

Actividad 1 - Cubos OLAP

Minería y Análisis de datos

Ingeniería en Desarrollo de Software

Tutor: Félix Acosta Hernández

Alumno: José Domingo Reyes Arroyo

Fecha: 04 de octubre de 2023

Índice

Índice.....	2
1 Introducción.....	3
2 Descripción.....	4
3 Justificación.....	4
4 Desarrollo:	5
4.1 Creación de la Base de Datos.....	6
4.2 Importación de la Base de Datos.....	13
4.3 Gráficos	18
4.4 Preguntas	20
5 Conclusión.....	23
Referencias.....	24

1 Introducción

La Minería de Datos es un término que se fijó en la década de 1990, la definición sencilla de este termino es considerada como el proceso de hurgar en los datos con la finalidad de encontrar anomalías, patrones y correlaciones en una gran cantidad de conjunto de datos proporcionada por una Base de Datos, este proceso de minería de los datos tiene el enfoque en los beneficios de la organización como la incrementación de los ingresos de la empresa, disminuir costos, mejorar la relación con los clientes, reducir riesgos, entre otros. La minería de datos comprende 3 grandes disciplinas entrelazadas, estas son la Estadística, la Inteligencia Artificial y Machine Learning (Análisis Predictivo).

En base a la minería de datos, existe un análisis multidimensional donde los datos se representan como dimensiones, las cuales se relacionan en jerarquías, un ejemplo de ello es el tiempo el cual también se puede jerarquizar en días, semanas, meses y años, etc. Así es como se puede dar una imagen multidimensional a esta información de acuerdo a los datos obtenidos en alguna medición que tenga correlación con el factor del tiempo, a este análisis multidimensional se le denomina Cubos de OLAP.

2 Descripción

Los cubos de OLAP es un tipo de análisis que nos permite tener múltiples dimensiones como su nombre lo indica, para obtener una idea mas concreta de lo que son los cubos OLAP y su función dentro de una Base de Datos multidimensional, cabe destacar que cada una de las dimensiones del cubo corresponde a una Jerarquía dentro de la Base de Datos, así pues con el uso de los cubos OLAP se amplían las posibilidades de las Bases de Datos relacionales, con lo cual se permite el procesamiento de una gran cantidad de información que no seria posible analizar sin el uso de este análisis multidimensional.

La aplicación de los cubos OLAP se considera como uno de los pilares que permiten o dan solución a la inteligencia de negocios, por la razón de que provee de información concreta a quienes están a cargo de la toma de decisiones, lo cual pueden hacer mediante métodos de navegación de los datos que les permitirá analizar y mantener una conversación fluida con los datos de que proporciona el sistema de la organización.

3 Justificación

Con base en lo observado de la minería de los datos y del Cubo OLAP, podemos definir a estos procesos como la gestión de una gran cantidad de información almacenada en las

bases de datos de la organización con la finalidad de obtener un análisis de dicha información enfocándola a obtener beneficios para la organización como los beneficios mencionados anteriormente, es por ello de suma importancia conocer el proceso de análisis o Minería de los datos para garantizar el alcance de dichos beneficios para la organización, además de ello se logra para cualquier organización tener un mejor control y un orden eficiente en la administración de sus Bases de Datos para la organización de la información.

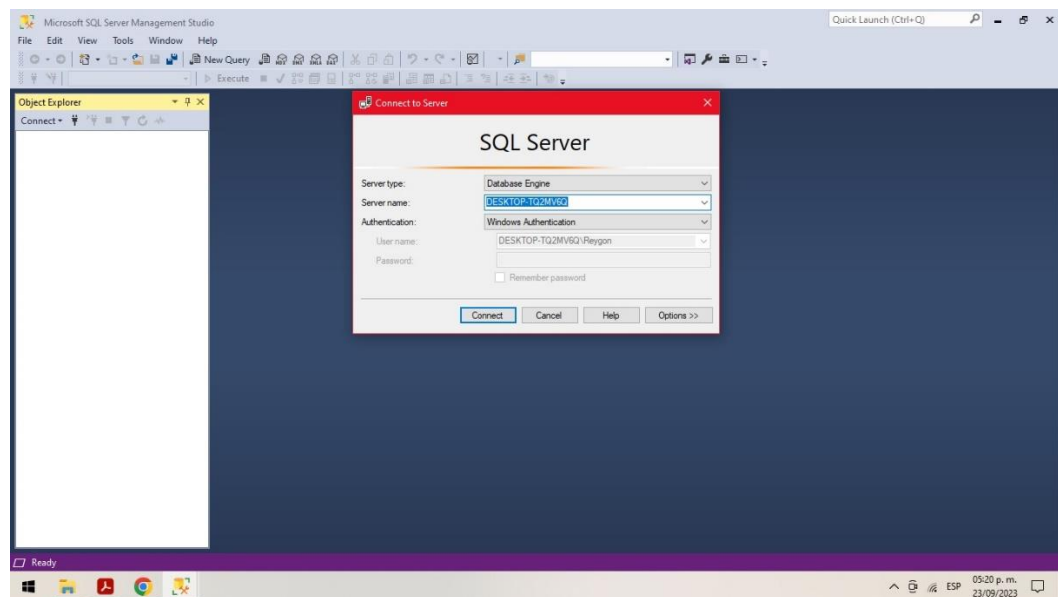
Por lo antes mencionado es de suma importancia el análisis y la minería de datos para cualquier organización que tenga una base de datos para administrar la información de sus clientes, productos, servicios, ingresos, egresos, etc. La implementación de esta tecnología de minería de datos, como los cubos OLAP, permite la administración, consolidación, profundización y la navegación en los datos provistos por la organización a través de las Bases de Datos.

4 Desarrollo:

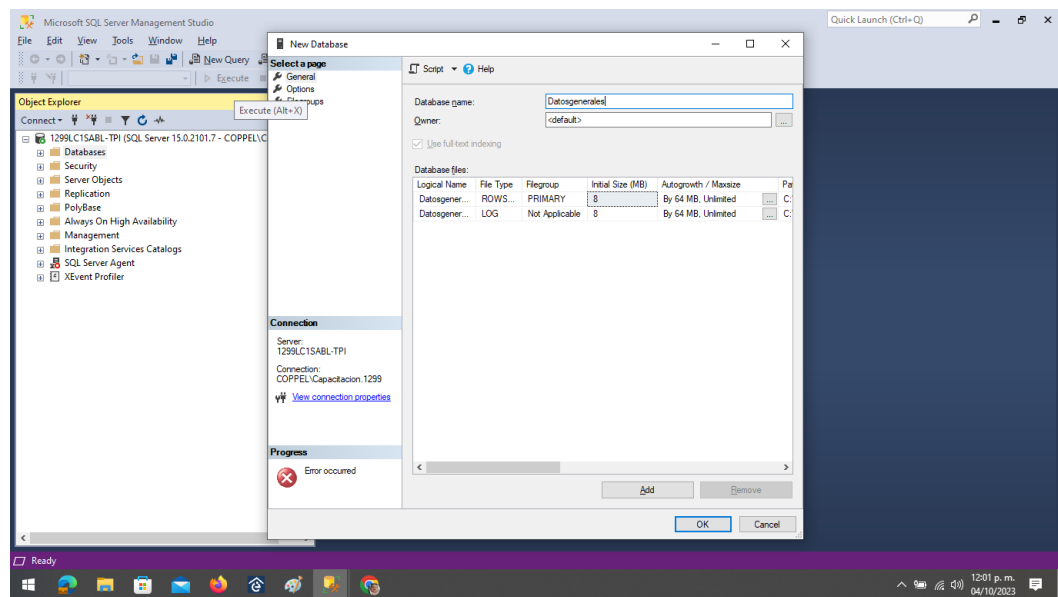
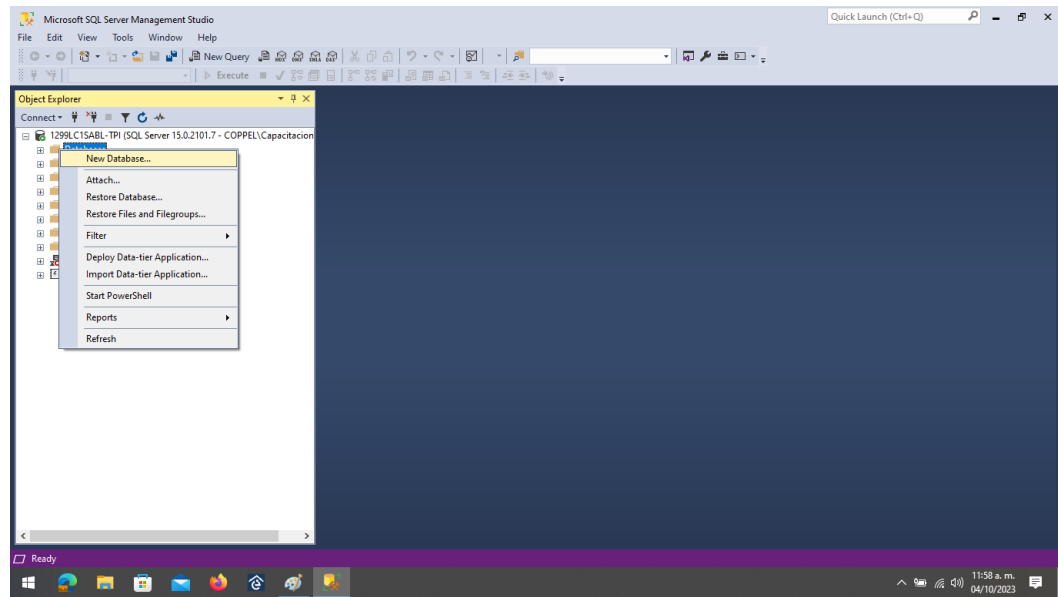
A continuación, se describe el proceso de aprendizaje que se llevo a cabo para administrar la información contenida en una Base de Datos:

