

Evaluación de tercer parcial

ADMINISTRACIÓN DE BASE DE DATOS

8IDS3

PRESENTA.

Cortes Soto Alex Saul. Hernández Cureño Marc Anthony.

NOMBRE DEL PROFESOR: Miguel Ángel Mendoza Flores FECHA DE ENTREGA: 30/Marzo/2025 Opción 1: Diseñar un esquema de base de datos que respete la integridad referencial (uso de claves primarias y foráneas), Crear las tablas necesarias e importar el dataset.

Crear la base de datos

```
CREATE DATABASE delitos db;
-- Conectar a la base de datos
\c delitos db;
   Crear tabla de alcaldías/municipios
CREATE TABLE ubicaciones (
  id SERIAL PRIMARY KEY,
  alcaldia VARCHAR(255),
  municipio VARCHAR(255)
);
   Crear tabla de delitos
CREATE TABLE delitos (
  id SERIAL PRIMARY KEY,
  nombre VARCHAR(255),
  categoria VARCHAR(255)
);
   Crear tabla de fiscalías
CREATE TABLE fiscalias (
  id SERIAL PRIMARY KEY,
```

nombre VARCHAR(255)

```
);
   Crear tabla de agencias
CREATE TABLE agencias (
  id SERIAL PRIMARY KEY,
  nombre VARCHAR(255)
);
   Crear tabla de unidades de investigación
CREATE TABLE unidades_investigacion (
  id SERIAL PRIMARY KEY,
  nombre VARCHAR(255)
);
   Crear tabla principal de carpetas de investigación
CREATE TABLE carpetas_investigacion (
  id BIGINT PRIMARY KEY,
  fecha_inicio DATE,
  hora_inicio TIME,
  fecha_hecho DATE,
  hora hecho TIME,
  delito_id INT REFERENCES delitos(id),
  fiscalia_id INT REFERENCES fiscalias(id),
  agencia_id INT REFERENCES agencias(id),
  unidad_id INT REFERENCES unidades_investigacion(id),
  ubicacion_id INT REFERENCES ubicaciones(id),
```

```
latitud DECIMAL(9,6),
  longitud DECIMAL(9,6)
);
Esto nos asegura la integridad y nos ayuda a optimizar las consultas.
Importar los archivos CSV en PostgreSQL
COPY ubicaciones(alcaldia, municipio)
FROM '/ruta/ubicaciones.csv'
DELIMITER ',' CSV HEADER;
COPY delitos(nombre, categoria)
FROM '/ruta/delitos.csv'
DELIMITER ',' CSV HEADER;
COPY fiscalias(nombre)
FROM '/ruta/fiscalias.csv'
DELIMITER ',' CSV HEADER;
COPY agencias(nombre)
FROM '/ruta/agencias.csv'
DELIMITER ',' CSV HEADER;
COPY unidades_investigacion(nombre)
FROM '/ruta/unidades_investigacion.csv'
```

DELIMITER ',' CSV HEADER;

COPY carpetas_investigacion(id, fecha_inicio, hora_inicio, fecha_hecho, hora_hecho, delito_id, fiscalia_id, agencia_id, unidad_id, ubicacion_id, latitud, longitud)

FROM '/ruta/carpetas_investigacion.csv'

DELIMITER ',' CSV HEADER;

1- Exactitud: Verificar errores tipográficos y duplicados en los folios

Consulta para encontrar folios duplicados:

SELECT id, COUNT(*)

FROM carpetas_investigacion

GROUP BY id

HAVING COUNT(*) > 1;

Ejemplo de salida:

Identificación.	Contar.
10234	2
20567	3

Significa que los folios 10234 y 20567 están duplicados y deben corregirse.

• Consulta para detectar errores tipográficos en alcaldías:

SELECT alcaldia, COUNT(*)

FROM ubicaciones

GROUP BY alcaldia

ORDER BY COUNT(*) DESC;

Alcaldía	Contar
Benito Juárez	1500
Benito Juárez	50
Cuauhtémoc	1800
Cuauhtémoc	30

El nombre "Benito Juárez" y "Benito Juarez" deben unificarse, al igual que "Cuauhtémoc" y "Cuauhtémoc".

2- Integridad: Datos nulos o incorrectos

• Consulta para detectar valores nulos en campos clave:

SELECT *

FROM carpetas_investigacion

WHERE fecha_inicio IS NULL OR hora_inicio IS NULL

OR fecha_hecho IS NULL OR hora_hecho IS NULL

OR delito_id IS NULL OR fiscalia_id IS NULL

OR agencia_id IS NULL OR unidad_id IS NULL

OR ubicacion id IS NULL;

Identifi cación	Fecha _inicio	Hora_ inicio	Fecha_ hecho	Hora_ hecho	Delit o_id	Fiscal ia_id	Agen cia_id	ID de uni dad	Ubicac ión_id
30567	NULO	14:30	15 de enero de 2024	10:00	2	5	8	3	12
40987	10 de marzo de 2024	NULO	09- 03- 2024	11:00	3	7	4	2	18

Estos registros contienen valores nulos y deben corregirse.

