|  |  |
| --- | --- |
| **Evaluación de riesgos en seguridad vial, uso de técnicas avanzadas de Machine Learning** | |
|  | |
|  | **Darwin Medardo Padilla Silva**  Máster Universitario en Ciencia de Datos  Área 4 – Aula 1    **Tutor/a de TF**  Joaquin Torres Sospedra  **Profesor/a responsable de la asignatura**  Carlos E. Jiménez    Fecha Entrega  12 de marzo de 2024 |

[Llicència de Creative Commons](http://creativecommons.org/licenses/by-nc-nd/3.0/es/)

Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](http://creativecommons.org/licenses/by-nc-nd/3.0/es/)

**B) GNU Free Documentation License (GNU FDL)**

Copyright © 2014 DARWIN MEDARDO PADILLA SILVA.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

**C) Copyright**

© (Darwin Padilla)

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

Ficha del Trabajo Final

|  |  |
| --- | --- |
| **Título del trabajo:** | Técnicas de aprendizaje automático en datos abiertos de siniestros de tránsito |
| **Nombre del autor/a:** | Darwin Padilla Silva |
| **Nombre del Tutor/a de TF:** | Joaquin Torres Sospedra |
| **Nombre del/de la PRA:** | Carlos E. Jiménez |
| **Fecha de entrega:** | marzo/2024 |
| **Titulación o programa:** | Máster Universitario en Ciencia de Datos |
| **Área del Trabajo Final:** | Área 4 – Aula 1 |
| **Idioma del trabajo:** | Castellano |
| **Palabras clave** | Aprendizaje automático, nivel de siniestralidad, datos abiertos . |
| **Resumen del Trabajo** | |
| La seguridad vial es un aspecto crucial para la gestión del tráfico y transporte en cualquier país, de aquí que los siniestros de tránsito son una preocupación constante en todos los países. Una alta incidencia de accidentes de tránsito impacta en la libre movilidad que tienen los ciudadanos en carreteras, pero además de eso y lo más grave se da en pérdidas humanas, lesiones irreversibles y todo tipo de costos asociados a estos siniestros. El reto está en adoptar enfoques innovadores y tecnológicos que permitan comprender los patrones y características de la siniestralidad en diferentes áreas geográficas. La siniestralidad de tránsito en Ecuador es un problema que año con año se agudiza, por diferentes factores, como la falta de control en escuelas de conducción, entrega irregular de licencias de conducir, falta de cultura vial, conducción imprudente acompañado de mal estado de las carreteras o un clima hostil. Lo que se pretende es no únicamente presentar la estadística descriptiva tradicional, sino más bien ahondar en esta problemática con un enfoque de aprendizaje automático, y que sean los datos los que señalen la problemática antes mencionada. Además de entender y encontrar aquellas variables que sean las que mejor expliquen este tipo de siniestralidad a nivel del país, además de identificar los lugares y horarios que presenten una alta probabilidad de ocurrencia de accidentes de tránsito. Este estudio se centra en la aplicación de técnicas de clustering para agrupar áreas con características similares en términos de siniestralidad vial. | |
| **Abstract** | |
|  | |

Index

[Capítulo 1: Introducción 2](#_Toc166666050)

[1.1. Contexto y justificación del Trabajo 2](#_Toc166666051)

[1.2. Objetivos del Trabajo 3](#_Toc166666052)

[1.3. Impacto en sostenibilidad, ético-social y de diversidad 3](#_Toc166666053)

[1.4. Enfoque y método seguido 4](#_Toc166666054)

[1.5. Planificación del trabajo 5](#_Toc166666055)

[1.6. Breve sumario de productos obtenidos 5](#_Toc166666056)

[Capítulo 2: Materiales y métodos 6](#_Toc166666057)

[Capítulo 3: Resultados 10](#_Toc166666058)

[Capítulo 4: Conclusiones y trabajos futuros 12](#_Toc166666059)

[Capítulo 5: Glosario 13](#_Toc166666060)

[Capítulo 6: Bibliografía 14](#_Toc166666061)

# Capítulo 1: Introducción

En los últimos años, la computación al igual que el manejo de grandes volúmenes de datos ha tenido un crecimiento exponencial, lo cual ha decantado en la revolución de diversas áreas de la vida como la conocíamos, de tal forma que hemos logrado tomar decisiones basados en información y no en “corazonadas” como se lo hacía anteriormente. La seguridad vial no debe ser una excepción, pues la aplicación práctica de técnicas de aprendizaje automático en la gestión y prevención de siniestros de tránsito se presenta como un campo prometedor, que a pesar de tener varios retos por delante, se pueden ofrecer soluciones innovadoras y eficientes para abordar estos desafíos que tienen que ver con la seguridad en las carreteras.

Los siniestros de tránsito son un peligro latente para la sociedad pues sus consecuencias van desde daños materiales a la propiedad privada, lesiones leves y graves y hasta la pérdida de vidas. La complejidad relacionada a estos incidentes se da por los diferentes factores que influyen en la siniestralidad, que pueden ser condiciones meteorológicas, cultura e infraestructura vial, entre otros factores, los cuales han empujado a la búsqueda de enfoques integrales, inteligentes y proactivos, que es donde ingresa el aprendizaje automático.

En resumen lo que se pretende con las técnicas de aprendizaje automático es que las máquinas (computadoras) tienen la capacidad de aprender de patrones y realizar predicciones o clasificaciones basadas en datos históricos. Aplicar este tipo de técnicas en datos de siniestros de tránsito no solo permitirán comprender mejor los factores de riesgo, sino que además facilitarán la identificación de patrones que podrían pasar desapercibidos por los métodos tradicionales o que son intuiciones que expertos en la materia lo detectaron pero no se han podido cuantificar.

## Contexto y justificación del Trabajo

La basta cantidad de información relacionada a siniestros de tránsito permitirá analizar y clasificar grandes conjuntos de datos, lo que resulta fundamental en la investigación posterior a los siniestros de tránsito. La reconstrucción de eventos y la identificación de las causas subyacentes son tareas complejas que pueden beneficiarse enormemente de la automatización proporcionada por los algoritmos de aprendizaje automático. Este enfoque no solo acelera el proceso de investigación, sino que también mejora la precisión al considerar una variedad de factores de forma simultánea.

A pesar de los beneficios evidentes, la implementación exitosa de técnicas de aprendizaje automático en la gestión de siniestros de tránsito, aún presenta varios desafíos, para lo cual la recopilación y la calidad de los datos son fundamentales, y la colaboración entre entidades gubernamentales, empresas privadas y la comunidad en general es esencial para garantizar la disponibilidad de conjuntos de datos significativos, además de fomentar una cultura de conducción responsable al identificar los potenciales factores, lugares y/u horarios de mayor incidencia de siniestralidad vial.

Además, la transparencia y la ética en el uso de datos deben ser consideraciones primordiales para garantizar la aceptación pública y la confianza en estas tecnologías y de forma especial a los resultados que este u otro estudio arrojen pues serán los insights para la toma de decisiones en aspectos como mejora de infraestructura vial, asignación de recursos económicos, operatividad para vehículos de emergencia, entre otros.

El resultado esperado es obtener un modelo que permita clasificar los lugares asociados en temas de siniestralidad, ya sea por lugares u horarios, cuya finalidad práctica se dé en que las autoridades de turno puedan realizar controles preventivos o mejor aún, anticipar la ocurrencia de estos siniestros considerando las áreas de alto riesgo.

