

# INFORME DETALLADO APLICATIVO DE STREAMLIT: ANÁLISIS DE DELITOS EN COLOMBIA

Universidad Santo Tomás  
Ricardo Vargas y Karen Suárez  
Análisis Exploratorio de Datos de Seguridad y Convivencia  
Consultoría e Investigación 2025-2

Fecha de Reporte: Noviembre 2025  
Periodo Analizado: 2018 - 2024

## 1. INTRODUCCIÓN

Este documento presenta un análisis detallado del aplicativo de Streamlit desarrollado por la Universidad Santo Tomás para el análisis de delitos en Colombia. El aplicativo, titulado "Análisis de Delitos en Colombia - Modelo Estrella 2018-2024", es una herramienta interactiva que permite explorar y visualizar datos de criminalidad a nivel nacional con granularidad municipal.

El aplicativo está diseñado bajo un modelo de arquitectura dimensional (Modelo Estrella) que facilita análisis OLAP (Online Analytical Processing) eficiente de los datos. Este enfoque permite a los usuarios ejecutar consultas rápidas, comprender patrones complejos y extraer conclusiones relevantes para la formulación de políticas públicas de seguridad.

## 2. CARACTERÍSTICAS GENERALES DEL APLICATIVO

### 2.1 Datos Fundamentales

- Total de Registros: 3,283,822
- Total de Columnas: 15
- Memoria Utilizada: 2006.04 MB
- Años de Cobertura: 7 (2018-2024)
- Tipos de Delito: 18 categorías
- Departamentos: 32
- Municipios: 1,017
- Tipos de Armas/Medios: 59 categorías
- Categorías de Género: 3 (Masculino, Femenino, No Reportado)

### 2.2 Objetivos del Análisis

Los objetivos principales del aplicativo son:

- Identificar patrones temporales en la delincuencia
- Analizar la distribución geográfica de delitos
- Caracterizar perfiles de víctimas
- Evaluar el uso de armas y medios empleados
- Apoyar la formulación de políticas públicas de seguridad

### 2.3 Estructura de Datos

- Período: 2018-2024
- Cobertura: Nacional (Colombia)
- Granularidad: Municipal
- Dimensiones: Temporal, Geográfica, Víctimas, Armas/Medios, Delitos
- Modelo: Estrella con tabla de hechos y múltiples tablas de dimensión

## 3. DESCRIPCIÓN DETALLADA DE SECCIONES Y FUNCIONALIDADES

### 3.1 Página Principal (Inicio)

La página principal del aplicativo muestra un resumen ejecutivo con información clave:

- Título: "ANÁLISIS DE DELITOS EN COLOMBIA"
- Subtítulo: "Modelo Estrella 2018-2024 | Universidad Santo Tomás"

Métricas Destacadas (Tarjetas de Información):

- Registros Totales: 3,283,822
- Memoria Utilizada: 2006.04 MB
- Años de Cobertura: 7
- Tipos de Delito: 18

Resumen Ejecutivo con dos secciones:

Objetivos del Análisis:

1. Identificar patrones temporales
2. Analizar distribución geográfica
3. Caracterizar perfiles de víctimas
4. Evaluar uso de armas
5. Apoyar políticas públicas

Estructura de Datos:

- Período: 2018-2024
- Cobertura: Nacional
- Granularidad: Municipal
- Dimensiones: Temporal, Geográfica, Víctimas, Armas, Delitos

- Modelo: Estrella

### 3.2 Información General de la Base de Datos

Esta sección proporciona un análisis exploratorio de la estructura y contenido de los datos.

