

Accidentes de Tránsito ocurridos en el Municipio de Bucaramanga

Marcela Duque Duque
✉ cheladuque650@gmail.com



**UNIVERSIDAD DE
SAN BUENAVENTURA
COLOMBIA**

Universidad de San Buenaventura Colombia

Facultad de Ingenierías
Ingeniería en sistemas cibernéticos
Medellín

Accidentes de Tránsito ocurridos en el Municipio de Bucaramanga

FUENTE DE DATOS

El conjunto de datos proviene de un archivo tabular (posiblemente CSV) con registros detallados sobre accidentes de tránsito, incluyendo variables como la fecha, la gravedad del accidente, el tipo de vehículos involucrados, la ubicación (barrio y comuna), la hora del incidente, la entidad que atendió y otras variables categóricas y cuantitativas.

Descripción del conjunto de datos: Este dataset incluye información de accidentes de tránsito en una ciudad durante el año 2012. Cada registro representa un accidente con atributos como:

- Fecha y hora del incidente
- Tipo de día (diurno/nocturno)
- Gravedad del accidente (con heridos, solo daños)
- Vehículos involucrados (automóvil, moto, bicicleta, etc.)
- Ubicación geográfica (barrio, comuna)
- Entidad que atendió el caso
- Propietario del vehículo (empresa o particular)

Contiene variables categóricas, numéricas y temporales, lo que permite un análisis multidimensional. La base de datos fue extraída en ([3. Accidentes de Transito ocurridos en el Municipio de Bucaramanga | Datos Abiertos Colombia](#))

Justificación del conjunto de datos: Este conjunto de datos es relevante por varias razones:

1. Impacto social: Los accidentes de tránsito afectan la salud pública, infraestructura y seguridad vial.
2. Análisis urbano: Permite identificar zonas de mayor riesgo en la ciudad.
3. Toma de decisiones: Puede ser usado por autoridades para mejorar políticas de movilidad y prevención.
4. Diversidad de variables: Incluye múltiples dimensiones (tiempo, ubicación, actores), lo cual enriquece el análisis estadístico.

Plan inicial de adquisición:

- Fuente: Datos obtenidos directamente de registros históricos de la Dirección de Tránsito Municipal o una institución pública abierta.
- Formato original: CSV o Excel, con delimitadores y codificaciones corregibles.
- Herramientas: Se utilizará PostgreSQL para almacenar y limpiar el dataset, y Python (pandas) para procesamientos adicionales.

PREGUNTAS DE ANÁLISIS CLAVE

- ¿En qué comunas o barrios ocurren más accidentes con heridos?
- ¿Cuál es la hora del día con mayor frecuencia de accidentes?
- ¿Qué tipo de vehículo se ve involucrado con más frecuencia?
- ¿Hay una relación entre el tipo de propietario del vehículo y la gravedad del accidente?
- ¿Existen patrones diferentes entre accidentes diurnos y nocturnos?

Hipótesis Iniciales:

- La mayoría de los accidentes con heridos involucran motocicletas.
- Los accidentes nocturnos tienen mayor gravedad.
- Los accidentes son más frecuentes entre semana que los fines de semana.

PLAN DE ANÁLISIS PRELIMINAR

Limpieza y preprocesamiento:

- Corregir errores de codificación (como "AÑO" → "AÑO").
- Transformar campos de fecha y hora a formato datetime.
- Convertir variables categóricas a formato estándar (ej. "Solo daños" → "Solo daños").
- Validar y completar valores nulos si es necesario.

Análisis exploratorio:

- Distribución de accidentes por comuna, barrio y gravedad.
- Análisis temporal por hora/día/mes.
- Conteo de accidentes por tipo de vehículo.
- Matriz de correlación entre variables cuantitativas.
- Visualización de tendencias y agrupamientos.

Técnicas estadísticas:

- Regresión logística (gravedad ~ tipo de vehículo, hora, día, etc.).
- Agrupamientos (k-means) de barrios con comportamiento similar.
- Árboles de decisión para identificar variables que predicen gravedad.

