

**Diseño de un dashboard para el análisis de los datos relacionados con los accidentes
ocurridos en la ciudad de Barranquilla**

José Luis Ordoñez Cueto

Liz Vanessa Fernandez Ebratt

Asesor

María Alejandra Varona Taborda

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI
Especialización en Ciencias de Datos y Analítica

2025

Resumen

Este proyecto tiene como objetivo principal identificar patrones y tendencias en los accidentes de tráfico ocurridos en Barranquilla, a partir del análisis de datos históricos de un período determinado. Se busca comprender las principales causas, características y factores de riesgo asociados con estos incidentes, mediante el uso de técnicas avanzadas de analítica de datos. A través de este análisis, se pretende establecer una base sólida para el diseño de estrategias preventivas y políticas públicas que mejoren la seguridad vial en la ciudad. Los hallazgos permitirán a las autoridades locales optimizar los recursos destinados a la prevención, reducir la frecuencia de los accidentes y minimizar los riesgos relacionados con el tráfico. Esta investigación se enfoca en el uso de herramientas de Big Data y Ciencia de Datos para generar conclusiones prácticas y significativas que contribuyan a la creación de un entorno más seguro y eficiente para la movilidad urbana.

Palabras claves: Análisis de datos, dashboard interactivo, python, MySQL, visualización de datos.

Abstract

This project aims to identify patterns and trends in traffic accidents that occurred in Barranquilla, based on the analysis of historical data over a specific period. It seeks to understand the main causes, characteristics, and risk factors associated with these incidents through the use of advanced data analytics techniques. Through this analysis, it is intended to establish a solid foundation for the design of preventive strategies and public policies that improve road safety in the city. The findings will allow local authorities to optimize resources allocated to prevention, reduce the frequency of accidents, and minimize traffic-related risks. This research focuses on the use of Big Data and Data Science tools to generate practical and meaningful conclusions that contribute to creating a safer and more efficient environment for urban mobility.

Keywords: Data analysis, interactive dashboard, python, MySQL, data visualization

Tabla de Contenido

Introducción	8
Descripción del Problema	9
Planteamiento del Problema.....	9
Sistematización del Problema	10
Justificación	11
Objetivos	13
Objetivo General	13
Objetivos Específicos.....	13
Marco de Referencia	14
Marco Contextual.....	14
Marco Teórico.....	14
Marco Conceptual	15
Marco Teórico.....	17
Metodología	18
Método	18
Fases del Proyecto	18
Tipo de Estudio	21
Recolección de Datos	21
Resultados	23
Información de los Datos en Bruto	23
Creación de la Base de Datos.....	23
Limpieza de Datos.....	26

Análisis Exploratorio de los Datos.....	28
Análisis de Clustering (K-Means).....	39
Mapa de Burbujas	40
Dashboard Interactivo	41
Conclusión	45
Recomendaciones	47
Referencias Bibliográficas	49

Lista de Figuras

Figura 1 <i>Base de Datos de Accidentes de Tránsito del Portal Datos Abiertos</i>	23
Figura 2 <i>Base de Datos en Supabase</i>	24
Figura 3 <i>Código para la Conexión a la Base de Datos</i>	25
Figura 5 <i>Datos Atípicos en la Variable Edad</i>	26
Figura 6 <i>Gráfico de Violín de la Variable Edad</i>	27
Figura 7 <i>Gráfico de Barras Número de Accidentes por Condición de Víctima</i>	28
Figura 8 <i>Gráfico de Barras Número de Accidentes por Gravedad</i>	28
Figura 9 <i>Gráfico de Barras Número de Accidentes por Clase</i>	29
Figura 10 <i>Gráfico de Barras Número de Accidentes por Sexo de la Víctima</i>	30
Figura 11 <i>Gráfico de Violín de la Variable Edad</i>	30
Figura 12 <i>Histograma de la Variable Edad de la Víctima</i>	31
Figura 13 <i>Gráfico de Barras Cantidad de Accidentes por Mes</i>	32
Figura 14 <i>Gráfico de Barras Cantidad de Accidentes por Día de la Semana</i>	33
Figura 15 <i>Gráfico de Barras Agrupadas por Tipo y Gravedad del Accidente</i>	33
Figura 16 <i>Diagrama de Caja y Bigotes de las Variables Edad y Gravedad</i>	34
Figura 17 <i>Matriz de Correlación</i>	35
Figura 18 <i>Matriz de Correlación de Spearman</i>	36
Figura 19 <i>Análisis Multivariado 1</i>	36
Figura 20 <i>Análisis Multivariado 2</i>	37
Figura 21 <i>Análisis Multivariado 3</i>	38
Figura 22 <i>Gráfico de K-Medias</i>	39
Figura 23 <i>Mapa de Burbujas</i>	40

Figura 24 <i>Dashboard 1</i>	41
Figura 25 <i>Dashboard 2</i>	41
Figura 26 <i>Dashboard 3</i>	42
Figura 27 <i>Dashboard 4</i>	42

Introducción

El análisis de datos se ha convertido en una herramienta esencial en el ámbito profesional, facilitando la extracción de conocimiento útil a partir de grandes volúmenes de información. Este proceso implica la recopilación, limpieza, transformación, análisis y visualización de datos para identificar patrones, tendencias y relaciones significativas. En el contexto de la seguridad vial, el análisis de datos es crucial para entender la dinámica de los accidentes de tráfico y tomar decisiones informadas que mejoren la seguridad de los ciudadanos. Según Provost y Fawcett (2013), “el análisis de datos permite a las organizaciones tomar decisiones basadas en evidencia, mejorando así la eficiencia y efectividad de sus operaciones” (p. 23), asimismo, Steinbach y Kumar (2019, p. 45) subrayan “que la visualización de datos es fundamental para comunicar los hallazgos de manera clara y efectiva, facilitando la comprensión de información compleja”.

Este proyecto se centra en el diseño de un dashboard interactivo para el análisis de datos relacionados con los accidentes ocurridos en la ciudad de Barranquilla. Se implementará un sistema de análisis avanzado que no solo permitirá identificar patrones y tendencias en los datos de accidentes, sino que también proporcionará a las autoridades locales las herramientas necesarias para mejorar la seguridad vial y reducir la incidencia de accidentes.

El trabajo de grado se ajusta perfectamente a la línea de investigación de "Gestión de Sistemas". El análisis, diseño e implementación de un sistema de información, con el objetivo de apoyar el desarrollo productivo, tecnológico y social de Barranquilla. Además, incluye los elementos clave de la gestión de sistemas como la planificación, dirección, control, evaluación y realimentación.

Descripción del Problema

La ciudad de Barranquilla enfrenta una alta incidencia de accidentes de tráfico que impacta negativamente en la seguridad vial y la calidad de vida de sus habitantes. Los datos relacionados con estos accidentes suelen estar fragmentados y no estructurados, dificultando su análisis. Esto genera una falta de conocimiento profundo sobre los patrones, factores contribuyentes y dinámicas que afectan la accidentalidad vial, lo que impide tomar decisiones informadas para prevenir estos eventos. Además, las entidades responsables carecen de herramientas avanzadas para analizar y visualizar datos, lo que limita su capacidad de respuesta y planificación estratégica.

Planteamiento del Problema

El problema identificado es la falta de herramientas avanzadas para el análisis y visualización de datos de accidentes de tráfico en Barranquilla. Actualmente, los datos se recopilan en formatos no estructurados, lo que dificulta su análisis y limita el acceso rápido a información relevante. Además, la ausencia de técnicas analíticas avanzadas y dashboards interactivos impide la identificación de patrones y factores contribuyentes, afectando negativamente la capacidad de las entidades locales para tomar decisiones informadas y mejorar la seguridad vial.

Sistematización del Problema

¿Qué herramientas y técnicas son necesarias para estructurar y analizar los datos de accidentalidad vial en Barranquilla?

¿Cómo pueden identificarse patrones y factores contribuyentes a los accidentes de tráfico mediante técnicas avanzadas de análisis de datos?

¿Qué características debe incluir un dashboard interactivo para facilitar la visualización y toma de decisiones basada en datos?

¿Qué impacto tendrá la implementación de un sistema integrado de análisis y visualización de datos en la seguridad vial de Barranquilla?

Justificación

El proyecto "Diseño de un dashboard para el análisis de los datos relacionados a los accidentes ocurridos en la ciudad de Barranquilla" se fundamenta en la necesidad crítica de proporcionar a las entidades locales herramientas avanzadas para el análisis y visualización de datos de accidentes. Este enfoque no solo mejora la eficiencia en la toma de decisiones, sino que también contribuye significativamente a la seguridad vial y la planificación urbana. Según McKinsey & Company (2016, p. 56), "las ciudades que utilizan análisis de datos para la gestión del tráfico pueden reducir los accidentes hasta en un 30%, mejorando así la seguridad y eficiencia del transporte urbano".

El análisis de datos es fundamental para comprender y mitigar los riesgos asociados con los accidentes de tráfico. Estudios han demostrado que la utilización de técnicas analíticas avanzadas puede identificar patrones de accidentes y factores contribuyentes, lo que facilita la implementación de medidas preventivas (Björkman & Sundqvist, 2020, p. 42). Además, Chen y Lin (2014, p 88) destacan que las herramientas de visualización de datos, como dashboards interactivos, son esenciales para comunicar de manera efectiva los resultados del análisis a las partes interesadas, permitiendo una respuesta rápida y adecuada a los problemas identificados.

La implementación de este proyecto tiene un impacto directo en la seguridad y bienestar de los ciudadanos de Barranquilla. Al proporcionar herramientas avanzadas para el análisis y visualización de datos de accidentes, se pueden identificar y abordar de manera más efectiva los factores que contribuyen a los accidentes de tráfico. Esto, a su vez, puede resultar en la reducción de la tasa de accidentes, mejorar la seguridad vial y optimizar la asignación de recursos para infraestructuras y medidas preventivas. Asimismo, la capacidad de realizar un análisis continuo y

actualizado permite a las autoridades adaptarse rápidamente a nuevas tendencias y desafíos, asegurando una gestión proactiva y eficiente de la seguridad vial en la ciudad.

