

Actividad 1 - Cubos OLAP Minería de datos Ingeniería en Desarrollo de Software

Tutor: Félix Acosta Hernández

Alumno: Jusi Ismael Linares Gutiérrez

Fecha: 23/09/2023

Índice

| Introducción | 2 |
|---|----|
| Introduccion | 3 |
| Descripción | 3 |
| Justificación | 3 |
| Desarrollo: | 4 |
| Creación de la base de datos | 4 |
| Importación de la base de datos a Excel | 5 |
| Gráficos | 8 |
| Salario de los empleados del más alto al más bajo | 8 |
| Ventas totales de las sucursales | 13 |
| Total de empleados acorde a su género | 17 |
| Preguntas | 19 |
| Conclusión | 21 |
| GitHub | 21 |

Introducción

En esta primera actividad de la materia minería y análisis de datos vamos a crear una base de datos para la empresa compuventas. Para esto se nos pide que cuente con 3 tablas con información de sus empleados y de sus sucursales, una vez creada esta base de datos deberemos ingresar 10 registros como mínimo a cada una de las tablas, cuando tengamos registros en nuestras tablas procederemos a sacar gráficos de nuestra información, ¿Cómo? Pues exportando la base de datos desde MySQL Server Management a Microsoft Excel para que con la información que se encuentra en nuestra base de datos seamos capaces de crear gráficos y así expresar de manera más sencilla nuestra información para cualquier persona. Como ejemplo uno de los gráficos nos pide visualizar a los empleados acorde a su género, por lo que, si en nuestros registros incluimos a 3 mujeres y a 7 hombres, esta información debe representarse de manera grafica como un 30% de mujeres y un 70% de hombres. Como parte final de la actividad se nos pide contestar preguntas con la información que obtengamos de nuestras gráficas para comprobar que las realizamos correctamente.

Descripción

Para esta primera actividad de nuestra materia minería y análisis de datos se nos pide la creación de una base de datos para la empresa CompuVentas que se dedican a la venta de hardware, el cliente (es decir, la empresa CompuVentas) nos especifica la cantidad de tablas (3 tablas) que desea en su base de datos, al igual que el nombre de las columnas de cada una de las tablas. Una vez tengamos creada la base de datos ingresaremos 10 registros a cada una de las tablas debido a que esa información no la especifica el cliente y ya que mas adelante necesitamos crear gráficos en Excel, los registros son una parte muy importante pues servirán para hacer las pruebas correspondientes. Cuando la base de datos este correctamente terminada y las 3 tablas cuenten con sus 10 registros cada una, procederemos a exportar la base de datos hacia Microsoft Excel para poder crear gráficos con dicha información, como salarios generales, ventas totales de las sucursales y total de empleados acorde a su género. Finalmente, la actividad nos pide contestar preguntas en base a los gráficos creados para verificar que realizamos todo correctamente.

Justificación

Lo que veremos a continuación en esta actividad 1 de la materia minería y análisis de datos nos será útil para nuestra formación profesional, pues casi cualquier programa hoy en día utiliza una base de datos, por lo que es muy importante saber cómo crear una desde cero y como utilizarla para manejar la información que contenga, desde solo leerla hasta saber modificarla (crear, eliminar y editar registros). Además cuando necesitemos graficar la información de nuestra base de datos necesitamos saber cómo hacerlo de manera rápida, claramente no podemos pasar el día copiando y pegando cada registro en Microsoft Excel para obtener las gráficas, por lo que es muy importante aprender a exportar nuestra base de datos hacia Excel para disponer de los registros de manera fácil y rápida, esta es una de las partes más importantes en esta actividad ya que nos ahorrara mucho tiempo para graficar nuestros datos. Debido a la facilidad y rapidez con la que podremos graficar nuestra información en Excel sin tener que pasar los registros uno a uno es que considero bastante importante la implementación de esta solución para nuestra actividad.

