

UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA
MODALIDAD PRESENCIAL



Facultad de Ingenierías y Arquitectura

COMPUTACIÓN

INFORME FINAL: PROYECTO INTEGRADOR

Docente:

Ing. Ángel Eduardo Encalada Encalada

Integrantes:

Ronal Rodrigo Álvarez Alvarado

Renata Alejandra Maldonado Bravo

Fecha de entrega:

28 de junio, 2024.

Loja – Ecuador

Contenido

1. Introducción	3
2. Desarrollo	3
2.1 Repositorio en GitHub	3
2.2 Datos base.....	4
2.3 Datos complementarios	7
Descripción y utilidad de cada dato.....	9
2.4 Posibles análisis a realizar	11
2.5 Variables complementarias:.....	11
1. Tasa Neta de Asistencia	11
2. Media del Índice de Desarrollo Humano.....	12
3. Tasa de Empleo Inadecuado	12
4. Tasa de Pobreza Extrema.....	13
5. Tráfico para Consumo Interno	13
2.6 Diseño conceptual	14
Diccionario de datos	15
2.7 Diseño lógico relacional.....	16
Diccionario de datos	16
2.8 Implementación y carga	17
2.9 Script SQL.....	17
2.10 Descarga en formato CSV.....	22
6. Conclusión	22
3. Bibliografía	24
4. Anexos.....	25
<i>Link del repositorio de GitHub</i>	<i>25</i>
Evidencia carga de datos	25
Capturas de Pantalla	28
Código.....	28

1. Introducción

La modelación de bases de datos juega un papel crucial en la organización y análisis de información relevante para diversas disciplinas. Este proyecto integrador se centra en la creación de una base de datos detallada de detenidos, abarcando el periodo comprendido entre los años 2016 y 2024. La motivación principal detrás de este proyecto radica en la aplicación práctica de los conocimientos adquiridos a lo largo del ciclo educativo, con el propósito de poner a prueba nuestras habilidades en la modelación de bases de datos y en el manejo efectivo de información compleja.

El objetivo general es desarrollar una estructura de base de datos robusta que permita almacenar y gestionar de manera eficiente los datos relacionados con detenidos durante el periodo mencionado. Esto no solo implica la creación de tablas y relaciones adecuadas, sino también la incorporación de variables adicionales relevantes para enriquecer el análisis posterior. Las mismas que serán identificadas y seleccionadas mediante una investigación adicional, asegurando que contribuyan significativamente al entendimiento de los patrones y tendencias en el sistema de detenciones.

La metodología empleada abarcará desde el uso inicial de los datos base hasta la implementación práctica del modelo de base de datos. Se utilizarán herramientas y técnicas adecuadas para la modelación conceptual, lógica y física de la base de datos, asegurando su coherencia y eficiencia. El alcance del proyecto incluirá la creación de consultas y análisis preliminares que permitan validar la funcionalidad y la utilidad del sistema diseñado.

2. Desarrollo

2.1 Repositorio en GitHub

[Link al repositorio.](#)

2.2 Datos base

Tabla 1

Descripción de Datos Base

Nombre del dato	Descripción	Tipo de dato	Año
Relacionados al detenido			
edad	Edad del detenido.	Int	2016 – 2024
sexo	Sexo del detenido/a.	Cadena/String	2016 – 2024
genero	Género del detenido/a.	Cadena/String	2023 – 2024
nacionalidad	Procedencia del detenido/a.	Cadena/String	2016 – 2024
estado_civil	Estado civil del detenido/a.	Cadena/String	2023 – 2024
nivel_instruccion	Nivel de educación del detenido/a.	Cadena/String	2023 – 2024
estatus_migratorio	Estatus migratorio del detenido/a.	Cadena/String	2023 – 2024
num_detenciones	Número total de detenciones del detenido.	Int	2022 – 2024
autoidentificacion_etnica	Autoidentificación étnica del detenido/a.	Cadena/String	2022 – 2024
codigo_iccs	Número de serie o identificador único e internacional que aparece en la tarjeta SIM del detenido/a.	Cadena/String	2023 – 2024
Relacionados con la aprehensión/detención			
fecha_detencion	Fecha en la que ocurrió la aprehensión/detención.	Date	2016 – 2024
hora_detencion	Hora en la que ocurrió la aprehensión/detención.	Time	2022 – 2024
tipo	Indica si la persona es aprehendida/detenida.	Cadena/String	2022 – 2024
condicion	Estado del aprehendido/detenido.	Cadena/String	2023 – 2024

movilizacion	Indica cómo se transportaba la persona en el momento de la aprehensión/detención.	Cadena/String	2023 – 2024
presunta_flagrancia	Si se encontró flagrantemente al aprehendido/detenido.	Cadena/String	2022 – 2024
presunta_modalidad	Indica la forma en la que la persona cometió la infracción o la ley que incumplió.	Cadena/String	2023 – 2024
flagrante_boleta	Indica si la aprehensión/detención de la persona es de tipo flagrante o boleta.	Cadena/String	2016 – 2018
Sub-relacionados con la aprehensión/detenidos			
arma	Indica el arma que se usó para cometer el crimen.	Cadena/String	2023 – 2024
tipo_arma	Indica de qué tipo es el arma que se usó para cometer el crimen.	Cadena/String	2023 – 2024
infraccion	Indica la infracción que se cometió de forma general.	Cadena/String	2022 – 2024
subinfraccion	Indica la sub-infracción de forma más detallada.	Cadena/String	2016 – 2024
Relacionados al lugar de aprehensión/detención			
codigo_provincia	Describe el código de la provincia donde ocurrió la detención/aprehensión.	Int	2023 – 2024
nombre_provincia	Describe el nombre de la provincia donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024

