



# Uniempresarial

FUNDACIÓN UNIVERSITARIA EMPRESARIAL  
DE LA CÁMARA DE COMERCIO DE BOGOTÁ

Siniestralidad vial en motociclistas: Análisis de datos abiertos para la localidad de Kennedy,  
2015-2020.

Opción de grado: Proyecto de investigación

Cindy Milena Prieto Reina

Ingeniería de Software – Promoción 6 JF

Fundación Universitaria Empresarial De La Cámara De Comercio De Bogotá

Bogotá, Colombia 2022

Siniestralidad vial en motociclistas: Análisis de datos abiertos para la localidad de  
Kennedy, 2015-2020

Opción de grado: Proyecto de investigación

Cindy Milena Prieto Reina

Profesor: Adán Beltrán Gómez

Ingeniería de Software – Promoción 6 JF

Fundación Universitaria Empresarial De La Cámara De Comercio De Bogotá

Bogotá, Colombia 2022

## **Resumen**

La Organización Mundial de la Salud hace un llamado no solo a las entidades gubernamentales sino también a cada individuo que hace parte de la sociedad, a tomar un tiempo de reflexión sobre la cantidad de pérdidas de vidas humanas en los accidentes viales, por este motivo la presente investigación tiene como fin conocer por medio de una de las ramas de la ingeniería de software que actualmente está tomando fuerza que es la analítica de datos, cuáles son las causas más frecuentes de muerte en la siniestralidad vial de la localidad de Kennedy y conocer cómo es el comportamiento de las edades de los motociclistas varones involucrados en estos trágicos eventos.

Palabras claves: análisis de datos, siniestralidad vial, estadística descriptiva, frecuencia.

## **Abstract**

The World Health Organization makes a call not only to government entities but also to each individual who is part of society, to take time to be aware of the amount of loss of human lives in road accidents, for this reason this research aims to know through one of the branches of software engineering that is currently gaining strength that is data analytics, identify what are the most common causes of death in road accidents where there was death in the town of Kennedy and know how is the behavior of the age of the people involved in these tragic events.

Key words: data analysis, road accident, descriptive statistics, frequency.

## CONTENIDO

### Contenido

1. Introducción	9
1.1. Formulación del problema	10
1.2. Pregunta de investigación	12
2. Justificación	12
3. Objetivos	13
3.1. Objetivo General	13
3.2. Objetivos Específicos	13
4. Marco referencial	13
4.1. Estado del arte	14
4.2. Marco Teórico	15
4.2.1. Análisis de datos	15
4.3. Marco Conceptual	20
4.3.1. Siniestralidad vial	21
4.4. Marco Legal	24
5. Propuesta metodológica	26
6. Resultados	30
6.1. Exploración de los datos	30
6.2. Exploración y selección de variables	31
6.2.1. Dataset siniestros	31
6.2.2. Dataset Actor Vial	33
6.2.3. Dataset Vehículos	35
6.2.4. Dataset Hipótesis	36
6.2.5. Dataset Diccionario	37
6.3. Preparación de los datos	38
6.4. Procesamiento y limpieza de los datos	39
6.5. Análisis de los datos	42
6.5.1. Análisis descriptivo sobre las edades en la localidad de Kennedy	43
6.5.2. Análisis descriptivo de las causas de mortalidad en Kennedy	46

6.5.3. Análisis descriptivo de las edades junto con las principales causas de mortalidad en Kennedy	50
7. Conclusiones y Recomendaciones	53
8. Referencia Bibliográfica	56

## Índice de Figuras

Figura 1 : Rango intercuartílico	18
Figura 2 : Paquetes de librerías usadas en R	38
Figura 3 : Cargue de los datasets de Excel a R	39
Figura 4 : Código para la unión de datasets en R Studio	39
Figura 5 : Código que filtra la base de datos en R Studio	40
Figura 6 : Nivel de gravedad por localidad	41
Figura 7 : Frecuencia de la cantidad de muertos por localidad	42
Figura 8 : Frecuencia absoluta de edades en motociclistas muertos en la localidad de Kennedy	44
Figura 9 : Frecuencia relativa de edades en motociclistas muertos en la localidad de Kennedy	44
Figura 10 : Distribución de la variable edad en motociclistas muertos en la localidad de Kennedy	46
Figura 11 : Causas de mortalidad en la localidad de Kennedy	48
Figura 12 : Causas más frecuentes de mortalidad en la localidad de Kennedy	49

## Índice de Tablas

Tabla 1 : Metodología por cada objetivo específico	27
Tabla 2 : Información sobre el dataset de siniestralidad vial	30
Tabla 3 : Ejemplo del dataset de Siniestros	33
Tabla 4 : Ejemplo del dataset de Actor Vial	34
Tabla 5 : Ejemplo del dataset de Vehículos	36
Tabla 6 : Ejemplo del dataset de Hipótesis	37
Tabla 7 : Ejemplo del dataset de Diccionario	38
Tabla 8 : Medidas de tendencia central de las edades en motociclistas muertos en la localidad de Kennedy	43
Tabla 9 : Medidas de dispersión de las edades en motociclistas muertos en la localidad de Kennedy	45
Tabla 10 : Frecuencias de las causas de accidentalidad vial en los motociclistas muertos en la localidad de Kennedy	47
Tabla 11 : Causas más frecuentes de mortalidad en la localidad de Kennedy	49
Tabla 12 : Medidas de tendencia central de las edades y la causa de exceso de velocidad	50
Tabla 13 : Medidas de tendencia central de las edades y la causa transitar entre vehículos	50
Tabla 14 : Medidas de tendencia central de las edades y la causa cruzar sin observar (peatón)	51
Tabla 15 : Medidas de tendencia central de las edades y la causa de semáforo en rojo	51
Tabla 16 : Medidas de tendencia central de las edades y la causa de desobedecer señales	52



## 1. Introducción

Según la tasa de mortalidad realizada por el DANE en el 2007 (Dane, 2007) el promedio de vida para los años 2015 al 2020 del hombre es hasta los 73 años y de la mujer hasta los 79 años, pero estas estadísticas no son posibles debido a factores externos como lo fue el COVID en el 2020, y el incremento de accidentes viales, los cuales de acuerdo con las estadísticas del 2014 realizadas por la Organización Mundial de la Salud (O. M. de la Salud, 2014), son la séptima causa de muerte a nivel global, y en el país representa la segunda causa de mortalidad con un 21%.

Para el Ministerio de Salud y Protección Social la seguridad vial hace parte de un tema de salud pública dado que en un accidente vial no solamente está involucrada la persona que ocasiona el accidente si no también la persona afectada por el mismo, es decir que en el peor de los escenarios son varias las personas involucradas en un siniestro vial. El boletín de prensa emitido por esta entidad gubernamental en el 2018 (MinisteriodeSalud, 2018), menciona la posibilidad de poder prevenir más de siete mil muertes en el año por accidentes viales.

El ministro Juan Pablo Uribe en este mismo boletín menciona que los accidentes viales son la principal causa de muerte de los jóvenes entre los 15 y los 19 años, motivo por el cual se deben crear estrategias de movilidad para prevenir estos fallecimientos puesto que las estadísticas del año 2016 en el país evidenciaron la muerte de 7.355 personas entre ciclistas, motociclistas y peatones.

A raíz de lo anterior el presente estudio tiene como objetivo utilizar los datos reales obtenidos por la Secretaría Distrital de Movilidad en el momento de ocurrencia de siniestros viales en la capital, con la intención de poder segmentar la ciudad para conocer el comportamiento de los fallecimientos en los motociclistas varones, identificando las causas y las

edades más frecuentes con el fin de poder brindar algunas recomendaciones a las entidades encargadas las cuales lleven a la disminución de muertes en la localidad escogida.

Por último, se realiza una descripción de forma panorámica sobre la estructura que tiene la presente investigación: en la introducción se aborda de forma breve la problemática a tratar y la formulación del problema, seguido de la justificación donde se explica el porqué de la investigación, se continúa con la sección de objetivos donde se encuentra el general y los específicos, posteriormente se encuentran los marcos referenciales los cuales son: el estado del arte el cual da un contexto del problema, el marco teórico que describe de forma abreviada las técnicas y el método utilizado en la investigación, el marco conceptual donde están los principales conceptos de siniestralidad vial y el marco legal en el cual están las leyes de la problemática que se trabaja.

Posteriormente está la sección de propuesta metodológica que contiene la forma como se van a trabajar los datos de la investigación, continuo con los resultados obtenidos al realizar la manipulación de la data y la discusión de la información encontrada, ya para terminar la sección de conclusiones y recomendaciones exponiendo la importancia de la técnica estadística y posibles estrategias de solución al problema.

### **1.1. Formulación del problema**

Debido al aumento en la población de Bogotá durante los últimos años y al cambio de reglamentación vehicular en la ciudad por parte de la Alcandía, hizo que una parte los ciudadanos tomaran la decisión de utilizar otro medio de transporte generando que la compra de motocicletas tuviera un aumento del 37,1% con respecto del año pasado 2021 (Semana, 2022).

Como consecuencia de esto se presentó un alza en la cantidad de accidentes viales donde los mayores protagonistas son los motociclistas, los cuales para año 2021 la capital se vio afectada con un total de 497 fallecidos un aumento con respecto al 2020 de 24.5% (Semana, 2022).

Adicional, Diego Roselli de la Pontificia Universidad Javeriana realizó un artículo de reflexión en el año 2014 donde destaca cifras desalentadoras en cuanto a siniestros viales, menciona que hubo 552 motociclistas fallecidos, de los cuales 488 fueron hombres cuyas edades más frecuentes fueron entre los 20 y 24 años (Roselli, 2017).

Teniendo en cuenta lo anterior en lo que va del año 2022, ocho años después del artículo de reflexión las cifras no han mejorado como lo muestra el periódico digital Infobae los fallecimientos en motociclistas están representados en un 85% por hombres cuyas edades oscilan entre los 20 y 35 años (Infobae, 2022), estas cifras son confirmadas por la revista digital Publimotos cuyos datos indican que durante enero y febrero del presente año ha habido 668 motociclistas muertos, donde 573 fueron hombres, con un recuento de 140 entre las edades de 20 a 25 y 117 contaban con la edad de 25 a 30 (Publimotos, 2022).

A lo largo de esta publicación se menciona que las causas que generan los accidentes viales son en un 45% el exceso de velocidad, el 42% no obedecer las señales de tránsito, un 4% la embriaguez y un 3% ir en contra vía (Publimotos, 2022).

De acuerdo con lo anterior, la presente investigación busca aplicar la estadística descriptiva en el dataset de la Secretaría Nacional de Movilidad cuyos datos fueron recolectados durante el periodo comprendido entre el 2015 y el 2020, con la finalidad de hallar cuál fue la localidad con mayor índice de muerte en siniestros viales de la ciudad de Bogotá, donde el resultado obtenido fue la localidad de Kennedy y adicional se busca identificar cuáles son las

edades y las causas más frecuentes de mortalidad en los varones motociclistas en localidad escogida.

## **1.2 Pregunta de investigación**

¿Cuáles fueron las edades y las causas de las muertes en los siniestros viales de los motociclistas hombres en la localidad de Kennedy durante el periodo 2015 al 2020?

## **2. Justificación**

Como se planteó anteriormente la accidentalidad vial es considerada un tema de salud pública, no solamente a nivel mundial si no también nacional, lo cual ha llevado a que los diferentes entes gubernamentales tomen medidas con respecto del nivel de fallecimientos y heridos en las vías del país.

Estas medidas y/o decisiones que toman los directivos encargados de la movilidad como son el Ministerio de Transporte, el Instituto Nacional de Vías (INVIAS), la Agencia Nacional de Infraestructura (ANI), la Agencia de Seguridad Vial (ANSV), lo hacen basados en cifras, las cuales van apoyadas con los resultados obtenidos por el departamento de sistemas, en donde un ingeniero de datos o un científico de datos, toman la información recolectada, la procesan y pueden mostrar sus hallazgos.

Al realizar análisis de datos sobre la data recogida una vez ocurran los siniestros viales, esta información es útil para conocer cómo se pueden prevenir estos hechos porque en algunos siniestros viales no solo ahí pérdidas materiales si no también vidas humanas.

Algunas de las decisiones que han tomado los encargados de la seguridad vial no solo a nivel nacional, sino también en la capital son leyes como se evidencia en el marco legal y campañas preventivas y manuales (Lota et al., 2021).

Esta investigación es útil para entidades gubernamentales y públicas interesadas en conocer cuáles son las causas que presentan más frecuencia de mortalidad en la localidad de Kennedy, con el fin de tomar medidas preventivas en las vías de tránsito, de igual modo, conocer las edades con más frecuencia para llevar capacitaciones adecuadas con el objetivo de minimizar la mortalidad en la localidad.

### **3. Objetivos**

#### **3.1 Objetivo General**

Analizar las edades y las causas de las muertes en los siniestros viales de los motociclistas hombres en la localidad de Kennedy durante el periodo 2015 al 2020, utilizando los datos abiertos suministrados por la Secretaría Distrital de Movilidad.

#### **3.2 Objetivos Específicos**

Describir los tres tipos principales de la estadística que se utilizan en el análisis de datos para darle el fundamento teórico al presente trabajo de grado.