Desarrollo:

Creación de la base de datos

Con estas líneas de código creamos la base de datos y posteriormente la utilizamos para que las siguientes consultas se ejecuten en esta base de datos

```
create database datos_Generales;
use datos Generales;
```

Creamos nuestra primer tabla la cual es datos personales, les asignamos los tipos que corresponden (varchar para los que llevan texto, int para los que llevan números), en cuanto al ID también le agregamos el código para que sea la llave primaria de la tabla y que el ID se asigne automáticamente de uno en uno al insertar cualquier registro

```
CREATE TABLE datosPersonales (
    ID int primary key NOT NULL IDENTITY(1, 1),
    Nombre varchar(60) NOT NULL,
    Direccion varchar(100) NOT NULL,
    Ciudad varchar(50) NOT NULL,
    num_Telefono bigint NOT NULL,
    Email varchar(60) NOT NULL,
    Genero varchar(10) NOT NULL);
```

A pesar de que en el PDF la tabla 2 es la de datos empleado, primero deberemos crear la tabla 3 sucursal pues la tabla 2 tiene 2 llaves foráneas que corresponden a la tabla 1 y 2 (datos personales y sucursal)

```
ICREATE TABLE sucursal (
    ID int primary key NOT NULL IDENTITY(1, 1),
    NombreSucursal varchar(60) NOT NULL,
    Ciudad varchar(50) NOT NULL,
    Direccion varchar(100) NOT NULL,
    CantidadVentasAnual int NOT NULL,
)
```

Una vez creadas las tablas 1 y 3 procederemos a crear la tabla datos empleados que contiene llaves foráneas de las tablas creadas anteriormente

```
CREATE TABLE datosEmpleado (
    Codigo int primary key NOT NULL IDENTITY(1, 1),
    FK_datosPersonales int NOT NULL,
    FK_sucursal int NOT NULL,
    Cargo varchar(30) NOT NULL,
    Horario varchar(25) NOT NULL,
    Salario int NOT NULL,
    FOREIGN KEY (FK_datosPersonales) REFERENCES datosPersonales(ID),
    FOREIGN KEY (FK_sucursal) REFERENCES sucursal(ID)
```

