

Actividad 1 - Cubos OLAP


Minería de Datos

Ingeniería en Desarrollo de Software

Tutor: Félix Acosta

Alumno: Homero Ramirez Hurtado

Fecha: 25 de Septiembre del 2023



Índice.

. Introducción.

. Descripción.

. Justificación.

. Desarrollo.

- Creación de la Base de Datos.
- Importación de la Base de Datos a Excel.
 - Gráficos.
 - Preguntas.

. Conclusión.

. Referencias.



Introducción.

A continuación iniciaremos un nuevo tema en la minería de datos y estos son los bloques OLAP los cuales analizaremos su función y como podemos emplearlos así como también las ventajas que tiene el usarlos a un lado de esto haremos una base de datos la cual haremos una transferencia a Excel y la graficaremos con los resultados de nuestra base de datos y así mismo resolviendo las preguntas que se verán a continuación después de haber realizado los pasos anteriores.

Descripción.

El procesamiento analítico en línea (OLAP) es una tecnología de software que puede utilizar para analizar datos empresariales desde diferentes puntos de vista. Las organizaciones recopilan y almacenan datos de múltiples fuentes de datos, como sitios web, aplicaciones, medidores inteligentes y sistemas internos. OLAP combina y agrupa estos datos en categorías para proporcionar información procesable para la planificación estratégica. Por ejemplo, un minorista almacena datos sobre todos los productos que vende, como el color, el tamaño, el costo y la ubicación. El minorista también recopila datos de compra de los clientes, como el nombre de los artículos pedidos y el valor total de las ventas, en un sistema diferente. OLAP combina los conjuntos de datos para responder preguntas como qué productos de color son más populares o cómo la colocación de productos afecta a las ventas.

Justificación.

OLAP permite que un usuario tenga varias vistas sobre un juego de datos. Entre esas vistas diferentes podemos pensar en particular en la implementación de paneles, informes o análisis varios. Las técnicas de análisis más conocidas permitidas por OLAP son las siguientes:

- Roll-up: esta operación consiste en reunir datos alrededor de un cubo. Ya sea eliminando algunas dimensiones de un cubo o por el contrario reuniendo datos.
- Drill-Down: esta operación consiste en afinar el análisis dentro del sistema de datos. Por ejemplo, si el usuario desea afinar su búsqueda geográfica, va a pasar de la dimensión “continente” a la dimensión “país”.
- Slice: esta operación consiste en construir un nuevo subdimensión a partir de una dimensión particular de otro cubo.

Creación de Base de Datos.

The screenshot displays the Microsoft SQL Server Enterprise Manager interface. The top menu bar includes File, Edit, View, Project, Query Designer, Tools, Window, and Help. The top toolbar contains various icons for file operations, query execution, and window management. The top status bar shows the current connection: 'LAPTOP-ED6K7OJ\Datos Generales - dbo.DATOSEMPLEO - Microsoft SQL Server Management Studio'.

The main window is divided into three panes:

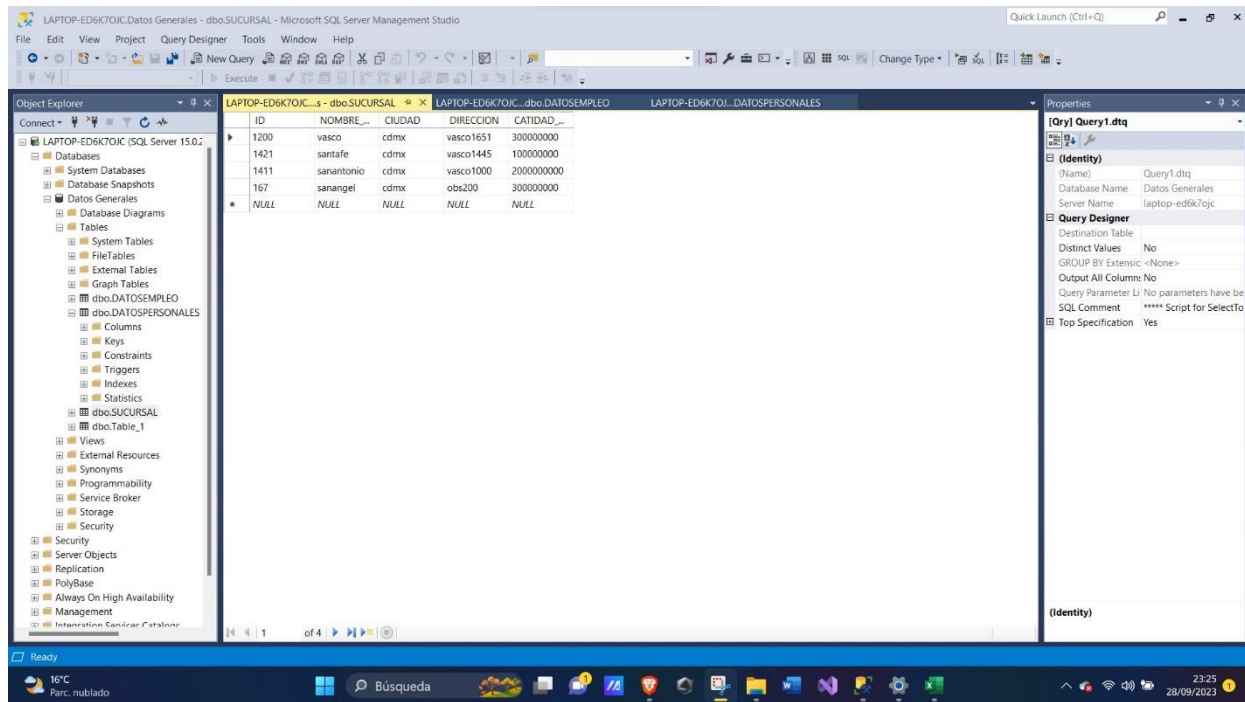
- Object Explorer:** Shows the database structure for 'LAPTOP-ED6K7OJ\Datos Generales'. The 'Tables' folder is expanded, showing a list of tables including 'dbo.DATOSEMPLEO', 'dbo.DATOSPERSONALES', 'dbo.SUCURSAL', and 'dbo.Table_1'.
- Query Designer:** Displays a query result for the table 'LAPTOP-ED6K7OJ\Datos Generales'. The query is a simple SELECT statement. The result set shows columns: CODIGO, FK_DATOSP..., FK_SUCURS..., CARGO, HORARIO, and SALARIO. The data is as follows:

CODIGO	FK_DATOSP...	FK_SUCURS...	CARGO	HORARIO	SALARIO
222	357	1200	ventas	8 horas	9000
333	951	1421	cajero	8 horas	6000
999	852	1411	ventas	8 horas	9500
777	789	0167	gerente	8 horas	16000
888	123	1200	auxiliar	8 horas	5500
111	456	1421	gerente	8 horas	32000
444	987	1411	cajero	8 horas	6500
666	975	0167	ventas	8 horas	10000
555	351	1200	ventas	8 horas	9850
1010	761	1200	ggg	5 horas	40000
NULL	NULL	NULL	NULL	NULL	NULL

