



**ESCUELA DE NEGOCIOS**

**MAESTRÍA EN INTELIGENCIA DE NEGOCIOS Y CIENCIA DE DATOS**

**ANÁLISIS SITUACIONAL DE MUERTES VIOLENTAS EN EL 2022 Y  
PREDICCIÓN DE MUERTES VIOLENTAS REGISTRADAS EN EL ECUADOR  
DURANTE EL PERIODO 2015-2022, MEDIANTE EL USO DE MODELOS DE  
PRONÓSTICO DE SERIES TEMPORALES.**

**Profesor  
Ing. Mario González. Pd. D**

**Autor  
María Belén Díaz Gómez  
Geovanny Andrés Pantoja Orbe**

**2023**

## RESUMEN

El presente proyecto tiene como propósito fundamental realizar un análisis situacional de las muertes violentas en el Ecuador y realizar su pronóstico, utilizando modelos de predicción de series de tiempo, que permitan identificar, analizar y cuantificar los diversos factores que inciden en la ocurrencia de muertes violentas en el país.

Al realizar el análisis exploratorio se identifica a nivel detallado el comportamiento de las muertes violentas en el Ecuador; como las muertes por año, por provincia, por tipo de muerte, por género, por etnia, estado civil, nacionalidad, etc.

Se realiza un análisis situacional de las muertes violentas identificando en algunos casos que existe correlación con variables de información socioeconómica como la pobreza extrema, trabajo inadecuado y analfabetismo que sugieren una posible relación directa para que las muertes violentas incrementen o decrementsen en el Ecuador.

Adicionalmente, se realiza un análisis predictivo de series temporales, mediante el cual se comparan los modelos: Aditivo (Prophet), ARIMA y SARIMA, evidenciando en los resultados que el modelo Aditivo es el más efectivo ya que presenta el más bajo Error Absoluto Medio.

Adicionalmente, los hallazgos presentados en este proyecto permitieron predecir de manera óptima las posibles muertes violentas en el país, mediante el uso de la tecnología específicamente realizando razonamiento probabilístico, para poder catalogar las diferentes causalidades, variables y como resultado obtener una correlación que permita asignar los recursos necesarios para contrarrestar la delincuencia en el país de los próximos tres años.

**Palabras Clave:** muertes violentas, series de tiempo, análisis exploratorio, análisis situacional, aditivo, ARIMA, SARIMA.

## **ABSTRACT**

The main purpose of this project is to carry out a situational analysis of violent deaths in Ecuador and make a forecast of violent deaths models using time series prediction, which allows the identification, analysis and quantification of the various factors that affect the occurrence of violent deaths in equator.

When carrying out the exploratory analysis, the behavior of violent deaths in Ecuador is identified at a detailed level; such as deaths per year, by province, by type of death, by gender, by ethnicity, marital status, nationality, etc.

A situational analysis of violent deaths is carried out, identifying in some cases that there are consequences with socioeconomic information variables such as extreme poverty, inadequate work and illiteracy that suggests a possible direct relationship for violent deaths to increase or decrease in Ecuador.

Additionally, a predictive analysis of time series is carried out, through which the models are compared: Additive (prophet), ARIMA and SARIMA, evidencing in the results that the Additive model is the most effective since it presents the lowest Mean Absolute Error.

In addition, the findings presented in this project made it possible to optimally predict possible violent deaths in the country, through the use of technology specifically developing a probabilistic study, to be able to catalog the different causalities, variables and as a result obtain a modification that allows the necessary resources to counteract crime in the country for the next three years.

**Keywords:** violent deaths, time series, exploratory analysis, situational analysis, additive, ARIMA, SARIMA.

## ÍNDICE DEL CONTENIDO

RESUMEN .....	2
ABSTRACT .....	3
1. INTRODUCCIÓN .....	1
2. REVISIÓN DE LITERATURA.....	3
3. IDENTIFICACIÓN DEL OBJETO DE ESTUDIO .....	10
4. PLANTEAMIENTO DEL PROBLEMA .....	11
5. OBJETIVO GENERAL .....	12
6. OBJETIVOS ESPECÍFICOS .....	12
7. JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA .....	13
8. RESULTADOS.....	40
8.1 ANÁLISIS DE MODELO ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS .....	40
8.2 PROPUESTA DE SOLUCIÓN .....	53
8.2.1 IMPLICACIONES ORGANIZACIONALES.....	53
8.2.2 DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN	55
9. CONCLUSIONES Y RECOMENDACIONES .....	59
9.1 CONCLUSIONES .....	59
9.2 RECOMENDACIONES .....	60
REFERENCIAS.....	62

## ÍNDICE DE TABLAS

Tabla 1: Transformación de Variables.....	18
Tabla 2: Categorización de variables Muertes Violentas.....	19
Tabla 3:Categorización de variables Tasa de Homicidios cada 100.000 hab. .	20
Tabla 4: Categorización de variables Tasa de Pobreza Extrema por Ingresos.	20
Tabla 5:Categorización de variables Tasa de Empleo Inadecuado.....	21
Tabla 6: Categorización de variables Tasa de Analfabetismo.....	21
Tabla 7:Comparativo de Muertes Violentas por Año y Provincias del Ecuador	24
Tabla 8: Ventajas y desventajas de los modelos Prophet, SARIMA y ARIMA:	36
Tabla 9:Cuadro comparativo ARIMA - Aditivo de Prophet - SARIMA.....	38
Tabla 10:resultado de los modelos.....	39
Tabla 11: Valores pronosticados y límites inferior y superior del intervalo de confianza .....	40
Tabla 12: Predicción Nacional.....	50
Tabla 13: Predicciones en la provincia de Guayas.....	50
Tabla 14: Predicciones en la provincia de Manabí .....	51
Tabla 15: Predicciones en la provincia de Pichincha .....	51
Tabla 16: Predicciones en la provincia de Los Ríos .....	51
Tabla 17: Predicciones en la provincia de Esmeraldas .....	52
Tabla 18: Comparación Predicción VS provincial (5) .....	52
Tabla 19: Homicidios de América Latina .....	67
Tabla 20:Socio-economic, built environment and mobility conditions associated with crime .....	68
Tabla 21:Identificación de patrones para redes de eventos espacio temporales .....	69
Tabla 22: Prediciendo el crimen en ciudades intermedias .....	70
Tabla 23: Aplicación de minería de datos en datos abiertos de Ecuador .....	71
Tabla 24: El homicidio y su caracterización situacional.....	72
Tabla 25: Detection of Homicide Trends in Colombia Using Machine Learning	73
Tabla 26: Modelación Del Número De Homicidios Vía.....	74

## ÍNDICE DE FIGURAS

Figura 1: Gráfico de cajas .....	18
Figura 2: Muertes violentas por año .....	22
Figura 3: Mapas Coropléticos de Muertes Violentas .....	23
Figura 4: Muertes violentas de las provincias con mayor participación a nivel nacional.....	25
Figura 5: Muertes violentas por Tipo de muerte en el año 2022 .....	26
Figura 6: Muertes violentas por Tipo de Área.....	26
Figura 7: Muertes violentas por Género en el año 2022 .....	27
Figura 8: Estado Civil. ....	27
Figura 9: Muertes violentas por día de la semana.....	28
Figura 10: Muertes violentas por fecha .....	29
Figura 11: Matriz de Correlación Nacional .....	30
Figura 12: Matriz de Correlación Provincial.....	31
Figura 13: Predicción diaria muertes violentas.....	41
Figura 14: Predicción anual muertes violentas.....	41
Figura 15: Predicción día de muertes violentas.....	42
Figura 16: Predicción mes de muertes violentas.....	42
Figura 17: Changepoints Nacional .....	43
Figura 18: Changepoints Guayas.....	43
Figura 19: Changepoints Manabí .....	44
Figura 20: Changepoints Pichincha.....	44
Figura 21: Changepoints Los Rios .....	45
Figura 22: Changepoints Esmeraldas .....	45
Figura 23: Datos Reales VS Pronosticados .....	46
Figura 24: Datos Reales VS Pronosticados Guayas .....	47
Figura 25: Datos Reales VS Pronosticados Manabí .....	47
Figura 26: Datos Reales VS Pronosticados Pichincha.....	48
Figura 27: Datos Reales VS Pronosticados Los Ríos .....	49
Figura 28: Datos Reales VS Pronosticados Esmeraldas .....	49

## 1. INTRODUCCIÓN

En el Ecuador, la tasa de homicidios ha sido una preocupación constante, según datos del Instituto Nacional de Estadísticas y Censos, en el año 2020 se registraron 1.372 homicidios en el país, lo que representa una tasa de 7,78 homicidios por cada 100.000 habitantes (INEC, 2021). Esta cifra representa un aumento del 0,95 en comparación con el año anterior. Sin embargo, continúa siendo una de las tasas más altas en comparación con los países de la región y de acuerdo a lo mencionado por (Datosmacro.com. (s. f.)) “En Ecuador se llevan a cabo al menos 4 asesinatos cada día”.

La violencia en el país incrementó en los últimos años, aduciendo principalmente a la íntima relación con los diferentes factores como son: pobreza, desigualdad social, falta de oportunidades, violencia intrafamiliar, presencia de grupos delictivos, actos de corrupción, narcotráfico, entre otros. A través de los medios de comunicación, se evidencia que la violencia en el país se concentra en ciertas regiones y sectores de la población, principalmente en los jóvenes y los habitantes de áreas urbanas marginales de la costa.

Por tal motivo, la tecnología desempeña un papel muy importante en la mitigación de las muertes violentas y los modelos predictivos o de pronóstico, pueden ser una herramienta útil para identificar zonas de mayor riesgo en donde se presenta la delincuencia con un alto índice de violencia, permitiendo a las autoridades enfocar sus recursos y esfuerzos en la prevención del delito. Estos modelos se basan en el análisis de grandes cantidades de datos para identificar patrones y tendencias que puedan indicar una mayor probabilidad de que ocurran los delitos.

En el caso del Ecuador, el uso de tecnologías como la inteligencia artificial y el análisis de datos, se han implementado en algunos programas de seguridad ciudadana. Por ejemplo, como menciona Salavarría, en la ciudad de Guayaquil se implementó, “un sistema de análisis de datos que utiliza algoritmos de

aprendizaje automático para predecir la generación de delitos” (2018), esto permite que las principales autoridades tengan una respuesta oportuna. Sin embargo, la falta de medidas efectivas por parte del Gobierno, impide un uso adecuado de los recursos, lo que resulta en un empeoramiento de la situación en el país, especialmente en la zona costera ecuatoriana, donde se observa un aumento alarmante de los casos de muertes violentas.

Finalmente, es importante tomar en cuenta que los modelos predictivos no son una solución mágica para contrarrestar la violencia y que su efectividad depende en gran medida de la calidad, procesamiento y precisión de los datos utilizados, así como, de la capacidad de las principales autoridades para reaccionar de manera oportuna y efectiva ante las posibles alertas generadas por el modelo. Además, es necesario que estas tecnologías se utilicen de manera responsable y ética, respetando los derechos y la privacidad de las personas involucradas en la recopilación y análisis de datos.



## 2. REVISIÓN DE LITERATURA

Las muertes violentas generadas en el país se clasifican en: homicidio, asesinato, femicidio, sicariato, entre otros. Por lo que es importante contar con una definición estandarizada, que brinde un concepto globalizado que permita mejorar la precisión y comparabilidad de los datos obtenidos para lo cual se utiliza definiciones obtenidas de organismos internacionales.

Adicional, para poder realizar un análisis mucho más objetivo y aumentar su precisión en la comparabilidad de los datos, principalmente de homicidio, la UNODC indica que se debe considerar lo siguiente:

“Si en un incidente fallecen dos personas, aquellos sistemas de informes basados en el incidente pueden reportar lo ocurrido como un homicidio y solamente se deben denunciar los que hayan sido finalizados, se debe evitar los casos donde la víctima sobrevivió a pesar de la evidencia de que el infractor tenía la intención de matarla.” (2013).

Esto es muy importante porque si se considera el “intento de asesinato” en la definición de “homicidio” a nivel estadístico la información variará por lo que el indicador de víctimas fallecidas será mayor al real. Adicionando que este tipo de investigaciones presentan variables muy específicas como sexo, edad de la víctima, generalmente la edad del infractor, el entorno y las armas utilizadas para cometer el crimen.

A continuación, se detalla los conceptos de las principales tipologías de muertes violentas presentados:

- Homicidio: el Código Orgánico Integral Penal (COIP) define como la muerte ilícita causada a una persona (2017). Y de acuerdo a la definición usada por la UNODC, el homicidio es “la muerte ilegal causada resueltamente a una persona por otra persona y presenta tres elementos que lo caracterizan
  - La muerte de una persona por parte de otra persona (elemento objetivo).

- El propósito del autor de matar o herir severamente a la víctima (elemento subjetivo).
  - El homicidio intencional es ilegal, por lo tanto, según la ley al autor como responsable de la muerte ilegal (elemento legal)” (2013):
- Asesinato: Es un delito contra la vida humana que consiste en matar a otras y será sancionada con pena privativa de la libertad de 22 a 26 años. (Código Orgánico Integral Penal (COIP), 2017) De acuerdo Código Penal español, en el artículo 139 se define al asesinato de la siguiente manera: "El que matare a otro concurriendo alguna de las circunstancias siguientes será castigado como reo del delito de asesinato:
- Con alevosía.
  - Por precio, recompensa o promesa.
  - Con ensañamiento, aumentando deliberada e inhumanamente el dolor del ofendido.
  - Para facilitar la comisión de otro delito o para evitar que se descubra.
  - Con ocasión, en el acto o por consecuencia de un delito contra la libertad sexual." (1995).
- Femicidio: El Código Orgánico Integral Penal (COIP), en el artículo 141, lo define como la persona que, como resultado de relaciones de poder manifestadas en cualquier tipo de violencia, dé muerte a una mujer por el hecho de serlo o por su condición de género, será sancionada con pena privativa de la libertad de 22 a 23 años. (2014)
- Sicariato: su significado etimológico es “hombre daga”, es el tipo de homicidio o asesinato agravado por el cobro de una remuneración económica a cambio de dar el servicio de matar a otra persona y de acuerdo al Diccionario de la Real Academia Española (RAE), se define al sicariato como "una práctica delictiva que consiste en el asesinato por encargo, generalmente con fines políticos o económicos". (2021) Es una práctica delictiva en la que una persona contrata a otra para que cometa un asesinato por encargo. También se puede denominar como "asesinato a sueldo" o "homicidio por encargo". El objetivo del sicariato es eliminar a

una persona específica, ya sea por motivos de venganza, dinero, política, entre otros.

Cabe recalcar que estos actos criminales son objeto de preocupación social y en términos de seguridad pública, un homicidio se refiere al acto de matar a otra persona. Y se pueden clasificar según la intención, contexto y crimen organizado, siendo: homicidios dolosos (el deseo de causar la muerte), homicidios culposos (producido por negligencia o imprudencia), homicidios domésticos, sicariato, entre otros.

Adicional, los homicidios pueden tener diversas causas y factores de riesgo asociados, como por ejemplo los que incluyen la presencia de armas de fuego, la violencia en el entorno social, la pobreza, la desigualdad, el consumo de drogas y alcohol, la falta de acceso a la educación y servicios de salud, entre otros que generan un impacto significativo en la sociedad y en las familias de las víctimas, generando temor, traumatismo y pérdida de vidas humanas. Sin considerar que los homicidios pueden tener un impacto económico negativo al afectar la seguridad y la estabilidad de las comunidades y del país. Y como menciona Angélica Escobar, “Para el caso específico de los homicidios, las variables que más destacan son la eficiencia de la policía, la educación, el desempleo y el género”. Adicionando que, en su estudio se encontró que la eficiencia de la policía y la educación tienen una relación inversa con la tasa de homicidios, es decir, que a medida que aumenta la eficiencia de la policía y el nivel de educación, disminuye la tasa de homicidios. Por otro lado, se encontró que el desempleo y el género tienen una relación positiva con la tasa de homicidios, lo que significa que a medida que aumenta el desempleo y la proporción de hombres en la población, aumenta la tasa de homicidios (2015).

De acuerdo a lo mencionado por los autores Mamadou Camara, Pierre Salama, en su investigación “Homicidios en América del Sur ¿Los pobres son peligrosos?”, los factores que desencadenan la violencia, se encuentran relacionados a "La evolución de la demografía, educación y de las estructuras

de la familia de cara a la modernización y a la urbanización (Todd, 2022) pueden conducir a un desencadenamiento de la violencia, no necesariamente con un objetivo político” (2004). También indica que relación entre la pobreza y los homicidios en América del Sur es compleja y puede estar influenciada por otros factores, principalmente la desigualdad y la presencia de grupos armados como las guerrilleras y narcotráfico e inclusive por razones de género. Tal como indica la UNODC en su estudio “Marco Estadístico para Medir el Homicidio de Mujeres y Niñas por Razones de Género”, manifiesta que, “el comportamiento de una mujer no se ajusta a las normas sociales o a los roles de género estereotipados”, siendo estos factores los desencadenantes de la violencia.

Adicional, en el estudio denominado “Detection of Homicide Trends in Colombia Using Machine Learning” detallan varios factores que pueden desencadenar los homicidios, algunos de estos factores incluyen:

- La presencia de bandas de crimen organizado enfocadas en negocios ilícitos, como el narcotráfico, que operan en la zona urbana.
- La falta de oportunidades laborales para los jóvenes, lo que repercute en el ingreso a pandillas o bandas de crimen organizado.
- Venganzas y atentados terroristas que impactan en la vida de la población inmersa en problemas de seguridad, especialmente en la primera etapa de la vida entre los 0 y 5 años. (Ordoñez Erazo, Pardo Calvache, & Cobos Lozada, 2020)

Por tal motivo, se consideró en realizar esta investigación y en el Anexo 1, se detalla algunas investigaciones relacionadas, en donde se incluyen los tipos de datos utilizados, metodologías para el análisis de datos, resultados e implicaciones gerenciales en dichos estudios.

## **METODOLOGÍAS PARA EL ANÁLISIS DE DATOS**

- **Análisis Exploratorio:** Se analizan los patrones generados a nivel del número de muertes violentas con sus diferentes características.
- **Análisis Descriptivo:** Se centra en el comportamiento de la tasa de homicidios durante el periodo estudiado, que abarca desde el año 2015 hasta el 2022. Se busca obtener una comprensión inicial de los factores que afectan esta tasa de homicidios.
- **Análisis de Correlación:** Se identifica las relaciones existentes entre el número de muertes violentas y las variables socioeconómicas.
- **Análisis de Series Temporales:** Se utilizan modelos predictivos, específicamente los modelos: aditivo de Prophet, ARIMA y SARIMA para evaluar el impacto de las muertes violentas generadas. Específicamente se centra en el modelo aditivo de Prophet.

## **TIPOS DE FUENTES DE DATOS UTILIZADAS**

- Datos oficiales del Ministerio de Gobierno que corresponden a muertes violentas generadas en el Ecuador en el periodo 2015-2022.
- Datos de proyecciones poblacionales del Ecuador en el periodo 2015 – 2022 de la Sistema Nacional de Información de la Secretaría Nacional de Planificación.
- Datos de la tasa extrema de pobreza por ingresos en el Ecuador en el período 2015-2022 del Ministerio de Inclusión.
- Datos de la tasa de empleo inadecuado en el Ecuador en el periodo 2015-2022 del Ministerio de Inclusión.
- Datos de la tasa de analfabetismo en el Ecuador en el período 2015-2022 del Instituto Nacional de Estadísticas y Censos.
- Investigaciones académicas previas sobre homicidios en Ecuador y su relación con el Big Data.

- Informes y estudios de organizaciones internacionales, como la Organización de las Naciones Unidas (ONU), el Banco Interamericano de Desarrollo (BID) y el Programa de las Naciones Unidas para el Desarrollo (PNUD), que puedan proporcionar datos y análisis sobre la violencia y la seguridad en el país.
- Artículos periodísticos y reportajes sobre homicidios.

### ***IMPLICACIONES GERENCIALES:***

Es importante destacar que la aplicación exitosa de los hallazgos requiere un liderazgo comprometido a nivel gubernamental y una gestión eficaz de los recursos y la capacidad de adaptación:

- Políticas de seguridad: Los resultados de la investigación pueden proporcionar información clave para aplicar políticas públicas de seguridad que sean eficientes y eficaces, lo que repercutiría en cambios en la asignación de recursos para la prevención de crímenes, generación de programas de rehabilitación y la mejora de la coordinación entre las fuerzas de seguridad como es la Policía Nacional y Fuerzas Armadas.
- Planificación estratégica: Los hallazgos pueden tener relación con la determinación de diferentes áreas siendo estas geográficas o demográficas que son consideradas como las más afectadas por la violencia, en específico de las muertes violentas, así como, permitirá identificar los grupos y factores de riesgo, permitiendo realizar una efectiva asignación de recurso para tener respuestas más enfocadas en problemas de violencia en el país.
- Colaboración y coordinación interinstitucional: Los resultados pueden impulsar la creación de alianzas estratégicas entre entidades gubernamentales, organizaciones no gubernamentales, sector privado, público y la sociedad civil para abordar de manera conjunta el problema.

