



FACULTAD DE CIENCIAS  
**EXACTAS**  
UNIVERSIDAD NACIONAL DEL CENTRO  
DE LA PROVINCIA DE BUENOS AIRES

# Informe Trabajo Práctico Final

Cursada 2025

**Autores:**

Fernández Lema, Brisa  
Knudsen, Benjamín  
Ramos, Lautaro

**Grupo:** 5

**Cátedra:** Fundamentos de la Ciencia de Datos

**Fecha de entrega:** 16/11/2025

# Índice

---

Índice .....	1
1. Introducción .....	3
2. Análisis exploratorio de los datos .....	5
2.1. Contexto de la entidad y del espacio del problema.....	5
2.1.1. Autopistas.....	5
2.2. Revisión de la calidad y estructura del dataset.....	8
2.3. Descripciones iniciales y observaciones de distribución.....	11
2.4 Primeras relaciones entre variables y motivación de hipótesis.....	11
3. Hipótesis planteadas y resolución.....	13
3.1. Hipótesis 1.....	13
3.1.1. Definición de la hipótesis .....	13
3.1.2. Estrategia de abordaje .....	13
3.1.3. Resultados obtenidos y discusión .....	14
3.2. Hipótesis 2.....	17
3.2.1. Definición de la hipótesis .....	17
3.2.2. Estrategia de abordaje .....	17
3.2.3. Resultados obtenidos y discusión .....	17
3.3. Hipótesis 3 .....	19
3.3.1. Definición de la hipótesis .....	19
3.3.2. Estrategia de abordaje .....	19
3.3.3. Resultados obtenidos y discusión .....	20
3.4. Hipótesis 4.....	22
3.4.1. Definición de la hipótesis .....	22
3.4.2. Estrategia de abordaje .....	22
3.4.3. Resultados obtenidos y discusión .....	22
3.5. Hipótesis 5 .....	24
3.5.1. Definición de la hipótesis .....	24
3.5.2. Estrategia de abordaje .....	24

3.5.3. Resultados obtenidos y discusión .....	24
3.6. Hipótesis 6.....	26
3.6.1. Definición de la hipótesis .....	26
3.6.2. Estrategia de abordaje .....	26
3.6.3. Resultados obtenidos y discusión .....	26
4. Conclusiones.....	28
4.1. Recomendaciones .....	29
Referencias.....	30

# 1. Introducción

---

En la ciudad de Buenos Aires, las autopistas urbanas constituyen vías clave para la movilidad diaria de cientos de miles de personas, integrando la circulación entre la Capital y el Gran Buenos Aires. La empresa Autopistas Urbanas Sociedad Anónima (AUSA) es la concesionaria encargada de administrar, mantener y operar varias de estas autopistas —entre ellas, la Autopista 25 de Mayo, la Autopista Perito Moreno y la Autopista Dellepiane— con el objetivo declarado de garantizar una circulación más rápida, segura y eficiente.

Dado este contexto, la siniestralidad vial en dichas autopistas representa un problema relevante tanto para la gestión urbana como para la seguridad pública: conocer sus patrones temporales, espaciales y de severidad es esencial para tomar decisiones basadas en datos que contribuyan a la reducción de accidentes y a la mejora de la experiencia vial.

En este trabajo analizaremos el conjunto de datos proporcionado por AUSA que contiene el registro de siniestros ocurridos en autopistas urbanas en Buenos Aires durante la segunda mitad del año 2022. El dataset contiene, entre otros atributos, la fecha, la hora, la autopista, la banda o ramal, el punto kilométrico, las condiciones meteorológicas, el estado de la superficie, así como el número de lesionados, fallecidos y los tipos de vehículos involucrados.

Este tipo de información revela su utilidad al permitir realizar análisis exploratorios para identificar cuándo y dónde ocurren más accidentes, cuáles son las condiciones más críticas, así como para evaluar la efectividad de medidas de control o infraestructura vial.

Informes técnicos destacan que la densidad de tránsito, las condiciones meteorológicas adversas y las características de las autopistas son factores que inciden significativamente en la siniestralidad vial, por lo que disponer de este tipo de datos mejora la capacidad de gestión.

En lo que respecta a la organización del trabajo, la estructura se divide en las siguientes secciones:

- **Análisis exploratorio de los datos:** En esta sección realizamos un primer recorrido descriptivo del conjunto de datos, evaluando su calidad, estructura y características iniciales. Incluye la revisión de valores nulos, tipos de variables, distribución básica de las principales variables, así como observaciones preliminares que permitieron motivar las hipótesis que luego se contrastan.
- **Hipótesis planteadas y resolución:** Aquí planteamos cada hipótesis (definición, estrategia de abordaje, resultados, discusión), vinculándolas con los hallazgos del análisis exploratorio, presentando gráficas, métricas y concluyendo si fueron o no confirmadas.
  - **Análisis univariado:** Se describen las distribuciones de cada variable relevante (por ejemplo: autopista, hora, condiciones meteorológicas, superficie, etc.) y se contrastan con hipótesis univariadas específicas.
  - **Análisis bivariado:** Se profundiza en las relaciones entre pares de variables, por ejemplo banda vs hora, autopista vs número de lesionados.

- **Análisis multivariado:** Se consideran interacciones de tres o más variables para evaluar hipótesis más complejas, tales como la relación entre tipo de siniestro, condiciones y banda.
- **Conclusiones:** Se resumen los hallazgos principales, se contrastan con las hipótesis planteadas y se sugieren líneas de mejora o investigación futura.

Con esta estructura, el trabajo pretende ofrecer una visión clara y fundamentada de la siniestralidad en autopistas urbanas de Buenos Aires durante el periodo analizado, con enfoque en aportar información útil para la gestión vial.

## 2. Análisis exploratorio de los datos

---

En esta sección realizamos un primer recorrido descriptivo del conjunto de datos proporcionado por Autopistas Urbanas S.A. (AUSA), con el fin de comprender su estructura, calidad y algunas características iniciales que nos permitieron plantear las hipótesis que se contrastan más adelante.

### 2.1. Contexto de la entidad y del espacio del problema

La empresa AUSA tiene a su cargo la administración, explotación, mantenimiento y conservación de la red de autopistas concesionadas en la Ciudad Autónoma de Buenos Aires, conforme a la Ley N° 3060/09.

Su rol estratégico en la movilidad urbana hace que los siniestros viales en las autopistas que gestiona sean un problema significativo para la seguridad vial, la gestión del tránsito y la infraestructura.

En particular, algunas de las autopistas analizadas —como la Autopista 25 de Mayo (AU1), la Autopista Perito Moreno (AU6) y la Autopista Dr. Arturo Umberto Illia— forman parte de una red que registra diariamente cientos de miles de vehículos.

Por ello, el estudio de los siniestros en estas vías —en términos de cuándo, dónde y en qué condiciones ocurren— permite aportar información relevante para optimizar políticas de gestión vial y medidas de seguridad.

#### 2.1.1. Autopistas

Sobre las autopistas más relevantes, podemos destacar las siguientes:

##### *Autopista 25 de Mayo (AU1)*

Esta autopista recorre la ciudad de Buenos Aires de oeste a este, conectando la zona de Parque Avellaneda / barrio de Liniers (a través del empalme con la Autopista Perito Moreno) con la zona del centro/puerto (hacia el empalme con la avenida Huergo, el Paseo del Bajo y la Autopista Buenos Aires-La Plata).

Fue inaugurada el 6 de diciembre de 1980 (Billiken, s.f.). Su longitud se estima en alrededor de 9,5 km para el tramo principal. Al ser una vía de alta densidad de tránsito que atraviesa zonas céntricas tiene mayor exposición al riesgo de siniestros (Wikipedia, s.f.)

##### *Autopista Perito Moreno (AU6)*

La Autopista Perito Moreno arranca en la avenida General Paz (continuación del Acceso Oeste) y termina en el empalme con la Autopista 25 de Mayo / Autopista Dellepiane, atravesando barrios como Liniers, Villa Luro y Parque Avellaneda. Es pieza clave de la movilidad entre el sur-oeste del conurbano y el centro de la ciudad. Su longitud aproximada es de 6 km, y por su trazado y volumen, también es probable que tenga una incidencia importante en el número de siniestros registrados.

### *Autopista Dr. Arturo Umberto Illia*

La Autopista Illia es otra de las vías principales concesionadas por AUSA. Tiene un rol clave como acceso al centro de la ciudad desde los barrios más al norte-este y conecta con otras autopistas principales. Su longitud aproximada es de 6.5km al norte, y de 4km al centro.

#### **Observaciones**

Durante el periodo de análisis, la concesionaria AUSA ha comenzado a implementar sistemas de peaje automático sin barreras y con lectura de patente, bajo su plan ‘Autopistas Inteligentes’. (Por ejemplo, en la Autopista Illia). Si bien muchos de esos cambios se intensificaron tras 2022, resulta relevante considerarlos como posibles factores de cambio en la dinámica vial y en la incidencia de siniestros (Autopistas Urbanas S.A., 2023).