## Objetivos del Trabajo

Objetivo General:

Analizar y agrupar áreas geográficas con características similares en términos de siniestralidad vial mediante técnicas de clustering, con el fin de identificar patrones y tendencias que contribuyan a la formulación de estrategias de prevención y mejora de la seguridad vial.

Objetivos Específicos:

* Aplicar técnicas de clustering como K-Means, para agrupar áreas geográficas con características similares en términos de siniestralidad además de evaluar la idoneidad de diferentes enfoques y parámetros de clustering.
* Investigar y analizar patrones emergentes en los grupos identificados, centrándose en las similitudes y diferencias clave entre las áreas agrupadas.
* Establecer relaciones causales entre variables identificadas y la frecuencia de siniestros.
* Evaluar la validez de los grupos formados mediante técnicas de validación interna y externa.

## Impacto en sostenibilidad, ético-social y de diversidad

No aplica

## Enfoque y método seguido

El enfoque que tiene el proyecto es aportar con un producto que permita identificar lugares y/u horarios conflictivos en el tema de siniestros de tránsito, mediante las siguientes estrategias:

1. Identificar el problema: Conocer la problemática e identificar de forma clara lo que se pretende resolver en los siniestros de tránsito.
2. Recopilación y exploración de datos: Identificar las fuentes primarias o secundarias a utilizarse, además de realizar una limpieza y procesamiento de los datos, ya sea que exista información faltante, valores atípicos o potenciales errores de recolección. Además de realizar un análisis exploratorio de la información para conocer su distribución, correlación o patrones existentes.
3. Selección de características: identificar las variables relevantes, con la finalidad de contribuir a la eficiencia computacional y mejora de la interoperabilidad del modelo.
4. División de datos: Segmentar los datos en conjunto de datos para entrenamiento, validación y prueba.
5. Elección, entrenamiento y evaluación del Modelo: Elegir algoritmos de aprendizaje automático apropiados acorde a la naturaleza del problema por resolver, los cuales pueden incluir algoritmos de clasificación, regresión o clustering. Luego se ajustan los hiperparámetros del modelo mediante técnicas como la validación cruzada para mejorar el rendimiento y la generalización. Para luego utilizar el conjunto de entrenamiento para entrenar el modelo seleccionado, ajustando los pesos y parámetros según sea necesario. Para la evaluación del modelo se utilizará el conjunto de validación con la finalidad de evaluar el rendimiento del modelo y realizar ajustes adicionales si es necesario. Finalmente se utilizará el conjunto de prueba para realizar una evaluación final del rendimiento del modelo y validar su capacidad predictiva.
6. Documentar y presentar el modelo: Registrar de manera detallada cada uno de los pasos utilizados y la metodología utilizada. Además de presentar de manera clara los resultados del modelo.

## Planificación del trabajo

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ACTIVIDAD TAREA** | **MARZO** | | | | **ABRIL** | | | | **MAYO** | | | | **JUNIO** | | | | **JUL** | **META PARCIAL** |
| **S 1** | **S 2** | **S3** | **S 4** | **S 1** | **S 2** | **S 3** | **S4** | **S 1** | **S 2** | **S 3** | **S 4** | **S 1** | **S 2** | **S 3** | **S 4** | **S 1** |  |
| Identificar el problema | X | X |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | Documento de a evaluar M1 |
| Recopilación y exploración de datos |  | X | X | X | X | X |  |  |  |  |  |  |  |  |  |  |  | Primeros resultados en IDE |
| Selección de características |  |  |  |  |  | X | X | X | X |  |  |  |  |  |  |  |  | Documento de a evaluar M2 |
| División de datos |  |  |  |  |  |  |  |  | X | X | X | X | X |  |  |  |  | Resultados alcanzado en el IDE |
| Elección, entrenamiento y evaluación del Modelo |  |  |  |  |  |  |  |  |  |  |  | X | X | X | X |  |  | Documento de a evaluar M4 |
| Documentar y presentar el modelo |  | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | Documento final |

## Breve sumario de productos obtenidos

.

# Capítulo 2: Materiales y métodos

Se define un accidente de tránsito como un evento imprevisto que ocurre en las vías, causado por factores humanos (como imprudencia o problemas de salud), mecánicos e inobservancia de las normas de seguridad, y que implica al menos un vehículo en movimiento, puede ser un automóvil, motocicleta o bicicleta. Estos accidentes pueden resultar en pérdidas de vidas o lesiones con personas, de forma directa o indirecta. Las muertes y lesiones por causa de accidentes de tránsito graves son un problema de salud pública, pues se estima que generan más de 1.2 millones de muertes y entre 20 a 50 millones de personas lesionadas en las carreteras de todo el mundo, lo declara la asociación médica mundial – AMM [10].

Los accidentes de tránsito fueron la séptima causa de muerte en ecuador en el año 2022 (3.593 que representa el 4,0%) de las 89.946 muertes ocurridas en ese año, en su gran mayoría comprometen a hombres, 2.973 muertes en accidentes de tránsito son hombres que representan el 82,7% [06].

Para mitigar esta problemática se han realizado diferentes tipos de esfuerzos, como por ejemplo, campañas de seguridad vial, dirigidas a peatones y conductores de todo tipo de vehículos, elaboración e implementación de aparatos tecnológicos que con inteligencia artificial tratan de advertir/evitar accidentes ya sea con sensores implementados en coches (Jabalquinto, 2020) o incluso generando autonomía en la conducción de vehículos [01], entre otros esfuerzos a nivel mundial.

A nivel internacional, se generó un análisis multidimensional de los datos sobre las muertes relacionadas con accidentes de tráfico en la provincia de Buenos Aires, se intentó categorizar los accidentes de tráfico, y luego examinarlos cómo se distribuían en diferentes áreas de la provincia mediante métodos de clasificación jerárquica. Los resultados revelaron una base de segmentación que podría guiar esfuerzos focalizados dirigidos a diferentes grupos de conductores [08].

En 2016, se llevó a cabo un estudio exploratorio, descriptivo y transversal que analizó la incidencia de accidentes de tránsito, su gravedad y mortalidad en las provincias y cantones de Ecuador, específicamente en la región amazónica. Utilizando datos oficiales secundarios, se identificaron los cantones con las tasas de mortalidad más altas. Estos hallazgos proporcionaron información valiosa para identificar áreas de mejora y respaldar la implementación de programas específicos y medidas correctivas en materia de seguridad vial [07].

Existe además un estudio interesante llamado “Comparison of traffic accident injury severity prediction models with explainable machine learning”, estudio en el cual se persigue comprender el o los factores que contribuyen en el aumento de la seguridad en red de tráfico, tratando de explicar los modelos complejos de aprendizaje automático y sus efectos en la predicción de accidentes. Además de construir modelos predictivos utilizando diferentes métodos de aprendizaje automático, y explicar los facotes más contribuyentes utilizando los valores de Shapley que se desarrollaron con base en la teoría de juegos. Árboles de Decisión, Redes Neuronales con Multicapa además de utilizar el perceptrón (MLP), el clasificador de vectores de soporte, el razonamiento basado en casos y el clasificador Naive Bayes para predecir la gravedad de las lesiones en accidentes. Encontrando que que el uso del cinturón, el consumo de alcohol y las infracciones de velocidad eran las características más efectivas y MLP proporcionó la mayor precisión entre todos los modelos predictivos aplicados [05].

