

# **Actividad 1 - Cubos OLAP**

## **Minería de Datos**

### **Ingeniería en Desarrollo de Software**

**Tutor: Félix Acosta Hernández**

**Alumno: Fernando Pedraza Garate**

**Fecha: 30 de Septiembre del 2023**

# Índice

---

## Etapa 1 – Cubos OLAP.

- Introducción. Pág. 3-4
- Descripción Pág. 5-6
- Justificación Pág. 7
- Desarrollo Pág.8-22
  - Creación de la base de datos
  - Importación de la base de datos a Excel
  - Gráficos
  - Preguntas
- Conclusión Pág. 23
- Referencias. pág. 24-25

# Introducción

---

El término KDD hace referencia a las siglas de Knowledge Discovery in Databases, que traducido al español significa el “Descubrimiento de Conocimiento en Bases de Datos”, este término fue mencionado por primera vez por Piatetsky-Shapiro en 1989 para describir el proceso por medio del cual se extraía de una base de datos información útil hasta ahora desconocida para quien realizaba el proceso, el cual involucra un trabajo extenso en minería de datos como uno de sus pasos más importantes.

El KDD implica un conjunto de etapas definidas para el tratamiento de los datos antes de aplicar las diferentes técnicas de minería de datos en la búsqueda de patrones ocultos en ellos para finalmente hacer el análisis de los patrones encontrados con el objeto de brindar nueva información y conocimiento.

Hasta hace unos años la minería de datos y el KDD eran mencionadas indiscriminadamente para referirse a la extracción de patrones de las variables en una base de datos, donde hasta hace poco menos de 10 años algunos investigadores comenzaron a utilizar el término KDD para referirse a la extracción de conocimiento proveniente de bases de datos como un macro proceso, mientras que la minería de datos era catalogada como el mecanismo de aplicación de algoritmos para extraer patrones presentados por los mismos.

Se dice entonces que el KDD es el proceso para la identificación de patrones válidos, nuevos, útiles y sobre todo comprensibles, que conllevan al descubrimiento de nuevo conocimiento, en cambio, la minería de datos se refiere a solo una de las etapas de dicho proceso, considerada por muchos como la más importante.

En resumen la minería de datos es el núcleo del proceso llamado “**KDD**, del inglés (Knowledge Discovery in Databases), Descubrimiento de conocimiento en bases de datos” el cual ayuda a encontrar un modelo valido, útil y entendible, que describa, permita explorar, y depurar los datos para encontrar patrones de comportamiento para así poder crear predicciones de acuerdo a la información proporcionada.

# Descripción.

---

En base al requerimiento de la empresa CompuVentas, empresa dedicada a la venta de hardware donde solicita crear una nueva base de datos llamada “datosgenerales” en Microsoft SQL Server Management, la cual deberá incluir las siguientes tablas y sus respectivas columnas:

Tabla 1:	Tabla 2:	Tabla 3:
datospersonales	Datosempleado	Sucursal
Columnas:	Columnas:	Columnas:
- ID	- Código	- ID
- Nombre	- FK_datospersonales	- Nombre sucursal
- Dirección	- FK_ Sucursal	- Ciudad
- Ciudad	- Cargo	- Dirección
- Num. Teléfono	- Horario	- Cantidad_ventas_anuales
- Email	- Salario	
- Género		

Para poder exportarla a Excel, que permita presentar las gráficas de los siguientes puntos especificados:

- Gráfica 1: Salarios generales
- Gráfica 2: Ventas totales de las sucursales.
- Gráfica 3. Total de empleados acorde a su género

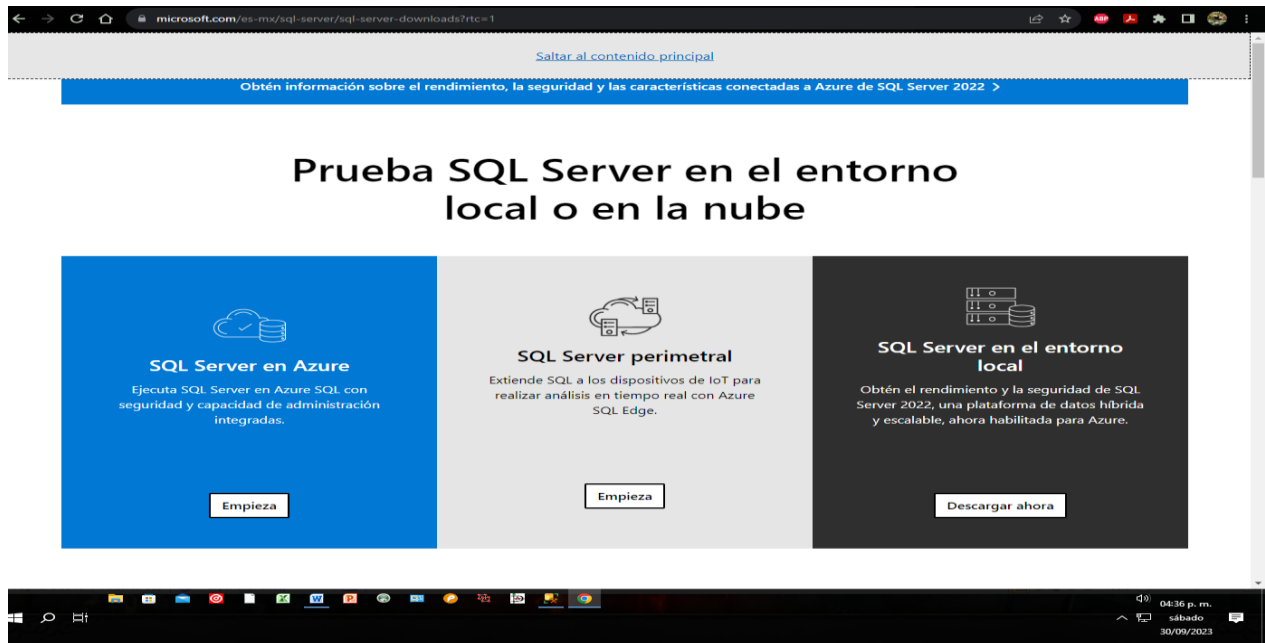
Una vez realizada la base de datos, se deberán ingresar como mínimo 10 registros por tabla y presentar la información en las gráficas mencionadas.

# Justificación.

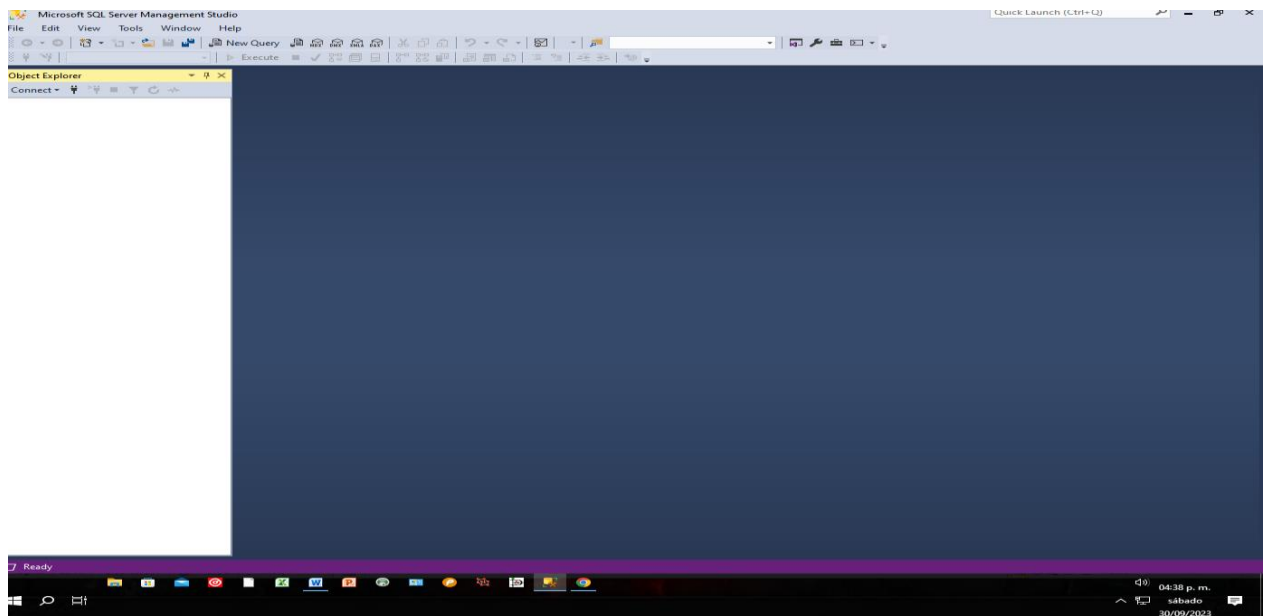
---

La implementación de esta terminología permite tomar decisiones más acertadas y objetivas para la resolución de cualquier evento que se presente ante cualquier situación, desde los financieros, hasta la implementación de una mejora, como el lanzamiento de nuevos productos al mercado, la creación de nuevos negocios o empresas, la innovación de cualquier producto, o simplemente el mejorar los procesos ya establecidos para evitar accidentes o incidencias, como fallas mecánicas, de infraestructura, incluyendo el recurso humano y su comportamiento, creando patrones de comportamiento que basados en estadísticas y resultados históricos puedan generar gráficos que permitan demostrar de forma sencilla, clara y entendible el porqué de cada posible decisión a considerar, logrando evitar daños irreparables con antelación, demostrando así los beneficios que se obtendrían al establecer metas y objetivos bien sustentados por medio de la minería de datos y al realizar los cambios de forma oportuna a los riesgos o situaciones que se llegaran a presentar con los cambios propuestos.

