Neutrosophic Computing and Machine Learning, Vol. 37, 2025
University of New Mexico

Patterns and Determining Factors of Homicides in Ecuador: An Open Data-Based Analysis Approach and Predictive Tools

Patrones y factores determinantes de los homicidios en Ecuador: Un Enfoque de Análisis Basado en Datos Abiertos y Herramientas Predictivas

Wladimir Alexander Palacios Zurita [1,2,*] Edwin José Moran Estrada [2], Alexander Javier Paredes Romero [3,] Ely Stefanie Parra Tauriz [4,] Patricia Stefania Morillo Montenegro [5], and Manuel Fabricio Reyes Wagnio [6]

[1] Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López ESPAM MFL, Manabí. Ecuador. apalacios@espam.edu.ec
[2] University Bolivarian of the Ecuador, Duran, Guayas, Ecuador. ejmorane@ube.edu.ec
[3] Universidad de Guayaquil, Guayas, Ecuador. ajparedesr@ube.edu.ec
[4] University Bolivarian of the Ecuador, Duran, Guayas, Ecuador. esparrat@ube.edu.ec
[5] University Bolivarian of the Ecuador, Duran, Guayas, Ecuador. psmorillom@ube.edu.ec
[6] Universidad de Guayaquil, Guayas, Ecuador. manuel.reyesw@ug.edu.ec

Summary. In recent years, Ecuador has experienced an alarming increase in deaths violent, especially in intentional homicides, which presented serious challenges for public safety. This study analyzed trends and patterns of intentional homicides in Ecuador between January and November 2024. The objective was to explore the nature of these homicides and demonstrate how data analysis tools, such as Orange, can be used for tackle issues of public safety and support the take of decisions informed. He study, of guy explanatory and graphic, employment a design of networks neuronal and analyzed a data panel covering eleven months of the year 2024, using a historical approach. Intentional homicides in the country showed an alarming upward trend, especially in urban areas and in the province of Guayas. The majority of the homicides were tasks with arms of fire, and common crime was the main motivation behind these crimes. The victims were mainly single men of mixed ethnicity, which indicated a pattern social and demographic that urgent intervention was required. Addressing the underlying causes of crime and strengthening public security policies were crucial to reducing the homicide rate. This analysis underscores the need to implement data-driven solutions and more effective security policies to curb violence and improve security in the country.

Keyword: Ecuador; homicides intentional; security public; orange; analysis of data

Resumen. En los últimos años, Ecuador ha experimentado un alarmante incremento de muertes violentas, especialmente de homicidios dolosos, lo que presenta serios desafíos para la seguridad ciudadana. Este estudio analizó las tendencias y patrones de los homicidios dolosos en Ecuador entre enero y noviembre de 2024. El objetivo fue explorar la naturaleza de estos homicidios y demostrar cómo herramientas de análisis de datos, como Orange, pueden ser utilizadas para abordar temas de seguridad ciudadana y apoyar la toma de decisiones informadas. El estudio, de tipo explicativo y gráfico, empleo un diseño de redes neuronales y analizó un panel de

datos cubriendo once meses del año 2024, utilizando un enfoque histórico. Los homicidios dolosos en el país mostraron una alarmante tendencia al alza, especialmente en las zonas urbanas y en la provincia del Guayas. La mayoría de los homicidios fueron tareas con armas de fuego, y la delincuencia común fue la principal motivación detrás de estos crímenes. Las víctimas eran principalmente hombres solteros de etnia mixta, lo que indicaba un patrón social y demográfico que requería una intervención urgente. Abordar las causas subyacentes de la delincuencia y reforzar las políticas de seguridad pública eran cruciales para reducir la tasa de homicidios. Este análisis subraya la necesidad de aplicar soluciones basadas en datos y políticas de seguridad más eficaces para frenar la violencia y mejorar la seguridad en el país.

Palabras Claves: Ecuador; homicidios intencionales; seguridad pública; naranja ; análisis de datos