Objetivos

Objetivo General

Diseñar un dashboard para el análisis de los datos relacionados a los accidentes ocurridos en la ciudad de Barranquilla.

Objetivos Específicos

Evaluar los datos históricos de accidentes viales en Barranquilla para identificar patrones significativos, factores recurrentes y tendencias clave, incluyendo variables como ubicación geográfica, horarios de mayor incidencia, condiciones del tráfico y tipos de accidentes.

Explorar y aplicar técnicas avanzadas de analítica de datos con el fin de descubrir correlaciones y factores de riesgo determinantes en la ocurrencia de accidentes viales, permitiendo una comprensión profunda de las causas subyacentes y su relación con el entorno urbano.

Proponer recomendaciones y estrategias preventivas basadas en los resultados del análisis de datos, enfocadas en mejorar la infraestructura vial, fortalecer la seguridad del tráfico y optimizar las políticas públicas, con el objetivo de reducir significativamente la tasa de accidentes viales en Barranquilla y promover una gestión vial eficiente y proactiva.

Marco de Referencia

Marco Contextual

Barranquilla, ubicada en la costa Caribe colombiana, es una ciudad con alta densidad poblacional y un flujo constante de tráfico vehicular debido a su importancia comercial y estratégica. Las principales rutas, como la Avenida Circunvalar, Vía 40 y Avenida Murillo, presentan desafíos significativos en términos de movilidad y seguridad vial. Estos problemas se ven agravados por deficiencias en señalización, iluminación y condiciones viales. Según la Secretaría de Tránsito y Transporte de Barranquilla (2023), la ciudad registra una alta incidencia de accidentes viales, especialmente en intersecciones peligrosas y áreas urbanas densamente pobladas.

Marco Teórico

Colombia presenta altos índices de accidentalidad vial, siendo una de las principales causas de muerte en el país. La Organización Mundial de la Salud (OMS) ubica a Colombia en la lista de países con mayor número de muertes por accidentes de tránsito. Barranquilla, como una ciudad con alto tráfico vehicular, no escapa a la realidad nacional. Evidencian la problemática de la accidentalidad en la ciudad.

Estudios previos: Se han desarrollado investigaciones locales y nacionales que analizan las causas de la accidentalidad en Barranquilla y otras ciudades, identificando factores como la velocidad, el estado de las vías, la falta de señalización y el consumo de alcohol. Estudios y estadísticas de la Secretaría de Tránsito y Transporte de Barranquilla (2023) respaldan estos hallazgos.

Análisis de la accidentalidad vial: El análisis de la accidentabilidad vial requiere de diferentes metodologías para comprender la complejidad del fenómeno. Una de las primeras etapas consiste en el análisis descriptivo, que busca obtener información básica sobre los accidentes, incluyendo la frecuencia, la gravedad y la ubicación. A partir de este análisis se puede avanzar hacia un análisis exploratorio, que busca identificar patrones y tendencias en los datos, como la identificación de zonas con mayor concentración de accidentes. Para profundizar en la comprensión de los eventos, se utiliza el análisis causal, que se enfoca en investigar las causas que desencadenan los accidentes, determinando si la causa principal es la velocidad, el alcohol, la distracción o alguna otra variable. Finalmente, el análisis predictivo, que utiliza técnicas de aprendizaje automático, permite predecir la probabilidad de accidentes en diferentes zonas, tipos de vehículos o momentos del día. Esta metodología combina la información histórica con análisis estadísticos para identificar tendencias y predecir futuros eventos (Vargas Montero, 2022; Shmueli, Patel, & Bruce, 2010).

Marco Conceptual

El análisis de accidentalidad en Barranquilla integra la información previa sobre seguridad vial, sistemas de información y el contexto específico de Barranquilla.

La seguridad vial se define como la protección de la vida y la salud de las personas en la vía pública, con el objetivo de reducir el riesgo de accidentes y sus consecuencias. Este enfoque integral prioriza la prevención y la gestión del riesgo para garantizar una movilidad segura (Ministerio de Transporte de Colombia, 2022).

Causas de la accidentalidad vial: Los accidentes de tránsito son el resultado de una interacción compleja entre factores humanos, vehículos y el entorno. Las causas principales se pueden clasificar en:

Factores humanos: Errores del conductor (exceso de velocidad, conducción bajo influencia de alcohol o drogas, distracciones, fatiga, falta de atención, conducción agresiva), falta de experiencia, errores de percepción, etc.

Factores del vehículo: Mantenimiento deficiente, fallas mecánicas, sistemas de seguridad inadecuados (frenos, llantas, luces), etc.

Factores ambientales: Clima adverso (lluvia, niebla, viento), visibilidad reducida, estado de las vías (pavimentos en mal estado, baches, falta de señalización), etc (Redalyc.org).

Consecuencias de la accidentalidad vial: La accidentabilidad vial tiene consecuencias devastadoras en la sociedad, impactando la vida de las personas y la economía de manera significativa. Cada año, miles de personas pierden la vida en accidentes de tránsito a nivel mundial. Además de la mortalidad, las lesiones que se producen en estos accidentes generan discapacidad física, psicológica y social, lo que afecta la calidad de vida de las personas. Estos eventos conllevan altos costos para el sistema de salud, desviando recursos que podrían utilizarse en otros ámbitos. Asimismo, la pérdida de productividad debido a la incapacidad laboral y la disminución de la productividad genera un impacto negativo en la economía. Las consecuencias de la accidentabilidad vial no se limitan a lo económico, también provocan dolor, sufrimiento, trauma y pérdida para las familias, comunidades y la sociedad en general (Organización Mundial de la Salud, 2022).

Prevención y gestión del riesgo: Los accidentes de tránsito son el resultado de una compleja interacción entre factores humanos, vehículos y el entorno. Las causas principales de estos eventos se pueden clasificar en tres categorías: factores humanos, factores del vehículo y factores ambientales. Los factores humanos incluyen errores del conductor, como el exceso de velocidad, la conducción bajo la influencia del alcohol o las drogas, las distracciones, la fatiga, la

falta de atención, y la conducción agresiva. También se incluyen la falta de experiencia y los errores de percepción. Los factores del vehículo se relacionan con el mantenimiento deficiente, las fallas mecánicas y los sistemas de seguridad inadecuados, como los frenos, las llantas y las luces. Finalmente, los factores ambientales abarcan el clima adverso, como la lluvia, la niebla y el viento, la visibilidad reducida y el estado de las vías, incluyendo pavimentos en mal estado, baches y la falta de señalización (ANSV, 2023).

Marco Teórico

Analítica de datos: Uso de técnicas como minería de datos y análisis predictivo para identificar factores de riesgo (Fayyad et al., 1996).

Gestión de la seguridad vial: Modelos que priorizan la prevención y mitigación de accidentes a través del análisis de factores humanos, vehiculares y ambientales (Organización Mundial de la Salud, 2022).

Visualización de datos: Herramientas como dashboards que combinan datos geoespaciales y gráficos dinámicos para apoyar decisiones estratégicas (Cairo, 2016; Few, 2012).

Metodología

Método

La metodología de este proyecto sigue un enfoque mixto, combinando métodos cuantitativos y cualitativos para analizar los accidentes de tráfico en Barranquilla, con el propósito de identificar patrones, factores de riesgo y tendencias que guíen la formulación de estrategias de prevención. Este enfoque asegura que el análisis sea integral, basado en datos relevantes y orientado a la generación de conocimiento práctico para la mejora de la seguridad vial, se compone de las siguientes fases:

Fases del Proyecto

Recopilación y preparación de datos: La recopilación y preparación de datos es un paso crucial para el análisis de la accidentabilidad vial en Barranquilla. Se inicia con la identificación de las fuentes de datos relevantes, incluyendo la Secretaría de Tránsito y Transporte, la Policía Nacional, hospitales y otras entidades que registren información sobre accidentes de tránsito. Posteriormente, se solicita acceso a las fuentes de datos identificadas y se descarga la información relevante, como la fecha del accidente, la ubicación, el tipo de accidente, los factores contribuyentes, la gravedad del accidente, las víctimas, etc. Para asegurar la calidad, consistencia y precisión de los datos, se realiza una limpieza y preprocesamiento, que incluye la corrección de errores de registro, la eliminación de duplicados y la complementación de valores faltantes. Finalmente, se establece un formato uniforme para las variables, asegurando la coherencia entre diferentes fuentes, lo que facilita la integración y el análisis de los datos.

1. **Análisis exploratorio de datos:** Una vez recopilados y preparados los datos, se realiza un análisis exploratorio para obtener una comprensión inicial de la información. Este análisis involucra el cálculo de estadísticas básicas para cada variable, como la media, la

desviación estándar, la frecuencia, etc., lo que proporciona una visión general de la distribución de los datos. A continuación, se utilizan herramientas de visualización de datos para identificar patrones y tendencias en la información, como la frecuencia de accidentes por tipo, ubicación o factor de riesgo. Finalmente, se emplea el análisis espacial, a través de herramientas de geovisualización, para identificar zonas de riesgo en la ciudad de Barranquilla. Este proceso permite visualizar la concentración de accidentes en diferentes áreas y determinar si existen patrones espaciales que sugieran la influencia de factores geográficos en la accidentabilidad.

2. fase: Diseño de la dashboard: El diseño de una dashboard para el análisis de la accidentabilidad vial en Barranquilla requiere una planificación cuidadosa que garantice su utilidad y eficacia. Se inicia definiendo los objetivos específicos de la dashboard y las necesidades de los usuarios, como la Secretaría de Tránsito, la Policía Nacional o investigadores. Se seleccionan indicadores relevantes, como la frecuencia de accidentes, la gravedad de las lesiones, la ubicación de los eventos y los factores de riesgo, para representar la información de manera efectiva. Se elige una combinación de visualizaciones, como mapas de calor, gráficos de barras, gráficos de líneas, tablas, etc., para presentar la información de forma clara y comprensible. Finalmente, se realizan pruebas y validaciones exhaustivas para asegurar que la dashboard funcione correctamente y que la información se presente de manera clara, efectiva y útil para los usuarios.