Estadísticas Generales (Tarjetas):

- Total Registros: 3,283,822
- Total Columnas: 15
- Departamentos: 32
- Municipios: 1,017
- Tipos Delito: 18
- Años Cobertura: 7
- Tipos Armas: 59
- Categorías Género: 3
- Memoria: 2006.0 MB

Detalle de Columnas (Tabla Interactiva):

La tabla muestra información sobre cada columna del dataset:

1. MUNICIPIO (object): 1,017 valores únicos, 0% nulos
2. GENERO (object): 3 valores únicos, 12.77% nulos
3. AGRUPA\_EDAD\_PERSONA (object): 4 valores únicos, 13.07% nulos
4. CANTIDAD (int64): 246 valores únicos, 0% nulos
5. TIPO\_DELITO (object): 18 valores únicos, 0% nulos
6. FECHA\_HECHO (object): 2,192 valores únicos, 12.81% nulos
7. DEPARTAMENTO (object): 32 valores únicos, 0% nulos
8. ARMAS\_MEDIOS (object): 59 valores únicos, 6.34% nulos
9. cod\_mpio (float64): 1,100 valores únicos, 0% nulos
10. tipo\_municipio (object): 2 valores únicos, 0% nulos
11. longitud (object): 1,099 valores únicos, 0% nulos
12. latitud (object): 1,100 valores únicos, 0% nulos
13. AÑO (int64): 7 valores únicos, 0% nulos
14. DP (int64): 33 valores únicos, 0% nulos
15. POBLACION\_TOTAL (int64): 6,894 valores únicos, 0% nulos

### 3.3 Análisis Temporal de Delitos

Esta sección proporciona un análisis de los patrones temporales en la ocurrencia de delitos.

Filtros Interactivos Disponibles:

- Años: Selección múltiple de 2018-2024
- Departamentos: Selección múltiple de todos los departamentos

- Tipos de Delito: Filtrado por categoría
- Top N: Configuración del número de elementos a visualizar

Métricas Clave:

- Registros Filtrados: 105,674
- Año Más Delitos: 2019 (17,332 registros)
- Año Menos Delitos: 2021 (9,574 registros)
- Variación Total: +4.4%

Visualización Principal - Evolución de Delitos por Año:

Gráfico de línea que muestra:

- Pico en 2019: ~17,000 delitos
- Descenso significativo en 2020: ~9,500 delitos
- Recuperación gradual 2021-2024: ~17,000 delitos

Drétrico de Distribución Porcentual por Año:

Gráfico de dona (Pie chart) que muestra la proporción de delitos por año:

- 2019: 16.4%
- 2024: 15.9%
- 2023: 15.3%
- 2018: 15.2%
- 2022: 14.7%
- 2020: 13.4%
- 2021: 9.06%

### 3.4 Análisis de Tipos de Delito

Esta sección analiza la distribución y frecuencia de diferentes tipos de delito.

Filtros Interactivos:

- Años: Selección múltiple
- Departamentos: Selección múltiple
- Armas/Medios: Categorización de los medios utilizados
- Top N: Número de delitos a mostrar

Métricas Principales:

- Registros Filtrados: 1,805
- Tipos de Delito Identificados: 2
- Delito Más Común: LESIONES PERSONALES (99.9%)
- Casos del Más Común: 1,803

Top 10 Tipos de Delito Más Frecuentes:

1. LESIONES PERSONALES: 99.9% (1,803 casos)
2. HOMICIDIO INTENCIONAL: 0.1% (2 casos)

#### Análisis Comparativo:

- Delitos con Mayor Frecuencia
- Delitos con Menor Frecuencia
- Top 5 Delitos - Distribución (gráfico de dona)

#### 3.5 Análisis de Armas y Medios

Esta sección examina los tipos de armas y medios utilizados en los delitos.