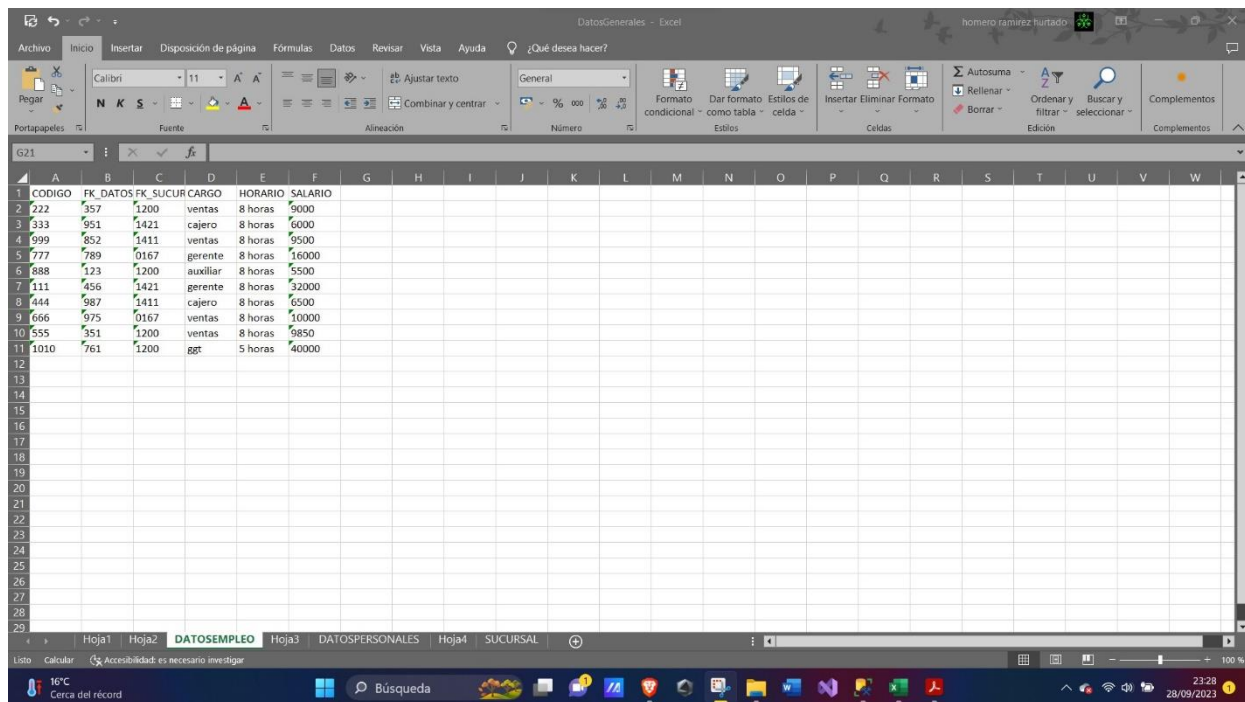
The bottom status bar shows the current page: '1 of 10'.

The right-hand pane shows the 'Properties' window for the selected table. The 'Identity' property is set to 'On', and the 'Destination Table' is 'dbo.DATOSEMPLEO'. The 'Query Designer' pane shows the query text: 'SELECT * FROM dbo.DATOSEMPLEO'.

[illegible]



Importacion de Base de Datos a Excel.



DatosGenerales - Excel

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda ¿Qué desea hacer?

Calibri 11 Fuente Alineación Número Formato condicional Dar formato como tabla Estilos de celda Insertar Eliminar Formato Autosuma Rellenar Ordenar y filtrar Buscar y seleccionar Complementos

Portapapeles

E21

ID	NOMBRE	DIRECCION	CIUDAD	NUM. TELEFONO	EMAIL	GENERO	H	I	J	K	L	M	N	O	P
357	ALEXA	VASCO DE QUIROGA	CIUDAD DE MEXICO		ALEXA@GMAIL.COM	MUJER									
951	LUIS	GALEANA	CIUDAD DE MEXICO		LUIS@GMAIL.COM	HOMBRE									
852	BRAYAN	JALAPA	CIUDAD DE MEXICO		BRAYAN@GMAIL.COM	HOMBRE									
789	SILVIA	LA CONCHITA	CIUDAD DE MEXICO		SILVIA@GMAIL.COM	MUJER									
123	OSCAR	PUEBLO NUEVO	CIUDAD DE MEXICO		OSCAR@GMAIL.COM	HOMBRE									
456	CARLOS	EJERCITO DE AGUA PRIETA	CIUDAD DE MEXICO		CARLOS@GMAIL.COM	HOMBRE									
987	BEATRIZ	SANTA FE	CIUDAD DE MEXICO		BEATRIZ@GMAIL.COM	MUJER									
975	ALEJANDRA	CAÑADA	CIUDAD DE MEXICO		ALEJANDRA@GMAIL.COM	MUJER									
351	IOVANNY	ESCUADRON 201	CIUDAD DE MEXICO		IOVANNY@GMAIL.COM	HOMBRE									
761	GERARDO	ESTRELLA	CIUDAD DE MEXICO		GERARDO@GMAIL.COM	HOMBRE									