Ahora que tenemos nuestras tablas creadas les insertaremos 10 registros a cada una empezando por la tabla datos personales

```
INSERT INTO datosPersonales VALUES ('Leonardo Sanchez Perez', 'Av. Miguel Ángel de Quevedo 19, Chimalistac, Álvaro Obregón, 01070 Ciudad de México, CDMX', 'Alvaro Obregon'
 INSERT INIO datosPersonales VALUES ('elonardo Sanchez Perez', Av. Miguel Angel de Quevedo 19, Chimalistac, Alvaro Obregon, 1070º Cludad de México, CDMX', Alvaro Obregon, 118SERT INTO datosPersonales VALUES ('Miguel Angel Gomez Corres', 'Balderas 72-Los Colonia Centro, Centro, Cauahtémoc, 06600 Ciudad de México, CDMX', 'Miguel Hidalgo INSERT INTO datosPersonales VALUES ('Donatelo Reyes Diaz', 'Calle Hipólito Taine 253, Polanco, Polanco V Secc, Miguel Hidalgo, 11560 Ciudad de México, CDMX', 'Miguel Hidalgo INSERT INTO datosPersonales VALUES ('Rafael Guzman Lopez', 'Av. Insurgentes Sur 3757-P. B, La Fama, Tlalpan, 14269 Ciudad de México, CDMX', 'Tlalpan', 7774567890, 'Rafael INNISERT INTO datosPersonales VALUES ('Abril Romero Flores', 'Av. Universidad, Chimalistac, Álvaro Obregón, 04318 Ciudad de México, CDMX', 'Vilvaro Obregón, '7775678901, 'Abril INSERT INTO datosPersonales VALUES ('Rebeca Vargas Salazar', '20 de Noviembre 134, Centro Histórico de la Cdad. de México, Centro, Cuauhtémoc, 06090 Centro, CDMX', 'Cuauhtémoc
 INISERT INIO datosPersonales VALUES ('Usai Ismael Linares Gutiérnez', 'Avenida Beres 58. Colonia Centro, Alcaldia Caushtémoc C.P. 86610, Ciudad de México.', 'Cuauhtémoc', 187898.
INSERT INIO datosPersonales VALUES ('Carol García Rodríguez', 'Av. 5 de Mayo 15, Centro Histórico de la CDMX, Cuauhtémoc, 86610 Ciudad de México, CDMX', 'Cuauhtémoc', 777898.
INSERT INIO datosPersonales VALUES ('Cristiano Ronaldo dos Santos Aveiro', 'Calz. de Tlalpan 3465, Sta. Úrsula Coapa, Coyoacán, 94650 Ciudad de México, CDMX', 'Cuouhtémoc', 7778123456, 'Capiègmon', 187897.
INSERT INIO datosPersonales VALUES ('Christopher Robert Evans', 'Av. P.º de la Reforma, Juárez, Cuauhtémoc, 86600 Ciudad de México, CDMX', 'Cuauhtémoc', 7778123456, 'capiègmon', 187897.
Después haremos lo mismo con la tabla sucursal donde registraremos 10 sucursales
  INSERT INTO sucursal VALUES ('CompuVentas Venustiano Carranza', 'Venustiano Carranza', 'Av. del Peñón 389-155, Moctezuma 2da Secc, Venustiano Carranza, 15530 Ciudad de México
  INSERT INTO sucursal VALUES ('CompuVentas Cuauhtémoc', 'Cuauhtémoc', 'Av Francisco I. Madero 39-Local 1-A, Centro Histórico de la CDMX, Cuauhtémoc, 06000 Ciudad de México', INSERT INTO sucursal VALUES ('CompuVentas Miguel Hidalgo', 'Miguel Hidalgo', 'Av. Marina Nacional 69, Tacuba, Miguel Hidalgo, 11410 Ciudad de México, CDMX', 2000000);
INSERT INTO sucursal VALUES ('CompuVentas Miguel Hidalgo', 'Miguel Hidalgo', 'Av. Marina Nacional 69, Tacuba, Miguel Hidalgo, 11410 Ciudad de México, CDMX', 2000000);
INSERT INTO sucursal VALUES ('CompuVentas Tlalpan', 'Tlalpan', 'Calz Acoxpa 430, Coapa, Ex-Hacienda Coapa, Tlalpan, 14300 Ciudad de México, CDMX', 2000000);
INSERT INTO sucursal VALUES ('CompuVentas Coyoacan', 'Coyoacân', 'Av. Universidad 1770, Romero de Terreros, Coyoacân, 04310 Ciudad de México, CDMX', 7000000);
INSERT INTO sucursal VALUES ('CompuVentas Iztapalapa', 'Iztapalapa', 'Canal de Tezontle 1512-Local L05, Dr Alfonso Ortiz Tirado, Iztapalapa, 9020 Ciudad de México, CDMX', 9500001;
INSERT INTO sucursal VALUES ('CompuVentas Kochimilco', 'Av. Jose Ma. Morelos 59, El Rosario, Xochimilco, 16070 Ciudad de México, CDMX', 8500001;
INSERT INTO sucursal VALUES ('CompuVentas Benito Juarez', 'Benito Juarez', 'Benito
Por último, insertaremos 10 registros en la tabla datos empleado
INSERT INTO datosEmpleado VALUES (1, 3, 'Director Financiero', '9:30 AM - 9:30 PM', 50000);
INSERT INTO datosEmpleado VALUES (2, 2, 'Empleado', '9:30 AM - 8:30 PM', 12000);
INSERT INTO datosEmpleado VALUES (3, 3, 'CEO', '9:30 AM - 9:30 PM', 80000);
INSERT INTO datosEmpleado VALUES (4, 4, 'Empleado', '9:30 AM - 8:30 PM', 12000);
INSERT INTO datosEmpleado VALUES (5, 8, 'Director Marketing', '9:30 AM - 8:30 PM', 40000);
INSERT INTO datosEmpleado VALUES (6, 2, 'Gerente', '9:30 AM - 9:30 PM', 25000);
INSERT INTO datosEmpleado VALUES (7, 1, 'Empleado', '9:30 AM - 8:30 PM', 12000);
INSERT INTO datosEmpleado VALUES (8, 5, 'Gerente', '9:30 AM - 9:30 PM', 25000);
```

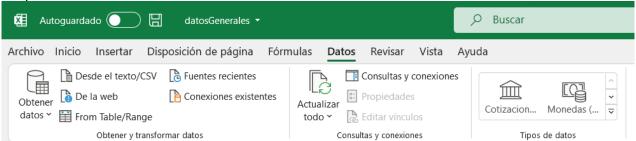
Una vez completada nuestra base de datos podemos continuar a exportar la base de datos a Excel para graficar nuestra información