PROUESTA DE VISUALIZACIÓN Y DASHBOARD

Funcionalidades interactivas:

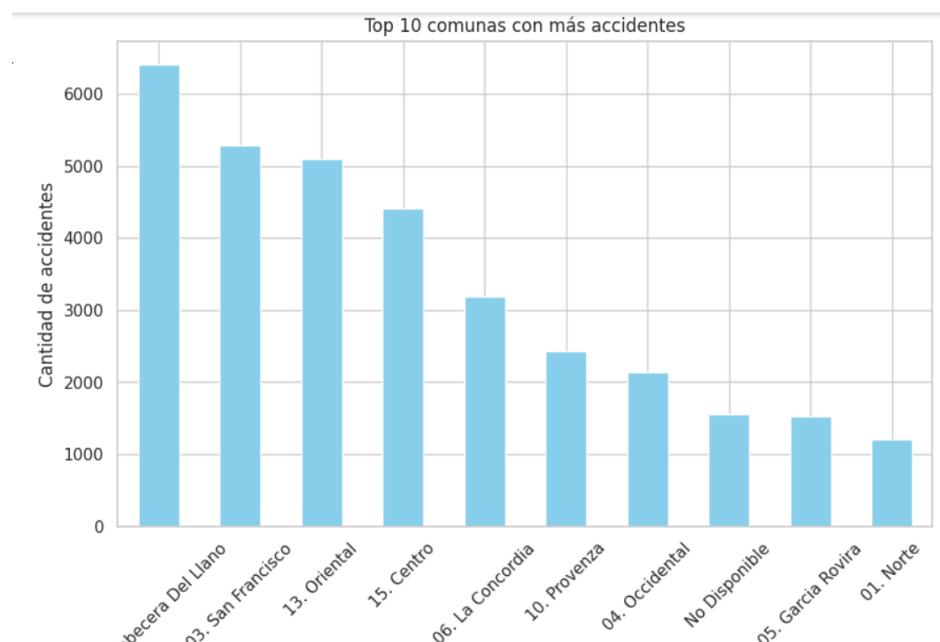
1. Filtros por:
 - Rango de fechas
 - Tipo de gravedad
 - Comuna / barrio
 - Tipo de vehículo
2. Selección múltiple de variables para análisis cruzado.
3. Tooltips informativos con datos por punto en el gráfico.
4. Mapa interactivo con zonas de mayor incidencia.
5. Gráficos de línea para evolución temporal y heatmaps por hora/día.

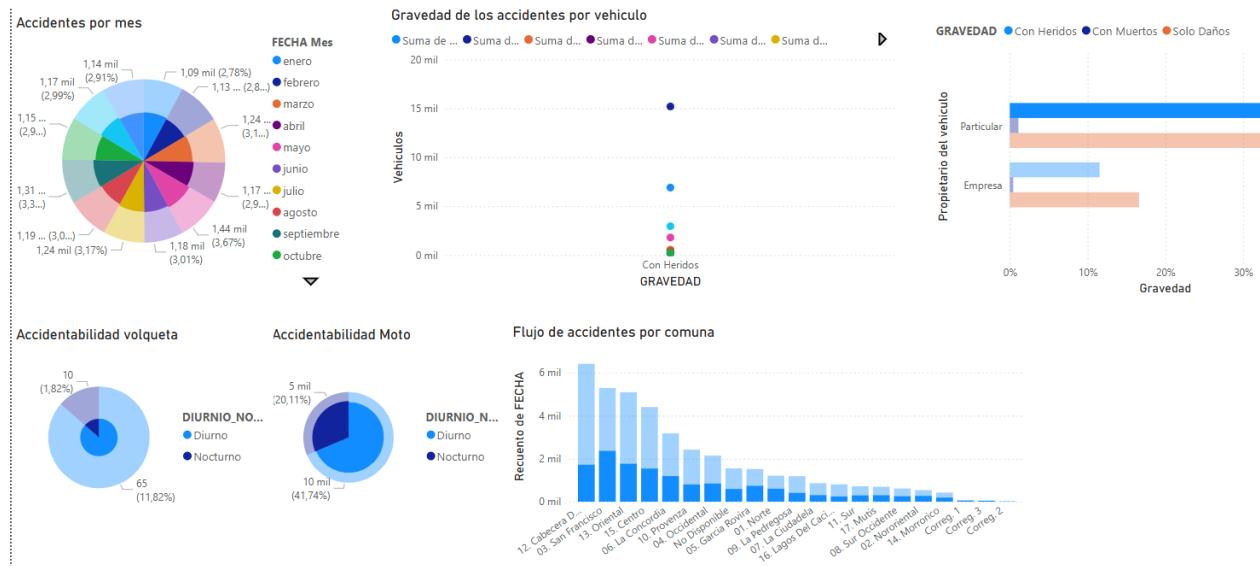
Herramientas / Librerías:

