

Actividad | 1 | Cubos OLAP

Minería de datos

Ingeniería en Desarrollo de Software



TUTOR: Félix Acosta Hernández

ALUMNO: Marco Antonio García Sánchez

FECHA: 07/06/2024

ÍNDICE

Portada	1
Índice	2
Introducción, descripción	3
Justificación	4
Desarrollo	5-7
Conclusión y referencias	8

INTRODUCCIÓN

La minería de datos es una técnica esencial en el análisis de grandes volúmenes de información, permitiendo extraer patrones y conocimientos ocultos que pueden ser críticos para la toma de decisiones estratégicas. En la era de la información, donde las organizaciones generan y almacenan datos a una velocidad sin precedentes, la capacidad de transformar estos datos en información útil se ha vuelto fundamental. La minería de datos ayuda a identificar tendencias, prever comportamientos futuros y descubrir relaciones complejas que no son evidentes a simple vista. Esto es particularmente importante en sectores como la salud, las finanzas, el marketing y la seguridad, donde los insights obtenidos pueden mejorar la eficiencia operativa, personalizar ofertas para clientes, detectar fraudes y salvar vidas. Además, la minería de datos facilita el desarrollo de modelos predictivos que pueden anticipar necesidades y riesgos, permitiendo a las empresas y organizaciones mantenerse competitivas y proactivas en un entorno cada vez más dinámico y orientado a los datos.

DESCRIPCIÓN

La empresa CompuVentas, dedicada a la venta de hardware pide crear una nueva base de datos llamada “datosgenerales” en Microsoft SQL Server Management con las siguientes tablas:

Tabla 1: datospersonales

Columnas:

- ID
- Nombre
- Dirección
- Ciudad
- Num. Teléfono
- Email
- Género

Tabla 2: datosempleado

Columnas:

- Código
- FK_datospersonales
- FK_Sucursal
- Cargo
- Horario
- Salario

Tabla 3: Sucursal

Columnas:

- ID
- Nombre sucursal
- Ciudad
- Dirección
- Cantidad_ventas_anuales

JUSTIFICACIÓN

La minería de datos es crucial en la era actual debido a varios factores que subrayan su importancia:

1. **Análisis de Grandes Volúmenes de Información:** La capacidad de manejar y analizar grandes conjuntos de datos es vital en un mundo donde las organizaciones generan y almacenan enormes cantidades de información diariamente.
2. **Extracción de Patrones y Conocimientos Ocultos:** La minería de datos facilita la identificación de patrones y relaciones ocultas dentro de los datos, proporcionando un conocimiento profundo que puede transformar la forma en que una organización opera.
3. **Mejora en la Toma de Decisiones:** Al convertir datos crudos en información útil, la minería de datos ayuda a las organizaciones a tomar decisiones más informadas y estratégicas, optimizando recursos y estrategias para mejorar el rendimiento y la eficiencia.
4. **Previsión de Comportamientos Futuros:** Las técnicas de minería de datos permiten prever tendencias y comportamientos futuros, lo cual es esencial para anticiparse a las necesidades del mercado y de los clientes, manteniéndose así competitivas.
5. **Eficiencia Operativa:** Identificar áreas de mejora en los procesos operativos es otra ventaja de la minería de datos, permitiendo aumentar la eficiencia y reducir costos.
6. **Modelos Predictivos:** Facilita el desarrollo de modelos predictivos que pueden anticipar necesidades y riesgos, manteniendo a las organizaciones proactivas y preparadas para el futuro.

DESARROLLO

A continuación, anexo las imágenes de la base de datos creada y también sus respectivas tablas de acuerdo a la información solicitada por la actividad:

```
create database DatosGenerales  
  
use datosgenerales
```