1. Introduction

The violence, especially in shape of intentional homicides, is one of the most complex and devastating problems for modern societies. Ecuador has experienced an alarming increase in levels of violence, particularly in relation to violent deaths, which that ha generated a growing concern among researchers, government authorities and the society in his set [1]. This The increase has highlighted the deep social problems and structural that underlying to the crime, highlighting the need urgent of tackle the causes fundamentals of this phenomenon (Zapata Mayorga & Kumar Vivas, 2024). This phenomenon affects not only the criminal justice system, but also fundamental aspects of public health and social well-being. Therefore, it is crucial to have accurate and up-to-date data that allows us to understand the underlying causes and dynamics of violence, in order to make informed decisions that contribute to the reduction of this problem [2]. The Government of Ecuador has promoted transparency in the area of public security through the Government of Ecuador's Open Data portal, where it makes various data sets available to citizens, including information on intentional homicides [3]. This resource offers key details, such as the date, geographic location, type of crime, and demographic characteristics of the victims. the victims, which allows a comprehensive analysis of the trends of homicide in the country. Public access to these data facilitates academic research and contributes to the improvement of policies public based in evidence (Ministry (Interior, 2025). This study uses the dataset "Intentional Homicides", available on the Open Data platform, with he purpose of analyze the tendencies and patterns of the homicides in Ecuador [4]. For it, The Orange program, a data analysis tool, will be used visual based in the mining of data and interactive visualization. Orange allows statistical analysis and predictive models to be built intuitively, without requiring advanced programming knowledge, which facilitates the exploration of large volumes of information. Through of this tool, HE will carry to cape a analysis descriptive of homicides in Ecuador, identifying correlations between geographic, temporal and demographic variables [5].

This work aims not only to explore the nature of intentional homicides in Ecuador, but also to demonstrate how accessible data analysis tools, such as Orange, can be used to address complex public safety problems and contribute to informed decision-making that improves the lives of victims. security in the country.

2. Methodology Computational

This data analysis was conducted within the framework of a datathon, using a set of 6,276 records of intentional homicides that occurred in Ecuador between January and November of 2024. The data, coming from from the public access database "Open Data of Ecuador" of the Ministry of Interior (2025), offer official, reliable and quality statistical information which aim to contribute to the design of public policies in citizen security, facilitating decision-making and optimizing territorial management [6]. The dataset includes variables such as: type of death, location of the incident (zone, sub-zone, district, circuit, province, canton, coordinates, place and type of place), date and time of the incident, type of weapon used, alleged motivation, and qualitative data on the victim (age, sex, gender, ethnicity, marital status, nationality, disability, profession, civil registration and education). For the analysis, the Orange Data Mining tool was selected for its ability to integrate visualization and predictive analysis in an efficient and accessible environment. The research adopted a quantitative approach to explore and analyze the information accurately.

2.1. Implementation in Orange

Data analysis began with the import of the database in CSV format into the Orange tool, using the "File" widget. From this central point, five distinct analysis flows were developed, each with a specific purpose. and a

configuration particular, besides HE added a flow of additional text processing, so the final pipeline developed in Orange had six analysis flows (See Annex 1). Below, the methodology and configuration of each flow is described, detailing the use of the widgets involved in the process [8]:

Flow 1: The workflow for loading analysis and initial data exploration was structured as follows:

1. **Data Info:** This widget was used to get an overview of the data, providing crucial information about the characteristics of the dataset before applying more advanced techniques. Specifically, the selected database has: 6275 rows and 34 columns, 21 variables They are categorical, 6 variables are numerical and 7 variables They are text. (Fig. 1) [8].

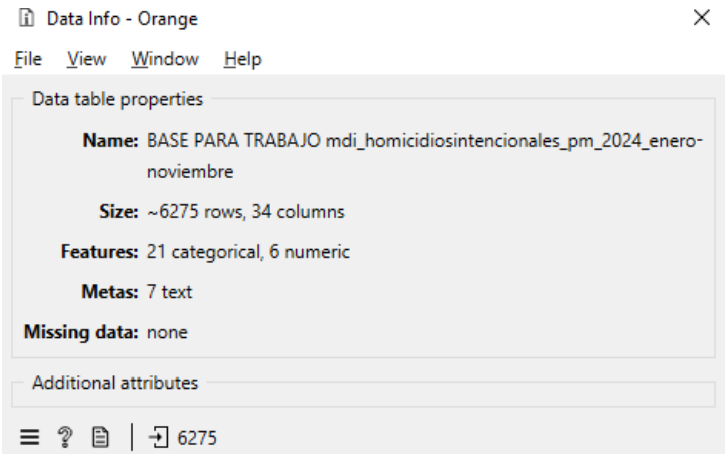

Figure 1. Widget Data Info - Flow 1

2. **Data Table:** This widget visualized the data in tabular format, making it easy to inspect and explore. In this study, HE confirmed that the base of data consisted of 6275 instances with complete data, without variables missing. HE noticed the presence of 27 characteristics and 7 meta attributes, although No HE identified a variable explicit-objective. "Data Table " allowed the verification and exploration of the values, both before as after of apply transformations or analysis (Fig. 2).