4.1 Creación de la Base de Datos

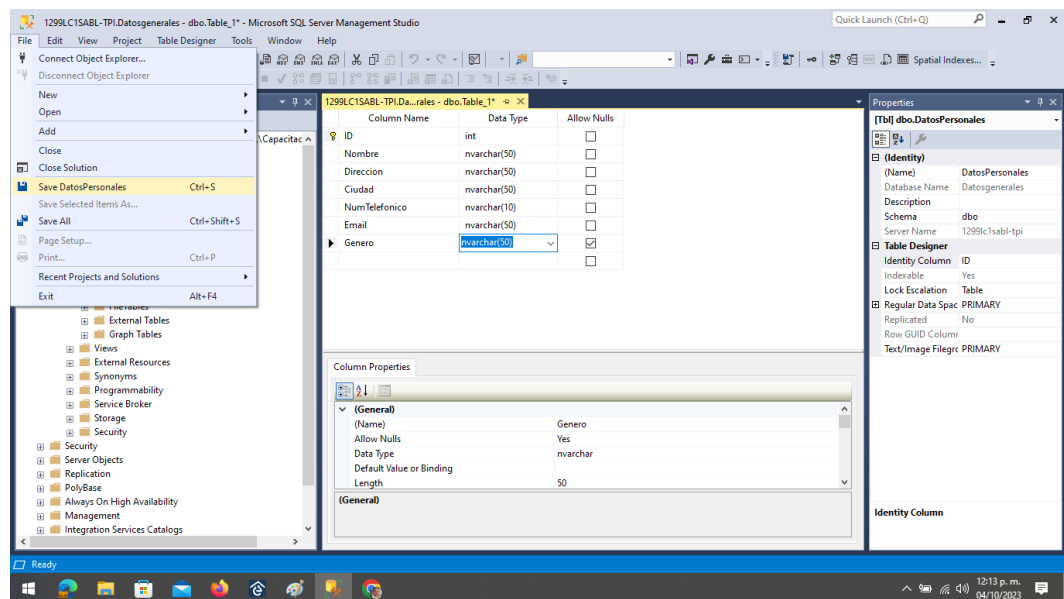
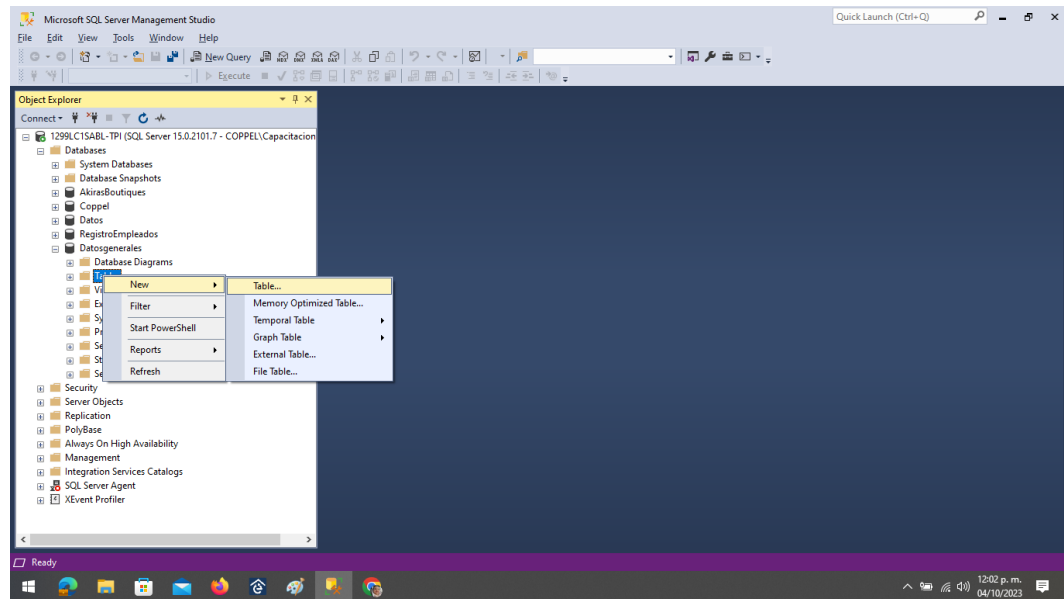
El proceso de creación de la base de datos conlleva la instalación del Gestor de Base de Datos MySQL Server, el cual ya se ha utilizado en actividades y materias anteriores por lo que no se requiere la explicación de la instalación del gestor, a continuación, se muestra el acceso al gestor y su ventana principal donde se genera la conexión con el servidor:

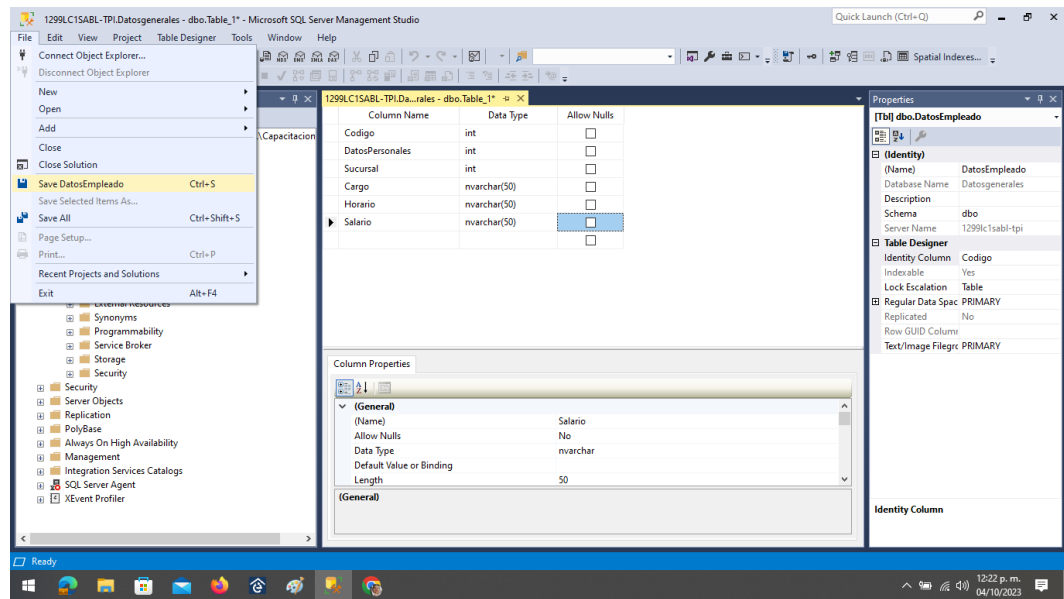
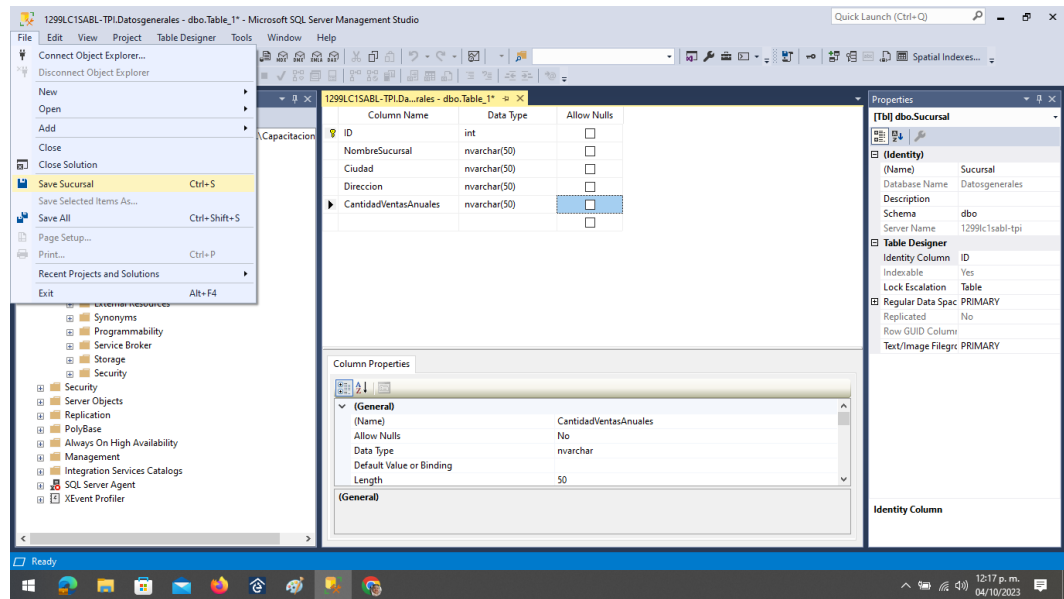


Una vez realizada la conexión con el servidor se procede a la creación de la base de datos:

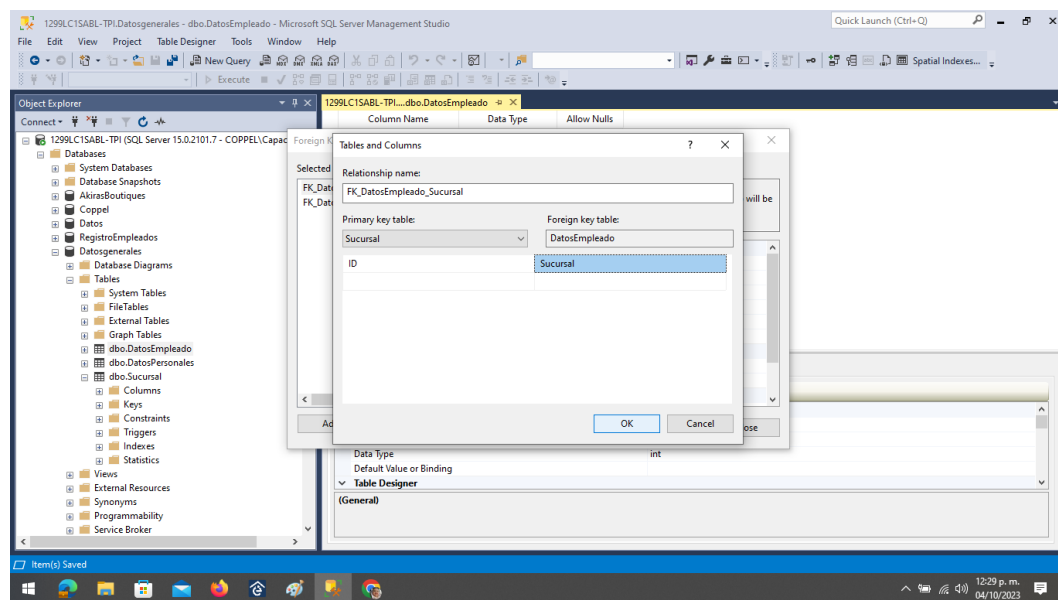
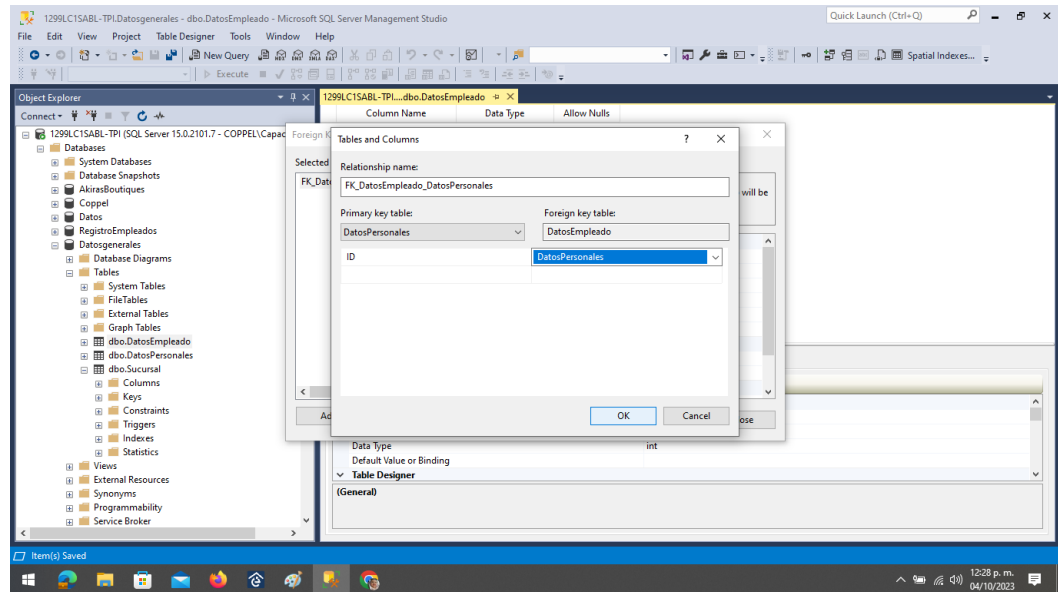


Posteriormente se crean las tablas correspondientes:

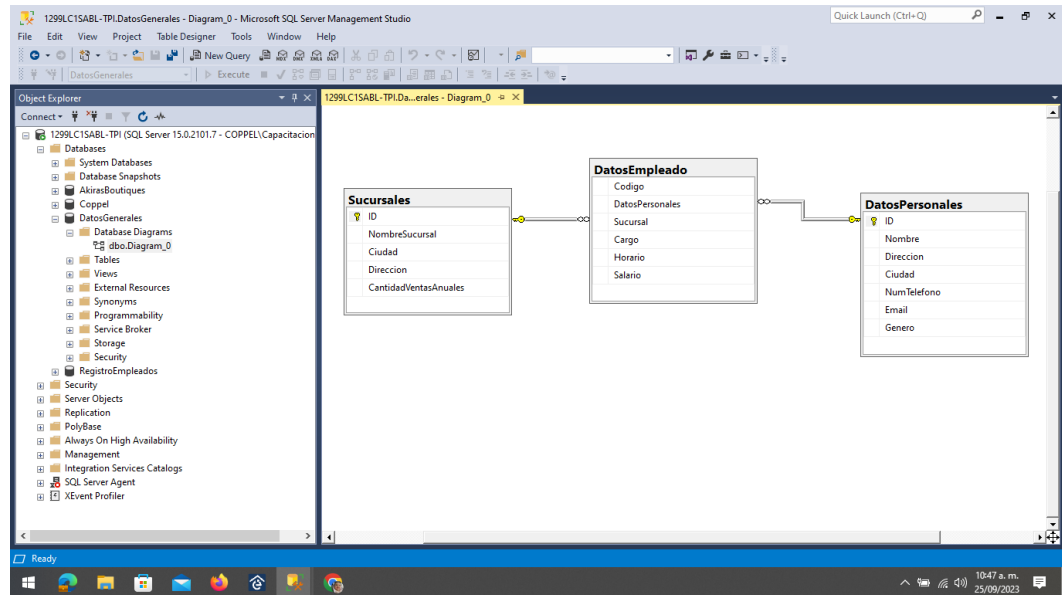




Después se crean las relaciones entre las llaves primarias y foráneas:



Se valida en el diagrama que estén establecidas las relaciones:



Y posteriormente se llenan los datos de las diferentes tablas, para lo cual solo se mostraran los datos registrados en cada tabla:

SQLQuery2.sql - 1299LC1SABL-TPI-DatosGenerales (COPPEL\Capacitacion.1299 (51)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

Connect +

1299LC1SABL-TPI (SQL Server 15.0.2101.7 - COPPEL\Capacitacion.1299 (51))

Databases

- System Databases
- Database Snapshots
- AkirasBoutiques
- Coppel
- DatosGenerales
 - Database Diagrams
 - dbo.Diagram_0
 - Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.DatosEmpleado
 - Columns
 - Keys
 - Constraints
 - Indexes
 - Statistics
 - dbo.DatosPersonales
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Service Broker
 - Storage

SQLQuery2.sql - 1299LC1SABL-TPI-DatosGenerales (COPPEL\Capacitacion.1299 (51))

SELECT * FROM [DatosGenerales].[dbo].[DatosEmpleado]

Results

	Codigo	DatosPersonales	Sucursal	Cargo	Horario	Salario
1	1	1	1	Asesor	9:00 a 19:00	7800
2	2	2	2	Asesor	9:00 a 19:30	7800
3	3	3	3	Gerente	9:00 a 20:00	20000
4	4	4	4	Tecnico	9:00 a 19:30	15000
5	5	5	5	Cajera	9:00 a 19:30	7800
6	6	6	6	Auxiliar de piso	9:00 a 20:00	7800
7	7	7	7	Asesor	9:00 a 20:00	7800
8	8	8	8	Gerente	9:00 a 20:00	20700
9	9	9	9	Promotor	9:00 a 20:00	7800
10	10	10	10	Cajero mixto	9:00 a 20:00	7800

Query executed successfully. 1299LC1SABL-TPI (15.0 RTM) COPPEL\Capacitacion.12... DatosGenerales 00:00:00 10 rows

SQLQuery2.sql - 1299LC1SABL-TPI-DatosGenerales (COPPEL\Capacitacion.1299 (51)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

Connect +

1299LC1SABL-TPI (SQL Server 15.0.2101.7 - COPPEL\Capacitacion.1299 (51))

Databases

- System Databases
- Database Snapshots
- AkirasBoutiques
- Coppel
- DatosGenerales
 - Database Diagrams
 - dbo.Diagram_0
 - Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.DatosEmpleado
 - Columns
 - Keys
 - Constraints
 - Indexes
 - Statistics
 - dbo.DatosPersonales
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Service Broker
 - Storage

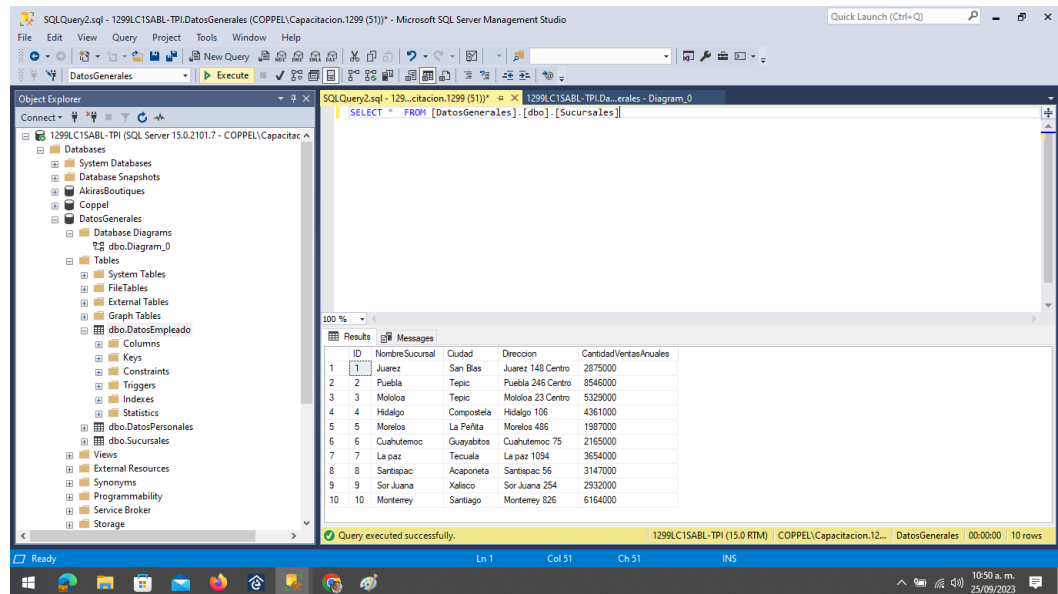
SQLQuery2.sql - 1299LC1SABL-TPI-DatosGenerales (COPPEL\Capacitacion.1299 (51))

SELECT * FROM [DatosGenerales].[dbo].[DatosPersonales]