# Desarrollo.

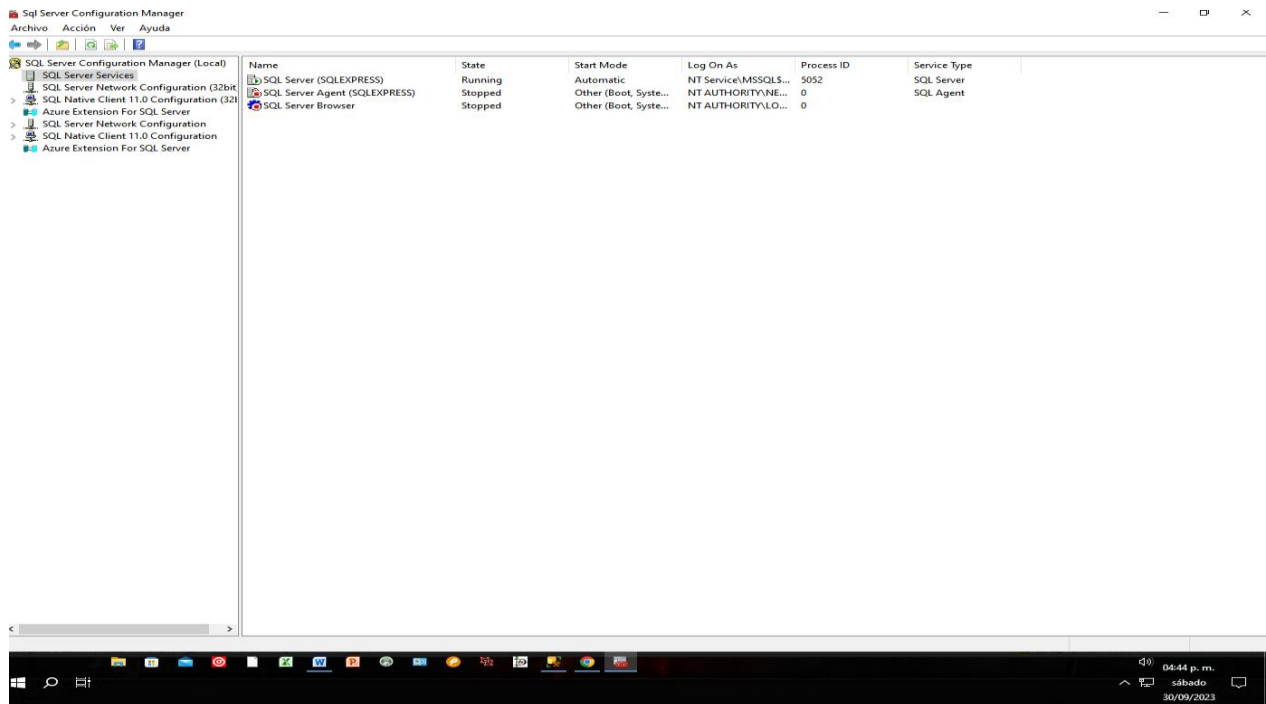


Instalamos y configuramos SQL Server desde su página oficial.

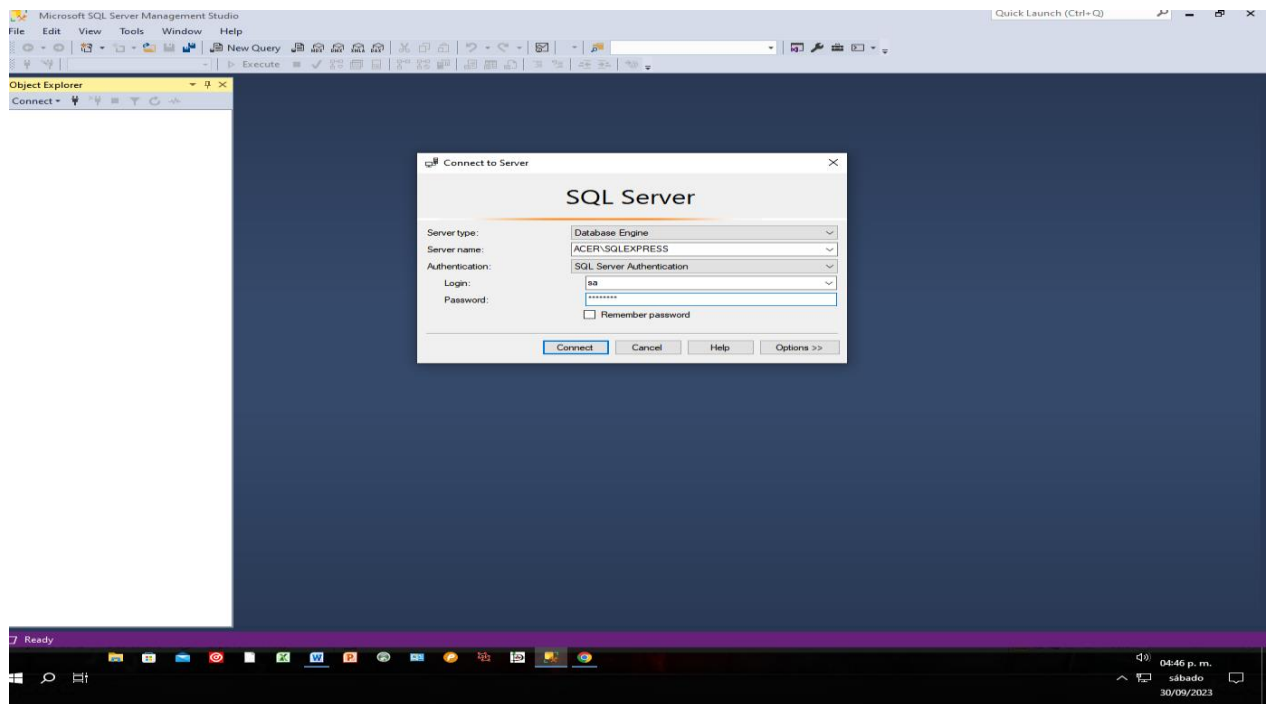


Una vez instalado y configurado SQL Server verificaos que el servidor se este ejecutando

