

# Análisis de Datos del Conflicto Armado en Colombia: Patrones, Tendencias y Correlaciones

## Integrantes del Equipo

- Santiago Rendón Castrillón
- José Daniel Orrego Ruiz
- Kelly Dayanna Hernández
- Osfary Rojas Vargas

## 1. Introducción

### 1.1 Descripción

El conflicto armado en Colombia ha generado múltiples hechos victimizantes que han afectado a diferentes poblaciones a lo largo del tiempo. Este proyecto tiene como objetivo analizar un conjunto de datos sobre el conflicto armado en el país, identificando patrones, tendencias y correlaciones entre los diferentes tipos de hechos victimizantes registrados. A través de un análisis exploratorio, se busca extraer información clave que ayude a comprender mejor la magnitud y evolución del conflicto.

Dado que inicialmente solo contamos con los datos sin una hipótesis definida, el análisis comenzará con una exploración de las variables disponibles para identificar posibles líneas de investigación. Se aplicarán técnicas de limpieza, visualización y modelado de datos para extraer insights relevantes.

## 2. Objetivos del Proyecto

1. **Explorar y comprender** la distribución de los hechos victimizantes en el conflicto armado.
2. **Analizar tendencias temporales y geográficas** para identificar patrones de violencia en el tiempo y el espacio.
3. **Detectar correlaciones** entre diferentes hechos victimizantes y su impacto en distintas poblaciones.
4. **Responder preguntas clave** mediante consultas exploratorias y visualizaciones interactivas.
5. **Evaluar la viabilidad de modelos predictivos** para anticipar tendencias en la ocurrencia de hechos victimizantes.

### 3. Preguntas de Análisis Exploratorio

- ¿Cuáles son los hechos victimizantes más frecuentes en el conflicto armado?
- ¿Cómo ha variado la cantidad de hechos victimizantes a lo largo del tiempo?
- ¿Cuáles son las regiones más afectadas por cada tipo de hecho victimizante?
- ¿Existe una correlación entre desplazamiento forzado y homicidios en ciertas regiones?
- ¿Se pueden identificar patrones estacionales en la ocurrencia de hechos victimizantes?
- ¿Qué municipios han experimentado el mayor número de desplazamientos forzados?
- ¿En qué regiones se han registrado más hechos de violencia contra mujeres o niños?
- ¿Podemos predecir la ocurrencia de ciertos tipos de violencia en función de datos históricos?
- ¿Qué factores están asociados con un mayor riesgo de desapariciones forzadas?
- ¿Cómo han evolucionado los ataques con minas antipersonal a lo largo del tiempo y en qué zonas son más recurrentes?

### 4. Metodología

El proyecto seguirá las siguientes etapas:

1. **Revisión y limpieza de datos** : Eliminación de valores nulos, estandarización y preparación del dataset.
2. **Análisis exploratorio** : Visualización de datos, generación de estadísticas descriptivas y formulación de hipótesis preliminares.
3. **Aplicación de técnicas de modelado de datos** : Identificación de patrones y correlaciones entre variables.
4. **Desarrollo de visualizaciones interactivas** : Creación de dashboards en Power BI para facilitar la interpretación de los resultados.
5. **Evaluación de modelos predictivos** : Aplicación de algoritmos de aprendizaje automático para identificar tendencias y posibles predicciones sobre hechos victimizantes.

### 5. Entregables

- **Diagrama entidad-relación (ER) y modelo relacional**
- **Script de creación y población de la base de datos en SQL**
- **Consultas SQL para responder preguntas clave**
- **Análisis de datos y visualizaciones en notebooks de Python**
- **Dashboard interactivo en Power BI**

Este análisis permitirá comprender mejor la dinámica del conflicto armado en Colombia y generar información útil para la toma de decisiones en el ámbito social y gubernamental.

## 6. Datos del Estudio

### 6.1 Columnas del Dataset

El dataset utilizado en este estudio contiene las siguientes columnas:

- **FECHA\_CORTE:** Fecha de corte del reporte.
- **NOM\_RPT:** Nombre del reporte.
- **COD\_PAIS:** Código del país.
- **PAIS:** Nombre del país.
- **COD\_ESTADO\_DEPTO:** Código del departamento.
- **ESTADO\_DEPTO:** Nombre del departamento.
- **PARAM\_HECHO:** Código del tipo de hecho victimizante.
- **HECHO:** Descripción del hecho victimizante.
- **SEXO:** Género de la víctima.
- **ETNIA:** Grupo étnico de la víctima.
- **DISCAPACIDAD:** Tipo de discapacidad de la víctima.
- **CICLO\_VITAL:** Rango de edad de la víctima.
- **PER\_OCU:** Número de personas ocupadas afectadas.
- **PER\_DECLA:** Número de personas declarantes.
- **PER\_UBIC:** Número de personas ubicadas.
- **PER\_SA:** Número de personas sin asistencia.
- **EVENTOS:** Número de eventos reportados.

### 6.2 Fuentes de Datos

Los datos utilizados en este estudio han sido obtenidos de fuentes oficiales y de acceso abierto:

- **Unidad de Víctimas:** [Publicación de Datos Abiertos](#)
- **Plataforma de Datos Abiertos de Colombia:** [Cifras de Víctimas por Hechos Departamental](#)

## 7. Diseño del esquema

El diseño de la base de datos se ha estructurado bajo un modelo relacional que permite la gestión eficiente de la información relacionada con los hechos victimizantes del conflicto armado en Colombia. Se han definido las siguientes tablas principales:

### 1. Pais

- `id_pais` (PK) → Identificador único del país.
- `nombre_pais` → Nombre del país.

### 2. Departamento

- `id_departamento` (PK) → Identificador único del departamento.
- `nombre_departamento` → Nombre del departamento.
- `id_pais` (FK) → Relación con la tabla Pais, indicando a qué país pertenece el departamento.

### 3. HechoVictimizador

- `id_hecho` (PK) → Identificador único del hecho victimizador.
- `nombre_hecho` → Nombre del hecho victimizador.
- `codigo_hecho` → Código que clasifica el hecho victimizador según su tipología.

### 4. Persona

- `id_persona` (PK) → Identificador único de la víctima.
- `sexo` → Género de la víctima.
- `etnia` → Grupo étnico al que pertenece la víctima.
- `discapacidad` → Información sobre discapacidad, si aplica.
- `ciclo_vital` → Clasificación de la víctima según su rango de edad.

## 5. EventoVictimizante

- `id_evento` (PK) → Identificador único del evento victimizante.
- `fecha_corte` → Fecha en la que se registró el evento.
- `id_departamento` (FK) → Relación con la tabla `Departamento`, indicando la ubicación del evento.
- `id_hecho` (FK) → Relación con la tabla `HechoVictimizante`, especificando el tipo de hecho registrado.
- `numero_eventos` → Cantidad de eventos de este tipo reportados en la fecha de corte.
- `personas_ocupadas_afectadas` → Número de personas ocupadas que fueron afectadas por el evento.
- `personas_declarantes` → Número de personas que declararon haber sido víctimas del evento.
- `personas_ubicadas` → Número de personas ubicadas después del evento.
- `personas_sin_asistencia` → Número de personas que no recibieron asistencia.

## 6. VictimaEvento

- `id_victima_evento` (PK) → Identificador único del registro de una víctima en un evento.
- `id_persona` (FK) → Relación con la tabla `Persona`, indicando quién fue la víctima.
- `id_evento` (FK) → Relación con la tabla `EventoVictimizante`, indicando en qué evento fue afectada la víctima.