Otro estudio que permite identificar un estudio de interés es “Predicción de la gravedad de los accidentes de tráfico mediante un novedoso algoritmo genético multiobjetivo”, el cual se menciona que existe el desafíos para predecir la gravedad de los accidentes de tráfico y se destacan en varios métodos de clasificación utilizados en la literatura existente a la fecha, como son redes neuronales artificiales, máquinas de vectores de soporte y árboles de decisión. Se señala que estos métodos tienen limitaciones, como la falta de interpretación en el caso de las redes neuronales y las SVM, y la baja precisión en el caso de los árboles de decisión convencionales. Se propone un método basado en reglas que se centra en las preferencias del usuario y utiliza un algoritmo genético multiobjetivo para optimizar e identificar reglas. El objetivo es permitir el uso del conocimiento de los usuarios, como la policía de tránsito y los ingenieros de carreteras, y lograr un equilibrio entre diferentes objetivos. El método se evalúa utilizando datos de accidentes de tráfico en la provincia de Teherán, Irán, durante un período de 5 años, y se encuentra que supera a otros métodos de clasificación en términos de precisión y rendimiento de las reglas [04].

Otro gran aporte que se pudo encontrar en el ámbito de la investigación denominado “Un nuevo enfoque de red neuronal espacio-temporal para la previsión de accidentes de tráfico” el cual da como objetivo prevenir de accidentes de tráfico con la finalidad de disminuir la pérdida de vidas, propiedades y económicas. Aunque es un desafío debido a la rareza de los accidentes, sus interdependencias temporales y espaciales, y la influencia del comportamiento humano, el uso creciente de datos ha permitido entrenar predictores de aprendizaje automático. Se menciona que las técnicas de aprendizaje profundo han mostrado mejoras significativas en la predicción, pero aún hay preguntas sobre su aplicabilidad y precisión. La propuesta es un nuevo marco de aprendizaje profundo espacio-temporal llamado XSTNN para predecir el número de accidentes de tráfico en diferentes vecindarios de Madrid, España. Este modelo ha demostrado un mejor rendimiento que otros modelos base y es adaptable a diferentes configuraciones temporales o espaciales. A pesar de los desafíos, el XSTNN se destaca por su capacidad para proporcionar una visión más profunda del problema y adaptarse a diversas situaciones, lo que sugiere que las redes neuronales espacio-temporales son prometedoras para la predicción de accidentes de tráfico en el futuro [03].

Finalmente el estudio llamado “Modelos forestales aleatorios para la predicción de accidentes de motocicleta utilizando big data basados en la conducción naturalista” el cual recolecta información de accidentes de tránsito de motocicletas a través de encuestas con cuestionarios y reportes policiales, incluyen características de los motociclistas y datos contextuales como el entorno vial. Utiliza big data en forma de patrones de trayectoria de vehículos obtenidos a través de GPS, junto con información autoinformada sobre accidentes de carretera y características de los motociclistas para predecir la probabilidad de que un motociclista esté involucrado en un accidente. Emplea un algoritmo de aprendizaje automático basado en Random Forest, utilizando una variedad de características derivadas de los datos de trayectoria, incluyendo características basadas en movilidad, eventos de aceleración, eventos de adelantamiento agresivo y características socioeconómicas de los motociclistas. Además, se determina la importancia relativa de estas características, mostrando que los eventos de adelantamiento agresivo tienen más impacto en los accidentes de motocicleta que otras categorías de características. El modelo desarrollado es útil para identificar motociclistas con mayor riesgo y para implementar medidas de seguridad dirigidas específicamente hacia ellos [09] .

Recopilación de datos

En lo que respecta a la siniestralidad de datos, en Ecuador el tema de tránsito es competencia de mancomunidades en Gobiernos Autónomos descentralizados (GADs), además de las Comisiones de Tránsito y Policía Nacional en ciertas arterias viales, Agencia Municipal de Tránsito de Quito, entre otros estamentos de control vial, todos estas instituciones son articuladas por la Agencia Nacional de Tránsito de la cual es una institución gubernamental, y cuyos datos que son ingresados en el Sistema Nacional de Estadísticas de Tránsito (SINET), mismo que consolida la información a nivel nacional, y entre sus objetivos se encuentran;

* *“Proporcionar información sistemática y actualizada respecto a los siniestros de tránsito, lesionados y fallecidos in situ en el Ecuador.*
* *Dar a conocer a los usuarios, las variables relevantes de siniestralidad vial, que constituye un insumo para la planificación de estrategias y política pública en materia de seguridad vial.*
* *Proveer al sector público y privado, una herramienta para el análisis de siniestralidad de tránsito.”*

La unidad de investigación que se presentan son los eventos de tránsito (siniestros y víctimas en sitio) registrados a nivel nacional cada mes, con desagregación geográfica a nivel provincial, cantonal y parroquial.

Además de indicar que el universo está comprendido por siniestros de tránsito ocurridos en todo el país, así mismo como el conjunto de personas lesionadas y fallecidas ocurrido y registrado en el lugar de los hechos.

La fuente de información es la página web institucional de la Agencia Nacional de Tránsito de la cual son las siguientes características:

* Link: <https://www.ant.gob.ec/visor-de-siniestralidad-estadisticas/>
* Periodo de análisis: enero de 2017 a marzo de 2024.
* Total de registros: 165017
* Fecha de corte: 16 de abril de 2024

A continuación se presentan las variables que contiene el data set:

Tabla No. 1 Descripción de variables del data set original

|  |  |
| --- | --- |
| **VARIABLE** | **DESCRIPCIÓN** |
| NUMERO | Identificador único de registro de información, secuencial |
| ANIO | Indica el año del incidente, es una variable numérica. |
| SINIESTROS | Variable alfanumérica con texto y número, es un código único por cada siniestro. |
| LESIONADOS | Número de lesionados, variable numérica. |
| FALLECIDOS | Número de fallecidos, variable numérica. |
| ENTE\_DE\_CONTROL | Describe la agencia de control, es una variable categórica. |
| LATITUD\_Y | Coordenada de latitud, variable numérica. |
| LONGITUD\_X | Coordenada de longitud, variable numérica. |
| DPA\_1 | Código de área administrativa, variable categórica. |
| PROVINCIA | Provincia, variable categórica. |
| DPA\_2 | Código de área administrativa 2, variable categórica. |
| CANTON | Cantón, variable categórica. |
| DPA\_3 | Código de área administrativa 3, variable categórica. |
| PARROQUIA | Parroquia, variable categórica. |
| DIRECCION | Dirección, variable categórica. |
| ZONA\_PLANIFICACION | Zona de planificación, variable categórica. |
| ZONA | Zona, variable categórica. |
| ID\_DE\_LA\_VIA | ID de la vía, variable categórica. |
| NOMBRE\_DE\_LA\_VIA | Nombre de la vía, variable categórica. |
| UBICACION\_DE\_LA\_VIA | Ubicación de la vía, variable categórica. |
| JERARQUIA\_DE\_LA\_VIA | Jerarquía de la vía, variable categórica. |
| FECHA | Fecha del incidente, variable categórica. |
| HORA | Hora del incidente, variable categórica. |
| PERIODO\_1 | Periodo del día, variable categórica. |
| PERIODO\_2 | Otro periodo del día, variable categórica. |
| DIA\_1 | Día de la semana, variable categórica. |
| DIA\_2 | Otro día de la semana, variable categórica. |
| MES\_1 | Mes, variable categórica. |
| MES\_2 | Otro mes, variable categórica. |
| FERIADO | Indicador de feriado, variable categórica. |
| CODIGO\_CAUSA | Código de causa, variable categórica. |
| CAUSA\_PROBABLE | Causa probable, variable categórica. |
| TIPO\_DE\_SINIESTRO | Tipo de siniestro, variable categórica. |
| TIPO\_DE\_VEHICULO\_1 | Tipo de vehículo involucrado, variable categórica. |
| SERVICIO\_1 | Servicio del vehículo, variable categórica. |
| AUTOMOVIL a TRICIMOTO | Conteo de tipos de vehículos, variables numéricas. |
| SUMA\_DE\_VEHICULOS | Suma de vehículos involucrados, variable numérica. |
| TIPO\_ID\_1 | Tipo de identificación, variable categórica. |
| EDAD\_1 | Edad, variable numérica. |
| SEXO\_1 | Género, variable categórica. |
| CONDICION\_1 | Condición del participante, variable categórica. |
| PARTICIPANTE\_1 | Rol del participante, variable categórica. |
| CASCO\_1 a CINTURON\_1 | Indicadores de seguridad, variables categóricas. |