Manipular la base de datos suministrada por la Secretaría Distrital de Movilidad para hacer uso de la estadística descriptiva.

Determinar las edades y las causas de las muertes más frecuente en los motociclistas de la localidad de Kennedy.

### **4. Marco referencial**

Esta sección cuenta con cuatro numerales el 4.1 es el estado del arte o marco contextual en el cual se describe la situación en la que se encuentra la accidentalidad vial de acuerdo con la Agencia Nacional de Seguridad Vial, en el numeral 4.2 se encuentra el marco teórico que contienen las teorías del trabajo, en el numeral 4.3 contiene el marco conceptual donde están los

conceptos más importantes del proyecto y el numeral 4.4 se encuentra el marco legal en el cual se escriben las leyes vigentes para evitar la siniestralidad vial.

#### **4.1. Estado del arte**

Como se mencionó con anterioridad es de suma importancia conocer las estadísticas sobre cómo va la accidentalidad vial en el país de acuerdo con las medidas que se han ido implementando por medio del establecimiento de leyes nacionales y departamentales, por ese motivo la agencia de seguridad vial en su página web, en la sección observación, cifras del año en curso aparecen las estadísticas de las personas lesionadas y muertas a nivel nacional, la página web es interactiva ya que cuenta con la opción de poder seleccionar un departamento, el municipio y las zonas de interés, mostrando datos estadísticos del año en curso comparándolo con el pasado (Agencia de Seguridad Vial, 2022).

Adicional a esto en la página oficial del ministerio de transporte, en la sección de publicaciones aparece el balance general de las personas que fallecieron en el año 2021, con un total de 7.270, de los cuales 4.312 son motociclistas, el director de la Agencia Nacional de Seguridad Vial hace un llamado a los colombianos que respeten las señales de tránsito, los límites de velocidad y a no manejar bajo efectos de bebidas alcohólicas (Ministerio de Transporte, 2022).

Juntamente con estas dos entidades, se une la entidad llamada Así Vamos en Salud, la cual tiene como propósito informar y mantener actualizada a la comunidad sobre el seguimiento que realiza a los diferentes entes gubernamentales en cuanto a normas, decretos e índices en salud y protección social, trabajando de la mano con diferentes entidades de salud, dándolo a conocer por medio de estadísticas y artículos. En su página web se encuentran cifras y recomendaciones sobre la accidentalidad vial a nivel nacional de forma dinámica, es decir una

vez seleccionado el departamento y el periodo da conocer los datos de forma dinámica (A. vamos Salud, 2022).

De acuerdo con esto, las fuentes de información oficiales que se consultaron no detallan con especificidad las localidades de Bogotá, como consecuencia de esto la presente investigación tiene como fin segmentar la ciudad de Bogotá para conocer el comportamiento de las causas de muertes y las edades de las personas que sufrieron siniestralidad en la localidad con índices más altos de mortalidad.

## **4.2. Marco Teórico**

En esta sección se describen las bases teóricas utilizadas para realizar la presente investigación sobre análisis de datos.

### **4.2.1 Análisis de datos**

El análisis de datos es la realización de un estudio profundo de la información que se tiene con el fin de obtener resultados y conclusiones las cuales permiten a los directivos de las empresas u organizaciones tomar decisiones (Westreicher, 2020b).

De acuerdo con esta definición para realizar un estudio profundo se necesita conocer el significado de análisis que es examinar de forma minuciosa y cuidadosa un objeto de tal forma que permita reconocer las partes de un todo (Ucha, 2022), por otro lado se tiene que es una acción la cual permite descomponer una situación en la suma de todos sus elementos, para entender como cada uno de estos se interrelaciona (Ramos, 2014), también se tiene que es un estudio de forma profunda en donde se identifican las bases, los fundamentos y cómo fue su surgimiento o creación de un objeto, situación u objeto de tal forma que permita sacar conclusiones de ese desglose de información obtenida (Martinez, 2021).

De acuerdo con las definiciones anteriores se puede percibir que el significado de la palabra análisis es la descripción de cada una de las partes de un objeto, ahora existen varios tipos de análisis de datos, algunos varían entre tres y cuatro, pero aproximadamente existen trece tipos: análisis descriptivo, análisis exploratorio, análisis de diagnóstico, análisis inferencial, análisis predictivo, análisis prescriptivo, análisis mecanicista, análisis de clústeres, análisis de cohortes, análisis de textos, análisis factorial, análisis causal y análisis de conjunto (Narvaez, 2022).

En este proyecto de investigación se describirán las tres más importantes que son el análisis descriptivo, el análisis prescriptivo y el análisis predictivo, en donde para el desarrollo de los objetivos se realizará el análisis descriptivo.

#### **4.2.1.1 Análisis descriptivo**

Es cualquier característica general de una población la cual es posible medir con valores las cuales se llaman variables cuantitativas o se puede describir usando diferentes modalidades llamadas variables cualitativas, algunos ejemplos de estas son la edad de un grupo de persona, su género o su estado civil (Posada Hernández, 2016).

Se busca describir como es el comportamiento de las tendencias en los datos de tal forma que permitan observar situaciones que produzcan nuevos hechos, su proceso es: recopilación de los datos, organización, tabulación y descripción del resultado, adicional una vez obtenido ese análisis descriptivo puede elaborarse un análisis cuantitativo (QuestionPro, 2019).

Se infiere que el análisis descriptivo es para resumir datos, existen cuatro tipos de estadísticos (Gonzalez, 2001) que se pueden utilizar que son:

- De centralización, los cuales indican el promedio de los datos y están conformados por la media, mediana y moda (Gonzalez, 2001) acompañadas de las tablas de

- frecuencia las cuales permiten llevar el registro de la cantidad de individuos que perteneces a una misma categoría o clase (Rustom, 2015) y que pueden ser representados mediante gráficos permitiendo ver de forma ilustrada los datos que se encuentran en la tabla de frecuencias, dependiendo del tipo de variable a representar (Rustom, 2015).
- De dispersión, indican que tan separados o unidos están los datos del valor central, si se encuentran muy separados del valor central significa que están muy dispersos, están conformados por: la varianza, la desviación típica y el coeficiente de variación (Gonzalez, 2001), se utilizan medidas en forma de resúmenes con el fin de que en la muestra se pueda obtener la posición de los datos mediante las medidas de posición o concentración y/o también llamadas medidas de dispersión. (Rustom, 2015).
  - De forma, están divididos en dos tipos, una es la asimetría las cuales indican si la distribución de los datos es simétrica o asimétrica (positiva o negativa), y la curtosis la cual indica como es el grado de distribución respecto a la normal, existen tres formas que son platicúrtica (la curtosis es menos a cero), mesocúrtica (la curtosis es igual a cero) y leptocúrtica (la curtosis es mayor a cero) (Gonzalez, 2001).
    - De posición, es la forma en la que se dividen los datos de tal forma que queden distribuidos igualmente, están los cuartiles (dividido en 4 secciones iguales), los deciles (dividido en 10 secciones iguales) y los percentiles (dividido en 100 secciones iguales) (Gonzalez, 2001).

Ampliando en este último estadístico, llamado rango intercuartílico se tiene que es una medida de dispersión la cual se obtiene de un conjunto de datos, esta medida expresa la distancia

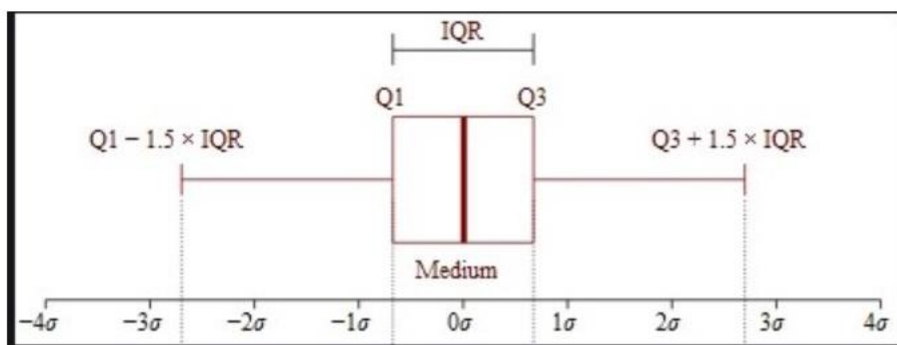
o la diferencia entre el tercer y primer cuartil, es decir, al graficar los datos por medio del diagrama de caja (este diagrama usa como medida central la mediana) la cual muestra la información de los cuartiles, al obtener la diferencia entre el primer y penúltimo cuartil sus resultados mostrarán que tan dispersos están los datos con respecto a la media (Paula Rodó, 2021).

Este rango también es una estimación estadística la cual permite ver la distribución de un conjunto de datos, donde al utilizarla se podrían eliminar los outlier o datos extremadamente alejados de la mediana (estadístico insensible a posibles irregularidades de valores extremos), se puede apreciar mejor los datos de dispersión utilizando los diagramas de caja (Serra, 2015)

El rango intercuartílico fue desarrollado por John Tukey, el cual fue el pionero en el uso del análisis de datos de exploración, este rango es una medida de variabilidad.

*Figura 1*

*Rango intercuartílico*



Nota. Distribución de los datos, usando

Fuente: Tomado de la página web de Oracle (Oracle, 2022)

Una vez dividido el conjunto de datos en cuatro partes iguales como lo muestra la figura 1, se realiza la diferencia entre el cuartil tercero y el primero (IQR) la cual es la medida de

variabilidad, que es la distancia entre el cuartil 1 o Q1 y el cuartil 3 o Q3, que son los bordes de la caja, también se aprecia en la figura1 que la mediana se encuentra en el centro, para conocer qué datos son atípicos se aplica el método de Tukey el cual indica que se debe multiplicar los datos por del primer cuartil con 1.5 ( $Q_1 - 1.5 * IQR$ ) o también los datos del tercer cuartil con 1.5 ( $Q_3 + 1.5 * IQR$ ) los cuales son los llamados bigotes de la caja, si el resultado de realizar estas dos operaciones queda fuera de los bigotes se han hallado los datos atípicos (Oracle, 2022)

#### 4.2.1.2 Análisis predictivo

El análisis predictivo es el resultado de aplicar análisis estadístico a los datos, en donde a partir de la información histórica se puede predecir cómo va a ser el comportamiento o las tendencias en el futuro de dichos datos (Ortega, 2022). Los pasos que se deben seguir para realizar análisis predictivo son:

- Tener claro que proyecto se va a realizar las ramas más frecuentes que usan análisis predictivos son la de seguros, finanzas, marketing y se está incluyendo el sector salud.
- Se deben recolectar los datos.
- Una vez se tienen los datos se debe analizar la información que se obtuvo.
- Se aplican técnicas para realizar análisis estadístico.
- El siguiente paso es escoger un modelo predictivo, es decir, se utilizan diferentes modelos de regresiones estadísticas, con el fin de tomar de los datos una variable respuesta y una o más variables explicativas, utilizando la técnica matemática de los mínimos cuadrados (Pereira González, 2010)
- Una vez escogido y probado el modelo se pasa a la etapa de producción o utilización del modelo por el área encargada

- Por último, se debe monitorear si el modelo cumple con los objetivos del proyecto.

La ventaja de realizar análisis predictivo es que se pueden descubrir cuáles son las tendencias en cualquiera de las ramas mencionadas con el fin de obtener mejores ingresos y mejorar el servicio al cliente y una desventaja de este análisis es su dependencia a los datos que se obtuvieron y su la información es de calidad (Ortega, 2022).

#### **4.2.1.3 Análisis prescriptivo**

El análisis prescriptivo es útil en el momento que se hayan realizado los dos tipos de análisis anteriores porque el objetivo de este análisis es hacer una combinación entre los datos cuantitativos y cualitativos para brindar mejores herramientas en la toma de decisiones de los directivos de las compañías (School, 2022).

Este tipo de análisis tiene su fundamento en técnicas estadísticas y matemáticas, investigación de operaciones y en el análisis predictivo, dado que busca encontrar cuáles son las limitaciones que tiene cada una de las variables que se predijeron en los datos con el fin de escoger el mejor resultado para poderlo ejecutar como por ejemplo en la parte productiva de las organizaciones para optimizar procesos y reducir costos (Velázquez, 2022)

Este análisis ayuda a evaluar las decisiones que se pueden tomar en los diferentes escenarios futuros (Tatic.net, 2021), de acuerdo con todo lo anterior a partir de los resultados obtenidos en el análisis prescriptivo lo encargados de tomar decisiones tienen diferentes escenarios para tomar decisiones estratégicas que beneficien a la compañía.

### **4.3 Marco Conceptual**

En esta sección se describen los conceptos utilizados a lo largo de la presente investigación sobre siniestralidad vial.

### 4.3.1 Siniestralidad vial

Para definir siniestralidad vial primero se tiene que conocer que es un siniestro, el cual se define como un hecho que no se puede predecir cuándo pasará en el tiempo el cual ocasiona daños a las organizaciones o las personas (Galán, 2021)

El artículo 1072 del código de comercio de Colombia lo define como la ejecución de un riesgo asegurado (República, 1996), ampliando la definición de riesgo es la probabilidad que un peligro se convierta en un daño (Concepto, 2020), la psicóloga Marie Drottz lo define como la probabilidad de padecer algún daño en específico (Martinez, 2022), la Organización Mundial de la Salud lo define como aquella situación en la cual este presente un factor de riesgo en el que una persona tenga alta probabilidad de padecer daños (Martinez, 2022).