- Base de datos: PostgreSQL
- Limpieza / análisis: Python (pandas, numpy)
- Visualización: Plotly, Seaborn, Matplotlib
- Dashboard: Dash (Python) o Power BI / Tableau

HALLAZGOS CLAVE

A continuación, se expondrán los hallazgos clave del análisis exploratorio y el dashboard que ayudaron con la resolución de las preguntas de análisis y la hipótesis inicial:

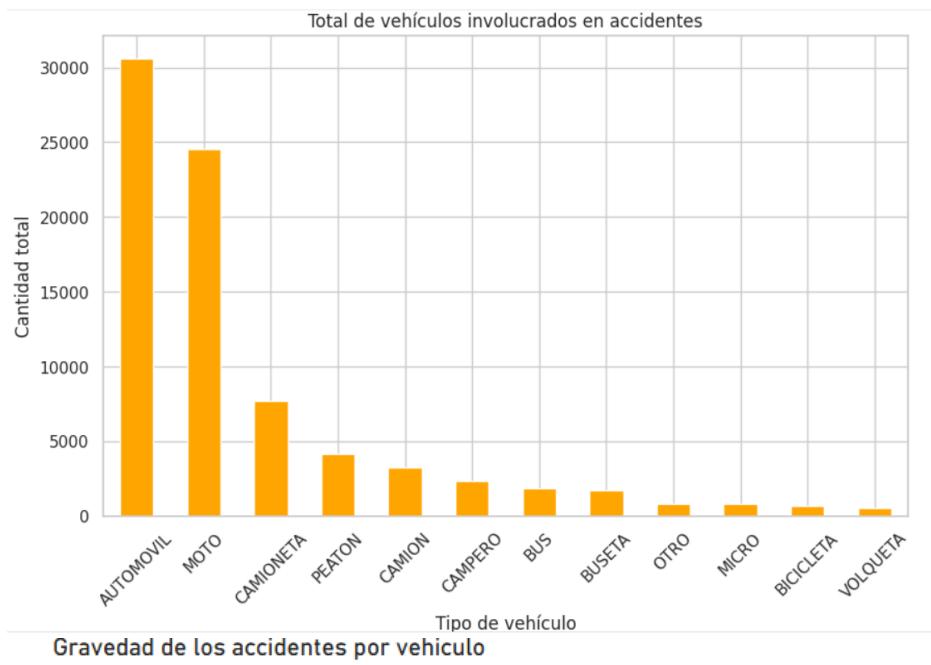




- En la primera grafica se puede observar que en las comunas Cabecera del Llano y San Francisco, es en las que se presentan mas accidentes en la ciudad de Bucaramanga. Ahora en el dashboard gracias a la grafica que especifica la gravedad del accidente y en la grafica que representa el flujo de accidentes por comuna se puede observar que cuando se selecciona que muestre solo los resultados de accidentes donde hay heridos San francisco seria la comuna con mas numero de heridos en accidentes y le seguirían las comunas Cabecera del Llano y Oriental.