*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a marzo de 2024*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Una vez descritas las variables originales y con las cuales se pretende trabajar, nos limitaremos a identificar la estadística descriptiva para posteriormente utilizar el o los modelos de machine learning, con el propósito de cubrir el objetivo del TFM.

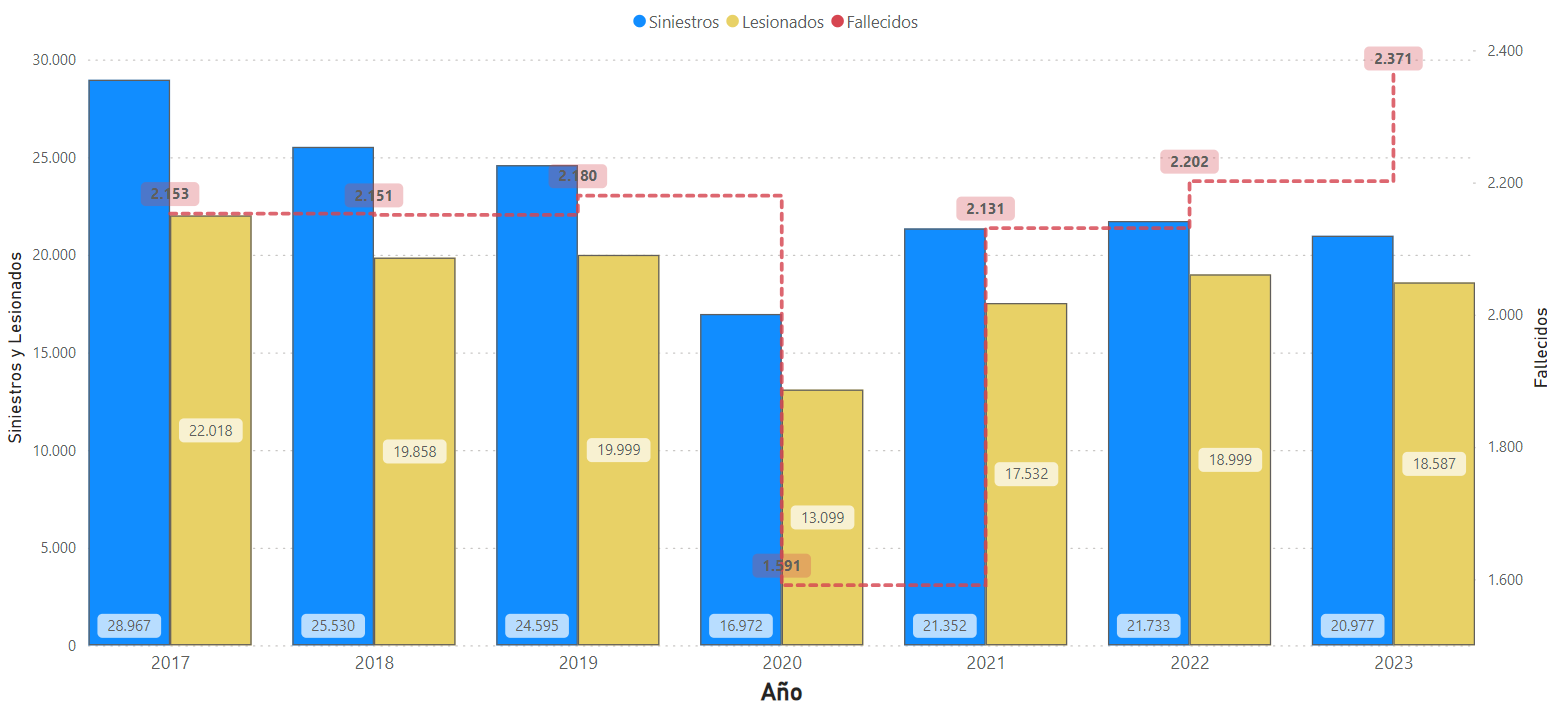
Con el propósito de trabajar con años cerrados, se utilizará información desde el año 2017 al año 2013, lo cual representa 160,126 registros en total con la fecha de corte 16 de abril de 2024.

# Capítulo 3: Resultados

Estadística descriptiva:

A partir de los 160,126 registros entre los años 2017 a 2023 se puede identificar lo siguiente:

Gráfico No. 1 Total de Siniestros, lesionados y fallecidos en Ecuador



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Entre los año de 2017 a 2023 se totalizaron 160.126 siniestros de tránsito a nivel nacional, los cuales dejaron 130.092 personas lesionadas y 14.779 personas fallecidas, todo esto in situ, estas cifras no muestran la realidad en un contexto de la gravedad que se tiene a nivel nacional, razón por la cual se la expresará en cifras diarias, dicho de otras forma se tuvo en promedio diario 63 siniestros de tránsito, mismos que como resultado dejaron 51 personas lesionadas y 6 personas fallecidas.

Además de indicar que el año 2017 es el año con un mayor número de siniestros y lesionados en promedio diario, 79 y 60 registros respectivamente, sin embargo la mayor cantidad de fallecidos en promedio diario se presentó en el año 2023, dejando un promedio de 7 personas fallecidas de forma diaria, así:

Tabla No. 2 Promedio diario de siniestros de tránsito, personas lesionadas y fallecidas

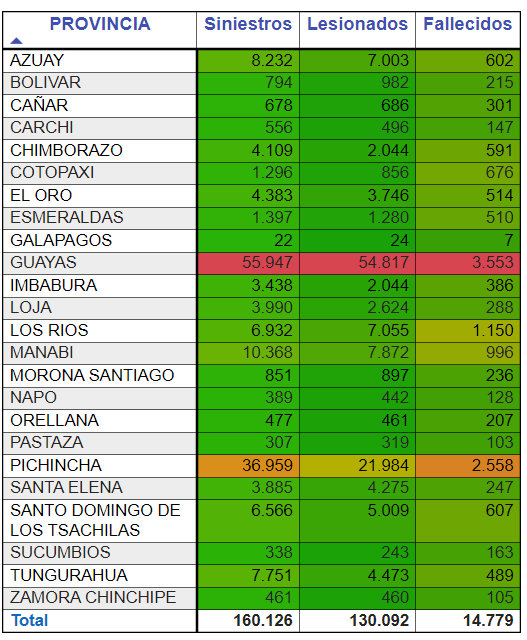
|  |  |  |  |
| --- | --- | --- | --- |
| **Año** | **Siniestros** | **Lesionados** | **Fallecidos** |
| 2017 | 79 | 60 | 6 |
| 2018 | 70 | 54 | 6 |
| 2019 | 67 | 55 | 6 |
| 2020 | 46 | 36 | 4 |
| 2021 | 58 | 48 | 6 |
| 2022 | 60 | 52 | 6 |
| 2023 | 58 | 51 | 7 |