Definiendo la siniestralidad vial es un evento impredecible asociado con el azar el cual se relaciona con circunstancias dolorosas como pérdida de la vida o lesiones en las personas involucradas (Distrito, 2019), se tiene también como un suceso de tránsito en el cual participa al menos un vehículo en movimiento ocasionando en al menos una persona la muerte o heridas (Autoridad Nacional de Licencias Ambientales (ANLA), 2013), adicional a esto el magistrado Miguel López lo define como las consecuencias de una situación en la que un vehículo quede de forma anormal dentro o fuera de las vías produciendo en las personas lesiones o muertes y generando daños a terceros (Carreras, 2015).

Teniendo como base las anteriores definiciones, se entiende por siniestro vial que es un conjunto probabilístico de eventos el cual es ocasionado por al menos un vehículo en donde las personas involucradas resulten lesionadas o muertas.

Es necesario tener clara la diferencia entre un siniestro y un accidente vial, ya que generalmente se confunden estos dos términos, la palabra accidente significa que es un evento

que sucede de manera inesperada, es involuntario (Pérez, 2015), por consiguiente un accidente vial es algo que no se espera por ejemplo que caiga un árbol sobre un carro (Seguros, 2022), lo que se diferencia del siniestro es que es un evento que pudo haber sido evitado, es decir, que son errores humanos en el momento de estar en la vía, como por ejemplo dos automóviles chocan porque uno iba a una velocidad mayor y no alcanzo a frenar.

Por otro lado, para que ocurra un siniestro vial debe existir todo un entorno el cual es la vía, el vehículo y la persona, seguido de esto existen tres pasos para que se lleve a cabo la conducción (Norza-Céspedes et al., 2014):

1. Técnicas relacionadas con el tipo de vehículo.
2. Leyes de tránsito las cuales normalizan la circulación vial.
3. La formación de la persona donde se evidencian sus aptitudes y actitudes

De igual forma es de vital importancia comprender que las personas tienen diferentes roles como lo es ser el conductor, un peatón, el pasajero o el acompañante, ya que un accidente es causa de un error humano (Norza-Céspedes et al., 2014).

Adicional a esto el estilo de vida y los hábitos influyen en la conducción ya que desobedecer las normas viales, un comportamiento imprudente, el stress de la cotidianidad, no tener un sueño adecuado hacen que la agresividad de las personas aumente generando alteraciones psicofísicas las cuales hacen que al conducir se pierda la precisión y ocasionen accidentes de tránsito (Norza-Céspedes et al., 2014).

Paralelamente con el factor individual humano, el factor colectivo y de las entidades gubernamentales tiene un papel importante ya que, por medio de campañas de concientización a la sociedad y la obligatoriedad de leyes de tránsito.

#### **4.3.1.1 Actor vial**

Para la Corporación Fondo de Prevención Vial son todas las personas que tienen una función específica en la vía, es decir, pueden ser transeúntes, peatones, conductores y pasajeros de vehículos automotores, ciclistas y motociclistas (Vial, 2011).

La red de movilidad empresarial complementa la anterior definición en que las personas cumplen un rol en la vía en el momento de desplazarse de un lugar a otro (RedEmpresarialdeSeguridadVial, n.d.)

Ampliando estos roles, se tiene por peatón aquella persona que camina en la zona destinada para personas, es decir, los andenes o aceras, los puentes peatonales y las cebras que están dibujadas en el suelo de la zona destinada a los vehículos, los motociclistas deben respetar las zonas destinadas para los peatones y los ciclistas por último, los conductores se clasifican dependiendo de su licencia de conducción la cual va de segundo a quinto grado (Publimotos, 2018).

Con base en estos conceptos se puede afirmar que un actor vial son las personas que hacen uso de algún vehículo para desplazarse de un lugar a otro y que cada actor vial no solo tiene un rol si no también tiene unas responsabilidades las cuales hacen que su cumplimiento evite siniestros viales.

#### **4.3.1.2 Vehículo**

La ley 769 del 2022 del código nacional de tránsito lo define como toda máquina y/o aparato que este sobre ruedas, el cual sirva para transportar cosas, animales o personas por vía terrestre de un punto A hacia un punto B (MinisteriodeTransporte, 2002).

Se tiene también que es cualquier máquina que permite el desplazamiento de un sitio a otro, permitiendo transportar cualquier tipo de objeto, planta, animal o persona (Pérez, 2022)

Ampliando un poco este concepto se tiene que hay diferentes tipos de vehículos que permiten el transporte dependiendo el medio, es decir, si es por aire, agua, el espacio o por tierra, donde a su vez se clasifica si tiene o no motor o por sus emisiones de dióxido de carbono ( $CO_2$ ). La clasificación de los que tienen motor es aquellos que necesitan propulsión para su funcionamiento como lo son las motocicletas, los carros, los especiales para el sector agrícola, obras y servicios, los vehículos que se encuentran en la categoría que no necesitan motor están los de tracción con animales, para personas con movilidad reducida, las bicicletas, los remolques o ciclomotores (Helloauto, n.d.)

Teniendo como base lo anterior un vehículo es aquella maquina la cual permite el desplazamiento de un sitio a otro, las que se usan con mayor frecuencia son las bicicletas, los carros y las motos.

#### **4.4 Marco Legal**

Las medidas que ha tomado la alcaldía de Bogotá con el fin de disminuir la accidentalidad vial y mejorar el bienestar de los ciudadanos han sido variadas en unas localidades como lo es Chapinero y Teusaquillo han colocado reducción en la calzada, en otras localidades como Barrios Unidos, Santa Fe y los Mártires se realizaron mejoras en las Ciclo Rutas (MovilidadBogota, 2018).

Se creó un Plan de Seguridad Vial el cual permite que entidades privadas y públicas trabajen de la mano con el fin de reducir la cantidad de siniestros viales donde hay involucrados accidentes fatales, en el cual se han incluido directrices dadas por la Organización Mundial de la Salud; todo esto registrado en el documento Visión Cero (Alcaldía Mayor de Bogotá, 2017b) en el cual se detallan las acciones del programa cuyos participantes son el Ministerio de Transporte,

la Policía Nacional, la Policía de Tránsito y Transporte, el Ministerio de Salud y Protección Social, la Dirección de Tránsito y Transporte (DITRA) y la Agencia Nacional de Seguridad Vial.

En el ámbito legal se han promulgado varias leyes y decretos como lo es el decreto 397 del 2010 por parte de la Alcaldía Mayor de Bogotá el cual habla sobre las acciones que se deben realizar para disminuir la morbilidad y mortalidad en las vías con sus respectivos responsables. (Alcaldía Mayor de Bogotá, 2010).

La ley 1503 del 2011 a nivel Nacional en donde menciona que las entidades gubernamentales en compañía de las privadas deben realizar una labor de educación vial a la población colombiana (Alcaldía Mayor de Bogotá, 2011).

El decreto 185 del 2012 por parte de la Alcaldía Mayor de Bogotá el cual habla sobre la importancia de la coordinación entre las entidades públicas y privadas las cuales deben realizar el seguimiento respectivo del Plan Distrital de Seguridad Vial (Alcaldía Mayor de Bogotá, 2012).

El decreto 2851 del 2013 a nivel Nacional en donde se reglamentan los artículos 3 al 13, 18 y 19 de la ley 1503 del 2011 y se incluye la participación de toda persona natural o jurídica que vendan bebidas alcohólicas las cuales deben realizar las acciones reglamentadas y trabajar en conjunto con las entidades responsables del seguimiento del plan de seguridad vial (Alcaldía Mayor de Bogotá, 2013).

El decreto 594 del 2015 por parte de la Alcaldía Mayor de Bogotá el cual habla sobre la importancia de vigilar el Plan de Movilidad Escolar (PME) y velar por el bienestar de la población infantil en las vías (Alcaldía Mayor de Bogotá, 2015).

El decreto 813 del 2007 por parte de la Alcaldía Mayor de Bogotá el cual va enfocado a la población que usa como medio de transporte una moto (Alcaldía Mayor de Bogotá, 2017a).

La resolución 122 del 2019 por parte de la secretaria Distrital de Movilidad en donde se establecen las directrices para la realización de auditorías en la seguridad vial de las vías urbanas (Alcaldía Mayor de Bogotá, 2019).

El acuerdo 834 del 2022 por parte del consejo de Bogotá en donde se incluye la importancia de cuidar a los animales que están en las vías, sensibilizar a la población en cuanto a la pérdida de vidas por la siniestralidad vial y animar a las entidades privadas y publicar a seguir trabajando con la ciudadanía en esta problemática (Alcaldía Mayor de Bogotá, 2022).

De acuerdo con la ley 1702 del 2013 la seguridad vial tiene un enfoque multidisciplinario ya que diferentes entes gubernamentales en compañía de organizaciones privadas elaboran las políticas y las acciones enfocadas en prevenir, controlar y disminuir el peligro de lesión o muerte sobre las personas que se transportan en los diferentes medios de transporte por las vías de la nación, las cuales su diseño debe ser el adecuado para permitir la circulación vial, la supervisión de la policía y la atención a las víctimas (Medellin, 2019).

## **5. Propuesta metodológica**

La presente investigación está enfocada en realizar análisis de datos sobre la recolección de información realizada por la Secretaría Distrital de Movilidad, cuyos resultados se encuentran en el Consolidado de Siniestros viales en Bogotá, que es una base de datos pública (Movilidad, 2021), obtenida de la página oficial de datos abiertos Bogotá, la cual cuenta con un registro de cuatrocientos veintidós mil cuatrocientos diecisiete (422.417) actores viales que tuvieron siniestralidad en el año 2015 al año 2020.

Los actores viales que se relacionan en la base de datos son: ciclistas, conductores, motociclistas, pasajeros/acompañantes y peatones los cuales se clasifican en tres tipos de

gravedad en el momento de ocurrir el siniestro vial, el primero indica que hubo muerte, el segundo relaciona las personas heridas y el tercero son daños en el lugar.

Las herramientas que se utilizan para realizar análisis de datos provienen de la matemática y la estadística (Westreicher, 2020a), la cual cuenta con seis etapas y/o pasos (Coursera, 2021) para llevar a cabo correctamente este proceso, las cuales son:

1. Preguntar: Cual es la necesidad o el objetivo de la empresa.
2. Preparar: Generar, recopilar, almacenar y administración de datos.
3. Procesar: Limpieza e integridad en los datos.
4. Analizar: Exploración, visualización y descripción de los datos.
5. Compartir: Comunicación e interpretación de los resultados.
6. Actuar: Tomar decisiones basadas en los datos.

En algunas organizaciones hacen una variación de estos pasos, obteniendo siete, que son preguntar, preparar, explorar, modelar, implementar, actuar y validar los datos que tienen almacenados y/o han recolectado por medio de encuestas.

*Tabla 1*

*Metodología por cada objetivo específico*

<b>OBJETIVO</b>	<b>METODOLOGÍA</b>
Obj1 – Describir los tres tipos principales de la estadística que se utilizan en el análisis de datos para darle el fundamento teórico al presente trabajo de grado.	<p>Paso 1. Búsqueda de fuentes de información.</p> <p>Paso 2. Caracterización de los conceptos acerca de los tipos de estadísticos.</p>

	Paso 3. Selección de la estadística descriptiva para aplicarla en la presente investigación.
Obj2 – Manipular la base de datos suministrada por la Secretaría Distrital de Movilidad para hacer uso de la estadística descriptiva.	<p>Paso 1. Exploración de los datos en Excel.</p> <p>Paso 2. Exploración y selección de variables de los datasets siniestros, actor vial, vehículo, hipótesis y diccionario en Excel.</p> <p>Paso 3. Preparación de los datos en R Studio.</p> <p>Paso 4. Procesamiento y limpieza de los datos en R Studio.</p>
Obj3 – Determinar las edades y las causas de las muertes más frecuentes en los motociclistas de la localidad de Kennedy.	Paso 5. Análisis de datos

Nota. Resumen metodología usada

Fuente: Elaboración propia

Para el desarrollo del primero objetivo que es describir los tres tipos principales de la estadística que se utilizan en el análisis de datos para darle el fundamento teórico al presente trabajo de grado. Los pasos para realizar este primer objetivo fueron:

- Paso 1: Búsqueda de fuentes de información para el acercamiento teórico, en dónde se revisaron páginas web, artículos y libros.

- Paso 2: Caracterización de los conceptos acerca de los tipos de estadísticos mediante una descripción del significado de la estadística descriptiva, predictiva y prescriptiva.
- Paso 3. Selección de la estadística descriptiva para aplicarla en la presente investigación, una vez realizado el paso dos.

Teniendo como base el procedimiento de cómo se debe realizar el análisis de datos, se describen los cinco procesos que se llevaron a cabo para cumplir con el segundo y tercer objetivo específico.