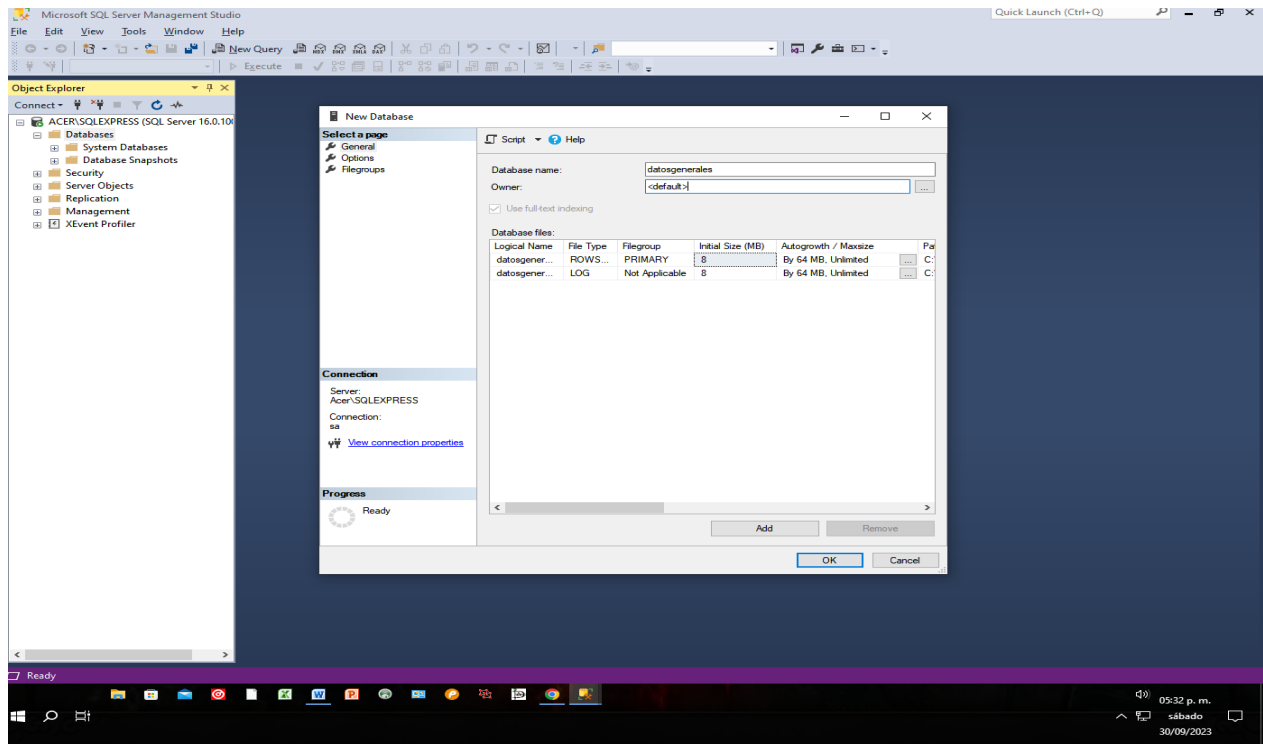




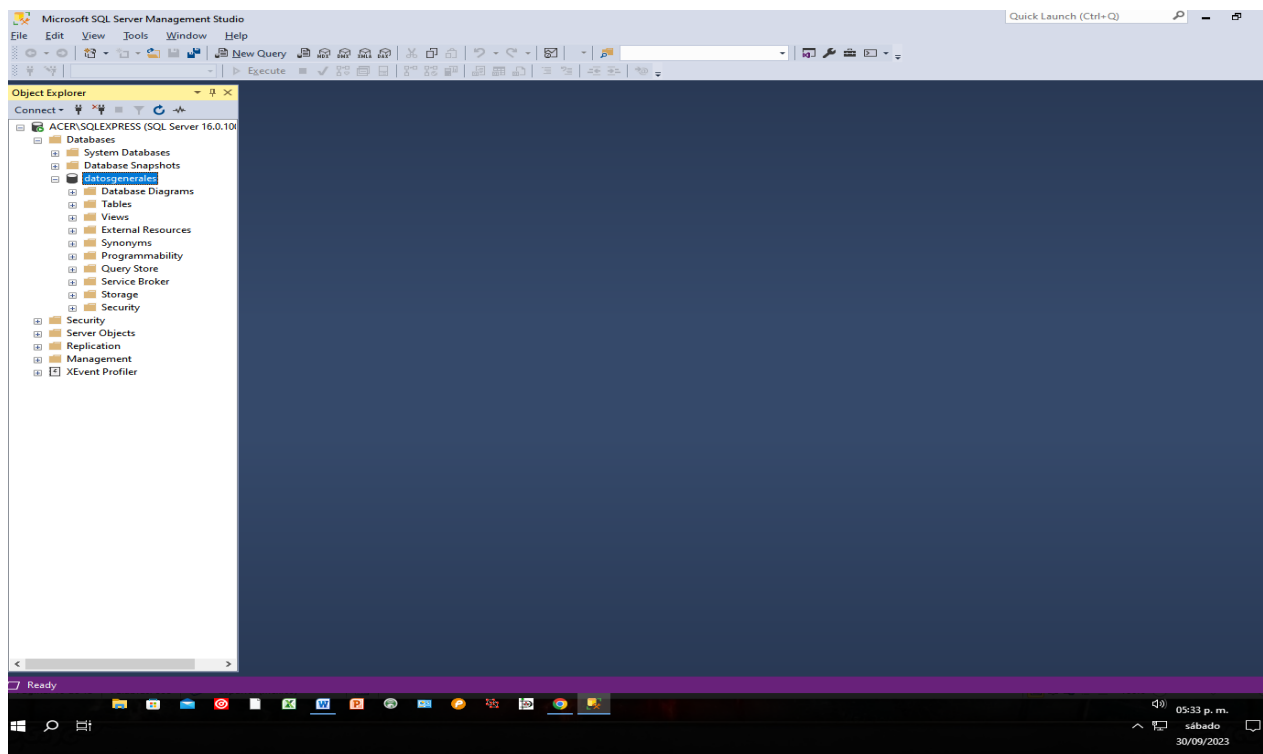
SQL Server debe estar corriendo

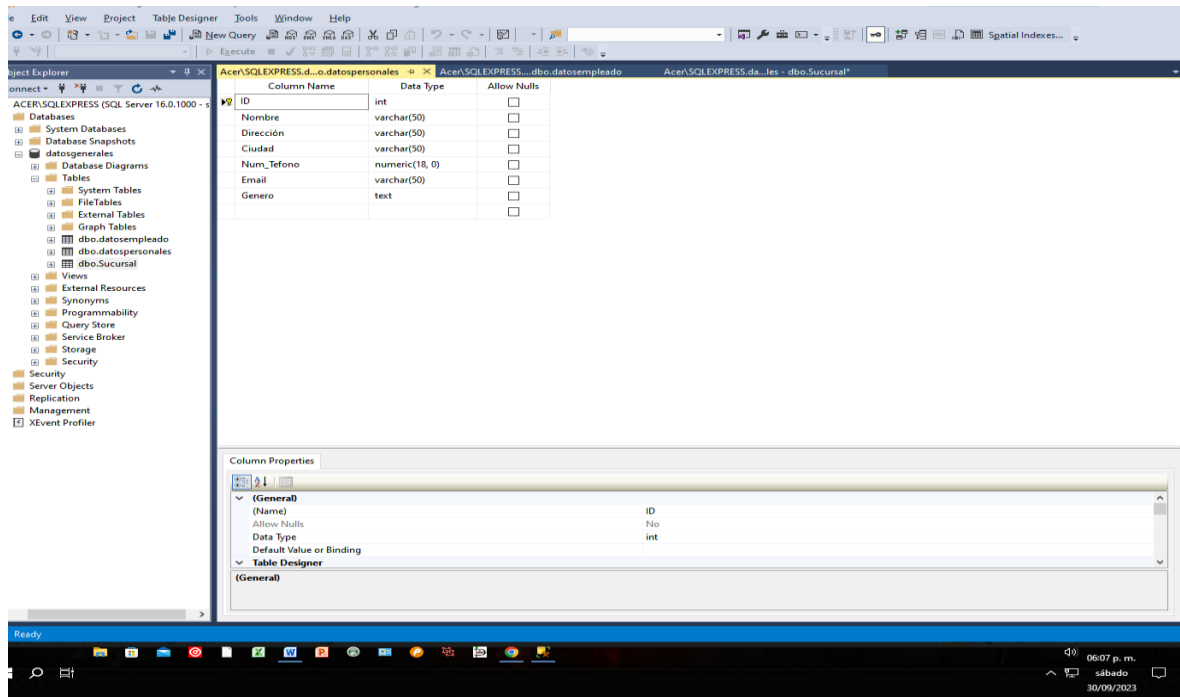


Autenticamos el ingreso al servidor.

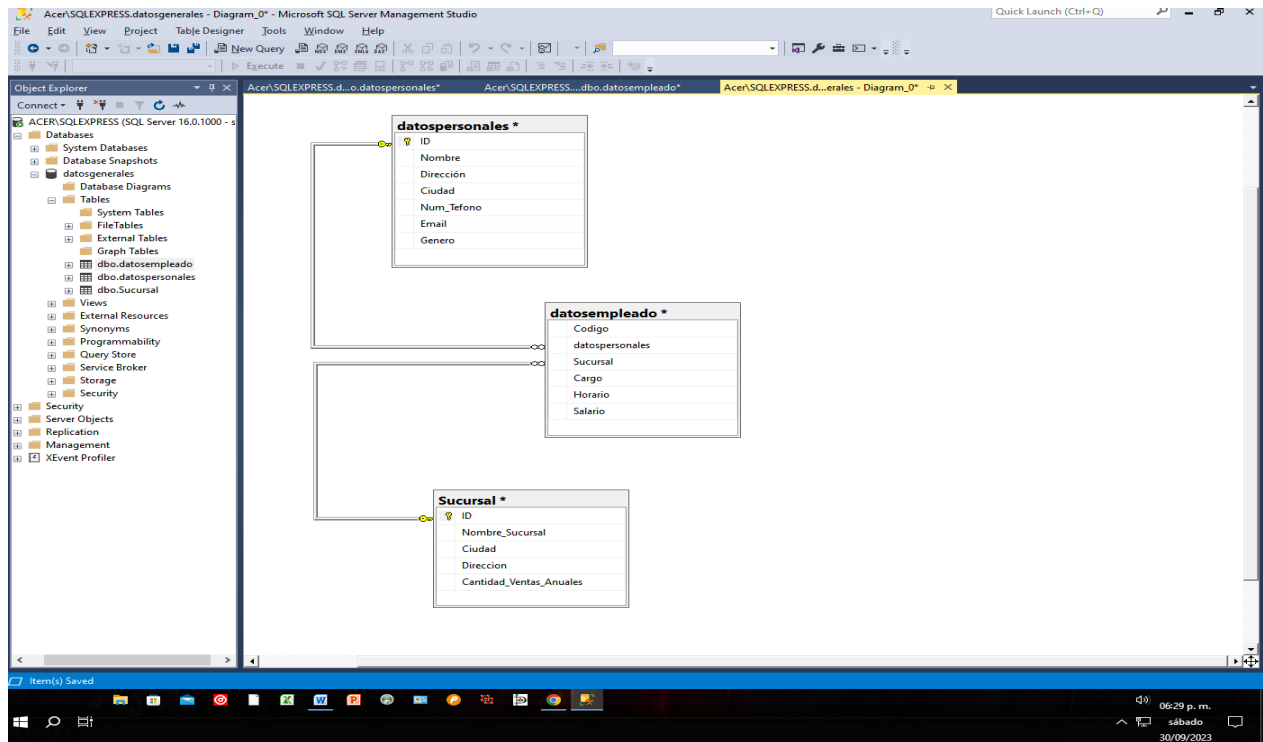


Se crea la base de datos como datosgenerales





Se crean las tablas solicitadas de datospersonales, datosempleado y Sucursal



Y las relacionamos

SQLQuery8.sql - ACER/SQLSERVER.datosgenerales (sa (58)) - Microsoft SQL Server Enterprise Manager

```

SELECT TOP (1000) [ID]
, [Nombre sucursal]
, [Ciudad]
, [Dirección]
, [Cantidad_ventas_anuales]
FROM [datosgenerales].[dbo].[SUCURSAL]

```

ID	Nombre sucursal	Ciudad	Dirección	Cantidad_ventas_anuales
1	Primavera	Apizaco	16 de Sept 3000	60000
2	Verano	Tlaxcala	Bvd Francisco I Madero 80	55000
3	Otoño	Santa Ana	Carr. Santa Ana 80	45000
4	Invierno	Tlaxco	Calle Principal 2	38000
5	Norte	Apizaco	Aquiles Serdan 4	62000
6	Sur	Zacatelco	Calle 1 #5	40000
7	Este	Huamantla	Mariano Matamoros 50	60000
8	Oeste	Huamantla	Rubies 46	58000
9	Luz	Contla	Pablo Picasso 85	28000
10	Noche	Nanacamilpa	Bvd Juárez 49	35000

Query executed successfully. ACER/SQLSERVER (16.0 RTM) sa (58) datosgenerales 00:00:00 10 rows

Se ingresan los datos en las tablas correspondientes

SQLQuery7.sql - ACER/SQLSERVER.datosgenerales (sa (56)) - Microsoft SQL Server Enterprise Manager