codigo_canton	Describe el código del cantón donde ocurrió la detención/aprehensión.	Cadena/String	2023 – 2024
nombre_canton	Describe el nombre del cantón donde ocurrió la detención/aprehensión.	Cadena/String	2016 – 2024
codigo_parroquia	Describe el código de la parroquia donde ocurrió la detención/aprehensión.	Cadena/String	2023 – 2024
nombre_parroquia	Describe el nombre de la parroquia donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
nombre_zona	Describe el nombre de la zona donde ocurrió la detención/aprehensión.	Cadena/String	2016 – 2024
nombre_subzona	Describe el nombre de la subzona donde ocurrió la detención/aprehensión.	Cadena/String	2016 – 2024
codigo_distrito	Describe el código del distrito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
nombre_distrito	Describe el nombre del distrito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
codigo_circuito	Describe el código del circuito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
nombre_circuito	Describe el nombre del circuito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
codigo_subcircuito	Describe el código del subcircuito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024

nombre_subcircuito	Describe el nombre del subcircuito donde ocurrió la detención/aprehensión.	Cadena/String	2022 – 2024
lugar	Describe el lugar donde ocurrió la detención/aprehensión de forma más específica.	Cadena/String	2022 – 2024
tipo_lugar	Clasifica el lugar donde ocurrió la detención/aprehensión de forma más específica.	Cadena/String	2023 – 2024

2.3 Datos complementarios

Tabla 2

Descripción de los Datos Complementarios

Nombre del dato	Descripción	Tipo de dato
media_tasa_educacionbasica_poblacion	Promedio de la tasa de población que ha completado la educación básica.	FLOAT
media_tasa_educacionbachillerato_poblacion	Promedio de la tasa de población que ha completado el bachillerato.	FLOAT
media_tasa_empleo_inadecuado	Promedio de la tasa de empleo inadecuado (trabajos que no cumplen con las condiciones mínimas).	FLOAT
media_tasa_pobreza_extrema	Promedio de la tasa de población que	FLOAT

	vive en condiciones de pobreza extrema.	
media_indice_dh	Promedio del índice de desarrollo humano (DH).	FLOAT
media_tasa_narcotrafico	Promedio de los kilogramos de narcóticos incautados.	FLOAT

La violencia y la criminalidad en Ecuador están aumentando a un ritmo sin precedentes. No solo se encuentra entre los diez países más violentos del mundo, sino que también tiene la tasa más alta de muertes violentas en América Latina, con 47,25 por cada 100 000 habitantes, lo que representa ocho veces más que en 2016, cuando experimentó su tasa más baja desde 1980. Durante menos de diez años, ha cambiado de ser el segundo país más seguro de América del Sur, después de Chile, a ser el más violento. (OECD, 2024)

Para el análisis y con el objetivo de proporcionar una visión más completa de las condiciones sociales, educativas y económicas para complementar el análisis de los datos base, se han incorporado datos complementarios obtenidos mediante el cálculo de la media de los valores registrados entre los años 2016 y 2024. Este enfoque permite suavizar las fluctuaciones anuales y proporcionar una perspectiva más estable y representativa de cada indicador.

La fórmula para encontrar este valor es:

$$media = \frac{\sum_{i=2016}^{2024} dato_i}{número\ de\ años}$$

Descripción y utilidad de cada dato

- **Promedio de la tasa de población que solo ha completado la educación básica:**

Este dato es crucial para entender el nivel educativo básico de la población detenida. Un mayor porcentaje de personas que solo han completado la educación básica puede indicar una falta de acceso a oportunidades educativas avanzadas, lo que a menudo está relacionado con mayores tasas de criminalidad. Mejorar el acceso a la educación fundamental puede contribuir a reducir la reincidencia y proporcionar mejores oportunidades laborales para los ex-detenedos.

Fuente: INEC, Censo Ecuador

<https://censoecuador.ecudatanalytics.com/>

- **Promedio de la tasa de población que solo ha completado el bachillerato:**

Evaluar este dato es esencial para comprender el nivel educativo medio de la población detenida. Aunque un mayor porcentaje de personas con educación secundaria puede reflejar una mayor cualificación laboral, también puede indicar que, a pesar de haber alcanzado este nivel educativo, no han logrado acceder a oportunidades que eviten la delincuencia. La finalización del bachillerato no garantiza automáticamente mejores condiciones de vida, pero puede proporcionar una base sobre la cual construir habilidades adicionales y mejorar la reintegración social y económica de los detenidos.

Fuente: INEC, Censo Ecuador

<https://censoecuador.ecudatanalytics.com/>

- **Promedio de la tasa de empleo inadecuado (trabajos que no cumplen con las condiciones mínimas):** Este dato ayuda a medir la calidad del empleo en la población. Una tasa alta de empleo inadecuado puede señalar problemas

estructurales en el mercado laboral, como la falta de trabajos decentes y la necesidad de políticas que promuevan empleos de calidad.

Fuente: Ministerio de Inclusión Económica y Social

<https://info.inclusion.gob.ec/index.php/caracterizacion-poblacion-objetivo-ancusrext/indicadores-laborales-ancusrext/tasa-de-empleo-inadecuado-ancusrext>

- **Promedio de la tasa de población que vive en condiciones de pobreza extrema:**
Sirve para evaluar y entender el nivel de pobreza severa en una población. Una alta tasa de pobreza extrema puede indicar la necesidad urgente de intervenciones sociales y políticas para mejorar las condiciones de vida y proporcionar una red de seguridad social. La pobreza extrema afecta gravemente la salud, la educación y las oportunidades de las personas.

Fuente: INEC, Compendio Estadístico

<https://www.ecuadorencifras.gob.ec/compendio-estadistico-2016/>

<https://censoecuador.ecudatanalytics.com/>

- **Promedio del índice de desarrollo humano (DH):** El índice de desarrollo humano es una medida compuesta de indicadores de esperanza de vida, educación e ingresos per cápita. Este índice proporciona una visión general del bienestar y la calidad de vida de la población, siendo un indicador para políticas de desarrollo sostenible. Un alto índice de DH refleja una mayor equidad y bienestar social.