Para el desarrollo del segundo objetivo el cual es manipular la base de datos suministrada por la Secretaría Distrital de Movilidad para hacer uso de la estadística descriptiva, la metodología utilizada fueron los pasos del uno al cuatro del análisis de datos, que es limpiar y procesar la base de datos, el cual esta descrito en el marco teórico. Los pasos para realizar este segundo objetivo fueron:

- Paso 1: Exploración de los datos, en donde se hace una tabla que contiene la información de la base de datos obtenida por la Secretaría Distrital de Movilidad.
- Paso 2: Descripción de cada una de las hojas de cálculo (Excel) contenida en el documento de la Secretaría, los cuales van desde el ítem 6.1.1 al 6.1.5
- Paso 3: Preparación de los datos, en donde se hace la descripción de cada uno de los pasos para adecuar en el Software de R Studio la base de datos de la Secretaría Distrital de Movilidad
- Paso 4: Procesamiento y limpieza de los datos, en donde se utilizan códigos de programación en R para conseguir de la base de datos la información necesaria para resolver la pregunta problema.

Por último, para el desarrollo del tercer objetivo el cual es determinar las edades y las causas de las muertes más frecuentes en los motociclistas de la localidad de Kennedy, la metodología utilizada es el paso cinco del análisis de datos, que es analizar la base de datos, el cual esta descrito en el marco teórico. Los pasos para realizar este tercer objetivo fueron:

Paso 5: Análisis de los datos, este paso lo dividí en tres numerales que son el 6.5.1, el 6.5.2 en los cuales se utiliza la estadística descriptiva para poder hacer el análisis de las edades y las causas de la siniestralidad vial en la localidad de Kennedy y el 6.5.3 donde se realizan las medidas de tendencia central de las cinco causas más frecuentes de siniestralidad vial.

## 6. Resultados

### 6.1. Exploración de los datos

Como se mencionó en la metodología de investigación se va a utilizar la base de datos abierta de la Secretaría Distrital de Movilidad, en donde se detalla en la página la siguiente información:

*Tabla 2*

*Información sobre el dataset de siniestralidad vial*

Forma de presentación	Tabla digital
Frecuencia de actualización	Anual
Palabras claves	Accidente, actor vial, atropello, choque, heridos, hipótesis, incidente, muertos, siniestro, solo daños, vial, vehículo
Enlace a otros recursos	<a href="https://datos.movilidadbogota.gov.co/">https://datos.movilidadbogota.gov.co/</a>
Fecha de creación del metadato	2019-09-06
Fecha de actualización metadato	2021-10-21

Categoría temática	Transporte
Idioma	Español
Alcance de la calidad de los datos	Conjunto de datos no geográfico
Reporte de la calidad	Validación interna en aplicativo SIGAT
Distribuidor	Secretaría Distrital de Movilidad
Alcance del metadato	Conjunto de datos no geográfico
Tipo de responsable	Custodio
Responsable del metadato	Secretaría Distrital de Movilidad
Link a servicio XML	<a href="/metadata/dataset?id=0b070626-fe5a-42dc-ae99-92601da166d9">/metadata/dataset?id=0b070626-fe5a-42dc-ae99-92601da166d9</a>

Nota. Descripción general del dataset sobre siniestralidad vial

Fuente: Elaboración propia obtenida de la pág. De datos abiertos

El formato del archivo se encuentra en un .xlsx, es decir, en un libro de Excel, primero se explora la información suministrada por la entidad pública, la cual cuenta con cinco hojas de cálculo, que son siniestros, actor\_vial, vehículos, hipótesis y diccionario, a continuación, se describen los dataset.

## 6.2. Exploración y selección de variables

### 6.2.1 Dataset siniestros

La información que contiene esta hoja de datos son 196.152 registros los cuales están en diez columnas, que son las siguientes:

- Código\_accidente: Es un número que está compuesto por 7 y 8 dígitos, el cual comienza por 4401419 y va hasta el 10537008 que lo da la Secretaría Distrital de Movilidad para realizar el registro del siniestro.

- Fecha: De acuerdo con la Secretaría Distrital de Movilidad las fechas de la obtención de los registros de los siniestros viales ocurrieron en un rango de fechas entre el 01/01/2015 hasta el 31/12/2020.
- Hora: Es el registro en hora militar de los siniestros viales ocurridos entre las 00:00:00 y las 23:59:00
- Gravedad: la Secretaría Distrital de Movilidad lo categorizo en tres niveles, uno (1) es el nivel más alto, que son los muertos, el dos (2) son los heridos y el tres (3) son solo daños.
- Clase: El rango de códigos es del 1 al 7 los cuales representan la forma como ocurrió el siniestro, es decir, del menor al mayor, son los siguientes: choque, atropello, volcamiento, caída de ocupante, incendio, otro y autolesión.
- Choque: El rango de códigos es del 1 al 4 los cuales representan quienes se chocaron, de menor a mayor son los siguientes: Vehículo, tren, semoviente y objeto fijo.
- Objeto\_fijo: El rango de códigos es del 1 al 11 los cuales representan cual fue el objeto fijo contra el cual hubo siniestro vial, de menor a mayor son los siguientes: muro, poste, árbol, barandas, semáforo, inmueble, hidrante, valla (señal), tarima (caseta), vehículo estacionado y otro.
- Dirección: Es la ubicación geográfica del lugar de los hechos donde ocurrió el siniestro vial.
- Codigo\_localidad: Se encuentran los códigos de las 20 localidades de la ciudad de Bogotá.
- Diseño\_lugar: El rango de códigos es del 1 al 13 los cuales son un complemento a la dirección, ya que da una pequeña descripción del lugar del siniestro vial, de menor a

mayor son los siguientes: tramo de vía, intersección, vía peatonal, paso elevado, paso inferior, paso a nivel, glorieta, puente, vía troncal, lote o predio, cicloruta, ponton y túnel.

Se tiene la siguiente tabla la cual muestra un ejemplo de los datos en el dataset:

*Tabla 3*

*Ejemplo del dataset de Siniestros*

<b>CODIGO_ACCIDENTE</b>	4401495
<b>FECHA</b>	02/01/2015
<b>HORA</b>	17:15:00
<b>GRAVEDAD</b>	2
<b>CLASE</b>	1
<b>CHOQUE</b>	4
<b>OBJETO_FIJO</b>	10
<b>DIRECCION</b>	CL 41-KR 11 SE 94
<b>CODIGO_LOCALIDAD</b>	4
<b>DISENO_LUGAR</b>	1

Nota. Registro del archivo de Excel de siniestros\_viales\_consolidados\_bogota\_dc.xlsx

Fuente: Elaboración propia obtenida del archivo de Excel de la Secretaría Distrital de Movilidad

Las columnas seleccionadas de este dataset para realizar el siguiente paso en R fueron el código de accidente, la gravedad del siniestro y el código de la localidad.

### **6.2.2 Dataset Actor Vial**

La información que contiene esta hoja de datos son 422.417 registros los cuales al restarlos con el dataset anterior se tienen 226.265 registros, se evidencia que algunos códigos de accidente están repetidos, dado que este dataset se llama actor\_vial, significa que el siniestro ocurrido puede estar involucrada una o dos personas, el dataset cuenta con ocho columnas que son:

- Código\_accidente: Su descripción se encuentra en el dataset de siniestros.

- Código\_accidentado: Es un número que está compuesto por 7 y 8 dígitos, donde el valor mínimo es 12549882 y el valor máximo es 2700326 que lo da la Secretaría Distrital de Movilidad para realizar el registro del siniestro.
- Fecha: Su descripción se encuentra en el dataset de siniestros.
- Condición: Es el rol de la persona dentro del siniestro, es decir, si es un ciclista, un conductor, un motociclista, un pasajero/acompañante o un peatón.
- Estado: Una vez ocurrido el siniestro vial es la condición en la cual quedo la persona afectada, es decir, si quedo herida, ilesa o muerta
- Edad: Es la edad de la persona involucrada en el siniestro su rango va de menos de un año hasta la edad de 120 años y se realizó un recuento de 7.911 registros etiquetados como sin información.
- Sexo: Es el género de la persona en el siniestro, es decir, masculino, femenino y hay un recuento de 9.639 registros sin información.
- Vehículo: Es un número que está compuesto por 8 dígitos y un dígito de verificación el cual indica si ubo una o dos personas involucradas en el siniestro vial, donde el valor mínimo es el 10446658-1 y el valor máximo es 4516903-1 que lo da la Secretaría Distrital de Movilidad para realizar el registro del siniestro y hay un recuento de 23.447 celdas que están vacías las cuales representan a los peatones.

Las columnas seleccionadas de este dataset para realizar el siguiente paso en R fueron el código de accidente, la edad y el sexo de la persona que sufrió el siniestro vial.

Se tiene la siguiente tabla la cual muestra un ejemplo de los datos en el dataset:

*Tabla 4*

*Ejemplo del dataset de Actor Vial*

<b>CODIGO_ACCIDENTE</b>	10446665
-------------------------	----------

<b>CODIGO_ACCIDENTADO</b>	12549897
<b>FECHA</b>	04/05/2018
<b>CONDICION</b>	MOTOCICLISTA
<b>ESTADO</b>	MUERTO
<b>EDAD</b>	23
<b>SEXO</b>	MASCULINO
<b>VEHICULO</b>	10446665-2

Nota. Registro del archivo de Excel de siniestros\_viales\_consolidados\_bogota\_dc.xlsx

Fuente: Elaboración propia obtenida del archivo de Excel de la Secretaría Distrital de Movilidad

### 6.2.3 Dataset Vehículos

La información que contiene esta hoja de datos son 371.606 registros los cuales están en siete columnas que son:

- Código\_accidente: Su descripción se encuentra en el dataset de siniestros.
- Fecha: Su descripción se encuentra en el dataset de siniestros.
- Vehículo: Es el código del vehículo asignado por la Secretaría Distrital de Movilidad.
- Clase: El rango de códigos es del 1 al 27 los cuales representan el tipo de vehículo que le ocurrió el siniestro, es decir, del menor al mayor, son los siguientes: automóvil, bus, buseta, camión (furgón), camioneta, campero, microbús, tractocamión, volqueta, motocicleta, M. Agrícola, M. industrial, bicicleta, motocarro, tracción animal, otro, motociclo, ambulancia, no identificado, bicitaxi, bus articulado, bus alimentador, tren, metro, remolque, cuatrimoto, semi-remolque y bus-biarticulado.
- Servicio: El rango de códigos es del 1 al 4 los cuales indican si el vehículo pertenece al sector público o privado, es decir, del menor al mayor, son los siguientes: oficial, público, particular y diplomático.
- Modalidad: El rango de códigos es del 1 al 12 los cuales indican la cantidad de pasajeros que puede transportar el vehículo, es decir, del menor al mayor, son los

- siguientes: mixto, carga - extrapesada, carga - extradimensionada, pasajeros - colectivo, pasajeros - individual, pasajeros - especial escolar, pasajeros - especial turismo, pasajeros - especial asalariado, pasajeros - especial ocasional, pasajeros - masivo, carga pesada y carga - mercancía peligrosa.
- Enfuga: Solo cuenta con dos valores un N (No) o un S (Si), lo que significa que el vehículo que participo en el siniestro se fugó o no.

Las columnas seleccionadas de este dataset para realizar el siguiente paso en R fueron el código de accidente y el vehículo.

Se tiene la siguiente tabla la cual muestra un ejemplo de los datos en el dataset:

*Tabla 5*

*Ejemplo del dataset de Vehículos*

<b>CODIGO_ACCIDENTE</b>	4401423
<b>FECHA</b>	01/01/2015
<b>VEHICULO</b>	4401423-1
<b>CLASE</b>	1
<b>SERVICIO</b>	2
<b>MODALIDAD</b>	5
<b>ENFUGA</b>	N

Nota. Registro del archivo de Excel de siniestros\_viales\_consolidados\_bogota\_dc.xlsx

Fuente: Elaboración propia obtenida del archivo de Excel de la Secretaría Distrital de Movilidad

### **6.2.4 Dataset Hipótesis**

La información que contiene esta hoja de datos son 233.820 registros los cuales están en tres columnas que son:

- Código\_accidente: Su descripción se encuentra en el dataset de siniestros.
- Fecha: Su descripción se encuentra en el dataset de siniestros.
- Código\_causa: Es el código con el cual la Secretaría distrital de movilidad los asigno, van desde el 90 hasta el 506.

Las columnas seleccionadas de este dataset para realizar el siguiente paso en R fueron el código de accidente, la edad y el código de la causa del siniestro vial.

Se tiene la siguiente tabla la cual muestra un ejemplo de los datos en el dataset:

*Tabla 6*

*Ejemplo del dataset Hipótesis*

<b>CODIGO_ACCIDENTE</b>	4401425
<b>FECHA</b>	01/01/2015
<b>CODIGO_CAUSA</b>	115

Nota. Registro del archivo de Excel de siniestros\_viales\_consolidados\_bogota\_dc.xlsx

Fuente: Elaboración propia obtenida del archivo de Excel de la Secretaría Distrital de Movilidad

### **6.2.5 Dataset Diccionario**

La información que contiene esta hoja de datos son 212 registros los cuales están en tres columnas que son:

- Campo: Menciona los datasets descritos anteriormente con su nombre.
- Código: Es el código que la Secretaría distrital de movilidad le asigno el cual lo menciona en los datasets anteriores.
- Descripción: Es el significado de cada uno de los códigos dados.