##### Top 10 Armas/Medios Más Utilizados:

1. SIN EMPLEO DE ARMAS: ~45,000 casos (representación visual en azul)
2. ARMA DE FUEGO: ~5,000 casos (amarillo)
3. NO REPORTADO: ~3,500 casos (rojo)
4. CONTUNDENTES: ~2,500 casos (rojo)
5. ARMA BLANCA/CORTOPUNZANTE: ~1,500 casos (rojo)
6. LLAMADA TELEFÓNICA: ~800 casos (rojo)
7. DIRECTA: ~600 casos (rojo)
8. REDES SOCIALES: ~400 casos (rojo)
9. VEHÍCULO: ~150 casos (rojo)
10. MOTO: ~100 casos (rojo)

#### Análisis Comparativo:

- Gráfico de barras horizontal mostrando frecuencia de uso
- Mapa de calor de Armas/Medios
- Evolución temporal del Top 10 Armas/Medios

#### 3.6 Análisis Geográfico

Esta sección proporciona un análisis espacial de la distribución de delitos a nivel departamental.

##### Evolución Temporal: Top 10 Departamentos

Gráfico de línea multicolor mostrando tendencias por departamento:

- AMAZONAS: Fluctuaciones entre 5,000-12,000 casos
- ANTIOQUIA: Variaciones entre 8,000-12,000 casos
- ARAUCA: Fluctuaciones entre 3,000-7,000 casos
- ARCHIPIÉLAGO DE SAN ANDRÉS: ~5,000 casos
- ATLÁNTICO: Variaciones entre 3,000-5,000 casos
- BOLÍVAR: ~3,000 casos
- BOYACÁ: ~2,500 casos
- CALDAS: ~2,000 casos
- CAQUETA: ~1,500 casos
- CASANARE: ~1,500 casos

Departamento Más Afectado: CUNDINAMARCA

Métricas del Departamento con Más Delitos:

- Total Delitos: ~60,000 casos
- Variación: Aumento significativo entre 2020-2024

Municipios Críticos:

- BOGOTÁ, D.C.: Punto de mayor concentración
- Otros municipios prioritarios identificados en el mapa

Top 10 Delitos por Departamento (Gráfico de Barras):

- AMENAZAS: Aproximadamente 60,000 casos
- DELITOS SEXUALES: ~500 casos
- HOMICIDIO INTENCIONAL: ~23,500 casos
- EXTORSIÓN: ~19,500 casos

Composición Delictiva:

Mapas de calor mostrando que la composición porcentual de delitos varía significativamente entre departamentos, validando la necesidad de estrategias de seguridad localizadas.

Distribución Top 10 Departamentos:

Visualización mediante gráfico de barras apiladas

Mapa de Árbol: Top 10 Departamentos

Representación jerárquica de la distribución geográfica

### 3.7 Perfil de Víctimas

Esta sección analiza las características demográficas de las víctimas de delitos.

Filtros Interactivos Disponibles:

- Departamentos: Selección múltiple
- Años: Rango temporal 2018-2024
- Tipos de Delito: Categorización
- Género de Víctimas

Métricas Generales de Víctimas:

- Víctimas Total: 105,439
- Género Predominante: MASCULINO
- Grupo de Edad Predominante: ADULTOS
- Porcentaje Género Predominante: 53.3%

Distribución de Víctimas por Género:

- Masculino: 53.3% (~56,200 víctimas)

- Femenino: 46% (~48,500 víctimas)
- No Reportado: <1%

Gráficos De Visualización:

- Gráfico de dona mostrando distribución por género
- Gráfico de barras con conteo de víctimas por género
- Mapa de calor de víctimas por género

Distribución por Grupo de Edad:

- ADULTOS: Grupo más afectado (~70,000+ víctimas)
- MENORES: Grupo de edad con riesgo (cantidad variable)
- ADOLESCENTES: Grupo intermedio
- No Reportado: Casos sin información de edad

Gráficos De Visualización Por Edad:

- Gráfico de dispersión (scatter plot) mostrando distribución por grupo
- Gráfico de barras comparativo entre grupos de edad
- Mapa de calor de víctimas por grupo etario

### 3.8 Modelo Estrella - Arquitectura Dimensional

La estructura del modelo estrella utilizado en el aplicativo es fundamental para entender la capacidad analítica.