INSERT INTO datosEmpleado VALUES (10, 9, 'Recursos Humanos', '9:30 AM - 9:30 PM', 30000);

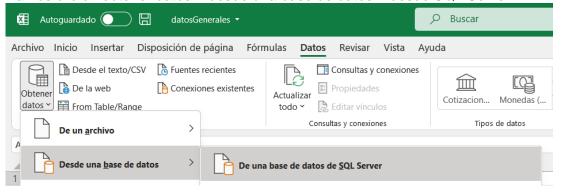
INSERT INTO datosEmpleado VALUES (9, 5, 'Empleado', '9:30 AM - 8:30 PM', 12000);

Importación de la base de datos a Excel

Para empezar, debemos abrir nuestro archivo de Excel y en la parte superior debemos dirigirnos al apartado de datos



Damos clic en obtener datos > desde una base de datos > desde SQL Server

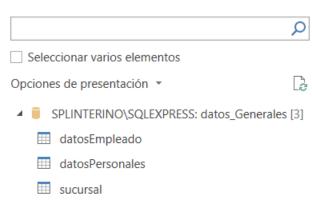


Nos abrirá esta venta donde llenaremos los datos que se nos piden

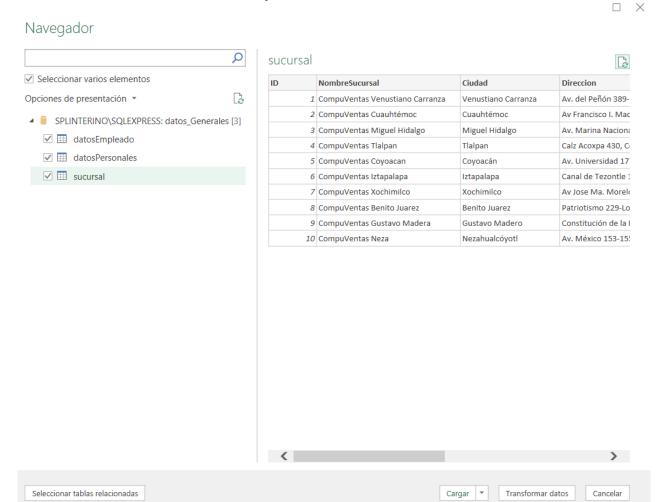


Antes que nada, debemos marcar la casilla de seleccionar varios elementos

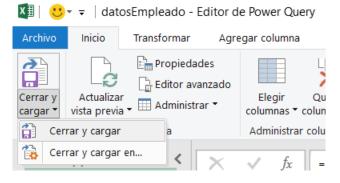
Navegador



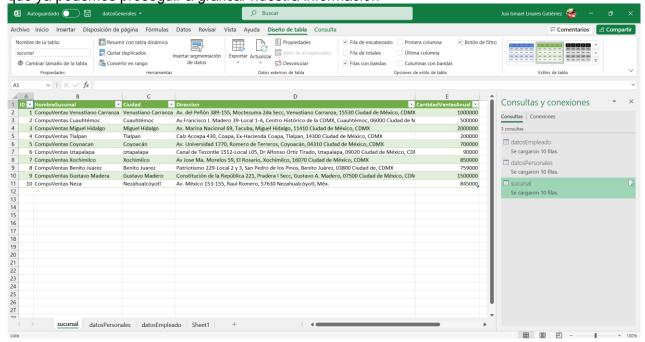
Seleccionamos todas nuestras tablas y damos clic en transformar datos



En la siguiente ventana que nos abre Excel iremos a Archivo > Cerrar y cargar y damos clic en la primer opción que es Cerrar y cargar



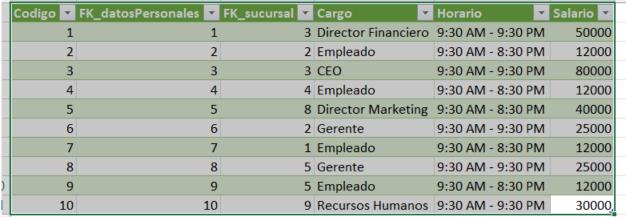
Ahora así se ve nuestro archivo de Excel con nuestras tablas ocupando una hoja cada una, por lo que ya podemos proseguir a graficar nuestra información