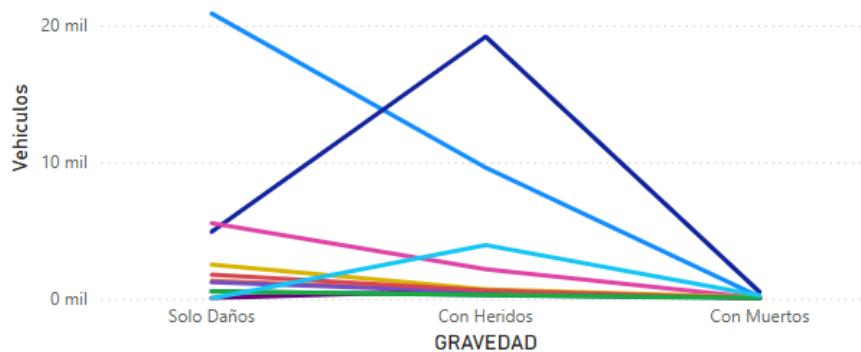


- En la gráfica del análisis exploratorio se puede observar como las horas en las que más accidentalidad por día hay son entre las 12 del medio día y la 1 de la tarde.



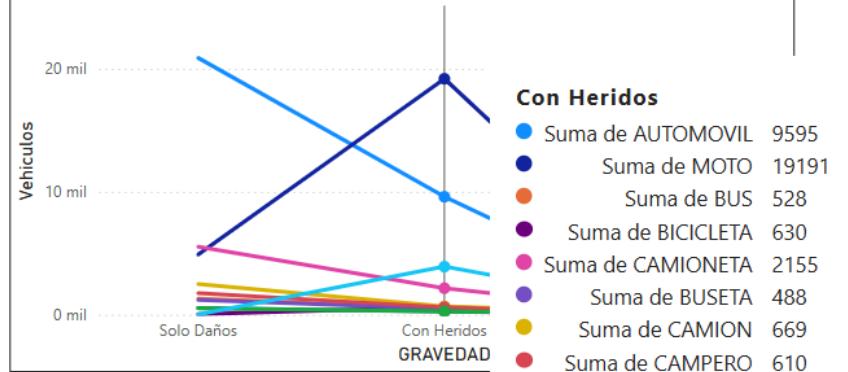
Gravedad de los accidentes por vehículo

● Suma de ... ● Suma de ... ● Suma d... ● Suma d... ● Suma d... ● Suma d... ● Suma d...



Gravedad de los accidentes por vehículo

● Suma de ... ● Suma de ... ● Suma d... ● Suma d... ● Suma d... ● Suma d... ● Suma d...