Si bien el dispositivo TelePASE era obligatorio para circular por las autopistas gestionadas por AUSA desde septiembre de 2020, el sistema de peaje *sin barreras* (free-flow) aún no estaba implementado en todos los tramos para 2022 (Autopistas Urbanas S.A., 2024). Esto implica que, aunque el método de pago podía estar modernizado, la presencia de cabinas y barreras podría variar entre autopistas y sentidos, lo cual debe considerarse como una posible influencia en los flujos de tránsito y en la ocurrencia de siniestros en ese periodo (Canal 26, 2025).

Sin embargo, la falta de información precisa sobre qué tramos ya operaban sin barreras en 2022 impide asumir un efecto uniforme de la modernización del peaje en los análisis. Pero es un dato que puede ser un punto de estudio futuro en una investigación sobre los siniestros de las autopistas bonaerenses.



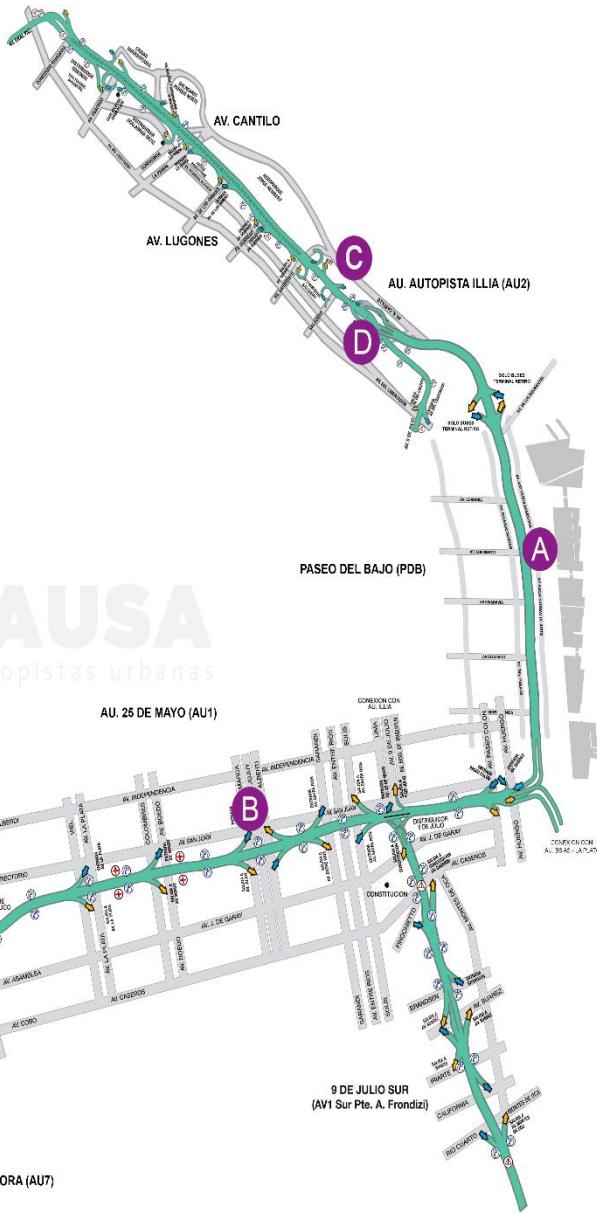
#### TelePASE OBLIGATORIO

- A Paseo del Bajo
- B Alberti
- C Salguero y Sarmiento
- D Illia y Retiro II
- E Parque Avellaneda sentido Centro y Provincia



#### TelePASE Y EFECTIVO

- F Dellepiane Centro y Aeropuerto



NUESTRAS AUTOPISTAS		
AUTOPISTAS	RECORRIDO	EXTENSIÓN
AU 9 de Julio Sur	de Av. 9 de Julio a Nuevo Puerto Pueyrredón - Avellaneda	3,5 km.
AU Presidente Cámpora	de Au. Dellepiane a Av. 27 de Febrero	3,6 km.
AU Dellepiane	de Au. 25 de Mayo y Pte. Cámpora a Gral. Paz y Richieri	4,2 km.
AU 25 de Mayo	de Au. Dellepiane y Au. Perito Moreno a Au. Bs As - La Plata	10,0 km.
AU Perito Moreno	de Av. 25 de Mayo a Gral. Paz y Acceso Oeste	6,0 km.
AU Illia al Norte	de Av. 9 de Julio a Cantilo	6,5 km.
AU Illia al Centro	de Lugones a 9 de Julio	4,0 km.
Av. Lugones	de Gral. Paz a Au. Illia	6,2 km.
Av. Cantilo	de Au. Illia a Gral. Paz	3,5 km.
Paseo del bajo	de Av. 25 de Mayo a Au. Illia	6,6 km.

Figura 1. Mapa de la red de autopistas gestionada por AUSA, e identificación de los puntos de "Peaje sin barrera" del programa "Autopistas Inteligentes" en la Ciudad de Buenos Aires.

## 2.2. Revisión de la calidad y estructura del dataset

El dataset entregado cubre el periodo junio 2022 hasta diciembre 2022. Contiene las siguientes variables:

- **FECHA:** Fecha del siniestro en formato YYYY-MM-DD. Es una variable de tipo cuantitativa que no tiene valores nulos, pero sí carece de registros en el mes de julio. La moda de esta variable es Septiembre (100 filas).
- **HORA:** Hora del siniestro en formato 24 horas. Es una variable cuantitativa discreta que no tiene tuplas nulas. El rango horario abarca desde las 0hs hasta las 23hs, y el horario donde más siniestros ocurrieron (la moda) es a las 7hs (38 registros/filas).
- **AUTOPISTA:** Nombre de la autopista donde ocurrió el siniestro. Es una variable cualitativa nominal. Las autopistas registradas son: AU 25 DE MAYO, AU DELLEPIANE, AU PERITO MORENO, AV. LUGONES, AU ILLIA, AU FRONDIZI, AV. CANTILLO, AU CAMPORA, PASEO DEL BAJO, TRANSICION RAMAL 2, DIST. 9 DE JULIO, TRANSICION RAMAL 3, DIST. DELLEPIANE, TRANSICION RAMAL 1, PTE. SCALABRINI ORTIZ. La moda de esta variable indica que la autopista con más siniestros es (spoiler alert): la Autopista 25 de mayo! (116 filas).
- **BANDA y/o RAMAL:** Indica si el siniestro ocurrió en la banda ascendente, descendente o en un ramal. Es una variable cualitativa nominal. Contiene valores nulos que no están representados como NaN y representan el 0.21% de la muestra. Esto se puede deber a una carga incorrecta de los datos. Respecto a los valores que puede tomar esta variable, estos se clasifican de la siguiente manera:
  - **BANDA ASCENDENTE:** Es el carril de circulación que va hacia el este, entrando a la Ciudad Autónoma de Buenos Aires.
  - **BANDA DESCENDENTE:** Es el carril de circulación que va hacia el oeste, saliendo de la Ciudad y dirigiéndose hacia el Gran Buenos Aires. La mayoría de los siniestros del dataset ocurren en este tipo de banda (registrado en 236 filas).
  - **RAMAL:** Es una vía de conexión. Es el tramo de calzada que permite a los vehículos entrar, salir o conectarse con otra autopista. No es la vía principal, sino la "rama" que se desprende o se une a ella.
    - **RAMAL DE INGRESO (o Entrada):** La vía que se toma desde una calle o avenida para incorporarse al flujo principal de la autopista.
    - **RAMAL DE EGRESO (o Salida):** La vía que se toma para abandonar la autopista y bajar a una calle o avenida.
    - **RAMAL DE ENLACE:** Es la vía que conecta dos autopistas diferentes entre sí, permitiendo el cambio de una a otra sin abandonarlas.
- **PK:** Punto kilométrico donde ocurrió el siniestro (en metros). Es una variable cuantitativa discreta que tiene un 3.78% de valores nulos. Además de transformar estos valores nulos a NAN, convertimos la variable de tipo object a Int64. Por otro lado, el valor máximo que puede tomar es 9800 (para la autopista 25 de mayo), el valor más frecuente registrado es 5400 (15 filas) y la mediana 3500. Hay que tener en cuenta, igualmente, que cada autopista tiene longitudes distintas.