*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Tabla No. 3 Total de siniestros viales, lesionados y fallecidos por Provincia

*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

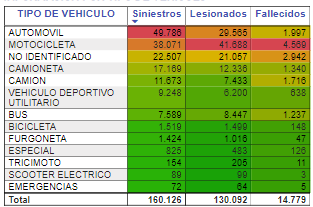
*Elaboración del autor*

Las provincias con mayor cantidad de siniestros viales y lesionados en el periodo de análisis son las Provincias de Guayas, Pichincha y Manabí, mismas que suman el 64.5% y el 65.1%, respectivamente, a nivel nacional. En cuanto al número de fallecidos son las provincias de Guayas, Pichincha y Los Ríos, sumando el de tal forma que estas provincias suman el 49.1% a nivel nacional, ver tabla anterior.

Una de las principales características que llama la atención es que de las 130,092 personas lesionadas el 62.5% ocurrieron en la zona urbana, y el restante 37.5% en la zona rural. Mientras que de las 14,779 personas fallecidas la mayoría corresponden a la zona rural, con el 63.6%, y el restante 36.4% fueron de la zona rural.

De los 160.126 siniestros viales se tiene que el automóvil y la motocicleta contienen el 54,9%, con el 31.1% y el 23.8% respectivamente. En cuanto a las personas lesionadas (130.092 personas) se invierten los papeles pues se tiene que la mayoría se genera en personas que utilizaban motocicletas, con el 32.1% seguido de personas que su siniestro se generó en un automóvil con el 22.7%, totalizando 54,8%. En lo que respecta a las 14.779 personas fallecidas, se presentó que la mayoría (30,9%) sucedió en motocicletas, seguida por vehículos no identificados con el 19.9%.

Tabla No. 4 Total de siniestros viales, lesionados y fallecidos por tipo de vehículo



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

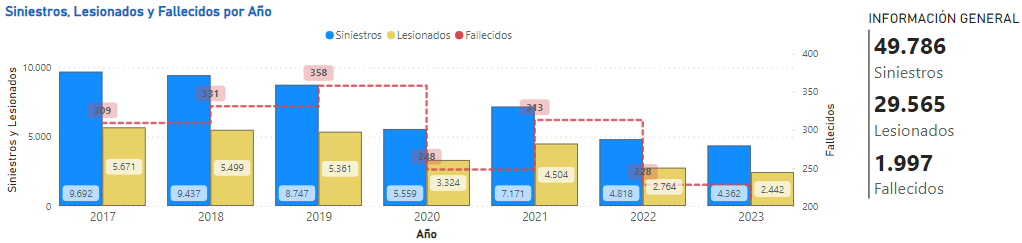
*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

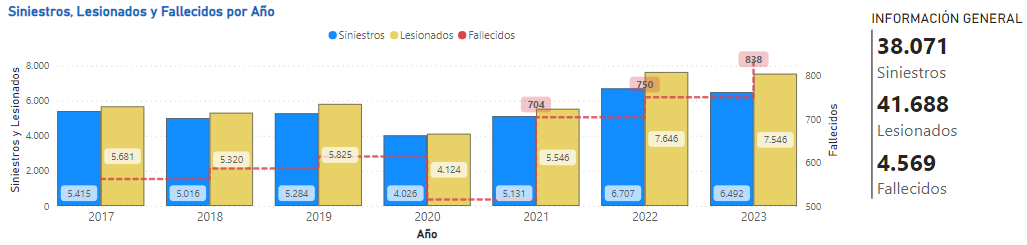
Considerando el análisis anterior, es importante poner en conocimiento la información que se genera por los tipos de vehículos con mayor cantidad de siniestros, lesionados y fallecidos, es así que se toman las tres principales categorías: automóvil, motociclistas y no identificados.

Gráfico No. 2 Total de Siniestros, lesionados y fallecidos en Ecuador, por tipo de vehículo, automóviles (a), motocicletas (b) y no identificados (c)

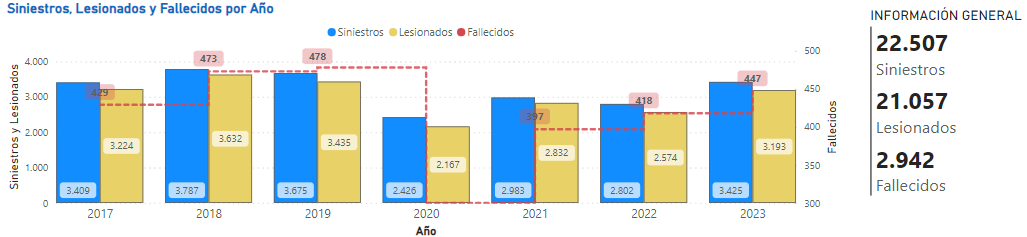
a)



b)



c)



*Fuente: Data set original de información*

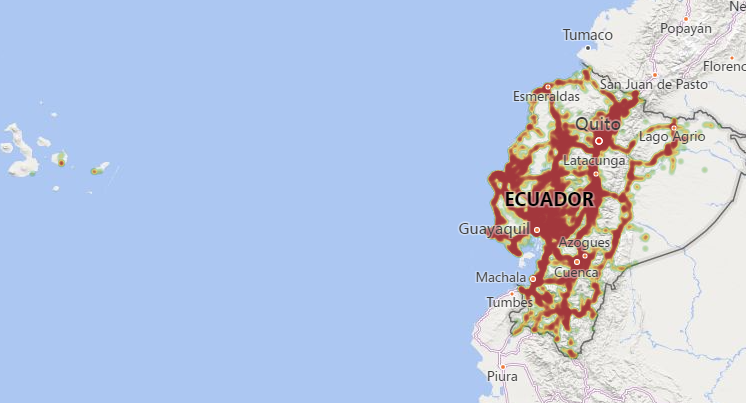
*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

La lectura de los cuadros anteriores es que la evolución de siniestros viales, lesionados y fallecidos ocurridos en vehículos presenta una importante disminución, a pesar de ser la de mayor frecuencia a lo largo de los datos históricos, mientras que con la información registrada en motocicletas, se puede apreciar todo lo contrario, pues posterior al año 2020, post pandemia, existen una tendencia al alza de este tipo de datos. En cuanto al tipo de vehículo no identificado se tiene que salvo el año 2020, pandemia, una distribución uniforme, es decir ausencia de incremento o decremento de las cifras.

La distribución que contiene la información a nivel nacional depende mucho del análisis por siniestros viales, personas lesionadas o fallecidas, para este caso se representa la cantidad de personas lesionadas, y su distribución a nivel nacional mediante un mapa de puntos calientes, lo que se puede observar es que las principales ciudades se presenta en las ciudades más pobladas como es Guayaquil, Quito, Manta, entre otras, así:

*Gráfico No. 3 Distribución espacial de personas lesionadas en Ecuador*

*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

De las personas lesionadas el 64,5% son hombres, además existe un 19% son mujeres y el restante 16,5% corresponde a personas no identificadas en su sexo.