Creamos nuestra base de datos y utilizamos el código **USE** para modificarla.

```
CREATE TABLE datosempleado (  
    codigo INT PRIMARY KEY,  
    datospersonales INT,  
    sucursal INT,  
    cargo NVARCHAR(100),  
    horario NVARCHAR(100),  
    salario DECIMAL(10, 2)  
);  
  
INSERT INTO datosempleado (Codigo, datospersonales, sucursal, Cargo, Horario, Salario)  
VALUES  
(1, 1, 101, 'Gerente', '9:00 AM - 5:00 PM', 5000.00),  
(2, 2, 102, 'Asistente', '10:00 AM - 6:00 PM', 3000.00),  
(3, 3, 103, 'Desarrollador', '9:00 AM - 6:00 PM', 4000.00),  
(4, 4, 104, 'Diseñadora', '8:00 AM - 4:00 PM', 3500.00),  
(5, 5, 105, 'Analista', '9:00 AM - 5:00 PM', 3800.00),  
(6, 6, 106, 'Soporte Técnico', '11:00 AM - 7:00 PM', 3200.00),  
(7, 7, 107, 'Marketing', '10:00 AM - 6:00 PM', 3400.00),  
(8, 8, 108, 'Ventas', '9:00 AM - 5:00 PM', 3600.00),  
(9, 9, 109, 'Contador', '8:00 AM - 4:00 PM', 3700.00),  
(10, 10, 110, 'Recursos Humanos', '9:00 AM - 5:00 PM', 3300.00);
```

Creamos e insertamos información en nuestra tabla de datos empleado.

```
CREATE TABLE sucursal (  
    ID INT PRIMARY KEY,  
    sucursal NVARCHAR(100),  
    ciudad NVARCHAR(100),  
    direccion NVARCHAR(200),  
    Cantidad_ventas_anuales INT  
);  
  
INSERT INTO Sucursal (ID, sucursal, Ciudad, Direccion, Cantidad_ventas_anuales)  
VALUES  
(101, 'Sucursal A', 'Ciudad de México', 'Calle Falsa 123', 100000),  
(102, 'Sucursal B', 'Guadalajara', 'Avenida Siempre Viva 742', 200000),  
(103, 'Sucursal C', 'Monterrey', 'Boulevard de los Sueños 1', 150000),  
(104, 'Sucursal D', 'Puebla', 'Calle Luna 456', 180000),  
(105, 'Sucursal E', 'Cancún', 'Avenida del Sol 789', 170000),  
(106, 'Sucursal F', 'Tijuana', 'Calle Estrella 101', 160000),  
(107, 'Sucursal G', 'Mérida', 'Avenida Primavera 202', 140000),  
(108, 'Sucursal H', 'Querétaro', 'Calle Otoño 303', 190000),  
(109, 'Sucursal I', 'León', 'Boulevard Invierno 404', 130000),  
(110, 'Sucursal J', 'Chihuahua', 'Calle Verano 505', 210000);
```

De la misma manera creamos e insertamos información en la tabla de sucursal.

```

CREATE TABLE datospersonales (
  ID INT PRIMARY KEY,
  nombre NVARCHAR(100),
  direccion NVARCHAR(200),
  ciudad NVARCHAR(100),
  telefono NVARCHAR(20),
  email NVARCHAR(100),
  genero NVARCHAR(10)
);

INSERT INTO datospersonales (ID, nombre, direccion, ciudad, telefono, email, genero)
VALUES
(1, 'Juan Pérez', 'Calle Falsa 123', 'Ciudad A', '555-1234', 'juan.perez@ejemplo.com', 'Masculino'),
(2, 'María Gómez', 'Avenida Siempre Viva 742', 'Ciudad B', '555-5678', 'maria.gomez@ejemplo.com', 'Femenino'),
(3, 'Carlos Ruiz', 'Boulevard de los Sueños 1', 'Ciudad C', '555-8765', 'carlos.ruiz@ejemplo.com', 'Masculino'),
(4, 'Ana Fernández', 'Calle Luna 456', 'Ciudad D', '555-2345', 'ana.fernandez@ejemplo.com', 'Femenino'),
(5, 'Luis Martínez', 'Avenida del Sol 789', 'Ciudad E', '555-3456', 'luis.martinez@ejemplo.com', 'Masculino'),
(6, 'Elena Rodríguez', 'Calle Estrella 101', 'Ciudad F', '555-4567', 'elena.rodriguez@ejemplo.com', 'Femenino'),
(7, 'Miguel Torres', 'Avenida Primavera 202', 'Ciudad G', '555-5679', 'miguel.torres@ejemplo.com', 'Masculino'),
(8, 'Laura Sánchez', 'Calle Otoño 303', 'Ciudad H', '555-6789', 'laura.sanchez@ejemplo.com', 'Femenino'),
(9, 'José Ramírez', 'Boulevard Invierno 404', 'Ciudad I', '555-7890', 'jose.ramirez@ejemplo.com', 'Masculino'),
(10, 'Carmen López', 'Calle Verano 505', 'Ciudad J', '555-8901', 'carmen.lopez@ejemplo.com', 'Femenino');

```

Por último la tabla de datos personales.