- **CONDICIONES METEOROLÓGICAS:** Estado del clima al momento del siniestro. Es una variable cualitativa nominal. Puede tomar los valores BUENO, LLUVIOSO, NIEBLA/BRUMA/HUMO. En este dataset, el tipo de clima “Bueno” es el más registrado para los siniestros (463 filas).
- **SUPERFICIE DE LA VÍA:** Estado de la superficie de la vía. Es una variable cualitativa nominal. Puede tomar los valores SECA, MOJADA/HÚMEDA, OTRA. El estado más registrado en el dataset es superficie de la vía: seca (451 filas).
- **LESIONADOS:** Cantidad de personas lesionadas en el siniestro. Es una variable cuantitativa discreta. Como estadística, el 65.2% de los siniestros no tienen lesionados (quizás posean fallecidos, pero no lesionados). Además, encontramos un outlier: ¡un siniestro presenta 14 lesionados!
- **FALLECIDOS:** Cantidad de personas fallecidas en el siniestro. Es una variable cuantitativa discreta. En el dataset el valor máximo obtenido de fallecidos en siniestros es 1.
- **TIPO DE SINIESTRO:** Clasificación del siniestro. Es una variable cualitativa nominal. Se registra como:
  - **COLISIÓN CON DOS O MÁS VEHÍCULOS:** Este es el siniestro prototípico en una autopista. Se define como el impacto entre dos o más vehículos que están en circulación (es decir, ambos o todos están en movimiento al momento inicial del impacto). Conocido comúnmente como choque en cadena.  
En la práctica, esto abarca la gran mayoría de los incidentes en horas pico. Incluye:
    - Colisiones por alcance: El vehículo A impacta por detrás al vehículo B. Es la tipología más frecuente en congestiones o frenadas súbitas.
    - Colisiones laterales: El impacto ocurre en los costados de los vehículos. Generalmente se debe a maniobras de cambio de carril (desplazamiento lateral) sin la debida precaución.
    - Choques múltiples (o en cadena): Involucran a tres o más vehículos, usualmente como resultado de un alcance inicial que se propaga.
  - **COLISIÓN CON OBSTÁCULO FIJO:** Se produce cuando un vehículo en movimiento impacta contra un objeto inamovible que forma parte de la infraestructura vial o de su entorno inmediato.  
En nuestra red, los obstáculos fijos más comunes son las defensas (guardarrail), las columnas de puentes o de cartelería, las cabinas de peaje, y los muros de contención de los viaductos (como en la AU 25 de Mayo o el Paseo del Bajo). Este tipo de siniestro suele estar asociado a la pérdida de control del vehículo por parte del conductor.
  - **OBSTÁCULO NO FIJO:** Esta categoría es más amplia. Se refiere al impacto con un objeto que se encuentra indebidamente sobre la calzada, pero que no es parte de la infraestructura.  
En la práctica, los ejemplos más claros incluyen:
    - Objetos en la calzada: Restos de neumáticos, partes de vehículos desprendidas, o carga caída de un camión.

- Vehículo detenido: Un impacto contra un vehículo que ya estaba detenido sobre un carril de circulación (sea por un desperfecto mecánico o por un siniestro previo). En este caso, el vehículo detenido no es un "participante" en movimiento, sino que actúa como un obstáculo.
- **SINIESTRO DE UN SOLO VEHÍCULO / SIN COLISIÓN:** Este es un término que agrupa incidentes donde un solo vehículo se ve involucrado y el evento inicial no es un impacto contra otro vehículo (aunque puede terminar en una colisión con un objeto fijo, cruzando las categorías). La característica principal es que la causa inicial es una acción o falla del propio vehículo.

Los ejemplos más comunes de esta categoría son:

- Despiste: El vehículo pierde la adherencia y se sale de la calzada (puede terminar en la banquina, en el césped, o en un carril contrario).
- Vuelco: El vehículo pierde su estabilidad y vuelca sobre la calzada, a menudo por una maniobra evasiva brusca, exceso de velocidad en una curva o, en el caso de vehículos pesados (camiones), por un desplazamiento de la carga.
- Incendio: Un incendio por falla mecánica que no involucra una colisión.
- Caída (para motos): Un motociclista que pierde la estabilidad y cae de su vehículo sin haber colisionado con nada.

Es importante notar que, en un siniestro real, estas categorías pueden superponerse. Un despiste (Siniestro de un solo vehículo) puede terminar en una colisión con un obstáculo fijo (guardarrail). Sin embargo, esta clasificación nos ayuda a ordenar el evento inicial que desencadenó el incidente para su correcto análisis (Agencia Nacional de Seguridad Vial, 2023).

- **OTRO:** Registra cualquier siniestro que no haya sido contemplado en las categorías anteriores (siniestros atípicos).

Por último, en nuestro dataset el tipo de siniestro más registrado es colisión con dos o más vehículos (289 filas).

- **MOTO:** Cantidad de motocicletas involucradas en el siniestro. Es una variable cuantitativa discreta.
- **LIVIANO:** Cantidad de vehículos livianos involucrados en el siniestro. Es una variable cuantitativa discreta, la cual presenta una distribución sesgada a la derecha. Cabe aclarar que le cambiamos el nombre a "LIVIANO" por el bienestar mental de los autores de este informe.
- **BUS:** Cantidad de autobuses involucrados en el siniestro. Es una variable cuantitativa discreta.
- **CAMION:** Cantidad de camiones involucrados en el siniestro. Es una variable cuantitativa discreta. El valor máximo para esta variable es 2 (probablemente un outlier).

Durante el análisis identificamos valores nulos en dos variables: BANDA Y/O RAMAL, y PK (punto kilométrico), pero tomamos la decisión de **no eliminar estas filas** porque, al ser un dataset pequeño, consideramos que toda la información que podamos obtener (aunque algunas no estén completas) es sustancial para el análisis de las hipótesis del trabajo, por lo que simplemente cambiamos estos valores (identificados con "-") a NAN. Además, pudimos notar que el mes de julio no tiene registros de siniestros para ninguna autopista, por lo que no pudimos inferir sobre este lapso de tiempo a la hora de realizar las pruebas y sacar conclusiones, lo que refuerza nuestra decisión de no eliminar los registros con valores en banda y/o ramal y/o pk nulos (todo suma). Pero, más allá de los puntos mencionados, nos encontramos con un dataset bastante limpio, y esto seguramente se deba a que AUSA debe hacer un preprocesamiento previo a la entrega del dataset al Gobierno de la Ciudad de Buenos Aires.

## 2.3. Descripciones iniciales y observaciones de distribución

Respecto a la distribución, las variables que indican la severidad del siniestro presentan distribuciones asimétricas (sesgo positivo), con la mayoría de los siniestros sin lesionados ni fallecidos, y con eventos de mayor magnitud representando casos excepcionales.

Además, observamos que la variable hora presenta picos de siniestros concentrados en lo que se conoce como "horas pico", es decir, el horario donde circulan más vehículos por la autopista. Esto lo mencionaremos mucho a la hora de justificar nuestras hipótesis posteriores.

## 2.4 Primeras relaciones entre variables y motivación de hipótesis

Antes de pasar al análisis formal, exploramos algunas relaciones preliminares:

- Se cruzó AUTOPISTA con LESIONADOS para ver si autopistas con más siniestros también presentan mayor número total de lesionados.
- Se observó HORA frente a BANDA y/o RAMAL para comprobar si ciertos horarios están asociados a bandas descendentes o ascendentes.
- También se exploró la incidencia de CONDICIONES METEOROLÓGICAS adversas sobre el número de fallecidos.

Estas observaciones iniciales permitieron plantear las hipótesis que se desarrollan más adelante, ya que se identificaron patrones sugestivos como: mayor concentración de siniestros en ciertas autopistas, incremento en horarios pico, posible dominancia de la banda descendente en el egreso de la ciudad y mayor severidad ante condiciones adversas.

En cuanto a las condiciones meteorológicas y la superficie de la vía, el análisis revela que la mayoría de los siniestros registrados en el periodo se produjeron bajo condiciones "BUENO" o "SECA" (es decir, buen clima y superficie seca), y no se observó un aumento claro de los siniestros asociados a lluvia, niebla o humedad. Esto indica que, en este conjunto de datos, las adversidades meteorológicas no parecen haber sido un factor determinante principal en la ocurrencia de siniestros (es decir, se podría prescindir de estas dos columnas para el análisis). Cabe señalar que el

año 2022 fue considerado un año muy seco para la Ciudad de Buenos Aires: por ejemplo, el acumulado de lluvias en el Observatorio Central de Buenos Aires fue de apenas ~735 mm, colocándolo como el séptimo año más seco desde 1908 (InfoMeteoBA, 2023).

Por lo tanto, aunque intuitivamente se esperaría que lluvia o superficies húmedas incrementen siniestros, en este dataset no se encontró evidencia fuerte de ello, por lo que se sugiere que otros factores —como el tráfico, la autopista, hora o banda— pueden tener un peso mayor en la siniestralidad vial que la condición del tiempo en este contexto particular.