3. Implementación y evaluación: Una vez desarrollada la dashboard, se implementa en un servidor web o en una plataforma de análisis de datos, lo que la hace accesible para los usuarios de forma eficiente. Para asegurar su utilidad e impacto, se realiza una evaluación continua, midiendo el uso de la herramienta, la satisfacción de los usuarios y su contribución a la mejora de la seguridad vial en Barranquilla. Esta evaluación permite identificar áreas de mejora

y optimizar la dashboard para que se convierta en una herramienta esencial para la toma de decisiones informadas y la implementación de estrategias de prevención de accidentes.

Roles y responsabilidades: Un proyecto de este tipo requiere un equipo con diferentes roles: el líder del proyecto, responsable de la coordinación general, la gestión de los recursos y la entrega final del proyecto. El analista de datos se encarga de la recopilación, limpieza, preprocesamiento y análisis de los datos necesarios. Un diseñador de dashboards se encarga de diseñar y construir la dashboard, utilizando herramientas de visualización de datos, mientras que el desarrollador web es responsable de la implementación de la dashboard en un servidor web o plataforma de análisis de datos. La colaboración entre estos roles es fundamental para el éxito del proyecto.

Recursos: Para llevar a cabo este proyecto, se requiere el acceso a bases de datos y sistemas de información que registren la accidentalidad en Barranquilla y software. El equipo humano también es crucial, incluyendo un líder del proyecto, un analista de datos, un diseñador de la dashboard y un desarrollador web. Para garantizar el éxito del proyecto, es indispensable establecer un cronograma realista que defina las fechas de inicio y fin de cada fase, asegurando la gestión eficiente del tiempo.

Gestión del proyecto: Para asegurar el éxito del proyecto, es fundamental mantener una comunicación efectiva entre los miembros del equipo, utilizando herramientas de colaboración y realizando reuniones periódicas. Además, se debe realizar un seguimiento del avance del proyecto, empleando herramientas de gestión de proyectos y reportes de progreso. Por último, es crucial garantizar la calidad de la información y la dashboard, realizando pruebas y revisiones periódicas para asegurar la precisión y la confiabilidad de los datos presentados.

Ética e impacto: Es fundamental que el uso de los datos y la dashboard sea ético y responsable, respetando la privacidad de las personas y evitando cualquier tipo de discriminación. La dashboard debe tener un impacto positivo en la seguridad vial de Barranquilla, ayudando a las autoridades a tomar decisiones más informadas y a desarrollar estrategias de prevención más efectivas para reducir la accidentalidad.

Documentación: Para finalizar el proyecto, se debe elaborar un informe final que describa el proyecto en detalle, incluyendo la metodología utilizada, los resultados obtenidos y las conclusiones alcanzadas.

Revisión y mejora: La dashboard debe ser una herramienta dinámica y adaptable. Se recomienda realizar revisiones periódicas para evaluar su utilidad y el impacto en la seguridad vial. Se debe actualizar la información y realizar mejoras en el diseño y la funcionalidad para asegurar su eficacia a largo plazo.

Tipo de Estudio

El proyecto se estructura de manera que cada etapa contribuya directamente al análisis profundo de la accidentabilidad vial en Barranquilla, desde la recopilación y preparación de datos hasta la implementación de estrategias basadas en evidencia. Se prioriza la calidad, la integridad y la consistencia de los datos, así como la claridad en la comunicación de los hallazgos para facilitar la toma de decisiones informadas.

Recolección de Datos

Se identifican fuentes confiables de información, como Datos Abiertos, para obtener datos relevantes sobre accidentes viales. Posteriormente, se realiza un proceso de limpieza y preprocesamiento que incluye la corrección de errores, la eliminación de duplicados y la

complementación de datos faltantes. Este paso garantiza que los datos estén listos para un análisis robusto.

Resultados

Información de los Datos en Bruto

Figura 1

Base de Datos de Accidentes de Tránsito del Portal Datos Abiertos

Fecha_Accidente	DIRECCION_ACCIDENTE	CONDICION_VICTIMA	GRAVEDAD_ACCIDENTE	CLASE_ACCIDENTE
2018 Jan 01 12:00:00 AM	CL 87 9H 24	Peaton	herido	Atropello
2018 Jan 01 12:00:00 AM	CLLE 119B CRA 11B	Peaton	herido	Atropello
2018 Jan 01 12:00:00 AM	CR 8 CL 41	Acompañante	herido	Choque
2018 Jan 01 12:00:00 AM	CR 8 CL 41	Conductor	herido	Choque
2018 Jan 01 12:00:00 AM	CR 8 CL 41	Motociclista	herido	Choque
2018 Jan 02 12:00:00 AM	CIRCUNVALAR CR 9G	Acompañante	herido	Choque
2018 Jan 02 12:00:00 AM	CIRCUNVALAR CR 9G	Motociclista	herido	Choque
2018 Jan 02 12:00:00 AM	CL 30 CR 30	Conductor	herido	Choque
2018 Jan 03 12:00:00 AM	CLLE 110 CRA 52	Motociclista	herido	Atropello
2018 Jan 03 12:00:00 AM	CLLE 110 CRA 52	Peaton	herido	Atropello

Nota. La imagen fue tomada de Datos Abiertos con corte al mes de enero de 2025

Los datos en bruto con los cuales se desarrolló el proyecto provienen del archivo "Accidentalidad en Barranquilla", disponible en la plataforma oficial Datos Abiertos de Colombia. Este conjunto de datos, compuesto por 15,746 registros, detalla accidentes de tránsito en la ciudad, incluyendo información sobre la fecha, ubicación, rol de las víctimas (como peatón, conductor o motociclista), gravedad del incidente, tipo de accidente, sexo, edad y número de afectados.

Creación de la Base de Datos

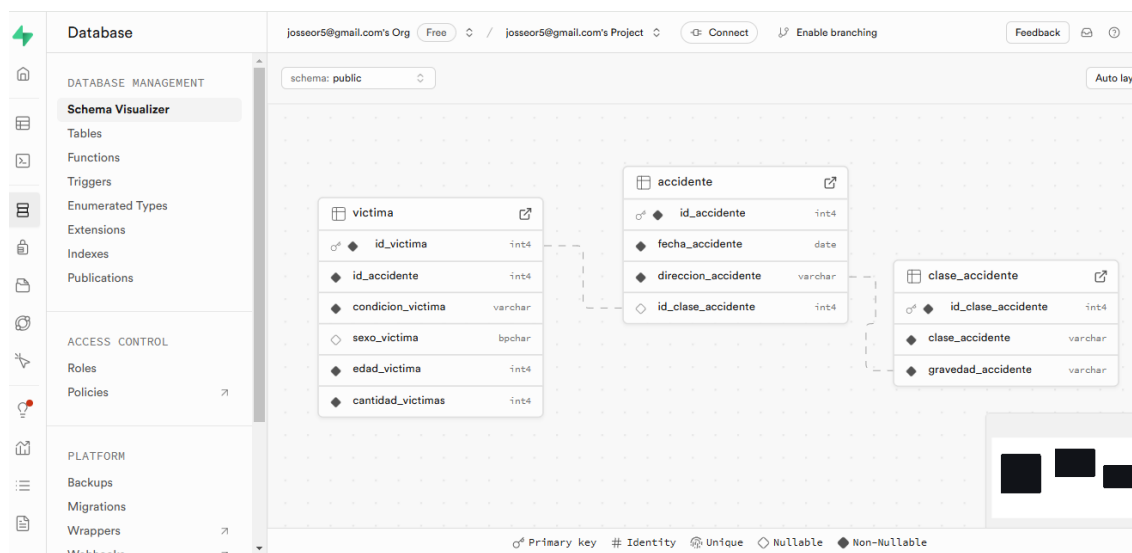
En este proyecto, se decidió crear una base de datos en lugar de trabajar directamente con archivos CSV para garantizar una gestión más eficiente y escalable de los datos. Las bases de datos relacionales permiten organizar la información de manera estructurada, establecer

relaciones claras entre las entidades (como accidentes, víctimas y tipos de accidentes) y mantener la integridad de los datos mediante restricciones como claves primarias y foráneas. Además, trabajar con bases de datos mejora el rendimiento en operaciones complejas, facilita el acceso concurrente y asegura que todos los datos estén centralizados, actualizados y protegidos contra inconsistencias.

Se eligió Supabase como la plataforma de base de datos debido a su simplicidad de configuración y capacidades avanzadas, como la generación automática de una API RESTful y un backend basado en SQL. Supabase ofrece un entorno administrado que permite gestionar datos sin necesidad de configurar servidores, además de características integradas como autenticación y control de acceso, ideales para proyectos colaborativos. También proporciona una interfaz gráfica para gestionar tablas, realizar consultas y configurar relaciones, acelerando el desarrollo.

Figura 2

Base de Datos en Supabase



Se crearon tres tablas principales para estructurar los datos de manera eficiente: `clase_accidente`, que contiene información sobre los tipos de accidentes y su gravedad con `id_clase_accidente` como clave primaria; `accidente`, que registra los detalles de cada accidente (fecha, dirección y tipo), vinculado a `clase_accidente` mediante una clave foránea; y `victima`, que almacena datos de las víctimas (condición, edad y cantidad) y se relaciona con la tabla `accidente` a través de otra clave foránea. Estas relaciones, definidas mediante claves primarias y foráneas, aseguran la integridad referencial y permiten realizar consultas complejas, como identificar víctimas asociadas a un tipo específico de accidente. Además, las tablas fueron normalizadas para evitar redundancias y garantizar una mejor organización de los datos.