Esto implica una gestión eficaz de las relaciones y la creación de redes de cooperación.

- **Sensibilización y educación:** La investigación puede contribuir a aumentar la conciencia sobre la violencia y sus causas, lo cual es fundamental para impulsar cambios a largo plazo, lo que sustentará la generación de campañas con programas educativos dirigidos a la comunidad en general o a grupos específicos (grupo de interés).

### 3. IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

El objeto de estudio son las muertes violentas ocurridas en el Ecuador y lo que se pretende, es proporcionar información relevante que contribuya a mejorar la seguridad pública en nuestro país; con los resultados obtenidos se desea disponer de una herramienta que permita a las autoridades identificar las áreas o sectores específicos, en donde es más probable que ocurran estos tristes acontecimientos. Con esto, se espera asignar recursos de una manera eficiente logrando una distribución más efectiva y que no exista desperdicios en los recursos de las fuerzas de seguridad principalmente de la Policía Nacional para prevenir estos delitos.

De igual forma se busca disminuir la tasa de homicidios al lograr prevenir los delitos que se generan en las diferentes provincias del Ecuador. Esta disminución impactará directamente en la imagen del país a nivel internacional de una manera positiva, porque la seguridad de la ciudadanía incrementará; la información resultante permitirá optimizar el uso de los recursos policiales y judiciales, lo que puede reducir los costos asociados a la investigación y resolución de homicidios.

Con esta investigación es posible identificar patrones y factores de riesgo, al analizar los datos y las variables que influyen en la ocurrencia de homicidios, permitiendo a las autoridades implementar políticas públicas y estrategias de prevención más efectivas.

Finalmente, la aplicación de un modelo predictivo de homicidios en Ecuador puede ser justificado por su potencial para mejorar la seguridad pública, reducir la tasa de homicidios, ahorrar recursos, identificar patrones y factores de riesgo, aplicando técnicas avanzadas de análisis de datos. Sin embargo, es importante tener en cuenta que cualquier modelo predictivo debe ser desarrollado e implementado con cuidado y ética, para evitar sesgos o discriminación en su uso y aplicación.



#### 4. PLANTEAMIENTO DEL PROBLEMA

El Ecuador históricamente ha enfrentado problemas como la corrupción, pobreza y bajo crecimiento como menciona (Ovalle Correa, 2019). Sin embargo, los altos niveles de delincuencia y crimen organizado, es algo nuevo que se presenta en el país, así como, la toma de territorio de bandas delictivas por las rutas del narcotráfico; esto repercute, en que el gobierno es incapaz de actuar y proteger a los ciudadanos.

Por lo que el aumento de la tasa de homicidios en Ecuador a la fecha se presenta como un problema grave que afecta la seguridad ciudadana y la imagen del país a nivel internacional. A pesar de los esfuerzos realizados por las autoridades y la sociedad civil, la violencia homicida sigue siendo un desafío importante en el país. En este contexto, surge la necesidad de un análisis situacional y predictivo que permita identificar patrones y factores de riesgo asociados a la ocurrencia de muertes violentas en Ecuador.

El objetivo principal de esta investigación es realizar un análisis situacional y predictivo que permita predecir la probabilidad de ocurrencia de homicidios en distintas áreas geográficas, además de identificar las posibles relaciones con las diferentes variables socioeconómicas y demográficas en las provincias del Ecuador en el periodo 2015-2022. Para lograr este objetivo, se llevará a cabo un análisis exploratorio de los datos disponibles relacionados con las muertes violentas; así como en otras fuentes secundarias relevantes.

El análisis situacional y predictivo se basará en técnicas avanzadas de análisis de datos y aprendizaje automático, y se evaluará su precisión y capacidad predictiva mediante pruebas y validaciones estadísticas. El resultado esperado es identificar un modelo predictivo que permita a las autoridades y la sociedad civil prevenir y reducir los homicidios, mejorando la seguridad ciudadana y promoviendo la paz y el bienestar social en el país.

## 5. OBJETIVO GENERAL

Realizar un pronóstico de las muertes violentas en el Ecuador mediante el uso de modelos de series temporales e identificar la relación entre las muertes violentas y las variables socioeconómicas en el país, con el fin de determinar los principales factores que contribuyen a su ocurrencia.

## 6. OBJETIVOS ESPECÍFICOS

- Identificar las principales causas que desencadenan las muertes violentas en las provincias de Ecuador, en particular de Esmeraldas, Manabí, Guayas, Los Ríos y Pichincha.
- Determinar a través de un análisis descriptivo, los principales factores que influyen a la ocurrencia de las muertes violentas para lograr definir grupos de poblaciones afectados, lugares de ocurrencia, temporalidad y factores socioeconómicos.
- Pronosticar las muertes violentas a nivel provincial, utilizando modelos de series temporales.
- Recomendar el uso del análisis situacional y predictivo a las entidades públicas competentes para que realicen gestiones logísticas, operacionales y tácticas correspondientes.

## 7. JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

En base a la situación actual, el presente estudio pretende realizar un análisis situacional y predictivo que permita establecer los principales factores desencadenantes de muertes violentas en el Ecuador. La importancia de este análisis se basa principalmente en el deseo de contrarrestar la delincuencia en el país y dar a conocer a las instituciones del gobierno, los principales factores y la creciente tasa de muertes violentas que se generará a futuro si en la actualidad no se aplican políticas públicas para solventarla.

Se escogió este tema, porque al hablar de homicidio connota a la búsqueda inherente de soluciones a diferentes fenómenos que se encuentran relacionados con la violencia, el rol de las armas, la desigualdad, asesinatos relacionados al género y asaltos que en la actualidad el Ecuador y el mundo está sufriendo. Por ello, es importante aprender, comprender y fortalecer la cultura de la prevención.

Es importante tener en cuenta que cualquier modelo predictivo debe ser usado y aplicado con cuidado y ética, para evitar sesgos o discriminación en su aplicación, por lo que la metodología utilizada es el análisis exploratorio y descriptivo para identificar las variables, en base al preprocesamiento de datos, selección de características, análisis correlacional y uso de modelos predictivos para identificar la proyección nacional y provincial.

Para el análisis predictivo se utilizan modelos de series temporales, ya que permiten analizar el comportamiento de los datos a lo largo del tiempo, es posible identificar patrones estacionales o tendencias a largo plazo. Además de generar pronósticos en el corto y largo plazo.

En ese sentido, se seleccionaron tres modelos usados en la predicción de series temporales, estos modelos a evaluar son: el modelo aditivo Prophet de Facebook, el modelo ARIMA y el modelo SARIMA.

El modelo aditivo Prophet permite realizar pronósticos precisos y de alta calidad en una variedad de series de tiempo. Y de acuerdo a lo mencionado por Harvey, A. C., & Shephard, N. "es un modelo de series de tiempo descomponible que

utiliza tres componentes principales: tendencia, estacionalidad y días festivos” (1993). El modelo a menudo funciona bien con los parámetros predeterminados. Además, cuenta con un sistema de medición y seguimiento de la precisión de los pronósticos, lo que permite a los analistas identificar cuándo se deben hacer ajustes al modelo o cuándo se debe utilizar un modelo completamente diferente. En general, el modelo aditivo Prophet es una solución práctica y escalable para producir pronósticos confiables y de alta calidad, incluso cuando se trabaja con una variedad de series de tiempo y se tiene una experiencia limitada en modelado de series de tiempo.

El modelo AutoRegressive Integrated Moving Average (ARIMA) es un modelo clásico para pronosticar series temporales. Esta combina componentes autoregresivos (AR), de media móvil (MA) y de diferenciación (I) para capturar las dependencias temporales y espaciales en la serie.

El modelo SARIMA, también conocido como ARIMA Seasonal, es una extensión del modelo ARIMA que utiliza elementos estacionales para hacer predicciones. Este modelo es particularmente útil para series temporales que tienen patrones estacionales evidentes.

## **SELECCIÓN DE LA BASE DE DATOS.**

Las bases de datos seleccionadas son fuentes secundarias que se obtuvieron del ministerio de Gobierno y corresponden al periodo 2015-2022, en donde se detallan las muertes violentas ocurridas en el Ecuador, las mismas se encuentran disponibles al público a través de la página web.<sup>1</sup>

Las principales variables que se detallan son: el tipo de muerte clasificado en homicidio, femicidio, asesinato y sicariato, la provincia, zona y subzona en donde se generaron los actos ilícitos, el área urbana, rural y el lugar en específico (vía pública, bares, colegios, hoteles, etc.), la fecha de infracción, el tipo de arma

---

<sup>1</sup> <https://www.datosabiertos.gob.ec>

utilizada incluye todo objeto cortopunzante y arma blanca, objeto del crimen y datos demográficos de la víctima.

Adicional, se extrajo información de las siguientes fuentes de datos:

- Datos de proyecciones poblacionales del Ecuador en el periodo 2015 – 2022 de la Sistema Nacional de Información de la Secretaría Nacional de Planificación.<sup>2</sup>
- Datos de la tasa de analfabetismo en el Ecuador en el período 2015-2022 del Instituto Nacional de Estadísticas y Censos.<sup>3</sup>
- Datos de la tasa de pobreza extrema por ingresos<sup>4</sup> y datos de la tasa de empleo inadecuado<sup>5</sup> en el Ecuador en el periodo 2015-2022 del Ministerio de Inclusión.

## **LIMPIEZA, PRE-PROCESAMIENTO Y/O TRANSFORMACIÓN DE DATOS.**

En cada etapa se realizó un tratamiento especial con la finalidad de presentar información precisa para obtener predicciones correctas, por tal motivo, para el pre - procesamiento de los datos y el modelamiento estadístico del estudio se utiliza Excel y Python, lo que permitirá analizar y predecir el comportamiento de las variables.

### *Limpieza*

Se valida que en las bases de datos no exista errores en los formatos de las variables temporales, que el nombre de las variables se encuentre correctamente

---

<sup>2</sup> <https://sni.gob.ec/proyecciones-y-estudios-demograficos>

<sup>3</sup> <https://aplicaciones3.ecuadorencifras.gob.ec/BIINEC-war/index.xhtml>

<sup>4</sup> <https://info.inclusion.gob.ec/index.php/caracterizacion-poblacion-objetivo-ancusrext/pobreza-y-desigualdad-ancusrext/tasa-de-pobreza-extrema-ancusrext>

<sup>5</sup> <https://info.inclusion.gob.ec/index.php/caracterizacion-poblacion-objetivo-ancusrext/indicadores-laborales-ancusrext/tasa-de-empleo-inadecuado-ancusrext>

escritas, eliminando errores tipográficos, datos duplicados y datos atípicos. Adicionalmente, se valida que los datos sean consistentes e íntegros.

### **PRE-PROCESAMIENTO:**

Tiene como objetivo principal obtener un data set que sea de calidad y útil para la investigación. A continuación, se detalla el pre-procesamiento en cada una de las bases de datos utilizadas:

- Base de datos de muertes violentas año 2022: La base original cuenta con 19 variables de tipos: categórica, temporal y geográfica. Para un mejor manejo de la base de datos se incorporan variables con mayor nivel de detalle a partir de las variables existentes, como son las variables: Hora, Día\_Semana, Distrito, Circuito y Cantidad; para obtener las variables Distrito y Circuito se realiza un cruce de datos con la base de Subcircuitos de Senplades disponible en la página web institucional. Esta base contiene mayor número de variables que la base de muertes años 2015-2022; por lo tanto, será utilizada para fines descriptivos debido a su mayor nivel de detalle de los datos.
- Base de datos de muertes violentas años 2015-2022: La base original cuenta con 13 variables, esta base de datos contiene menos variables con menor detalle de información que la base del año 2022 y contiene variables de tipos: categórica, temporal y geográfica. Se incorporan variables con mayor nivel de detalle a partir de las variables existentes, como son las variables: Día\_Semana, Provincia y Cantidad. Esta base de datos se utilizará para generar los modelos predictivos y realizar el análisis de correlación con las variables socioeconómicas.
- Datos de variables socioeconómicas: En estos sets de datos (tasa de pobreza extrema, tasa de empleo inadecuado y tasa de analfabetismo) se dispone de una variable temporal "Año", una variable geográfica "Provincia" y otra variable numérica "Tasa" que no requieren ningún

preprocesamiento. Este set de datos se utilizará en el análisis correlacional.

- Datos de tasa de homicidios cada 100.000 habitantes: Se genera un set de datos con tres variables “Año”, “Provincia” y “Tasa” a partir de la base de datos de proyecciones poblacionales y el número de homicidios a nivel nacional y provincial. Este set de datos se utilizará en el análisis correlacional.

Posteriormente en Python se cargan las librerías correspondientes y se cargan las bases de datos para realizar los análisis correspondientes.

## **TRANSFORMACIÓN DE DATOS**

En la base de datos de muertes violentas 2015-2022 están identificados como una muerte violenta en cada registro, por lo cual es necesario hacer una agrupación de datos a nivel temporal diario.

Mediante un gráfico de caja se identifica que en la base de datos existen datos atípicos, de manera más evidente en los años 2021 y 2022, considerando que el promedio se encuentra en 4,73, la mediana en 3, el tercer cuartil en 6 y la desviación estándar en 5,42; en ese sentido es necesario realizar un procesamiento para el tratamiento de estos datos:

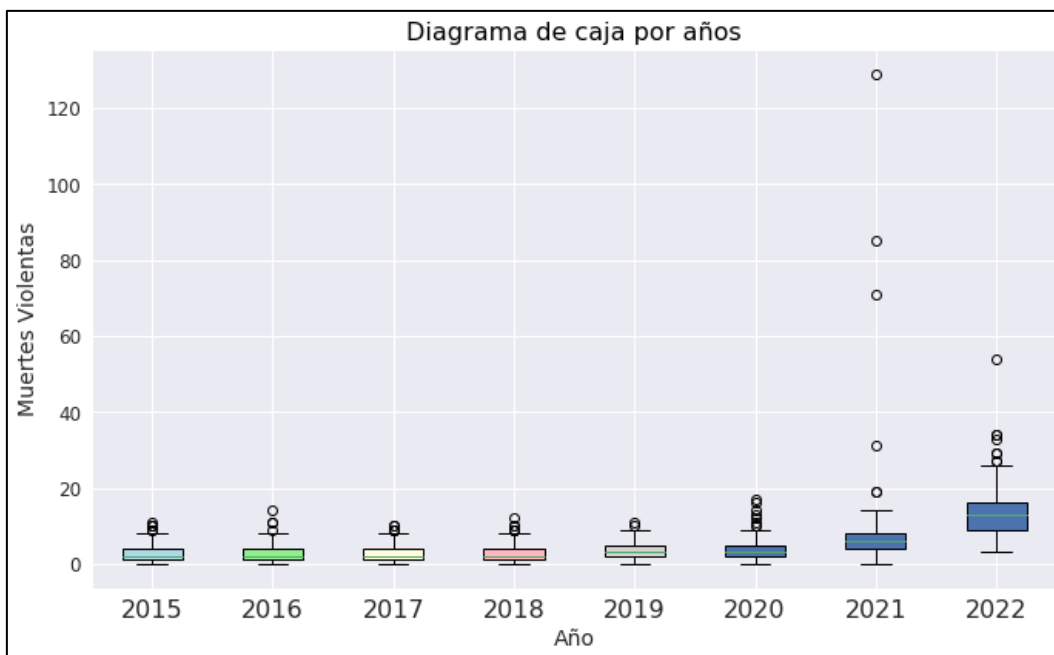


Figura 1: Gráfico de cajas  
Fuente: Elaboración propia

Los datos que superen el percentil 99, se reemplazan por el promedio general de muertes violentas, esto con el objetivo de que los outliers no afecten en la posterior aplicación de los modelos predictivos. Se puede apreciar que los datos con muertes violentas superior a 21 fueron transformados.

**Tabla 1: Transformación de Variables**

Detalle	Datos Iniciales	Datos Modificados	Variación Absoluta
Conteo	2922	2922	0
Promedio	4,73	4,46	-0,27
Desviación Estándar	5,42	3,99	-1,43
Mínimo	0	0	0
25%	2	2	0
50%	3	3	0
75%	6	6	0
Máximo	129	21	-108

Fuente: Elaboración propia

Es importante mencionar que el modelo aditivo de Prophet impone la estricta condición de que las columnas de entrada deben denominarse “ds” con su abreviatura timestamp (columna que contiene las fechas) y la variable “y”



(variable objetivo o variables que se desea predecir). Así pues, se debe cambiar el nombre de las columnas del marco de datos.

## IDENTIFICACIÓN Y DESCRIPCIÓN DE VARIABLES

El análisis descriptivo y predictivo se centran en la variable dependiente “Cantidad” que corresponde al número de Muertes Violentas, el resto de las variables son independientes, de acuerdo con la categorización de variables que se muestra a continuación:

**Tabla 2: Categorización de variables Muertes Violentas**

Variable	Descripción del campo	Tipo de Variable
<b>Tipo de muerte</b>	<b>Homicidio:</b> La persona que mate a otra.	Categoría Object Independiente
	<b>Asesinato:</b> La persona que mate a otra.	
	<b>Femicidio:</b> Violencia de muerte a una mujer	
	<b>Sicariato:</b> Persona que mate a otra por precio u otra forma de beneficio, para sí o un tercero.	
<b>Zona</b>	Extensión de terreno cuyos límites se encuentran definidos.	Categoría /Object Independiente
<b>Subzona</b>	Porciones de terreno pequeñas que se encuentran dentro de las zonas.	Categoría /Object Independiente
<b>Provincia</b>	Extensión de terreno cuyos límites están determinados geográficamente.	Categoría /Object Independiente
<b>Área del Hecho</b>	Limitación geográfica de acuerdo a su ubicación: Urbano y Rural.	Categoría /Object Independiente
<b>Lugar de hecho</b>	Área o espacio donde se ha desarrollado la muerte de una persona por causas estructurales de violencia.	Categoría /Object Independiente
<b>Fecha Infracción</b>	Es la fecha en la cual se suscita un hecho delictivo	Data time 64 Independiente
<b>Año</b>	Año calendario	Numérica/ Int 64 Independiente
<b>Mes</b>	Mes del calendario	Categoría /Object Independiente
<b>Día_Semana</b>	Día de la semana	Categoría /Object Independiente
<b>Tipo Arma</b>	Pudiendo ser arma blanca, sustancias, etc.	Categoría /Object Independiente

<b>Presunta motivación (observación)</b>	Probable causa de muerte	Categórica /Object Independiente
<b>Sexo/Género</b>	Seres animados racionales, sea varón o mujer y el no determinado que hace referencia a la variable que describe el desconocimiento de las características biológicas presentes en la víctima.	Categórica /Object Independiente
<b>Estado Civil</b>	Condición de una persona en función de si tiene o no pareja registrada en el Registro Civil.	Categórica /Object Independiente
<b>Nacionalidad</b>	Condición que reconoce a una persona la pertenencia a un estado o nación.	Categórica /Object Independiente
<b>Cantidad</b>	Número de muertes violentas generadas en el Ecuador en el periodo 2015-2022	Numérica Int 64 Dependiente

*Fuente: (Código Orgánico Integral Penal, COIP, 2021)/ Elaboración propia*

**Tabla 3: Categorización de variables Tasa de Homicidios cada 100.000 hab.**

<b>Variable</b>	<b>Descripción del campo</b>	<b>Tipo de Variable</b>
<b>Año</b>	Período de doce meses que comienza el día 1 de enero y finaliza el 31 de diciembre.	Numérica/ Int 64
<b>Provincia</b>	Extensión de terreno cuyos límites se encuentran definidos.	Categórica /Object Independiente
<b>Tasa</b>	Se refiere a una medida que indica la frecuencia o proporción de homicidios en relación con una población determinada durante un período específico.	Numérica/ Int 64

*Fuente: (Código Orgánico Integral Penal, COIP, 2021)/ Elaboración propia*

**Tabla 4: Categorización de variables Tasa de Pobreza Extrema por Ingresos.**

<b>Variable</b>	<b>Descripción del campo</b>	<b>Tipo de Variable</b>
<b>Año</b>	Período de doce meses que comienza el día 1 de enero y finaliza el 31 de diciembre.	Numérica/ Int 64

<b>Provincia</b>	Extensión de terreno cuyos límites se encuentran definidos.	Categórica /Object Independiente
<b>Tasa</b>	Se refiere a una medida que indica el porcentaje de la población que vive en condiciones de pobreza extrema, basada en el nivel de ingresos.	Numérica/ Int 64

*Fuente: (ENEMDU, Diciembre 2009-2022)/ Elaboración propia*

**Tabla 5: Categorización de variables Tasa de Empleo Inadecuado.**

<b>Variable</b>	<b>Descripción del campo</b>	<b>Tipo de Variable</b>
<b>Año</b>	Período de doce meses que comienza el día 1 de enero y finaliza el 31 de diciembre.	Numérica/ Int 64
<b>Provincia</b>	Extensión de terreno cuyos límites se encuentran definidos.	Categórica /Object Independiente
<b>Tasa</b>	Se refiere a una medida que indica el porcentaje de personas empleadas en trabajos que no satisfacen completamente sus habilidades, calificaciones o necesidades laborales.	Numérica/ Int 64

*Fuente: (ENEMDU, 2007-2022)/ Elaboración propia*

**Tabla 6: Categorización de variables Tasa de Analfabetismo.**

<b>Variable</b>	<b>Descripción del campo</b>	<b>Tipo de Variable</b>
<b>Año</b>	Período de doce meses que comienza el día 1 de enero y finaliza el 31 de diciembre	Numérica/ Int 64
<b>Provincia</b>	Extensión de terreno cuyos límites se encuentran definidos.	Categórica /Object Independiente
<b>Tasa</b>	Se refiere a una medida que indica el porcentaje de personas mayores de cierta edad (generalmente 15 años o más) que no saben leer ni escribir en un nivel básico.	Numérica/ Int 64

*Fuente: (ENEMU, 2014-2021)/ Elaboración propia*

## VISUALIZACIÓN DE VARIABLES

En esta etapa se realiza el análisis descriptivo de los datos con la finalidad de identificar patrones o tendencias en las variables; principalmente se utilizan técnicas de visualización.