# 3. Hipótesis planteadas y resolución

---

## 3.1. Hipótesis 1

### 3.1.1. Definición de la hipótesis

La hipótesis planteada es: “**La mayor cantidad de siniestros se reportan en la autopista ‘25 de Mayo’.**”

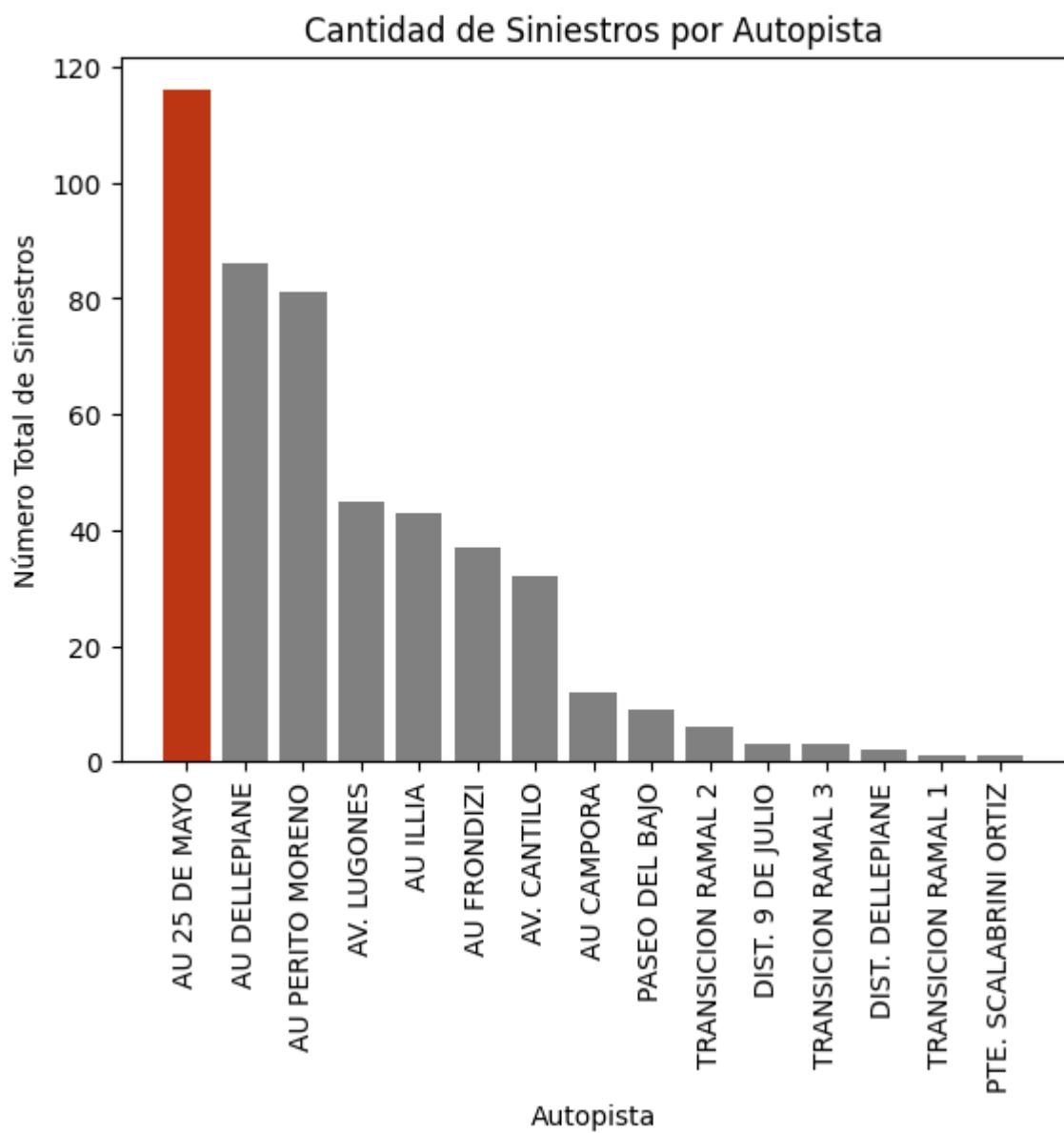
Esta hipótesis es univariada, ya que la variable de estudio será la cantidad de siniestros por autopista, y tiene utilidad práctica porque permite focalizar recursos de prevención vial y mantenimiento en la vía (Autopista 25 de Mayo) que aparentemente concentra la mayor frecuencia de siniestros, lo que puede indicar un punto crítico para la seguridad vial urbana.

### 3.1.2. Estrategia de abordaje

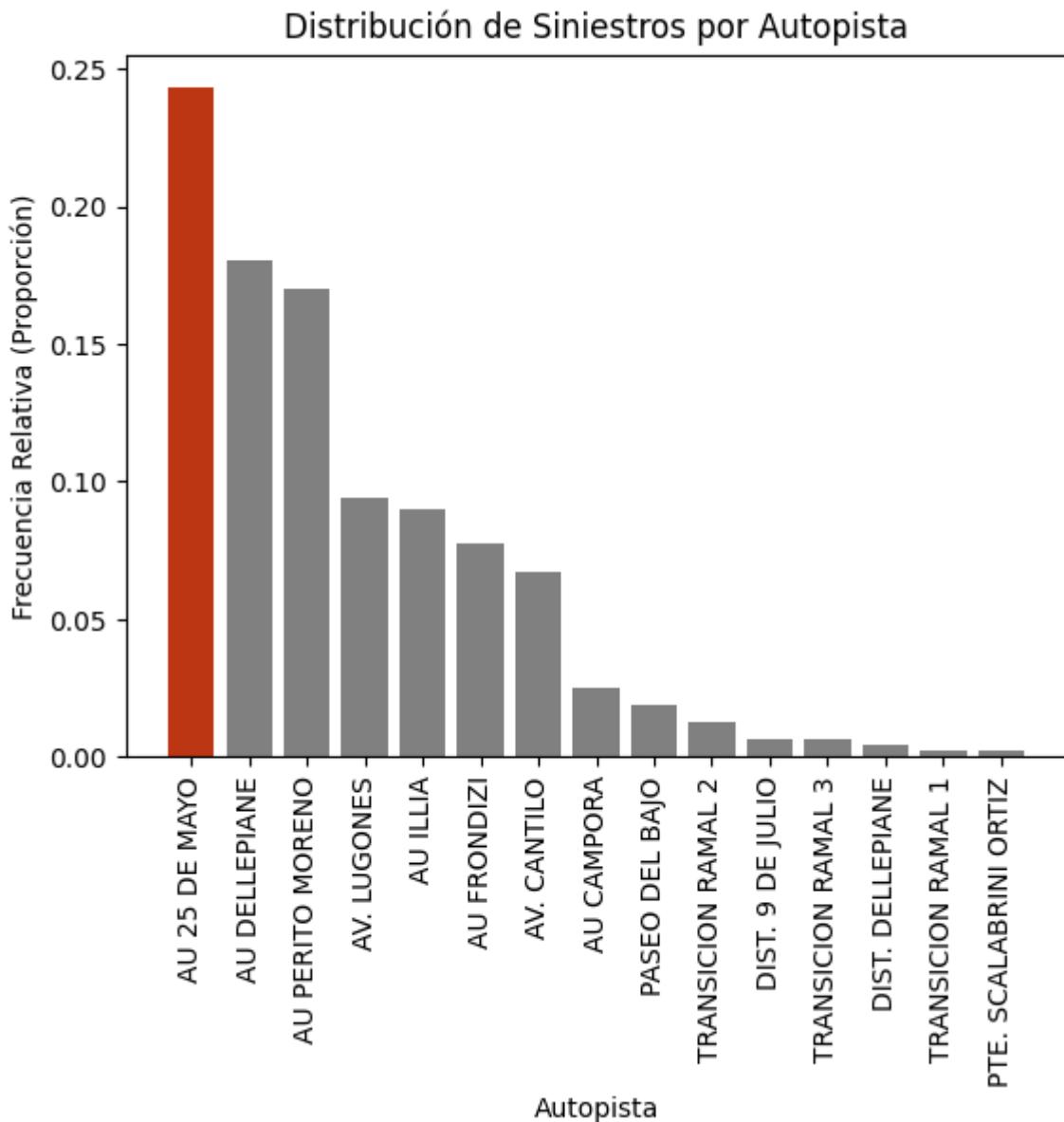
Para verificar esta hipótesis se llevó a cabo el siguiente abordaje:

- Se agrupó el dataset por la variable AUTOPISTA, contando la cantidad de siniestros por autopista.
- Se generaron dos gráficos de barras: uno con la cantidad de siniestros, y otro con las frecuencias relativas (proporción de siniestros por autopista), donde el eje X refleja las autopistas y el eje Y la cantidad/proporción de siniestros.
- Se evaluó cuál autopista presenta el mayor recuento.
- Como referencia contextual mencionada anteriormente, se consultó que la Autopista 25 de Mayo es una de las vías de mayor tránsito de la ciudad (longitud aproximada de 9,5 km, conecta Dellepiane con Av. Huergo).

### 3.1.3. Resultados obtenidos y discusión



*Figura 2: Gráfico de barras para visualizar la cantidad de siniestros por autopista.*



*Figura 3: Gráfico de barras para visualizar la frecuencia relativa (proporción) de siniestros por autopista.*

El gráfico de barras muestra que la Autopista 25 de Mayo presenta el mayor número de siniestros en el periodo analizado (junio-diciembre 2022), seguida por la Autopista Dellepiane, y la Autopista Perito Moreno. En porcentajes, el 24.31% de los siniestros ocurren en la 25 de mayo; el 18.03% en la autopista Dellepiane, y el 16.98% en la Perito Moreno.