Figure 2. Widget Data Table - Flow 1

3. **Feature Statistics [11]:** This widget is used to inspect and find potentially interesting features in the given dataset (Orange Data Mining, n.d.). For the purposes of this study, it was used to facilitate exploration, ID of patterns and the detection of outliers, where a specific color scheme was applied to the variable "guy of death": Murder (light blue), Femicide (red), Homicide (green) and Sicariato (orange). This scheme allowed visualize and clearly relate the statistics of the variables in the context of the analysis of Intentional Homicides. Additionally, " Feature Statistics " provided metrics such as: Distribution, Mode, Mean, Dispersion, Minimum

and Maximum (Fig. 3).

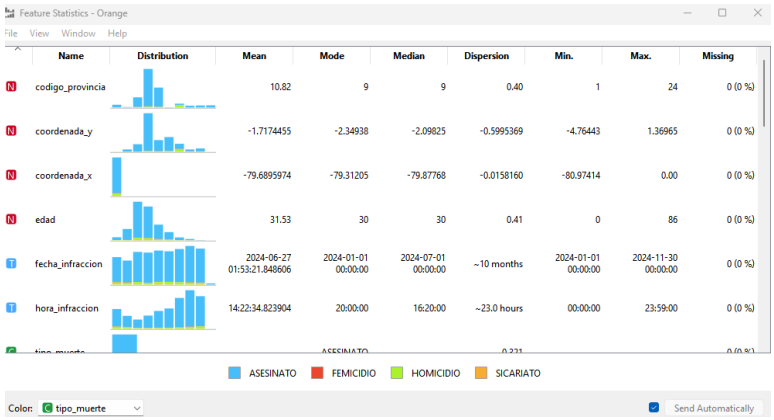

Figure 3. Widget Feature Statistics - Flow 1

4. **Box Plot :** This widget was used to visualize the distribution of numerical variables using box plots showing the interquartile range (IQR), with a line inside representing the median (Q2). The “whiskers” extend from the box to the minimum and maximum values (Neto et al., 2017). Thus, “Box Plot ” was instrumental in identifying variability, detecting outliers, and obtaining an overview of how numerical features are distributed in the Intentional Homicides dataset. In this case, the variable “type of death” was used as the main filter to analyze the distribution of other variables, including "Province", "Alleged motivation", "Weapon", "Sex" and "State civil".

Flow 2: He flow of job for he analysis of geocoding HE development of the following manner:

1. **Geocoding:** This widget was used for transform the coordinates geographic (latitude and (length) into information about the administrative region to which they belong. To do this, the coordinates present in the columns "Y coordinate" and "X coordinate" were decoded, obtaining as a result the name of the first-level administrative subdivision. The main objective was to enrich the data with relevant geographic information, thus facilitating its understanding for analysis and visualization. (Fig. 4).

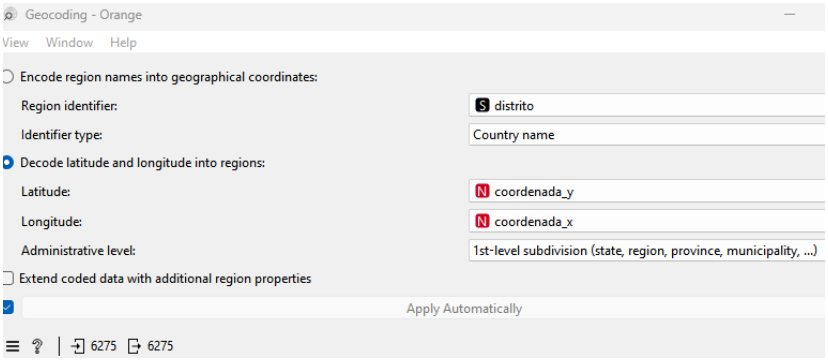

Figure 4. Widget Geocoding - Flow 2

2. **Geo Map:** This widget was used to visualize the data geographic in a map. HE They took as input the geographic coordinates ("Y coordinate" and "X coordinate") and were colored the points according to the values in the "presumed motivation" column. The fundamental purpose was to visually represent the spatial distribution of the data and to identify possible geographic patterns related to the different categories of motivation (Fig. 5).