Figura 3

Código para la Conexión a la Base de Datos

```
import psycopg2
import pandas as pd

# Datos de conexión
host = "db.anpdyfjelmahoberfv.supabase.co"
port = 5432
database = "postgres"
user = "postgres"
password = "Jose12345" # Reemplaza con tu contraseña

# Establecer la conexión
try:
    conn = psycopg2.connect(
        host=host,
        port=port,
        database=database,
        user=user,
        password=password
    )
    print("Conexión exitosa a la base de datos")

# Crear un cursor para interactuar con la base de datos
cursor = conn.cursor()

# Consulta para obtener datos de la tabla victima y las otras tablas relacionadas
query = """
SELECT
    v.id_victima,
    v.condicion_victima,
    v.sexo_victima,
    v.edad_victima,
    v.cantidad_victimas,
    a.id_accidente,
    a.fecha_accidente,
    a.direccion_accidente,
    c.id_clase_accidente,
    c.clase_accidente,
    c.gravedad_accidente
FROM victima v
LEFT JOIN accidente a ON v.id_accidente = a.id_accidente
LEFT JOIN clase_accidente c ON a.id_clase_accidente = c.id_clase_accidente;
"""

# Ejecutar la consulta
cursor.execute(query)

# Obtener los resultados
rows = cursor.fetchall()

# Crear un DataFrame de pandas con los resultados
df = pd.DataFrame(rows, columns=[
    "id_victima", "condicion_victima", "sexo_victima", "edad_victima", "cantidad_victimas",
    "id_accidente", "fecha_accidente", "direccion_accidente",
    "id_clase_accidente", "clase_accidente", "gravedad_accidente"
])
```

La conexión a la base de datos de Supabase se realizó utilizando la biblioteca `psycopg2`, que interactúa con bases de datos. Los parámetros necesarios incluyen el host (URL proporcionada por Supabase), el nombre de la base de datos, el usuario y la contraseña. La función `get_connection()` establece y devuelve la conexión activa. Para ejecutar consultas, la función `run_query()` utiliza `pd.read_sql_query()` de `pandas`, que toma la consulta SQL y la convierte automáticamente en un `DataFrame`, facilitando el análisis y la visualización de datos.

Limpieza de Datos

Figura 4

Datos Atípicos en la Variable Edad

```

Valores mayores a 100 años:
  fecha_accidente  direccion_accidente  condicion_victima \
291      2018-02-13      CALLE 17 CARRERA 8      Peaton
2646     2019-01-16              CR 8 CL 38B      Motociclista
5018     2020-02-10              CL 51B CR 2      Motociclista
5203     2020-03-07              CL 87 CR 50      Motociclista
5393     2020-04-30              CR 2G 51E 131     Motociclista
5661     2020-08-04              CL 42 43 131     Ciclista
6735     2021-03-12              CR 43B CL 9      Peaton
7150     2021-06-13              CL 88 5 SUR 110    Peaton
7202     2021-06-21              CL 45 CR 26      Motociclista
10969    2022-12-05    CARRERA 44 CON CALLE 30     Peaton
12174    2023-05-18    CALLE 119 CARRERA 31      Ciclista

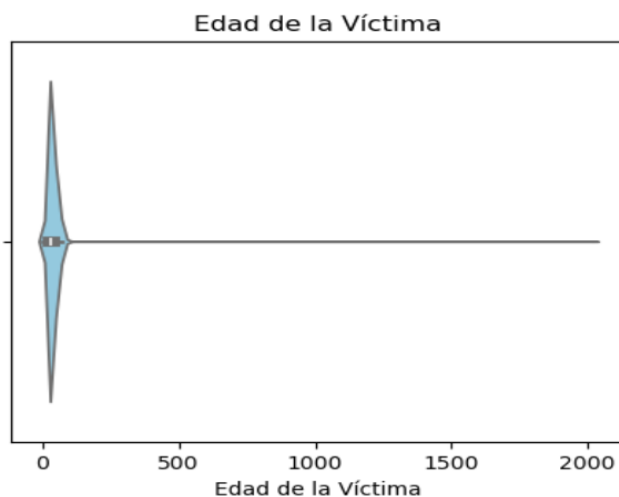
  gravedad_accidente  clase_accidente  sexo_victima  edad_victima \
291      herido      Atropello      M      2018
2646     herido      Choque      F      2019
5018     herido      Choque      M      2020
5203     herido      Choque      M      2021
5393     herido      Choque      F      2021
5661     herido      Choque      M      2021
6735     muerto      Otro      M      121
7150     herido      Atropello      M      122
7202     herido      Choque      M      2022
10969    muerto      Atropello      F      2023
12174    herido      Choque      F      2024

  cantidad_victimas
291      1
2646     1
5018     1
5203     1
5393     1
5661     1
6735     1
7150     1
7202     1
10969    1
12174    1

```

Figura 5

Gráfico de Violín de la Variable Edad.

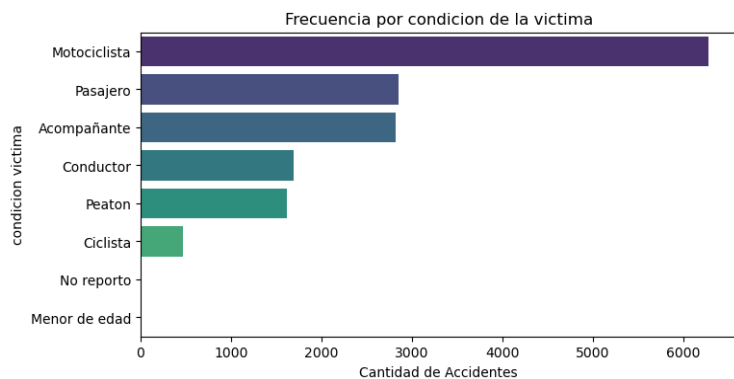


En el proceso de limpieza de datos, se identificaron y trataron valores nulos mediante la eliminación de filas incompletas y el relleno de datos faltantes con valores predeterminados. Además, se eliminaron columnas irrelevantes para el análisis utilizando drop, y se llevó a cabo la conversión de tipos de datos apropiados, asegurando que las columnas, como las de fechas y números, tuvieran formatos consistentes y adecuados para el análisis. También se eliminaron registros duplicados para evitar redundancias en los resultados. Adicionalmente, se identificaron valores inusuales en las edades, específicamente registros con edades mayores a 100 años, los cuales se decidió eliminar para mejorar la calidad y la coherencia del conjunto de datos. Estos pasos permitieron preparar un conjunto de datos limpio, consistente y listo para el análisis y visualización.

Análisis Exploratorio de los Datos

Figura 6

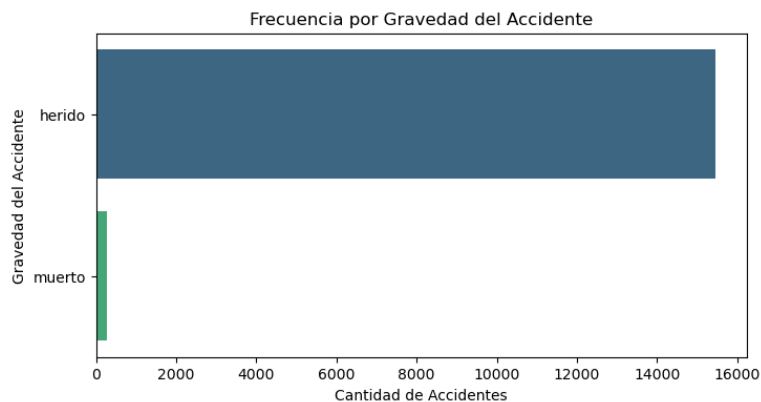
Gráfico de Barras Número de Accidentes por Condición de Víctima



La gráfica evidencia que los motociclistas son las principales víctimas de accidentes, seguidos por pasajeros, acompañantes, conductores y peatones, con ciclistas y menores de edad como poblaciones vulnerables. Esto sugiere la necesidad de implementar campañas de educación vial dirigidas especialmente a motociclistas, promoviendo el respeto por las normas de tránsito y el uso de cascos, así como concienciar a pasajeros y conductores sobre la seguridad.

Figura 7

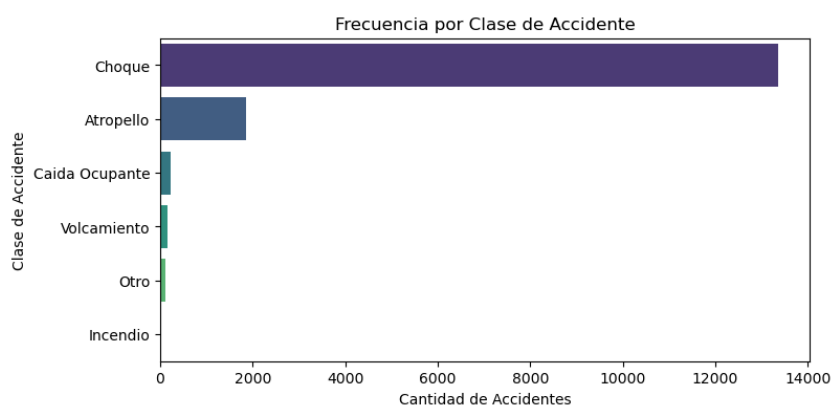
Gráfico de Barras Número de Accidentes por Gravedad



La mayoría de los accidentes en Barranquilla resultan en heridos (98.22%), con un pequeño porcentaje resultando en muertes (1.78%). Se destaca la necesidad de un sistema de atención integral y prevención de muertes. El análisis de datos avanzado es crucial para identificar patrones y factores de riesgo en la seguridad vial, proponiéndose un sistema unificado de datos y campañas de capacitación y sensibilización.

Figura 8

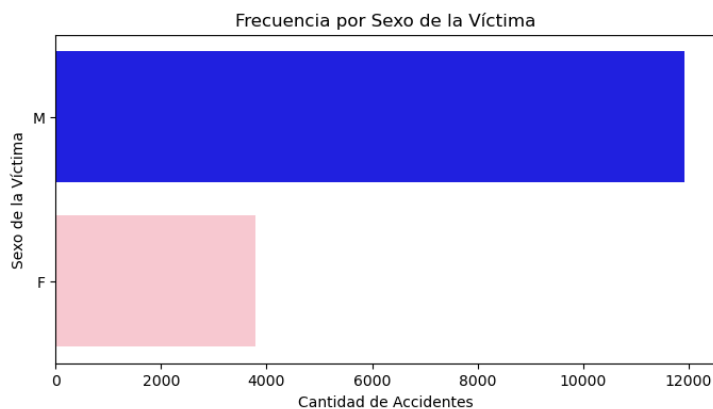
Gráfico de Barras Número de Accidentes por Clase



La gráfica muestra que los choques son, con gran diferencia, el tipo de accidente más frecuente, seguidos por los atropellos, mientras que otros tipos como caídas de ocupantes, volcamientos e incendios tienen menor incidencia. Esto sugiere que las entidades de tránsito deben priorizar acciones para reducir los choques mediante la instalación de sistemas de control de velocidad, mejoras en la señalización vial y campañas de concientización sobre el respeto a los límites de velocidad.

