



## Actividad | #1 | Cubos OLAP

### Minería de datos

---

Ingeniería en Desarrollo de Software

TUTOR: Félix Acosta Hernández

---

ALUMNO: Alberth Alexis Rodríguez Astorga

---

FECHA: 15-02-2026



## **Indice**

Introducción.....03

Descripción.....03

Justificación.....03

Desarrollo.....04

- a) Creación de la base de datos
- b) Importación de la base de datos a Excel
- c) Gráficos
- d) Preguntas

Conclusión.....07

Referencias.....07

## **Introducción**

La minería de datos es una herramienta muy importante para analizar grandes cantidades de información, ya que permite descubrir patrones y conocimientos que no se ven fácilmente.

Hoy en día, las organizaciones generan y almacenan datos constantemente, por lo que poder convertir toda esa información en algo útil para tomar mejores decisiones se ha vuelto una necesidad para ellas.

A través de la minería de datos es posible identificar tendencias, predecir comportamientos y encontrar relaciones entre distintos datos que a simple vista no se detectan, esto resulta especialmente valioso en áreas como la salud, las finanzas, el marketing y la seguridad, donde la información obtenida puede ayudar a mejorar procesos, ofrecer mejores servicios, prevenir fraudes e incluso salvar vidas.

Además, permite crear modelos predictivos que anticipan riesgos y necesidades, ayudando a las empresas a ser más competitivas y a adaptarse a un entorno cada vez más basado en datos.

## **Descripción**

En esta actividad trabajaremos con la creación y manejo de una base de datos en Microsoft SQL Server.

Primero se creara una base de datos llamada “datosgenerales” y dentro de ella diseñaremos tres tablas: “datospersonales”, “datosempleado” y “sucursal”.

Cada tabla fue estructurada con sus respectivas columnas, claves primarias y llaves foráneas para relacionarlas correctamente.

Después ingresaremos al menos diez registros en cada tabla para contar con información suficiente para realizar consultas.

Posteriormente, se elaboraran distintos queries para obtener datos específicos, como los salarios de los empleados, las ventas totales de las sucursales y el total de empleados según su género, con esta información ya desglosada identificaremos mediante graficas de Excel ciertas tendencias para contestar preguntas.

## **Justificación**

La minería de datos permite manejar y analizar grandes cantidades de información, algo esencial en un mundo donde las organizaciones generan datos constantemente.

Algunas ventajas de la minería de datos:

Facilita la identificación de patrones y relaciones ocultas, proporcionando conocimientos valiosos que pueden transformar la forma en que una empresa opera, gracias a esto, se mejora la toma de decisiones, ya que los datos crudos se convierten en información útil que optimiza recursos y estrategias, mejorando el rendimiento.

También permite prever tendencias y comportamientos futuros, lo que ayuda a anticiparse a las necesidades del mercado y mantener una ventaja competitiva.

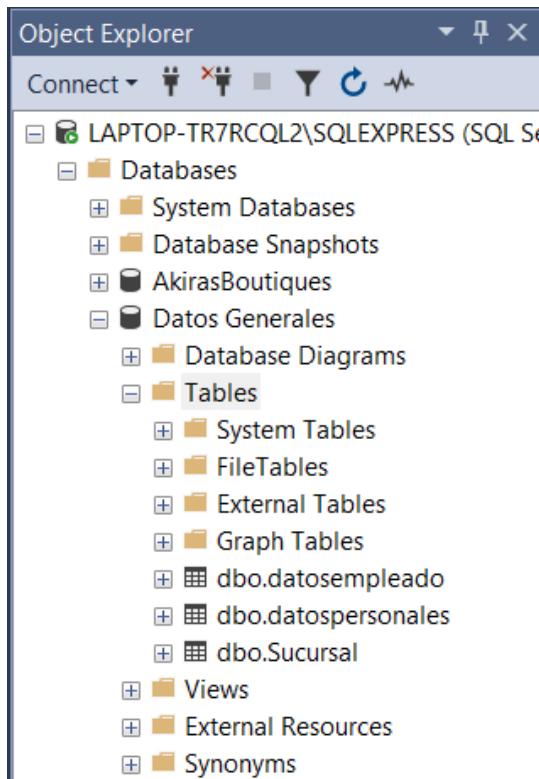
Otro beneficio importante es la mejora en la eficiencia operativa, al identificar áreas de mejora y reducir costos.