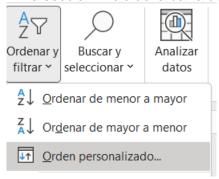
Gráficos

Salario de los empleados del más alto al más bajo

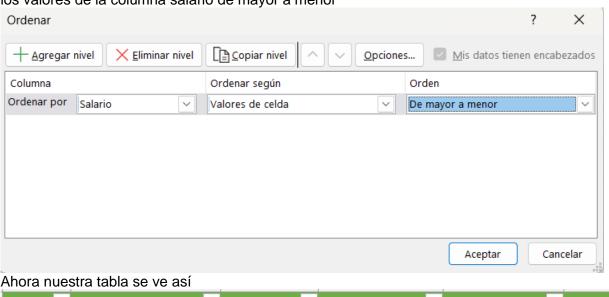
El primer gráfico que se nos pide es "Salario de los empleados del más alto al más bajo" por lo que primero seleccionamos todas las celdas de nuestra tabla datosEmpleado



En la sección Inicio de la barra de herramientas buscamos Ordenar y filtrar > Orden personalizado



En esta ventana seleccionamos las opciones que necesitamos, en mi caso lo quiero ordenar por los valores de la columna salario de mayor a menor

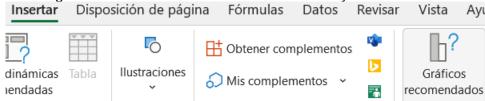




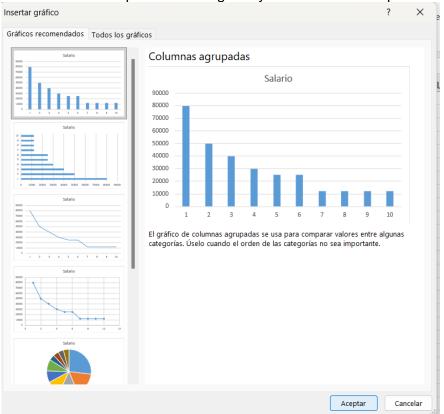
Ahora solo debemos seleccionar las celdas de la columna salario

| Codigo 💌 | FK_datosPersonales 💌 | FK_sucursal 💌 | Cargo | Horario 🔻 | Salario 🛂 |
|----------|----------------------|---------------|---------------------|-------------------|-----------|
| 3 | 3 | 3 | CEO | 9:30 AM - 9:30 PM | 80000 |
| 1 | 1 | 3 | Director Financiero | 9:30 AM - 9:30 PM | 50000 |
| 5 | 5 | 8 | Director Marketing | 9:30 AM - 8:30 PM | 40000 |
| 10 | 10 | 9 | Recursos Humanos | 9:30 AM - 9:30 PM | 30000 |
| 6 | 6 | 2 | Gerente | 9:30 AM - 9:30 PM | 25000 |
| 8 | 8 | 5 | Gerente | 9:30 AM - 9:30 PM | 25000 |
| 2 | 2 | 2 | Empleado | 9:30 AM - 8:30 PM | 12000 |
| 4 | 4 | 4 | Empleado | 9:30 AM - 8:30 PM | 12000 |
| 7 | 7 | 1 | Empleado | 9:30 AM - 8:30 PM | 12000 |
| 9 | 9 | 5 | Empleado | 9:30 AM - 8:30 PM | 12000 |
| | | | | | |

Nos dirigimos a la barra de herramientas > Insertar y seleccionamos Gráficos recomendados