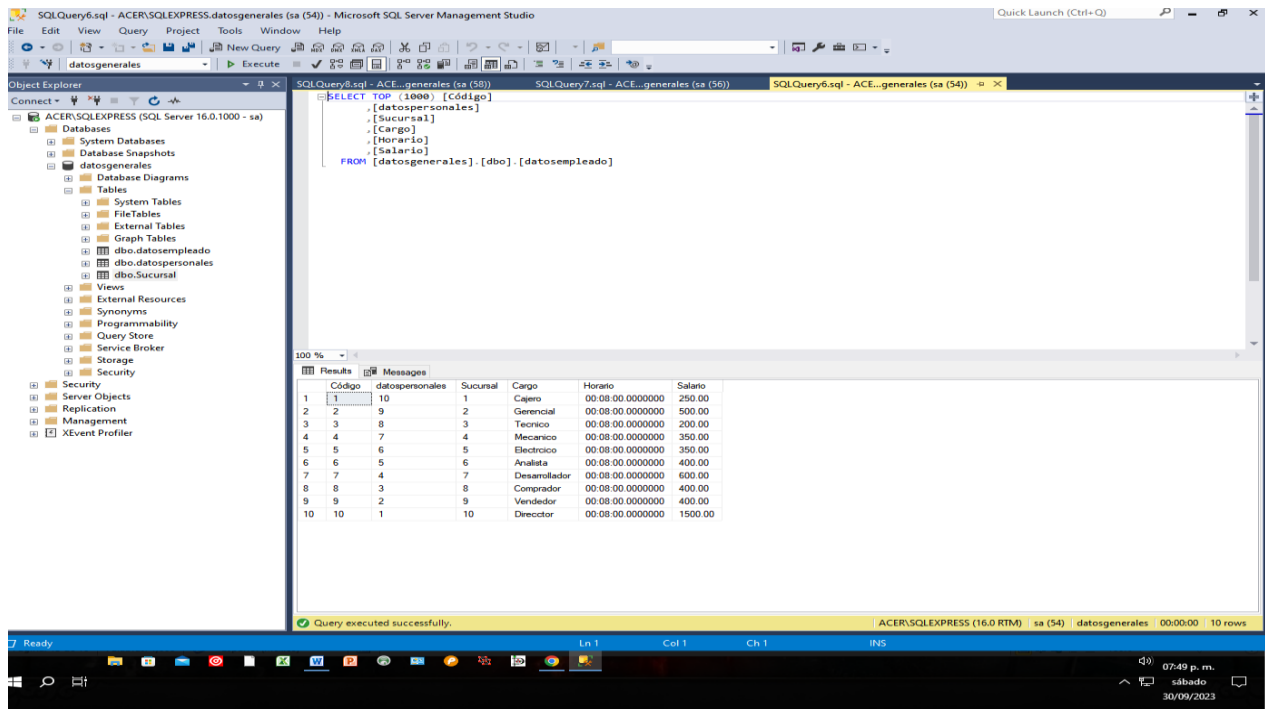
```

SELECT TOP (1000) [ID]
, [Nombre]
, [Dirección]
, [Ciudad]
, [Num_Teléfono]
, [Email]
, [Género]
FROM [datosgenerales].[dbo].[datospersonales]

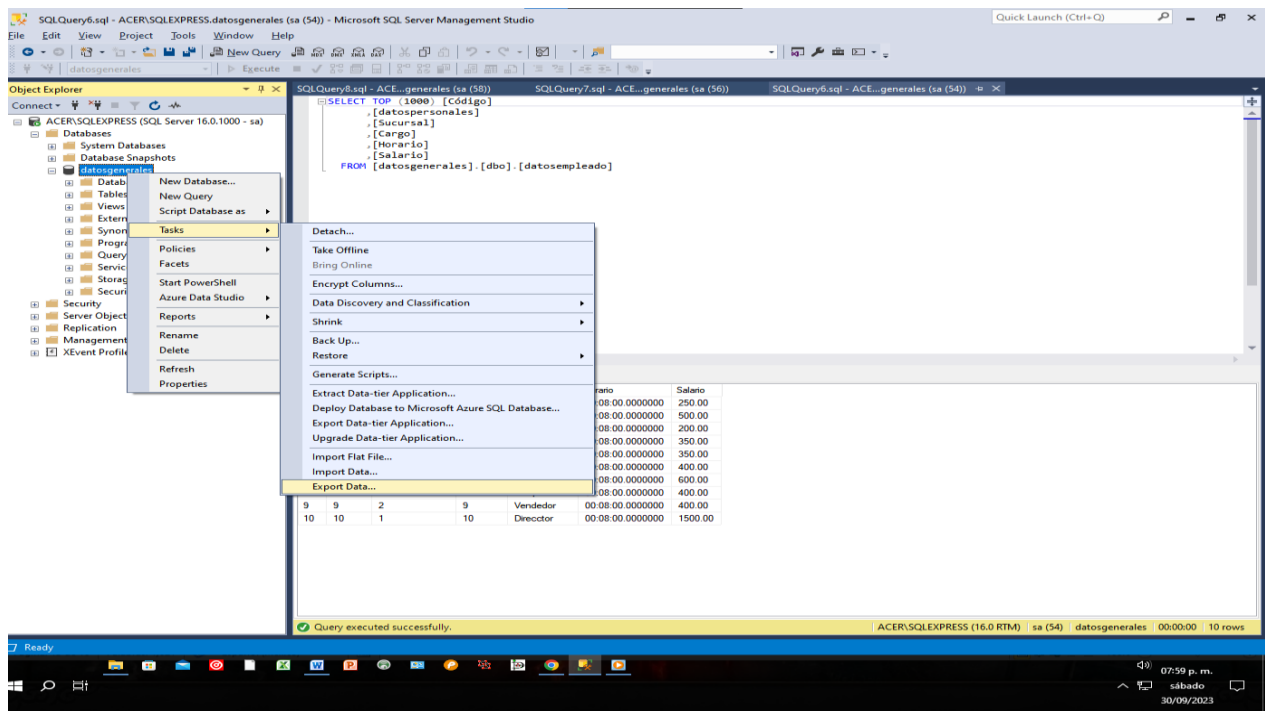
```

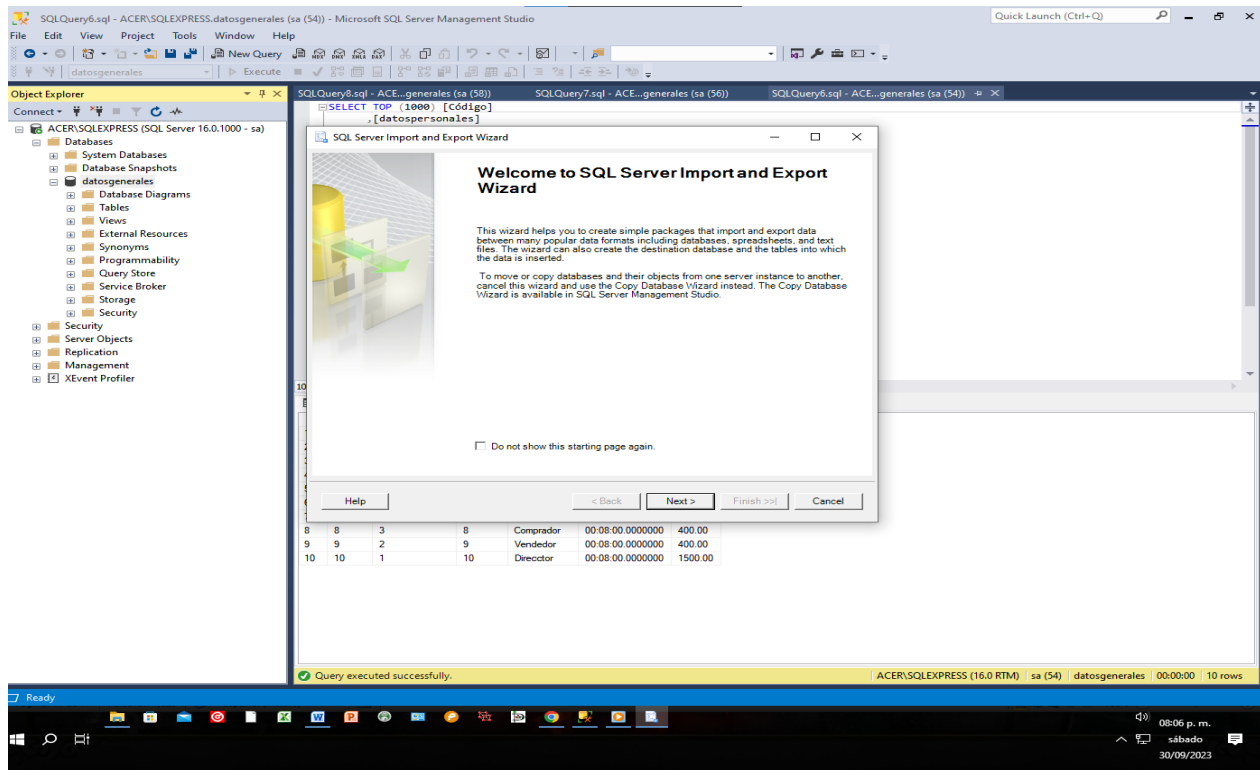
ID	Nombre	Dirección	Ciudad	Num_Teléfono	Email	Género
1	Fernando Pedraza	Calle La Luz # 50	Apizaco	1234567891	user@user1.com	Masculino
2	Laura Rivas	Abolides # 4	Tlaxcala	1234567892	user@user2.com	Femenino
3	Ramon Arenas	Francisco I Madero 86	San Miguel	1234567893	user@user3.com	Masculino
4	Silvia Jimenez	Alvaro Obregon 200	Santa Ana	1234567894	user@user4.com	Femenino
5	Rosa Castro	Jose Maria 22	Tlaxco	1234567895	usee@user5.com	Femenino
6	Ricardo Lozano	Alende 90	San Matias	1234567896	usee@user6.com	Masculino
7	Citlak Gomez	Bvd 16 de Sept 900	Puebla	1234567897	usee@user7.com	Femenino
8	Carlos Hernandez	Bvd Norte 35	Puebla	1234567898	usee@user8.com	Masculino
9	Ivonne Cid	Fracc Grasoales 26	Apizaco	1234567899	usee@user9.com	Femenino
10	Yoko Escalante	Fracc San Rafael 12	Tlaxcala	1234567890	usee@user10.com	Femenino

Query executed successfully. ACER/SQLSERVER (16.0 RTM) sa (56) datosgenerales 00:00:00 10 rows