Las columnas seleccionadas de este dataset para realizar el siguiente paso en R fueron el código de accidente y la descripción del siniestro vial.

De acuerdo con el análisis realizado anteriormente todas las hojas cuentan con un código de accidente el cual se va a utilizar en R para la unión de la información.

Se tiene la siguiente tabla la cual muestra un ejemplo de los datos en el dataset:

Tabla 7

Ejemplo del dataset Diccionario

<b>HOJA</b>	SINIESTROS
<b>CAMPO</b>	GRAVEDAD
<b>CODIGO</b>	1
<b>DESCRIPCION</b>	Con Muertos

Nota. Registro del archivo de Excel de siniestros\_viales\_consolidados\_bogota\_dc.xlsx

Fuente: Elaboración propia obtenida del archivo de Excel de la Secretaría Distrital de Movilidad

### 6.3. Preparación de los datos

El proceso de preparación de datos se realizó en el programa R Studio, de la siguiente manera:

1. Cargar las librerías necesarias en R Studio

Figura 2:

Paquetes de librerías usados de R

```

1 ## Cargando paquetes
2 library(readxl) # Para cargar archivos de excel
3 library(tidyverse) # Libreria para para el procesamiento de datos
4 library(dplyr) # Para manipular datos
5 library(data.table) # Procesamiento de datos
6 library(agricolae) # Manejo de histogramas
7 library(ggplot2) # Para graficar
8

```

Nota. Software R Studio para análisis de datos

Fuente: Elaboración propia

2. Cargue del archivo de Excel en el software de R.

Figura 3:

*Cargue de los datasets de Excel a R*

```

14 ##### Cargar los datasets
15
16 setwd('C:/Users/Milena/Documents/Grado')
17 siniestros = read_excel('siniestros_viales_consolidados_bogota_dc.xlsx',
18                         sheet = 'SINIESTROS')
19 actor = read_excel('siniestros_viales_consolidados_bogota_dc.xlsx',
20                   sheet = 'ACTOR_VIAL')
21 vehiculos = read_excel('siniestros_viales_consolidados_bogota_dc.xlsx',
22                        sheet = 'VEHICULOS')
23 hipotesis = read_excel('siniestros_viales_consolidados_bogota_dc.xlsx',
24                        sheet = 'HIPOTESIS')
25 diccionario = read_excel('siniestros_viales_consolidados_bogota_dc.xlsx',
26                           sheet = 'DICCIONARIO')
27

```

Nota. Software R Studio para análisis de datos

Fuente: Elaboración propia

Con `setwd`, se le indica al software de R la ruta donde está ubicado el archivo de Excel obtenido de la Secretaría Distrital de Movilidad, la funcionalidad de `read_excel` permite que el software de R pueda leer el archivo de Excel, dado que el archivo cuenta con cinco hojas y/o datasets, se cargan en R las cuatro que se van a utilizar, guardando en la variable de cada uno de los nombres la respectiva hoja de cálculo.

#### 6.4. Procesamiento y limpieza de los datos

1. Realizar la unión de los datasets.

Figura 4:

*Código para unión de datasets en R Studio*

```

54 ##### Uniendo datos
55 datos = merge(siniestros,actor,
56               by = c('CODIGO_ACCIDENTE', 'FECHA'),
57               all = T)
58 #422.564
59

```

Nota. Software R Studio para análisis de datos

Fuente: Elaboración propia

La función en R merge sirve para la unión de dos datasets, en donde se van a unir primero la información guardada en la variable siniestros y actor cargadas anteriormente, la línea 30, en by se especifica el nombre de las columnas para hacer la unión, las cuales como se hizo la descripción anteriormente aparecen en ambos dataset, en la línea 31 all = T significa que debe unirse la información en ambos dataset, obteniendo en datos un total de 422.564 registros y la línea 33 indica.

De la misma forma se realiza la unión del dataset obtenido con la hoja de cálculo donde están los vehículos, obteniendo 422.567 registros, y al hacer la unión con la hoja de hipótesis se obtuvieron 509.442 registros, la hoja de cálculo de diccionario no se une a estas ya que es donde se encuentra el significado de los códigos usados por la Secretaría Distrital de Movilidad.

2. Una vez con los datos en un solo data frame, se procede a hacer el filtro de la información, teniendo en cuenta los datos y la información obtenida en la introducción del presente trabajo, es decir, se filtra la información para que solo este como actor vial los motociclistas cuyo genero sea masculino.

*Figura 5:*

*Código que filtra la base de datos en R Studio*

```

81 ##### Filtrar BD solo motos y hombres
82 datos_f = datos_f %>%
83   filter(CONDICION == 'MOTOCICLISTA'
84         & SEXO == 'MASCULINO')
85 #60.137
86

```

Nota. Software R Studio para análisis de datos

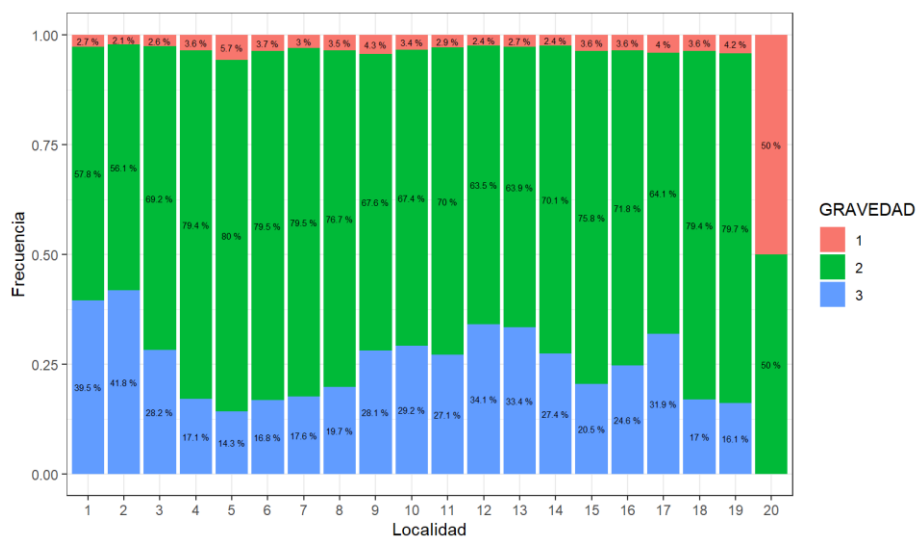
Fuente: Elaboración propia

El símbolo `%>%` llamado pipe es un operador que sirve para concatenar operaciones de la librería `dplyr`, donde el significado del anterior código es que almacene en la variable `datos_f`, el dataframe que tenga filtrada la columna condición que es motociclista y la columna sexo que son las personas de género masculino, obteniendo un resultado de 60.137 registros.

- Para explorar como es el comportamiento de la siniestralidad en cada la localidad, se realizó un gráfico, en el cual se muestran los 60.137 registros, los cuales los porcentajes que aparecen en cada barra significa la distribución categórica de la gravedad en cada una de las localidades.

*Figura 6:*

*Nivel de gravedad por localidad*



Nota. Gráfica de frecuencias absolutas por cada localidad y su nivel de gravedad

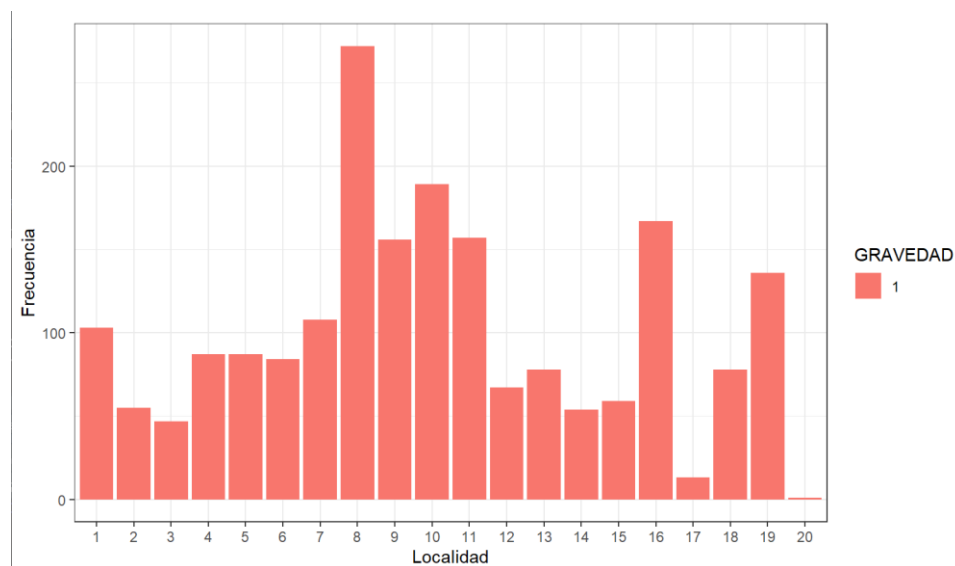
Fuente: Elaboración propia

Según la nomenclatura utilizada por la Secretaría Distrital de Movilidad, la gravedad 1 hace referencia a los siniestros que tuvieron muertes, la 2 donde hubieron heridos y la 3 donde se presentaron solo daños.

Dado que solo se quiere conocer la localidad que tuvo la mayor cantidad de siniestralidad donde hubo muertes, se realiza el filtro indicando solo la gravedad 1.

*Figura 7*

*Frecuencia de la cantidad de muertos por localidad*



Nota. Gráfica de frecuencias absolutas por cada localidad

Fuente: Elaboración propia

De acuerdo con el gráfico anterior se evidencia que la localidad con mayor número de muertes en los siniestros viales es Kennedy (8) con 272 registros.

## 6.5. Análisis de los datos

En esta sección se encuentran tres numerales, el 6.5.1 donde se realiza un análisis descriptivo sobre las edades de los varones que tuvieron muerte en el siniestro vial de Kennedy,

el numeral 6.5.2 el cual analiza las causas de los siniestros de muertes en estos motociclistas y el 6.5.3 donde se encuentra relacionado el análisis de las edades junto con sus causas.

### 6.5.1. Análisis descriptivo sobre las edades en la localidad de Kennedy

Con los resultados obtenidos en el numeral anterior, de las 272 personas muertas se obtuvieron las medidas de posición y de tendencia central que son las siguientes:

*Tabla 8*

*Medidas de tendencia central de las edades en motociclistas muertos en la localidad de Kennedy*

Mínimo	18.00
Q1	22.00
Q2=Mediana	27.00
Media	29.46
Q3	34.00
Q4	72.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

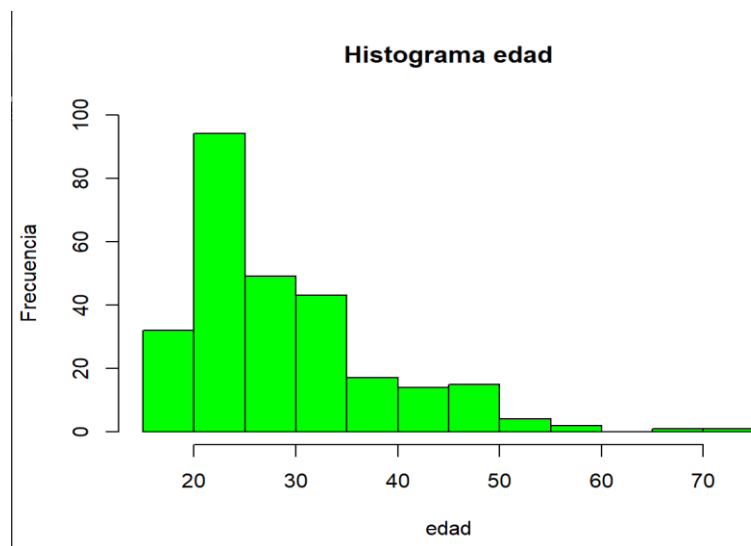
De acuerdo con la tabla 8, se puede apreciar que la edad mínima de las víctimas fatales en accidentes viales se ha presentado personas de 18 años, adicional a esto, se analiza que le edad promedio de personas que han muerto en la localidad de Kennedy está entre 29 y 30 años ya que la media de las edades fue de 29.46

La mediana de edades de víctimas mortales es de 27 años la cual se acerca a la edad promedio y la persona con mayor edad que perdió la vida en un siniestro vial fue de 72 años.