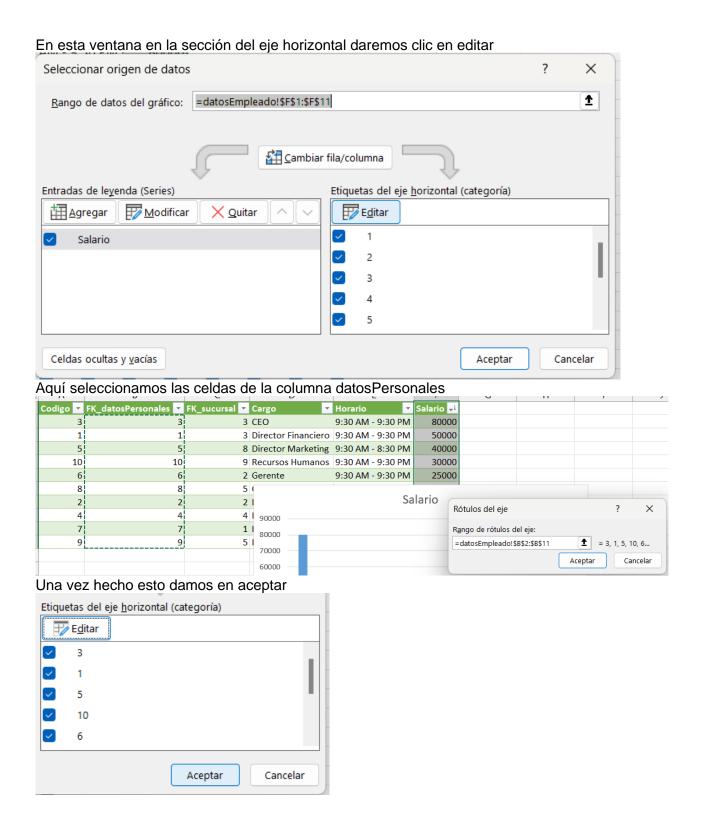


Seleccionamos el que más nos agrade y damos clic en aceptar

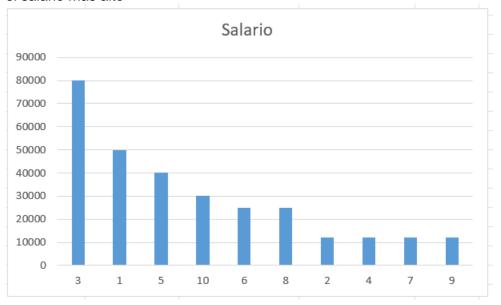


Ya que seleccionamos el gráfico daremos clic derecho en el para seleccionar la opción Seleccionar datos...



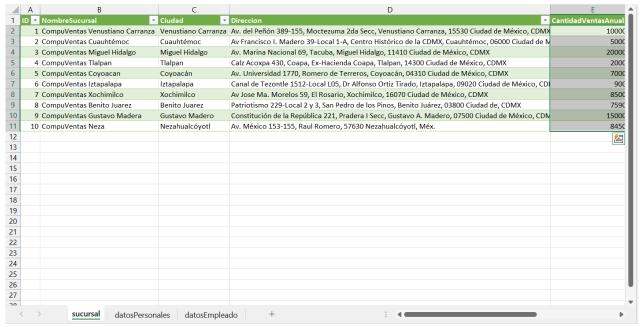


Nuestro gráfico nos quedo de la siguiente manera, por lo que sabemos quien es el empleado con el salario mas alto