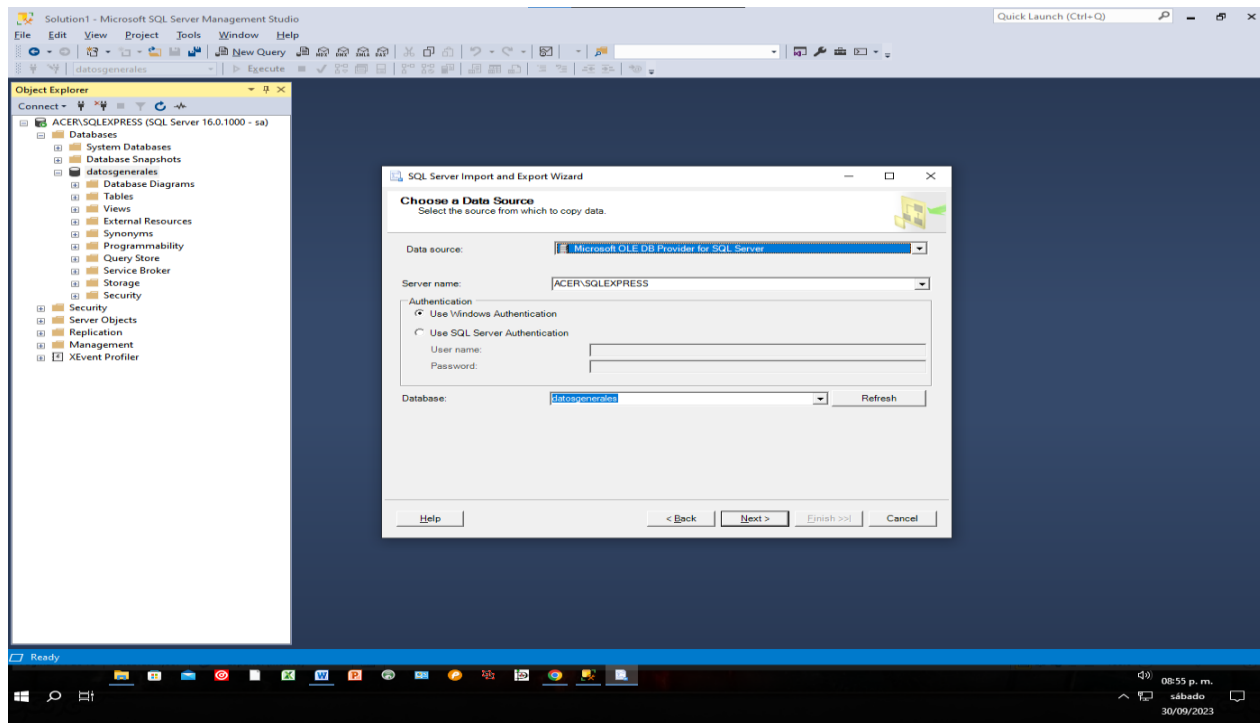


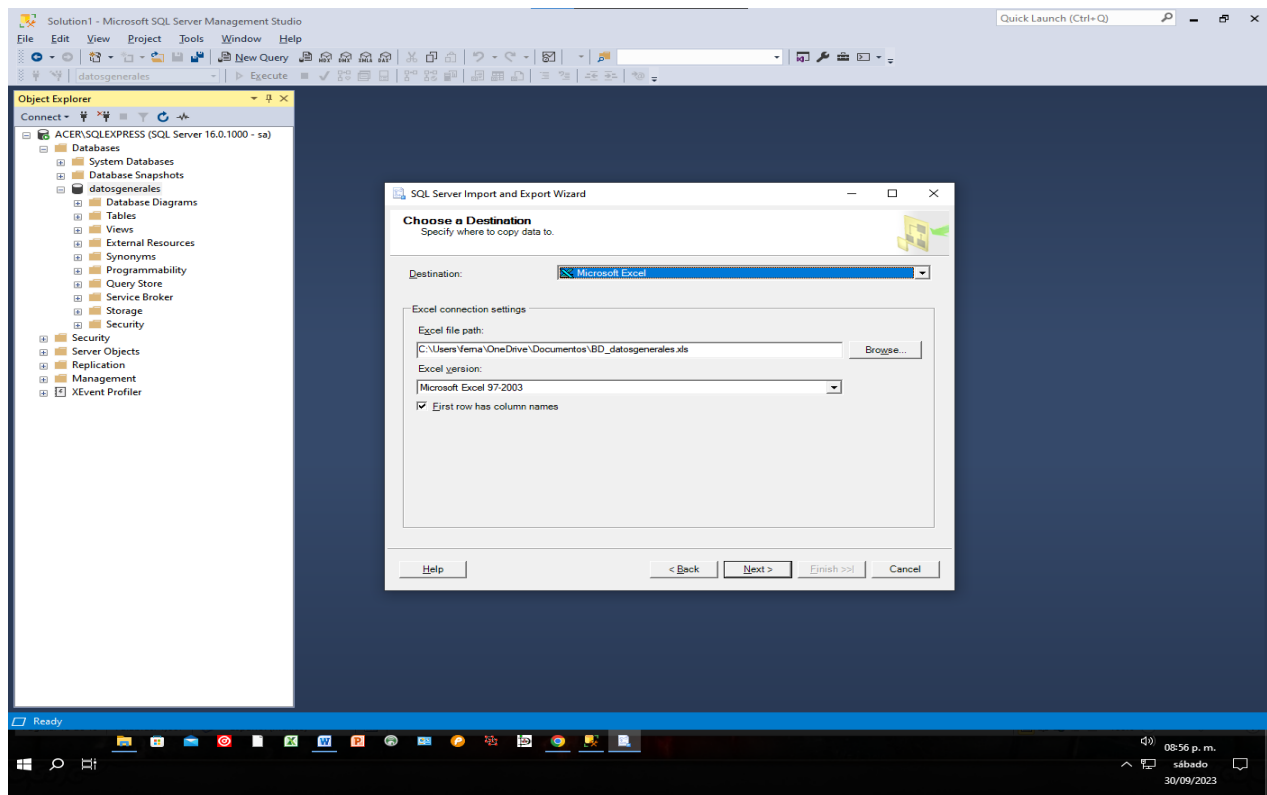
En base a lo solicitado se exporta la base de datos a Excel



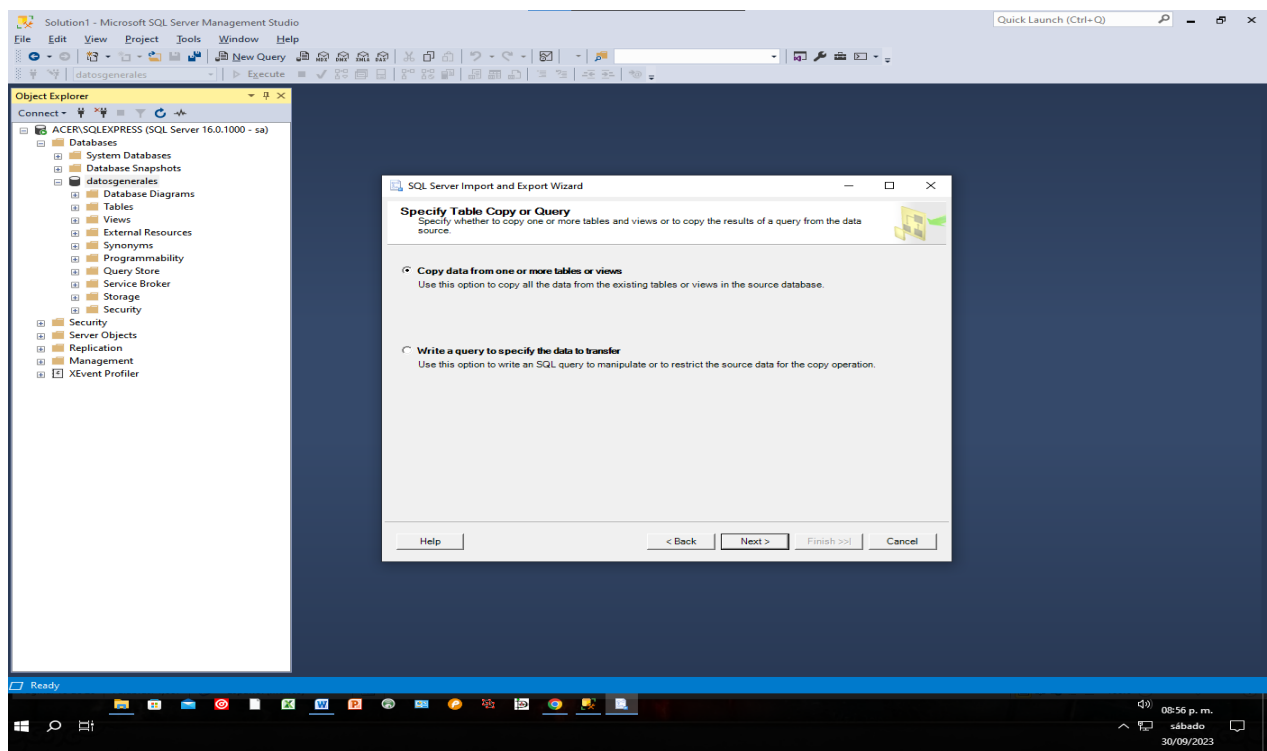


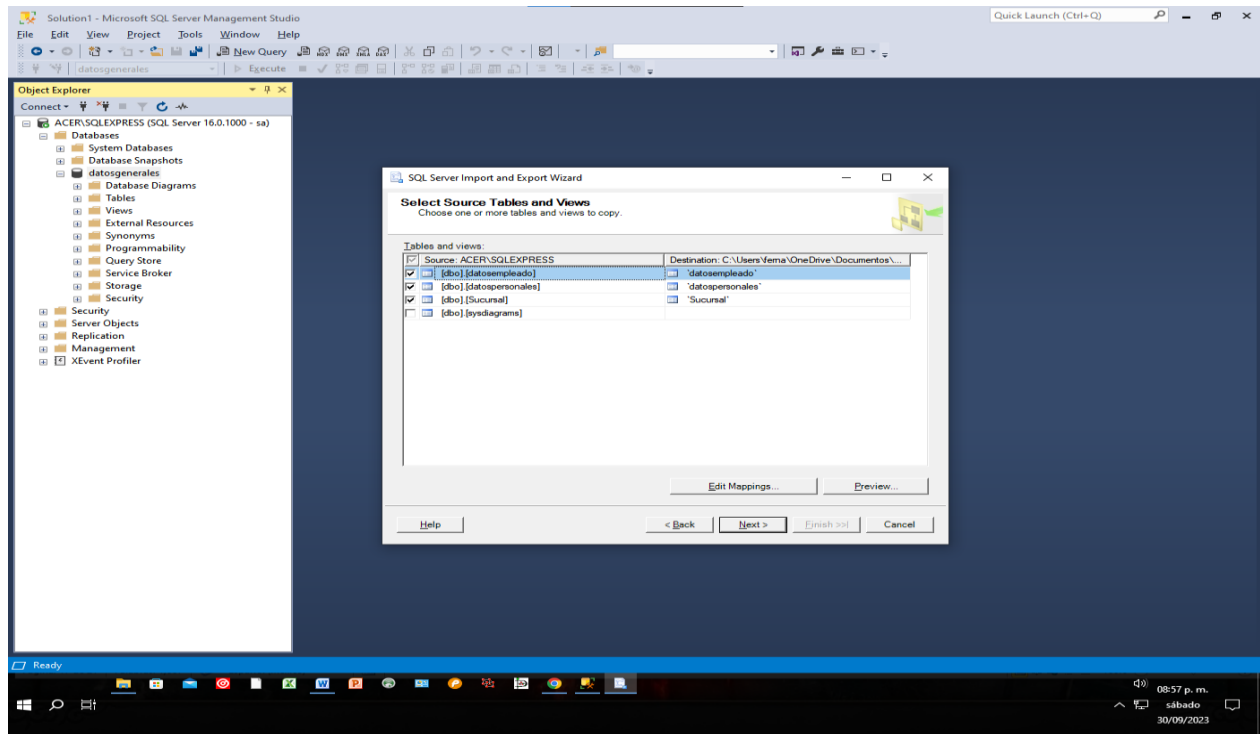
Aparece el asistente y se selecciona de donde se exportara la información