### Análisis Descriptivo / Exploratorio

#### Muertes Violentas por Años

En el siguiente gráfico se observa el número de muertes violentas en el Ecuador por años, en el cual se evidencia que existe una tendencia creciente en el número de muertes violentas, es así que para el año 2022 se presenta un incremento del 357% respecto al año base 2015:

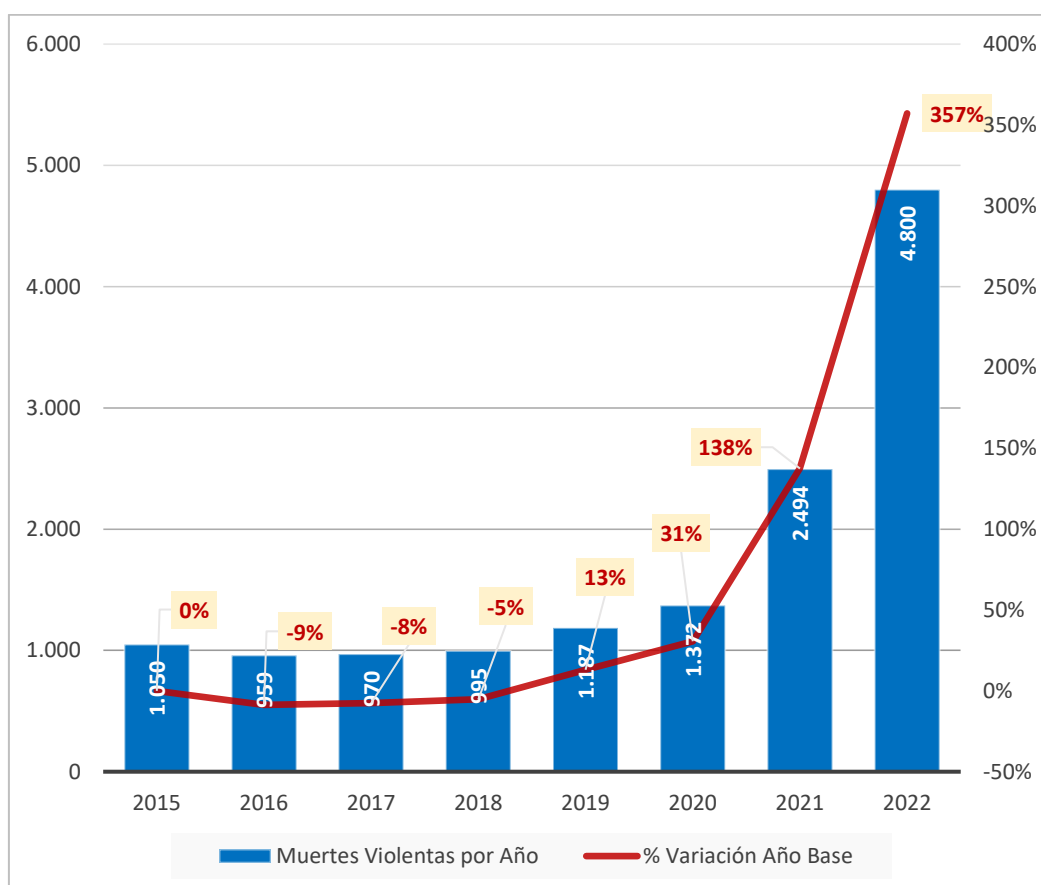


Figura 2: Muertes violentas por año  
Fuente: Elaboración propia.

## Mapas Coropléticos de Muertes Violentas en el Ecuador

En los siguientes mapas se puede apreciar que existe un incremento de las muertes violentas en el año 2022 respecto al año 2015, principalmente en las provincias de la costa ecuatoriana.

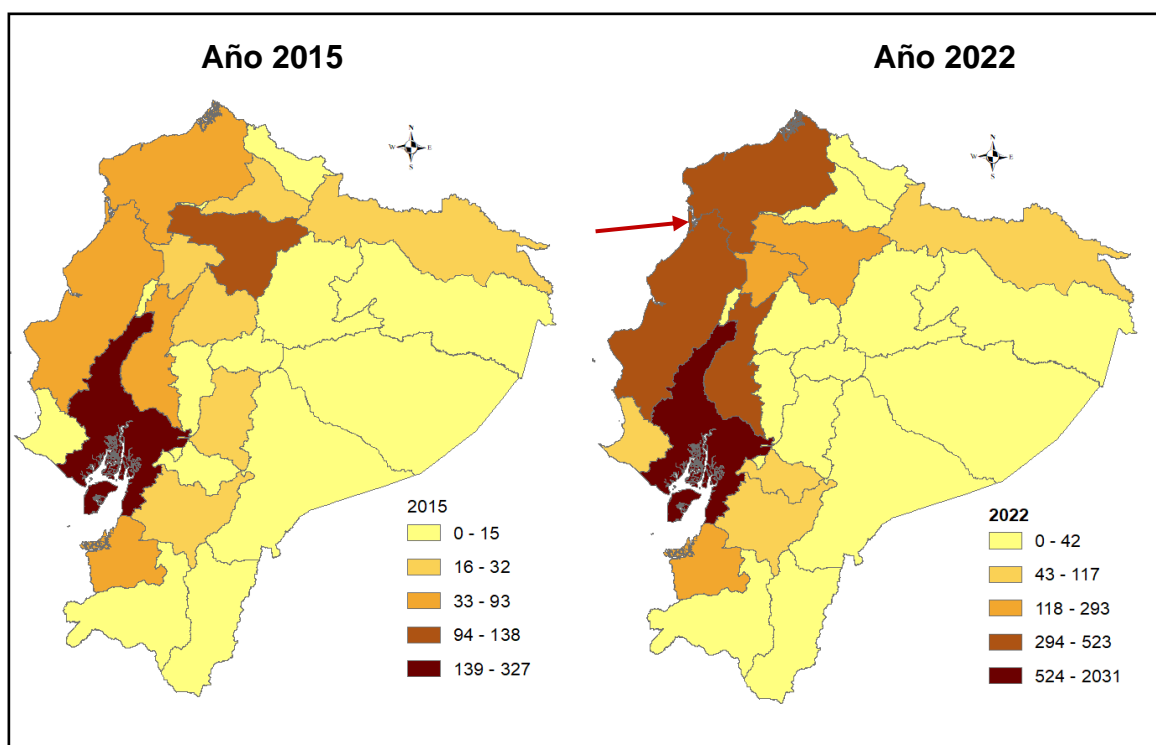


Figura 3: Mapas Coropléticos de Muertes Violentas  
Fuente: Elaboración propia

## Comparativo de Muertes Violentas por Año y Provincias del Ecuador

En 18 de las 24 provincias del Ecuador, se presenta un incremento porcentual en el número de muertes violentas del año 2022 respecto al año 2015, especialmente en las provincias de Cañar, Santa Elena, Santo Domingo de los Tsáchilas, Guayas, Manabí y Esmeraldas.

**Tabla 7: Comparativo de Muertes Violentas por Año y Provincias del Ecuador**

PROVINCIA	2015	2016	2017	2018	2019	2020	2021	2022	% Variación 2022-2015
AZUAY	28	35	44	38	24	33	59	74	164%
BOLIVAR	6	9	7	12	9	12	7	12	100%
CAÑAR	11	6	8	4	6	17	41	113	927%
CARCHI	9	8	5	8	5	7	6	12	33%
CHIMBORAZO	21	8	9	15	13	13	15	15	-29%
COTOPAXI	18	20	21	11	28	28	47	30	67%
EL ORO	63	62	56	54	87	95	146	293	365%
ESMERALDAS	93	74	63	79	67	80	151	517	456%
GALAPAGOS	0	0	1	0	0	1	0	0	
GUAYAS	327	293	286	291	409	528	1192	2031	521%
IMBABURA	22	15	12	22	25	22	18	16	-27%
LOJA	13	13	12	11	9	10	10	25	92%
LOS RIOS	85	76	77	94	111	110	197	413	386%
MANABI	91	90	100	84	113	119	195	523	475%
MORONA SANTIAGO	6	8	6	8	10	11	6	9	50%
NAPO	6	4	6	2	2	3	5	18	200%
ORELLANA	13	9	12	20	9	5	11	23	77%
PASTAZA	8	2	2	5	9	2	8	8	0%
PICHINCHA	138	136	129	134	157	159	160	218	58%
SANTA ELENA	15	17	13	8	8	15	46	117	680%
SANTO DOMINGO DE L	29	25	31	36	27	36	71	207	614%
SUCUMBIOS	32	28	37	32	30	33	40	81	153%
TUNGURAHUA	9	17	14	12	17	18	21	42	367%
ZAMORA CHINCHIPE	3	3	6	7	2	1	6	3	0%
ZONA NO DELIMITADA	4	1	13	8	10	14	36	0	-100%
<b>TOTAL</b>	<b>1.050</b>	<b>959</b>	<b>970</b>	<b>995</b>	<b>1.187</b>	<b>1.372</b>	<b>2.494</b>	<b>4.800</b>	<b>357%</b>

Fuente: Elaboración propia

### Muertes violentas de las provincias con mayor participación a nivel nacional

En el período del 2015 al 2022, aproximadamente el 74% de las muertes violentas se concentra en 5 provincias, siendo Guayas la provincia con el mayor porcentaje (38,7%), seguido de Manabí (9,5%), Pichincha (8,9%), Los Ríos (8,4%) y Esmeraldas (8,1%).

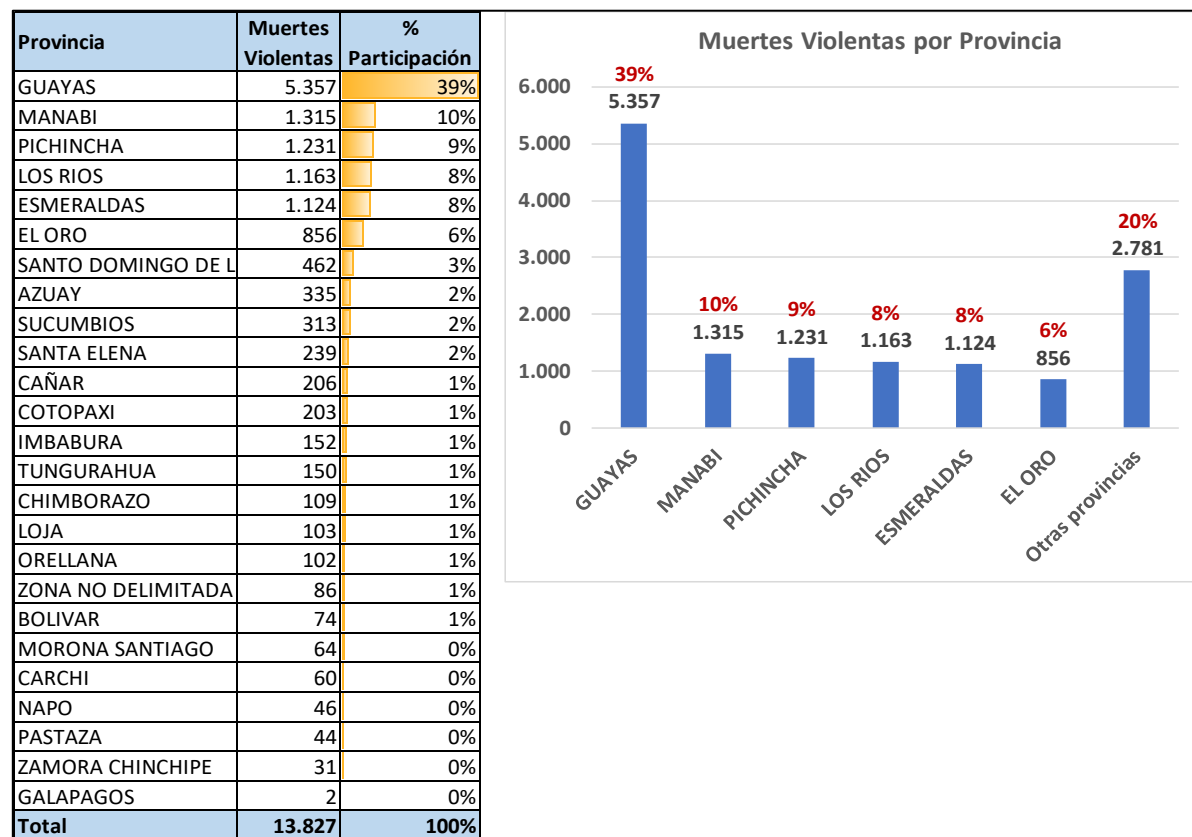


Figura 4: Muertes violentas de las provincias con mayor participación a nivel nacional  
Fuente: Elaboración propia

## Muertes violentas por Tipo de muerte en el año 2022

La distribución de muertes violentas en el año 2022 se presenta de acuerdo a la siguiente distribución: Asesinato (90,6%), Homicidio (6,8%), Femicidio (1,6%) y Sicariato (1%).

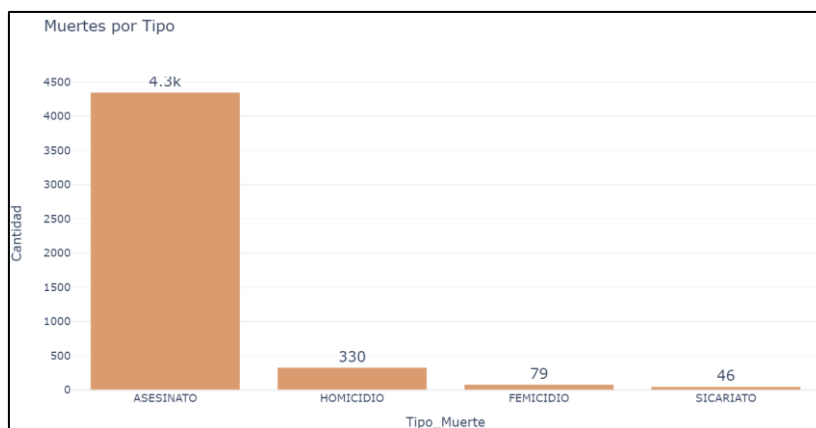


Figura 5: Muertes violentas por Tipo de muerte en el año 2022  
Fuente: Elaboración propia

## Muertes violentas por Tipo de Área

En el año 2022 las muertes violentas se distribuyen en un 82,1% en el área urbana y el 17,9% en el área rural.

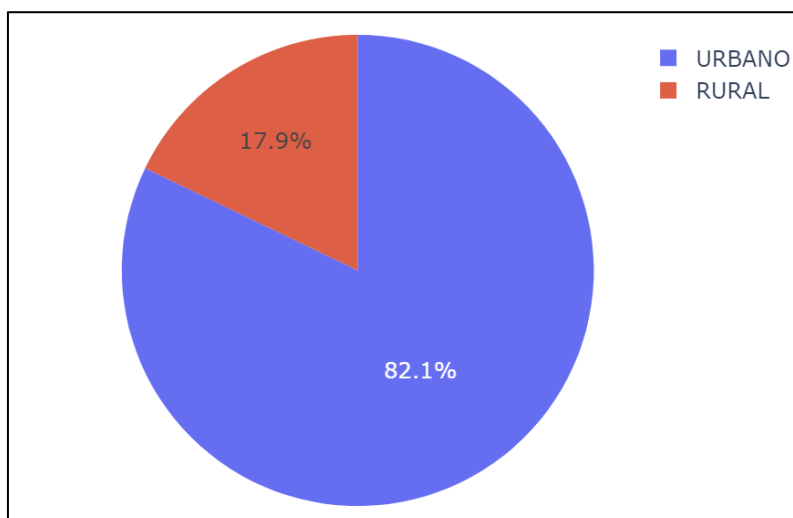


Figura 6: Muertes violentas por Tipo de Área  
Fuente: Elaboración propia



## Muertes violentas por Género en el año 2022

En el año 2022 las muertes violentas tienen como víctimas a los hombres con un 91,2%, a las mujeres con un 8,56% y no determinados con 0,23%.

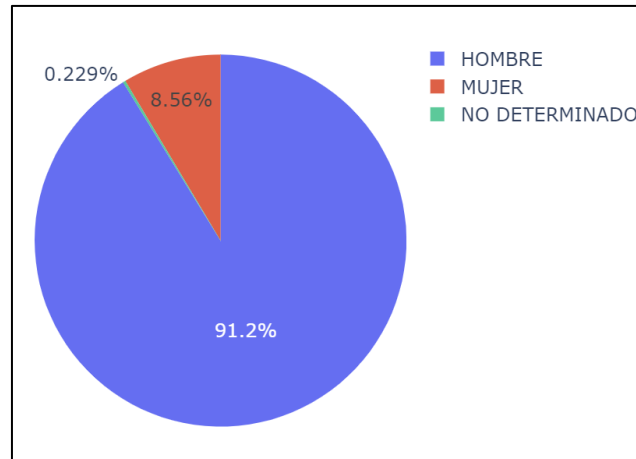


Figura 7: Muertes violentas por Género en el año 2022  
Fuente: Elaboración propia

## Muertes Violentas por Estado Civil en el año 2022

Las muertes violentas en el año 2022 se presentan en mayor porcentaje en personas solteras (75,3%), casados (10,2%), no determinados (9,3%) y las otras categorías en porcentajes menores.

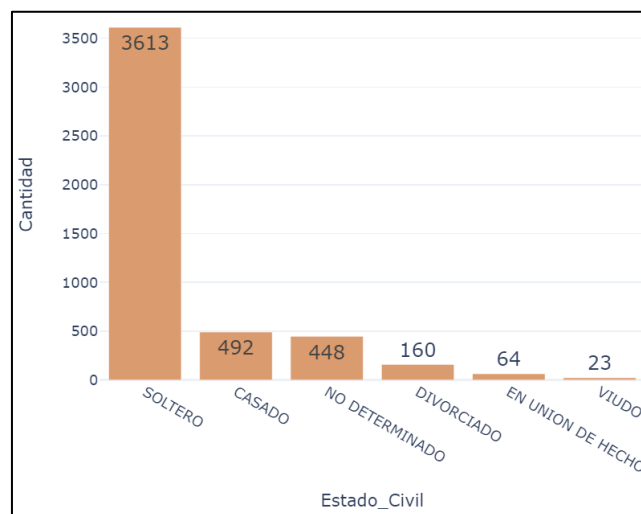


Figura 8: Estado Civil.  
Fuente: Elaboración propia

## Muertes violentas por día de la semana

Al realizar el análisis en busca de estacionalidad anual, mensual y semanal, primero se verificó que en todos los días de la semana cuentan con muestras similares y se realizó el gráfico para visualizar el histórico por días de la semana y se puede observar que las muertes violentas en el período de 2015 al 2022 tienen una mayor concentración en el día domingo.

