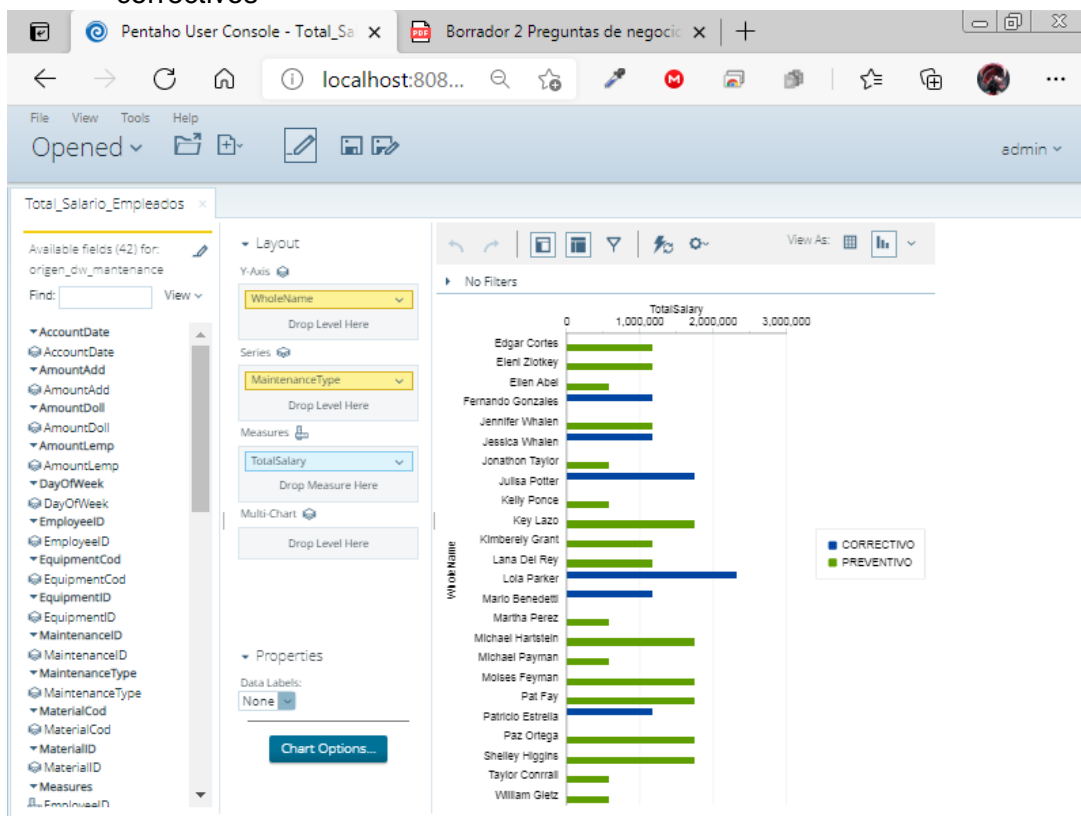
8. Conocer las horas invertidas al trimestre en mantenimientos preventivos



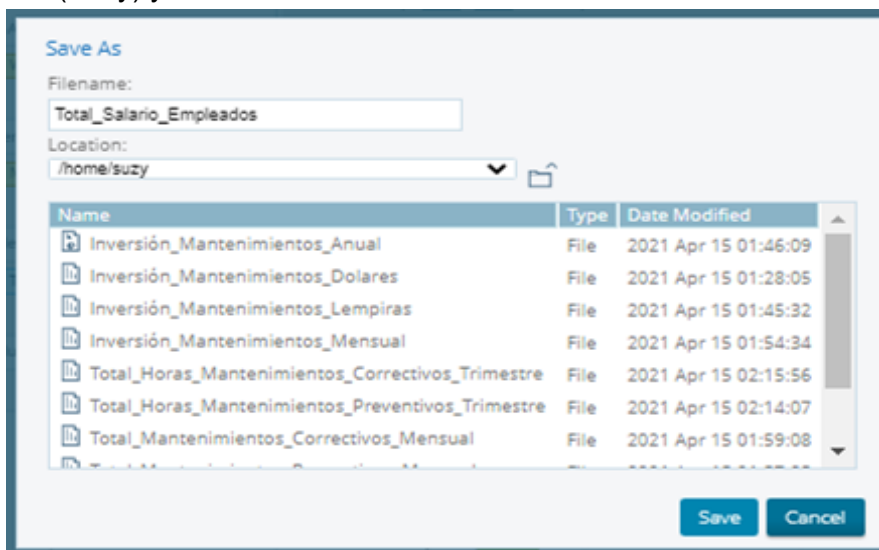
9. Conocer las horas invertidas al trimestre en mantenimientos correctivos