Fuente: Ministerio de Inclusión Económica y Social, , Compendio Estadístico

<https://censoecuador.ecudatanalytics.com/>

Índice de Desarrollo Humano en Ecuador

- **Promedio de los kilogramos de drogas incautadas para consumo dentro del país:** Este dato es fundamental para evaluar la magnitud del problema del narcotráfico en una región. Una alta cantidad de drogas incautadas puede indicar

una prevalencia significativa del narcotráfico y el consumo de drogas dentro del país, lo cual está asociado con altos niveles de violencia, inseguridad y problemas sociales. Este indicador es crucial para diseñar y evaluar políticas de combate al narcotráfico y reducción del consumo de drogas.

Fuente: Observatorio Ecuatoriano de Crimen Organizado

<https://oeco.padf.org/visualizador-de-datos-resultados-antinarcoticos/>

2.4 Posibles análisis a realizar

Algunas posibles ideas de análisis que se podrían realizar con los datos incluyen:

- **Análisis de tendencias:** Identificación de tendencias en los datos de detenciones a lo largo del tiempo.
- **Análisis geográfico:** Mapeo de las detenciones para identificar áreas con mayores incidencias.
- **Análisis demográfico:** Estudio de las características demográficas de las personas detenidas.
- **Análisis de correlación:** Exploración de posibles correlaciones entre diferentes variables (por ejemplo, tipo de delito y ubicación).

2.5 Variables complementarias:

1. Tasa Neta de Asistencia

Posible Consulta: Relación entre la Tasa de Asistencia Escolar y la Delincuencia por Provincia

Descripción: Evaluar si las provincias con diferentes tasas de asistencia escolar tienen variaciones en el número y tipo de delitos.

Ejemplo:

- Comparar la tasa de asistencia en Azuay (91.6% en educación básica, 79.5% en bachillerato) con el número de delitos relacionados con drogas en esa provincia.
- Si Azuay tiene una alta tasa de asistencia escolar y un bajo número de delitos relacionados con drogas, podría sugerir una correlación negativa entre educación y criminalidad.

2. Media del Índice de Desarrollo Humano

Posible Consulta: Correlación entre el Índice de Desarrollo Humano y la Incidencia Delictiva

Descripción: Analizar cómo el índice de desarrollo humano (IDH) en cada provincia se relaciona con la cantidad de delitos.

Ejemplo:

- Comparar el IDH de Azuay (0.801) con el número de delitos cometidos allí.
- Si Azuay, con un alto IDH, tiene menos delitos en comparación con Bolívar (IDH 0.712333333), podría indicar que un mayor desarrollo humano se asocia con menor criminalidad.

3. Tasa de Empleo Inadecuado

Posible Consulta: Impacto del Empleo Inadecuado en la Criminalidad

Descripción: Investigar si las provincias con una alta tasa de empleo inadecuado tienen un mayor número de delitos.

Ejemplo:

- Comparar la tasa de empleo inadecuado en Carchi (57%) con el número de delitos relacionados con tráfico de drogas en esa provincia.

- Si Carchi, con una alta tasa de empleo inadecuado, también muestra un alto número de delitos, esto podría sugerir una relación entre desempleo inadecuado y criminalidad.

4. Tasa de Pobreza Extrema

Posible Consulta: Relación entre la Pobreza Extrema y la Actividad Delictiva

Descripción: Evaluar si las provincias con diferentes tasas de pobreza extrema tienen variaciones en el número y tipo de delitos.

Ejemplo:

- Comparar la tasa de pobreza extrema en Bolívar (10.2%) con el número de delitos relacionados con robo en esa provincia.
- Si Bolívar tiene una alta tasa de pobreza extrema y también muestra altos niveles de delitos como robos, esto podría sugerir una correlación entre pobreza extrema y ciertos tipos de delitos.

5. Tráfico para Consumo Interno

Posible Consulta: Cantidad de Drogas Incautadas y su Relación con la Delincuencia

Descripción: Comparar la cantidad de drogas incautadas en cada provincia con el número de delitos relacionados con drogas.

Ejemplo:

- Comparar la cantidad de drogas incautadas en Carchi (2753.47 kg) con el número de delitos relacionados con tráfico de drogas en esa provincia.
- Si Carchi, con una alta incautación de drogas, muestra un alto número de delitos relacionados, esto podría indicar la eficacia o ineficacia de las medidas de control de drogas en la reducción de la criminalidad.

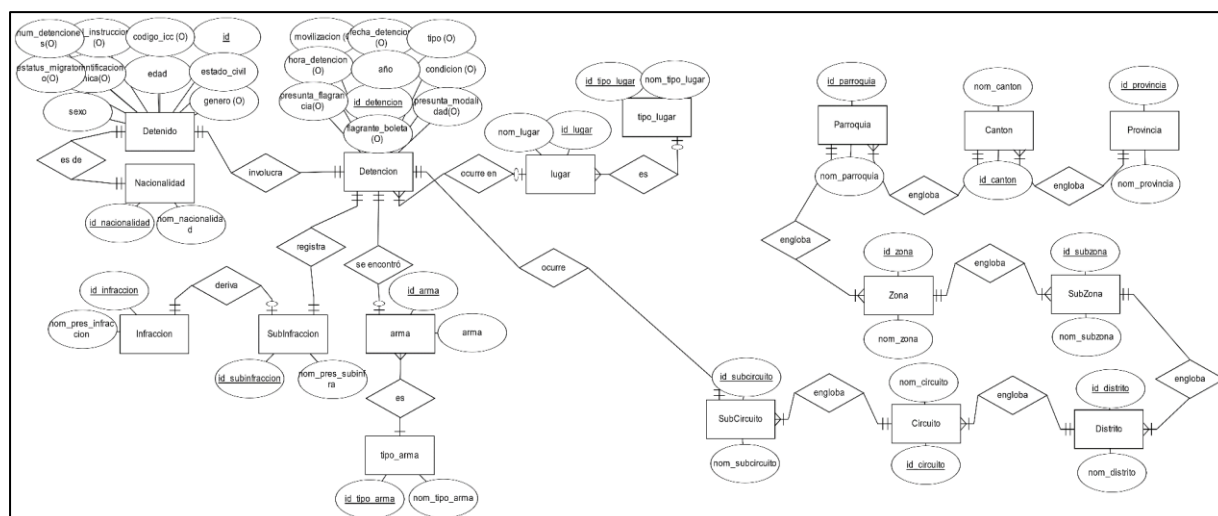
Los datos del Barómetro de las Américas de 2023, una encuesta probabilística de cobertura nacional que se realiza cada dos años en más de 26 países del continente, muestran que alrededor del 36% de los ecuatorianos en edad de votar fueron víctimas de algún delito en los 12 meses previos a la encuesta. Esto significa que alrededor de 5 millones de personas en Ecuador habrían sido objeto de robo, hurto, agresión, fraude, chantaje, extorsión, amenazas u otros delitos. Estas estadísticas indican que cada hora, más de 500 ciudadanos ecuatorianos son víctimas de la violencia.