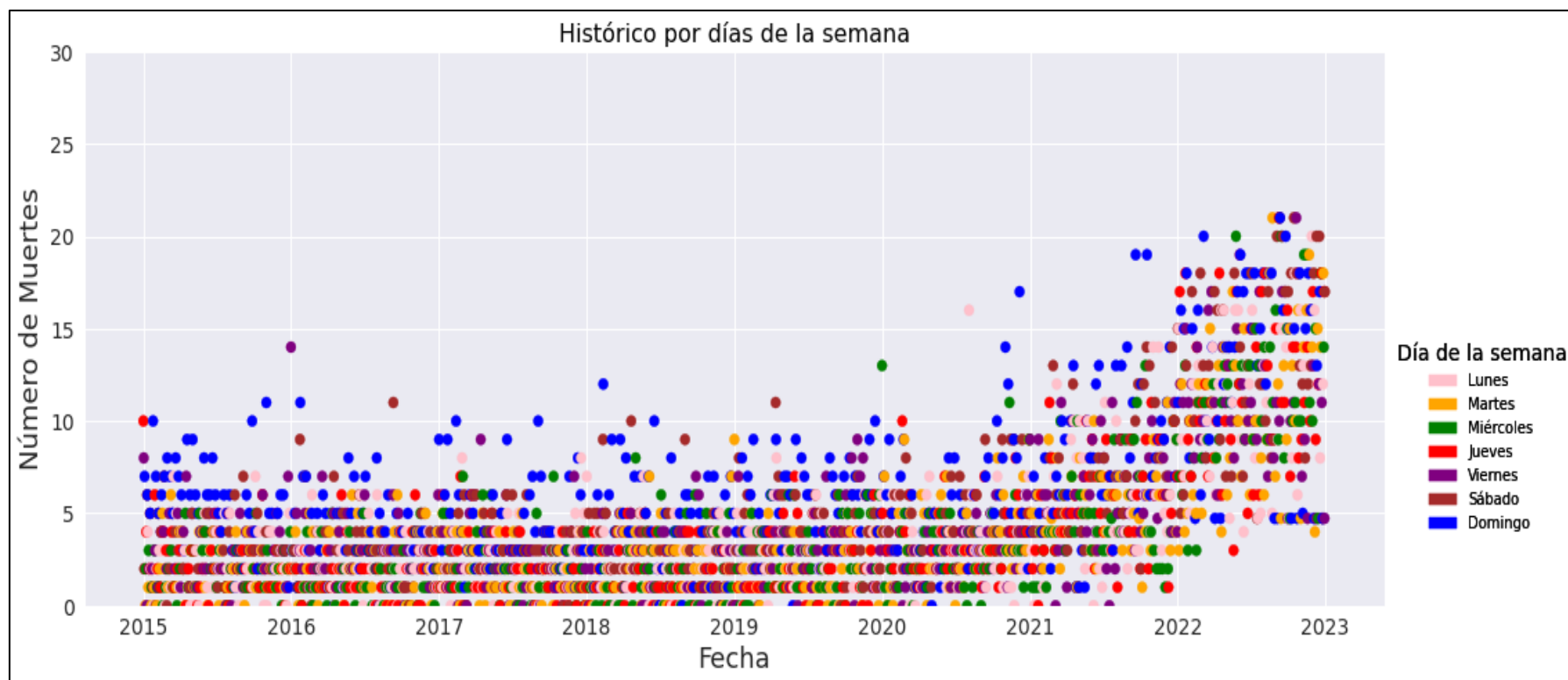


Figura 9: Muertes violentas por día de la semana  
Fuente: Elaboración propia

### Muertes violentas por fecha

En el período del año 2015 al año 2022 las muertes violentas presentan una tendencia creciente y presenta fechas puntuales con ciertos picos, como el 28 de septiembre de 2021 (129), el 23 de febrero de 2021 (85), el 12 de noviembre de 2021 (71) y el 09 de mayo de 2022 (54); en estas fechas se produjeron amotinamientos en las cárceles del país que originaron un alto número de muertes violentas y otras fechas específicas correspondientes a feriados como navidad, fin de año, fiestas de Quito, etc.

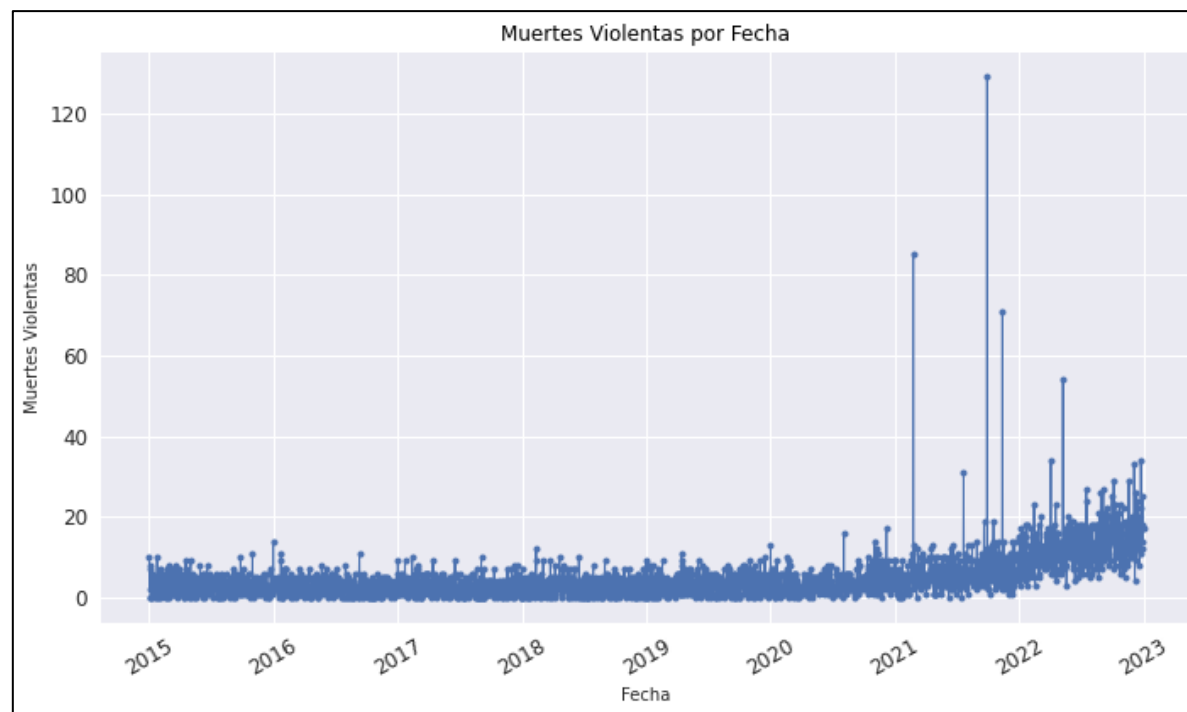


Figura 10: Muertes violentas por fecha  
Fuente: Elaboración propia

## Matriz de Correlación Nacional

Se realiza un análisis situacional para identificar si existe una correlación entre la tasa de muertes violentas y las variables socioeconómicas de tasa de empleo inadecuado, tasa de pobreza extrema, y tasa de analfabetismo con datos históricos del año 2015 al año 2022.

En la matriz de correlación Nacional se evidencia que existe una correlación positiva entre la tasa de muertes violentas y la Tasa de Empleo Inadecuado, esto podría significar que si existe un mayor porcentaje de empleo inadecuado las muertes violentas incrementan.

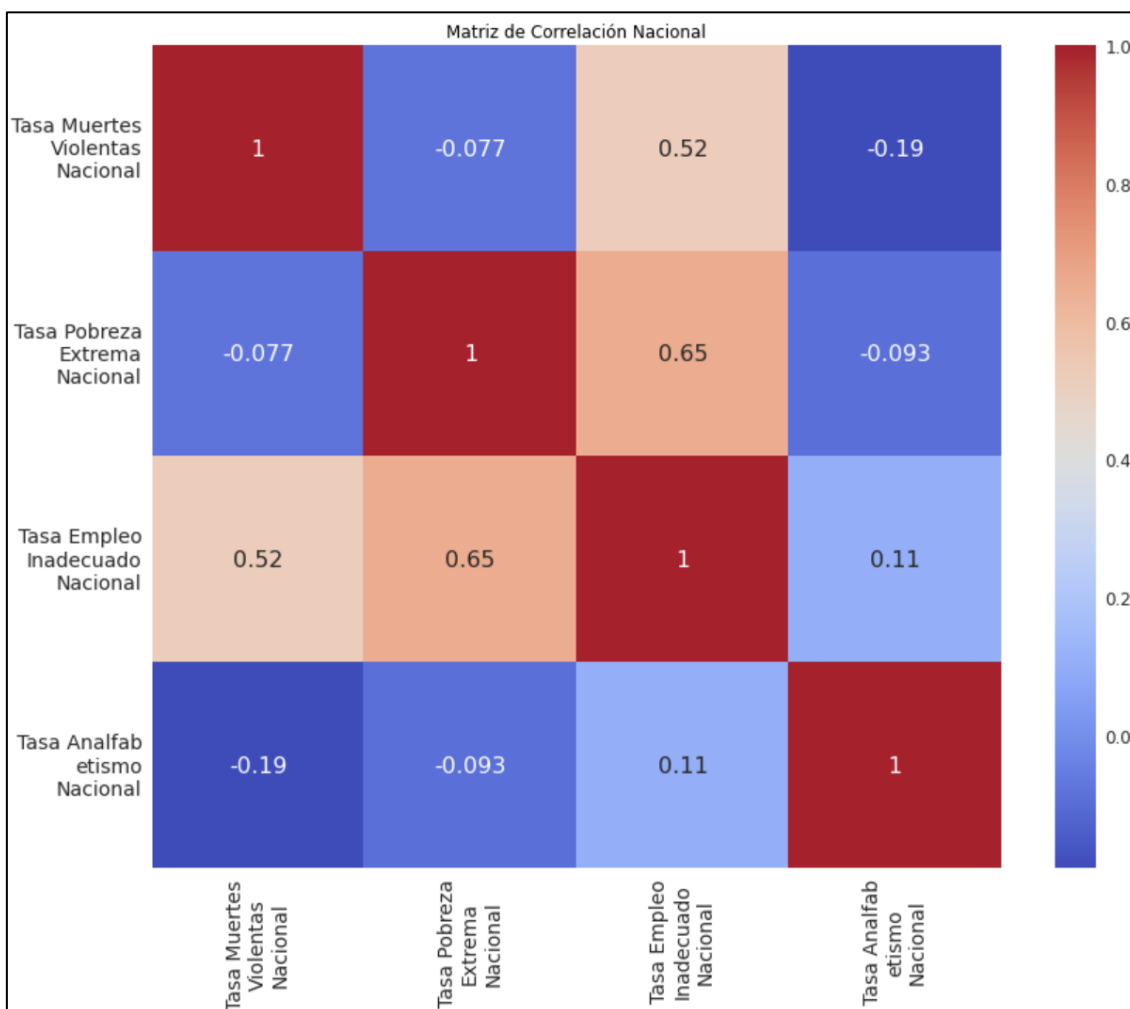


Figura 11: Matriz de Correlación Nacional  
Fuente: Elaboración propia

## Matriz de Correlación Provincial

Se realiza un análisis de las 5 provincias con el mayor porcentaje de muertes violentas en el Ecuador: Guayas (38,7%), Manabí (9,5%), Pichincha (8,9%), Los Ríos (8,4%) y Esmeraldas (8,1%). Estas provincias tienen un comportamiento similar a los resultados nacionales, a excepción de la provincia de Los Ríos, en la cual se presenta una correlación positiva entre la tasa de muertes violentas y la tasa de empleo inadecuado, esto podría significar que si existe un mayor porcentaje de empleo inadecuado las muertes violentas incrementan. Además de una correlación positiva entre la tasa de muertes violentas y la tasa de analfabetismo, esto podría significar que si existe un mayor porcentaje de analfabetismo las muertes violentas incrementan.

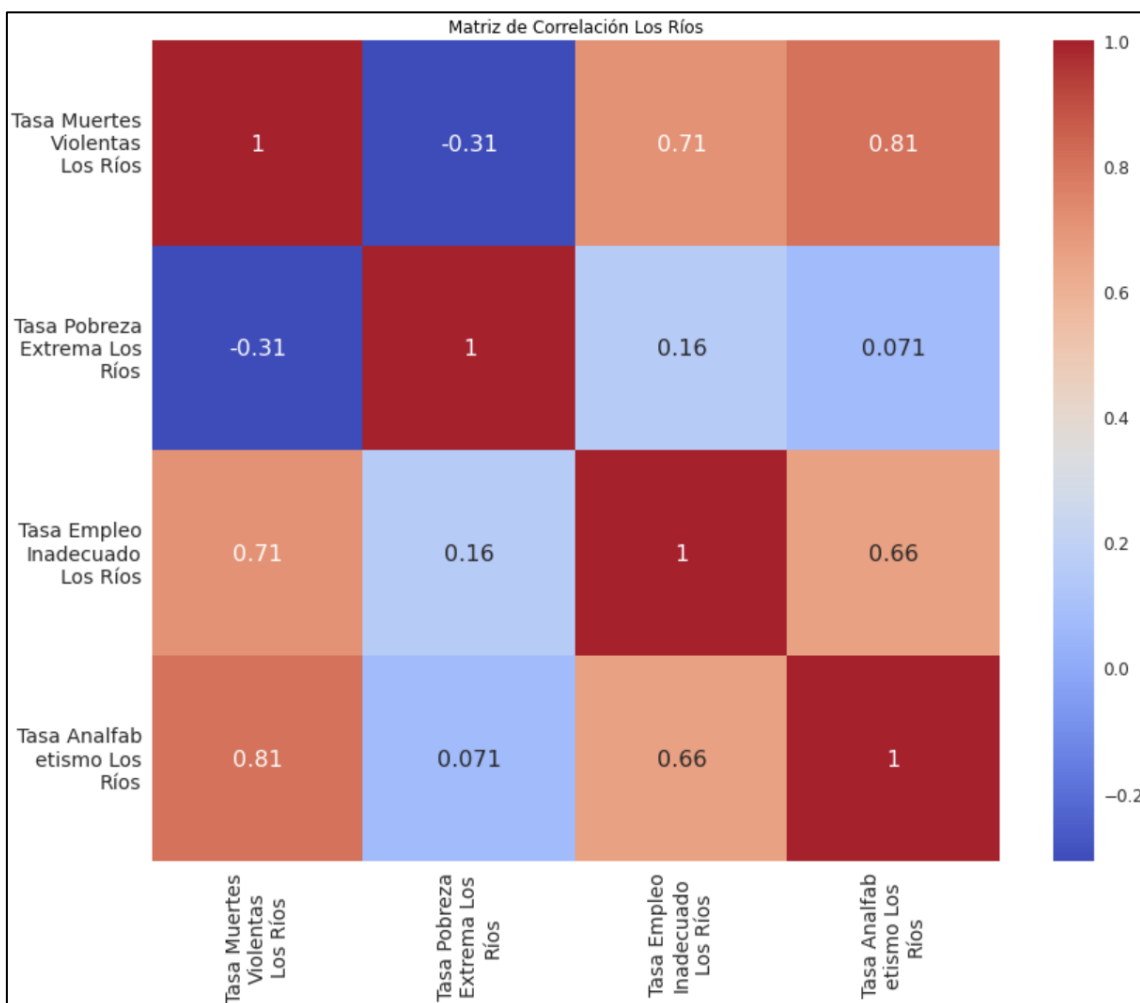


Figura 12: Matriz de Correlación Provincial  
Fuente: Elaboración propia

## **SELECCIÓN DEL MODELO ESTADÍSTICO**

Esta etapa es importante porque se define el modelo a utilizar, considerando que existen numerosos algoritmos de predicción de series temporales que se agrupan en “modelos estadísticos” y “de redes neuronales. En los modelos estadísticos, se indica que las relaciones subyacentes entre los datos son lineales, mientras que los modelos neuronales, descubren relaciones no lineales. Es importante definir que es una **serie temporal** y de acuerdo a lo mencionado por Hariharan, “Son un conjunto de datos que están organizados en función del tiempo, se puede presentar en forma de puntos, ya sea indexados, listados o representados en gráficos, y se ordenan de acuerdo a tiempos predefinidos. A diferencia de los datos de muestras aleatorias, los datos de series temporales son más predecibles y pueden contener información adicional que podemos extraer.” (2020).

La finalidad del proyecto es utilizar un modelo para realizar un pronóstico de una serie de datos de tiempo e inspeccionar los resultados (muertes violentas 2015-2022) y los principales modelos que cubren la necesidad de la investigación como ARIMA (Media Móvil Integrada Autorregresiva), SARIMA y el modelo aditivo Prophet.

Que como se mencionó son entre los métodos estadísticos que más destacan para calcular la regresión lineal. La familia de modelos ARIMA (Autoregressive Integrated Moving Average) con AR, MA, ARMA, ARIMA, ARIMAX, SARIMAX y finalmente el modelo Aditivo Prophet. A continuación, la definición de los modelos:

### **El modelo ARIMA**

Los modelos AR son autorregresivos y explican una serie de tiempo determinada en función de los valores pasados, modelando las series de tiempo que no son estacionales y que muestre patrones y no sea únicamente un "ruido blanco aleatorio", entendiéndose ruido blanco como “una sucesión de variables

aleatorias (proceso estocástico) con esperanza (media) cero, varianza constante e independientes para distintos valores de  $t$  (covarianza nula)” (Arce & Mahía, 2001).

De acuerdo con lo mencionado por Arce & Mahía, se define a un modelo como autorregresivo cuando “la variable intrínseca de un período ‘ $t$ ’ es explicada por las observaciones generados en periodos anteriores, dicho de otra manera, asume que los valores pasados de una serie temporal tienen un impacto en los valores futuros de esa misma serie.

En el caso de procesos estacionarios con distribución normal, la teoría estadística de los procesos estocásticos (sucesión de variables aleatorias  $Y_t$  ordenadas, en las que “ $t$ ” puede tomar cualquier valor entre  $-/+$  infinito). Bajo determinadas condiciones previas, toda  $Y_t$  puede expresarse como una combinación lineal de sus valores pasados (parte sistemática) más un término de error (innovación)” (2001). En términos más simples “es un modelo que se utiliza para predecir tendencias futuras en una serie de datos de tiempo” (Hariharan, 2020).

El modelo ARIMA ( $p,d,q$ ) se caracteriza por los términos:

- $p$ : orden de término AR (Auto regresivo): Número de rezagos de “ $y$ ” que se utiliza como predictores.
- $q$ : orden de término MA (Media móvil): número de errores de pronóstico retrasados.
- $d$ : diferenciaciones necesarias para que la serie de tiempo sea estacionaria.

Se escribe como:

$$y(t) = c + \phi_1 \cdot y(t - 1) + \phi_2 \cdot y(t - 2) + \dots + \phi_p \cdot y(t - p) + \theta_1 \cdot \epsilon(t - 1) + \theta_2 \cdot \epsilon(t - 2) + \dots + \theta_q \cdot \epsilon(t - q) + \epsilon(t)$$

*Modelo ARIMA ( $p,d,q$ ) Fuente: (Arce & Mahía, 2001)*

Donde:

$y(t)$  es el tiempo observado en el tiempo  $t$ .

$c$  es una constante o intercepto.

$\phi_1, \phi_2, \dots, \phi_p$  son los coeficientes de la parte autoregresiva del modelo.

$y(t - 1), y(t - 2), \dots, y(t - p)$  son los valores pasados de la serie.

$\theta_1, \theta_2, \dots, \theta_q$  son los coeficientes de la parte de promedio móvil del modelo.

$\epsilon(t - 1), \epsilon(t - 2), \epsilon(t - q)$  son los errores pasados de la serie.

$\epsilon(t)$  es el error aleatorio en el tiempo  $t$ .

### **El modelo aditivo de PROPHET**

El modelo aditivo de Prophet, de acuerdo a lo mencionado por Model Differently, “es una herramienta de pronóstico de series temporales desarrollada por Facebook que se basa en un enfoque aditivo en el que los efectos estacionales y las tendencias son modelados por componentes individuales; estos componentes incluyen un componente de tendencia, componentes estacionales (como diarios, semanales y anuales) y días festivos que pueden afectar los datos” (Análisis y Predicción de Series Temporales con FB Prophet Python, 2022).

Una de las principales características del modelo aditivo de Prophet es su capacidad para manejar datos faltantes y anomalías. Puede manejar de manera efectiva los valores perdidos y las interrupciones en los datos, lo que lo hace adecuado para datos del mundo real que a menudo son imperfectos. Adicional, el modelo aditivo de Prophet se basa en la suma o producto de distintas componentes: la tendencia, las componentes estacionales, la componente de festivos y las componentes del resto de regresores.

El efecto del resto de variables, ya sean estacionales o del resto de regresores, puede ser aditivo o multiplicativo.

Esto se traduce en que la componente se suma o se multiplica a la de tendencia.



- Modelo completamente aditivo:

$$y(t)=g(t)+s(t)+h(t)+r(t)+\epsilon t$$

- Modelo con estacionalidad multiplicativa:

$$y(t)=g(t)\cdot s(t)\cdot h(t)\cdot \epsilon t$$

*Modelo Aditivo de Prophet / Fuente: (Arce & Mahía, 2001)*

Donde:

$y(t)$  es la predicción,

$g(t)$  es la componente tendencia,

$s(t)$  es la estacionalidad,

$h(t)$  es la componente de los festivos,

$r(t)$  son los regresores y

$\epsilon t$  es el término de ruido blanco.