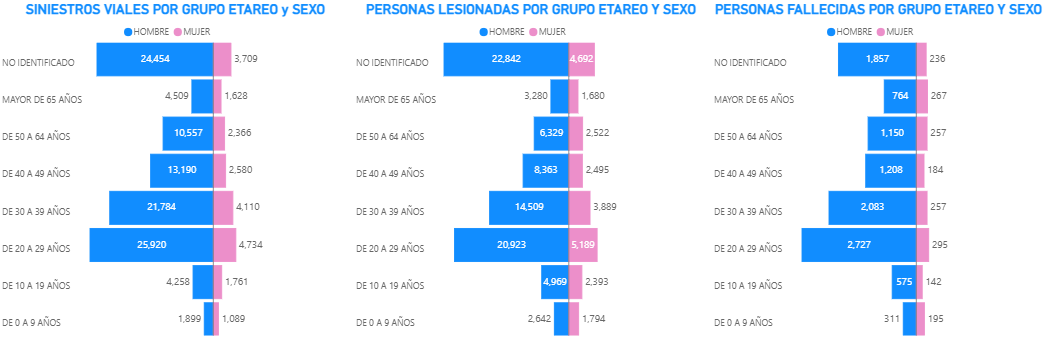
En cuanto a las personas fallecidas se tiene que un 15,4% corresponde a personas que no han podido ser identificadas, sin embargo el 12,4% corresponden a víctimas mujeres y el mayor porcentaje corresponde a personas fallecidas del sexo hombre con el 72,2%.

Además es importante destacar que tanto las personas que se encuentran entre los 20 y 39 años, personas jóvenes, son las más propensas para sufrir algún tipo de lesión o incluso fallecimiento por causa de un siniestro vial, tanto para hombres como para mujeres, esto se puede identificar de mejor forma en el siguiente gráfico.

Además de indicar que en lo que corresponde al tipo de siniestros vial el más frecuente se da en choque lateral (28,0%), atropellos (14,0%) y estrellamiento (13,0%), captando el 55% del total de siniestros. Para las personas lesionadas según el tipo de siniestro el más frecuente es choque lateral (31,1%), seguido por atropellos (16,6%) y pérdida de pista (9,0%) totalizando el 56,7% de personas lesionadas por estos tipos de siniestros. En lo que corresponde a personas fallecidas tres tipos de siniestros concentran el 49%, así los atropellos lideran con el 20,5%, choque frontal con el 15,1% y el 13,4% se da en perdida de pista. Esto se puede observar en la tabla inferior.

*Gráfico No. 4 Distribución por sexo, según grupo etario de (a) personas lesionadas y (b) personas fallecidas en Ecuador*

*a) b)*



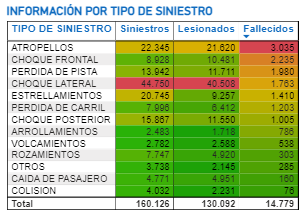
*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Tabla No. 5 Total de siniestros viales, lesionados y fallecidos por tipo de vehículo



*Fuente: Data set original de información*

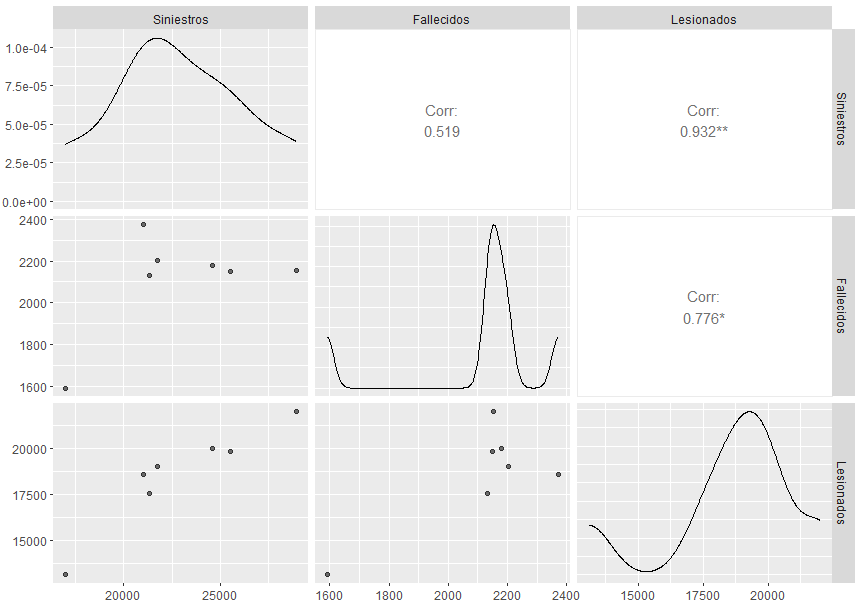
*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Hasta ahora se han evaluado información numérica, el cual puede generar cierta incertidumbre al momento de sacar conclusiones, pues la cantidad de siniestros de tránsito, lesiones y fallecimientos en personas tienen una correlación año por año, ver el siguiente gráfico.

*Gráfico No. 5 Gráfico de correlación lineal entre siniestros viales, personas lesionadas y fallecidas.*



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Como se puede apreciar en el gráfico de correlación existente entre las tres variables, total de siniestros de tránsito, personas lesionadas y fallecidas, se tiene que entre las dos primeras variables existe una correlación lineal alta 0,932, al analizar la información entre las personas lesionadas y fallecidas, se tiene una correlación lineal de 0,776, mientras que entre el total de siniestros y personas fallecidas es de 0,519.

Uno de los principales factores que puede afectar el análisis numérico es el parque automotor y la población que tiene cada uno de los territorios en análisis, por esta razón se utilizarán dos tasas que permiten realizar una homologación en este tipo de medidas, así:

* Tasa de fallecimientos por cada 1’000,000 habitantes.
* Tasa de fallecimientos por cada 100 lesiones.

**Tasa de fallecimientos por cada 1’000,000 habitantes**

Fórmula:

Donde:

*TF: Tasa de personas fallecidas por cada millón de habitantes*

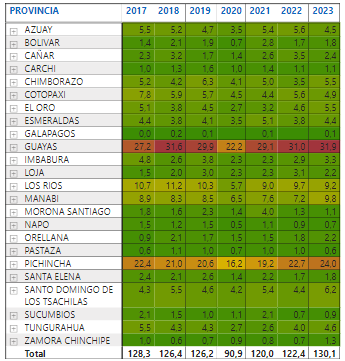
*# PFST: Número personas fallecidas en siniestros de tránsito en un año*

*PT: Población total en ese mismo año*

*Interpretación: Se generaron # fallecimientos en siniestros de tránsito por cada millón de habitantes.*

Calcular esta tasa tiene por objetivo comparar eventos o incidencias en poblaciones de diferentes tamaños, como es nuestro caso, además de normalizar los datos brutos, ya que al analizar los datos absolutos son difíciles de comparar entre una región u otra, pudiendo generar conclusiones erradas y por ende a tomar decisiones que no aporten a la construcción de política pública que permita mitigar o combatir la problemática en análisis, además de la facilidad de comunicación e identificación de tendencias, en nuestro caso es conocer la cantidad de fallecidos por cada millón de habitantes de cada provincia que conforma el Ecuador, así tenemos:

Tabla No. 6 Tasa de fallecidos en accidentes de tránsito por cada millón de habitantes, por año de ocurrencia



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

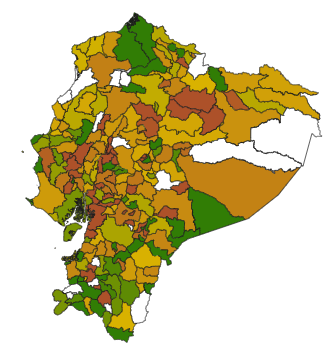
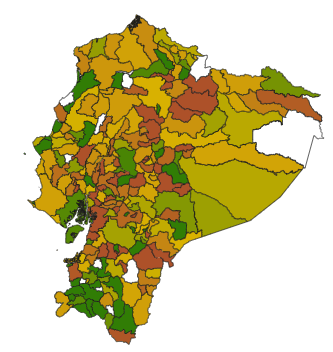
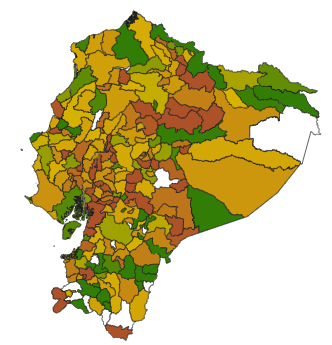
*Elaboración del autor*

Al observar la tabla anterior se tiene que todos los años de estudio sobrepasan las 3 cifras, excepto el año 2020, pandemia, es decir que por cada millón de habitantes han ocurrido más de 120 personas fallecidas por accidentes de tránsito.

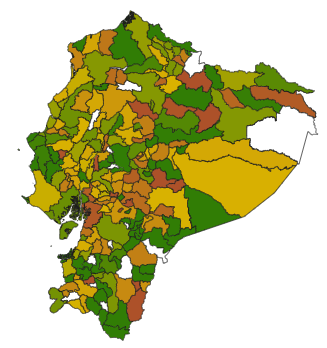
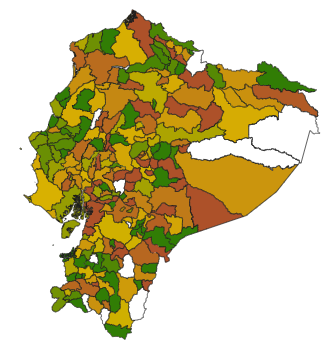
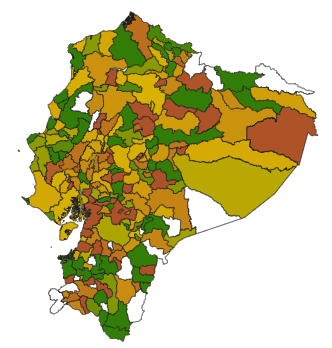
Además de visualizar que las provincias con la mayor tasa en la serie inicial, años 2017 al 2019, son Guayas, Pichincha y Los Ríos, para el año 2020, a nivel nacional existe un decremento significativo en todas las provincias del país, sin embargo a partir del año 2021 las cifras conservan una tendencia incremental en las Provincias de Guayas y Pichincha, y una distribución uniforme en la provincia de Los Ríos. A continuación se presenta la distribución parroquial de las tasas de fallecidos por accidentes de tránsito por cada millón entre los años 2017 a 2023.

*Gráfico No. 6 Mapas cantonales - Tasa de fallecidos por cada millón de habitantes*

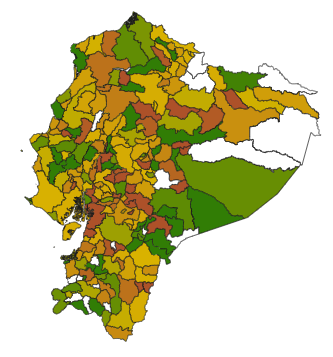
a) Año 2017 b) Año 2018 c) Año 2019

d) Año 2020 e) Año 2021 f) Año 2022

g) Año 2023



*Fuente: Data set original de información*

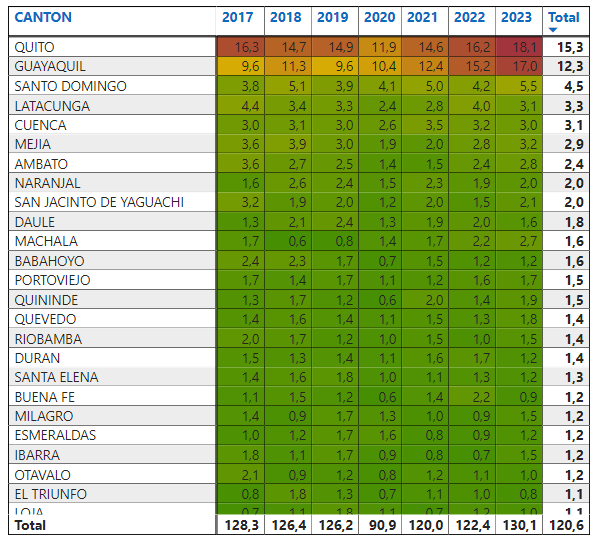
*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

*\* El color verde significa la tasa de menor valor, mientras que el color rojo representa la tasa con mayor valor.*

Tabla No. 7 Top de tasa de fallecidos en accidentes de tránsito por cada millón por año y Cantón de ocurrencia



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

Lo que se puede apreciar de forma superficial es que la distribución cantonal que se tiene en los diferentes años de análisis existen distribuciones diferentes, sin embargo existen cantones marcadas que presentan siempre una tasa alta de fallecimientos por cada millón de habitantes, siendo las principales, Quito, Guayaquil, Santo Domingo, Latacunga, Cuenca, Mejía, entre otros cantones, los cuales se detallan a continuación.

**Tasa de fallecimientos por cada 100 lesiones**

Fórmula:

Donde:

*TFxL : Tasa de personas fallecidas por cada 100 personas lesionadas*

*# F : Número de personas fallecidas en siniestros de tránsito*

*# L : Número de personas lesionadas en siniestros de tránsito*

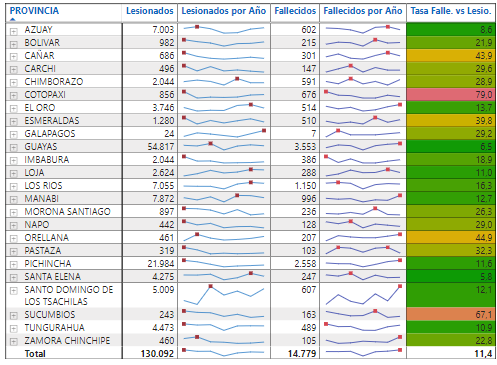
Interpretación: Se han registrado ## fallecidos por cada 100 personas lesionadas en siniestros de tránsito.

El objetivo que persigue este indicador es poner en conocimiento la gravedad de las personas fallecidas por cada 100 personas lesionadas, es decir, pueden existir lugares donde existan tasas menores, es decir que los siniestros de tránsito no sean graves, mientras que en otros lugares como Cotopaxi, que es la provincia con la mayor TFxL, lo cual no era apreciable si únicamente se utilizan los valores absolutos.