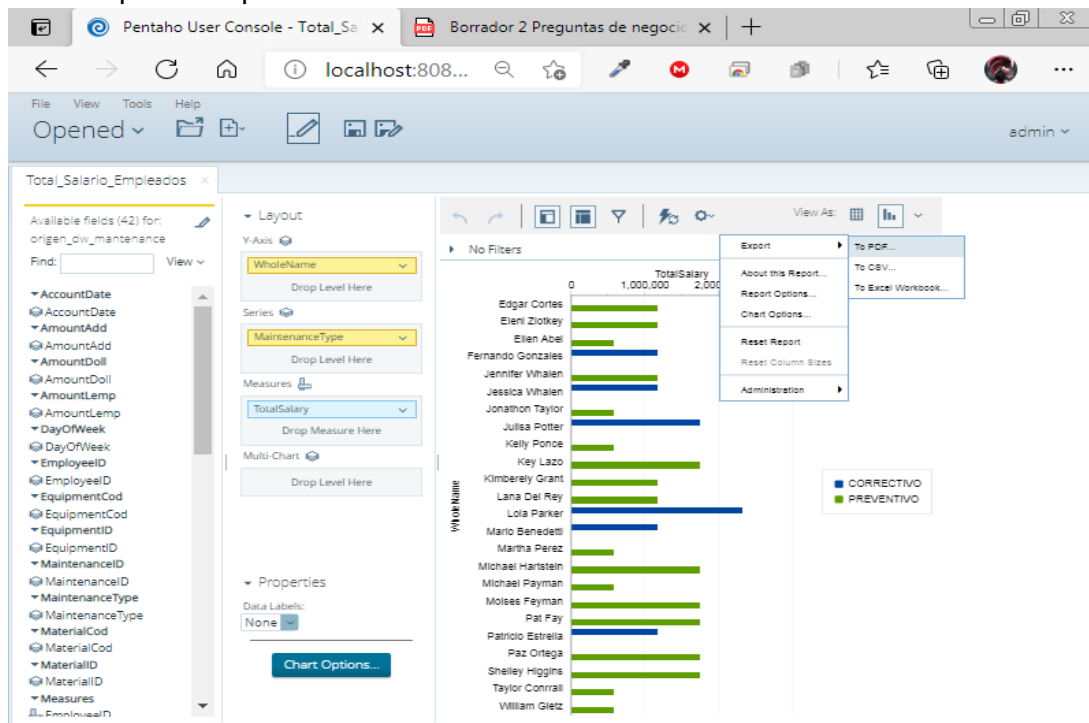
10. Promedio del salario del empleado que efectúa mantenimientos preventivos y correctivos



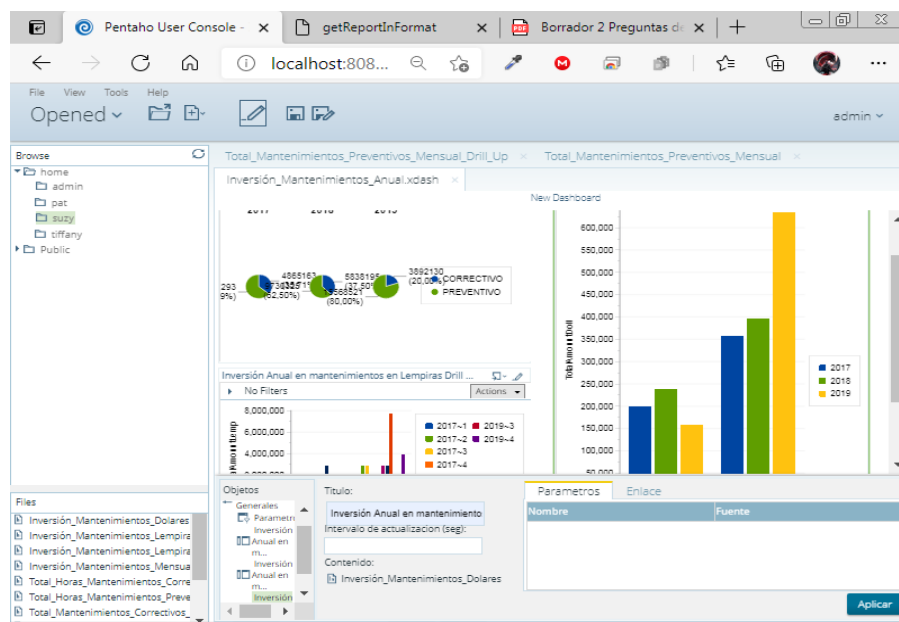
11- Guardar los reportes y tableros de control en un usuario en específico, para ello se debe hacer clic en el icono de Guardar, luego definir un nombre, seleccionar un usuario específico (Suzy) y dar clic en “Save”