Una de las ventajas que tiene el modelo aditivo de Prophet es que es un modelo de fácil descomposición, lo que permite que la interpretabilidad sea alta, adicionando a que se facilita la extracción de los coeficientes de cada una de sus componentes, por tal motivo se recomienda su uso, para el estudio de estacionalidades, tendencias, días festivos y demás variables.

## **SARIMA**

Refleja la característica de variación estacional en series temporales. La serie de tiempo original  $\{Y_t\}$  utiliza un operador de retardo  $B$ , para procesar SARIMA  $(p, d, q) \times (P, D, Q)_s$ .

Un modelo SARIMA, puede escribirse como:

$$\phi(B)\phi_p(B^S)(1-B)^d(1-B^S)^DY_t = \theta_q(B)\Theta_Q(B^S)\epsilon_t$$

$B$  es el operador de retraso (definido como  $BkY_t = Y_t - k$ )

$$\phi_p(B) = 1 - \phi_1B - \phi_2B^S - \dots - \phi_pB^p$$

$$\Phi_p(B^S) = 1 - \phi_sB^S - \phi_{2s}B^{2s} - \dots - \phi_{ps}B^{Ps}$$

$$\theta_q(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$$

$$\Theta_Q(B^S) = 1 - \theta_sB^S - \theta_{2s}B^{2s} - \dots - \theta_{Qs}B^{Qs}$$

Modelo Aditivo de Prophet / Fuente: (Miranda, 2021)

Donde:

$\phi(B)$  y  $\theta(B)$  son polinomios de orden  $p$  y  $q$ ,

$\Phi(B^S)$  y  $\Theta(B^S)$  son polinomios en  $B^S$  de grados  $P$  y  $Q$ , respectivamente:

$p$  es el orden de autorregresión no estacional;

$d$  es el número de diferencias regulares;

$q$  es el orden de la media móvil no estacional;

$P$  es el orden de autorregresión estacional;

$D$  es el número de diferencias estacionales;

$Q$  es el orden de la media móvil estacional; y

$S$  es la duración de la temporada.

A continuación, se detalla las ventajas y desventajas de los modelos Prophet, SARIMA y ARIMA:

**Tabla 8: Ventajas y desventajas de los modelos Prophet, SARIMA y ARIMA:**

Modelo	VENTAJAS	DESVENTAJAS
ARIMA	<ul style="list-style-type: none"> <li>Flexibilidad en la modelización, porque permite modelar componentes de tendencia y estacionalidad en series temporales, siendo muy útil para diferentes situaciones.</li> <li>Históricamente es el más utilizado, es un modelo clásico en el análisis de series de tiempo, proporciona una base</li> </ul>	<ul style="list-style-type: none"> <li>Este modelo asume que la serie de tiempo es estacionaria, lo que ocasiona un limitante cuando la serie de tiempo presenta cambios de tendencia o una estacionalidad no lineal.</li> <li>Requieren técnicas adicionales para ajustar el impacto que representa los datos atípicos</li> <li>La selección manual de órdenes que requiere ARIMA es un proceso que</li> </ul>

	<p>sólida de conocimiento y recursos.</p> <ul style="list-style-type: none"> <li>• Se puede interpretar fácilmente y claro los coeficientes, facilitando la lectura entre las variables y sus relaciones con la finalidad de tomar buenas decisiones.</li> </ul>	<p>demanda tiempo en la elección de los órdenes p, d y q.</p>
Prophet	<ul style="list-style-type: none"> <li>• Fácil implementación porque es una biblioteca de código abierto desarrollada por Facebook, presenta un fácil diseño e implementación e incluso puede ser utilizado por personas sin experiencia en series de tiempo.</li> <li>• El manejo de tendencias y estacionalidad lo realiza automáticamente, identificando patrones de tendencia y estacionalidad en los datos, lo que lo convierte en adecuado al momento de predecir series de tiempo que presente comportamientos no lineales.</li> <li>• Manejo de datos faltantes y atípicos, debido a la robustez que presenta frente a la aparición de los mismos, utiliza métodos de interpolación y técnicas de modelado estadístico para tratar con ellos.</li> </ul>	<ul style="list-style-type: none"> <li>• Limitado en datos de alta frecuencia: Prophet no se desempeña bien con datos de alta frecuencia, ya que está más orientado hacia datos de frecuencia diaria o mensual.</li> <li>• Requiere ajuste de hiperparámetros para obtener resultados óptimos.</li> </ul>
SARIMA	<ul style="list-style-type: none"> <li>• Permite modelar tanto al componente estacional como los componentes de tendencia y estacionalidad en una serie de tiempo, por lo que es recomendado para modelar datos con patrones estacionales complejos.</li> <li>• Flexibilidad en la modelización, se puede ajustar a diferentes patrones de series de tiempo.</li> <li>• Presenta modelos interpretables, lo que permite comprender las relaciones entre las variables.</li> </ul>	<ul style="list-style-type: none"> <li>• Requiere estacionariedad.</li> </ul>

*Fuente: Elaboración propia*

Adicional, se detalla un cuadro comparativo de los modelos:

**Tabla 9:Cuadro comparativo ARIMA - Aditivo de Prophet - SARIMA**

	ARIMA	Prophet	SARIMA
Utilidad	Análisis y pronóstico de series temporales.	Análisis y pronóstico de series temporales.	Pronostica y analiza series temporales que presentan patrones estacionales y tendencias.
Enfoque de modelado	Modelo estadístico que se basa en la teoría de autocorrelación y estacionariedad de los datos.	Modelo desarrollado por Facebook que utiliza un enfoque aditivo y busca capturar tanto las tendencias no lineales como las componentes estacionales y los efectos de días festivos.	Modelo estadístico y econométrico utilizado para predecir las series temporales, presenta componentes autoregresivos (AR), de media móvil (MA) e integrados (I).
Estacionalidad	Se aborda a través del modelo SARIMA (ARIMA estacional) al incluir componentes estacionales en el modelo.	Se modela de forma automática al agregar componentes estacionales (diarios, semanales y anuales) en el enfoque aditivo.	Modela la estacionalidad en los datos, tomando en cuenta patrones estacionales (fluctuaciones) que son incorporados en los análisis, dando como resultados pronósticos más precisos.
Requerimientos de estacionariedad	Implica que la serie temporal sea estacionaria o se pueda hacer estacionaria mediante la diferenciación.	Puede manejar series temporales que no son estrictamente estacionarias y realiza internamente una descomposición aditiva para separar la tendencia y la estacionalidad.	La media y la varianza de la serie no varían significativamente a lo largo del tiempo, lo que facilita el modelado y la predicción.
Interpretabilidad y ajuste automático	Es un modelo más tradicional lo que permite que su interpretación sea mucho más fácil, al igual que los ajustes manuales de los parámetros.	Proporciona una interfaz simple y realiza automáticamente muchas tareas como la identificación de tendencias y estacionalidades.	Ofrece una interpretación clara de los coeficientes y proporciona métodos automáticos para seleccionar los órdenes adecuados de AR, MA, I.

*Fuente: (Arce & Mahía, 2001)/ Elaboración propia*

En base a las comparaciones realizadas, ventajas y desventajas que presenta cada modelo, se puede concluir que ARIMA y SARIMA son un modelo estadístico que se centra en relaciones lineales y estacionariedad, mientras que el modelo aditivo de Prophet es un modelo más moderno y accesible que busca capturar tendencias no lineales y efectos estacionales.

De acuerdo a los resultados obtenidos de los diferentes modelos, el modelo aditivo de Prophet es el que mejor desempeño muestra en términos de precisión.

**Tabla 10: resultado de los modelos**

<b>Modelo</b>	<b>MAE</b>	<b>MSE</b>	<b>RMSE</b>	<b>R<sup>2</sup></b>
Prophet	2.276.523	8.646.264	2.940.453	0.457928
ARIMA	4.034.597	30.818.300	5.551.423	-0.312940
SARIMA	7.165.651	74.907.608	8.654.918	-2.191.260

*Fuente: Elaboración propia*

El modelo aditivo de Prophet presenta el menor valor de MAE (2.276523), lo que indica que tiene la menor discrepancia promedio entre las predicciones y los valores reales. El modelo ARIMA tiene un MAE de 4.034597, mientras que el modelo SARIMA tiene el mayor MAE de 7.165651.

El MSE (Mean Squared Error) y al RMSE (Root Mean Squared Error), para el modelo aditivo de Prophet tiene los valores más bajos, lo que implica que sus predicciones tienen un mejor ajuste a los datos reales. El modelo ARIMA tiene un MSE de 30.818300 y un RMSE de 5.551423, mientras que el modelo SARIMA muestra un MSE de 74.907608 y un RMSE de 8.654918.

Respecto al R<sup>2</sup>, el modelo aditivo de Prophet tiene el valor más alto de 0.457928, lo que indica que explica mejor la variabilidad de los datos en comparación con los otros modelos. El modelo ARIMA tiene un R<sup>2</sup> de -0.312940 y el modelo SARIMA tienen el valor más bajo de R<sup>2</sup> de -2.191260, probablemente estos modelos no se ajustan bien a los datos.

## 8. RESULTADOS

### 8.1 ANÁLISIS DE MODELO ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS

Se implementa el modelo aditivo Prophet para el pronóstico de series temporales, estableciendo un 95% de confianza para realizar el pronóstico. Se establecen 36 meses como la cantidad de períodos futuros que se desea pronosticar.

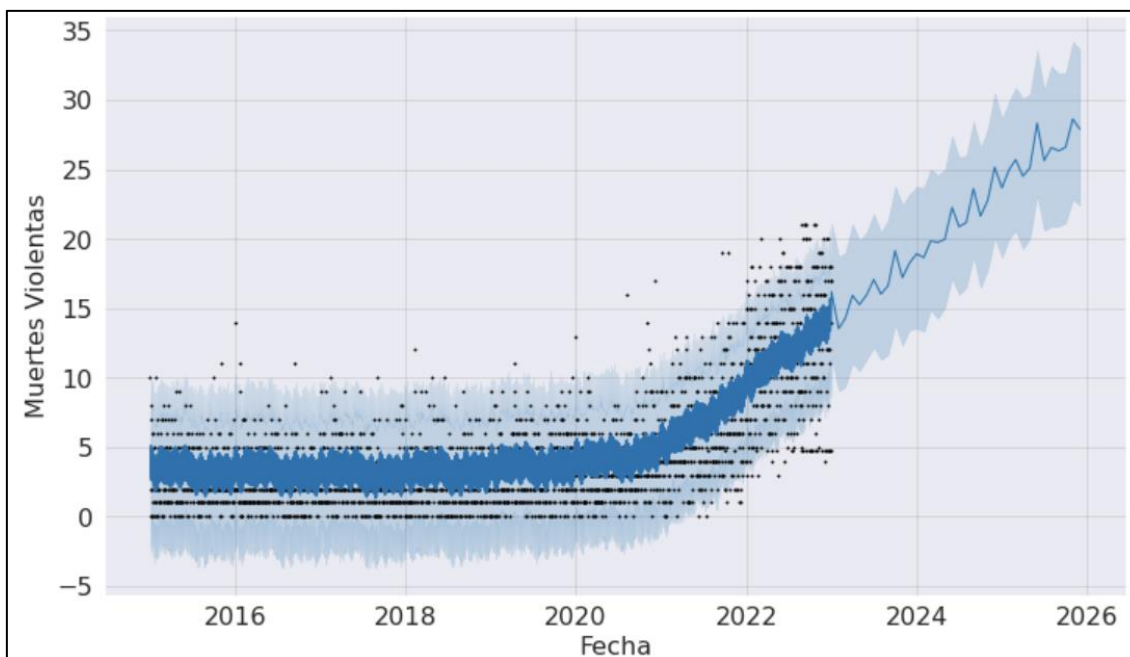
Al implementar el modelo se obtienen los valores pronosticados y límites inferior y superior del intervalo de confianza asociado con el pronóstico.

**Tabla 11: Valores pronosticados y límites inferior y superior del intervalo de confianza**

Fecha	yhat (Valores pronosticados)	yhat_lower (Límite inferior del intervalo de confianza)	yhat_upper (Límite superior del intervalo de confianza)
1/1/2015	0,29	-1,07	1,64
2/1/2015	0,36	-0,97	1,89
3/1/2015	0,39	-1,04	1,87
4/1/2015	0,59	-0,97	2,07
5/1/2015	0,28	-1,13	1,67
...	...	...	...
1/8/2025	2,81	1,42	4,32
1/9/2025	2,68	1,19	4,08
1/10/2025	2,74	1,32	4,18
1/11/2025	2,95	1,50	4,37
1/12/2025	2,94	1,51	4,35

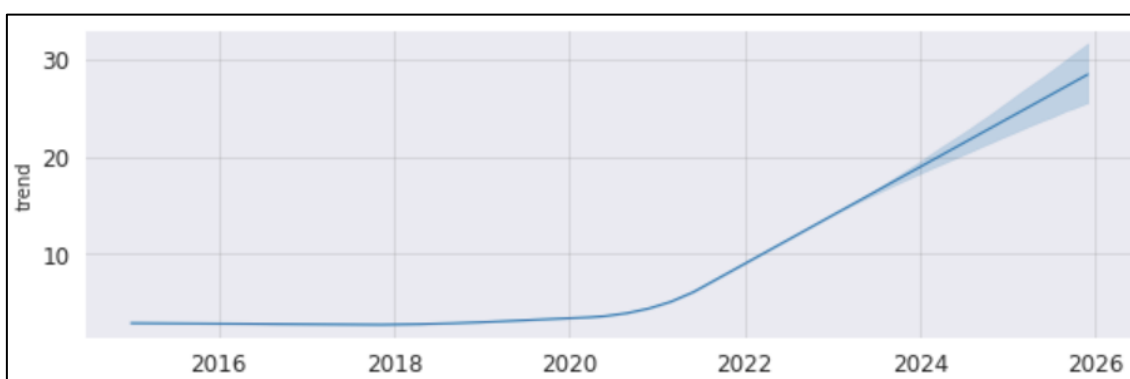
*Fuente: Elaboración propia*

En el gráfico de predicción diaria, se puede evidenciar que existe una tendencia creciente en las muertes violentas para los años 2023 a 2025:



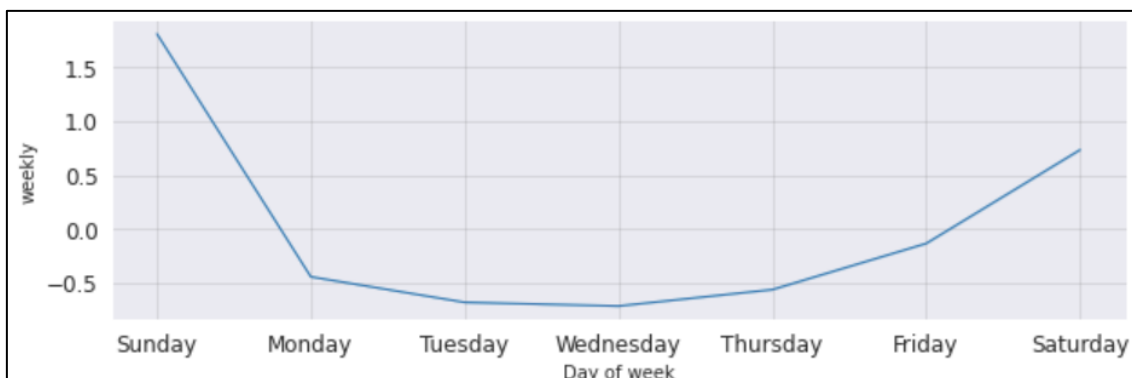
*Figura 13: Predicción diaria muertes violentas  
Fuente: Elaboración propia*

En la predicción anual presenta una tendencia creciente e irá incrementando de manera lineal a lo largo del tiempo y se evidencia que a partir del año 2024 existe una mayor dispersión en las posibles predicciones.

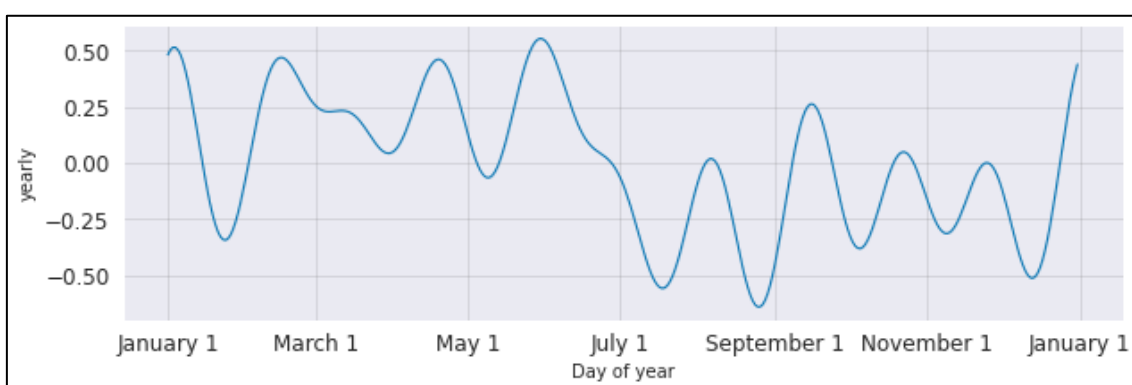


*Figura 14: Predicción anual muertes violentas  
Fuente: Elaboración propia*

De acuerdo con las predicciones por día de la semana, los días de la semana que mayor número de muertes violentas presentan son domingo y sábado.



*Figura 15: Predicción día de muertes violentas  
Fuente: Elaboración propia*

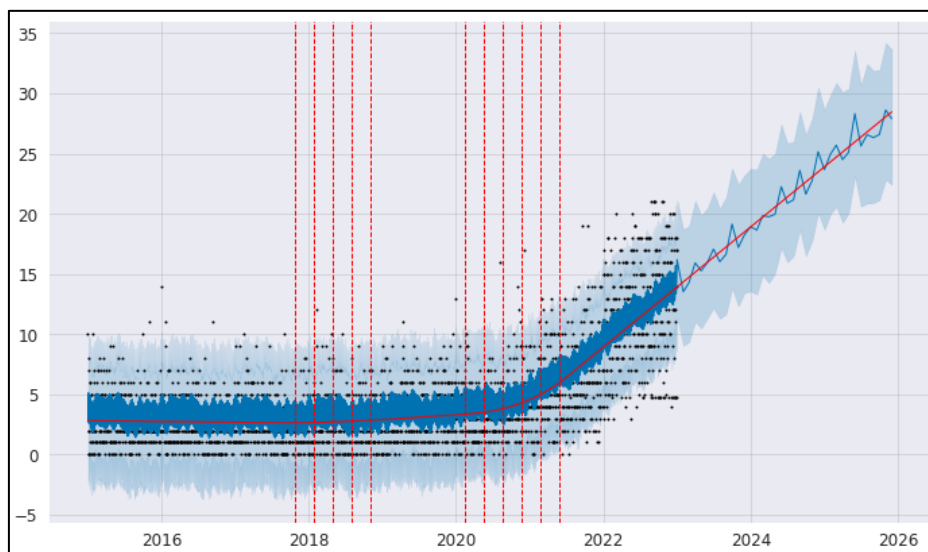


*Figura 16: Predicción mes de muertes violentas  
Fuente: Elaboración propia*

De acuerdo con las predicciones mensuales, los meses con mayor número de muertes violentas son diciembre, seguido de los meses: junio, septiembre, octubre y noviembre.

El modelo aditivo de Prophet incorpora puntos de cambio (changepoints) en los cuales las series temporales presentan cambios significativos en la serie de tiempo, de manera predeterminada el modelo identifica estos cambios y ajusta la tendencia, la estacionalidad y otros componentes para adaptar a la serie de tiempo pronosticada. Se realizan pruebas, con fechas específicas que tienen un alto número de muertes violentas y se evidencia que no existe una variación significativa en los resultados.