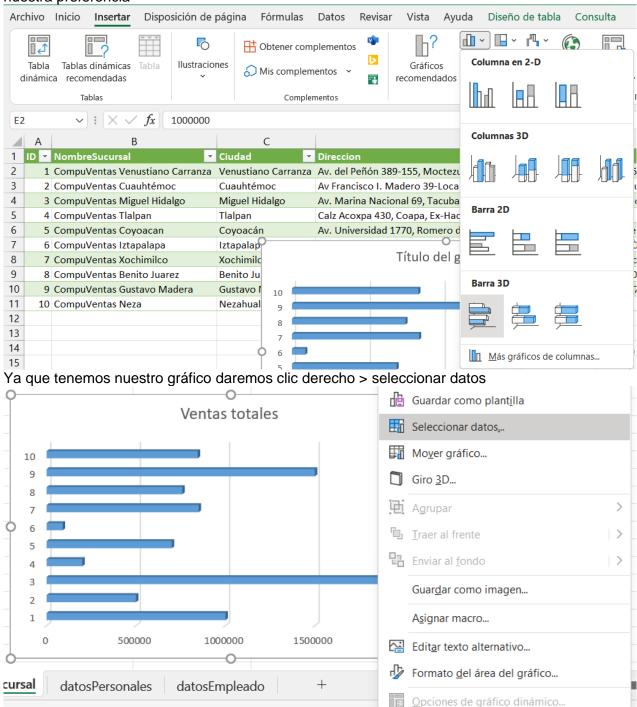


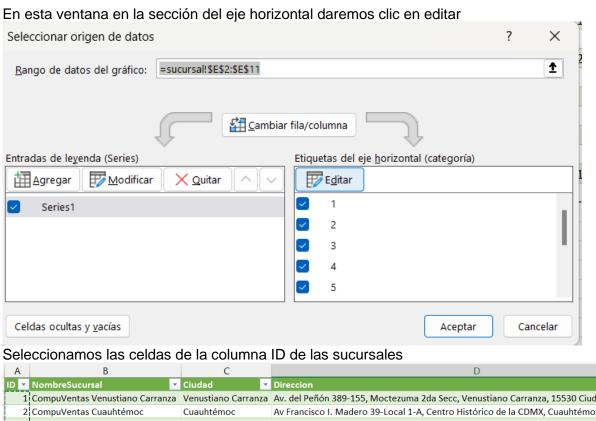
Ventas totales de las sucursales

Nos dirigimos a la hoja de sucursal donde seleccionaremos las celdas de la columna ventas anuales



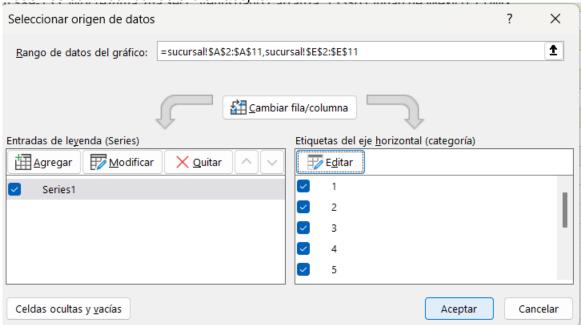
En la barra de herramientas nos dirigimos a la sección insertar > gráficos y seleccionamos el de nuestra preferencia



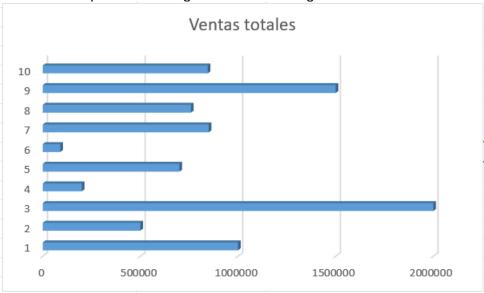




Podemos continuar dando clic en aceptar

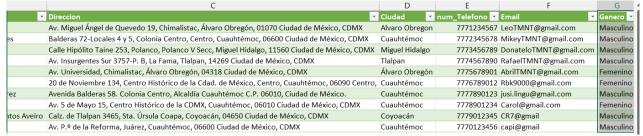


Así es como quedo nuestro grafico sobre los ingresos de las distintas sucursales



Total de empleados acorde a su género

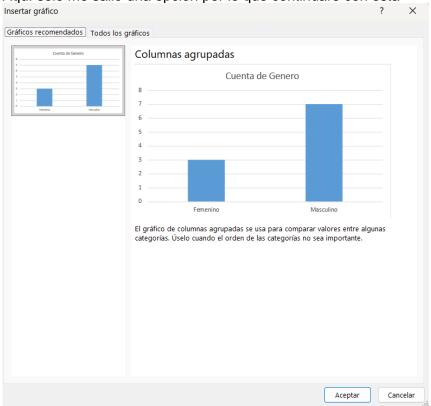
Para finalizar las gráficas que se nos piden iremos a la hoja datosPersonales y seleccionaremos las celdas de la columna Genero



En la barra de herramientas nos dirigimos a gráficos recomendados



Aquí solo me salió una opción por lo que continuare con esta



Nos quedo de la siguiente manera, por lo que ahora ya sabemos cuantos empleados hay por cada



Preguntas

¿Quién es la persona que gana más?