ad Moto

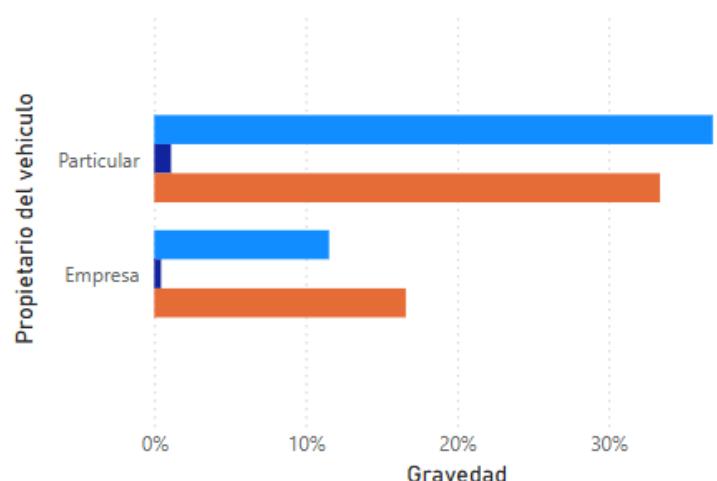
Flujo de accidentes por comur

Con Heridos

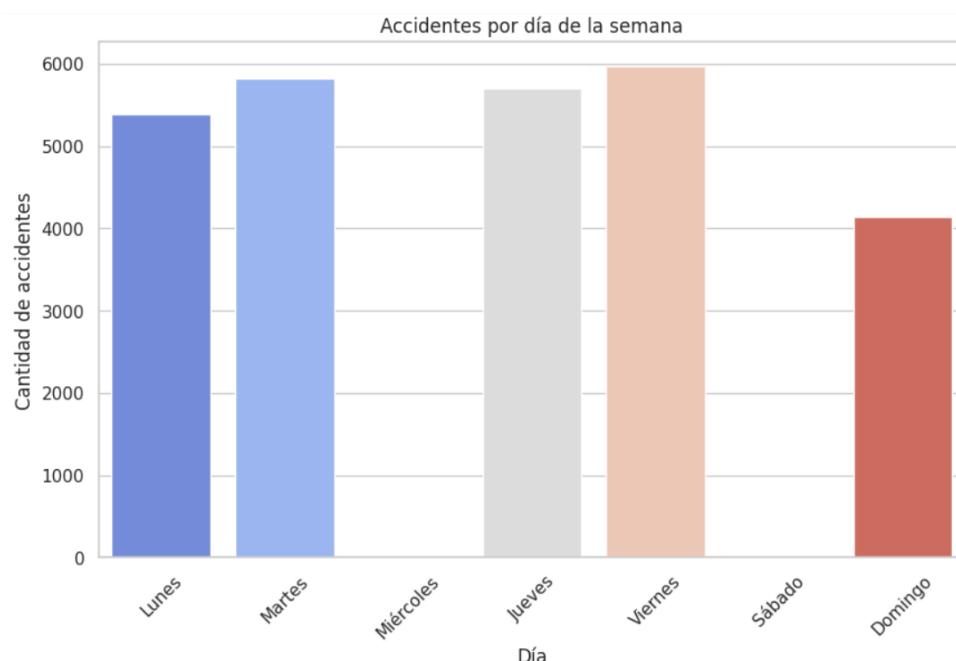
Suma de AUTOMOVIL	9595
Suma de MOTO	19191
Suma de BUS	528
Suma de BICICLETA	630
Suma de CAMIONETA	2155
Suma de BUSETA	488
Suma de CAMION	669
Suma de CAMPERO	610
Suma de MICRO	248
Suma de OTRO	270
Suma de PEATON	3909

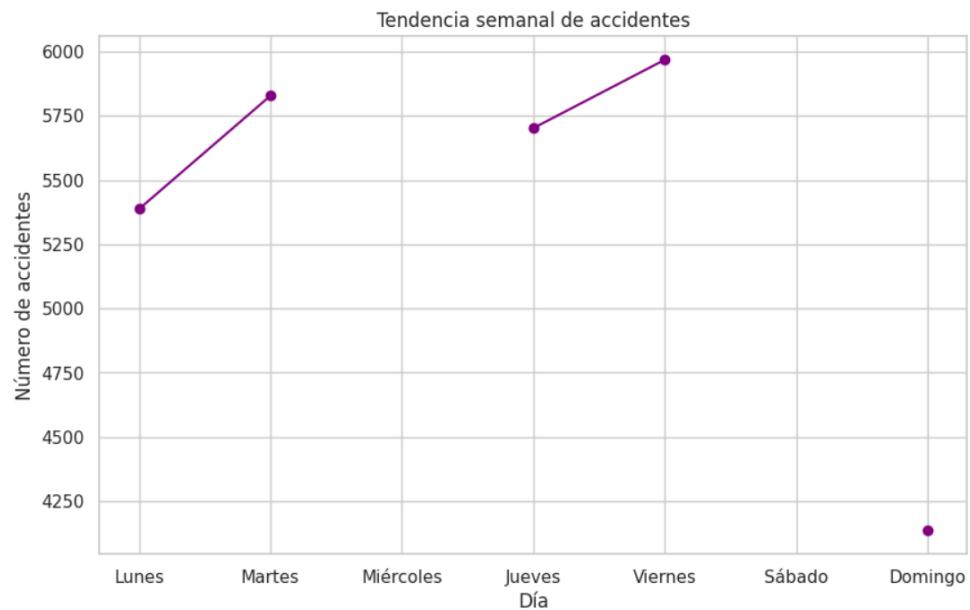
- La primera grafica de análisis exploratorio nos ayudo a identificar el tipo de vehículo que se accidentaba en la ciudad de Bucaramanga con mas frecuencia, dando como resultado que sería el automóvil. Y la segunda sacada del dashboard es importante para poder identificar cual es el vehículo que involucra mas heridos en sus accidentes y se puede apreciar que las motocicletas son el vehículo con más accidentes que presenta heridos.