Finalmente, la minería de datos facilita la creación de modelos predictivos que permiten a las organizaciones estar preparadas y ser proactivas frente a posibles riesgos y cambios.

## Desarrollo

Para propósitos de la actividad se estarán utilizando algunas herramientas que hemos aplicado en otras materias, ya tenemos el Microsoft SQL Server Management previamente instalado y operando:

Creamos el servidor en el que estaremos trabajando y alojando nuestras tablas:



Crearemos las tablas e insertaremos la información en esta utilizando comando en SQL:

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) Query Editor window. The code pane contains the following SQL script:

```
CREATE TABLE datospersonales (
    ID INT PRIMARY KEY IDENTITY(1,1),
    Nombre VARCHAR(100),
    Direccion VARCHAR(150),
    Ciudad VARCHAR(100),
    Numtelefono VARCHAR(20),
    Email VARCHAR(100),
    Genero VARCHAR(20)
);

CREATE TABLE Sucursal (
    ID INT PRIMARY KEY IDENTITY(1,1),
    NombreSucursal VARCHAR(100),
    Ciudad VARCHAR(100),
    Direccion VARCHAR(150),
    Cantidad_ventas_anuales INT
);

CREATE TABLE datosempleado (
    Codigo INT PRIMARY KEY IDENTITY(1,1),
    FK_datospersonales INT,
    FK_Sucursal INT,
    Cargo VARCHAR(100),
    Horario VARCHAR(100),
    SUELDO DECIMAL(10,2),
    FOREIGN KEY (FK_datospersonales) REFERENCES datospersonales(ID),
    FOREIGN KEY (FK_Sucursal) REFERENCES Sucursal(ID)
);

INSERT INTO datospersonales (Nombre, Direccion, Ciudad, NumTelefono, Email, Genero) VALUES
('Juan Perez', 'Av. Reforma 123', 'Ciudad de Mexico', '5551234567', 'juan.perez@gmail.com', 'Masculino'),
('Maria Lopez', 'Calle Hidalgo 45', 'Guadalajara', '331234567', 'maria.lopez@gmail.com', 'Femenino'),
('Carla Sanchez', 'Av. Juarez 567', 'Monterrey', '8641234567', 'carla.sanchez@gmail.com', 'Masculino'),
('Ana Hernandez', 'Calle Mexico 234', 'Puebla', '2221234567', 'ana.hernandez@gmail.com', 'Femenino'),
('Luis Gomez', 'Av. Central 456', 'Toluca', '7221234567', 'luis.gomez@gmail.com', 'Masculino'),
('Sofia Ramirez', 'Calle Norte 234', 'Merida', '9991234567', 'sofia.ramirez@gmail.com', 'Femenino'),
('Pedro Torres', 'Av. Sur 876', 'Tijuana', '6641234567', 'pedro.torres@gmail.com', 'Masculino'),
('Laura Hernandez', 'Calle Este 321', 'Queretaro', '4421234567', 'laura.hernandez@gmail.com', 'Femenino');

10 rows affected
Completion time: 2024-02-15T10:10:51.8141180+01:00
```

The status bar at the bottom indicates "Query executed successfully." and shows the completion time as 2024-02-15T10:10:51.8141180+01:00.