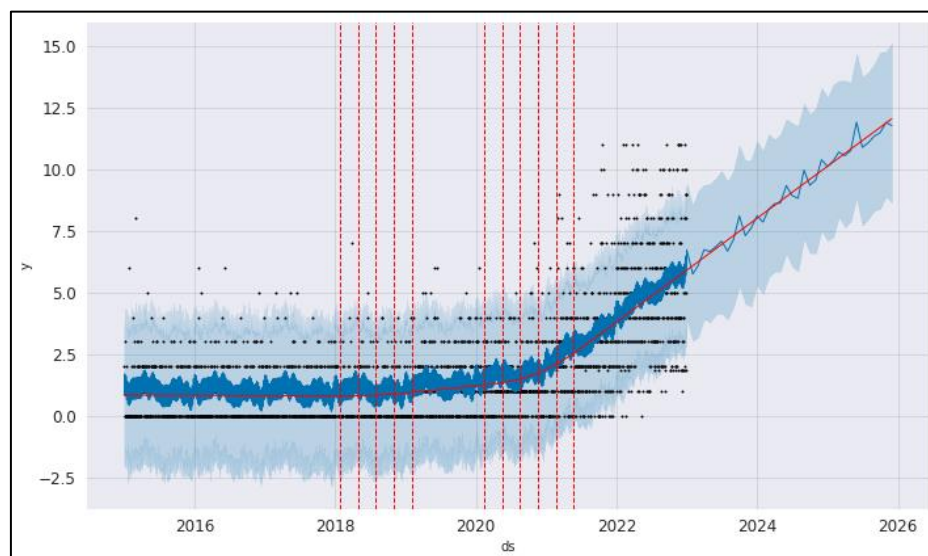




*Figura 17: Changepoints Nacional  
Fuente: Elaboración propia*

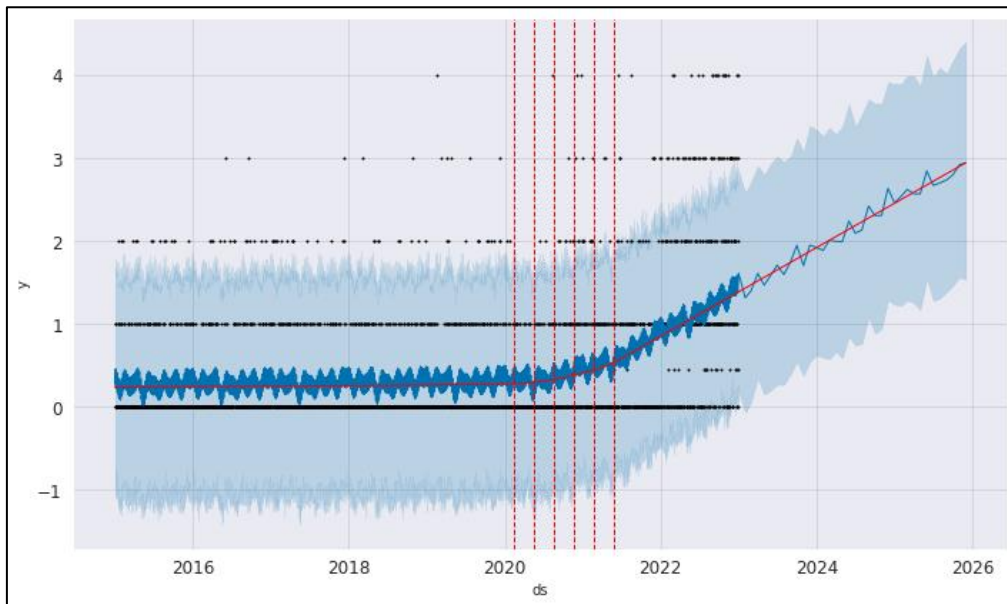
Luego de realizar un análisis predictivo a nivel Nacional, se realiza el mismo análisis con las cinco provincias que tienen el mayor porcentaje de participación de muertes violentas a nivel nacional: Guayas (38,7%), Manabí (9,5%), Pichincha (8,9%), Los Ríos (8,4%) y Esmeraldas (8,1%).

En la provincia de Guayas se utilizan puntos de cambio que se identifican de manera automática entre los años 2018 y 2022 y que sirven como insumo para mejorar la precisión del modelo.



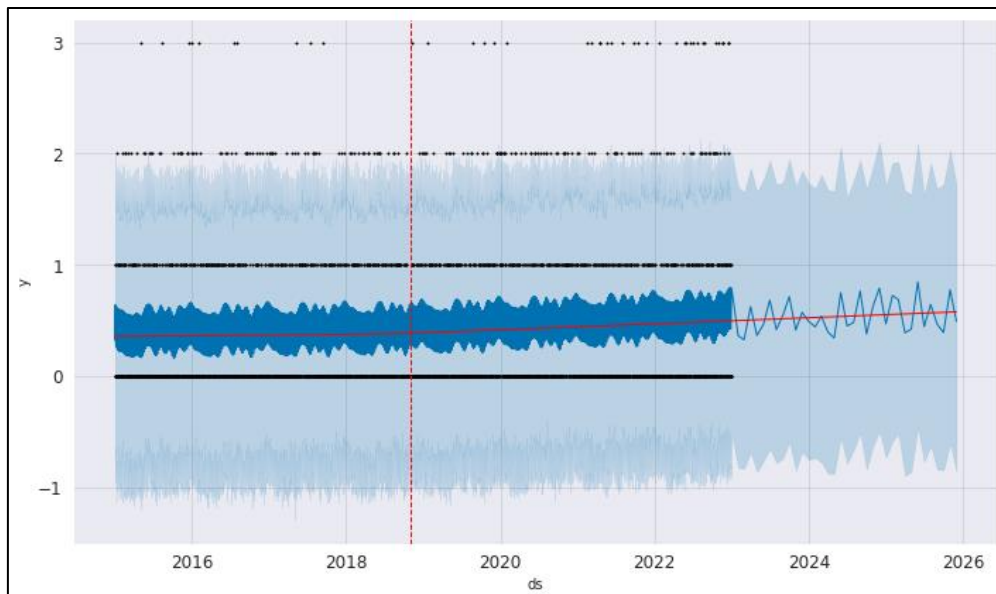
*Figura 18: Changepoints Guayas  
Fuente: Elaboración propia*

En la provincia de Manabí se utilizan seis puntos de cambio que se identifican de manera automática entre los años 2020 y 2022.



*Figura 19: Changepoints Manabí*  
*Fuente: Elaboración propia*

En la provincia de Pichincha se identifica automáticamente un punto de cambio entre los años 2018 y 2019 que sirve como insumo para mejorar la precisión del modelo.



*Figura 20: Changepoints Pichincha*  
*Fuente: Elaboración propia*

En la provincia de Los Ríos entre los años 2020 y 2022 se identifican puntos de cambio entre los años 2020 y 2022.

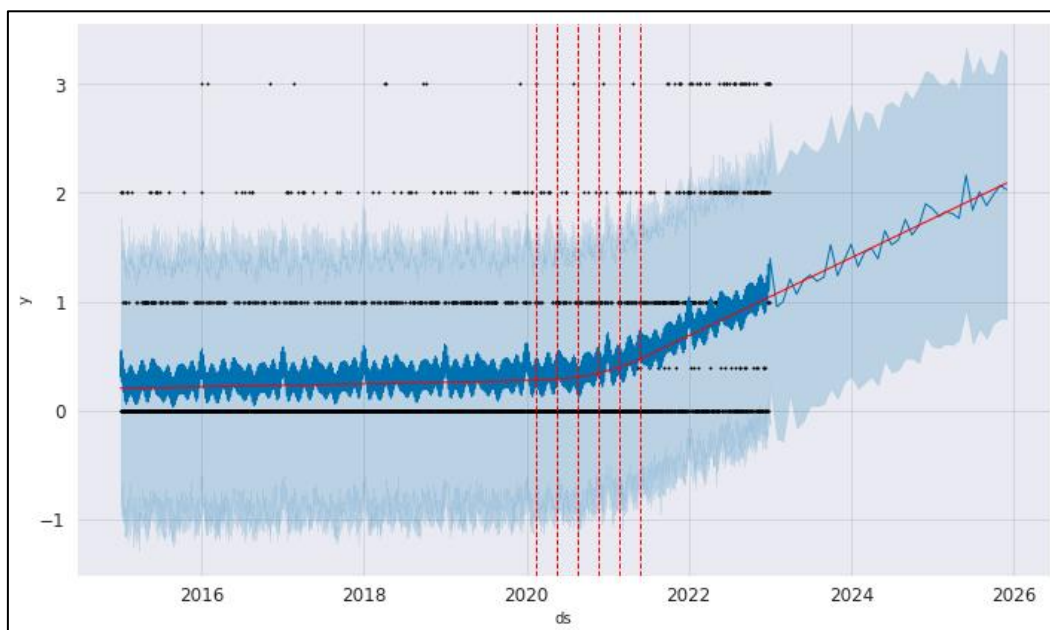


Figura 21: Changepoints Los Ríos  
Fuente: Elaboración propia

En la provincia de Esmeraldas se utilizan puntos de cambio que se identifican de manera automática entre los años 2020 y 2022 y que sirven como insumo para mejorar la precisión del modelo.

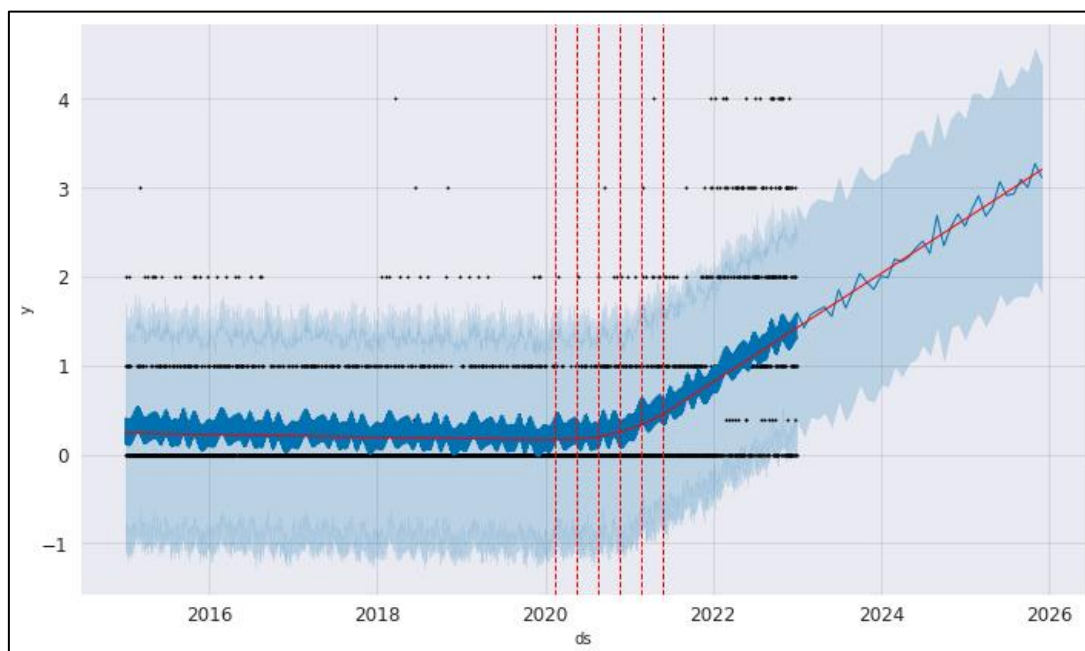
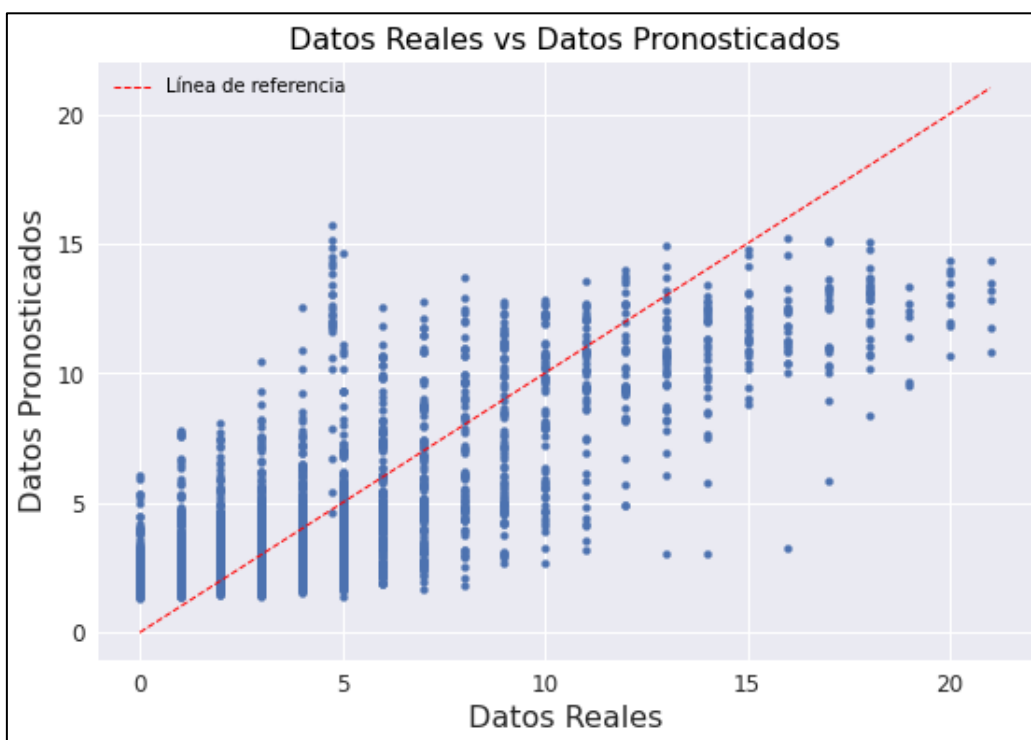


Figura 22: Changepoints Esmeraldas  
Fuente: Elaboración propia

### Gráfico comparativo de datos reales vs. Pronosticados

A nivel nacional el comportamiento de la predicción es relativamente bueno a nivel general, un poco menos en las colas. Predice valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo.



*Figura 23: Datos Reales VS Pronosticados  
Fuente: Elaboración propia*

En la provincia de Guayas el comportamiento de la predicción es relativamente bueno a nivel general. Predice valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo.

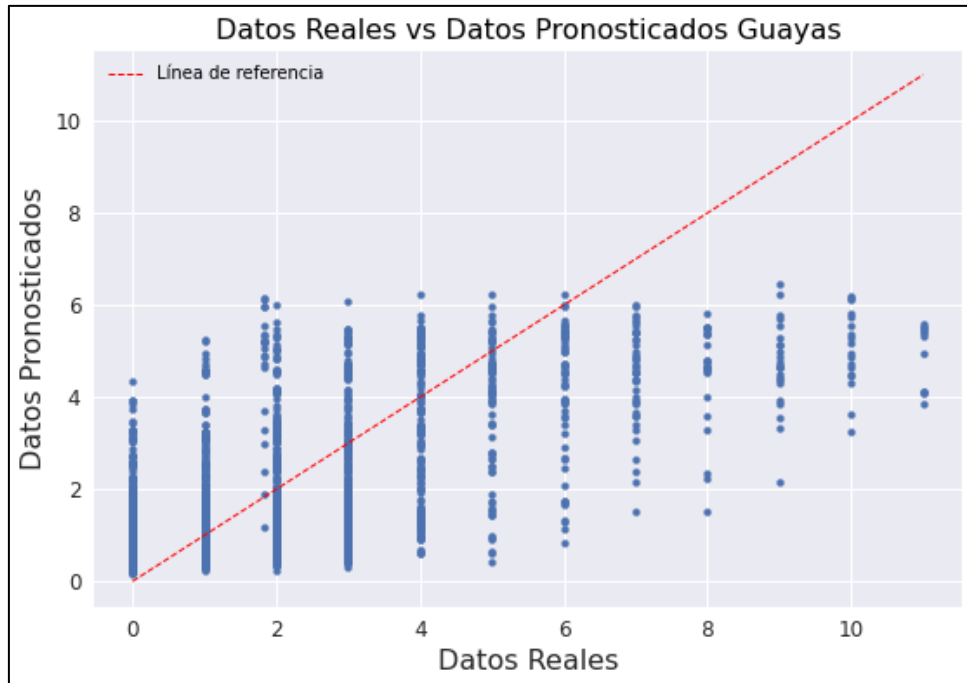


Figura 24: Datos Reales VS Pronosticados Guayas  
Fuente: Elaboración propia

En la provincia de Manabí el comportamiento de la predicción presenta valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo. Existen puntos dispersos y alejados de la línea de referencia, puede indicar discrepancias significativas entre los valores pronosticados y los valores reales.

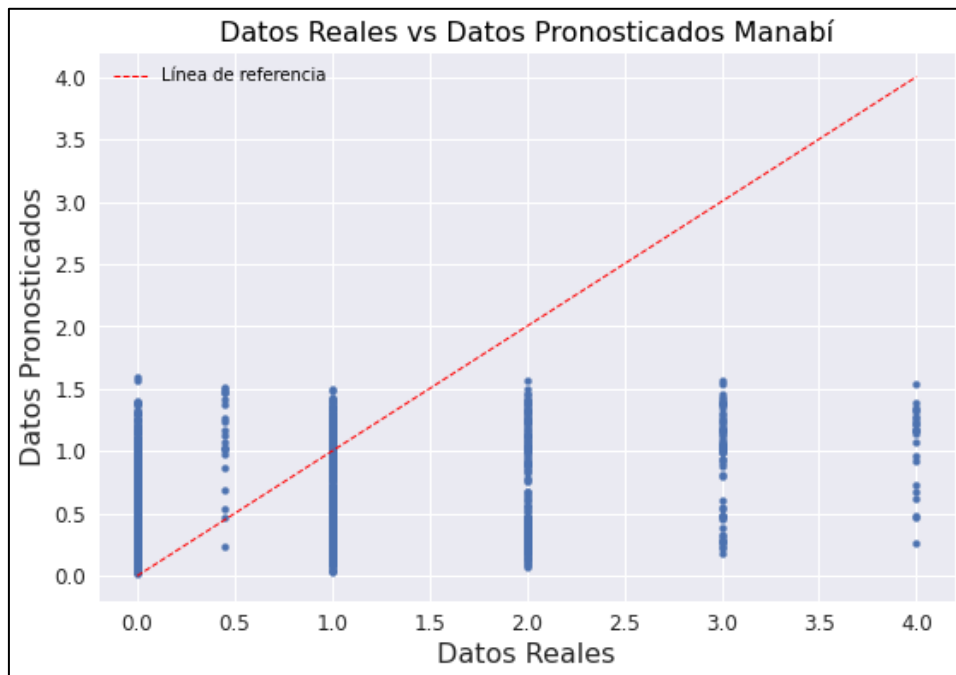


Figura 25: Datos Reales VS Pronosticados Manabí  
Fuente: Elaboración propia

En la provincia de Pichincha el comportamiento de la predicción presenta valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo. Existen puntos dispersos y alejados de la línea de referencia, puede indicar discrepancias significativas entre los valores pronosticados y los valores reales.

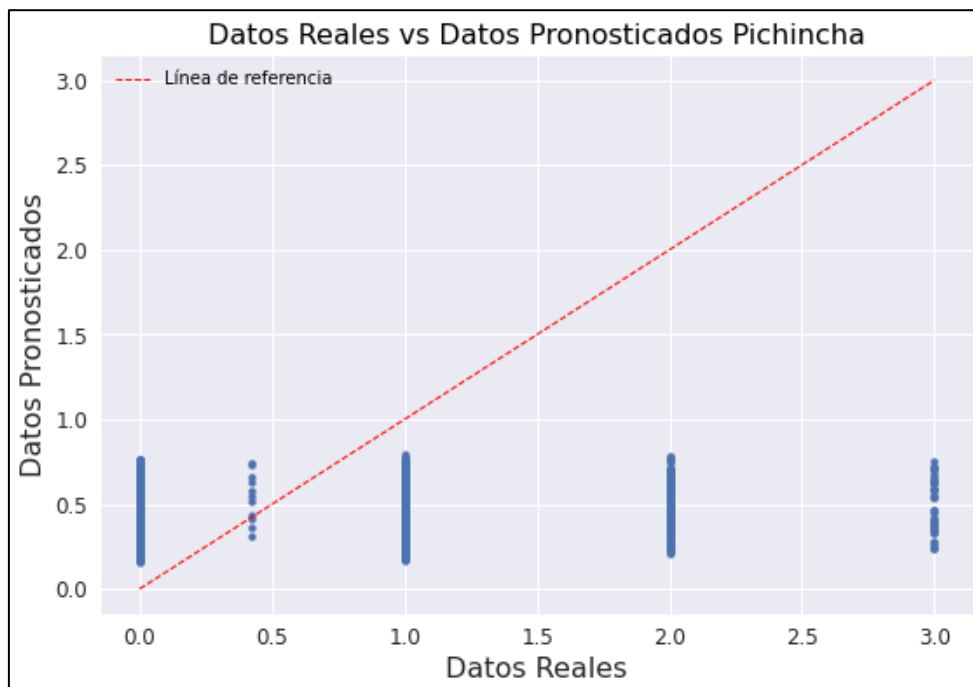


Figura 26: Datos Reales VS Pronosticados Pichincha  
Fuente: Elaboración propia

En la provincia de Los Ríos el comportamiento de la predicción presenta valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo. Existen puntos dispersos y alejados de la línea de referencia, puede indicar discrepancias significativas entre los valores pronosticados y los valores reales.