GRAVEDAD ● Con Heridos ● Con Muertos ● Solo Daños



- Gracias a la gráfica sobre la gravedad del accidente según el tipo de propietario del vehículo (particular o de empresa), es posible comparar y analizar si existe una relación entre estas dos variables. Se observa que los vehículos particulares presentan un mayor porcentaje de accidentes con heridos (36,89%) en comparación con los accidentes que solo implican daños materiales (33,40%). Esto indica que los accidentes en los que se ven involucrados vehículos particulares tienden a generar más heridos que únicamente daños materiales en la ciudad de Bucaramanga. Por otro lado, en el caso de los vehículos de empresa, se evidencia que el porcentaje de accidentes con solo daños materiales (16,62%) es mayor que el de accidentes con heridos (11,55%). Además, el porcentaje de accidentes con personas fallecidas también es más alto en los vehículos particulares (1,11%) que en los de empresa (0,45%). Esto permite concluir que sí existe una relación entre el tipo de propietario del vehículo y la gravedad del accidente. Los vehículos particulares están más asociados a accidentes con consecuencias más graves, como heridos y fallecidos, mientras que los vehículos de empresa tienden a estar involucrados en accidentes de menor gravedad, principalmente con daños materiales.





- En las dos graficas anteriores del análisis exploratorio se puede apreciar que la accidentalidad es mas frecuente en semana que los fines de semana. El Dashboard del Proyecto Final ha sido diseñado para analizar y visualizar datos relacionados con la accidentalidad vial en la ciudad de Bucaramanga. Su objetivo principal es facilitar la exploración interactiva de las variables más relevantes en los siniestros viales, permitiendo identificar patrones, relaciones y posibles focos de atención.

DESCRIPCIÓN DEL DASHBOARD

Funcionalidades principales

- Filtros interactivos: El panel cuenta con filtros por año, tipo de vehículo, tipo de accidente, gravedad del siniestro, día de la semana, hora del día, entre otros. Estos filtros permiten al usuario segmentar la información según sus intereses o necesidades específicas.
- Gráficas dinámicas: Las visualizaciones se actualizan automáticamente al interactuar con cualquier filtro o gráfico del dashboard. Esto permite observar cómo cambian las tendencias y proporciones al modificar las condiciones de análisis.
- Segmentación por tipo de propietario del vehículo: Una funcionalidad destacada es la comparación entre vehículos particulares y vehículos de empresa, permitiendo observar diferencias en la gravedad de los accidentes según el tipo de propietario.

Cómo usar el dashboard

- Interactúa con los gráficos: Al hacer clic sobre una categoría dentro de una gráfica (por ejemplo, "Con Heridos" en un gráfico de barras), el resto de los gráficos se ajustarán para mostrar solo la información relacionada con esa selección.
- Analiza relaciones y patrones: Las gráficas permiten comparar visualmente variables como la gravedad del accidente, el tipo de vehículo y la localización, lo cual facilita el análisis de correlaciones y tendencias.

Utilidad de las visualizaciones interactivas

- Las visualizaciones interactivas son clave para una exploración profunda de los datos, ya que:
- Facilitan la identificación de patrones ocultos.
- Permiten realizar análisis comparativos de forma ágil.
- Brindan autonomía al usuario para personalizar su análisis sin necesidad de conocimientos técnicos avanzados.
- Mejoran la comprensión de los datos mediante elementos visuales intuitivos.

CONCLUSIONES

- Mayor accidentalidad en zonas céntricas y de alto flujo.
- Motocicletas y vehículos particulares son los más involucrados.
- Los accidentes diurnos son más frecuentes, pero menos graves que los nocturnos.

CONSIDERACIONES ÉTICAS

- Privacidad: No se manejan datos personales identificables, lo cual protege la privacidad.
- Sesgos: Pueden existir sesgos en el registro de accidentes (por ejemplo, subregistro de accidentes menores o sesgos por comuna).
- Limitaciones: El dataset solo cubre un año y puede no representar tendencias actuales. No incluye información meteorológica, estado de las vías, o consumo de sustancias.