Podemos consultar las tablas que hemos creado con los siguientes queries:

The screenshot shows a SQL query window titled "SQLQuery1.sql - not connected". It contains three separate SELECT statements. The first statement selects Nombre and de.Salario from datosempleado inner joined with datospersonales. The second statement selects NombreSucursal and Cantidad\_ventas\_anuales from Sucursal. The third statement counts Total\_Employees from datosempleado inner joined with datospersonales grouped by Genero. Below the queries is a results grid showing data from the Sucursal table. The status bar at the bottom indicates the user is disconnected.

```
SELECT
    dp.Nombre,
    de.Salario
FROM datosempleado de
INNER JOIN datospersonales dp
    ON de.FK_datospersonales = dp.ID;

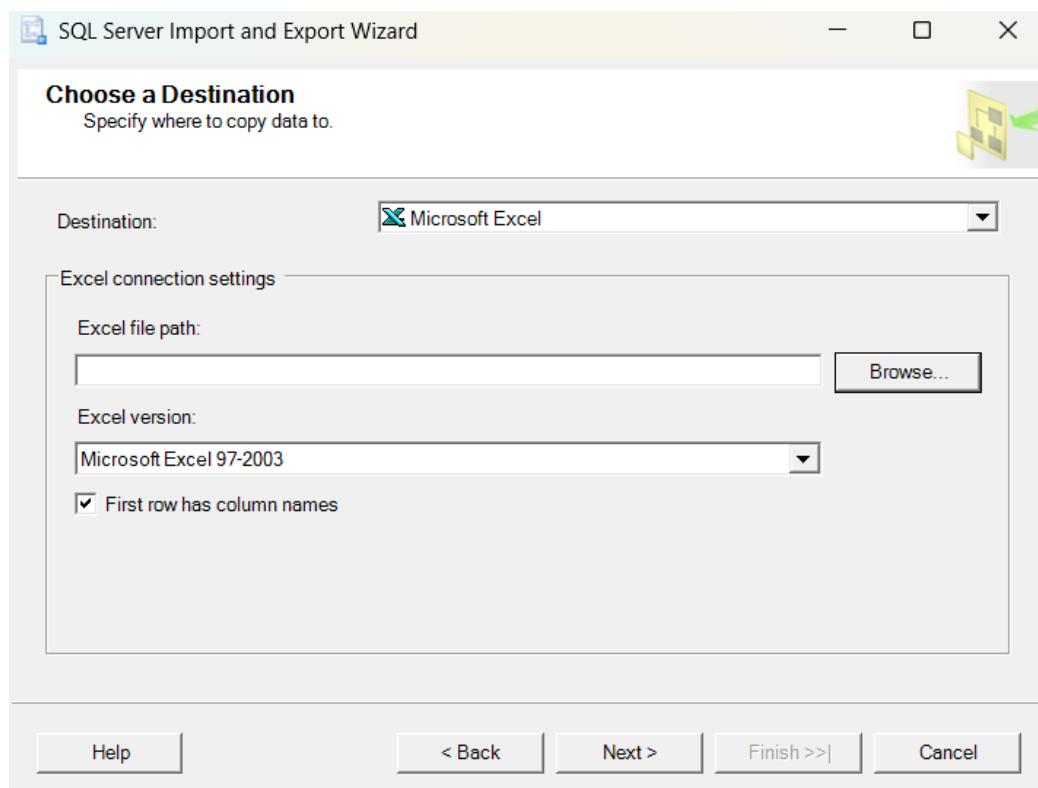
SELECT
    NombreSucursal,
    Cantidad_ventas_anuales
FROM Sucursal;

SELECT
    dp.Genero,
    COUNT(*) AS Total_Employees
FROM datosempleado de
INNER JOIN datospersonales dp
    ON de.FK_datospersonales = dp.ID
GROUP BY dp.Genero;

SELECT
    de.Codigo,
    dp.Nombre,
    dp.Genero,
    dp.Ciudad AS Ciudad_Persona,
    dp.Email,
    s.NombreSucursal,
    s.Ciudad AS Ciudad_Sucursal,
    de.Cargo,
    de.Horario,
    de.Salario,
    s.Cantidad_ventas_anuales
FROM datosempleado de
INNER JOIN datospersonales dp
    ON de.FK_datospersonales = dp.ID
```