Obteniendo el siguiente histograma de frecuencias absolutas:

Figura 8

*Frecuencia absoluta de edades en motociclistas muertos en la localidad de Kennedy*



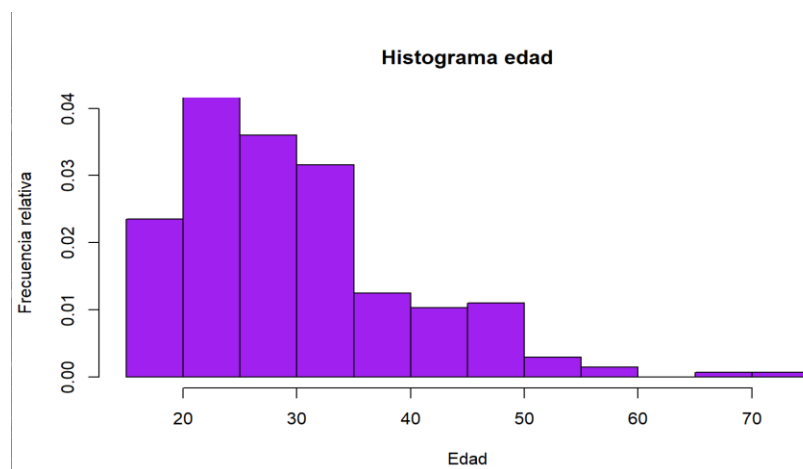
Nota. Gráfica de frecuencias absolutas de las edades de los motociclistas muertos en la localidad de Kennedy

Fuente: Elaboración propia

Y el diagrama de frecuencias relativas:

Figura 9

*Frecuencia relativa de edades en motociclistas muertos en la localidad de Kennedy*



Nota. Gráfica de frecuencias relativas de las edades de los motociclistas muertos en la localidad de Kennedy

Fuente: Elaboración propia

Adicional a esto, las medidas de dispersión que se obtuvieron fueron:

*Tabla 9*

*Medidas de dispersión de las edades en motociclistas muertos en la localidad de Kennedy*

Desviación estándar	9.629972
Varianza	92.73637
Coefficiente de varianza (CV)	0.3268879

Nota. Descripción del comportamiento de dispersión de la edad de los motociclistas muertos en Kennedy

Fuente: Elaboración propia

Ahora, se puede apreciar que las medidas de dispersión de la tabla 3 junto con la figura 9 explican el comportamiento de la dispersión de las edades, teniendo presente el concepto de rango intercuartílico descrito en el marco conceptual, se percibe que la dispersión de los datos es de 9.6 años, esto quiere decir que la distribución de frecuencias con respecto a la edad media es de 9.6, debido a que la edad mínima es de 18 años y la edad máxima es de 72 años.

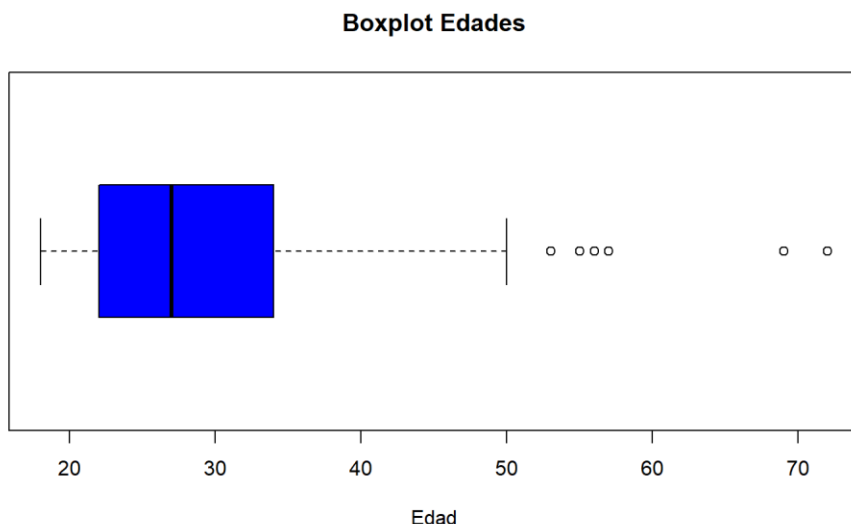
Dado que la varianza es una característica numérica de las variables aleatorias, se utiliza el coeficiente de variación para entender mejor la forma en que varía la edad de las personas que han muerto en esta localidad.

Por lo tanto, la edad varía con respecto a la media en un 32.68%, es decir, que la dispersión de las edades con respecto a la media es alta.

A continuación, se presentan el diagrama de caja (boxplot).

*Figura 10*

*Distribución de la variable edad en motociclistas muertos en la localidad de Kennedy*



Nota. Gráfica de los datos mínimo y máximo, valores outliers y cuartiles

Fuente: Elaboración propia

En el diagrama de caja de la figura 9 muestra que el valor de 72 años (dato máximo de las edades en Kennedy) se trata posiblemente de un dato atípico (outlier), este dato afecta el valor de la edad promedio y genera una distribución sesgada a la derecha como se puede apreciar en los histogramas, además, los valores con mayores frecuencias absolutas y relativas se encuentran entre los 20 y 40 años, es decir, que hay un alto número personas jóvenes que han muerto en la localidad de Kennedy.

### **6.5.2. Análisis descriptivo de las causas de mortalidad en Kennedy**

Con los resultados obtenidos en los numerales anteriores, a continuación, se muestra la tabla de frecuencias en R Studio sobre la causalidad de las muertes en la localidad de Kennedy.

Tabla 10

*Frecuencias de las causas de accidentalidad vial en los motociclistas muertos en la localidad de Kennedy*

Cod_causa	Frecuencia	Descripción
97	1	Transitar por vías prohibidas (ciclista o motociclista)
123	1	No respetar prelación de intersecciones o giros (conductor en general)
133	1	Subirse al andén o vías peatonales (conductor en general)
305	1	Obstáculo en la vía (vía)
401	1	Pasar semáforo en rojo (peatón)
93	2	Transitar distante de la acera u orilla de la calzada (ciclista o motociclista)
106	2	Adelantar invadiendo carril del mismo sentido en zigzag (conductor en general)
122	2	Girar bruscamente (conductor en general)
127	2	Transitar en contravía (conductor en general)
132	2	No respetar prelación (conductor en general)
139	2	Impericia en el manejo (conductor en general)
143	2	Poner en marcha un vehículo sin precauciones (conductor en general)
303	2	Superficie lisa (vía)
402	2	Salir por delante de un vehículo (peatón)
406	2	Cruzar en diagonal (peatón)
102	3	Adelantar por la derecha (conductor en general)
103	3	Adelantar cerrando (conductor en general)
115	3	Embriaguez o droga (conductor en general)
404	3	Transitar por la calzada (peatón)
410	3	Cruzar en estado de embriaguez (peatón)
306	4	Huecos (vía)
121	10	No mantener distancia de seguridad (conductor en general)
114	12	Embriaguez aparente (conductor en general)
112	13	Desobedecer señales (conductor en general)
142	13	Semáforo en rojo (conductor en general)
409	19	Cruzar sin observar (peatón)
98	21	Transitar entre vehículos (ciclista o motociclista)
116	28	Exceso de velocidad (conductor en general)
411	49	Otras (peatón)
157	63	Otra (conductor en general)

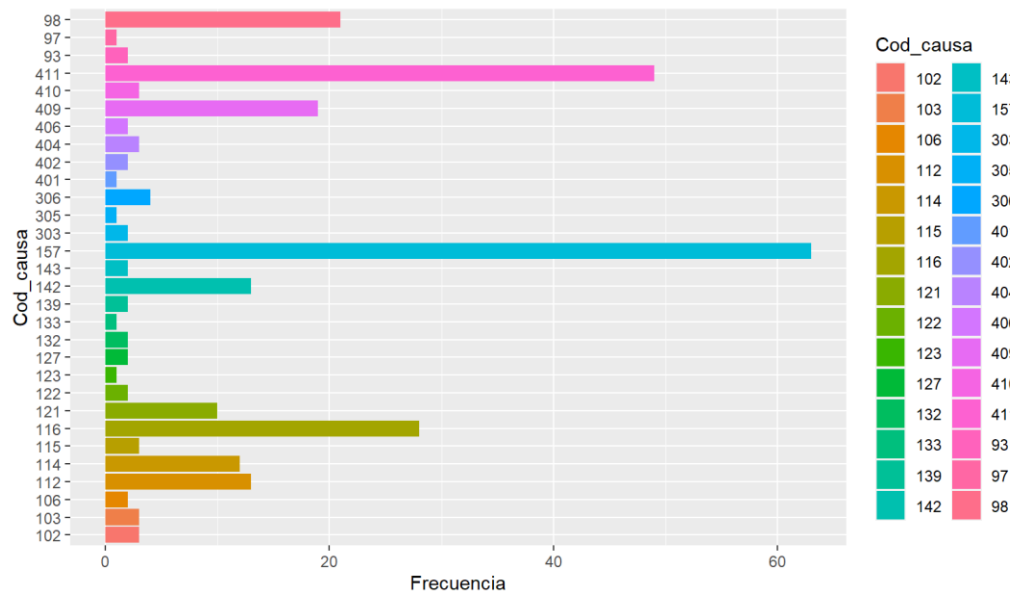
Nota. Causas de muerte en la accidentalidad vial de los motociclistas masculinos en Kennedy

Fuente: Elaboración propia

El diagrama de barras con los códigos de las causas se presenta a continuación:

*Figura 11*

*Causas de mortalidad en la localidad de Kennedy*



Nota. Gráfica de frecuencias absolutas sobre el código de causa de muertes

Fuente: Elaboración propia

Con base en la tabla de frecuencias absolutas y en el diagrama de barras, se puede notar que la mayor causa del accidente es el código 157 el cual se asocia con algún comportamiento indebido del conductor con 63 casos, en segunda posición se encuentra el código 411 el cual se asocia con malos comportamientos de los peatones contando con 49 casos.

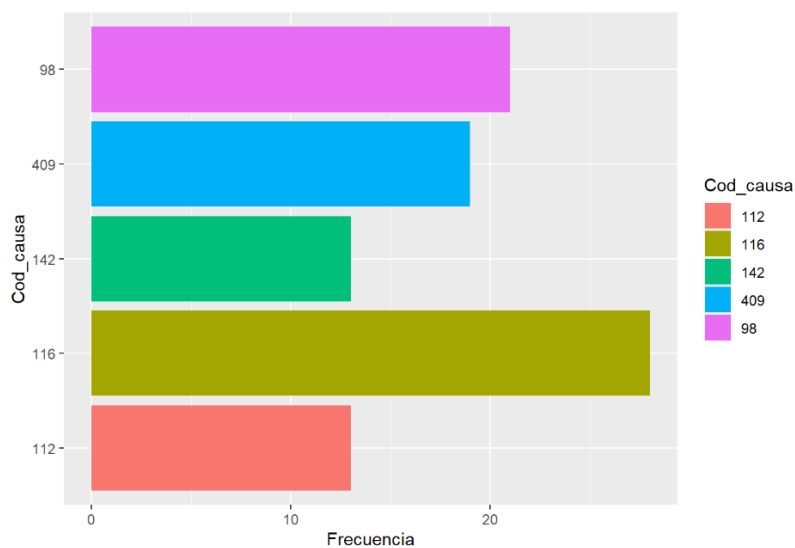
Sigue el código 116 el cual se asocia con el exceso de velocidad del motociclista con 28 casos, seguido del código 98 el cual se asocia con transitar entre vehículo o también llamado sigsageo con 21 casos y en la quinta posición de las causas más frecuentes de muerte el código 409 el cual se asocia con los peatones imprudentes que cruzan la calle sin mirar contando con 19 casos.

Dado que las dos causas más frecuentes no dan ninguna información acerca de las causas de mayor muerte en la localidad, ya que la descripción de los códigos 157 y 411 se relacionan a otros relacionadas con los conductores y los peatones, se toman las siguientes dos causas de muerte por siniestralidad vial en la localidad de Kennedy que están en la lista, las cuales son: el código 142 que es pasarse el semáforo en rojo y el código 112 que es desobedecer las señales de tránsito con la misma cantidad de muertes.

Para apreciar mejor estos resultados, se obtiene el diagrama de barras con las causas más frecuentes y se tiene la descripción de los códigos en la tabla N. 11:

*Figura 12*

*Causas más frecuentes de mortalidad en la localidad de Kennedy*



*Tabla 11*

116	28	Exceso de velocidad
98	21	Transitar entre vehículos
409	19	Cruzar sin observar (peatón)
142	13	Semáforo en rojo
112	13	Desobedecer señales

Nota. Gráfica de las cinco principales causas de muerte en la localidad de Kennedy

Fuente: Elaboración propia.

### 6.5.3. Análisis descriptivo de las edades junto con las principales causas de mortalidad en Kennedy

Teniendo en cuenta los dos numerales anteriores donde se obtuvo que la edad mínima de la persona que falleció en un siniestro vial fue de 18 años y que la causa más frecuente es el exceso de velocidad se va a analizar ahora las medidas de tendencia central en las cinco causas principales de accidentalidad vial de los motociclistas en la localidad de Kennedy en donde hubo muerte.

Para la primera causa más frecuente de muerte que es el exceso de velocidad representada por el código 116 con 28 casos de personas fallecidas se obtuvieron los siguientes datos

*Tabla 12*

*Medidas de tendencia central de las edades y la causa de exceso de velocidad*

Mínimo	19.00
Q1	21.75
Q2=Mediana	24.00
Media	26.29
Q3	29.25
Q4	46.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

La edad mínima de las víctimas fatales con exceso de velocidad fue de 19 años, la edad promedio fue de 26 años y la persona con mayor edad fallecida fue de 46 años.