Hoja1 Hoja2 DATOSEMPLEO Hoja3 DATOSPERSONALES Hoja4 SUCURSAL

Listo Calcular Accesibilidad: es necesario investigar

16°C Cerca del récord

Búsqueda

23:28 28/09/2023

DatosGenerales - Excel

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda ¿Qué desea hacer?

Calibri 11 Fuente Alineación Número Formato condicional Dar formato como tabla Estilos de celda Insertar Eliminar Formato Autosuma Rellenar Ordenar y filtrar Buscar y seleccionar Complementos

Portapapeles

F16

ID	NOMBRE_DE_SI	CIUDAD	DIRECCION	CATIDAD_VENTAS_ANUALES	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1200	vasco	cdmx	vasco1651	300000000															
1421	santafe	cdmx	vasco1445	100000000															
1411	sanantonio	cdmx	vasco1000	2000000000															
167	sanangel	cdmx	obs200	300000000															

Hoja1 Hoja2 DATOSEMPLEO Hoja3 DATOSPERSONALES Hoja4 SUCURSAL

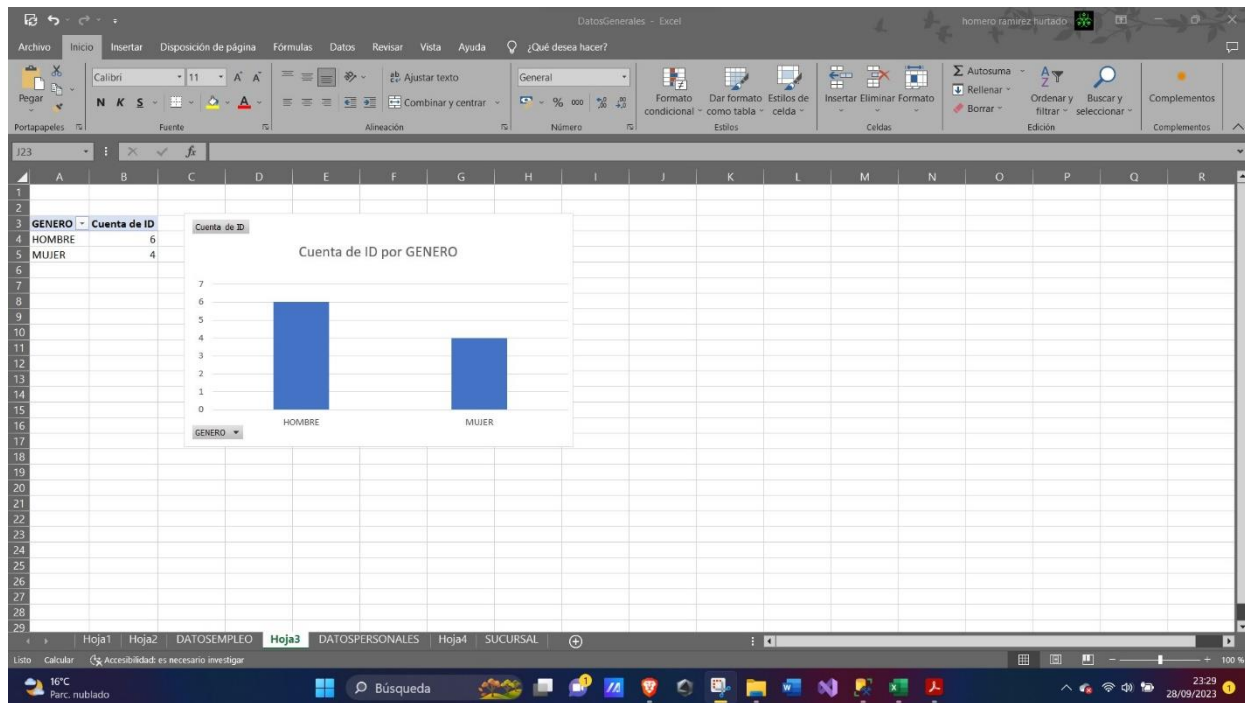
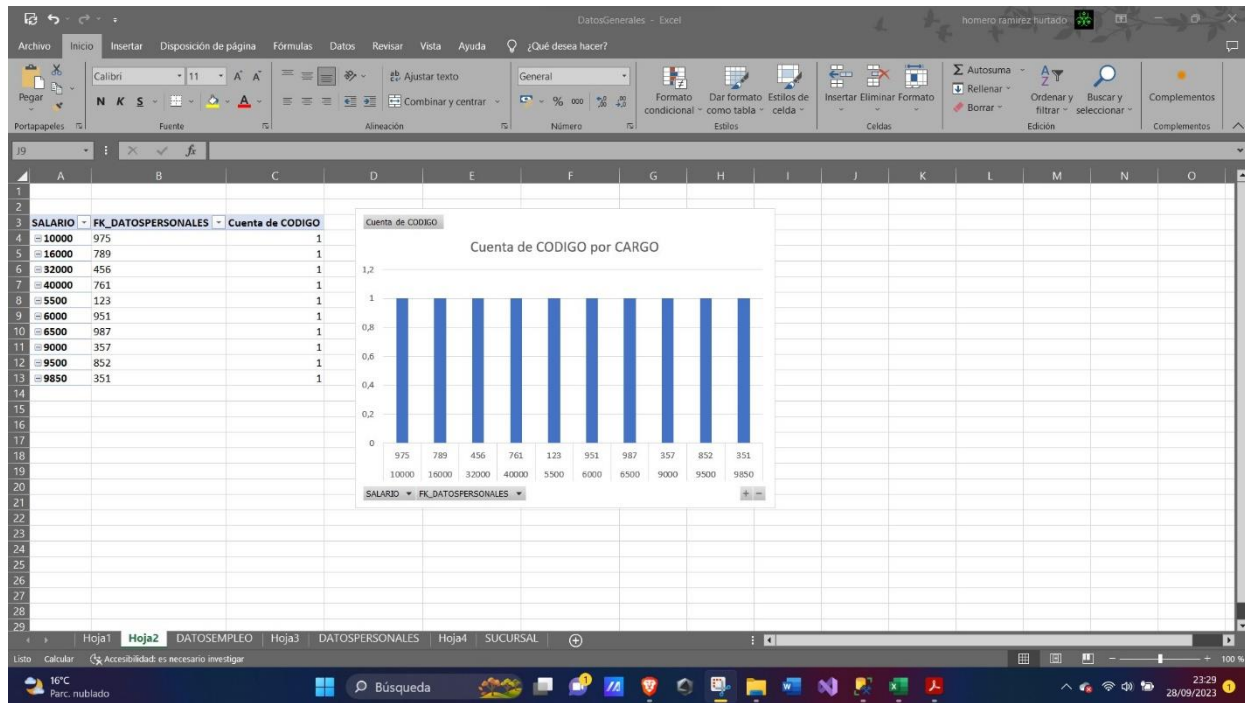
Listo Calcular Accesibilidad: es necesario investigar