Figura 9

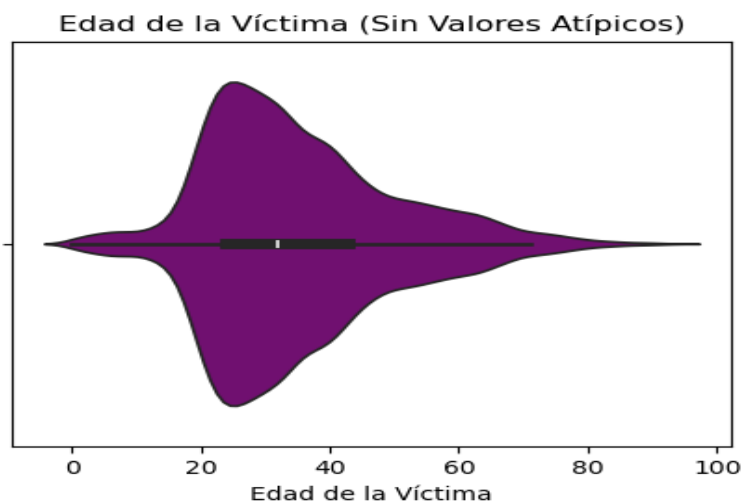
Gráfico de Barras Número de Accidentes por Sexo de la Víctima



La mayoría de las víctimas de accidentes en Barranquilla son hombres (75.68%), mientras que las mujeres representan un 24.05%. Un pequeño porcentaje de accidentes no tiene datos sobre el género de la víctima, representando el 0.27%. Esta disparidad sugiere que hay factores relacionados con el género que deben ser estudiados para entender por qué los hombres son más propensos a ciertos tipos de accidentes.

Figura 10

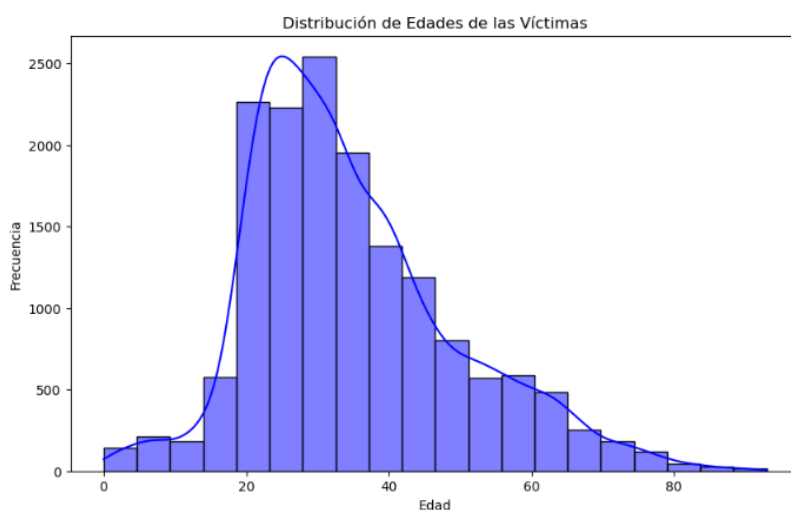
Gráfico de Violín de la Variable Edad



El análisis del histograma y el gráfico de violín sobre la distribución de edades de las víctimas, tras eliminar valores atípicos, muestra una concentración principal en edades jóvenes (cercanas a 0 años y entre 20-50 años). La mediana está cerca de 30 años, con una ligera asimetría hacia edades menores. La limpieza de datos permitió una representación más precisa, aunque los rangos amplios dificultan detalles específicos. Destaca la alta frecuencia en edades tempranas, sugiriendo un análisis más profundo en este grupo.

Figura 11

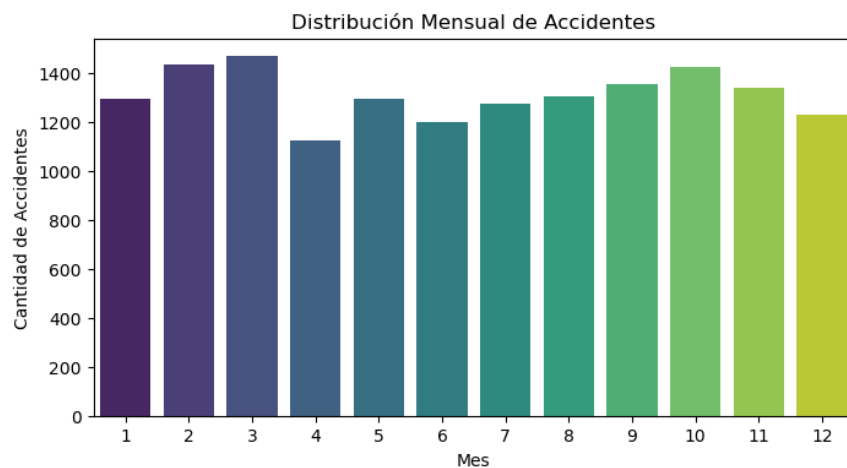
Histograma de la Variable Edad de la Víctima



Estos datos sugieren que las campañas de prevención y educación vial podrían enfocarse más en los jóvenes adultos, quienes parecen ser los más vulnerables a los accidentes de tráfico. Además, es importante considerar estrategias específicas para proteger a los grupos de edad más jóvenes y mayores, aunque su incidencia sea menor.

Figura 12

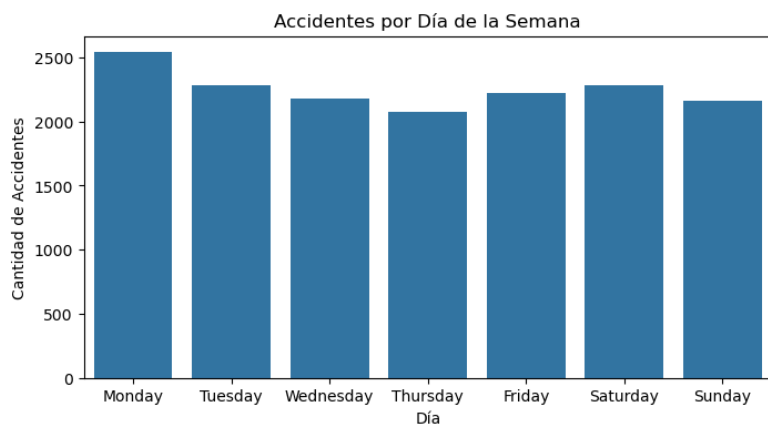
Gráfico de Barras Cantidad de Accidentes por Mes



La gráfica muestra que la cantidad de accidentes se mantiene relativamente constante durante el año, con ligeros picos en meses como febrero y marzo, posiblemente debido a mayor movilidad por vacaciones o eventos especiales, mientras que diciembre presenta una disminución, probablemente por campañas de seguridad vial efectivas. Se recomienda a las entidades de tránsito reforzar las campañas de prevención y los controles en meses críticos como febrero y marzo, investigar las causas específicas detrás de los picos de accidentes, y extender las estrategias exitosas de diciembre a otros meses del año. Además, es clave mantener un monitoreo constante para identificar patrones emergentes y ajustar las acciones preventivas de manera oportuna.

Figura 13

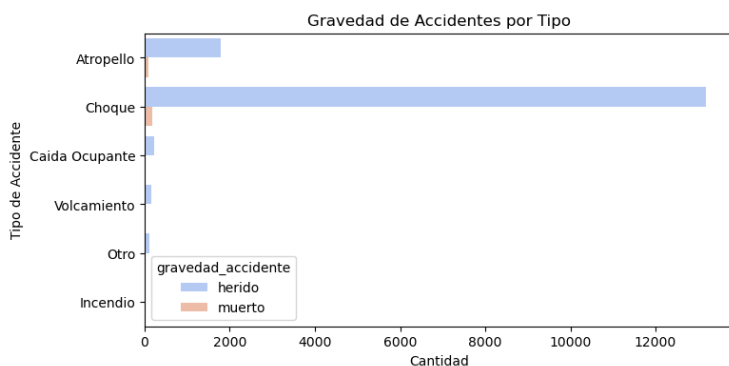
Gráfico de Barras Cantidad de Accidentes por Día de la Semana



La gráfica muestra que los lunes tienen la mayor cantidad de accidentes, probablemente por el inicio de la semana laboral, estrés y tráfico intenso, mientras que de martes a sábado los accidentes son estables y disminuyen los domingos por menor tráfico. Se recomienda implementar campañas y controles específicos para los lunes, monitorear patrones regulares durante la semana y garantizar la seguridad los domingos en zonas recreativas y vías rápidas. Además, promover horarios laborales flexibles ayudaría a reducir la congestión en días críticos.