Para la segunda causa más frecuente de muerte que es transitar entre los vehículos o el llamado sigsageo representado por el código 98 con 21 casos de personas fallecidas se obtuvieron los siguientes datos:

*Tabla 13*

*Medidas de tendencia central de las edades y transitar entre vehículos*

Mínimo	19.00
Q1	22.00

Q2=Mediana	28.00
Media	29.67
Q3	37.00
Q4	49.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

La edad mínima de las víctimas fatales que transitaron entre los vehículos fue de 19 años, la edad promedio fue entre los 29 y 30 años y la persona con mayor edad fallecida fue de 49 años.

Para la tercera causa más frecuente de muerte se ve involucrada no solo la persona que va manejando la motocicleta si no también la imprudencia de los peatones está causa está representada por el código 409 con 19 casos de personas fallecidas se obtuvieron los siguientes datos:

*Tabla 14*

*Medidas de tendencia central de las edades y la causa cruzar sin observar (peatón)*

Mínimo	18.00
Q1	22.50
Q2=Mediana	27.00
Media	32.42
Q3	40.00
Q4	53.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

La edad mínima de las víctimas fatales que murió por la imprudencia del peatón fue de 18 años, la edad promedio fue de 32 años y la persona con mayor edad fallecida fue de 53 años.

Para la cuarta causa más frecuente de muerte es cruzar el semáforo en rojo, representada por el código 142 con 13 casos de personas fallecidas se obtuvieron los siguientes datos

*Tabla 15*

*Medidas de tendencia central de las edades y la causa de semáforo en rojo*

Mínimo	18.00
--------	-------

Q1	20.00
Q2=Mediana	21.00
Media	24.23
Q3	25.00
Q4	43.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

La edad mínima de las víctimas fatales que murió por cruzarse el semáforo en rojo fue de 18 años, la edad promedio fue de 24 años y la persona con mayor edad fallecida fue de 43 años.

Y la última causa más frecuente de muerte que es desobedecer las señales de tránsito representada por el código 112 con 13 casos de personas fallecidas se obtuvieron los siguientes datos

*Tabla 16*

*Medidas de tendencia central de las edades y la causa de desobedecer señales*

Mínimo	20.00
Q1	24.50
Q2=Mediana	25.00
Media	28.69
Q3	29.00
Q4	56.00

Nota. Descripción de la distribución de la muestra de los hombres motocicletas muertos en siniestros viales

Fuente: Elaboración propia

La edad mínima de las víctimas fatales que murió por desobedecer las señales de tránsito fue de 20 años, la edad promedio fue de entre los 28 y 29 años y la persona con mayor edad fallecida fue de 56 años.

Al realizar un análisis de las cinco causas más frecuentes de siniestralidad vial donde hubo muerte por parte de los varones motociclistas en la localidad de Kennedy se obtuvo que:

- El promedio de edades en estas causas va desde 24 años a 32 años.
- La edad mínima del motociclista muerto fue de 18 años y la edad máxima fue de 56 años.

- Los más jóvenes mueren por no obedecer las señales de tránsito, en especial no respetar los semáforos en rojo.
- Aunque entre los adultos se ve que no obedecen las señales, los datos muestran que han muerto por la imprudencia de los peatones por pasar las calles sin observar.
- Por transitar entre los vehículos y el exceso de velocidad han muerto motociclistas entre los 26 y 29 años.

Teniendo el conocimiento de los resultados obtenidos al comparar las edades vs. Causas las entidades encargadas del tránsito pueden tomar medidas preventivas enfocadas a esta población joven con el fin de minimizar las muertes en la localidad de Kennedy de la capital.

## **7. Conclusiones y Recomendaciones**

Debido a la importancia de la transformación digital y su implementación en los diferentes tipos de empresas ha hecho que la ingeniería de software sea vista de forma multidisciplinaria puesto que mejora los procesos internos de las compañías ya sean públicas o privadas.

La rama de la ingeniería de software que está tomando fuerza desde la popularización del Big Data, es la llamada análisis de datos, la cual utiliza como herramienta principal la estadística, donde para el desarrollo de este proyecto de investigación se utilizó la estadística descriptiva, dado que permite conocer como es el comportamiento de los datos que se tienen, al seguir su debido proceso y lograr que las directivas puedan tomar decisiones basada en los datos.

Una vez terminado de realizar el proceso de análisis de datos, comenzando con la limpieza de datos, la segmentación y la aplicación de técnicas estadísticas sobre la base de datos dada por la Secretaría Distrital de Movilidad, al segmentar la información con la variable edad de

los hombres motociclistas muertos en los siniestros viales, los resultados arrojados fueron similares a los dados por las entidades públicas como es la Agencia Nacional de Seguridad Vial (ANSV).

Esto se afirma porque en los resultados obtenidos, se obtuvo que la edad promedio de los motociclistas varones que mueren en accidentes viales en la localidad de Kennedy está entre los 29 y 30 años, y la segunda posición a nivel nacional es el rango de edad entre los 25 a los 30 años, lo cual da a entender que la localidad de Kennedy sigue la tendencia de los resultados a nivel nacional.

De la misma manera sucede con las causas de ocurrencias de siniestralidad vial, dado que la investigación muestra que a nivel nacional la principal causa con el 45% de las muertes viales en motociclistas fueron por el exceso de velocidad, en donde la localidad de Kennedy ocupa el tercer lugar.

Las recomendaciones para reducir el exceso de velocidad basados en que la edad promedio es de 26 años (categoría joven-adulto) y con base en lo que se mencionó en el concepto de siniestralidad vial donde la personalidad de los individuos influye mucho al momento de conducir, es de vital importancia que las empresas que emiten el pase de conducción persuadan en las capacitaciones las graves consecuencias de no respetar los límites de velocidad junto con la necesidad de saber controlar las emociones debido a que emociones como la ira hacen que lleve al individuo a acelerar de forma inconsciente, adicional estos motociclistas pueden adquirir dispositivos y/o aplicaciones que los alerten en el momento que estén excediendo los límites establecidos de velocidad.

De igual manera se recomienda que para evitar el incumplimiento de las leyes de tránsito la Policía Nacional podrían incrementar operativos en las vías con mayor índice de siniestralidad

vial y de incumplimiento a las normas viales permitiendo también controlar la velocidad de los motociclistas, por parte de la alcaldía subir los partes no solo por exceso de velocidad sino también por desobedecer las señales de tránsito, instalar radares y cámaras que capten el mal comportamiento de los conductores y el mejoramiento de la infraestructura vial con más señalización y mejora en las calles reduciendo el número de huecos en las vías.

De acuerdo con los resultados obtenidos la limitación que se tuvo en la investigación fue la base de datos debido a no haber especificidad en la información, en otras palabras la base de datos suministrada por la Secretaría Distrital de Movilidad tiene una variable en la sección de causas de mortalidad llamada: Otra (conductor en general), que fue la causa principal de mortalidad en la localidad de Kennedy con una frecuencia de 63 motociclistas pero en la descripción de la tabla dada por la Secretaria Nacional de Movilidad no se encuentra información adicional de que causas pertenecen al conjunto de "otra", de igual forma sucedió con la segunda causa con mayor frecuencia, como consecuencia a esto se tomó la posición tercera como la primer causa frecuente de mortalidad para poder obtener las cinco causas más frecuentes de mortalidad vial en los motociclistas varones de la localidad.

Por consiguiente, se recomienda a la Secretaría Distrital de Movilidad y/o a las entidades encargadas de gestionar datos de transporte ya sean públicas o privadas que al momento de colocar los dataset abiertos al público las variables sean más específicas para evitar ambigüedades, donde especifiquen tomando esta investigación como ejemplo para la entidad que significan otras causas y/o cómo hacen la medición de este conjunto de causas.

Para líneas futuras de investigación a los investigadores se les propone aplicar las mismas técnicas aquí implementadas en otras localidades de la capital, con el fin de generar un análisis

comparativo entre ellas para que las administraciones públicas de cada localidad puedan tomar medidas colectivas y tener una ciudad con menor cantidad de motociclistas muertos.

## 8. Referencia Bibliográfica

AgenciadeSeguridadVial. (2022). *Cifras año en curso | ANSV. 2022.*

<https://ansv.gov.co/es/observatorio/estadísticas/cifras-ano-en-curso>

Alcaldía Mayor de Bogotá. (2010). *Decreto 397 de 2010 Alcaldía Mayor de Bogotá, D.C.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=40354>

Alcaldía Mayor de Bogotá. (2011). *Ley 1503 de 2011 Nivel Nacional.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=45453&dt=S>

Alcaldía Mayor de Bogotá. (2012). *Decreto 185 de 2012 Alcaldía Mayor de Bogotá, D.C.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=47065&dt=S>

Alcaldía Mayor de Bogotá. (2013). *Decreto 2851 de 2013 Nivel Nacional.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=55853>

Alcaldía Mayor de Bogotá. (2015). *Decreto 594 de 2015 Alcaldía Mayor de Bogotá, D.C.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=64245&dt=S>

Alcaldía Mayor de Bogotá. (2017a). *Decreto 813 de 2017 Alcaldía Mayor de Bogotá, D.C.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=73425&dt=S>

Alcaldía Mayor de Bogotá. (2017b). *Plan Distrital de Seguridad Vial 2017-2026. 23.*

Alcaldía Mayor de Bogotá. (2019). *Resolución 122 de 2019 Secretaría Distrital de Movilidad.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=83286&dt=S>

Alcaldía Mayor de Bogotá. (2022). *Acuerdo 834 de 2022 Concejo de Bogotá, D.C.*

<https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=121642&dt=S>

Autoridad Nacional de Licencias Ambientales (ANLA). (2013). *Glosario Glosario Glosario Glosario Glosario*. 3–8.

Carreras, E. J. A. (2015). Conceptualizando los siniestros viales desde la perspectiva de la Criminología Vial. *Spokein*, 7, 2–14.

<https://dialnet.unirioja.es/servlet/articulo?codigo=5001975>

Concepto. (2020). *Seguridad - Qué es, concepto de riesgo, amenaza y vulnerabilidad*.

<https://concepto.de/seguridad/>

Coursera. (2021). *Del problema a la acción: las seis fases del análisis de datos* | Coursera.

<https://www.coursera.org/learn/formula-preguntas-para-tomar-decisiones-basadas-en-datos/supplement/Tbsdh/del-problema-a-la-accion-las-seis-fases-del-analisis-de-datos>

Dane. (2007). *PROYECCIONES DE POBLACIÓN 2005-2020 2 de 224 Fecha de elaboración:*

*Septiembre de 2007.*

Distrito, S. de E. del. (2019). *PROTOCOLO DE VICTIMAS DE SINIESTROS* | Secretaría de Educación del Distrito.

[https://www.educacionbogota.edu.co/porta\\_institucional/porta\\_institucional/gestion-educativa/protocolo-de-victimas-de-siniestros](https://www.educacionbogota.edu.co/porta_institucional/porta_institucional/gestion-educativa/protocolo-de-victimas-de-siniestros)

Galán, J. S. (2021). *Siniestro - Qué es, definición y concepto* | 2022 | *Economipedia*.

<https://economipedia.com/definiciones/siniestro.html>

Gonzalez, M. A. M. (2001). *Estadística - Análisis descriptivo*.

Helloauto. (n.d.). *¿Qué es un Vehículo?* - *Glosario de mecánica* | Hello Auto. Retrieved

December 18, 2022, from <https://helloauto.com/glosario/vehiculo>

Infobae. (2022). *Van en aumento las muertes de motociclistas y otros actores viales en lo que va del 2022* - *Infobae*. <https://www.infobae.com/america/colombia/2022/06/02/van-en->

aumento-las-muertes-de-motociclistas-y-otros-actores-viales-en-lo-que-va-del-2022/

Lota, L. F., Rodríguez, M. I., & Alba González Bustos, F. (2021). *Primera Edición Enero 2021*  
*Agencia Nacional de Seguridad Vial-ANSV.*

Martinez, A. (2021). *¿Qué es un Análisis? » Su Definición y Significado [2022].*

<https://conceptodefinicion.de/analisis/>

Martinez, A. (2022). *¿Qué es un Riesgo? » Su Definición y Significado [2022].*

<https://conceptodefinicion.de/riesgo/>

Medellin, A. de. (2019). *0261-vision-cero.pdf.*

MinisteriodeSalud. (2018). *La seguridad vial es un asunto vital de salud pública.*

<https://www.minsalud.gov.co/Paginas/La-seguridad-vial-es-un-asunto-vital-de-salud-publica.aspx>

MinisteriodeTransporte. (2002). *Glosario.* <https://www.mintransporte.gov.co/glosario/v/>

MinisteriodeTransporte. (2022). *ANSV presenta balance oficial de siniestralidad vial 2021.*

<https://www.mintransporte.gov.co/publicaciones/10578/ansv-presenta-balance-oficial-de-siniestralidad-vial-2021/>

Movilidad, S. distrital de. (2021). *Datos Abiertos Bogotá.*

<https://datosabiertos.bogota.gov.co/dataset/siniestros-viales-consolidados-bogota-d-c>

MovilidadBogota. (2018). *Medidas para mejorar la seguridad vial y la movilidad en las localidades de Teusaquillo y Chapinero | Secretaría Distrital de Movilidad.*