Por lo tanto, **nuestra hipótesis ha sido confirmada**.

El hecho de que la Autopista 25 de Mayo sea la que más cantidad de siniestros presenta puede deberse a que es una de las principales vías de acceso y egreso a la Ciudad de Buenos Aires, con alto volumen de tránsito y conexión directa con otras autopistas (Perito Moreno, Dellepiane, 9 de Julio Sur). Además, su extensión y cantidad de carriles aumentan la probabilidad de incidentes en

comparación con autopistas de menor tránsito o extensión, por lo que los resultados mostrados en el gráfico son coherentes con la densidad vehicular y la extensión de la autopista respecto de las demás.

Es importante destacar que, aunque la autopista registre más siniestros, esto no implica que cada siniestro sea más grave. Habría que realizar un análisis posterior para evaluar la cantidad de lesionados o fallecidos por autopista.

**Resultado Hipótesis 1: Confirmada.**

## 3.2. Hipótesis 2

### 3.2.1. Definición de la hipótesis

La hipótesis es: “**La mayor cantidad de siniestros ocurre entre las 6 y las 8 de la mañana.**”

Acá presentamos otra hipótesis univariada, donde la variable de estudio es HORA. Nos interesa principalmente analizar qué ocurre entre las 6 y las 8 de la mañana, ya que este rango horario corresponde al inicio del ingreso laboral y escolar, por lo que entender ese patrón tiene valor para planificar intervenciones en horarios críticos.

### 3.2.2. Estrategia de abordaje

- Se examinó la variable HORA y se construyó un gráfico de líneas que muestra la cantidad de siniestros por hora del día.
- Se identificó el rango horario de 6 a 8 como “**hora pico matutina**” y se comparó visualmente con el resto del día.
- Se verificó que tanto en el gráfico como en los recuentos numéricos, dicho rango presenta el máximo o uno de los máximos de siniestros.
- Se reconoce que las “horas pico” están asociadas a mayor densidad vehicular, distracción del conductor y mayor probabilidad de incidentes (*rush hour*) (Wikipedia, s.f.).

### 3.2.3. Resultados obtenidos y discusión



Figura 4: Gráfico de líneas para visualizar la cantidad de siniestros por hora.

El gráfico de siniestros por hora evidencia un aumento significativo en el periodo entre las 6:00 y las 8:00, corroborando el patrón esperado. Respecto del total, obtuvimos que el 19.28% de los

siniestros ocurren en este horario pico matutino. En la hipótesis 5 observaremos que esto está directamente relacionado con el aumento de siniestros en la banda ascendente en dicho horario, lo cual es coherente debido a que en estos carriles la gente ingresa a la ciudad para ir a trabajar o llevar a los chicos al colegio.

Por otro lado, calculamos la mediana de siniestros por hora, que resultó en **20.0**, lo cual implica que la mitad de las horas analizadas presentan 20 o menos siniestros. Las horas pico, en cambio, superan claramente este valor de referencia (por ejemplo, la franja de 6 a 8h registra más de 25 siniestros por hora) lo que confirma que esos horarios se encuentran por encima del nivel “típico” del día. Este comportamiento también sugiere que la distribución de siniestros por hora está sesgada hacia valores altos en ciertos momentos, lo que refuerza la necesidad de focalizar intervenciones en las franjas que exceden la mediana (podrían considerarse horas de mayor riesgo).

En consecuencia, **se confirma la hipótesis**.

Esto indica que ese rango horario es crítico en términos de siniestralidad vial, lo cual puede orientar estrategias de prevención, señalización y mantenimiento.

**Resultado Hipótesis 2: Confirmada.**

### 3.3. Hipótesis 3

#### 3.3.1. Definición de la hipótesis

Nuestra tercera hipótesis es: “**Las autopistas con más siniestros tienden a registrar más lesionados**”.

Esta hipótesis bivariada busca validar si existe una relación directa entre el volumen total de siniestralidad de una autopista (variable AUTOPISTA como frecuencia) y el número total de lesionados que esta genera (variable LESIONADOS como suma).

Este planteamiento es relevante, ya que sugiere que la frecuencia de siniestros en una autopista puede estar vinculada con una mayor exposición al riesgo (más vehículos circulando, más interacciones) y, por tanto, con un mayor número total de lesionados. En términos de gestión vial, identificar autopistas con este patrón puede ayudar a priorizar intervenciones.

#### 3.3.2. Estrategia de abordaje

Para poder comparar las autopistas entre sí, fue necesario transformar el dataset. Se realizó una operación de agrupación (groupby) por la variable categórica AUTOPISTA. Para cada una de las autopistas, se calcularon dos métricas agregadas:

1. El conteo total de siniestros (contando las filas).
2. La suma total de LESIONADOS.

Con este nuevo dataset agregado, se emplearon dos técnicas para validar la hipótesis:

1. Se generó un **gráfico de dispersión (scatterplot)** para observar la forma, dirección y fuerza de la relación entre el “Total de Siniestros” (eje X) y el “Total de Lesionados” (eje Y). Adicionalmente, se coloreó el punto de la “AU 25 DE MAYO” en rojo para contrastar con las demás, con el mismo color con el que veníamos coloreando a esta autopista para destacarla. El resto de las autopistas tienen colores diversos pero no tan llamativos.
2. Se calculó el **coeficiente de correlación de Pearson (r)** para cuantificar numéricamente la fuerza y dirección de la asociación lineal entre estas dos variables numéricas.

### 3.3.3. Resultados obtenidos y discusión

El análisis visual y estadístico arrojó resultados concluyentes.

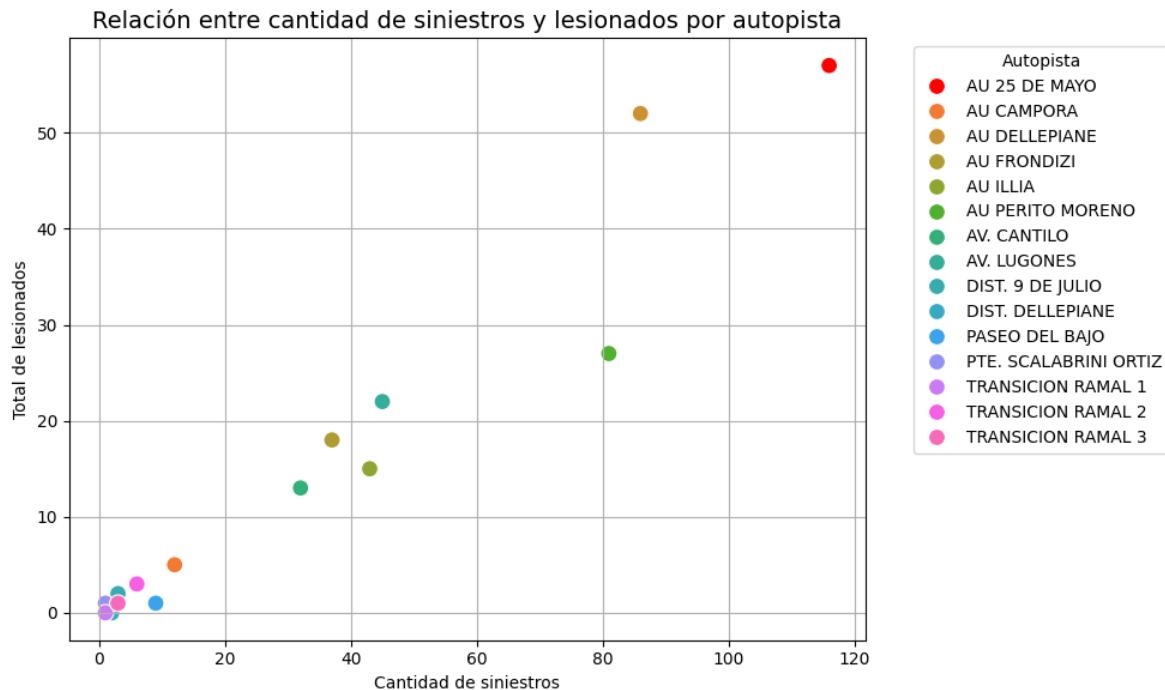


Figura 5: Gráfico de dispersión de Siniestros Totales vs. Lesionados Totales por Autopista.

El gráfico de dispersión (Figura 4) muestra una **clara y fuerte tendencia lineal positiva**. Los puntos (cada uno representando una autopista) se alinean casi perfectamente desde la esquina inferior izquierda hasta la superior derecha.

La “AU 25 DE MAYO” (punto rojo) se sitúa como la autopista con el mayor número de siniestros y, coherentemente, el mayor número de lesionados totales, muy por encima del resto.

El análisis cuantitativo confirmó esta observación visual. El cálculo del coeficiente de correlación de Pearson arrojó un valor de: **r = 0.97**

Un valor tan cercano a +1 indica una asociación lineal positiva casi perfecta. Esto significa que a medida que aumenta el número de siniestros en una autopista, el número total de lesionados aumenta de manera casi directamente proporcional.