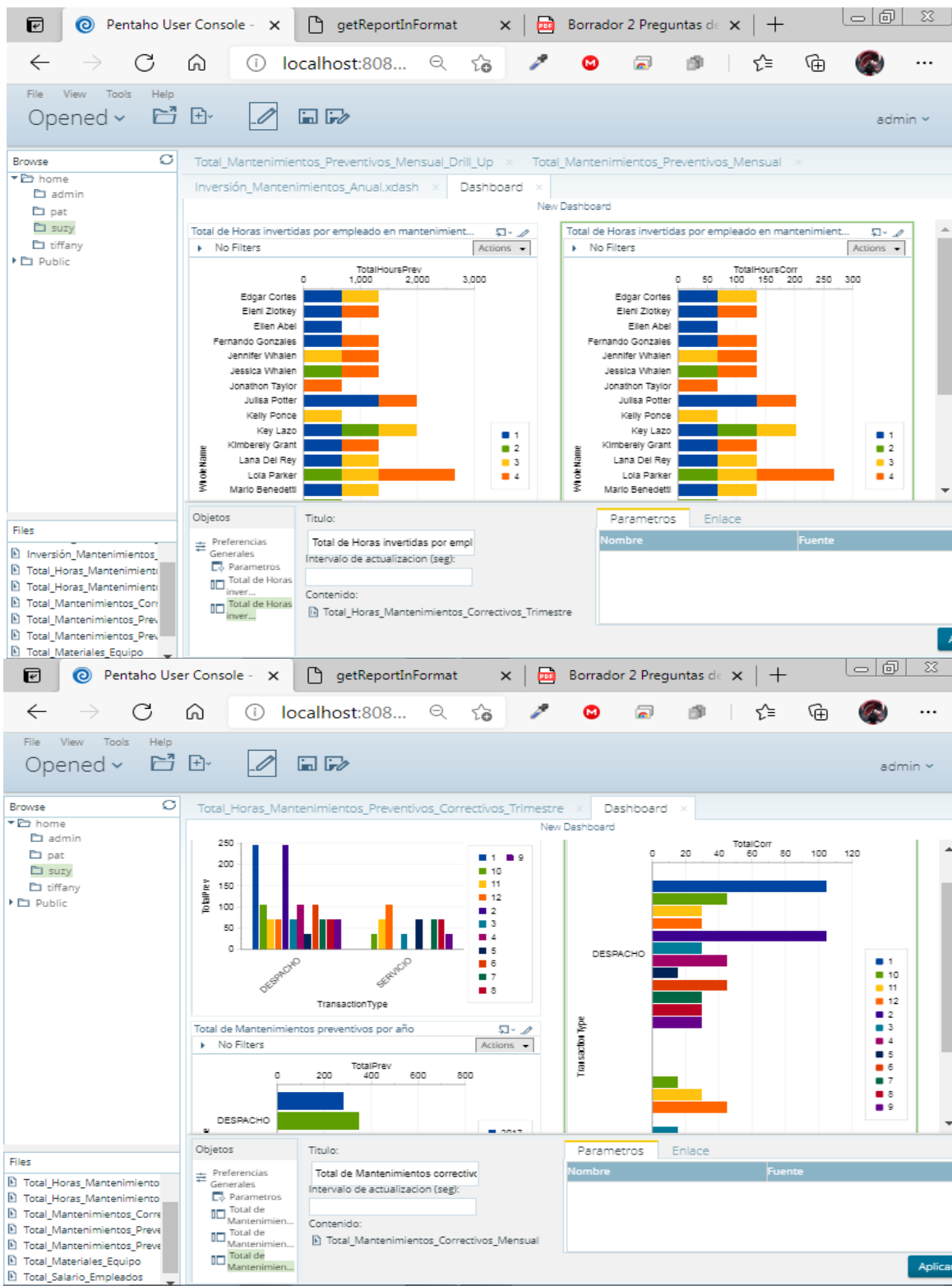


12- Para convertir los reportes a distintos formatos como PDF, CSV o una hoja de Excel, Se debe hacer clic en el icono de Ajustes, Seleccionar "Export" y seleccionar el formato que se requiera.



12- Para la creación de los tableros de control se deben agrupar los reportes asociables en una plantilla que cumpla con lo que se busca mostrar asignándole un título a cada reporte a mostrar.