Figura 14

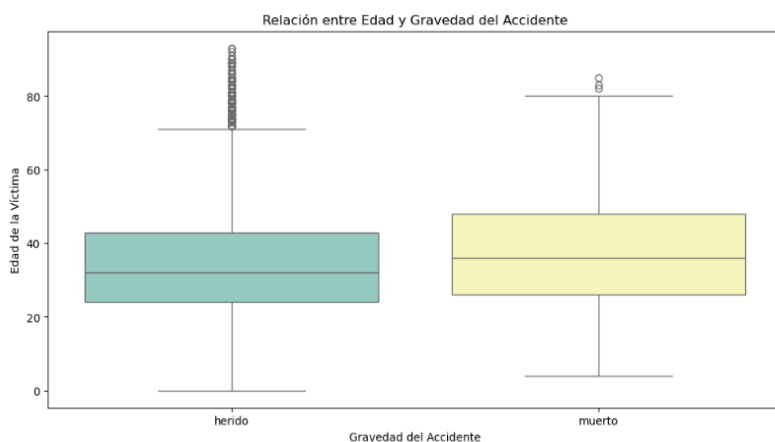
Gráfico de Barras Agrupadas por Tipo y Gravedad del Accidente



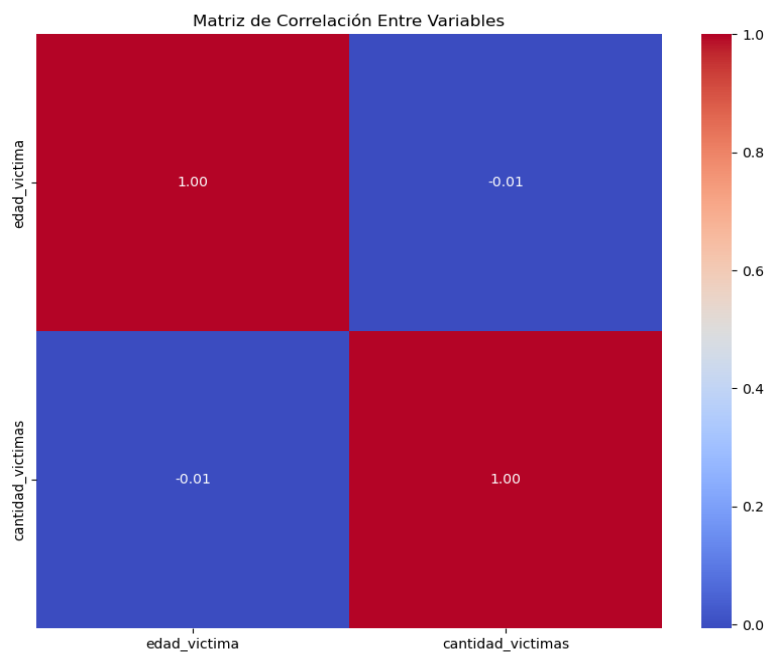
El gráfico muestra la distribución de la gravedad de los accidentes según el tipo. Los "choques" son el tipo de accidente más común, con una gran cantidad de heridos y una menor proporción de fallecidos. Otros tipos de accidentes, como atropellos, caídas de ocupantes, volcamientos, incendios y "otros," tienen una frecuencia significativamente menor. Entre ellos, los atropellos destacan con una proporción ligeramente mayor de fallecidos en comparación con los demás tipos.

Figura 15

Diagrama de Caja y Bigotes de las Variables Edad y Gravedad

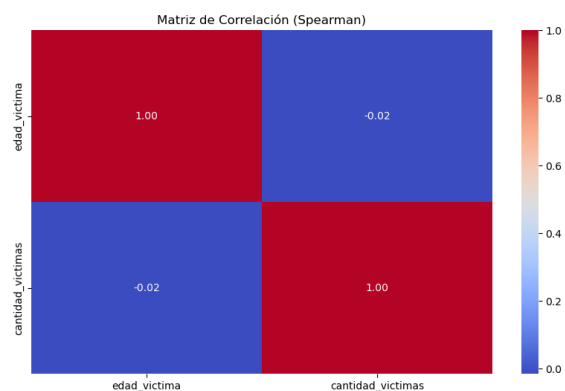


La gráfica muestra que los accidentes afectan principalmente a personas en edad laboral (30-40 años), pero los adultos mayores son más vulnerables a fallecer. Se recomienda a las entidades de control implementar campañas de sensibilización para jóvenes y adultos, enfocar medidas específicas para proteger a los adultos mayores, mejorar la infraestructura vial con cruces y señalización seguros, y fomentar el uso de tecnologías de asistencia al conductor. Además, es crucial analizar continuamente los datos para identificar patrones y optimizar las estrategias preventivas.

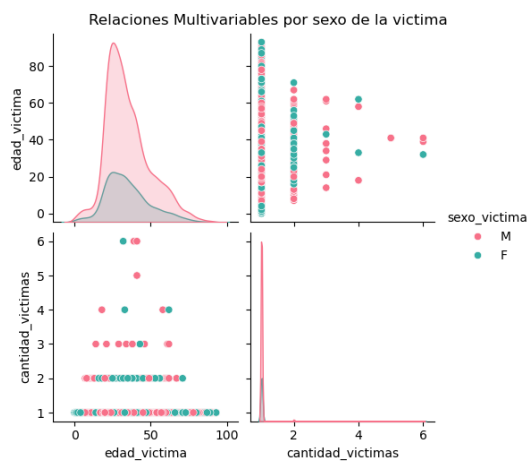
Figura 16*Matriz de Correlación*

La matriz indica que no hay una relación lineal significativa entre la edad de las víctimas y la cantidad de víctimas por accidente. Esto significa que estas variables no varían de manera predecible una con respecto a la otra, por lo que cambios en la edad no influyen directamente en el número de víctimas por accidente.

Debido a que no se encontró una relación lineal significativa, se realizó una matriz de correlación de Spearman para identificar posibles relaciones no lineales entre la edad de las víctimas y la cantidad de víctimas por accidente. Este método evalúa relaciones monotónicas, proporcionando un análisis más robusto en caso de que las variables no sigan una distribución lineal.

Figura 17*Matriz de Correlación de Spearman*

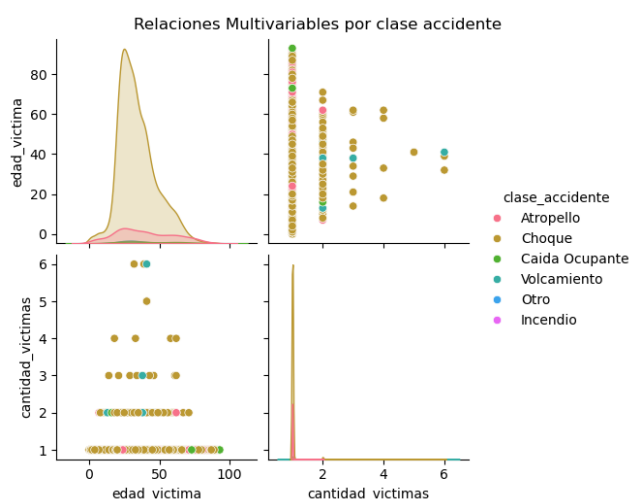
La matriz de correlación de Spearman confirma que no existe una relación monotónica significativa entre la edad de las víctimas y la cantidad de víctimas por accidente, con valores cercanos a 0. Esto indica que estas variables no están relacionadas de forma directa ni no lineal. Para las entidades de control, esto sugiere que deben enfocarse en otras variables más relevantes, como el tipo de accidente, las condiciones viales o factores temporales, para diseñar estrategias preventivas efectivas.

Figura 18*Análisis Multivariado 1*

La gráfica muestra relaciones entre la edad de las víctimas, la cantidad de víctimas y el sexo (M y F). Se observa que la mayoría de las víctimas tienen edades entre 20 y 40 años, con un pico pronunciado en este rango, y son principalmente de sexo masculino. Además, la cantidad de víctimas por incidente suele ser baja (1 o 2), siendo rara una mayor cantidad. Para las entidades de tránsito, se recomienda enfocar campañas de sensibilización y educación vial en conductores y adultos jóvenes, especialmente hombres, ya que son los más afectados.

Figura 19

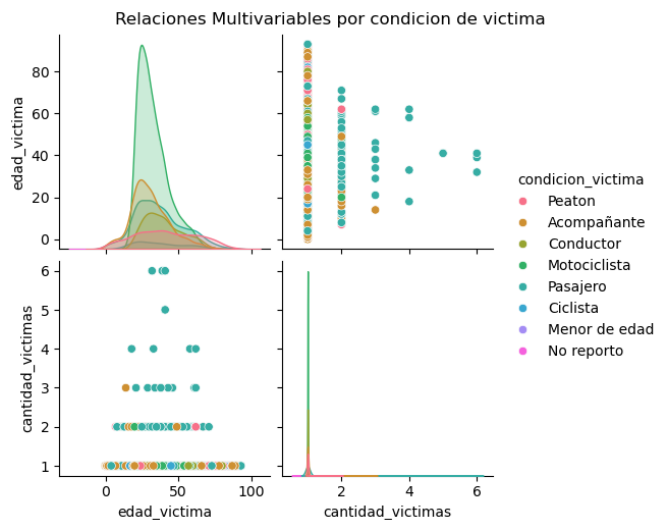
Análisis Multivariado 2



La gráfica muestra que los choques y atropellos son los accidentes más frecuentes, especialmente en víctimas jóvenes (20-40 años). Los incidentes con múltiples víctimas son raros y se concentran en choques y volcamiento. Se recomienda implementar campañas de educación vial, mejorar cruces peatonales, controlar la velocidad y reforzar medidas de seguridad vehicular para reducir estos tipos de accidentes.

Figura 20

Análisis Multivariado 3

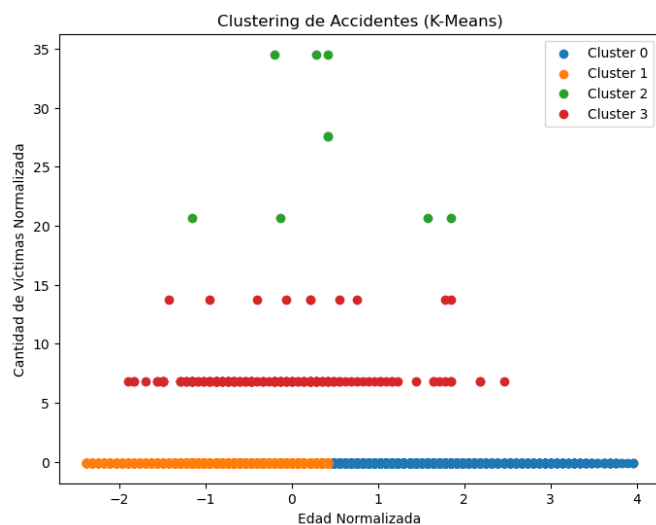


La gráfica muestra que las víctimas más comunes son conductores, peatones y motociclistas, principalmente jóvenes entre 20 y 40 años. La cantidad de víctimas por incidente suele ser baja, mayormente de una o dos personas. Los peatones y menores de edad tienen menor representación, pero siguen siendo vulnerables. Para las entidades de tránsito, se recomienda fortalecer la seguridad vial con campañas dirigidas a conductores y motociclistas, implementar infraestructura segura para peatones y ciclistas, y fomentar la educación vial desde edades tempranas para reducir accidentes y proteger a los grupos más vulnerables.