En consecuencia, el número total de lesionados parece ser un reflejo directo de la frecuencia de siniestros en cada autopista.

Nuestra justificación inicial para esta relación fue el **volumen de tránsito**: más vehículos implican mayor exposición al riesgo. Sin embargo, esta explicación puede ser enriquecida.

Un informe de AUSA, citado por el diario La Nación (2020), identifica los “puntos negros” de la red. Es de notar que los tramos más peligrosos identificados pertenecen, precisamente, a la **AU 25 DE MAYO**.

**de Mayo**, nuestro principal punto de interés y el mayor outlier en nuestro gráfico. El informe destaca el enlace de la AU 25 de Mayo con la AU Dellepiane y la bajada a la 9 de Julio Sur.

La causa identificada no es solo el volumen, sino la **complejidad de la infraestructura**. Estos "embudos" y "enlaces" obligan a los conductores a realizar maniobras abruptas de cruce de carriles (conocidas como *weave*), lo que incrementa drásticamente la probabilidad de colisiones (La Nación, 2024).

Por lo tanto, la AU 25 de Mayo no solo acumula más lesionados por tener más siniestros, sino que su diseño en puntos críticos genera las condiciones de riesgo que facilitan esas colisiones. Esto refuerza la idea de que la alta siniestralidad (y sus consecuentes lesionados) está fuertemente ligada a la configuración física de la autopista en sus enlaces clave.

**Resultado Hipótesis 3: Confirmada.**

## 3.4. Hipótesis 4

### 3.4.1. Definición de la hipótesis

La siguiente hipótesis que planteamos es la siguiente: “**La probabilidad de colisión con más de dos vehículos aumenta en autopistas ascendentes durante el horario pico matutino (6-8 am)**”.

Esta hipótesis bivariada recae sobre dos variables: BANDA Y/O RAMAL, y HORA, filtrando por colisión con más de dos vehículos (que es un tipo de siniestro). Además, se fundamenta en la dinámica del flujo de ingreso laboral y escolar (identificado en la H2). Se espera que la congestión vehicular se acumule en las bandas **ascendentes** (dirección al centro de la ciudad) durante este rango horario.

La alta densidad reduce la distancia de seguridad entre vehículos, y se teoriza que esto incrementa la probabilidad de colisiones en cadena (clasificadas como “COLISION CON DOS O MAS VEHICULOS”).

### 3.4.2. Estrategia de abordaje

Para validar esta hipótesis, se aisló el fenómeno de interés. Primero, se filtró el dataset para incluir únicamente:

1. Siniestros del tipo "COLISION CON DOS O MAS VEHICULOS".
2. Siniestros ocurridos en las bandas principales "ASCENDENTE" o "DESCENDENTE".

A continuación, se creó una variable categórica binaria (HORA\_PICO) que identifica si el siniestro ocurrió dentro (1) o fuera (0) del rango pico matutino (6-8 hs).

Se utilizó un **Test Chi-Cuadrado ( $\chi^2$ ) de Independencia** para determinar si existe una asociación estadísticamente significativa entre las dos variables categóricas: BANDA y/o RAMAL (Ascendente/Descendente) y HORA\_PICO (Pico/No Pico).

- **Hipótesis Nula ( $H_0$ ):** La banda y el horario pico matutino son independientes. No hay asociación.
- **Hipótesis Alternativa ( $H_1$ ):** Existe una asociación entre la banda y el horario pico matutino.

### 3.4.3. Resultados obtenidos y discusión

El Test Chi-Cuadrado ( $\chi^2$ ) arrojó un estadístico  $g = 9.697$  y un **p-valor = 0.0018**.

Dado que el p-valor (0.0018) es significativamente menor que el umbral de significancia ( $\alpha = 0.05$ ), **se rechaza la hipótesis nula**. Esto confirma que existe una asociación estadísticamente significativa entre la banda de circulación y el horario pico matutino para las colisiones múltiples.

El análisis de la **tabla de contingencia** (Figura 5) permite interpretar la naturaleza de esta asociación, **confirmando la dirección planteada en la hipótesis**.

BANDA Y/O   HORA_PICO RAMAL	0	1
ASCENDENTE	93	35
DESCENDENTE	138	19

*Figura 6: Tabla de Contingencia de colisiones múltiples vs. Banda y Pico Horario Matutino (6-8 am).*

La tabla evidencia una frecuencia de colisiones múltiples significativamente mayor en la banda **ASCENDENTE** durante el rango de 6 a 8 AM, en comparación con la banda descendente en el mismo horario.

Este hallazgo se alinea perfectamente con la H2 (pico matutino) y la lógica del tránsito (flujo de ingreso). La alta densidad en la banda ascendente genera las condiciones de "ondas de tráfico" (o *traffic shockwaves*), donde frenadas abruptas en congestión se propagan hacia atrás, causando colisiones por alcance, la principal causa de choques múltiples (National Highway Traffic Safety Administration, NHTSA, s.f.).

Estudios sobre dinámica de tráfico (Kerner, B. S., 2009) confirman que el estado de "tráfico congestionado", común en las horas pico, es inherentemente inestable y propenso a estas colisiones en cadena, a menudo exacerbadas por distracciones (uso de celular) y no mantener la distancia de seguridad (CECAITRA, s.f.).

**Resultado Hipótesis 4: Confirmada.**

## 3.5. Hipótesis 5

### 3.5.1. Definición de la hipótesis

Nuestra hipótesis 5 es la siguiente: "**Existe una asociación estadísticamente significativa entre las condiciones meteorológicas y el tipo de siniestro.**"

Esta hipótesis bivariada busca determinar si las CONDICIONES METEOROLÓGICAS (agrupadas en "buenas" y "adversas") influyen en la probabilidad de que un siniestro sea una COLISION CON DOS O MAS VEHICULOS (colisión múltiple) o de otro tipo (siniestro único).

La justificación intuitiva es que las condiciones adversas (lluvia, niebla) reducen la visibilidad y la adherencia, lo que podría incrementar la frecuencia de colisiones más complejas, como las múltiples.

### 3.5.2. Estrategia de abordaje

Para validar esta hipótesis, fue necesario realizar una ingeniería de variables para agrupar las categorías y obtener frecuencias esperadas suficientes para el análisis estadístico:

1. **Variable TIPO DE SINIESTRO:** La transformamos en otra variable, TIPO\_AGRUPADO, con dos valores:
  - "colision multiple": (Solo "COLISION CON DOS O MAS VEHICULOS")
  - "siniestro unico": (Agrupa "SINIESTRO DE UN SOLO VEHICULO", "OBSTACULO FIJO", etc.)
2. **Variable CONDICIONES METEOROLÓGICAS:** Se recodificó en CONDICON\_AGRUPADA con dos valores:
  - "buenas": (Solo "BUENO")
  - "adversas": (Agrupa "LLUVIOSO", "NIEBLA/BRUMA/HUMO")

Se utilizó un **Test Chi-Cuadrado ( $\chi^2$ ) de Independencia** para determinar si existe una asociación estadísticamente significativa entre estas dos nuevas variables categóricas.

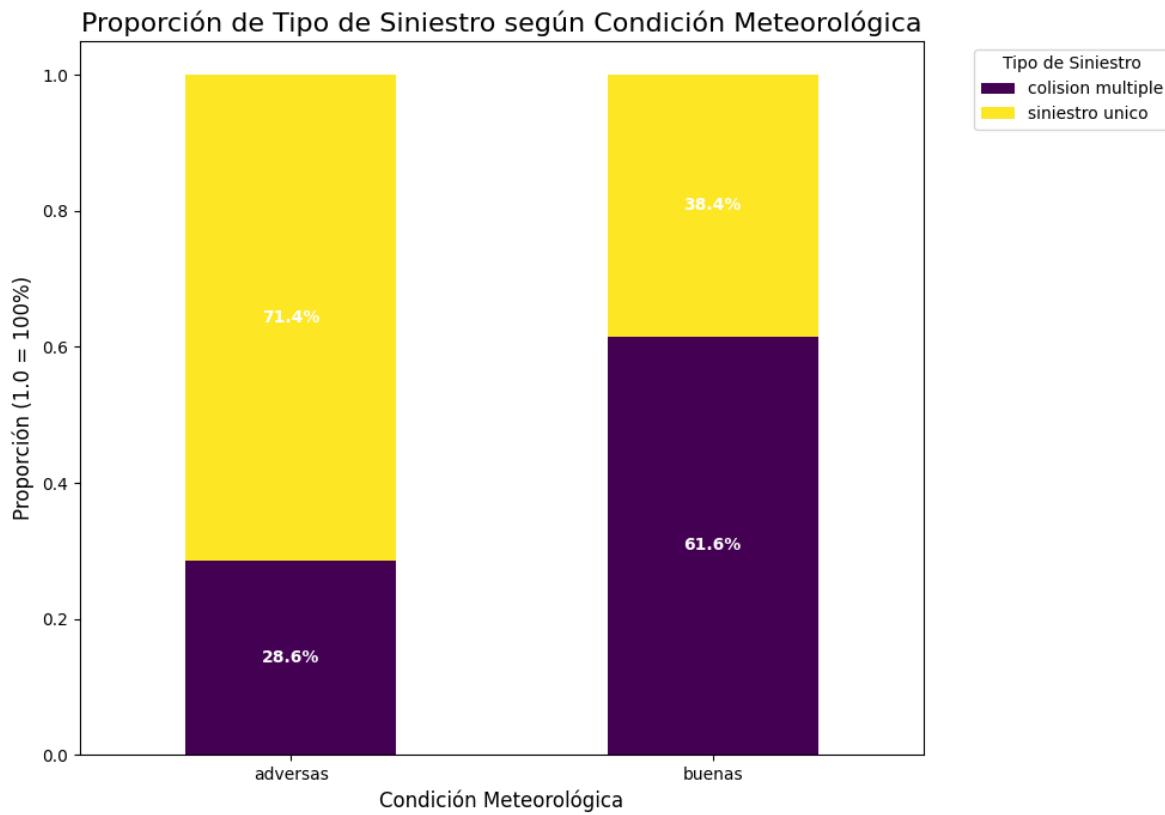
- **Hipótesis Nula ( $H_0$ ):** La condición meteorológica y el tipo de siniestro agrupado son independientes.
- **Hipótesis Alternativa ( $H_1$ ):** Existe una asociación entre la condición meteorológica y el tipo de siniestro agrupado.