Codigo	Nombre	Genero	Ciudad_Persona	Email	NombreSucursal	Ciudad_Sucursal	Cargo	Horario	Salario	Cantidad_ventas_anuales
1	Juan Perez	Masculino	Ciudad de Mexico	juan.perez@gmail.com	Sucursal Centro	Ciudad de Mexico	Gerente	9:00-18:00	25000.00	500000
2	Maria Lopez	Femenino	Guadalajara	maria.lopez@gmail.com	Sucursal Norte	Monterrey	Vendedor	8:00-16:00	12000.00	350000
3	Carlos Sanchez	Masculino	Monterrey	carlos.sanchez@gmail.com	Sucursal Sur	Guadalajara	Tecnico	10:00-19:00	15000.00	420000
4	Ana Martinez	Femenino	Puebla	ana.martinez@gmail.com	Sucursal Este	Puebla	Vendedor	9:00-17:00	11000.00	280000

Ya creadas las tablas necesarias para la actividad, procederemos a exportar la base de datos a una hoja de calculo dinámica(Excel):



Esa es una manera, en este caso como yo estoy manejando libre office, no puedo exportarlo de esa forma, entonces solo usare una forma muy sencilla enviándola a un formato CSV para su manipulación en una hoja de calculo para poder graficar:

The screenshot shows a LibreOffice Calc window with two tabs: 'Results' and 'Messages'. The 'Results' tab displays a table of employee data:

Codigo	Nombre	Genero	Ciudad_Persona
1	Juan Perez	Masculino	Ciudad de Mexico
2	Maria Lopez	Femenino	Guadalajara
3	Carlos Sanchez	Masculino	Monterrey
4	Ana Martinez	Femenino	Puebla
5	Luis Gomez	Masculino	Toluca
6	Sofia Ramirez	Femenino	Merida
7	Pedro Torres	Masculino	Tijuana
8	Laura Hernandez	Femenino	Queretaro
9	Miguel Diaz	Masculino	Leon
10	Daniela Cruz	Femenino	Cancun

The 'Messages' tab contains two SQL queries:

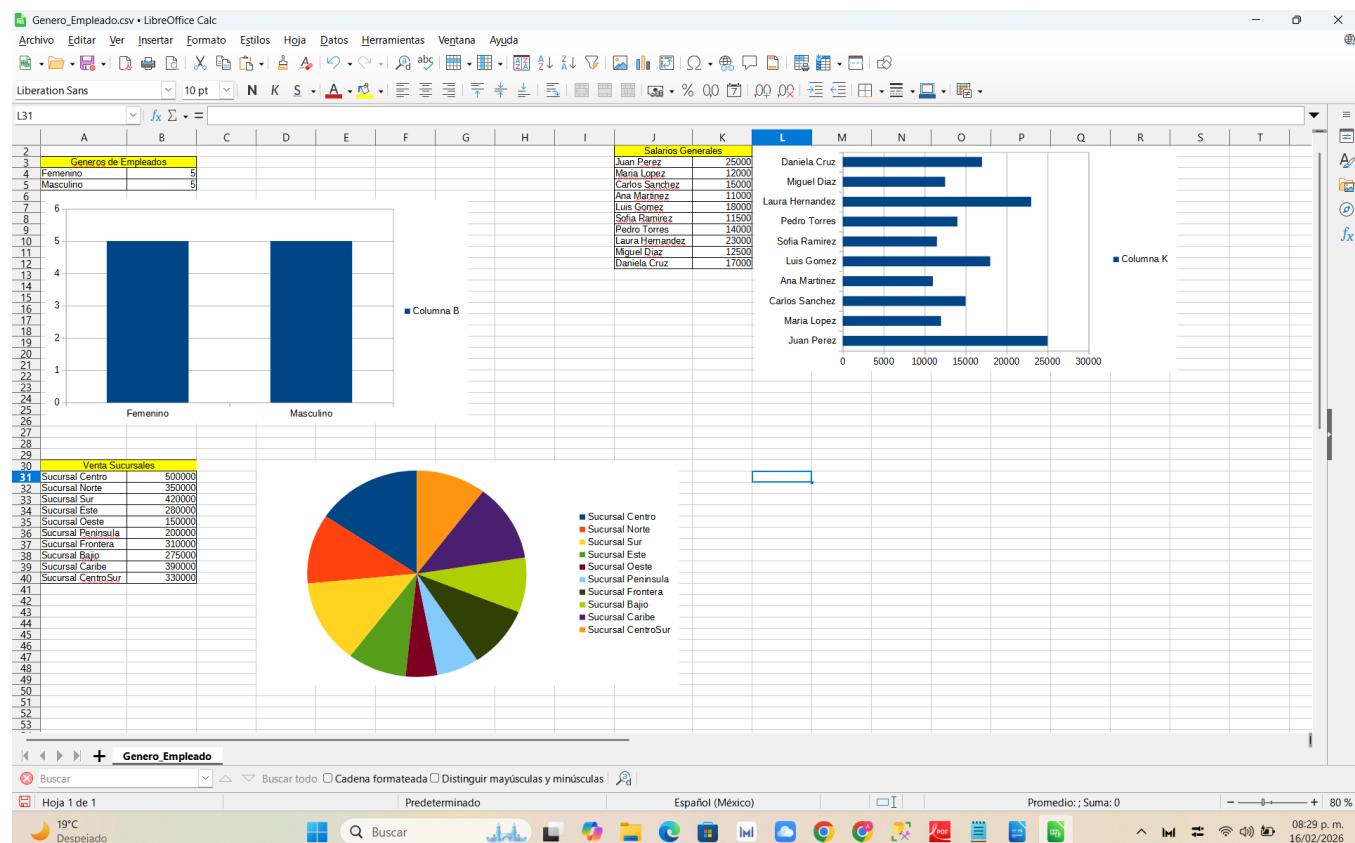
```

SELECT
    NombreSucursal,
    Cantidad_ventas_anuales
FROM Sucursal;

SELECT
    dp.Genero,
    COUNT(*) AS Total_Empleados
FROM datosempelado de
INNER JOIN datospersonales dp
    ON de.FK_datospersonales = dp.ID
GROUP BY dp.Genero;
  