<https://www.movilidadbogota.gov.co/web/node/2545>

Narvaez, M. (2022). *12 tipos de análisis de datos y cómo utilizarlos.*

<https://www.questionpro.com/blog/es/tipos-de-analisis-de-datos-2/>

Norza-Céspedes, E. H., Granados-León, E. L., Useche-Hernández, S. A., Romero-Hernández,

- M., & Moreno-Rodríguez, J. (2014). Componentes descriptivos y explicativos de la accidentalidad vial en Colombia: incidencia del factor humano. *Revista Criminalidad*, 56(1), 157–187. [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S1794-31082014000100009&lng=en&nrm=iso&tlng=es](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-31082014000100009&lng=en&nrm=iso&tlng=es)
- Oracle. (2022). *IQR (Rango intercuartílico)*.  
[https://docs.oracle.com/cloud/help/es/pbcs\\_common/PFUSU/insights\\_metrics\\_IQR.htm#PFUSU-GUID-CF37CAEA-730B-4346-801E-64612719FF6B](https://docs.oracle.com/cloud/help/es/pbcs_common/PFUSU/insights_metrics_IQR.htm#PFUSU-GUID-CF37CAEA-730B-4346-801E-64612719FF6B)
- Ortega, C. (2022). *Análisis predictivo: Qué es y cómo implementarlo*.  
<https://www.questionpro.com/blog/es/analisis-predictivo/>
- Paula Rodó. (2021). *Rango intercuartílico - Qué es, definición y concepto | 2022 | Economipedia*. <https://economipedia.com/definiciones/rango-intercuartilico.html>
- Pereira González, A. (2010). Analisis predictivo de datos mediante tecnicas de regresion estadistica. *Article*, 1, 40–61.
- Pérez, P. (2015). *Definición de accidente de tránsito - Qué es, Significado y Concepto*.  
<https://definicion.de/accidente-de-transito/>
- Pérez, P. (2022). *Definición de vehículo - Qué es, Significado y Concepto*.  
<https://definicion.de/vehiculo/>
- Posada Hernández, G. J. (2016). Elementos básicos de estadística descriptiva para el análisis de datos. In *Fondo Editorial Universidad Católica Luis Amigó*.  
[http://www.funlam.edu.co/uploads/fondoeditorial/120\\_Ebook-elementos\\_basicos.pdf](http://www.funlam.edu.co/uploads/fondoeditorial/120_Ebook-elementos_basicos.pdf)
- Publimotos. (2018). *Conozcamos los actores viales*.  
<https://publimotos.com/index.php/especiales/tips-y-curiosidades/2099-conozcamos-los-actores-viales>

Publimotos. (2022). *668 motociclistas muertos en el 2022 | Accidentes en Colombia*.

<https://publimotos.com/mactualidad/19-mundo/colombia/5399-668-motociclistas-muertos-en-el-2022-accidentes-en-colombia>

QuestionPro. (2019). *¿Qué es el análisis descriptivo?*

<https://www.questionpro.com/blog/es/analisis-descriptivo/>

Ramos, G. (2014). *Definición de Analizar - Qué es y Concepto*. <https://enciclopedia.net/analizar/>

RedEmpresarialdeSeguridadVial. (n.d.). *¿Qué es un actor vial? | Red Empresarial de Seguridad*

*Vial*. Retrieved January 2, 2023, from

<https://redempresarial.movilidadbogota.gov.co/?q=content/¿qué-es-un-actor-vial>

República, P. de la. (1996). *Código De Comercio De Colombia*.

Roselli, D. (2017). *El valor social de los motociclistas muertos en Colombia*.

<https://www.redalyc.org/journal/817/81759607026/html/>

Rustom, A. (2015). Estadística descriptiva, Probabilidad e Inferencia - Una visión conceptual y

Aplicada. In *Current Opinion in Critical Care* (Vol. 21, Issue 5).

Salud, A. vamos. (2022). *Consolidado Nacional | Así Vamos en Salud - indicadores en salud*

*normatividad derechos*. <https://www.asivamosensalud.org/indicadores/consolidado-nacional>

Salud, O. M. de la. (2014). *Estadísticas Sanitarias Mundiales*.

[http://apps.who.int/iris/bitstream/10665/131953/1/9789240692695\\_spa.pdf](http://apps.who.int/iris/bitstream/10665/131953/1/9789240692695_spa.pdf)

School, T. (2022). *Tipos de análisis de datos: todas las claves | Tokio*.

<https://www.tokioschool.com/noticias/analisis-datos-tipos/>

Seguros, H. (2022). *¿Son lo mismo siniestro vial y accidente? ¿Cuál es la diferencia? |*

*Hipotecario Seguros*. <https://hipotecarioseguros.com.ar/blog/nota/accidente-y-siniestro->

cual-es-la-diferencia

Semana. (2022). *Compra de motos en Bogotá se disparó 35,9% por el pico y placa extendido.*

<https://www.semana.com/nacion/articulo/atencion-se-disparo-compra-de-carros-y-motos-en-bogota-por-pico-y-placa-extendido/202253/>

Serra, B. R. (2015). *Rango intercuartílico.*

<https://www.universoformulas.com/estadistica/descriptiva/rango-intercuartilico/>

Tatic.net. (2021). *Análisis descriptivo, predictivo y prescriptivo: relaciones, diferencias y*

*relevancia - Tatic - From data to Value.* <https://tatic.net/es/blog/analisis-descriptivo-predictivo-y-prescriptivo-relaciones-diferencias-y-relevancia/>

Ucha, F. (2022). *Definición de Analizar » Concepto en Definición ABC.*

<https://www.definicionabc.com/general/analizar.php>

Velázquez, A. (2022). *¿Qué es el análisis prescriptivo?*

<https://www.questionpro.com/blog/es/analisis-prescriptivo/>

Vial, C. F. de P. (2011). *Actor de la Vía | ANSV.* [https://ansv.gov.co/es/atencion-](https://ansv.gov.co/es/atencion-ciudadania/glosario/actor-de-la)

[ciudadania/glosario/actor-de-la](https://ansv.gov.co/es/atencion-ciudadania/glosario/actor-de-la)

Westreicher, G. (2020a). *Análisis de datos - Qué es, definición y concepto | 2022 |*

*Economipedia.* <https://economipedia.com/definiciones/analisis-de-datos.html>

Westreicher, G. (2020b). *Análisis de datos - Qué es, definición y concepto | 2023 |*

*Economipedia.* <https://economipedia.com/definiciones/analisis-de-datos.html>

**\*Estimado estudiante diligencie todos los campos de la ficha.**

**TITULO COMPLETO DEL TRABAJO DE GRADO**

Siniestralidad vial en motociclistas: Análisis de datos abiertos para la localidad de Kennedy, 2015-2020

**AUTORES**

**Apellidos completos**

Prieto Reina

**Nombres completos**

Cindy Milena

**TUTOR DE TRABAJO DE GRADO**

**Apellidos completos**

Beltrán Gómez

**Nombres completos**

Adán

**EVALUADOR TRABAJO DE GRADO**

**Apellidos completos**

Arias Galindo

León Villanueva

**Nombres completos**

Yeik Exneider

Alexander

**PROGRAMA ACADÉMICO**

**Nombre del programa**

Ingeniería de software

**Tipo de programa  
(marque con una x)**

Pregrado

X

Especialización

Maestría

**Linea de Investigación**

**Modalidad de Grado  
(marque con una x)**

Misión académica

Proyecto de Investigación

X

Emprendimiento

Semillero de Investigación

Diplomado

**CIUDAD**

Bogotá

**AÑO DE PRESENTACIÓN DEL  
TRABAJO DE GRADO**

2023

**NÚMERO DE PÁGINAS**

61

**PALABRAS CLAVES**

**Español**

**Inglés**

Análisis de datos

Data Analysis

Siniestralidad vial

Road Accident

Estadística Descriptiva


Descriptive Statistics

Frecuencia

Frequency

**RESUMEN (Máximo 250 palabras)**

La Organización Mundial de la Salud hace un llamado no solo a las entidades gubernamentales sino también a cada individuo que hace parte de la sociedad, a tomar un tiempo de reflexión sobre la cantidad de pérdidas de vidas humanas en los accidentes viales, por este motivo la presente investigación tiene como fin conocer por medio de una de las ramas de la ingeniería de software, la analítica de datos cuáles son las causas más frecuentes de muerte en siniestralidad vial de la localidad de Kennedy y conocer cómo es el comportamiento de las edades de los motociclistas varones involucrados en estos trágicos eventos.

	<b>LICENCIA Y AUTORIZACIÓN DE LOS ESTUDIANTES PARA PUBLICAR Y COMPARTIR LA CONSULTA Y USO EN EL REPOSITIO.</b>	<b>Código:</b> INV-FM-013
		<b>Versión:</b> 1.0
		<b>Página:</b> 1 de 3

**Bogotá D.C.,** 20 de febrero del 2023

**Términos aplicados a la licencia general para publicar y difundir obras en el repositorio institucional de la Cámara de Comercio de Bogotá.**

El (los) suscrito (s) a continuación declara(n) que es el titular (es) de los derechos de propiedad intelectual objeto de la presente cesión en relación con la obra aquí descrita

Cindy Milena Prieto Reina con C.C. No 1.014.200.365

actuando en calidad de autor(es) de la (obra): Siniestralidad vial en motociclistas: Análisis de datos abiertos para la localidad de Kennedy, 2015-2020

Hago entrega a UNIEMPRESARIAL de una copia trabajo académico en formato digital o electrónico.

**Objeto y fines de la cesión:**


Bajo los términos de la Decisión Andina 351, la Ley 23 de 1982 y demás normas aplicables y con el fin de dar la máxima difusión posible de la obra mencionada a través del repositorio institucional y demás herramientas de visibilidad académica existentes de la Universidad, el (los) autor(es) cede(n) a la Fundación Empresarial Uniempresarial de **forma no exclusiva, limitada y gratuita** para la presente licencia y de la licencia de Creative Commons con que se publica para:

- Publicar la obra de forma abierta y con acceso libre en el Repositorio institucional de la Universidad en el formato en el que lo requiera (impreso, digital, electrónico o cualquier otro conocido o por conocer) y en las demás herramientas de visibilidad académica y científica existentes y futuras por el plazo máximo legal y con ámbito universal. Además, entiende(n) que circulará en internet a un alcance mundial.
- Autorizar a la Universidad para incluir la obra en los índices y buscadores que estimen necesarios para promover su difusión.
- A transferir y convertir la obra a cualquier formato y medio para propósitos de preservación digital.

**Derechos del autor:**

El autor como titular de la obra y por aceptar depositar el documento en el repositorio tiene derecho a:

- A que sea identificada claramente su nombre como autor o como propietario del documento.

	<b>LICENCIA Y AUTORIZACIÓN DE LOS ESTUDIANTES PARA PUBLICAR Y COMPARTIR LA CONSULTA Y USO EN EL REPOSITIO.</b>	<b>Código:</b> INV-FM-013
		<b>Versión:</b> 1.0
		<b>Página:</b> 1 de 3

- Comunicar y dar visibilidad a su obra en esta y posteriores versiones en los medios que considere oportunos.
- Solicitar la retirada de la obra por causa justificada. Para tal efecto es importante que se comunique con el área encargada.
- Recibir notificaciones por parte de terceras personas a fin de reclamación de su obra relativas al derecho de autor.
- A reclamar sus derechos a terceras personas sobre sus derechos autor.

### **Deberes del autor:**

El autor se compromete a:

- A no infringir ningún derecho legal de propiedad industrial, intelectual o de cualquier otra a terceros.
- Garantizar que su obra cuenta con los derechos a la intimidad, a la imagen a otros.
- A asumir toda la responsabilidad legal, eximiendo a la Universidad por los daños causados a terceros.

Para constancia se firma el presente documento en la ciudad de Bogotá, el año 2023 del mes de febrero a los 20 días.

### **FIRMA**



**Firma:**

**C.C.** 1.014.200.365

Otorgando licencia o autorización de uso sobre la misma, para que en los términos de la Decisión Andina 351, la Ley 23 de 1982 y demás normas aplicables, realice los actos de explotación de los derechos patrimoniales y de manera especial, para que la divulgue, reproduzca, comunique al público y la ofrezca en préstamo al público. La presente licencia o autorización se extiende no solo a la fijación en medio o formato físico, analógico o material, sino también al medio virtual, electrónico, óptico, usos de red, Internet, extranet, intranet, repositorio institucional y demás formatos conocidos o por conocer.

El autor de la obra manifiesta de igual manera que la obra objeto de esta licencia o autorización de uso es creación original y que se realizó sin infringir los derechos de autor que le correspondan a terceros.

PARÁGRAFO: Si llegase a presentarse cualquier tipo de reclamación o acción por parte de un tercero en cuanto a los derechos de autor sobre la obra en mención, asumiré la responsabilidad, dejando indemne a UNIEMPRESARIAL y saliendo en defensa de los derechos aquí autorizados.

Para constancia se firma el presente documento en la ciudad de Bogotá, el año 2023 del mes de febrero a los 20 días.

**FIRMA**



Firma:

C.C. 1.014.200.365