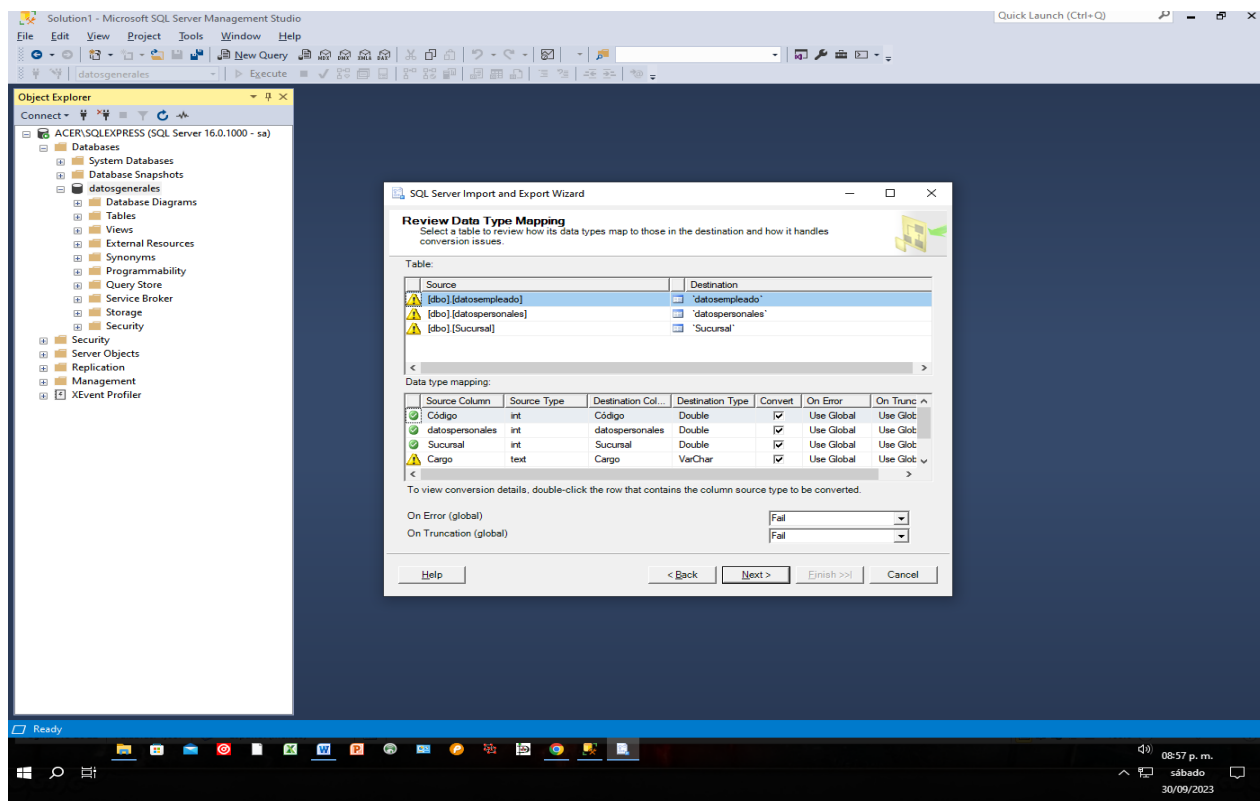


Se selecciona el formato en el que se guardara la información y la ruta destino

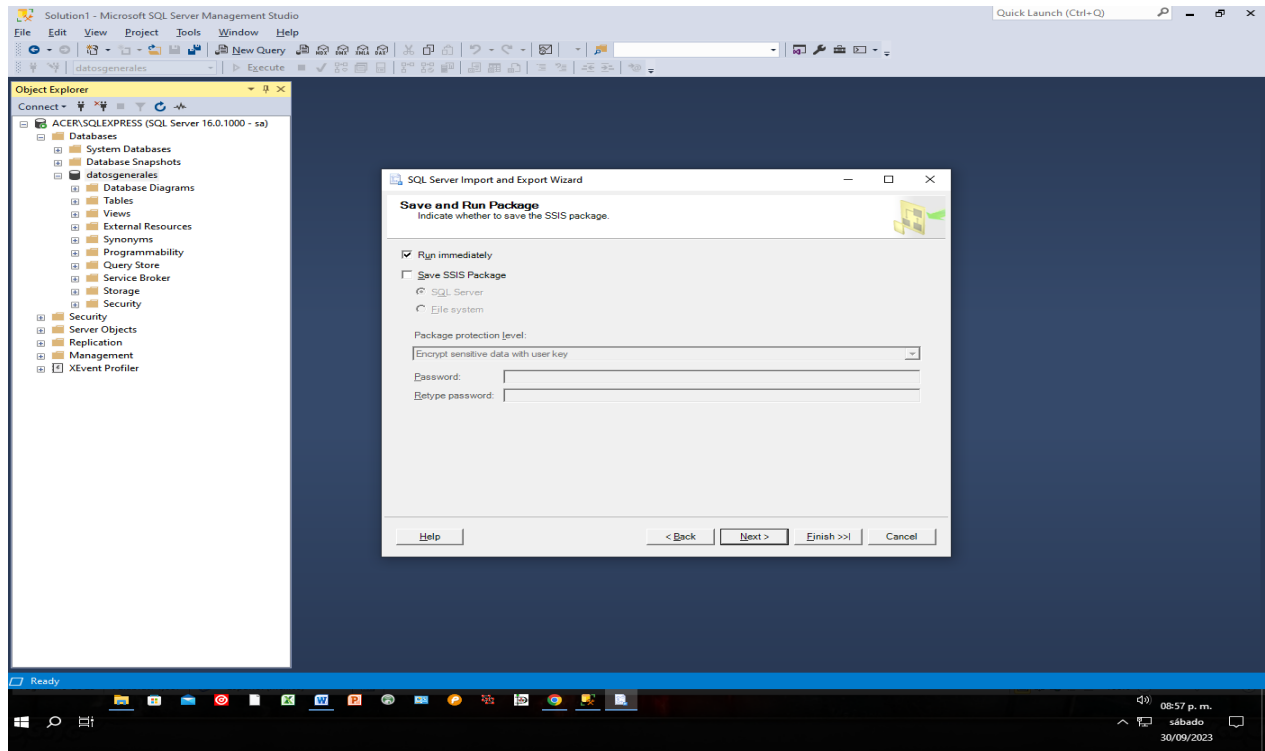




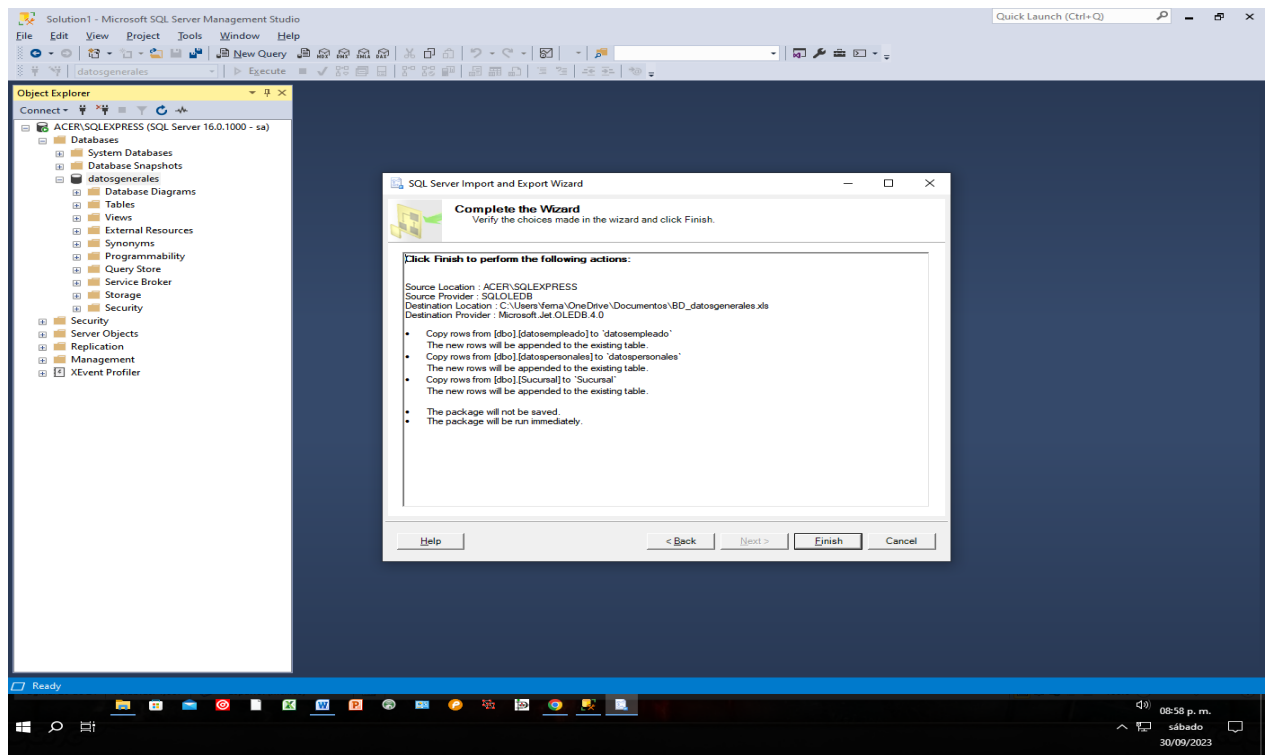
Se seleccionan las tablas a exportar

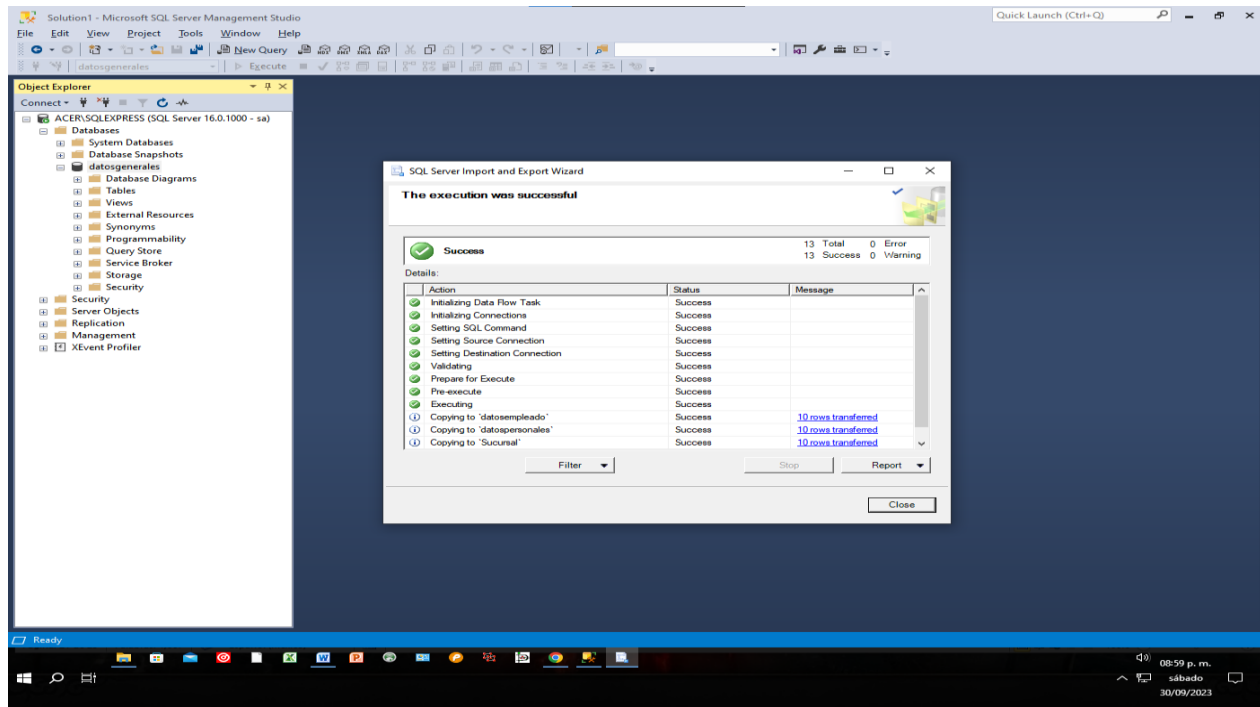




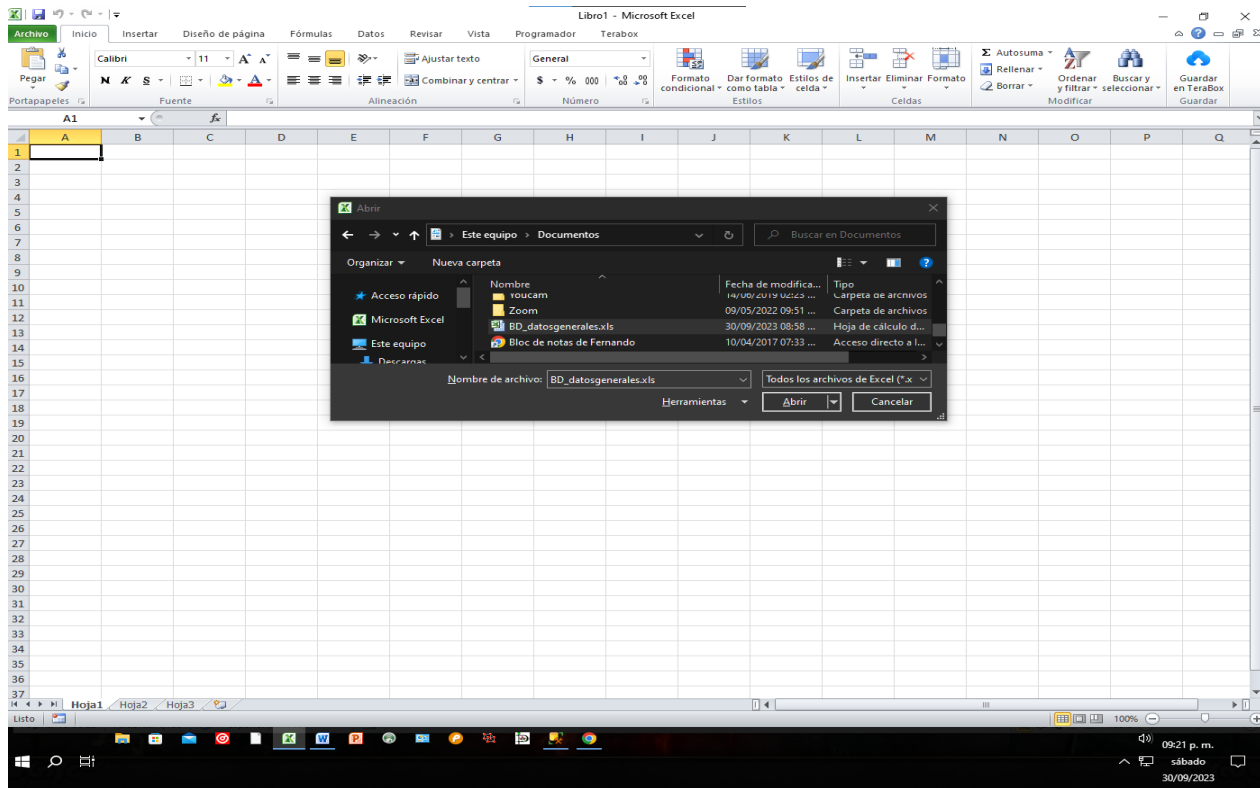


Se deja correr el proceso hasta que finalice





Abrimos Excel para abrir el archivo creado de la BD y visualizar las tablas



BD\_datosgenerales.xls [Modo de co

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista Programador Terabox

Tabla dinámica Tabla Imagen Imágenes prediseñadas Formas SmartArt Captura Columna Línea Circular Barra Área Dispersión Otros

Tablas Ilustraciones Gráficos

	A	B	C	D	E	F	G
1	ID	Nombre sucursal	Ciudad	Dirección	Cantidad_ventas_anuales		
2	1	Primavera	Apizaco	16 de Sept 3000	60000		
3	2	Verano	Tlaxcala	Blvd Francisco I Madero 80	55000		
4	3	Otoño	Santa Ana	Carr. Santa Ana 80	45000		
5	4	Invierno	Tlaxco	Calle Principal 2	38000		
6	5	Norte	Apizaco	Aquiles Serdan 4	62000		