Análisis de Clustering (K-Means)

Figura 21

Gráfico de K-Medias

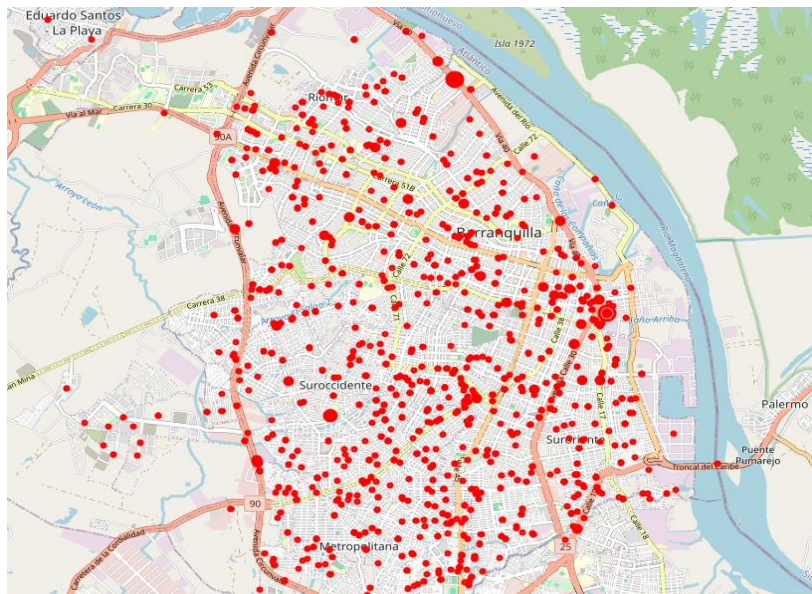


En la gráfica se identifican cuatro clusters basados en la edad normalizada de las personas involucradas y la cantidad de víctimas. El Cluster 0 (azul) agrupa accidentes menores en edades bajas, el Cluster 1 (naranja) refleja eventos leves en personas de mayor edad, el Cluster 2 (verde) destaca casos graves con un alto número de víctimas y un rango de edad variado, y el Cluster 3 (rojo) incluye la mayoría de los accidentes con un número medio de víctimas. Se recomienda focalizar campañas de prevención en los grupos con mayor incidencia de víctimas (Clusters 2 y 3), implementar controles y educación vial en zonas de riesgo, e investigar las causas de los accidentes graves para diseñar políticas más efectivas.

Mapa de Burbujas

Figura 22

Mapa de Burbujas



Se generó un mapa de burbujas para visualizar la distribución geográfica de los accidentes de tránsito en Barranquilla, Colombia. Para ello, se ajustaron las direcciones de los registros agregando la ciudad y el país, y se obtuvieron sus coordenadas geográficas mediante la librería geopy. Luego, se eliminaron los registros sin coordenadas válidas para asegurar la precisión del análisis. Con la información procesada, se creó un mapa interactivo con Folium, centrado en el punto medio de las coordenadas obtenidas. En este mapa, los accidentes fueron representados con burbujas rojas, cuyo tamaño fue proporcional a la cantidad de víctimas en cada incidente. Finalmente, el mapa se guardó en un archivo HTML, permitiendo una visualización clara de la distribución y gravedad de los accidentes en la ciudad.

Dashboard Interactivo

Figura 23

Dashboard 1

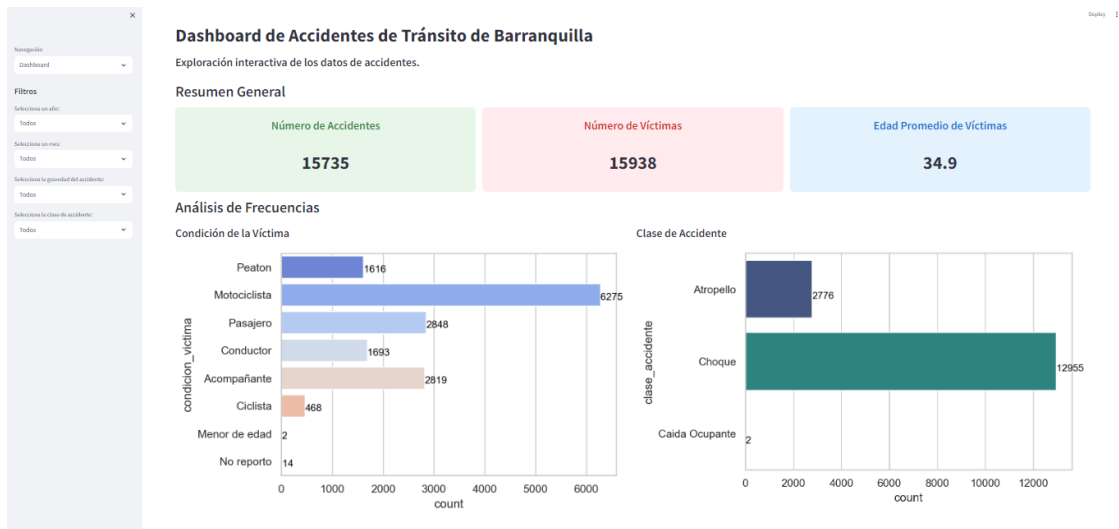


Figura 24

Dashboard 2

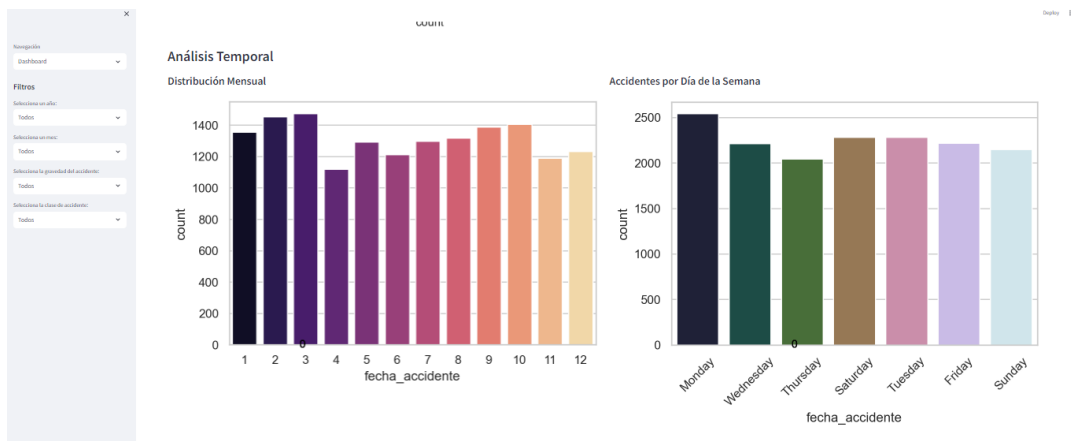


Figura 25

Dashboard 3

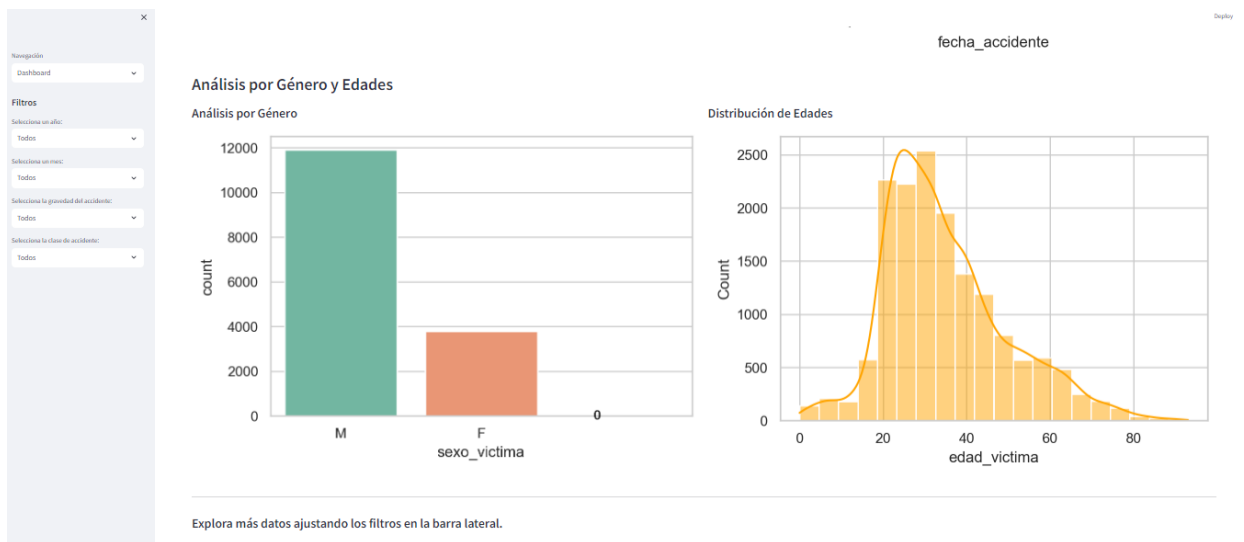
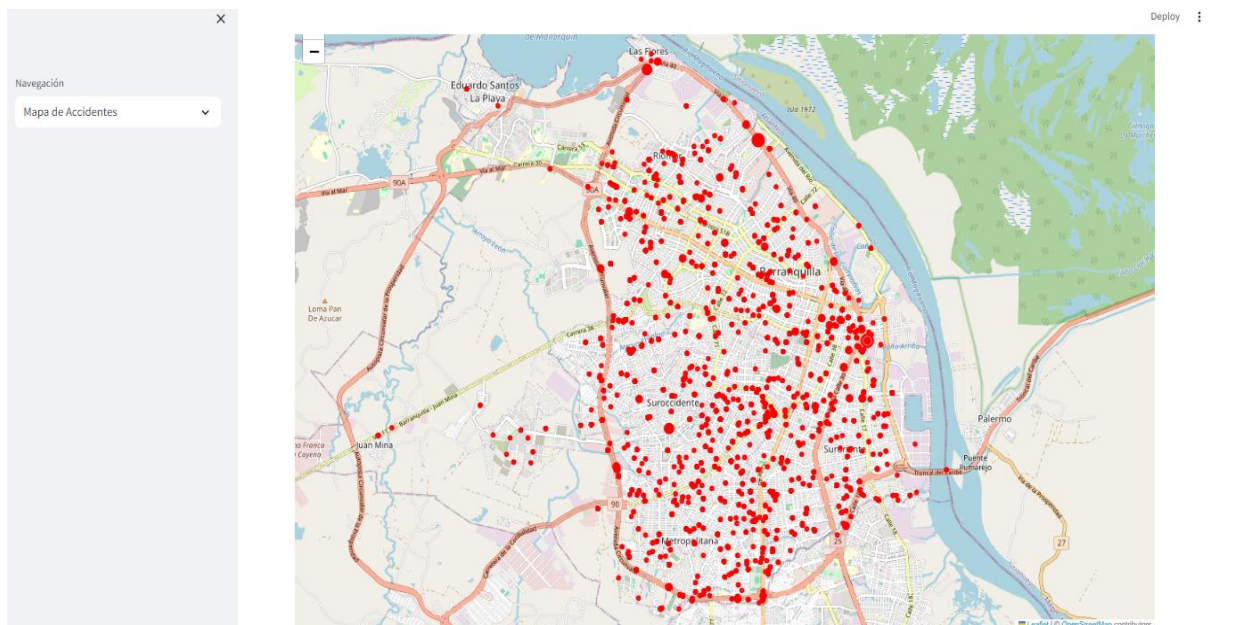