“Ecuador terminó el año 2023 con un total de 7.878 crímenes ocurridos hasta el 31 de diciembre de 2023, según cifras oficiales, lo que significa una tasa de 46,5 asesinatos, homicidios, femicidios y sicarios por cada cien mil habitantes. Es la tasa más alta de la historia nacional” (Ramos, 2024)

2.6 Diseño conceptual

Ilustración 1

Modelo conceptual



Diccionario de datos

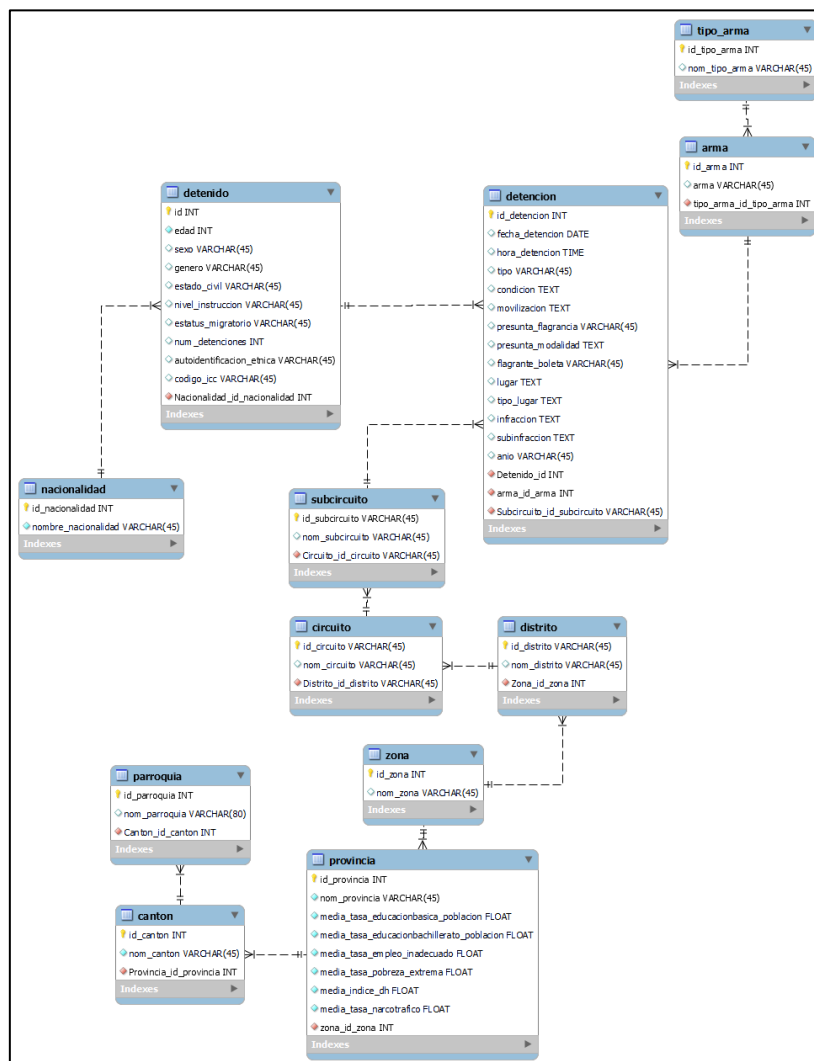
Tabla 3*Diccionario de datos del Modelo Relacional*

Tipo Entidad/Relación	Atributo	Restricción de Dominio Adicional	Valor por defecto
Detencion	tipo	Solo los valores de “DETENIDO”, “APREHENDIDO” y “SIN DATO”.	“SIN DATO”
Detencion	presunta_flagrancia	Solo los valores de “SI”, “NO” y “SIN DATO”.	“SIN DATO”
Detencion	flagrante_boleta	Solo los valores de “FLAGRANTE”, “BOLETA” y “SIN DATO”.	“SIN DATO”
Detenido	num_detenciones	Valores mayores o iguales a 0, se usará “-1” para indicar “SIN DATO”.	-1

2.7 Diseño lógico relacional

Ilustración 2

Modelo lógico - relacional



Diccionario de datos

Tabla 4

Diccionario de datos del Modelo Lógico

Tabla	Campo	Restricción específica de dominio	Valor por defecto
detencion	tipo	Solo los valores de “DETENIDO”, “APREHENDIDO” y “SIN DATO”.	“SIN DATO”

detencion	flagrante_boleta	Solo los valores de “FLAGRANTE”, “BOLETA” y “SIN DATO”.	“SIN DATO”
detencion	presunta_flagrancia	Solo los valores de “SI”, “NO” y “SIN DATO”.	“SIN DATO”
detenido	num_detenciones	Valores mayores o iguales a 0, se usará “-1” para indicar “SIN DATO”.	-1

2.8 Implementación y carga

Carga mediante código:

- Se utilizó principalmente este método para la carga de grandes volúmenes de datos, ya que resultó ser más rápido que hacerlo a través de Workbench. Aunque Workbench es más intuitivo, el proceso tomaba más tiempo. Al realizar la carga mediante código, logramos reducir significativamente el tiempo que normalmente tardaría.