7	6	Sur	Zacatelco	Calle 1 #5	40000		
8	7	Este	Huamantla	Mariano Matamoros 50	60000		
9	8	Oeste	Huamantla	Robles 46	58000		
10	9	Luz	Contla	Pablo Picasso 85	28000		
11	10	Noche	Nanacamilpa	Blvd Juarez 49	35000		
12							

BD\_datosgenerales.xls [M

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista Programador

Tabla dinámica Tabla Imagen Imágenes prediseñadas Formas SmartArt Captura Columna Línea Circular Barra Área Dispersión

Tablas Ilustraciones Gráficos

	A	B	C	D	E	F	G	H
1	Código	datospersonales	Sucursal	Cargo	Horario	Salario		
2	1	10	1	Cajero	00:08:00.0000000	\$250.00		
3	2	9	2	Gerencial	00:08:00.0000000	\$500.00		
4	3	8	3	Tecnico	00:08:00.0000000	\$200.00		
5	4	7	4	Mecanico	00:08:00.0000000	\$350.00		
6	5	6	5	Electricico	00:08:00.0000000	\$350.00		
7	6	5	6	Analista	00:08:00.0000000	\$400.00		
8	7	4	7	Desarrollador	00:08:00.0000000	\$600.00		
9	8	3	8	Comprador	00:08:00.0000000	\$400.00		
10	9	2	9	Vendedor	00:08:00.0000000	\$400.00		
11	10	1	10	Direcctor	00:08:00.0000000	\$1,500.00		
12								
13								
14								