Al analizar la tabla que se indica en el párrafo anterior, se puede apreciar, como ya se mencionó, que la Provincia de Cotopaxi presenta la tasa más alta de este indicador en los años de estudio, es decir al tener 676 personas fallecidas y realizar el cociente con las 856 personas lesionadas en el mismo periodo de tiempo, da como resultado el valor de 79, considerando la interpretación se tiene que por cada 100 personas lesionadas en siniestros de tránsito han fallecido 79 personas, lo cual es un valor extremadamente algo y da señales que cuando existen siniestros de tránsito son realmente graves. Le sigue la provincia de Sucumbíos, con una tasa de 67,1, la cual resulta del cociente entre las 163 personas fallecidas en siniestros de tránsito para los 243 siniestros de tránsito acaecidos en los años de estudio.

Para mayor detalle se pone en conocimiento la siguiente tabla:

Tabla No. 8 Tasa de fallecidos por 100 personas lesionadas por provincia



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

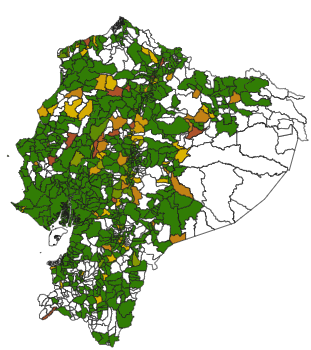
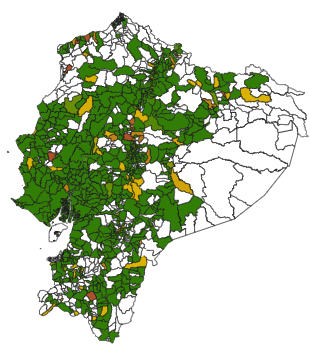
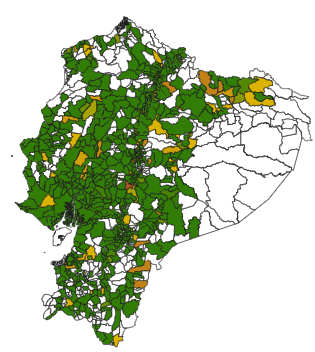
*Elaboración del autor*

La tabla anterior permite conocer la gravedad de los siniestros de tránsito acaecidos en cada una de las provincias, además de identificar la tendencia que estas presentan en el tiempo, años, con un punto en el valor de mayor frecuencia para personas lesionadas como para fallecidos.

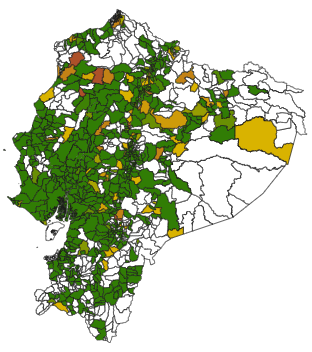
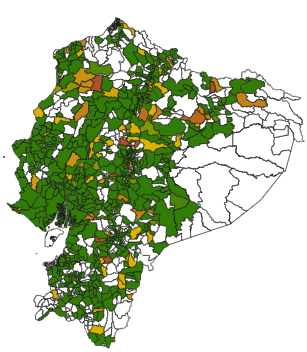
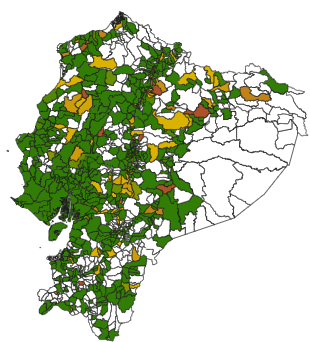
Adicional se pone en conocimiento la tasa explicada en párrafos anteriores en temas territoriales, de tal forma que se puede apreciar la distribución que esta tiene a lo largo de tiempo en todo el territorio nacional, siendo cantones específicos que tienen una incidencia considerable en varias parroquias del Ecuador, así:

*Gráfico No. 7 Mapas parroquiales - Tasa de fallecidos por cada 100 lesionados*

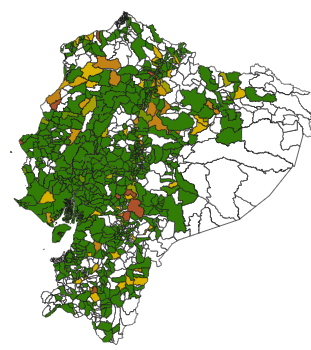
*a) Año 2017 b) Año 2018 c) Año 2019*



*d) Año 2020 e) Año 2021 f) Año 2022*



*g) Año 2023*



*Fuente: Data set original de información*

*Periodo de análisis: enero 2017 a diciembre de 2023*

*Fecha de corte: 16 abril de 2024*

*Elaboración del autor*

# Capítulo 4: Conclusiones y trabajos futuros

# Capítulo 5: Glosario

# Capítulo 6: Bibliografía

[01] *Coches autónomos 2024 | ¿Qué son y cómo funcionan?* (s. f.). Carwow. https://www.carwow.es/coches-autonomos#gref

[02] Clarke, D. D., Forsyth, R., & Wright, R. (1998). Machine learning in road accident research: decision trees describing road accidents during cross-flow turns. *Ergonomics*, *41*(7), 1060-1079. https://doi.org/10.1080/001401398186603

[03] De Medrano, R., & Aznarte, J. L. (2021). A New Spatio-Temporal Neural Network Approach for Traffic Accident Forecasting. *Applied Artificial Intelligence*, *35*(10), 782-801. https://doi.org/10.1080/08839514.2021.1935588

[04] Hashmienejad, S. H., & Hasheminejad, S. M. H. (2017). Traffic accident severity prediction using a novel multi-objective genetic algorithm. *International Journal Of Crashworthiness*, *22*(4), 425-440. https://doi.org/10.1080/13588265.2016.1275431

[05] Jabalquinto, J. (2020, 30 septiembre). Todas las nuevas tecnologías y aplicaciones para evitar accidentes de tráfico. *Autopista*. https://www.autopista.es/noticias-motor/todas-nuevas-tecnologias-aplicaciones-evitar-accidentes-trafico\_209425\_102.html

[06] INEC - INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS. (2022). *Registro Estadístico de Defunciones Generales de 2022*. https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion\_y\_Demografia/Defunciones\_Generales\_2022/Principales\_resultados\_EDG\_2022.pdf.

[07] *Galarza, L., Merino, P., Algora, A. & Gómez, A. (2017), Estudio geoespacial de los accidentes de tránsito en la Región Amazónica Ecuatoriana. CienciaAmérica 6 (2), 21-26*

[08] Montes, s. a. m., m. (2021). Muertes por siniestros de tránsito en la provincia de buenos aires en 2017: un análisis mediante métodos de clasificación jerárquica. *Revista argentina  de salud pública*, *10*, 13-34. http://www.scielo.org.ar/pdf/rasp/v13/1853-810X-rasp-13-51.pdf

[09] Outay, F., Adnan, M., Gazder, U., Baqueri, S. F. A., & Awan, H. H. (2023). Random forest models for motorcycle accident prediction using naturalistic driving based big data. *International Journal Of Injury Control And Safety Promotion*, *30*(2), 282-293. https://doi.org/10.1080/17457300.2022.2164310

[10] *WMA - The World Medical Association-Declaración de la AMM sobre los Accidentes de Tránsito*. (s. f.). WMA - The World Medical Association-Declaración de la AMM Sobre los Accidentes de Tránsito. https://www.wma.net/es/policies-post/declaracion-de-la-amm-sobre-los-accidentes-de-transito/