Figura 26

Dashboard 4



Se desarrolló un dashboard interactivo con Streamlit para analizar los accidentes de tránsito en Barranquilla, Colombia. Para ello, se estableció una conexión con una base de datos SQL alojada en Supabase, desde donde se extrajeron los datos de víctimas, accidentes y sus características. Luego, se realizaron transformaciones en los datos, como la conversión de fechas, la eliminación de edades atípicas y la aplicación de filtros dinámicos por año, mes, gravedad y tipo de accidente. El dashboard presentó un resumen general con métricas clave, como el número de accidentes, la cantidad total de víctimas y la edad promedio de los afectados, utilizando tarjetas con estilos personalizados. Además, se incluyeron diversas visualizaciones con Matplotlib y Seaborn para analizar la distribución de los accidentes según la condición de la víctima, la clase de accidente, la distribución mensual y semanal, así como el desglose por género y edades.

Adicionalmente, se incorporó un módulo de visualización geoespacial donde se integró un mapa interactivo de accidentes. Este mapa, generado previamente con Folium, mostró la ubicación de los accidentes en Barranquilla, representados con burbujas rojas cuyo tamaño fue proporcional a la cantidad de víctimas en cada incidente. El archivo HTML del mapa fue cargado en el dashboard, permitiendo a los usuarios explorar de manera visual la concentración y distribución geográfica de los accidentes reportados. Finalmente, se proporcionó una barra lateral con opciones de navegación y filtros, permitiendo una exploración más detallada de los datos.

El despliegue del dashboard se realizó utilizando Streamlit Community Cloud, una plataforma gratuita para alojar aplicaciones de Streamlit. Primero, se creó un repositorio en GitHub donde se subieron los archivos esenciales: el código principal del dashboard (`app.py`) y el archivo de dependencias (`requirements.txt`). Luego, en la plataforma de Streamlit Community Cloud, se vinculó el repositorio seleccionando la opción de "Deploy an app". Tras elegir la rama

y el archivo principal, la plataforma instaló automáticamente las dependencias y configuró el entorno. Finalmente, se generó un enlace público que permitió acceder y compartir el dashboard de forma inmediata.

Conclusión

El análisis de la accidentalidad vial en Barranquilla, basado en técnicas avanzadas de ciencia de datos y analítica predictiva, ha permitido una comprensión profunda de los factores que influyen en la ocurrencia de accidentes de tránsito en la ciudad. A través del uso de herramientas como minería de datos, análisis exploratorio y modelos predictivos, se han identificado patrones significativos que permiten a las autoridades y tomadores de decisiones diseñar estrategias de prevención más efectivas y focalizadas. La implementación de un dashboard interactivo desarrollado en Python y MySQL ha demostrado ser una herramienta clave en la gestión y visualización de la información, permitiendo no solo una interpretación clara y dinámica de los datos, sino también la posibilidad de identificar en tiempo real las zonas y condiciones de mayor riesgo, optimizando la toma de decisiones y la asignación de recursos.

Los hallazgos de este estudio revelan que los motociclistas y jóvenes adultos son los grupos más vulnerables, siendo los más afectados en la accidentalidad vial. Además, se determinó que los choques representan el tipo de accidente más frecuente y que los lunes presentan un mayor número de incidentes, lo que sugiere una correlación con la movilidad laboral y la congestión vehicular al inicio de la semana. Estos resultados resaltan la importancia de establecer medidas preventivas dirigidas a los actores viales más expuestos, reforzar las campañas de seguridad vial y fortalecer la regulación del tráfico en horarios y puntos críticos.

Otro aspecto clave identificado en este estudio es la necesidad de una integración eficiente de datos provenientes de diversas fuentes, lo que permitiría mejorar la calidad y disponibilidad de la información para un análisis más robusto. La utilización de técnicas de Big Data y analítica avanzada ha evidenciado que el manejo de información estructurada y

actualizada es un pilar fundamental para diseñar políticas públicas y estrategias de mitigación con base en evidencia empírica.

En términos de impacto, este proyecto no solo aporta valor en el ámbito académico y tecnológico, sino que también establece un modelo innovador de análisis de datos para la seguridad vial, que puede ser replicado en otras ciudades con problemáticas similares. Su implementación práctica tiene el potencial de reducir la frecuencia de los accidentes de tránsito, mejorar la movilidad urbana y contribuir al desarrollo de una infraestructura vial más segura y eficiente.

Recomendaciones

Fortalecimiento de los sistemas de información:

Implementar una plataforma unificada que integre datos de todas las entidades relacionadas con la seguridad vial en Barranquilla.

Estandarizar el formato de recolección y almacenamiento de datos para facilitar su análisis y compatibilidad con herramientas de inteligencia artificial y analítica avanzada.

Estrategias de prevención basadas en datos:

Desarrollar campañas dirigidas a motociclistas y jóvenes adultos, considerando su alta incidencia en los accidentes.

Reforzar la seguridad en los días y horarios con mayor cantidad de incidentes, como los lunes y las horas pico.

Implementar estrategias para reducir choques y atropellos, mejorando la señalización vial y aumentando controles de velocidad en zonas críticas.

Optimización del uso del dashboard:

Integrar el dashboard con sistemas de alerta temprana para notificar a las autoridades sobre zonas de alto riesgo.

Permitir la actualización en tiempo real de los datos, mejorando su utilidad para la toma de decisiones inmediatas.

Capacitar a funcionarios de tránsito y seguridad vial en el uso del dashboard para maximizar su impacto en la reducción de accidentes.

Monitoreo y mejora continua:

Evaluar periódicamente el impacto del dashboard y ajustar sus funcionalidades según nuevas necesidades y tendencias de accidentalidad.

Ampliar el análisis a nuevas variables, como condiciones climáticas y factores socioeconómicos, para una comprensión más profunda de la accidentalidad vial.

Articulación con políticas públicas:

Incorporar los hallazgos de este estudio en planes estratégicos de movilidad y seguridad vial en Barranquilla.

Fomentar la colaboración entre gobierno, academia y sector privado para garantizar la sostenibilidad y actualización constante del sistema de análisis de datos.

Fomento de la investigación continua:

Utilizar los datos y herramientas de este proyecto como base para futuras investigaciones en movilidad y seguridad vial.

Explorar modelos predictivos avanzados que permitan estimar el riesgo de accidentes en diferentes contextos urbanos, optimizando la planificación y prevención vial.

Link : . <https://dash-accidentes-baq-wwzpyzc7jfte25rx9az3k8.streamlit.app/>

Referencias Bibliográficas

- Cairo, A. (2016). *The Truthful Art: Data, Charts, and Maps for Communication*. New Riders.
- Davenport, T. H., & Patil, D. J. (2014). *Data Scientist: the Sexiest Job of the 21st Century*.
Harvard business review, 92(5), 70-76.
- Tufte, E. (2001). *Data Visualization: A Handbook for Data Exploration and Communication*.
Sage Publications.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *From Data Mining to Knowledge
Discovery in Databases*. AI magazine, 17(3), 37-37.
- Fernández, D. C. (2014). *Desarrollo de Bases de Datos: Casos Prácticos desde el Análisis a la
Implementación* (2ª ed. act., Vol. 10). RA-MA.
- Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*.
MorganKaufmann.
- Katal, A., Wazid, M., & Goudar, R. H. (2013). *Big Data: Issues, Challenges, Tools and Good
Practices*. En 2013 Sixth International Conference on Contemporary Computing (IC3)
(pp. 404-409). IEEE.
- Ministerio de Transporte de Colombia. (2022). *Plan Nacional de Seguridad Vial*.
[\[https://ansv.gov.co/es/escuela/9760#:~:text=%C2%A1Bienvenido%20a%20conocer%20el%20Plan,y%20aumentando%20la%20capacidad%20institucional.\]](https://ansv.gov.co/es/escuela/9760#:~:text=%C2%A1Bienvenido%20a%20conocer%20el%20Plan,y%20aumentando%20la%20capacidad%20institucional.)
- Moses, B., Gavish, L., & Vorwerck, M. (2022). *Data Quality Fundamentals*. " O'Reilly Media, Inc."
- Organización Mundial de la Salud (OMS). (2022). *Informe Mundial Sobre la Seguridad Vial*.
[\[https://www.who.int/es/news/item/13-12-2023-despite-notable-progress-road-safety-remains-urgent-global-issue\]](https://www.who.int/es/news/item/13-12-2023-despite-notable-progress-road-safety-remains-urgent-global-issue)

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know About Data Mining & Data-Analytic Thinking*. O'Reilly Media, Inc.

Secretaría de Tránsito y Transporte de Barranquilla. (2023). *Estadísticas de Accidentes de Tránsito*. [<https://www.barranquilla.gov.co/page/2?s=estadisticas>]

Vargas Montero, F. (2022). *Análisis de Datos de Accidentalidad Vial de la Ciudad de Bogotá a Partir de Datos Abiertos y Datos Obtenidos desde Redes Sociales*. [Trabajo de Grado inédito]. Universidad Nacional de Colombia.