3- Disponibilidad: Consultas eficientes

• Consulta para delitos más frecuentes:

SELECT d.nombre, COUNT(*) AS cantidad

FROM carpetas_investigacion ci

JOIN delitos d ON ci.delito_id = d.id

GROUP BY d.nombre

ORDER BY cantidad DESC

LIMIT 10;

Delito	Cantidad
Robo a transeúnte	5000
Fraude electrónico	3200
Robo de vehículo	2800

• Consulta para casos por alcaldía:

SELECT u.alcaldia, COUNT(*) AS casos

FROM carpetas_investigacion ci

JOIN ubicaciones u ON ci.ubicacion_id = u.id

GROUP BY u.alcaldia

ORDER BY casos DESC;

• Ejemplo de salida:

Alcaldía	Casos
Iztapalapa	4500
Gustavo A. Madero	3800
Cuauhtémoc	3200

4- Confiabilidad: Coherencia de fechas

 Consulta para verificar si la fecha del hecho es posterior a la fecha de inicio:

SELECT id, fecha_inicio, fecha_hecho, (fecha_inicio - fecha_hecho) AS dias_diferencia

FROM carpetas_investigacion

WHERE fecha_inicio < fecha_hecho;

Identificación	Fecha_inicio	Fecha_hecho	Diferencia de días
12345	25 de marzo de 2024	28 de marzo de 2024	-3

Este registro es un error porque la fecha de investigación ocurre antes de la fecha del delito.

5- Utilidad: Análisis de delitos y zonas de riesgo

• Consulta para delitos más frecuentes por categoría:

SELECT d.categoria, d.nombre, COUNT(*) AS cantidad

FROM carpetas_investigacion ci

JOIN delitos d ON ci.delito id = d.id

GROUP BY d.categoria, d.nombre

ORDER BY cantidad DESC;

Ejemplo de salida:

Categoría	Delito	Cantidad
Violencia	Robo con violencia	4200
Cibernético	Fraude electrónico	3200
Contra propiedad	Robo de autopartes	2900

Consulta para zonas con más delitos:

SELECT u.alcaldia, COUNT(*) AS casos

FROM carpetas_investigacion ci

JOIN ubicaciones u ON ci.ubicacion_id = u.id

GROUP BY u.alcaldia

ORDER BY casos DESC

LIMIT 10;

Alcaldía	Casos
Iztapalapa	4500
Cuauhtémoc	3800
Benito Juárez	2800

6- Accesibilidad: Índices y vistas

• Índice para acelerar búsquedas por alcaldía:

CREATE INDEX idx_alcaldia ON carpetas_investigacion(ubicacion_id);

• Vista para delitos frecuentes:

CREATE VIEW vista_delitos_frecuentes AS

SELECT d.nombre, COUNT(*) AS cantidad

FROM carpetas_investigacion ci

JOIN delitos d ON ci.delito_id = d.id

GROUP BY d.nombre;

• Consulta en la vista:

SELECT * FROM vista_delitos_frecuentes;

Ejemplo de salida:

Delito	Cantidad
Robo con violencia	4500
Fraude electrónico	3200

7- Pertinencia: Evaluar la relevancia de los datos

Los datos permiten:

- Identificar zonas peligrosas.
- Conocer los delitos más reportados.
- Evaluar eficacia de fiscales y agencias.
- ♣ Tomar decisiones de seguridad pública.

8- Usabilidad: Consulta fácil para usuarios

• Consulta para delitos por alcaldía:

SELECT ci.id, ci.fecha_inicio, d.nombre AS delito, u.alcaldia, u.municipio

FROM carpetas_investigacion ci

JOIN delitos d ON ci.delito_id = d.id

JOIN ubicaciones u ON ci.ubicacion_id = u.id

WHERE u.alcaldia = 'Benito Juárez'

ORDER BY ci.fecha_inicio DESC;

Ejemplo de salida:

Identificación	Fecha_inicio	Delito	Alcaldía Municipio
98765	25 de marzo de	Robo con	Benito CDMX
	2024	violencia	Juárez
87654	24 de marzo de	Fraude	Benito CDMX
	2024	electrónico	Juárez

Conclusión.

El análisis de calidad de datos en la base de delitos nos permitió evaluar diversos aspectos fundamentales. En cuanto a la exactitud, identificamos posibles errores tipográficos y duplicados en los folios, lo que evidencia la necesidad de limpieza y estandarización. La integridad se vio afectada por valores nulos en campos claves y fechas incorrectas, lo que podría impactar en la confiabilidad de los registros. Evaluamos la disponibilidad mediante consultas eficientes que permitieron conocer los delitos más frecuentes y su distribución geográfica, asegurando que la información sea accesible y útil para la toma de decisiones. La confiabilidad de los datos se comprobó al detectar inconsistencias temporales, como fechas de investigación anteriores a la fecha del delito. La utilidad se destacó en la clasificación de delitos por categorías y en la identificación de zonas de alto riesgo, información clave para las autoridades. Para mejorar la accesibilidad, se propone la creación de índices y vistas que optimizan la consulta de datos. La pertinencia se evidencia en la capacidad del sistema para responder preguntas relevantes sobre seguridad pública, mientras que la usabilidad se mejoró mediante consultas amigables que facilitan el acceso a la información. En conclusión, este análisis no solo permite detectar y corregir inconsistencias en los datos, sino que también optimiza su aprovechamiento, garantizando información confiable y accesible para la toma de decisiones estratégicas en materia de seguridad.