Results

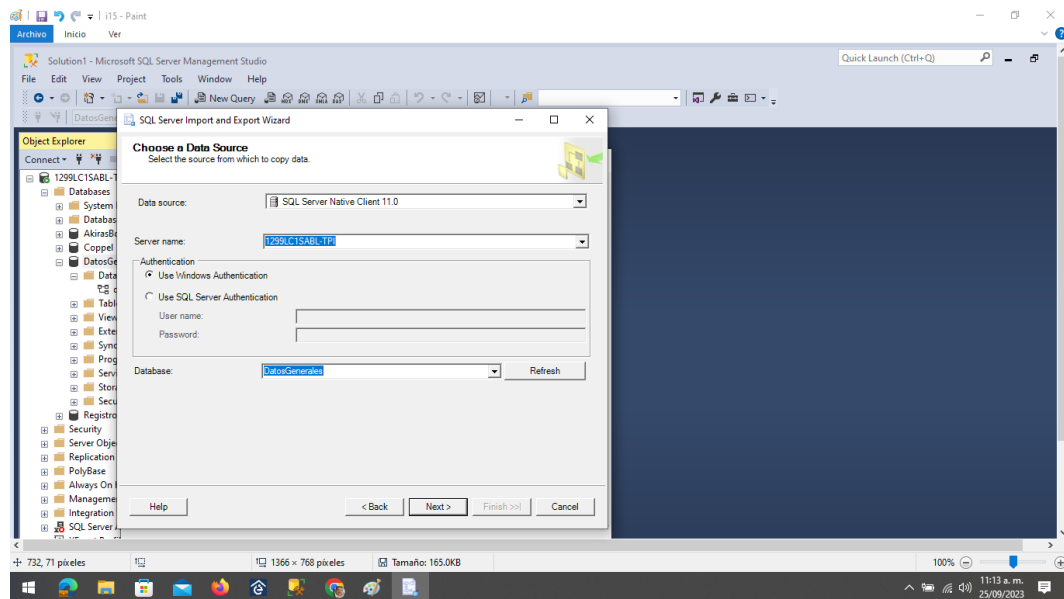
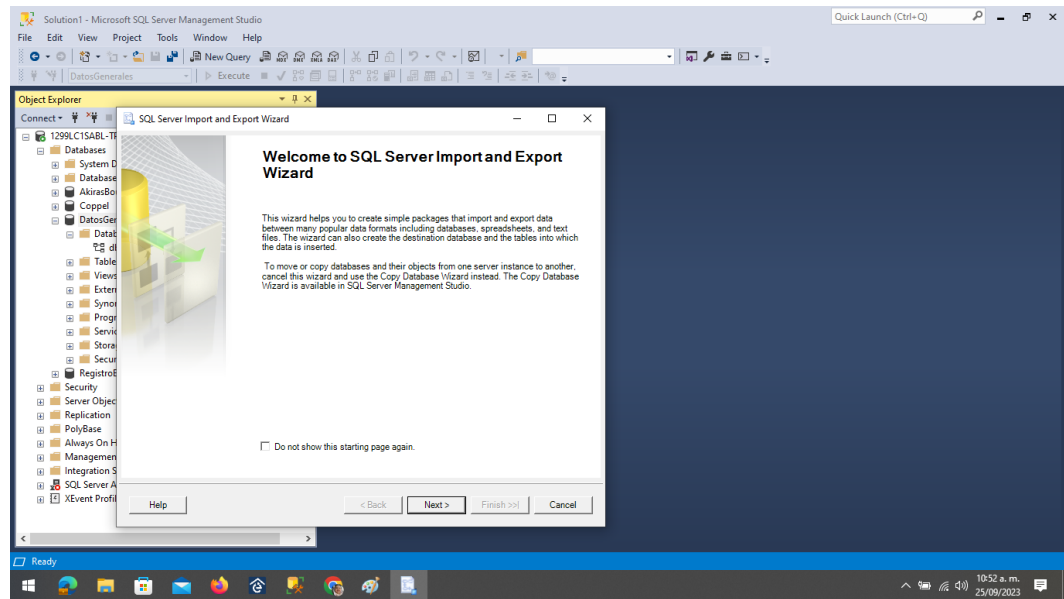
	ID	Nombre	Direccion	Ciudad	Num Telefono	Email	Genero
1	1	Jose Domingo Reyes Amoy	Campeche 2 Centro San Blas	San Blas	3231179298	drckaug@gmail.com	Masculino
2	2	Griedda Huzar Chavez	Anadno neno 829 Zapata	Tapic	3111347632	ghuizar@coppel.com	Femenino
3	3	Rosa Francia Uanos Estrada	Coloso 37 Fovissate	Tapic	3111567832	rlanos@coppel.com	Femenino
4	4	Daniel Contreras Barrera	Andador Coamiles 6 Fraccionamiento las brisas	Compostela	3222325679	dcontreras@coppel.com	Masculino
5	5	Rahim Escobar Llanas	paredes 357 Benito Juarez	La Peñita	3241346581	rescobar@coppel.com	Masculino
6	6	Riorbertha Driantes Ramirez	Singata 339 La libertad	Guayabitos	3271571537	fdorantes@coppel.com	Femenino
7	7	Olga Carolina Virgen Rodarte	Bravo 422 Centro	Tecuala	3211192139	ovirgen@coppel.com	Femenino
8	8	Diana Vaney Robles Guzman	Callejon del Rey 3 La Playita	Acaponeta	3212319446	drobles@coppel.com	Femenino
9	9	Barry Ludin Medina Ojeda	Prolongacion Victoria 1656 Andarellas	Xalisco	3111115437	bmedina@coppel.com	Femenino
10	10	Belem contreras arellano	Silao 49 Centro	Santiago	3241547633	bcontreras@coppel.com	Femenino

Query executed successfully. 1299LC1SABL-TPI (15.0 RTM) COPPEL\Capacitacion.12... DatosGenerales 00:00:00 10 rows

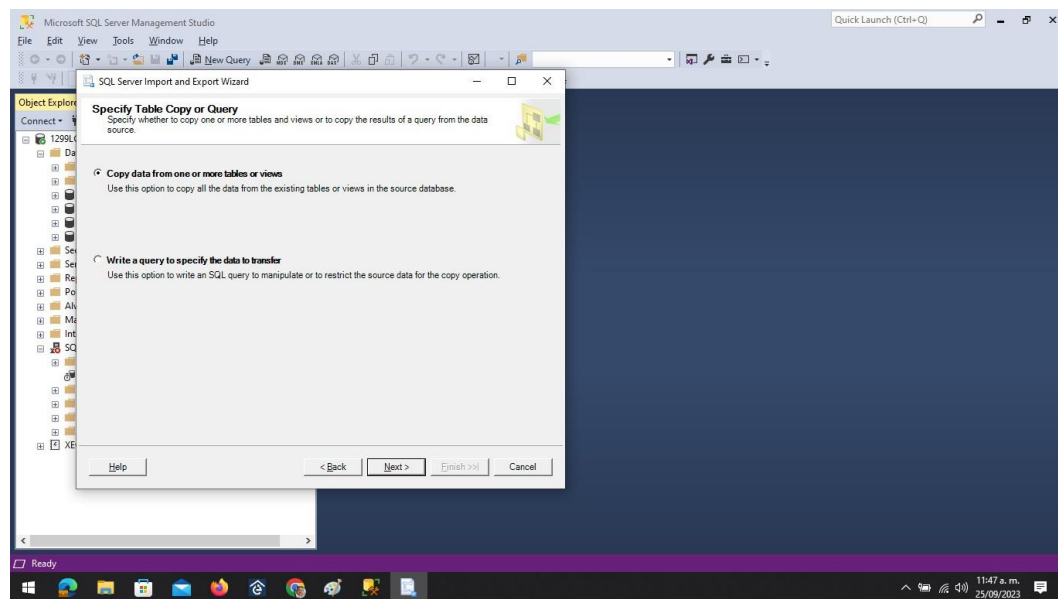
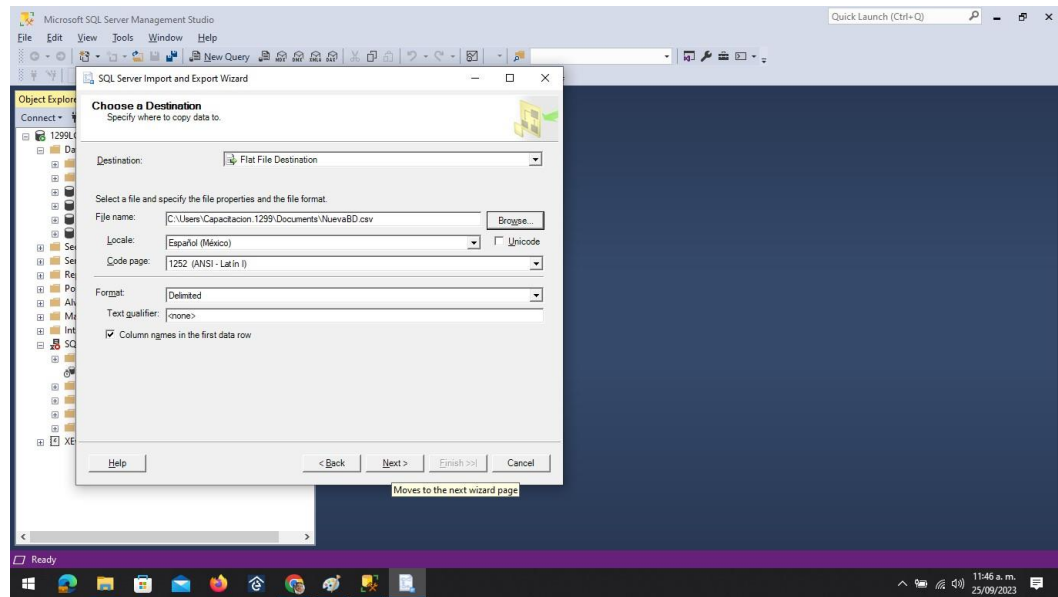