2.9 Script SQL

```
-- Crear un nuevo usuario.
CREATE USER 'pi_user'@'localhost' IDENTIFIED BY 'renata15';

-- Conceder permisos.
GRANT ALL PRIVILEGES ON pi_db.* TO 'pi_user'@'localhost';

-- Aplicar los cambios
FLUSH PRIVILEGES;

-- Crear la base de datos y tablas
CREATE DATABASE IF NOT EXISTS `pi_db`
  DEFAULT CHARACTER SET = utf8mb4
  COLLATE = utf8mb4_0900_ai_ci;
```

```

USE `pi_db`;

CREATE TABLE IF NOT EXISTS `tipo_arma` (
  `id_tipo_arma` INT NOT NULL AUTO_INCREMENT,
  `nom_tipo_arma` VARCHAR(45) NULL DEFAULT NULL,
  PRIMARY KEY (`id_tipo_arma`))
ENGINE = InnoDB
AUTO_INCREMENT = 1
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `arma` (
  `id_arma` INT NOT NULL AUTO_INCREMENT,
  `arma` VARCHAR(45) NULL DEFAULT NULL,
  `tipo_arma_id_tipo_arma` INT NOT NULL,
  PRIMARY KEY (`id_arma`),
  INDEX `fk_arma_tipo_arma1_idx` (`tipo_arma_id_tipo_arma` ASC) VISIBLE,
  CONSTRAINT `fk_arma_tipo_arma1`
    FOREIGN KEY (`tipo_arma_id_tipo_arma`)
    REFERENCES `tipo_arma` (`id_tipo_arma`))
ENGINE = InnoDB
AUTO_INCREMENT = 1
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `zona` (
  `id_zona` INT NOT NULL,
  `nom_zona` VARCHAR(45) NULL DEFAULT NULL,
  PRIMARY KEY (`id_zona`))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `provincia` (
  `id_provincia` INT NOT NULL,
  `nom_provincia` VARCHAR(45) NOT NULL,
  `media_tasa_educacionbasica_poblacion` FLOAT NOT NULL,
  `media_tasa_educacionbachillerato_poblacion` FLOAT NOT NULL,
  `media_tasa_empleo_inadecuado` FLOAT NOT NULL,
  `media_tasa_pobreza_extrema` FLOAT NOT NULL,
  `media_indice_dh` FLOAT NOT NULL,
  `media_tasa_narcotrafico` FLOAT NOT NULL,
  `zona_id_zona` INT NOT NULL,
  PRIMARY KEY (`id_provincia`),
  INDEX `fk_provincia_zona1_idx` (`zona_id_zona` ASC) VISIBLE,
  CONSTRAINT `fk_provincia_zona1`
    FOREIGN KEY (`zona_id_zona`)
    REFERENCES `zona` (`id_zona`)
    ON DELETE NO ACTION

```

```

ON UPDATE NO ACTION)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `canton` (
  `id_canton` INT NOT NULL,
  `nom_canton` VARCHAR(45) NOT NULL,
  `Provincia_id_provincia` INT NOT NULL,
  PRIMARY KEY (`id_canton`),
  INDEX `fk_Canton_Provincial_idx` (`Provincia_id_provincia` ASC) VISIBLE,
  CONSTRAINT `fk_Canton_Provincial`
    FOREIGN KEY (`Provincia_id_provincia`)
    REFERENCES `provincia` (`id_provincia`)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `distrito` (
  `id_distrito` VARCHAR(45) NOT NULL,
  `nom_distrito` VARCHAR(45) NULL DEFAULT NULL,
  `Zona_id_zona` INT NOT NULL,
  PRIMARY KEY (`id_distrito`),
  INDEX `fk_Distrito_Zonal_idx` (`Zona_id_zona` ASC) VISIBLE,
  CONSTRAINT `fk_Distrito_Zonal`
    FOREIGN KEY (`Zona_id_zona`)
    REFERENCES `zona` (`id_zona`)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `circuito` (
  `id_circuito` VARCHAR(45) NOT NULL,
  `nom_circuito` VARCHAR(45) NULL DEFAULT NULL,
  `Distrito_id_distrito` VARCHAR(45) NOT NULL,
  PRIMARY KEY (`id_circuito`),
  INDEX `fk_Circuito_Distrito1_idx` (`Distrito_id_distrito` ASC) VISIBLE,
  CONSTRAINT `fk_Circuito_Distrito1`
    FOREIGN KEY (`Distrito_id_distrito`)
    REFERENCES `distrito` (`id_distrito`)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `nacionalidad` (
  `id_nacionalidad` INT NOT NULL AUTO_INCREMENT,
  `nombre_nacionalidad` VARCHAR(45) NOT NULL,
  PRIMARY KEY (`id_nacionalidad`)