Arquitectura del Modelo Dimensional:

Este modelo permite análisis OLAP (Online Analytical Processing) eficiente de los datos de criminalidad en Colombia.

Tabla de Hechos Principal: fact\_delitos

Métricas en la tabla de hechos:

- CANTIDAD: Número de delitos (medida principal)
- latitud, longitud: Coordenadas geográficas
- FECHA\_HECHO: Fecha de ocurrencia del delito

Claves Foráneas (FK) que conectan a dimensiones:

- fecha\_key -> dim\_tiempo
- ubicacion\_key -> dim\_ubicacion
- victimas\_key -> dim\_victima
- arma\_key -> dim\_arma\_medio
- delito\_key -> dim\_tipo\_delito

Tablas de Dimensión:

## 1. dim\_tiempo (Dimensión Temporal)

Atributos:

- fecha\_key (PK)
- AÑO
- Casos de Uso: Análisis temporal, tendencias anuales

## 2. dim\_ubicacion (Dimensión Geográfica)

Atributos:

- ubicacion\_key (PK)
- DEPARTAMENTO
- MUNICIPIO
- CODIGO\_DANE
- Casos de Uso: Análisis geográfico, focos geográficos

## 3. dim\_victima (Dimensión de Víctimas)

Atributos:

- victimas\_key (PK)
- GÉNERO
- AGRUPA\_EDAD\_PERSONA
- ESCOLARIDAD (si aplica)
- Casos de Uso: Perfilamiento de víctimas, vulnerabilidad demográfica

## 4. dim\_arma\_medio (Dimensión de Armas y Medios)

Atributos:

- arma\_key (PK)
- ARMAS\_MEDIOS
- CATEGORIA\_ARMA
- Casos de Uso: Análisis de medios empleados, control de armas

## 5. dim\_tipo\_delito (Dimensión de Tipos de Delito)

Atributos:

- delito\_key (PK)
- TIPO\_DELITO
- CATEGORIA\_DELITO
- Casos de Uso: Análisis de concentración delictiva, prioridades de intervención

Ventajas del Modelo Estrella:

- Consultas rápidas: Estructura optimizada para queries OLAP
- Fácil de entender: Diseño intuitivo y escalable
- Optimizado para BI: Preparado para herramientas de Business Intelligence
- Flexible para análisis: Permite creámentos complejos de atributos
- Escalable: Puede crecer con nuevos años y datos

## 3.9 Hallazgos Principales y Conclusiones

El aplicativo resumen un conjunto de hallazgos clave derivados del análisis integral.

## 1. TENDENCIA TEMPORAL

Variación de Delitos:

Se observa una variación del -8.6% entre el año inicial y el final del período, indicando una tendencia general.

Picos Temporales:

El año con más delitos fue 2019 (17,332 registros), lo que requiere un análisis específico de los factores desencadenantes.

## 2. CONCENTRACIÓN DELICTIVA

Delito Predominante:

El "HURTO A PERSONAS" es el delito más frecuente en el período analizado, siendo el principal foco de atención.

Concentración de Frecuencia:

El Top 5 de delitos concentra aproximadamente el 69.8% del total de casos, confirmando que la intervención debe priorizar estas categorías.

## 3. FOCOS GEOGRÁFICOS

Departamento Más Afectado:

El CUNDINAMARCA lidera el número absoluto de delitos en la región, requiriendo intervención prioritaria.

Municipios Críticos:

El municipio de BOGOTÁ, D.C. es el punto de mayor concentración y requiere una intervención prioritaria por ser el punto de mayor concentración.

Composición Delictiva Variable:

Los mapas de calor muestran que la composición porcentual de delitos varía significativamente entre departamentos, lo que valida la necesidad de estrategias de seguridad localizadas.

## 4. USO DE MEDIOS

Arma/Medio Común:

El "SIN EMPLEO DE ARMAS" es el medio más utilizado, lo cual guía las acciones de control.