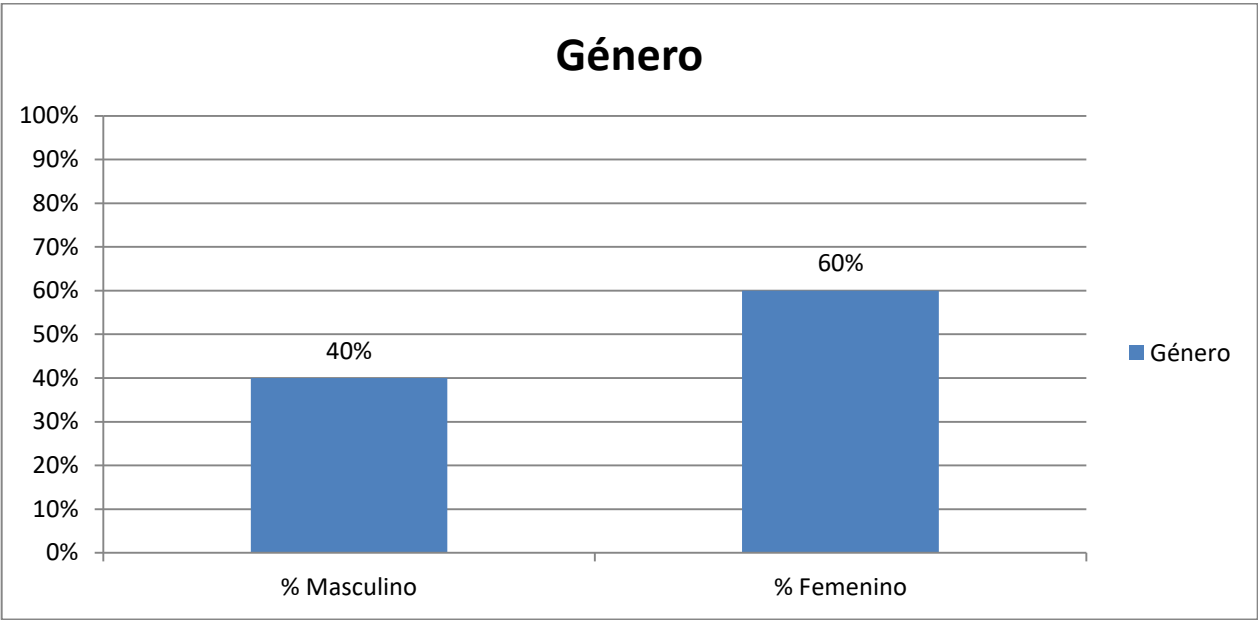
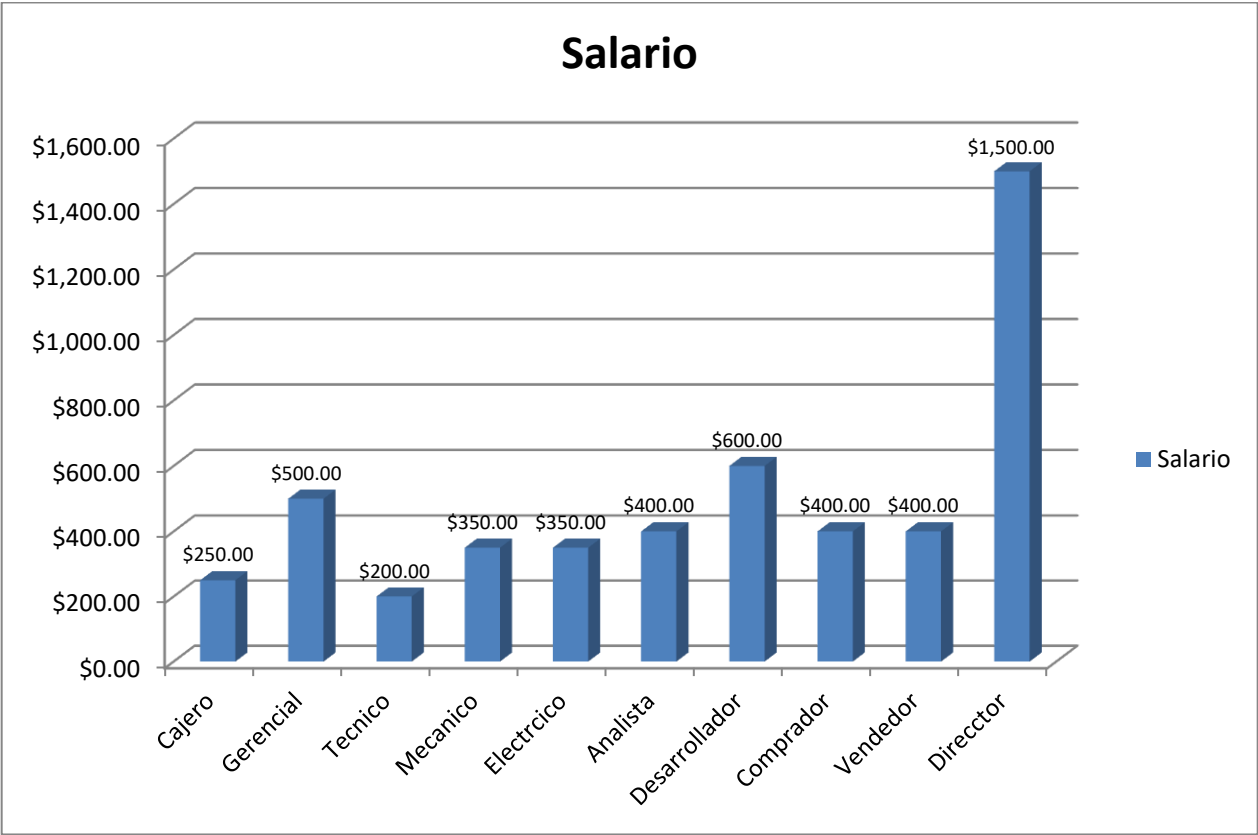
BD\_datosgenerales.xls [Modo de co

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista Programador Terabox

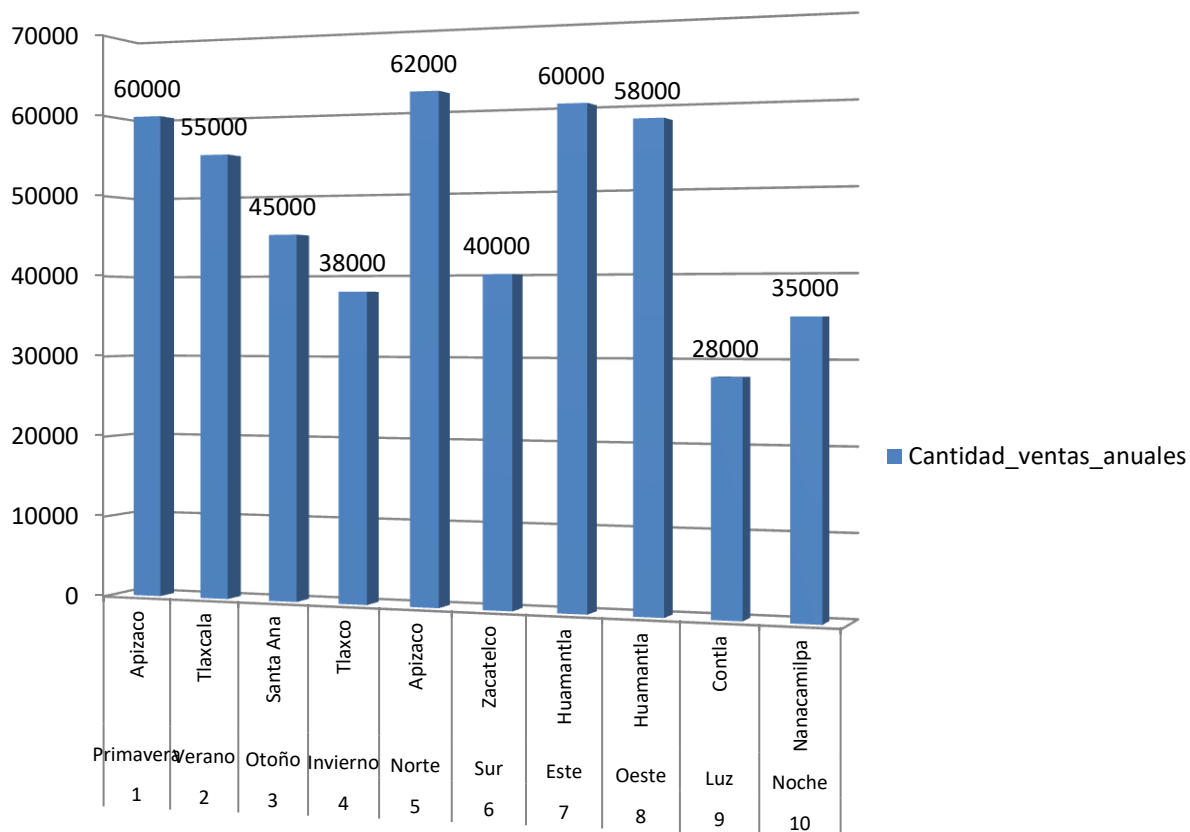
Tabla dinámica Tabla Imagen Imágenes prediseñadas Formas SmartArt Captura Columna Línea Circular Barra Área Dispersión Otros

Tablas Ilustraciones Gráficos

	A	B	C	D	E	F	G
1	ID	Nombre sucursal	Ciudad	Dirección	Cantidad_ventas_anuales		
2	1	Primavera	Apizaco	16 de Sept 3000	60000		
3	2	Verano	Tlaxcala	Blvd Francisco I Madero 80	55000		
4	3	Otoño	Santa Ana	Carr. Santa Ana 80	45000		
5	4	Invierno	Tlaxco	Calle Principal 2	38000		
6	5	Norte	Apizaco	Aquiles Serdan 4	62000		
7	6	Sur	Zacatelco	Calle 1 #5	40000		
8	7	Este	Huamantla	Mariano Matamoros 50	60000		
9	8	Oeste	Huamantla	Robles 46	58000		
10	9	Luz	Contla	Pablo Picasso 85	28000		
11	10	Noche	Nanacamilpa	Blvd Juarez 49	35000		
12							
13							



## Cantidad\_ventas\_anuales



¿Quién es la persona que gana más?

El director es la persona que tiene el mayor ingreso

¿Cuál es la sucursal que gana menos anualmente?

El personal técnico es quien tiene el menor ingreso

¿Cuántos empleados son hombres?

Solamente el 40%

# Conclusión.

---

Al realizar este tipo de actividades y generar historiales de comportamiento se permite crear bases de datos, que en la minería de datos a futuro ayudan, de forma representativa en la toma de decisiones, basados en los resultados presentados de forma gráfica y sencilla, y al concentrar la información en la base de datos se crean alternativas de funcionamiento, innovación, desarrollo, crédito financiero o simplemente estadístico, para el mantenimiento y monitoreo del buen funcionamiento de procesos, equipos, conductas, inclusive de recurso humano, mostrando las áreas de oportunidad existentes para mejorar, desarrollar o corregir dichos patrones existentes, evitando afectar de forma irreparable los resultados esperados.

¿Qué aprendo?

Como se puede uno beneficiar al tomar decisiones con objetividad, basado en hechos históricos reales por medio de los patrones creados con el tiempo.

Enlace Git-Hub: <https://github.com/Chifer888/Mineria-de-Datos.git>

# Referencias

---

Agencia, F. (2021). Descubrimiento de Conocimiento en Bases de Datos – KDD. *Salud Electrónica*. <https://saludelectronica.com/descubrimiento-de-conocimiento-en-bases-de-datos-kdd/>

Sergio Alejandro Campos - EXCELeINFO. (2021, 9 abril). *Cómo instalar y configurar SQL SERVER paso a paso* [Vídeo]. YouTube.  
<https://www.youtube.com/watch?v=mA1qoWdNCOE>

*Descargas de SQL Server / Microsoft*. (s. f.). <https://www.microsoft.com/es-mx/sql-server/sql-server-downloads?rtc=1>

Franklin García. (2020, 17 febrero). *Como crear BASE DE DATOS en SQL SERVER desde cero* 😊 [Vídeo]. YouTube. <https://www.youtube.com/watch?v=fyvEhDgKl7E>

Franklin García. (2020b, febrero 25). ● *Crear TABLAS en SQL Server con y sin CODIGO* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=rHcB0KGoysw>



Emanuel De Jesus Santana. (2021, 23 octubre). *Solventar error al exportar una base de datos a Excel* [Video]. YouTube. <https://www.youtube.com/watch?v=abuQ4jjNTAM>

jotajotavm. (2017, 22 febrero). *SQL Server / 51 - Exportar a Excel desde SQL Server* [Video]. YouTube. [https://www.youtube.com/watch?v=-tPfgJ\\_tc5U](https://www.youtube.com/watch?v=-tPfgJ_tc5U)

Mas que solo Excel. (2021, 7 marzo). 🔥📊 *GRAFICAR SEXO MASCULINO y FEMENINO EN EXCEL* 📊🔥 [Video]. YouTube. <https://www.youtube.com/watch?v=xglBeVN1sno>