Preguntas del negocio utilizadas

En la teórica del Modelo de Productividad Total Operativo se puede hacer uso de tecnologías de la información junto al concepto de Data Warehouse para poder cumplir con los 4 procedimientos del ciclo MPT, iniciando por demostrar las carencias que tiene el método actual en donde la automatización es mínima, para así empezar a optimizar la toma de decisiones hasta lograr una transformación digital orgánica.

1. Conocer la inversión que se hace en lempiras por órdenes de mantenimiento al año.
2. Conocer la inversión que se hace en dólares por órdenes de mantenimiento al año.
3. Cantidad invertida en mantenimientos preventivo al trimestre.
4. Cantidad invertida en mantenimiento correctivos al trimestre.
5. Cantidad de ordenes por mantenimientos preventivos al mes.
6. Cantidad de ordenes por mantenimientos correctivos al mes.
7. Conocer la cantidad de materiales utilizados a partir del equipo.
8. Conocer las horas invertidas al trimestre en mantenimientos preventivos.
9. Conocer las horas invertidas al trimestre en mantenimientos correctivos.
10. Total, del salario devengado por empleado que efectúa mantenimientos preventivos y correctivos.

Explicación de métricas utilizadas

A partir de las preguntas de negocio se pueden detectar varias métricas, las cuales permitirán un análisis a profundidad de los datos recopilados sin la necesidad de realizar consultas complejas. Principalmente se deben observar datos cruciales que serán el punto de partida para la automatización de todo el proceso que corresponde a la realización de los mantenimientos dentro de una planta de refrigeración.

Las métricas por utilizar, junto a su explicación son:

1. **Inversión total en lempiras por mantenimientos:**
Al obtener el total de lo que se invierte en mantenimientos en Lempiras, se abre la posibilidad de asociarlo a una dimensión del cubo OLAP que esté relacionada con el tiempo y de esta forma se responde la pregunta de negocio.



2. Inversión total en dólares por mantenimientos:

Al obtener el total de lo que se invierte en mantenimientos en dólares, al igual que en la métrica anterior se abre la posibilidad de asociarlo a una dimensión del cubo OLAP que esté relacionada con el tiempo y de esta forma se responde la pregunta de negocio.

3. Tiempo total en horas invertidas en mantenimientos preventivos:

El conteo de las horas invertidas en un tipo de mantenimiento específico además de responder la pregunta de negocio al asociar esta métrica con una dimensión de tiempo del cubo OLAP, también puede ser utilizada para desglosar la cantidad de horas que un empleado dedica a dicho trabajo por lo tanto se pueden extraer datos como el salario pagado a dicho empleado.

4. Tiempo total en horas invertidas en mantenimientos correctivos:

Al igual que en la métrica anterior, el conteo de las horas invertidas en un tipo de mantenimiento específico abre la posibilidad de asociarla con una dimensión de tiempo del cubo OLAP.

5. Cantidad total de mantenimientos preventivos efectuados:

El conteo de los mantenimientos puede ser asociado a distintas dimensiones del cubo OLAP como tiempo, empleado o equipo, de esta forma además de responder la pregunta de negocio se pueden generar reportes más extensos a partir de las asociaciones mencionadas.

6. Cantidad total de mantenimientos correctivos efectuados:

Al igual que en la métrica anterior se pueden obtener una gran variedad de reportes que respondan a más preguntas de negocio o problemáticas especuladas.

7. Cantidad total de materiales utilizados en mantenimientos:

La obtención de todos los materiales utilizados en mantenimientos permite que dicho dato pueda ser asociado a dimensiones de tiempo y equipo, de esta forma además de responder la pregunta de negocio se pueden observar patrones que servirán para optimizar el proceso de los mantenimientos.

8. Cantidad total de equipos que recibieron mantenimientos preventivos:

Puede asociarse a otras dimensiones del cubo OLAP como el tiempo y empleado, de esta forma se responde la pregunta de negocio y da la posibilidad de poder generar informes acerca de los empleados que le han realizado mantenimiento preventivo a un equipo en específico.

9. Cantidad total de equipos que recibieron mantenimientos correctivos:

Esta métrica permite que, a la hora de la realización de los reportes, se le asocie con otra dimensión del cubo OLAP como ser el tiempo y así observar patrones temporales donde los mantenimientos correctivos puedan ser minimizados en cantidad lo máximo posible.

10. Salario total pagado a los empleados que efectúan mantenimientos:

Con esta métrica se responde la pregunta de negocio planteada.