```

A 'Save Grid Results' dialog is open, showing a file path 'Materias > Minería y Análisis de Datos'. It lists three files: 'Genero\_Empelado', 'Salarios\_Generales', and 'Ventas\_Sucursales'. The 'Nombre de archivo:' field is empty, and the 'Tipo:' dropdown is set to 'CSV (Comma delimited) (\*.csv)'. Buttons for 'Save' and 'Cancelar' are at the bottom.

Graficamos la información del servidor:



Utilizamos diferentes graficas para ser mas vistoso.

Preguntas:

- ¿Quién es la persona que gana más?

En base a los registros proporcionados, Juan Pérez es la persona que gana más, con un salario de \$25,000.00 (según la tabla de empleados).

- ¿Cuál es la sucursal que gana menos anualmente?

La sucursal que gana menos anualmente es Sucursal Oeste, con \$150,000 en ventas anuales (según la tabla de sucursales).

- ¿Cuántos empleados son hombres?

De los 10 empleados que ingresamos, son 5 hombres.

## Conclusión

En esta actividad, se aprendió a crear y gestionar una base de datos en Microsoft SQL Server, donde se diseñaron tres tablas interconectadas para almacenar información de empleados y sucursales.

A través de consultas SQL, se extrajeron datos específicos como salarios, ventas anuales y el total de empleados por género.

Además, se generaron gráficas para visualizar esta información de manera clara y ordenada, facilitando el análisis y la toma de decisiones.

Las gráficas de salarios, ventas y género nos permitieron entender mejor la distribución de los datos, identificar tendencias y relaciones importantes entre los diferentes aspectos del negocio.

Esta actividad no solo refuerza habilidades en manejo de bases de datos y consultas SQL, sino también en la interpretación de datos y la creación de informes visuales útiles para la toma de decisiones estratégicas.

## Referencias

**McDonald, K. (2006). *SQL Server 2005 for Dummies*. Wiley Publishing.**

**Microsoft. (2021). *Microsoft SQL Server Management Studio [Software]*. Recuperado de <https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms>**

**Página de GitHub: <https://github.com/AlexisRodriguez1988/Mineria-de-datos->**