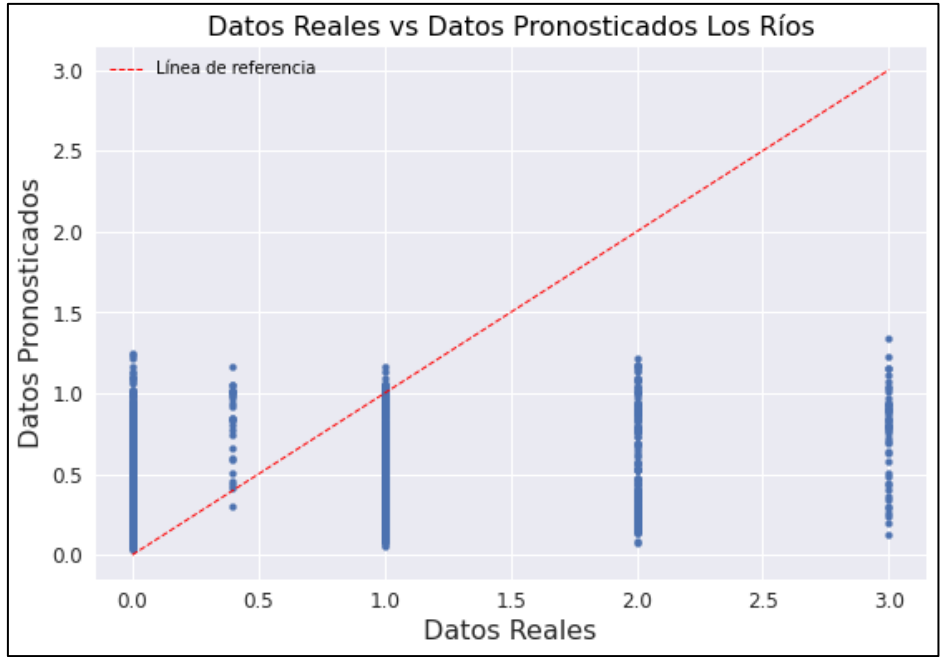


Figura 27: Datos Reales VS Pronosticados Los Ríos  
Fuente: Elaboración propia

En la provincia de Esmeraldas el comportamiento de la predicción presenta valores más altos cerca de cero y a medida que se acerca al máximo de la cola derecha, empieza a predecir más bajo. Existen puntos dispersos y alejados de la línea de referencia, puede indicar discrepancias significativas entre los valores pronosticados y los valores reales.

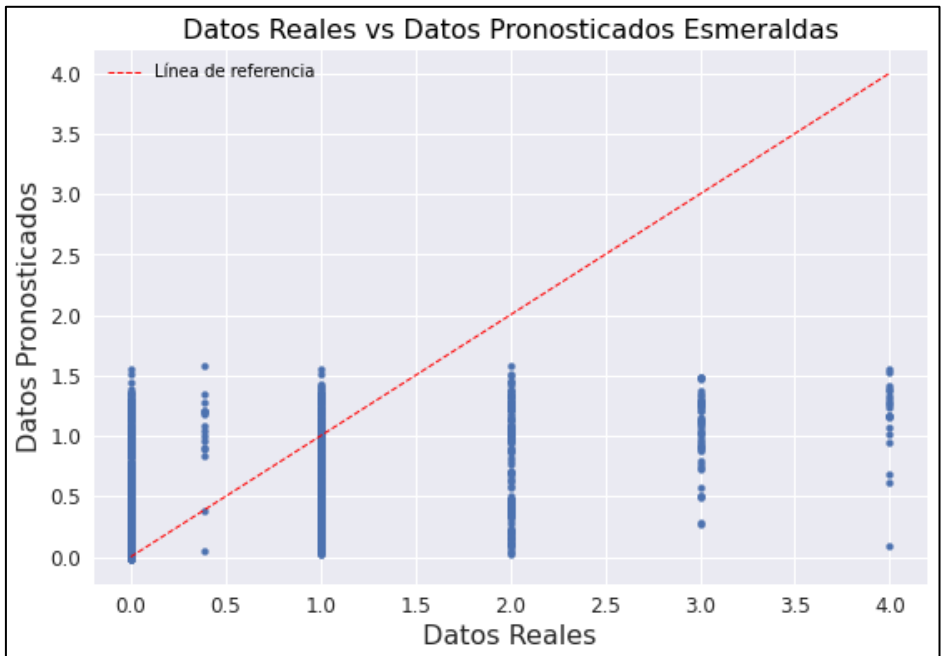


Figura 28: Datos Reales VS Pronosticados Esmeraldas  
Fuente: Elaboración propia

### Porcentaje de variación de las predicciones (2025) respecto al año inicial (2015) y final (2022) de los datos reales

De acuerdo con los resultados de las predicciones nacionales, en el año 2025 se incrementará un 844% de muertes violentas respecto al año 2015 y un 129% respecto al año 2022.

**Tabla 12: Predicción Nacional**

NACIONAL			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	2,77		
2022	11,42	312%	
2025	26,15	844%	129%

*Fuente: Elaboración propia*

De acuerdo con los resultados de las predicciones en la provincia de Guayas, en el año 2025 se incrementará un 1.233% de muertes violentas respecto al año 2015 y un 128% respecto al año 2022.

**Tabla 13: Predicciones en la provincia de Guayas**

GUAYAS			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	0,83		
2022	4,85	484%	
2025	11,06	1233%	128%

*Fuente: Elaboración propia*

De acuerdo con los resultados de las predicciones en la provincia de Manabí, en el año 2025 se incrementará un 1.025% de muertes violentas respecto al año 2015 y un 141% respecto al año 2022.



**Tabla 14: Predicciones en la provincia de Manabí**

MANABÍ			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	0,24		
2022	1,12	367%	
2025	2,70	1025%	141%

Fuente: Elaboración propia

De acuerdo con los resultados de las predicciones en la provincia de Pichincha, en el año 2025 se incrementará un 56% de muertes violentas respecto al año 2015 y un 17% respecto al año 2022.

**Tabla 15: Predicciones en la provincia de Pichincha**

PICHINCHA			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	0,36		
2022	0,48	33%	
2025	0,56	56%	17%

Fuente: Elaboración propia

De acuerdo con los resultados de las predicciones en la provincia de Los Ríos, en el año 2025 se incrementará un 814% de muertes violentas respecto al año 2015 y un 123% respecto al año 2022.

**Tabla 16: Predicciones en la provincia de Los Ríos**

LOS RÍOS			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	0,21		
2022	0,86	310%	
2025	1,92	814%	123%

Fuente: Elaboración propia

De acuerdo con los resultados de las predicciones en la provincia de Esmeraldas, en el año 2025 se incrementará un 1.157% de muertes violentas respecto al año 2015 y un 161% respecto al año 2022.

**Tabla 17: Predicciones en la provincia de Esmeraldas**

ESMERALDAS			
Año	Promedio muertes diarias	Porcentaje de variación respecto al año 2015	Porcentaje de variación respecto al año 2022
2015	0,23		
2022	1,12	382%	
2025	2,92	1157%	161%

Fuente: Elaboración propia

Comparando los resultados a nivel Nacional y de las 5 provincias analizadas, se evidencia que la provincia con el mayor porcentaje de incremento en la predicción de muertes violentas en el año 2025 respecto al año 2022 es Esmeraldas (166%), seguido de Manabí (141%), Nacional (129%), Guayas (128%), Los Ríos (123%) y por último la provincia de Pichincha (17%).

**Tabla 18: Comparación Predicción VS provincial (5)**

Provincia	Porcentaje de variación respecto al año 2022
NACIONAL	129%
GUAYAS	128%
MANABÍ	141%
PICHINCHA	17%
LOS RÍOS	123%
ESMERALDAS	161%

Fuente: Elaboración propia

## 8.2 PROPUESTA DE SOLUCIÓN

### 8.2.1 IMPLICACIONES ORGANIZACIONALES

Este proyecto busca plantear diferentes soluciones con herramientas que permitirán abordar esta problemática que, a nivel de país tiene alto impacto. A continuación, se detallan algunas opciones de acuerdo a las diferentes perspectivas con las que se realizó el análisis y que se podrían implementar o serviría de guía para las que las organizaciones pertinentes, aprovechen al máximo el poder del análisis de datos en su toma de decisiones:

➤ **Planificación estratégica:**

Los resultados obtenidos respaldan la existencia de una conexión entre la determinación de áreas afectadas por la violencia, ya sea geográfica o demográficamente. Por lo tanto, es crucial realizar una asignación eficiente de recursos para abordar de manera más precisa los problemas de violencia en el país principalmente en las provincias en donde se registra mayor porcentaje de muertes violentas, en este caso son: Guayas, Manabí, Pichincha, y los Ríos.

➤ **Políticas de seguridad:**

Los resultados de la investigación proporcionan información clave para aplicar políticas públicas de seguridad que sean eficientes y eficaces, por ejemplo, en la zona geográfica considerada como de alto impacto y que se encuentra nuestro grupo de interés se podría diseñar espacios públicos seguros, la instalación de sistemas de vigilancia, el mejoramiento del alumbrado y la promoción de la participación comunitaria, de igual forma aplicar estrategias de prevención, que se centren a los principales factores de riesgo que se identificaron e implementando programas de prevención del delito, educación en valores y resolución de conflictos, fortalecimiento

comunitario, mejora del sistema de justicia penal, promoción de políticas de control de armas, entre otros.

➤ **Colaboración y coordinación interinstitucional:**

En base a los resultados sería posible impulsar la creación de alianzas estratégicas entre entidades gubernamentales, organizaciones no gubernamentales, sector privado, público y las comunidades que son afectadas, permitiendo aprender de las experiencias, para generar soluciones innovadoras y compartir buenas prácticas.

En base a los resultados sería posible impulsar la creación de alianzas estratégicas entre entidades gubernamentales, organizaciones no gubernamentales, sector privado, público y las comunidades que son afectadas, permitiendo aprender de las experiencias, para generar soluciones innovadoras y compartir buenas prácticas.

Incentivar la generación de campañas con programas educativos dirigidos a la comunidad en general o a grupos específicos (grupo de interés).

➤ **Relación entre estrategia, innovación y competitividad:**

Independientemente del modelo utilizado para realizar el análisis, la estrategia direcciona y da el enfoque para realizar el análisis, la innovación implica la adopción de enfoques novedosos y creativos en su aplicación, y la competitividad se logra al utilizar el análisis realizado de manera efectiva en comparación con otros enfoques.

- Estrategia: implicó la planificación y toma de decisiones sobre la elección y uso de este modelo que permite pronosticar series de tiempo y comprender los homicidios. Esto incluye la selección de variables relevantes, la determinación de los objetivos y la identificación de acciones para abordar la prevención de homicidios. Una estrategia sólida y bien definida es fundamental

para guiar el análisis y garantizar que los resultados obtenidos sean utilizados de manera efectiva para la toma de decisiones.

- **Innovación:** Principalmente se refiere a la adopción de diferentes enfoques, para utilizar el modelo aditivo de Prophet. Implica el uso de nuevas variables predictoras, la experimentación con diferentes configuraciones y parámetros del modelo, la incorporación de técnicas y herramientas complementarias como Python y ArcGis, que permitieron una mejor comprensión de los patrones y las tendencias relacionadas con los homicidios.
- **Competitividad:** se refiere a la capacidad de utilizar esta herramienta de manera efectiva para obtener ventajas frente a otros enfoques o estudios, aprovechando al máximo su capacidad de modelado y pronóstico de manera precisa. La competitividad se logra al utilizar el análisis del modelo aditivo de Prophet de manera estratégica e innovadora para obtener información valiosa que pueda informar las decisiones de prevención de homicidios de manera más efectiva que otros modelos como los descartados anteriormente SARIMA y ARIMA.

### **8.2.2 DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN**

Al realizar el análisis exploratorio de los datos a nivel nacional, se identificó la estacionalidad anual, mensual y semanal de los mismos y se aplicó el análisis con el modelo aditivo de Prophet, considerando la estricta condición de renombrar las columnas, se realizó la agrupación a nivel de Cantidad de Muertes Violentas y Fechas, con una predicción de 36 meses desde enero de 2023 a diciembre de 2025. En donde se identifica lo siguiente:

- La proyección anual que presenta una tendencia creciente de muertes violentas generadas en los diferentes años, indica que se irá aumentando de manera lineal a lo largo del tiempo.
- El día de la semana en donde se alcanza su valor máximo son fines de semana, principalmente en sábado y domingo.
- Finalmente muestra que el mayor número de muertes violentas se generan durante los meses diciembre, junio, septiembre, octubre y noviembre.

Adicional, respecto a los puntos de cambio el modelo aditivo de Prophet de manera predeterminada identifica estos cambios y ajusta la tendencia, la estacionalidad y otros componentes para adaptar a la serie de tiempo pronosticada. Se realizaron pruebas con fechas específicas con valores altos, pero debido a la gran cantidad de datos no se aprecian cambios significativos en los resultados.

AL realizar el análisis desagregado se evidencia que existen provincias en donde el nivel de muertes violentas generadas es mayor, las principales son Guayas con el 38.7%, Manabí con el 9.5%, Pichincha con el 8.9%, Los Ríos con el 8.4% y Esmeraldas con el 8.1%. y se detalla lo siguiente:

- Entre el 01 de enero de 2015 al 31 de diciembre de 2022 se presentaron un total de 5.357 muertes violentas en la provincia de Guayas, lo que representa el 38.7% a nivel nacional. Se evidencia que a finales del 2021 y todo el 2022 existió un incremento de muertes violentas generadas y el día en el que se comente más actos violentos es el domingo y los meses donde existe un incremento son septiembre, octubre y noviembre.
- Entre el 01 de enero de 2015 al 31 de diciembre de 2022 se presentaron un total de 1.315 muertes violentas en la provincia de Manabí, lo que representa el 9.5% a nivel nacional. Se evidencia que en todo el 2022 existió un incremento de muertes violentas generadas y los días en los

que se comente más actos violentos son sábado y domingo. Los meses donde existe un incremento son noviembre, diciembre y enero.

- Entre el 01 de enero de 2015 al 31 de diciembre de 2022 se presentaron un total de 1.231 muertes violentas en la provincia de Pichincha, lo que representa el 8.9% a nivel nacional. Se evidencia que a mediados del 2022 y todo lo que lleva el 2023, se presenta un incremento de muertes violentas generadas y los días en los que se comente más actos violentos son viernes y sábado. Los meses donde existe un incremento son noviembre y diciembre.
- Entre el 01 de enero de 2015 al 31 de diciembre de 2022 se presentaron un total de 1.163 muertes violentas en la provincia de Los Ríos, lo que representa el 8.4% a nivel nacional. Se evidencia que a mediados del 2022 y todo lo que lleva el 2023, se presenta un incremento de muertes violentas generadas y el día en el que se comente más actos violentos es el domingo. Los meses donde existe un incremento son enero y junio.
- Entre el 01 de enero de 2015 al 31 de diciembre de 2022 se presentaron un total de 1.124 muertes violentas en la provincia de Esmeraldas, lo que representa el 8.1% a nivel nacional. Se evidencia que a mediados del 2022 y todo lo que lleva el 2023, se presenta un incremento de muertes violentas generadas y el día en el que se comente más actos violentos es el domingo. Los meses donde existe un incremento son marzo, junio y septiembre.

Realizando el análisis con la información de la predicción a nivel provincial, se identifica lo siguiente:

- En la provincia de Esmeraldas en el año 2015 se registró un promedio de 0,23 muertes violentas diarias, incrementando a 1,33 muertes diarias en el año 2022 y con la proyección al 2025 incrementaría a 3,84 muertes diarias.

- En la provincia del Guayas en el año 2015 se registró un promedio de 0,86 muertes violentas diarias, incrementando a 5,59 muertes diarias en el año 2022 y con la proyección al 2025 incrementaría a 12,69 muertes diarias.
- En la provincia de Los Ríos en el año 2015 se registró un promedio de 0,22 muertes violentas diarias, incrementando a 1,11 muertes diarias en el año 2022 y con la proyección al 2025 incrementaría a 2.69 muertes diarias
- En la provincia de Pichincha en el año 2015 se registró un promedio de 0,37 muertes violentas diarias, incrementando a 0.59 muertes diarias en el año 2022 y con la proyección al 2025 incrementaría a 0.97 muertes diarias.
- Comparando los resultados a nivel Nacional y de las 5 provincias analizadas, se evidencia que la provincia con el mayor porcentaje de incremento en la predicción de muertes violentas en el año 2025 respecto al año 2022 es Esmeraldas (188%) y la provincia que presenta el menor incremento es Pichincha (64%).



## 9. CONCLUSIONES Y RECOMENDACIONES

### 9.1 CONCLUSIONES

De acuerdo con el análisis exploratorio se identifica que los años 2021 y 2022, son los que presentan el mayor número de homicidios; las provincias con el mayor porcentaje de homicidios a nivel nacional son: Guayas Manabí y Pichincha; la mayor cantidad de muertes se presentan en fin de semana y el mes con el mayor porcentaje de muertes violentas es diciembre.

Para el caso del análisis predictivo se analizaron tres modelos: aditivo de Prophet, ARIMA y SARIMA; siendo el modelo aditivo el que menor MAE (media de las diferencias absolutas entre los valores reales y los valores predichos) presentó. Al realizar un gráfico comparativo de los valores reales vs. valores pronosticados, se evidencia que existe una tendencia similar permitiendo una alineación adecuada al modelo.

Respecto al análisis provincial realizado, existen variaciones particulares en los resultados, pero en todas las provincias se presenta una tendencia creciente de muertes violentas.

Al realizar una correlación de la tasa de homicidios en Ecuador con las variables socioeconómicas, se evidencia que a nivel Nacional existe una correlación positiva con la Tasa de Empleo Inadecuado y para el caso de la provincia de Los Ríos, existe una correlación positiva con la Tasa de Empleo Inadecuado y la Tasa de Analfabetismo.

Realizando una comparación de los datos nacionales con la desegregación en provincia, se identifica que con esta tabla se pueden sacar esos datos, por ejemplo: En la provincia de Esmeraldas en el año 2015 se registró un promedio de 0,23 muertes diarias, incrementando a 1,12 muertes diarias en el año 2022 y con la proyección al 2025 incrementaría a 2,92 muertes diarias

## 9.2 RECOMENDACIONES

Para prevenir la violencia y reducir la tasa de homicidios en Ecuador, se pueden tomar varias medidas como se recomienda a continuación:

- Fortalecer las instituciones encargadas de hacer cumplir la ley y la justicia, principalmente la Policía Nacional, al igual que desarrollar programas educativos y de conciencia social que aborden la violencia y promuevan la resolución pacífica de conflictos.
- Implementar políticas públicas que aborden las causas subyacentes de la violencia, como la pobreza, la falta de oportunidades, crimen organizado, la falta de acceso a servicios básicos y drogadicción.
- Se recomienda utilizar el modelo aditivo de Prophet porque aplica logaritmos matemáticos más profundos que permiten pronosticar las series de tiempo y ayudan a tomar en cuenta los efectos de las rotaciones por ejemplo festividades como navidad, fin de año, entre otros; ahorrando tiempo en la implementación de algoritmos para realizar el pronóstico.
- Fomentar la innovación mediante la aplicación de técnicas avanzadas de aprendizaje automático y análisis de datos, como el uso de redes neuronales, algoritmos de agrupamiento o técnicas de minería de datos, para descubrir patrones ocultos y complejos en los datos de homicidios.
- La incorporación de modelos de predicción de homicidios en el análisis de políticas de gobierno ofrece una oportunidad para promover la innovación en la seguridad pública. Estos modelos proporcionan información valiosa para la toma de decisiones estratégicas, permitiendo una asignación más efectiva de recursos y la implementación de políticas preventivas. Al adoptar enfoques basados en datos y nuevas tecnologías, el gobierno puede fortalecer sus políticas de seguridad para abordar de manera más

efectiva el problema de las muertes violentas y promover un país más seguro.

- Para mejorar aún el estudio, se recomienda analizar otros modelos de predicción de series temporales que brinden la oportunidad de encontrar el mejor modelo que se ajuste al tipo de datos, detectar anomalías, diversificar estrategias de pronóstico y realizar una comparación de rendimiento. De igual forma, ampliar el periodo de estudio para obtener información más precisa de las tendencias y patrones a lo largo del tiempo. Finalmente, incluir variables adicionales relevantes para capturar mejor los factores que influyen en la tasa de homicidios, siendo estos, culturales, políticos o históricos.