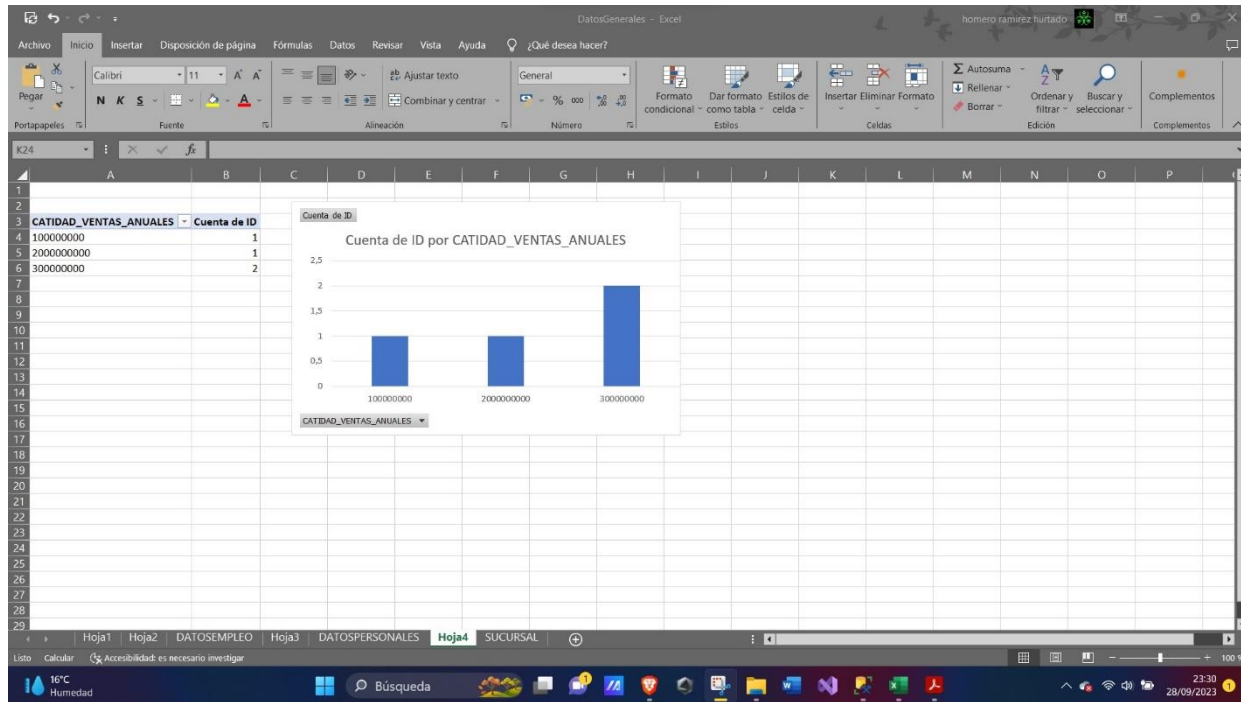
16°C Cerca del récord

Búsqueda

23:28 28/09/2023

Graficos.





Preguntas.

¿Quién es la persona que gana más?

R: Gerardo.

¿Cuál es la sucursal que gana menos anualmente?

R: Sante Fe.

¿Cuántos empleados son hombres?

R: 6 hombres.

Conclusión.

En conclusión, es importante realizar este tipo de procesos ya que con ello podemos observar perfiles los cuales podemos identificar como por ejemplo sueldos, gustos etc. Y con ello podemos ser mas precisos en cuanto a poder ofrecer un servicio, producto o incluso conocer mejor al usuario, empresa o cualquier tipo de cosa que queramos saber con el fin de poder ser mas precisos en lo que se desee realizar ya que estos datos son muy importantes ya sea para consumo o bien realizar estadísticas con ello podemos conocer mejor cualquier tipo de dato, con ello aprendí a realizar esta actividad y espero poder seguir aprendiéndolo y ponerlo en práctica.