Conclusiones

- La correcta gestión de los datos obtenidos en los distintos procesos del negocio, permiten generar información valiosa que puede ser interpretada y analizada de forma eficaz, dando como resultado técnicas que optimizan la toma de decisiones internas del negocio, donde el concepto de Data Warehouse cumple con los requisitos previos para dar una solución directa a la correcta gestión de los datos que se convierten en información útil.
- Los problemas representados como pérdidas en los negocios que se abstienen o presentan resistencia a la hora de emprender la transformación digital puede quedar evidenciados con análisis de información puntual, información que en donde la mayoría de los casos, el negocio ya posee, por lo tanto es importante adaptar la cultura de las empresas a un cambio progresivo hacia una estructura de negocio con procesos automatizados donde conceptos como el Data Warehouse puede ser implementado desde el inicio de dicha transformación digital.
- Las empresas con maquinaria de tipo industrial mal gestionadas no solo representan pérdidas subyacentes para el negocio, sino también agravan problemas como el cambio climático, derroche de recursos naturales y derroche de recursos no perecederos. Por lo que la adaptación al mercado digital junto con una gestión progresivamente basada en procesos automatizados es la opción correcta para adoptar lo antes posible, puesto que las tecnologías se renuevan a un ritmo acelerado y eventos como pandemias, guerras mundiales y la colonización espacial solo aceleran la velocidad con la que se actualizan las tecnologías de la información.
- Las PYMES que adoptan tecnologías de información para la gestión de sus recursos tienen estadísticamente un porcentaje mayor de ganancias, comparado a cuando no disponían de procesos automatizados para realizar distintas acciones dentro del negocio. Las tecnologías asociadas a la Inteligencia de negocios suelen ser una respuesta a las problemáticas que presentan en negocios de menor escala como las PYMES, donde conceptos como el Data Warehouse se aplica en una medida más limitada pero de igual forma eficiente para poder llevar un control centralizado de los datos y poder gestionar las decisiones como cual producto se vende más en que épocas del año, o cual producto es que más escasea en ciertos periodos de tiempo, dando como resultado una óptima priorización de las decisiones del negocio.



Recomendaciones

- En este caso hipotético donde se tomó como referencia a empresas productoras con plantas industriales donde se tuviesen cámaras de enfriamiento como ser Cadeca o Zigma Alimentos, las cuales a la hora de realizar mantenimientos correctivos tienen que detener operaciones hasta solucionar el problema, lo que deja minuto a minuto pérdidas que pueden llegar a ser millonarias si el problema persiste o no se trata a tiempo; con lo cual conceptos como el Data Warehouse evidencian estas carencias tomando como referencia los datos que se pueden obtener de las operaciones que se ejecutan en dichas empresas. La implementación de tecnologías relacionadas a la inteligencia de negocios como las que involucran el concepto de Data Warehouse deberían de ser implementadas al iniciar un proceso de transformación tecnológica en empresas de este calibre.
- Las bases de datos OLTP y OLAP de este proyecto en concreto pueden ser ampliadas en cuanto a entidades y campos, principalmente para evidenciar los gastos en tiempo presente que se generan en mantenimientos preventivos y correctivos. Con datos reales y una exploración a profundidad de todos los elementos involucrados en una empresa del rubro específico de este proyecto se abre la posibilidad de ser utilizado para una aplicación real en campo.
- Como institución, al disponer de tanto alumnado en la carrera de Ingeniería en sistemas, este tipo de proyectos pueden ser orientados a rubros que presenten realmente carencias en su gestión de la transformación digital, donde exista la oportunidad de mejora que lo provee un análisis de datos puntual utilizando el concepto de Data Warehouse, dando una serie de rubros recomendados a la hora de la asignación del proyecto.



Bibliografía

- *Consultores y Outsourcing de Software a Medida en Honduras*. (2021, 18 febrero). Go Consultores. <https://www.goconsultores.com/>
- BI-DSS Technology | Líderes en soluciones tecnológicas. (2021, 1 enero). BI-DSS Technology. <https://www.bi-dss.com/>
- Tecnología al Servicio de su Organización. (2021, 1 enero). IT4 Honduras. <https://www.it4honduras.net/>
- Transformation step reference. (2021, 25 marzo). Pentaho Documentation. https://help.pentaho.com/Documentation/9.1/Products/Transformation_step_reference
- Kenneth C. Laudon y Jane P. Laudon, *Sistemas de Información Gerencial*, Décimo Segunda Edición, Always Learning Pearson, 2011.
- Laudon, K. C., & Laudon, J. P. (2012). *Sistemas de Información Gerencial* (Décimosegunda ed.). México: Pearson Education.