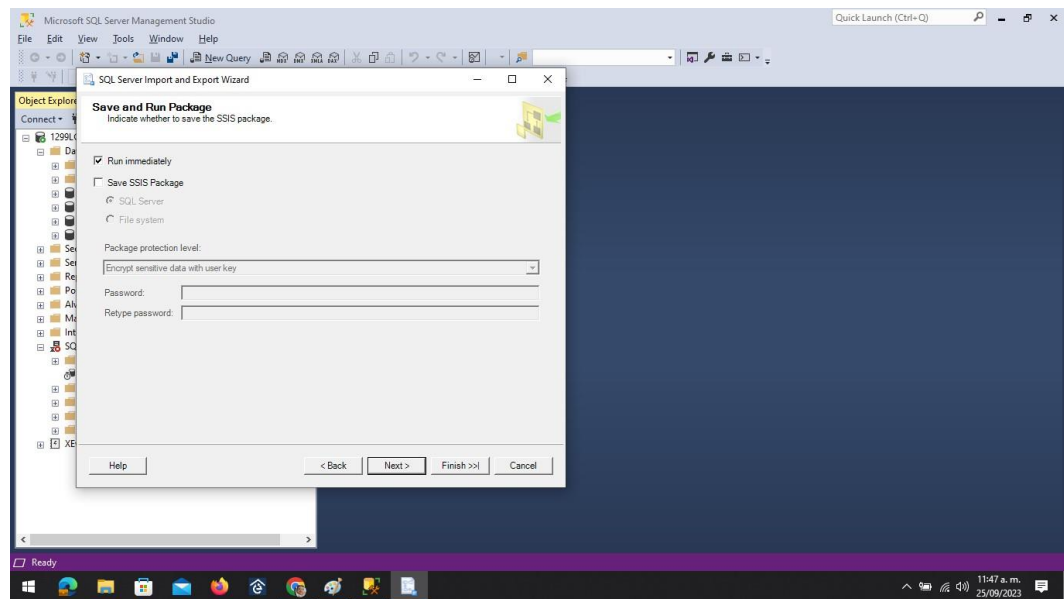
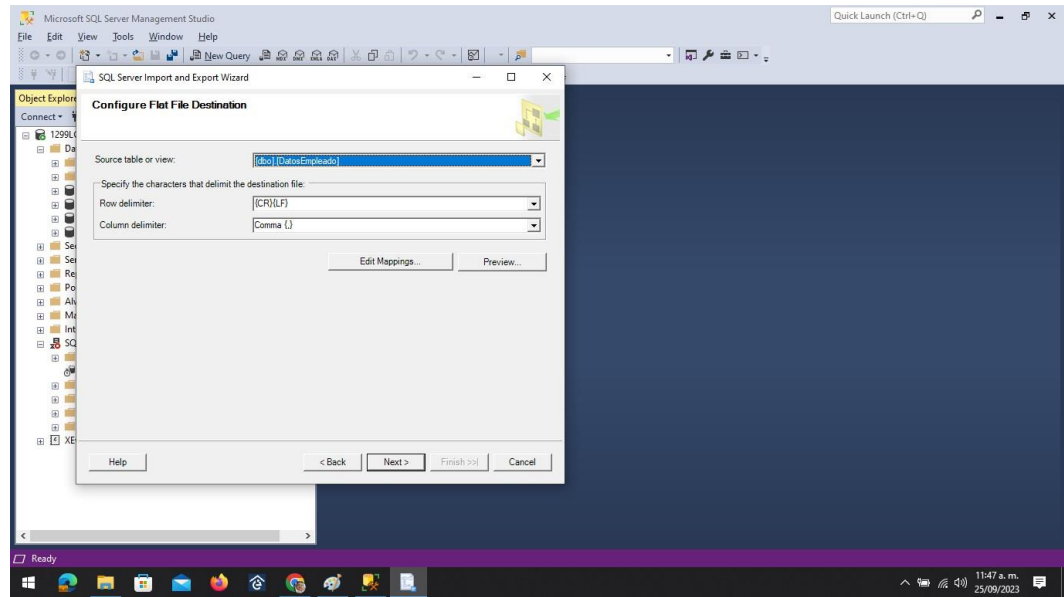
4.2 Importación de la Base de Datos

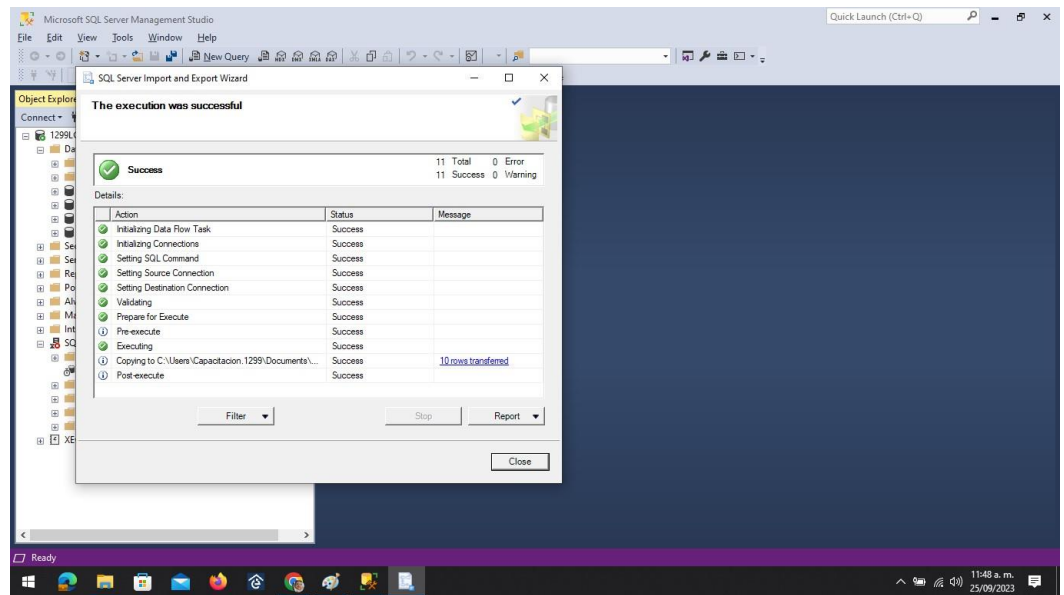
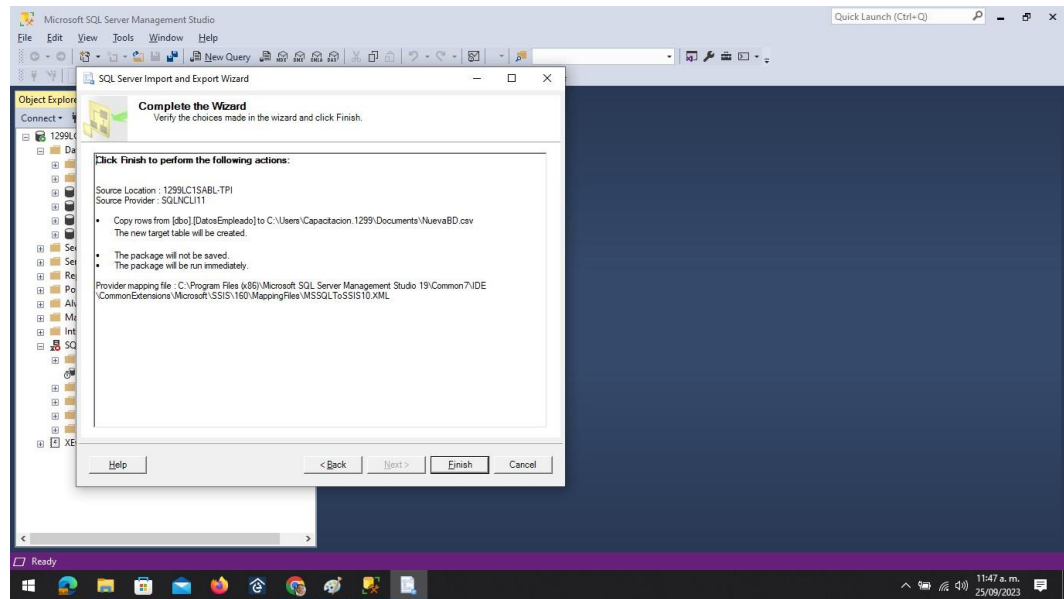
Después de haber elaborado la Base de datos con sus tablas y registros, se realiza la importación de esta a Excel:



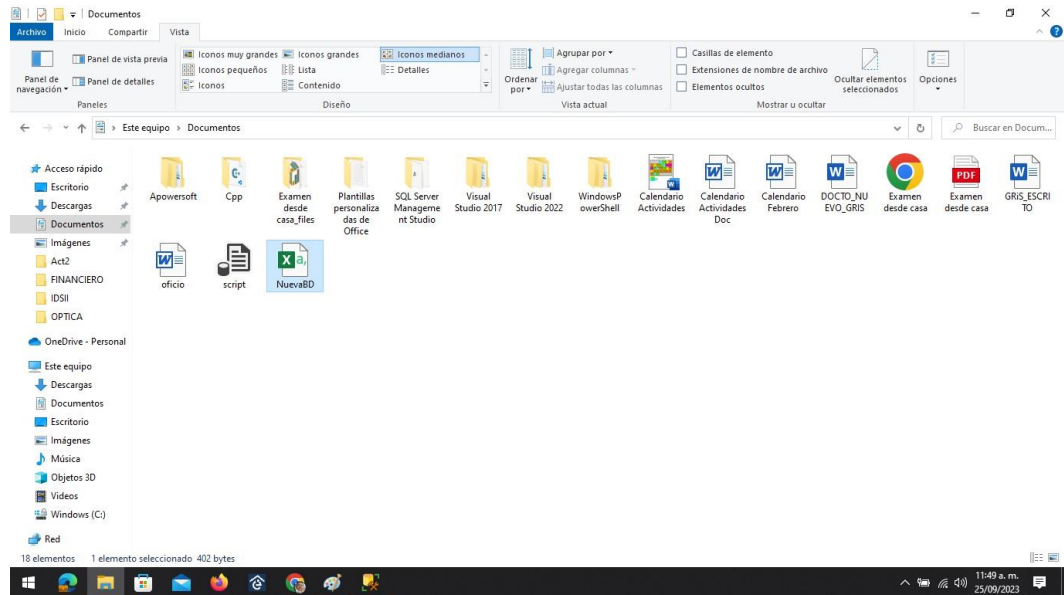
Aquí se inicia el asistente de exportación y se selecciona servidor y BD.







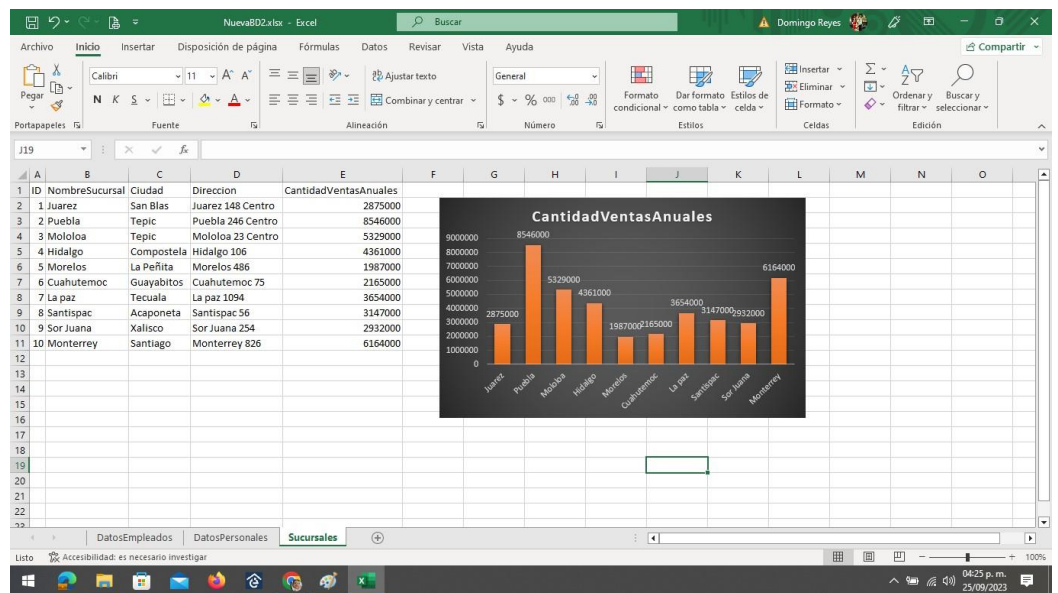
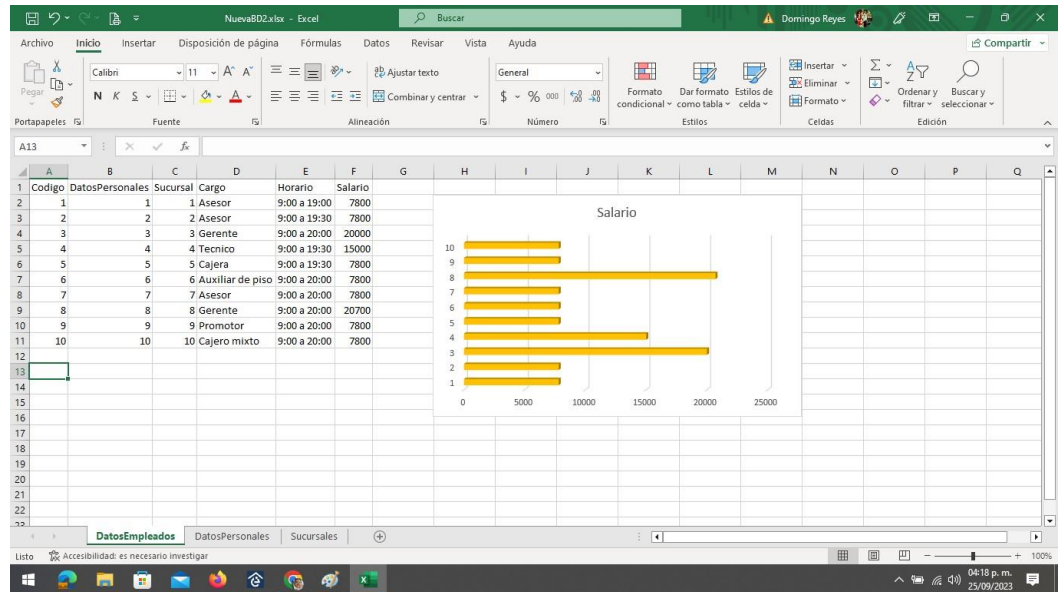
Solo se selecciona el destino y en seguida se le da siguiente hasta concluir.

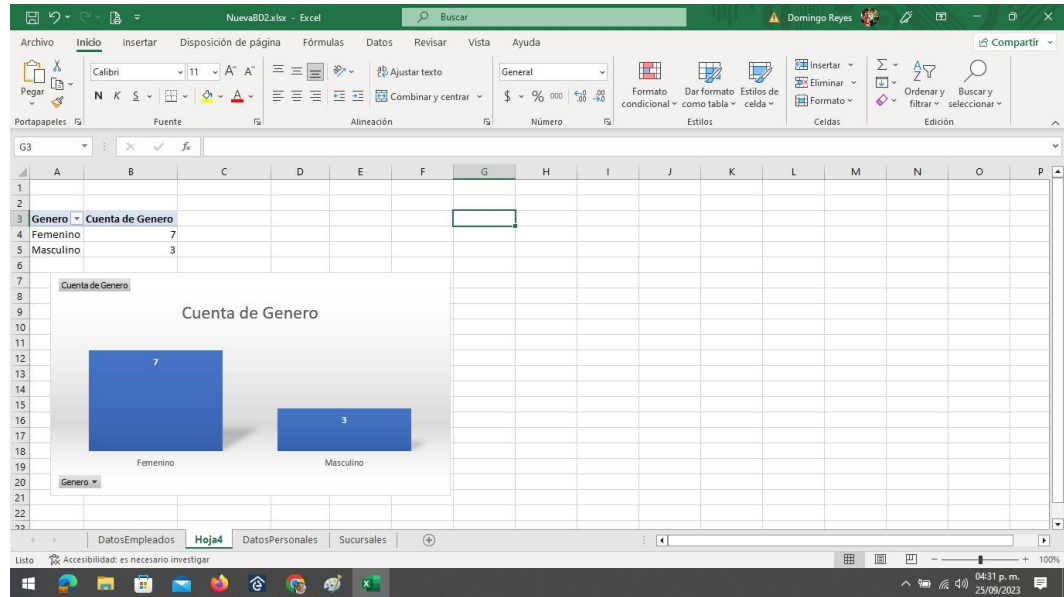


Una vez concluido el proceso tendremos en la carpeta destino, el archivo correspondiente a la base de datos exportada, que se puede abrir con la aplicación de Excel y posteriormente darle formato adecuado para que tenga una buena presentación

4.3 Gráficos

En cada una de las tablas se selecciona la grafica conforme la columna que corresponda:





En este último se filtró la información para seleccionar solo la información que se graficará, se crea por automático una hoja nueva en el mismo libro.