## REFERENCIAS

- Gabaldón , L. (09 de Agosto de 2018). Homicide and its Situational Characterization: Between Quantitative and Qualitative. *Dilemas: Revista de Estudios de Conflicto e Controle Social*, 361-384. Obtenido de <https://www.redalyc.org/journal/5638/563866237001/html/>
- (RAE), R. (2021). *Sicariato*. Obtenido de <https://dle.rae.es/sicariato>
- Arango Londoño, D., Ortega Lenis, D., & Olaya Ochoa, J. (2009). Modelación del número de homicidios Via Regresión de Paoisson. *Heurística*, 81-90. Obtenido de <https://bibliotecadigital.univalle.edu.co/bitstream/handle/10893/6105/Heuristica16-A07.pdf>
- Arce, R., & Mahía, R. (mayo de 2001). *Modelos Arima*. Obtenido de <https://gc.scalahed.com/recursos/files/r161r/w24503w/Box-Jenkins.PDF>
- Código Orgánico Integral Penal (COIP). (2017). *Clasificación Internacional de Delitos con Fines Estadísticos*. Ficha Metodológica, Oficinas de las Naciones Unidas contra la Droga y el Delito. Obtenido de [https://www.ecuadorencifras.gob.ec/documentos/web-inec/Sistema\\_Estadistico\\_Nacional/Planificacion\\_Estadistica/Plan\\_Nacional\\_de Desarrallo\\_2017\\_2021/OBJETIVOS/Objetivo-1/1.23-FM-T\\_homicidios.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Sistema_Estadistico_Nacional/Planificacion_Estadistica/Plan_Nacional_de Desarrallo_2017_2021/OBJETIVOS/Objetivo-1/1.23-FM-T_homicidios.pdf)
- Código Orgánico Integral Penal, COIP. (2021). *Registro Oficial Suplemento 180 de 10-feb.-2014*. Obtenido de [https://www.defensa.gob.ec/wp-content/uploads/downloads/2021/03/COIP\\_act\\_feb-2021.pdf](https://www.defensa.gob.ec/wp-content/uploads/downloads/2021/03/COIP_act_feb-2021.pdf)
- Colina Vargas, A., & Espinoza Mina, M. (11 de Marzo de 2022). Aplicación de minería de datos en datos abiertos de Ecuador: Delitos. . *UCV HACER Revista de investigación y cultura*, 11(1). doi:<https://doi.org/10.18050/RevUCVHACER.v11n1a8>
- Datosmacro.com. (s. f.). (s.f.). Ecuador—Homicidios Intencionados 2020. Recuperado el 11 de 06 de 2023, de <https://datosmacro.expansion.com/demografia/homicidios/ecuador>

- De Nadai, M., Xu, Y., & Letouzé, E. (17 de 08 de 2020). *Socio-economic, built environment, and mobility conditions associated with crime: a study of multiple cities*. Obtenido de <https://doi.org/10.1038/s41598-020-70808-2>
- Differently, M. (26 de abril de 2022). *Análisis y Predicción de Series Temporales con FB Prophet Python*. Obtenido de [https://www.modeldifferently.com/2022/04/analisis\\_prediccion\\_ts\\_prophe t/](https://www.modeldifferently.com/2022/04/analisis_prediccion_ts_prophe t/)
- ENEMDU. (2007-2022). *Tasa de empleo inadecuado a nivel nacional*. Ministerio de Inclusión. Obtenido de <https://info.inclusion.gob.ec/index.php/caracterizacion-poblacion-objetivo-ancusrext/indicadores-laborales-ancusrext/tasa-de-empleo-inadecuado-ancusrext>
- ENEMDU. (Diciembre 2009-2022). Obtenido de <https://info.inclusion.gob.ec/index.php/caracterizacion-poblacion-objetivo-ancusrext/pobreza-y-desigualdad-ancusrext/tasa-de-pobreza-extrema-ancusrext>
- (2014-2021). *ENEMU*. Ecuador en cifras. Obtenido de <https://aplicaciones3.ecuadorencifras.gob.ec/BIINEC-war/index.xhtmll>
- Escobar Cabrera , A. (31 de 08 de 2015). *Biblioteca Digital Universidad del Valle* . doi:<http://hdl.handle.net/10893/8675>
- español, C. P. (1995). *Asesinato*. Madrid, España: Boletín Oficial del Estado de España.
- Forero Sanabria, A. M., Bohorquez Castañeda, M. P., Rentería Ramos, R. R., & Mateu,, J. (2022). doi:<https://doi.org/10.1007/s41109-021-00442-y>
- Gelvez Ferreira, J., Nieto-Rodríguez, M.-P., & Rocha-Ruiz, C.-A. (Septiembre de 2022). Prediciendo el crimen en ciudades intermedias: un modelo de “machine learning” en Bucaramanga, Colombia. *Revista Latinoamericana de Estudios de Seguridad*(34), 83-98. doi:<https://doi.org/10.17141/urvio.34.2022.5395>
- Hariharan, K. (23 de enero de 2020). *Time Series Forecasting—ARIMA vs Prophet*. *Analytics Vidhya*. Obtenido de <https://medium.com/analytics-vidhya/time-series-forecasting-arima-vs-prophet-5015928e402a>

- Harvey, A., & Shephard, N. (1993). Forecasting at Scale. *Handbook of statistics*, 11, 261-302.
- Hernández Brigas, H. (julio de 2022). CEPAL ORG. Obtenido de <https://www.cepal.org/es/publicaciones/47698-homicidios-america-latina-caribe-magnitud-factores-asociados>
- Hyndman, R.J., & Athanasopoulos, G. (2018). "Forecasting: Principles and Practice.". 2nd. Recuperado el 07 de 06 de 2023, de <https://otexts.com/fpp2/accuracy.html>
- INEC. (2017). *Indicadores de pobreza y desigualdad*. @ecuadorencifras. Obtenido de [https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2017/Junio/062017\\_Pobreza%20VF.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2017/Junio/062017_Pobreza%20VF.pdf)
- INEC. (2018). *Plan Nacional de Desarrollo 2017\_2021*. Quito: Instituto Nacional de Estadística y Censos. Obtenido de [https://www.ecuadorencifras.gob.ec/documentos/web-inec/Sistema\\_Estadistico\\_Nacional/Planificacion\\_Estadistica/Plan\\_Nacional\\_de\\_Developmento\\_2017\\_2021/OBJETIVOS/Objetivo-1/1.23-FM-T\\_homicidios.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Sistema_Estadistico_Nacional/Planificacion_Estadistica/Plan_Nacional_de_Developmento_2017_2021/OBJETIVOS/Objetivo-1/1.23-FM-T_homicidios.pdf)
- INEC. (2021). *Programa Nacional de Estadística 2021-2025*. Quito.
- INEC. (2022). *Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU)*. @ecuadornecifras. Obtenido de [https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2022/Agosto\\_2022/202208\\_Mercado\\_Laboral.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2022/Agosto_2022/202208_Mercado_Laboral.pdf)
- Mamadou, C., & Salama, P. (2004). Homicidios en América del Sur: ¿Los pobres son peligrosos? *Revista de Economía Institucional*, 6(10), 24. Obtenido de <http://www.scielo.org.co/pdf/rei/v6n10/v6n10a7.pdf>
- Miranda, C. (2021). *Modelización de Series Temporales modelos*. Granada. Obtenido de [https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM\\_MIRANDA\\_CHINLLI\\_CARLOS.pdf](https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_MIRANDA_CHINLLI_CARLOS.pdf)
- Ordoñez Erazo, H., Pardo Calvache, C., & Cobos Lozada, C. (30 de Diciembre de 2020). Detection of Homicide Trends in Colombia Using Machine

- Learning. *revista.ingenieria@uptc.edu.co*, 29(54).  
doi:<https://doi.org/10.19053/01211129.v29.n54.2020.11740>.
- Ovalle Correa, B. H. (2019). Análisis de la causalidad de las muertes violentas y su. *Espirales Revista Multidisciplinaria de(3)*, 2-13. Obtenido de <https://doi.org/10.31876/re.v3i26.447>
- Penal, C. O. (2014). *Femicidio*. Quito, Ecuador: Asamblea Nacional.
- Salavarría Peña, O. (02 de 2018). *SMART CITY: DIAGNOSTICO DE LA CIUDAD DE GUAYAQUIL (ECUADOR)*. Obtenido de <https://riunet.upv.es/bitstream/handle/10251/98485/Smart%20City.%20Diagnostico%20de%20la%20ciudad%20de%20Guayaquil%20%28Ecuador%29.pdf>
- Todd, E. (2022). *Après l'empire, essai sur la décomposition du système américain*,. París.
- UNODC. (11 de 06 de 2013). *Oficina de las Naciones Unidas contra la Droga y el Delito*. Obtenido de Hacia una definición estandarizada del homicidio intencional para propósitos: [https://www.unodc.org/documents/data-and-analysis/statistics/Homicide/homicide2013/Intentiona\\_homicide\\_Spanish.pdf](https://www.unodc.org/documents/data-and-analysis/statistics/Homicide/homicide2013/Intentiona_homicide_Spanish.pdf)
- UNODC. (11 de 06 de 2013). *Oficina de las Naciones Unidas contra la Droga y el Delito*. Obtenido de Hacia una definición estandarizada del homicidio intencional para propósitos: [https://www.unodc.org/documents/data-and-analysis/statistics/Homicide/homicide2013/Intentiona\\_homicide\\_Spanish.pdf](https://www.unodc.org/documents/data-and-analysis/statistics/Homicide/homicide2013/Intentiona_homicide_Spanish.pdf)

# ANEXOS



## Anexo 1

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
1	Héctor Hiram Hernández Bringas (2022)	Homicidios en América Latina y el Caribe: magnitud y factores asociados	Las variables independientes a considerar son: 1. Porcentaje de población en situación de pobreza que sobrevive con menos de 5,50 USD al día en 2015. 2. Índice de Gini que mide la desigualdad en cada país. 3. Porcentaje anual de crecimiento de la población urbana (promedio 2000-2018). 4. Porcentaje de adolescentes que no asisten a la escuela en el año 2015. 5. Porcentaje de desempleo entre las personas de 15 a 24 años en 2015 6. Índice de estado de derecho en 2017. 7. Consumo promedio anual de litros de alcohol en varones mayores de 15 años en 2015, y	La pobreza y desigualdad explican las altas tasas de homicidio, en la medida en que reducen las oportunidades laborales, académicas y de consumo, además de la frustración social asociada con estos factores.	Países de América Latina y el Caribe con los 202 países sobre los que ofrecen información las Naciones Unidas (2021)). y de la Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC, 2013)	Se realizaron análisis de correlación (coeficiente de correlación de Pearson) de las tasas de homicidios (tasas promedio entre los años 2015 y 2018) con respecto a distintas variables que denotan los factores aparentemente explicativos de la violencia letal.	Se incluyó una caracterización de la situación particular que existe, en materia de violencia y homicidios, en los países de la región mayormente afectados.	Determinar la tasa de homicidio en los países de América Latina y el Caribe

Tabla 19: Homicidios de América Latina

Fuente: H. Hernández Bringas - 2022/Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
2	De Nadai, M., Xu, Y., Letouzé, E., González, M. C., & Lepri, B. (2020)	Socio-economic, built environment, and mobility conditions associated with crime: a study of multiple cities	El estudio analiza cómo diferentes factores, incluyendo características socioeconómicas, ambientales, construidas y de movilidad, se correlacionan con la actividad criminal en cuatro ciudades diferentes: Bogotá, Boston, Chicago y Los Ángeles.	La comprensión limitada de los factores que influyen en la actividad criminal en diferentes ciudades y culturas.	* American Community Survey (ACS) para los bloques censales, población, empleo y pobreza de las ciudades estadounidenses. * Departamento Administrativo Nacional de Estadística (DANE) para los datos del censo general de Bogotá en 2005. * Sistema de Identificación Símbol III de 2014 para los datos de pobreza en Bogotá.	Utiliza un modelo bayesiano para explorar cómo los delitos violentos y de propiedad están relacionados con diferentes factores en las ciudades estudiadas. También utiliza coeficientes $\beta$ estandarizados para revelar cómo las características se correlacionan con la actividad criminal en diferentes ciudades.	Indican que los coeficientes $\beta$ varían significativamente entre las diferentes ciudades. Además, una población más densa se correlaciona positivamente con la actividad criminal en Bogotá y Chicago, pero negativamente en Boston y Los Ángeles.	Una mejor comprensión de cómo diferentes factores, incluyendo características socioeconómicas, ambientales, construidas y de movilidad, se correlacionan con la actividad criminal en diferentes ciudades.

**Tabla 20: Socio-economic, built environment and mobility conditions associated with crime**

Fuente: Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
3	Forero Sanabria, Alan Miguel; Bohorquez Castañeda, Martha Patricia; Rentería Ramos, Rafael Ricardo; Mateu, Jorge. (2022)	Identification of patterns for space-time event networks	Identificar y analizar patrones espaciotemporales para comprender mejor los fenómenos subyacentes en diversas áreas de investigación.	El análisis de patrones puntuales espaciotemporales en diversas áreas de investigación, como la delincuencia, los terremotos, la ecología y la epidemiología.	Datos proporcionados por la Policía Metropolitana de Medellín (Colombia).	Siguen la construcción de redes de Davies y Marchione y construyen una serie temporal de redes de eventos. Esta metodología se puede aplicar a cualquier tipo de eventos interactivos que evolucionan en el espacio y el tiempo.	La aplicación de redes de eventos para la detección de patrones espaciotemporales demuestra la capacidad de capturar diferentes tipos de configuraciones que generan estos patrones.	Utilizar herramientas como la identificación de patrones espaciotemporales por bandas temporales para analizar los datos

**Tabla 21: Identificación de patrones para redes de eventos espacio temporales**

Fuente: Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
4	Gelvez Ferreira, Juan David; Nieto-Rodríguez, María-Paula; Rocha-Ruiz, Carlos-Andrés (2022)	Prediciendo el crimen en ciudades intermedias: un modelo de "machine learning" en Bucaramanga, Colombia	Predecir el delito en Bucaramanga y determinar si se podrían desarrollar modelos similares en otras ciudades del país para mejorar la gestión de los recursos de las autoridades policiales.	El crimen en ciudades intermedias y si la política pública se concentrase en herramientas de predicción del delito.	Sistema de Información Estadístico, Delincuencial, Contravencional y Operativo de la Policía Nacional (SIEDCO).  Se utilizó las siguientes variables: modalidad del delito, departamento, municipio, fecha (mes, día, año), edad de la víctima, género de la víctima, zona urbana o rural, hora del día en el que ocurrió el hecho, código de identificación, país de la víctima, latitud y longitud del hecho y número de casos ocurridos en el hecho.	El modelo elegido en el desarrollo del estudio es el modelo KNN, que presentó mejores resultados de clasificación y tiempos de ejecución.	Se encontraron buenos resultados al utilizar la frecuencia semanal y que los mejores resultados de la predicción del crimen en Bucaramanga se dieron al utilizar los modelos con relaciones espaciales de grafos por semanas.	Depende del usuario y su disposición para priorizar una métrica frente a otra, pues se puede lograr un recall mucho más alto, pero a costo de una precisión relativamente más baja.

**Tabla 22: Prediciendo el crimen en ciudades intermedias**

Fuente: Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
5	Colina Vargas, Alejandra ; Espinoza Mina, Marcos; (2022)	Aplicación de minería de datos en datos abiertos de Ecuador : Delitos	Identificar patrones significativos relacionados con incidentes delictivos y agrupar distintos tipos de delitos dentro de las categorías existentes para ampliar la comprensión de los problemas delictivos y respaldar el diseño de intervenciones preventivas.	Falta de disponibilidad de datos abiertos de interés social y la poca exigencia para su publicación en Ecuador	Indicadores de seguridad ciudadana, las series históricas de robos y homicidios intencionales tomados del Ministerio de Gobierno de Ecuador.	Consistió en la aplicación de técnicas de minería de datos para extraer y analizar los datos obtenidos del Ministerio de Gobierno de Ecuador. Se identificó el modelo de "regresión algorítmica" como el algoritmo cuyos valores superan el 90% de probabilidad de detección de patrones delictivos, en este caso, delitos de homicidios y robos principalmente.	Los resultados indican que el modelo de "regresión algorítmica" es el que tiene una mayor capacidad predictiva para detectar patrones delictivos en estos seis cantones: Durán, Esmeraldas, Guayaquil, Machala, Quevedo y Quito.	Se demuestra que la minería de datos es una herramienta útil para analizar grandes cantidades de datos y extraer información significativa para la toma de decisiones y la definición de políticas y estrategias para disminuir los niveles de delincuencia en el país.

**Tabla 23: Aplicación de minería de datos en datos abiertos de Ecuador**

Fuente: Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
6	Luis Gerardo Gabaldón (2018)	Homicide and its Situational Characterization: Between Quantitative and Qualitative "	Revisar perspectivas sobre oportunidades y dinámicas de crimen, violencia interpersonal y explicaciones macroestructural es aplicadas a homicidios como un evento situacional, a fin de proponer un abordaje integrada para su explicación e intervención.	La necesidad de una mejor comprensión de la dinámica y caracterización situacional del homicidio. Al igual que, se requiere proporcionar una base sólida para el diseño de políticas públicas efectivas para prevenir el homicidio.	El estudio utiliza datos cuantitativos emergentes de investigación es de víctimas y datos cualitativos emergentes de agentes de homicidio.	Se basa en fortalezas de investigación consistentes en el tiempo en el ámbito latinoamericano, integrando diversas perspectivas cuantitativas, derivadas de encuestas de victimización, y perspectivas cualitativas, derivadas de entrevistas a victimarios, para comprender mejor la dinámica y caracterización situacional del homicidio.	Los resultados de la regresión logística para casos de homicidios (n=423), contrastados con los casos de lesiones y robos agrupados (n=3.017), donde se analizan variables como edad y género, relaciones laborales, distribución espacio temporal de la victimización y dinámica situacional del encuentro victimal.	Permite una mejor comprensión de la dinámica y caracterización situacional del homicidio, lo que puede ayudar a los responsables de políticas públicas a diseñar intervenciones más efectivas para prevenir este tipo de violencia.

**Tabla 24: El homicidio y su caracterización situacional**

Fuente: Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
7	Hugo-Armando Ordoñez-Eraso; César-Jesús Pardo-Calvache; Carlos Alberto Cobos-Lozada (2020)	Detection of Homicide Trends in Colombia Using Machine Learning	Detectar tendencias en los homicidios en Colombia y proporcionar información útil para la toma de decisiones por parte de los organismos gubernamentales.	Alta y creciente tasa de homicidios en Colombia en las últimas décadas, lo que ha llamado la atención de los gobiernos y organismos gubernamentales encargados de controlar el crimen violento	Se utilizaron datos de homicidios en Colombia desde el año 2000 hasta el año 2018, obtenidos del Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia y contienen información sobre la ubicación geográfica, la fecha y la causa de muerte de cada homicidio registrado en Colombia durante ese período.	Se divide en tres etapas (preprocesamiento de datos, selección de características y modelado) <b>1. Preprocesamiento de datos:</b> limpieza y transformación de los datos para eliminar valores atípicos y asegurar la calidad de los datos. <b>2. Selección de características:</b> Se utiliza técnicas estadísticas para identificar las variables más relevantes para el modelo. <b>3. Etapa de modelado:</b> técnicas de aprendizaje automático para construir modelos predictivos que permitieran detectar tendencias en los homicidios en Colombia	Los datos muestran que la tasa de homicidios en las zonas rurales y urbanas tiene un enorme impacto que es catalogado como una crisis de salud pública.	Evaluar la eficacia del modelo utilizando métricas como el error cuadrático medio y el coeficiente de determinación.

**Tabla 25: Detection of Homicide Trends in Colombia Using Machine Learning**

Fuente: Hugo-Armando Ordoñez-Eraso; César-Jesús Pardo-Calvache; Carlos Alberto Cobos-Lozada/ Elaboración propia

	Autor	Tema	Objetivos de la investigación	Problema abordado	Fuente de datos utilizados	Metodología implementada	Resultados	Implicaciones
8	David Arango Londoño Javier Olaya Ochoa Delia Ortega Lenis (2009)	Modelación Del Número De Homicidios Vía Regresión De Poisson	Modelar los homicidios en la ciudad de Cali utilizando la regresión de Poisson y analizar cómo se relacionan con algunas características particulares como pueden ser el género o la edad.	No se reconocen los homicidios como uno de los principales problemas socioeconómicos que afectan la calidad de vida en todas sus dimensiones, en los niveles psicológico, emocional y de convivencia, y en la mayor parte de los sectores sin importar los estratos socio demográficos y económicos.	Se utilizaron algunas bases de datos construidas por el instituto de investigación CISALVA en Cali en los años 2007 y 2008.	Se utilizó un modelo de regresión de Poisson para estudiar el número de homicidios a partir de las covariables: grupo etario, género y tipo de arma. Este modelo es adecuado para datos de conteo, como el número de homicidios, y permite modelar la relación entre las covariables y la variable respuesta.	Indican que el riesgo de ser víctima de homicidio es mucho más alto en los hombres entre los 20 y los 30 años que en las mujeres y que más probablemente el homicidio será cometido con un arma de fuego	Puede tener un impacto negativo en la economía y el desarrollo social de la ciudad, ya que puede disuadir a las empresas y a los turistas de invertir o visitar la ciudad.

**Tabla 26: Modelación Del Número De Homicidios Vía**

Fuente: Elaboración propia