Patrones Regionales:

El mapa de calor (Armas vs. Departamentos) permite identificar departamentos donde el "ARMA DE FUEGO" tiene una penetración porcentual inusualmente alta.

## 5. VULNERABILIDAD DE VÍCTIMAS

Género Dominante:

Las víctimas son predominantemente de género MASCULINO.

Grupo de Edad:

El grupo de edad ADULTOS es el más afectado, lo cual informa la focalización de campañas de prevención.

## CONCLUSIONES GENERALES

Los hallazgos presentados en el análisis evidencian patrones claros que requieren atención diferenciada:

1. Priorización Estratégica: Los recursos deben concentrarse en los delitos y zonas de mayor incidencia.
2. Enfoque Territorial: Las estrategias deben adaptarse a las características específicas de cada departamento.
3. Prevención Focalizada: Las campañas preventivas deben dirigirse a los grupos poblacionales más vulnerables.
4. Control de Medios: Las políticas de control de armas deben considerar los patrones regionales identificados.
5. Monitoreo Continuo: Es fundamental mantener sistemas de seguimiento para detectar cambios en las tendencias.

## 4. FUNCIONALIDADES TÉCNICAS DEL APLICATIVO

### 4.1 Navegación

El aplicativo utiliza un menú de navegación lateral que permite acceder a 8 secciones principales:

1. Inicio - Página principal con resumen ejecutivo
2. Información General - Detalles de la base de datos
3. Análisis Temporal - Patrones por año
4. Tipos de Delito - Distribución de categorías delictivas

5. Armas y Medios - Evaluación de medios empleados
6. Análisis Geográfico - Distribución espacial
7. Perfil de Víctimas - Características demográficas
8. Modelo Estrella - Arquitectura y estructura de datos
9. Hallazgos Principales - Conclusiones clave

Cada sección es accesible mediante botones de radio interactivos.

#### 4.2 Componentes Interactivos

Filtros Dinámicos:

- Selectores de años con opciones múltiples
- Selectores de departamentos
- Categorización de delitos
- Filtros de armas/medios
- Control de cantidad de elementos mostrados

Marcadores (Badges/Cards):

- Tarjetas informativas con métricas clave
- Valores calculados en tiempo real
- Indicadores de tendencia

#### 4.3 Visualizaciones

Tipos de Gráficos Implementados:

1. Gráficos de Línea: Evolución temporal de delitos, tendencias por departamento
2. Gráficos de Barras: Comparativas de frecuencia, top-n de categorías
3. Gráficos de Dona/Pie: Distribución porcentual, composición por dimensión
4. Mapas de Calor (Heatmaps): Relaciones entre dos dimensiones (ej: Armas vs. Departamentos)
5. Mapas de Árbol (Treemaps): Representación jerárquica de distribución geográfica
6. Mapas Geográficos: Distribución espacial de delitos a nivel municipal
7. Gráficos de Dispersión: Análisis bivariado de víctimas

#### 4.4 Tablas Interactivas

Características:

- Tablas con información descriptiva

- Opciones de descarga (CSV)
- Búsqueda y filtrado integrado
- Paginación de resultados
- Vista previa de datos

#### 4.5 Estilo y Diseño

Características Visuales:

- Branding: Logo de la Universidad Santo Tomás
- Paleta de colores profesional (azul, rojo, amarillo, etc.)
- Tipografía clara y legible
- Espaciado consistente
- Componentes responsivos

#### 4.6 Funcionalidad de Exportación

- Descarga de datos en formato CSV
- Descarga de visualizaciones como imágenes
- Generación de reportes integrados en la plataforma