Ahora que los datos fueron ingresados validamos con un query si los datos se ingresaron de manera correcta.

```

select * from sucursal
select * from datospersonales
select * from datosempleado

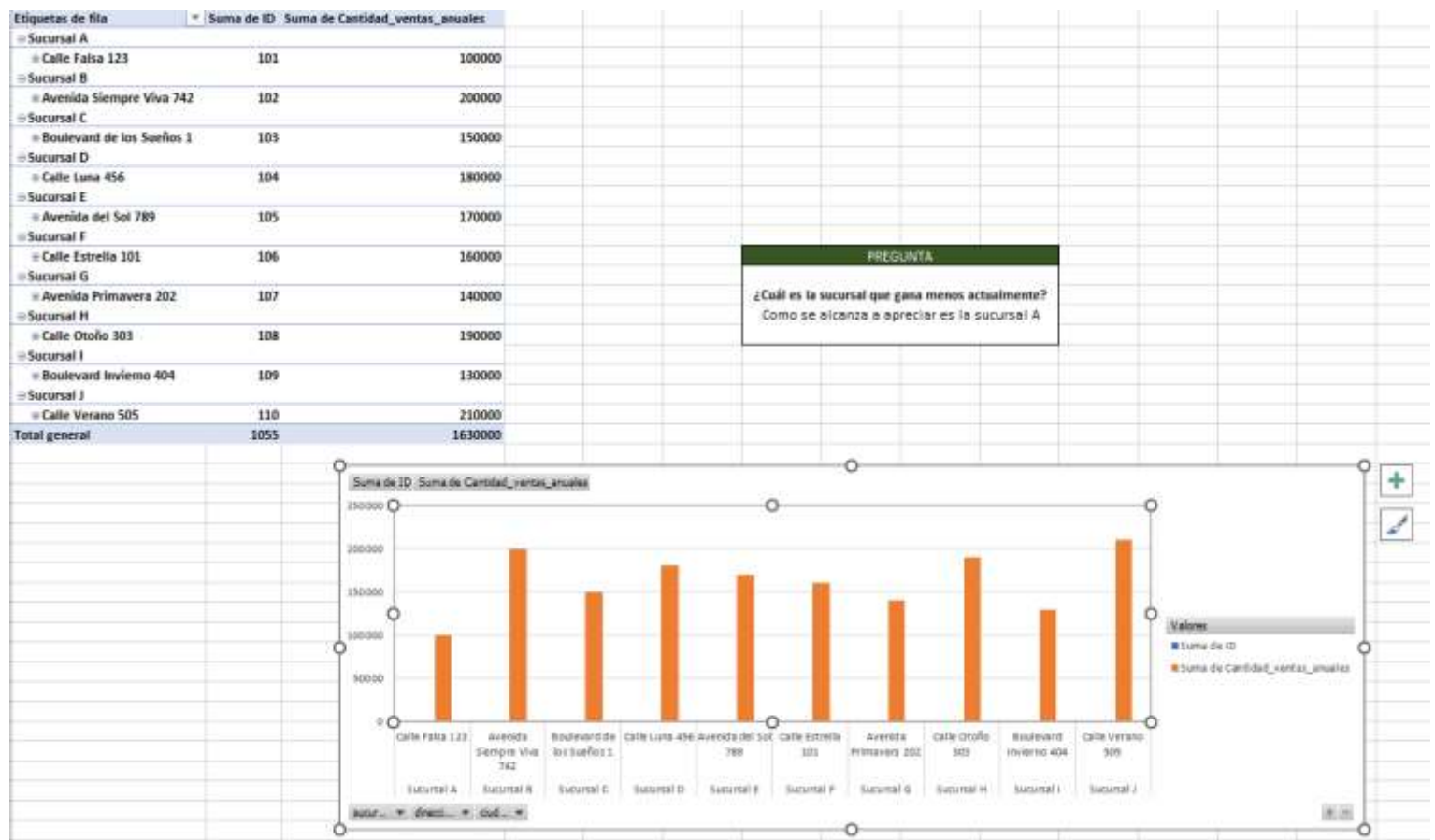
```