```

```

ENGINE = InnoDB
AUTO_INCREMENT = 1
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `detenido` (
  `id` INT NOT NULL AUTO_INCREMENT,
  `edad` INT NOT NULL,
  `sexo` VARCHAR(45) NULL DEFAULT NULL,
  `genero` VARCHAR(45) NULL DEFAULT NULL,
  `estado_civil` VARCHAR(45) NULL DEFAULT NULL,
  `nivel_instruccion` VARCHAR(45) NULL DEFAULT NULL,
  `estatus_migratorio` VARCHAR(45) NULL DEFAULT NULL,
  `num_detenciones` INT NOT NULL DEFAULT '-1',
  `autoidentificacion_etnica` VARCHAR(45) NULL DEFAULT NULL,
  `codigo_icc` VARCHAR(45) NULL DEFAULT NULL,
  `Nacionalidad_id_nacionalidad` INT NOT NULL,
  PRIMARY KEY (`id`),
  INDEX `fk_Detenido_Nacionalidad_idx` (`Nacionalidad_id_nacionalidad` ASC)
VISIBLE,
  CONSTRAINT `fk_Detenido_Nacionalidad`
    FOREIGN KEY (`Nacionalidad_id_nacionalidad`)
    REFERENCES `nacionalidad` (`id_nacionalidad`))
ENGINE = InnoDB
AUTO_INCREMENT = 1
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `subcircuito` (
  `id_subcircuito` VARCHAR(45) NOT NULL,
  `nom_subcircuito` VARCHAR(45) NULL DEFAULT NULL,
  `Circuito_id_circuito` VARCHAR(45) NOT NULL,
  PRIMARY KEY (`id_subcircuito`),
  INDEX `fk_Subcircuito_Circuito1_idx` (`Circuito_id_circuito` ASC)
VISIBLE,
  CONSTRAINT `fk_Subcircuito_Circuito1`
    FOREIGN KEY (`Circuito_id_circuito`)
    REFERENCES `circuito` (`id_circuito`))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `detencion` (
  `id_detencion` INT NOT NULL AUTO_INCREMENT,
  `fecha_detencion` DATE NULL DEFAULT NULL,
  `hora_detencion` TIME NULL DEFAULT NULL,
  `tipo` VARCHAR(45) NOT NULL DEFAULT 'SIN DATO',
  `condicion` TEXT NULL DEFAULT NULL,
  `movilizacion` TEXT NULL DEFAULT NULL,

```

```

`presunta_flagrancia` VARCHAR(45) NOT NULL DEFAULT 'SIN DATO',
`presunta_modalidad` TEXT NULL DEFAULT NULL,
`flagrante_boleta` VARCHAR(45) NOT NULL DEFAULT 'SIN DATO',
`lugar` TEXT NULL DEFAULT NULL,
`tipo_lugar` TEXT NULL DEFAULT NULL,
`infraccion` TEXT NULL DEFAULT NULL,
`subinfraccion` TEXT NULL DEFAULT NULL,
`anio` VARCHAR(45) NULL DEFAULT NULL,
`Detenido_id` INT NOT NULL,
`arma_id_arma` INT NOT NULL,
`Subcircuito_id_subcircuito` VARCHAR(45) NOT NULL,
PRIMARY KEY (`id_detencion`),
INDEX `fk_Detencion_Detenidol_idx` (`Detenido_id` ASC) VISIBLE,
INDEX `fk_Detencion_armal_idx` (`arma_id_arma` ASC) VISIBLE,
INDEX `fk_Detencion_Subcircuitol_idx` (`Subcircuito_id_subcircuito` ASC)
VISIBLE,
CONSTRAINT `fk_Detencion_armal`
  FOREIGN KEY (`arma_id_arma`)
  REFERENCES `arma` (`id_arma`),
CONSTRAINT `fk_Detencion_Detenidol`
  FOREIGN KEY (`Detenido_id`)
  REFERENCES `detenido` (`id`),
CONSTRAINT `fk_Detencion_Subcircuitol`
  FOREIGN KEY (`Subcircuito_id_subcircuito`)
  REFERENCES `subcircuito` (`id_subcircuito`)),
CONSTRAINT `tipo_restrict`
  CHECK (`tipo` IN ('DETENIDO', 'APREHENDIDO', 'SIN DATO')),
CONSTRAINT `flagrante_boleta_restrict`
  CHECK (`flagrante_boleta` IN ('FLAGRANTE', 'BOLETA', 'SIN DATO')),
CONSTRAINT `presunta_flagrancia_restrict`
  CHECK (`presunta_flagrancia` IN ('SI', 'NO', 'SIN DATO'))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

CREATE TABLE IF NOT EXISTS `parroquia` (
  `id_parroquia` INT NOT NULL AUTO_INCREMENT,
  `nom_parroquia` VARCHAR(80) NULL DEFAULT NULL,
  `Canton_id_canton` INT NOT NULL,
  PRIMARY KEY (`id_parroquia`),
  INDEX `fk_Parroquia_Canton1_idx` (`Canton_id_canton` ASC) VISIBLE,
  CONSTRAINT `fk_Parroquia_Canton1`
    FOREIGN KEY (`Canton_id_canton`)
    REFERENCES `canton` (`id_canton`))
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8mb4
COLLATE = utf8mb4_0900_ai_ci;

ALTER TABLE `detenido`

```

```
MODIFY COLUMN `num_detenciones` INT NOT NULL DEFAULT '-1',  
ADD CONSTRAINT `num_detenciones_restrict`  
CHECK (`num_detenciones` >= -1);  
  
SET SQL_MODE=@OLD_SQL_MODE;  
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;  
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;
```

2.10 Descarga en formato CSV

[Link para descargar csv.](#)

6. Conclusión

El presente proyecto se ha centrado en unificar y analizar exhaustivamente múltiples fuentes de datos sobre detenciones desde 2016 hasta 2024. A través de la aplicación de diversas herramientas y tecnologías, se logró convertir, normalizar y unificar datos procedentes de distintos años y formatos en una estructura coherente y analizable. Este proceso abarcó:

Preparación y Conversión de Datos: Se utilizó csvkit para convertir hojas de Excel a CSV y scripts de partición para manejar datos de múltiples años.

Normalización de Esquemas: Se crearon y adaptaron esquemas específicos para asegurar consistencia en las columnas a lo largo de los diferentes años.

Procesamiento con Spark y Python: Se emplearon Spark y Scala para datos iniciales, y Python con Pandas para datos más recientes, garantizando limpieza y transformación de datos para compatibilidad.

Uso del Formato Parquet: Se implementó Parquet para almacenamiento eficiente y rápido acceso a los datos unificados.