### 3.5.3. Resultados obtenidos y discusión

El Test Chi-Cuadrado ( $\chi^2$ ) arrojó un estadístico  $\chi^2 = 4.8869$  y un **p-valor = 0.0271**.

Dado que el p-valor (0.0271) es menor que el umbral de significancia ( $\alpha = 0.05$ ), **se rechaza la hipótesis nula**. Esto confirma que existe una asociación estadísticamente significativa entre la condición meteorológica y el tipo de siniestro.

Para interpretar la naturaleza de esta asociación, se generó un gráfico de barras apiladas al 100% (Figura 7).



*Figura 7: Proporción de Tipo de Siniestro (agrupado) según la Condición Meteorológica.*

El análisis del gráfico arroja un resultado contraintuitivo, que contradice la justificación inicial:

- En condiciones "**buenas**", el **61.6%** de los siniestros son **colisiones múltiples**.
- En condiciones "**adversas**", solo el **28.6%** de los siniestros son colisiones múltiples (la gran mayoría, 71.4%, son "siniestros únicos").

**Discusión:** Los datos muestran que la probabilidad de una colisión múltiple es significativamente mayor durante el buen tiempo. Esto puede explicarse por la psicología del conductor y la velocidad.

Bajo condiciones "buenas", los conductores tienden a circular a mayor velocidad y con menores niveles de alerta. Este exceso de confianza, combinado con un posible mayor volumen de tráfico, incrementa el riesgo de colisiones por alcance que derivan en choques múltiples.

Por el contrario, en condiciones "adversas" (lluvia, niebla), los conductores instintivamente reducen la velocidad y aumentan la distancia de seguridad (Fuente: CECAITRA, s.f.). Esto disminuye la probabilidad de colisiones en cadena, pero aumenta la de "siniestros únicos", como el hidroplaneo (*hydroplaning*) o la pérdida de control que termina en un despiste o choque contra un obstáculo fijo (Fuente: NHTSA, s.f.).

**Resultado Hipótesis 5: Confirmada.**

## 3.6. Hipótesis 6

### 3.6.1. Definición de la hipótesis

Nuestra primer hipótesis multivariada es la siguiente: "**Existen diferencias significativas en la mediana de LESIONADOS entre los distintos TIPOS DE SINIESTRO, pero esta diferencia es particularmente pronunciada en siniestros que involucran al menos una moto en comparación con aquellos que no involucran ninguna.**"

Esta hipótesis analiza la interacción de tres variables:

- LESIONADOS (Numérica, dependiente)
- TIPO DE SINIESTRO (Categórica, independiente)
- MOTO (Categórica binaria, moderadora)

La justificación es la vulnerabilidad inherente del motociclista. Se espera que la diferencia de gravedad (mediana de lesionados) entre un "SINIESTRO DE UN SOLO VEHICULO" y una "COLISION" sea mucho mayor si hay una moto involucrada, debido a la falta de protección física (chasis, airbag) que sí tiene un automóvil (Fuente: *Organización Mundial de la Salud, OMS*).

### 3.6.2. Estrategia de abordaje

Para testear si la presencia de una moto (MOTO) modera la relación entre TIPO DE SINIESTRO y LESIONADOS, se siguió la siguiente estrategia:

1. **Ingeniería de Variables:** Se creó una columna binaria `con_moto`, asignando (1) a los siniestros con `MOTO > 0` y (0) a los siniestros con `MOTO == 0`.
2. **Segmentación del Dataset:** La muestra total se dividió en dos subconjuntos independientes: `df_con_moto` ( $N=167$ ) y `df_sin_moto` ( $N=310$ ).
3. **Prueba Estadística:** Se aplicó el **Test de Kruskal-Wallis** por separado a cada subconjunto. Este test no paramétrico es la alternativa robusta al ANOVA y es adecuado ya que la variable LESIONADOS no sigue una distribución normal (como se infiere del outlier de 14 lesionados detectado previamente).

Se plantearon dos hipótesis nulas:

- **$H_0$  (Test 1 - Con Moto):** Las medianas de LESIONADOS son iguales para todos los TIPO DE SINIESTRO dentro del grupo *con moto*.
- **$H_0$  (Test 2 - Sin Moto):** Las medianas de LESIONADOS son iguales para todos los TIPO DE SINIESTRO dentro del grupo *sin moto*.

La hipótesis  $H_6$  se confirmaría si el Test 1 (Con Moto) resultara significativo ( $p < 0.05$ ) y el Test 2 (Sin Moto) no, o si ambos fueran significativos pero el estadístico H del Test 1 fuera sustancialmente mayor.

### 3.6.3. Resultados obtenidos y discusión

La ejecución de ambas pruebas arrojó los siguientes resultados:

### Test 1: Siniestros CON MOTO (N=167)

- Estadístico H: 2.1531
- **p-valor: 0.7076**

### Test 2: Siniestros SIN MOTO (N=310)

- Estadístico H: 6.1409
- **p-valor: 0.1050**

Ambos p-valores (0.7076 y 0.1050) son mayores al umbral de significancia ( $\alpha = 0.05$ ).

La interpretación de estos resultados es directa:

1. **Test 1 (Con Moto):** Con un p-valor de 0.7076, no se encontró evidencia estadística para rechazar la hipótesis nula. Esto significa que, cuando una moto ya está involucrada, el TIPO DE SINIESTRO (ej. 'COLISION' vs 'OBSTACULO FIJO') no parece tener un efecto significativo en la mediana de lesionados.
2. **Test 2 (Sin Moto):** Con un p-valor de 0.1050, tampoco se encontró evidencia suficiente para afirmar que el TIPO DE SINIESTRO afecte la mediana de lesionados en siniestros donde no participan motos.

Dado que ninguno de los grupos mostró una diferencia significativa, la hipótesis de que el efecto es "particularmente pronunciado" en motos no puede sostenerse. Contrario a la intuición, la presencia de una moto no parece *moderar* la relación entre el tipo de siniestro y su gravedad. Esto podría sugerir que, en el contexto de autopistas, cualquier tipo de siniestro ya posee un nivel de gravedad basal que no se ve alterado significativamente por la *categoría* del siniestro, aunque sí por la *participación* de un vehículo vulnerable.

**Resultado Hipótesis 6: Refutada.**

## 4. Conclusiones

---

Luego de realizar el análisis exploratorio y la contrastación de hipótesis, pudimos obtener una visión más clara sobre cómo, cuándo y en qué condiciones se producen los siniestros viales en las autopistas gestionadas por AUSA durante la segunda mitad del 2022. Aunque el dataset era relativamente pequeño, y además estaba incompleto en algunos meses (por ejemplo, no había registros del mes de julio), la información disponible permitió identificar patrones consistentes y responder a las preguntas planteadas al inicio del trabajo.

En primer lugar, la estructura del dataset estaba bastante limpia, lo cual facilitó el preprocesamiento. Aun así, varias variables esenciales eran cualitativas, y las cuantitativas eran en su mayoría discretas, con distribuciones fuertemente sesgadas hacia la derecha. Esta combinación limitó la cantidad de tests paramétricos que pudimos usar y condicionó especialmente el desarrollo de hipótesis multivariadas. Muchas veces tuvimos que recurrir a versiones no paramétricas (como Kruskal-Wallis) o a recodificaciones categóricas para poder analizar algo significativo.