De acuerdo a nuestro grafico el primer ID que se muestra es el 3

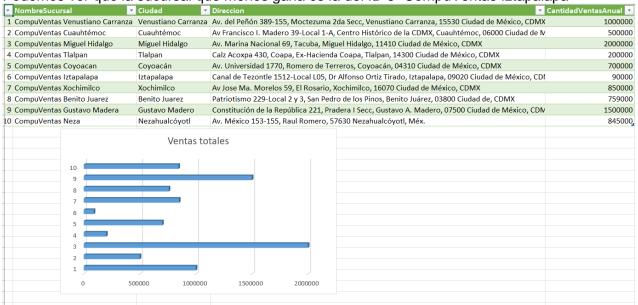
| Codigo 🔽 FK_datosP | ersonales 💌 | FK_sucursal 💌 | Cargo | Horario 💌 | Salario 🚽 |
|--------------------|---|---------------|---------------------|-------------------|-----------|
| 3 | 3 | 3 | CEO | 9:30 AM - 9:30 PM | 80000 |
| 1 | 1 | . 3 | Director Financiero | 9:30 AM - 9:30 PM | 50000 |
| 5 | 5 | 8 | Director Marketing | 9:30 AM - 8:30 PM | 40000 |
| 10 | 10 | 9 | Recursos Humanos | 9:30 AM - 9:30 PM | 30000 |
| 6 | 6 | 2 | Gerente | 9:30 AM - 9:30 PM | 25000 |
| 8 | 8 | 5 | Gerente | 9:30 AM - 9:30 PM | 25000 |
| 2 | 2 | | Empleado | 9:30 AM - 8:30 PM | 12000 |
| 4 | 4 | | Empleado | 9:30 AM - 8:30 PM | |
| 7 | 7 | | Empleado | 9:30 AM - 8:30 PM | |
| 9 | 9 | 5 | Empleado | 9:30 AM - 8:30 PM | 12000 |
| | 70000 — 60000 — 50000 — 40000 — 30000 — | н | | | |
| | 20000 — 10000 — 0 — | | | | |
| | 3 | 1 5 | 10 6 8 | 2 4 7 | 7 9 |

Y en a tabla datosPersonales este ID pertenece al empleado Donatelo



¿Cuál es la sucursal que gana menos anualmente?

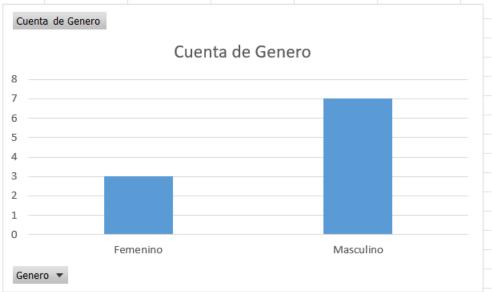
Podemos ver que la sucursal que menos gana es la del ID 6 "CompuVentas Iztapalapa



¿Cuántos empleados son hombres?

De acuerdo a nuestro grafico podemos apreciar que hay 3 empleadas y 7 empleados en los 10

registros que insertamos en la tabla



Conclusión

En esta actividad, hemos tenido la oportunidad de comprender la importancia fundamental de las bases de datos en el ámbito empresarial. Una base de datos no es más que un almacén organizado de información valiosa que, en este caso, ha sido bautizado como "datosgenerales" para la empresa CompuVentas.

A lo largo de esta actividad, aprendimos a crear y gestionar eficazmente esta base de datos en Microsoft SQL Server Management. Hemos organizado datos personales y laborales, lo cual es crucial para comprender la estructura y el funcionamiento de la empresa.

la verdadera importancia de esta actividad radica en la capacidad de extraer información significativa de estos datos y representarla visualmente a través de gráficos. La "Gráfica 1: Salarios generales" nos proporciona una visión clara de la distribución salarial en la empresa, mientras que la "Gráfica 2: Ventas totales de las sucursales" nos permite identificar las sucursales más productivas. Por último, la "Gráfica 3: Total de empleados por género" arroja luz sobre la igualdad de género en la empresa. Este tipo de análisis es esencial para tomar decisiones estratégicas informadas, como ajustar salarios, asignar recursos y garantizar un ambiente de trabajo equitativo. En resumen, además de lo aprendido como desarrolladores de software (crear una base de datos, exportarla a Excel y graficar los datos) también hemos comprendido que los datos obtenidos nos pueden resultar útiles para mejorar la eficacia y eficiencia de la empresa basándonos en los datos obtenidos, como la posibilidad de ajustar sueldos entre hombre y mujeres, reducir jornadas laborales, entre otros.

GitHub

https://github.com/JusiLinGu/Practicas_UMI