Diseño Conceptual y Lógico: Se desarrolló un Modelo Entidad-Relación (E-R) transformado en un Modelo Relacional para estructurar la base de datos.

Generación e Implementación del Script DDL: Se creó un script SQL para implementar la base de datos en MySQL.

Preparación y Carga de Datos: Se importaron datos base y complementarios en la base de datos MySQL.

Generación de un Archivo CSV: Se creó un archivo CSV con todas las variables relevantes para análisis.

El resultado obtenido es una base de datos MySQL completamente implementada y cargada con los datos finales, lista para análisis detallados. Esta base de datos facilitará la realización de análisis de tendencias, geográficos, demográficos y de correlación, proporcionando valiosos *insights* sobre las detenciones a lo largo de los años.

3. Bibliografía

OECD. (28 de Junio de 2024). *¿El paraíso perdido? Tráfico de armas de fuego y violencia en Ecuador*. Obtenido de OECD: <https://oecd.padf.org/el-paraiso-perdido-trafico-de-armas-de-fuego-y-violencia-en-ecuador/>

Ramos, X. (2 de Enero de 2024). *7.878 crímenes en 2023, solo 584 resueltos: ¿Qué está pasando en Ecuador?* Obtenido de El Universo: <https://www.eluniverso.com/noticias/seguridad/el-91-de-los-crimenes-ocurridos-durante-2023-estan-bajo-investigacion-es-decir-no-han-sido-resueltos-nota/>

reeenatamc/ProyectoIntegrador: Repositorio del Proyecto Integrador de Base de Datos. (n.d.). Retrieved June 29, 2024, from <https://github.com/reeenatamc/ProyectoIntegrador>

4. Anexos

Link del repositorio de GitHub

<https://github.com/reenatamc/ProyectoIntegrador>

Link para descargar CSV

<https://drive.google.com/drive/folders/1JGDS0WNvXIIMe4yxRojmzTLL8-VKZag?usp=sharing>

Evidencia carga de datos

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Table: detencion_copia

Columns:

- id_detencion
- fecha_detencion
- hora_detencion
- tipo
- condicion
- movilizacion
- presunta_flagancia
- presunta_modalidad
- flagrante_boleta
- lugar
- tipo_lugar
- inf

Result Grid

id_detencion	fecha_detencion	hora_detencion	tipo	condicion	movilizacion	presunta_flagancia	presunta_modalidad	flagrante_boleta	lugar	tipo_lugar	inf
9991	2016-03-12	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9992	2016-03-13	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9993	2016-03-14	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9994	2016-03-14	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9995	2016-03-14	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9996	2016-03-14	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9997	2016-03-01	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9998	2016-03-08	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
9999	2016-03-10	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN
10000	2016-03-16	00:00:00	SIN DATO	SIN DATO	SIN DATO	SI	SIN DATO	SIN DATO	SIN DATO	SIN DATO	SIN

Output

Action Output

Time: 22:27:19 Action: SELECT COUNT(*) AS total_filas FROM detencion_copia LIMIT 0, 10000

Message: 1 row(s) returned

Duration / Fetch: 0.000 sec / 0.000 sec

Time: 22:28:15 Action: SELECT * FROM definalproy.detencion_copia LIMIT 0, 10000

Message: 10000 row(s) returned

Duration / Fetch: 0.000 sec / 0.063 sec

Tabla “Detenido”:

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: pi_db

Result Grid

COUNT(*)
0

Output

Action Output

Time: 09:44:37 Action: SELECT COUNT(*) FROM DETENIDOS LIMIT 0, 10000

Message: Error Code: 1146, Table 'pi_db.detenido' doesn't exist

Duration / Fetch: 0.016 sec

Time: 09:45:12 Action: SET SQL_MODE=@OLD_SQL_MODE;

Message: 1 row(s) returned

Duration / Fetch: 0.156 sec / 0.000 sec

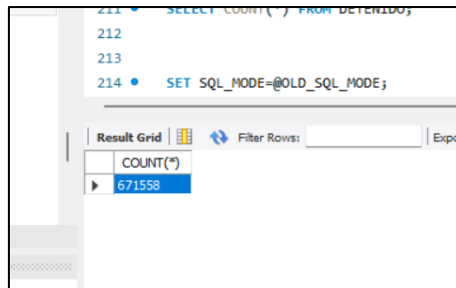
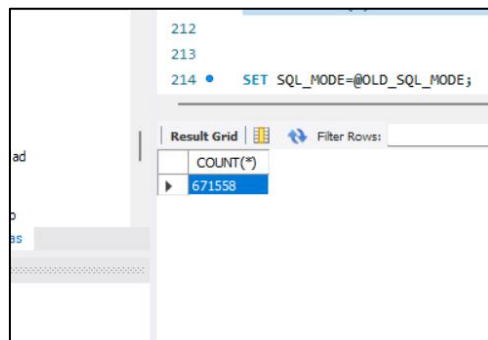
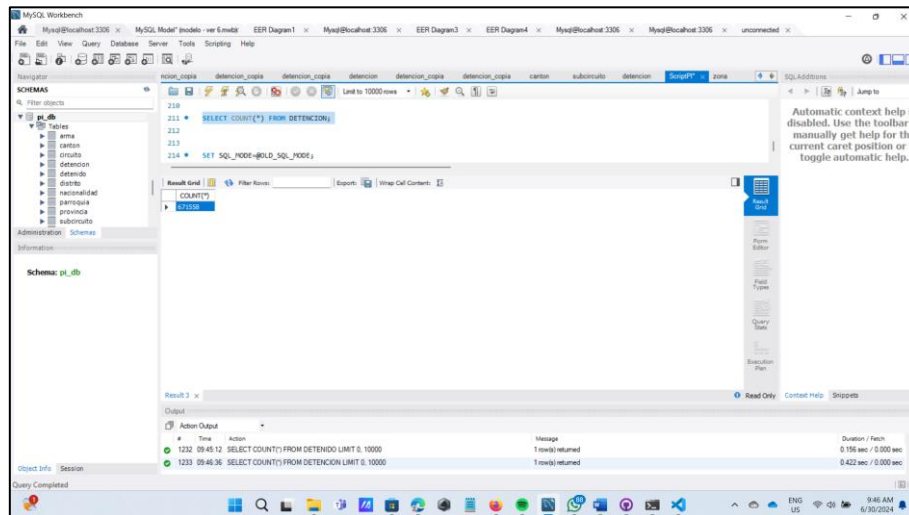
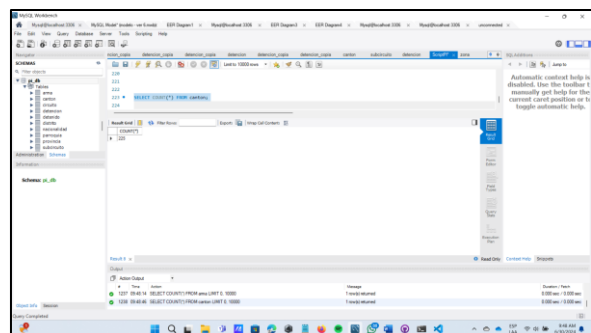
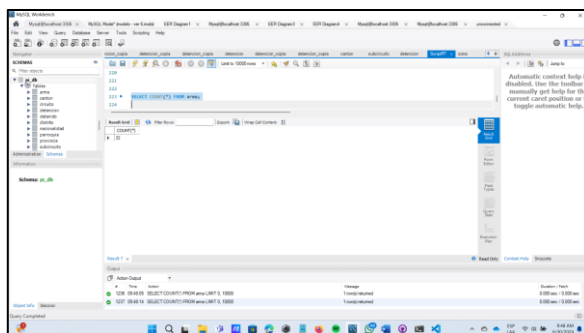