	ID	sucursal	ciudad	direccion	Cantidad_ventas_anuales
1	101	Sucursal A	Ciudad de México	Calle Falsa 123	100000
2	102	Sucursal B	Guadalajara	Avenida Siempre Viva 742	200000
3	103	Sucursal C	Monterrey	Boulevard de los Sueños 1	150000
4	104	Sucursal D	Puebla	Calle Luna 456	180000
5	105	Sucursal E	Cancún	Avenida del Sol 789	170000
6	106	Sucursal F	Tijuana	Calle Estrella 101	160000
7	107	Sucursal G	Mérida	Avenida Primavera 202	140000
8	108	Sucursal H	Querétaro	Calle Otoño 303	190000

	ID	nombre	direccion	ciudad	telefono	email	genero
1	1	Juan Pérez	Calle Falsa 123	Ciudad A	555-1234	juan.perez@ejemplo.com	Masculino
2	2	María Gómez	Avenida Siempre Viva 742	Ciudad B	555-5678	maria.gomez@ejemplo.com	Femenino
3	3	Carlos Ruiz	Boulevard de los Sueños 1	Ciudad C	555-8765	carlos.ruiz@ejemplo.com	Masculino
4	4	Ana Fernández	Calle Luna 456	Ciudad D	555-2345	ana.fernandez@ejemplo.com	Femenino
5	5	Luis Martínez	Avenida del Sol 789	Ciudad E	555-3456	luis.martinez@ejemplo.com	Masculino
6	6	Elena Rodríguez	Calle Estrella 101	Ciudad F	555-4567	elena.rodriguez@ejemplo.com	Femenino
7	7	Miguel Torres	Avenida Primavera 202	Ciudad G	555-5679	miguel.torres@ejemplo.com	Masculino
8	8	Laura Sánchez	Calle Otoño 303	Ciudad H	555-6789	laura.sanchez@ejemplo.com	Femenino

	codigo	datospersonales	sucursal	cargo	horario	salario
1	1	1	101	Gerente	9:00 AM - 5:00 PM	5000.00
2	2	2	102	Asistente	10:00 AM - 6:00 PM	3000.00
3	3	3	103	Desarr...	9:00 AM - 6:00 PM	4000.00
4	4	4	104	Diseña...	8:00 AM - 4:00 PM	3500.00
5	5	5	105	Analista	9:00 AM - 5:00 PM	3800.00
6	6	6	106	Soport...	11:00 AM - 7:00 PM	3200.00
7	7	7	107	Marketi...	10:00 AM - 6:00 PM	3400.00
8	8	8	108	Ventas	9:00 AM - 5:00 PM	3600.00
9	9	9	109	Contador	8:00 AM - 4:00 PM	3700.00
10	10	10	110	Recurs...	9:00 AM - 5:00 PM	3300.00

Posteriormente de todo esto, ingresamos los datos a excel. Creando una tabla dinamica y colocando los datos de nuestro servidor. Al hacer esto la info se desglosa de la siguiente manera como ejemplo:



A continuación, realizaremos un análisis de datos de acuerdo a las tablas que tenemos, sin embargo esto se anexará en el otro documento Excel con el fin de contestar ciertas preguntas.

CONCLUSIÓN

La conclusión del análisis de datos resalta su importancia fundamental en la toma de decisiones informadas y estratégicas en diversas industrias. A través del análisis de datos, las organizaciones pueden extraer información valiosa de grandes volúmenes de datos, identificando patrones y tendencias ocultas que impulsan la innovación y la eficiencia operativa. Este proceso no solo mejora la comprensión del comportamiento del mercado y del cliente, sino que también facilita la previsión de tendencias futuras, permitiendo a las empresas adaptarse proactivamente a los cambios del entorno.

El análisis de datos optimiza la gestión de recursos, reduce costos y mitiga riesgos, proporcionando una ventaja competitiva significativa. En sectores como la salud, finanzas, marketing y seguridad, el análisis de datos es crucial para mejorar los resultados y la satisfacción del cliente. En resumen, el análisis de datos es una herramienta indispensable que transforma datos en conocimientos prácticos, impulsando la toma de decisiones basada en evidencia y fortaleciendo la posición estratégica de las organizaciones en un mundo cada vez más orientado a la información.

REFERENCIAS

Gil, J. (2023, 16 noviembre).
Tableau: la importancia del análisis de datos.
Blog de Hiberus.

<https://www.hiberus.com/crecemos-contigo/tableau-la-importancia-del-analisis-de-datos/>