### 5. REQUISITOS TÉCNICOS Y ESPECIFICACIONES

#### 5.1 Stack Tecnológico

- Framework: Streamlit (Python)
- Base de Datos: Dataset con 3,283,822 registros
- Lenguaje: Python 3.x
- Librerías Principales:
  - \* pandas: Manipulación de datos
  - \* plotly: Visualizaciones interactivas
  - \* streamlit: Framework principal
  - \* numpy: Cálculos numéricos
  - \* geopandas/folium: Mapas geográficos (probable)

#### 5.2 Requisitos del Sistema

- Memoria: Mínimo 2GB RAM (aplicativo usa 2006.04 MB)
- Procesador: Moderno (Intel i5 o equivalente)
- Navegador Web: Moderno (Chrome, Firefox, Safari, Edge)
- Conexión: Internet para streaming de aplicativo

#### 5.3 Rendimiento

- Carga del Aplicativo: Rápida (~2-5 segundos)
- Actualización de Gráficos: Instantánea en aplicativos filtrados
- Manejo de Eventos: Responsivo
- Escalabilidad: Optimizada para los 3.2M de registros

## 6. FORTALEZAS Y OPORTUNIDADES

### 6.1 Fortalezas

1. Cobertura Completa: Datos nacionales con granularidad municipal
2. Período Extendido: 7 años de datos permiten identificar tendencias
3. Interactividad: Filtros dinámicos facilitan exploración
4. Visualizaciones Avanzadas: Múltiples tipos de gráficos
5. Modelo Bien Estructurado: Arquitectura dimensional optimizada
6. Accesibilidad: Interfaz web fácil de usar
7. Documentación: Explicación clara de objetivos y estructura

### 6.2 Oportunidades de Mejora

1. Previsión: Agregar modelos de machine learning para predicción
2. Comparativas Regionales: Benchmarking entre municipios
3. Dashboard Ejecutivo: Resumen más compacto para toma de decisiones rápida
4. Exportación Avanzada: Generación de reportes PDF automatizados
5. Métricas Calculadas: Métricas per cápita, tasas de criminalidad
6. Alertas: Notificaciones de cambios significativos en tendencias
7. Comentarios: Capacidad de añadir notas contextuales
8. Versionamiento: Control de cambios en los datos

## 7. RECOMENDACIONES PARA LA IMPLEMENTACIÓN

### 7.1 Para Instituciones Públicas

1. Adopción Oficial: Implementar como herramienta oficial para análisis de seguridad
2. Actualización Periódica: Mantener datos actualizados mensualmente
3. Capacitación: Entrenar a analistas en el uso de la plataforma
4. Integración: Conectar con sistemas de información existentes

### 7.2 Para Investigadores y Académicos

1. Publicación: Resultados en revistas de investigación
2. Colaboración: Trabajar con otras universidades
3. Validación: Verificar hallazgos con expertos en criminología

#### 4. Extensión: Agregar nuevas dimensiones de análisis

##### 7.3 Para Ciudadanía

1. Transparencia: Publicar datos abiertos
2. Acceso Público: Permitir consultas limitadas sin autenticación
3. Educación: Usar para comunicación de seguridad pública

## 8. CONCLUSIONES FINALES

El aplicativo de Streamlit para el análisis de delitos en Colombia representa una herramienta sofisticada y bien diseñada para la exploración de datos de criminalidad. Su arquitectura basada en el modelo estrella, la variedad de visualizaciones interactivas y la capacidad de filtrado dinámico lo convierten en un instrumento valioso para:

- Tomadores de decisiones en seguridad pública
- Investigadores y académicos
- Planificadores territoriales
- Gestores de políticas públicas

Los hallazgos derivados del análisis proporciona evidencia sólida para la formulación de intervenciones estratégicas, enfoque territorial y prevención focalizada.

Con las mejoras sugeridas, el aplicativo podría expandir su impacto aún más, proporcionando predicciones, alertas y análisis más profundos.

La Universidad Santo Tomás ha desarrollado una herramienta que establece un estándar de excelencia en análisis de datos de seguridad en Colombia.