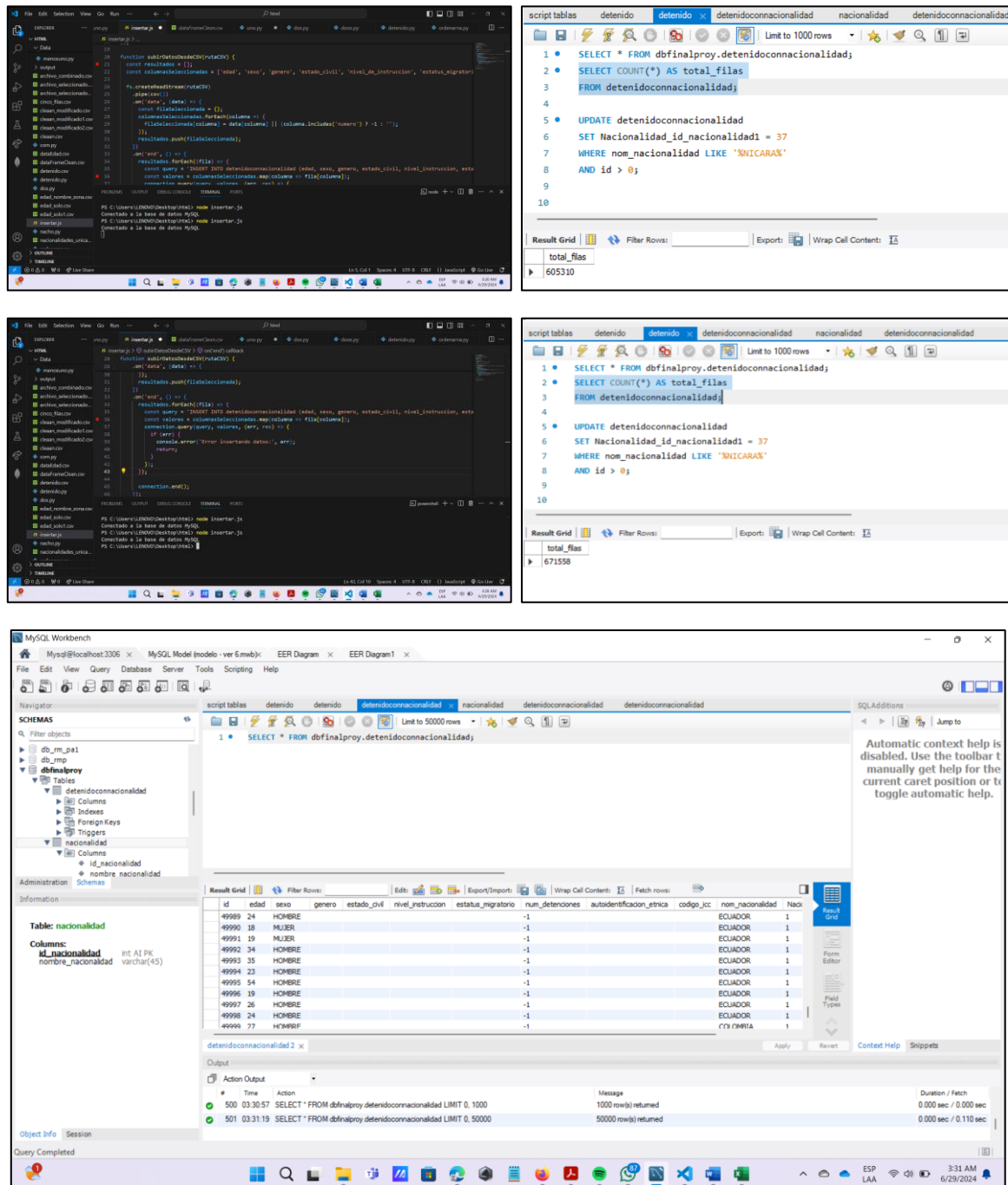
Tabla “Detención”:



Demás tablas:



Capturas de Pantalla



Código

Véase en [GitHub](#).