4.4 Preguntas

A continuación, con la ayuda de las tabla y gráficas, se dará respuesta a las siguientes preguntas:

1. ¿Quién es la persona que gana más?

Para contestar esta pregunta, en las mismas tablas que se generaron en Excel de la exportación, se realiza un filtro y nos proporciona la siguiente información:

Table: ExpandRecordColumn(#"Se expandió DatosPersonales(DatosPersonales)", "Sucursales", ("NombreSucursal"), ("Sucursales.NombreSucursal"))

	Sucursal	Cargo	Horario	Salario	DatosPersonales(DatosPersonales).Nombre	Sucursales.NombreSucursal
1	8	Gerente	9:00 a 20:00	20700	Diana Vianey Robles Guzman	Santispac
2	3	Gerente	9:00 a 20:00	20000	Rosa Francia Llanos Estrada	Mololoa
3	4	Tecnico	9:00 a 19:30	15000	Daniel Contreras Barrera	Hidalgo
4	5	Cajera	9:00 a 19:30	7800	Rahim Escobar Llamas	Morelos
5	6	Auxiliar de piso	9:00 a 20:00	7800	Floriberta Drantes Ramirez	Cuahuatmoc
6	7	Asesor	9:00 a 20:00	7800	Olga Carolina Virgen Rodarte	La paz
7	9	Promotor	9:00 a 20:00	7800	Banny Ludin Medina Ojeda	Sor Juana
8	10	Cajero mixto	9:00 a 20:00	7800	Belem contreras arellano	Monterrey
9	1	Asesor	9:00 a 19:00	7800	Jose Domingo Reyes Arroyo	Juarez
10	2	Asesor	9:00 a 19:30	7800	Griselda Huzar Chavez	Puebla

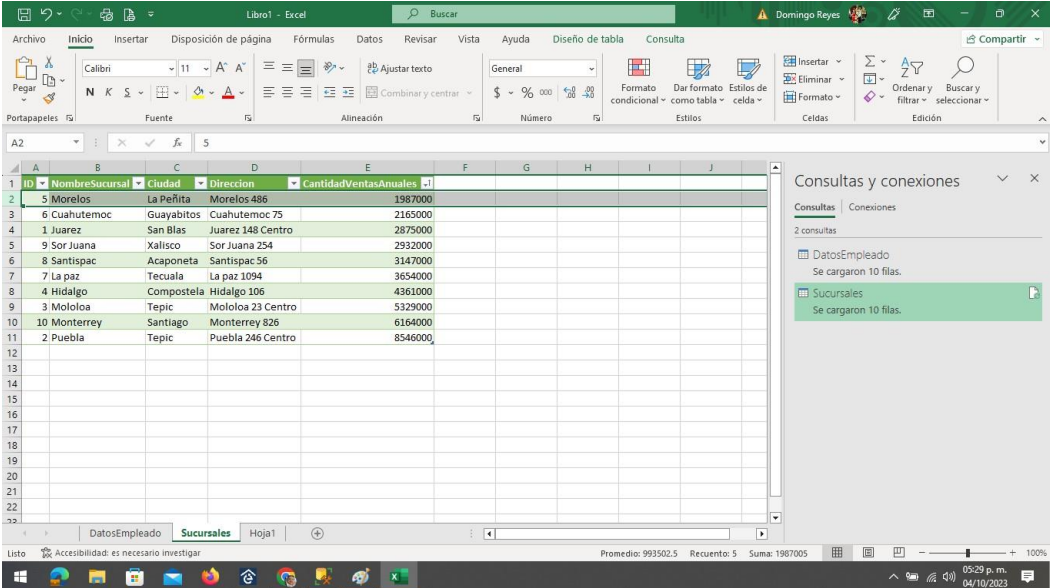
9 COLUMNAS, 10 FILAS Generación de perfiles de columnas basada en las 1000 primeras filas

VISTA PREVIA DESCARGADA A LAS 05:23 P. M. 05:24 p. m. 04/10/2023

Donde nos indica que Diana vianey Robles Guzmán tiene el Sueldo Mas alto conforme a lo que se organizo en la Base de Datos obteniendo la información de las ultimas 2 columnas que se relaciona con la llave primaria y foránea.

2. ¿Cuál es la sucursal que gana menos anualmente?

De la misma manera al importar la tabla se genera un filtro que nos mostrara la información de menor a mayor en base a la cantidad de ventas anuales:



ID	NombreSucursal	Ciudad	Direccion	CantidadVentasAnuales
5	Morelos	La Peñita	Morelos 486	1987000
6	Cuahutemoc	Guayabitos	Cuahutemoc 75	2165000
1	Juarez	San Blas	Juarez 148 Centro	2875000
9	Sor Juana	Xalisco	Sor Juana 254	2932000
8	Santispac	Acaponeta	Santispac 56	3147000
7	La paz	Tecuala	La paz 1094	3654000
4	Hidalgo	Compostela	Hidalgo 106	4361000
3	Mololoa	Tepic	Mololoa 23 Centro	5329000
10	Monterrey	Santiago	Monterrey 826	6164000
2	Puebla	Tepic	Puebla 246 Centro	8546000

Donde se observa que la Sucursal Morelos de la Ciudad La Peñita, es la que tiene el ingreso en ventas más bajo.

3. ¿Cuántos empleados son hombres?

Para encontrar esta información dentro de la tabla exportada de la base de datos se utiliza la función de contar.si en la columna de género:

ID	Nombre	Dirección	Ciudad	NumTelefono	Email	Género
2	Griselda Huzar Chavez	Amado nervo 829 Zapata	Tepic	3111347692	ghuizar@coppel.com	Femenino
3	Rosa Francia Llanos Estrada	Colosio 37 Fovissste	Tepic	3111567832	rllanos@coppel.com	Femenino
4	Floribertha Drantes Ramirez	Singaíta 339 La libertad	Guayabitos	3271571537	fdorantes@coppel.com	Femenino
5	Olga Carolina Virgen Rodarte	Bravo 422 Centro	Tecuala	3211192139	ovirgen@coppel.com	Femenino
6	Diana Vianey Robles Guzman	Callejon del Rey 3 La Playita	Acaponeta	3212318946	drobles@coppel.com	Femenino
7	Banny Ludin Medina Ojeda	Prologacion Victoria 1656 Andareñas	Xalisco	3111115437	lbmedina@coppel.com	Femenino
8	Belem contreras arellano	Silao 49 Centro	Santiago	3241547693	bcontreras@coppel.com	Femenino
9	Jose Domingo Reyes Arroyo	Campeche 2 Centro San Blas	San Blas	3231179298	drksug@gmail.com	Masculino
10	Daniel Contreras Barrera	Andador Coamiles 6 Fraccionamiento las brisas	Compostela	3222325679	dcontreras@coppel.com	Masculino
11	Rahim Escobar Llamas	paredes 397 Benito Juarez	La Peñita	3241346581	rescobar@coppel.com	Masculino

Dando como resultado 3 Hombres.

5 Conclusión

Con la elaboración de esta actividad se logra comprender el uso de tablas y Bases de datos en MySQL Server Management, así como la exportación de esta base de datos hacia el

programa de Excel para su análisis a través de las graficas creadas, estas graficas si se unifican se puede generar una tabla OLAP con la cual se obtendrá una información mas detallada dentro de la misma Base de Datos, con esta actividad se logra comprender el uso de Bases de Datos, así como su manipulación, organización y análisis de los datos obtenidos de esta, se generan las Graficas correspondientes para mostrar la información requerida dependiendo de lo que se quiera realizar con esta información.

En general se obtienen los conocimientos necesarios sobre el uso de Bases de Datos y sobre las Tablas OLAP en cuanto a la minería de datos se refiere, además se adquieren los conocimientos específicos para realizar un análisis de los datos obtenidos en cualquier base de datos.

Referencias

Stark, K. (2022, 18 agosto). *Cubos OLAP de información para la toma de decisiones - Evaluando software*. Evaluando Software.

<https://www.evaluandosoftware.com/cubos-olap-informacion-la-toma-decisiones/>

¿Qué es la minería de datos? (s. f.). SAS.

https://www.sas.com/es_mx/insights/analytics/data-mining.html