También notamos que había variables muy relevantes que no estaban presentes, como velocidad del vehículo, uso de casco, uso de cinturón, alcoholemia, distancia de frenado, o el límite de velocidad específico según autopista. Son datos que, si existieran o si la concesionaria los hiciera públicos, permitirían análisis mucho más robustos, especialmente en temas de severidad y causalidad. Incluso un simple archivo TXT que explique a fondo el significado de cada columna habría facilitado muchísimo la interpretación (especialmente para quienes no vienen del dominio vial).

A pesar de estas limitaciones, pudimos llegar a conclusiones claras a partir de las hipótesis planteadas:

- H1 y H3 mostraron que la Autopista 25 de Mayo es la que concentra más siniestros y más lesionados, algo coherente con su tráfico, su extensión y los enlaces complejos que tiene. Esto se observó tanto en los gráficos como en los indicadores estadísticos.
- H2 confirmó que los picos horarios entre las 6 y 8 AM son críticos, y esto coincidió con las observaciones sobre densidad vehicular y flujo de ingreso a la ciudad. La mediana de siniestros por hora (20) permitió cuantificar esta diferencia: las horas pico exceden notablemente ese valor.
- Respecto a H4, pudimos verificar que durante el horario pico matutino las colisiones múltiples se concentran en la banda ascendente, lo cual tiene sentido considerando la existencia de “ondas de tráfico” y frenadas bruscas en la circulación hacia el centro porteño.
- En cuanto a las condiciones meteorológicas (H5), los resultados fueron contraintuitivos: las colisiones múltiples ocurrieron mayormente en buen clima. Esto refuerza la idea de que la “falsa sensación de seguridad” puede ser más peligrosa que las condiciones adversas, donde los conductores suelen manejar con mayor cautela.

- La única hipótesis refutada fue la H6, donde esperábamos encontrar un efecto significativo de la participación de motos sobre la gravedad del siniestro dependiendo del tipo de choque. Con este dataset, no se observaron diferencias estadísticamente significativas.

Es importante remarcar que el tamaño reducido del dataset (menos de 500 registros) fue una limitación real para las hipótesis multivariadas: con más datos, especialmente con varios años consecutivos y más variabilidad en condiciones meteorológicas, podrían descubrirse patrones más sutiles o incluso no lineales.

En cuanto al proceso, algo que nos resultó muy útil fue usar un LLM como Gemini para aclarar conceptos de siniestralidad real, terminología vial y el funcionamiento de las autopistas. Más allá del análisis en sí, aportó contexto experto que no estaba disponible en el dataset y que ayudó a interpretar mejor varios resultados.

#### **4.1. Recomendaciones**

A partir de lo aprendido y de los resultados obtenidos, creemos que hay varias líneas de acción que podrían mejorar tanto la gestión vial como futuros análisis:

Por un lado, sería ideal que AUSA o los organismos públicos publiquen dataset más completos, con variables cuantitativas clave (velocidad, climatología exacta, distancia entre vehículos, uso de casco/cinturón, alcoholemia, tipo de maniobra, etc.). Esto permitiría pasar de análisis descriptivos a modelos predictivos o causales con mayor precisión.

En relación con la operación vial, nuestros resultados muestran que las horas pico, especialmente en la banda ascendente, son momentos críticos. Intervenciones como señalización dinámica, alertas de frenado, o estrategias para reducir embudos y cambios de carril abruptos podrían ayudar a disminuir colisiones múltiples. Algo parecido aplica para la AU 25 de Mayo, donde los enlaces complicados explican buena parte de su siniestralidad: rediseñar señalética y rampas podría reducir los puntos conflictivos.

Finalmente, aunque las condiciones climáticas no parecieron tener un impacto fuerte en este dataset, esto se debe más al año extremadamente seco que fue 2022 en Buenos Aires que a una inexistencia de efecto. Sería importante monitorear este punto con años más lluviosos, porque la literatura internacional sí demuestra una relación entre clima adverso y severidad del siniestro.

En conjunto, el trabajo nos permitió comprender mejor los patrones de siniestros en autopistas urbanas y detectar las limitaciones y oportunidades reales para el análisis. Una base de datos mejorada y más extensa sin duda permitiría avanzar hacia modelos más completos que contribuyan a mejorar la seguridad vial en la Ciudad de Buenos Aires.

# Referencias

---

**Autopistas Urbanas S.A.** (2023). *En la autopista Illia ya funciona el peaje sin barrera.* <https://www.ausa.com.ar/sections/novedades.php?idNov=2023-1>

**Autopistas Urbanas S.A.** (2024). *Desde el 1 de mayo, los usuarios que aún no cuenten con TelePASE tendrán una nueva modalidad de pago online.* <https://www.ausa.com.ar/sections/novedades.php?idNov=2024-5>

**Autopistas Urbanas S.A.** (s.f.). *Autopistas.* Recuperado el 16 de noviembre de 2025, de <https://www.ausa.com.ar/Autopistas.html>

**Autopista 25 de Mayo.** (s.f.). En **Wikipedia**. Recuperado el 16 de noviembre de 2025, de [https://es.wikipedia.org/wiki/Autopista\\_25\\_de\\_Mayo](https://es.wikipedia.org/wiki/Autopista_25_de_Mayo)

**Canal 26.** (2025, 30 de octubre). *Se terminó el pago en efectivo en los peajes: dejará de existir y cambia el tránsito de millones de argentinos.* <https://www.canal26.com/general/2025/10/30/se-termino-el-pago-en-efectivo-en-los-peajes-dejara-de-existir-y-cambia-el-transito-de-millones-de-argentinos/>

¿En qué año se inauguró la Autopista 25 de Mayo? (s.f.). **Billiken**. Recuperado el 16 de noviembre de 2025, de <https://billiken.lat/mi-pais/en-que-ano-se-inauguro-la-autopista-25-de-mayo>

**Ente Único Regulador de los Servicios Públicos de la Ciudad de Buenos Aires.** (s.f.). *Autopistas urbanas.* Recuperado el 16 de noviembre de 2025, de <https://entedelaciudad.gov.ar/servicios-publicos/autopistas-urbanas/>

**Gobierno de la Ciudad de Buenos Aires.** (s.f.). *Autopistas Urbanas S.A.* Recuperado el 16 de noviembre de 2025, de <https://buenosaires.gob.ar/infraestructura/autopistas-urbanas-sa>

**Gobierno de la Ciudad de Buenos Aires.** (s.f.). *Metrobus AU 25 de Mayo.* Recuperado el 16 de noviembre de 2025, de <https://buenosaires.gob.ar/movilidad/metrobus/metrobus-au-25-de-mayo>

**Agencia Nacional de Seguridad Vial.** (2023). *Glosario de términos de seguridad vial.*

**InfoMeteoBA.** (2023, febrero). *Resumen del año 2022 en Buenos Aires.* [Weblog]. <https://infometeoba.blogspot.com/2023/02/resumen-del-ano-2022-en-buenos-aires.html>

**La Nación.** (2020, 11 de febrero). *Tránsito: cuáles son los tramos más peligrosos de las autopistas.* <https://www.lanacion.com.ar/buenos-aires/transito-cuales-son-tramos-mas-peligrosos-autopistas-nid2332428/>

**La Nación.** (2024, 12 de mayo). *En Boedo y San Telmo, los dos puntos con mayor cantidad de siniestros viales a lo largo de las...* <https://www.lanacion.com.ar/buenos-aires/en-boedo-y->

[san-telmo-los-dos-puntos-con-mayor-cantidad-de-siniestros-viales-a-lo-largo-de-las-nid12052024/](https://www.santelmo.org.ar/nid12052024/)

*Rush hour.* (s.f.). En **Wikipedia**. Recuperado el 16 de noviembre de 2025, de [https://en.wikipedia.org/wiki/Rush\\_hour](https://en.wikipedia.org/wiki/Rush_hour)

**Tatarletti, P., & López, D.** (2016). *Evolución de los siniestros viales sobre las autopistas de la Ciudad de Buenos Aires* [Presentación de paper]. XVIII Congreso Argentino de Vialidad y Tránsito, Buenos Aires, Argentina. <http://congresodevialidad.org.ar/congreso2016/TRA/TRA-218.pdf>

**CECAITRA (Centro de Experimentación y Seguridad Vial, Argentina).** (s.f.). Recuperado de <https://www.cecaitra.org.ar/>

**National Highway Traffic Safety Administration (NHTSA).** (s.f.). *Driver Safety: Driving in Rain*. Recuperado el 16 de noviembre de 2025, de [https://www.nhtsa.gov/\\_driver-safety/driving-rain](https://www.nhtsa.gov/_driver-safety/driving-rain)

**National Highway Traffic Safety Administration (NHTSA).** (s.f.). *Vehicle Safety, Statistics*. Recuperado de <https://www.nhtsa.gov/>

**Kerner, B. S.** (2009). *Introduction to Modern Traffic Flow Theory and Control*. Springer.

**Organización Mundial de la Salud (OMS).** (s.f.). *Traumatismos causados por el tránsito: Motociclistas.* Recuperado de <https://www.who.int/es/news-room/factsheets/detail/road-traffic-injuries>