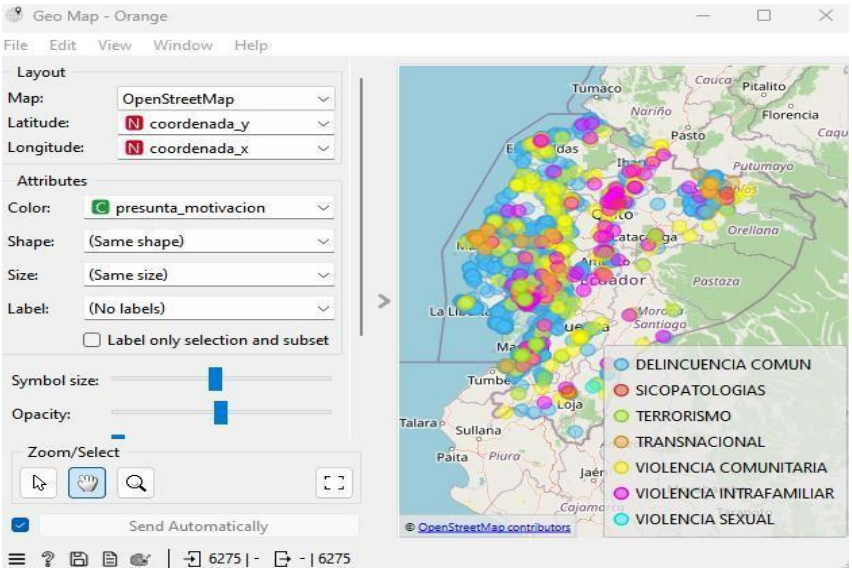

Figure 5. Widget Geo Map - Flow 2

3. **Data Table:** This widget was used to visualize the results of the process of geocoding in tabular format. Provided a detailed view of the new columns generated with administrative information (first-level subdivision), along with the original columns of coordinates and the presumed motivation. It allowed for accurate inspection of the geocoded data, ensuring its integrity before proceeding with the analysis.

Flow 3: The workflow for developing the classification model (Naive Bayes) was carried out of as follows:

1. **Select Columns:** This widget was used to filter and select key variables from the dataset, focusing the analysis on the most relevant characteristics. In this case, variables such as geographic location, type of crime, age and gender of victims, among others, were selected. These characteristics were considered determinants to understand intentional homicide trends. in Ecuador between January and November 2024.

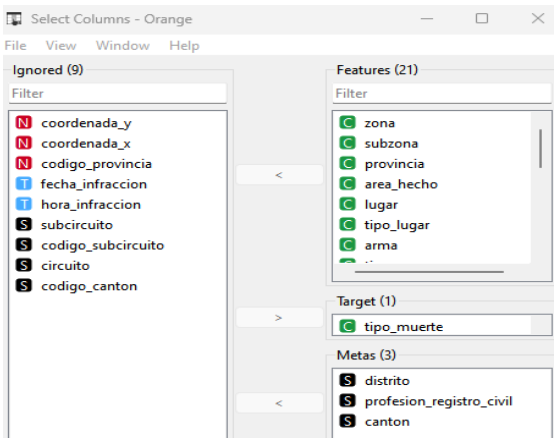

Figure 6. Widget Select Columns - Flow 3

2. **Naive Bayes:** After the selection of the key variables, the model was trained. classification. HE employment he sorter Naive Bayes, a algorithm probabilistic that assumes the dependence between the characteristics of the data set (Hofmeyr et al., 2024). This model was used to predict the probability that a homicide intentional occur low certain conditions, providing a perspective to analyze the factors that influence the homicide rate in Ecuador (Fig 7).

Figure 7. Widget Naive Bayes - Flow 3

3. **Confusion Matrix:** This widget was used to visualize the confusion matrix, which shows the correct and incorrect predictions of the model. It facilitated the analysis of the model's performance, identifying the types of errors (false positives and false negatives).

4. **Feature Importance:** To better understand the impact of each variable on homicide prediction, the " Feature " widget was used. Importance ". This widget calculated the influence of each feature in the model, allowing to determine which factors are most determining in the occurrence of intentional homicides.

5. **Test and Score:** Once the model was trained, it was necessary to evaluate its performance. To do this, the "Test and Score" widget was used, which performed cross-validation and calculated metrics such as precision , recall , and accuracy of the model. The results indicated that the Naive model Bayes showed reasonable accuracy, suggesting its usefulness for predicting the probability of homicide under certain conditions.

6. **Data Table:** This widget HE used for introduce a board interactive with the instances of the data and their corresponding model predictions. It facilitated the verification and analysis detailed of the results of the model, showing the predictions along with the values royal.

Flow 4: He flow of job for he analysis exploratory of data and views included:

1. **Select Columns:** The following was used " Select " widget Columns " to select the variables of interest for exploratory analysis. Unlike other flows, this widget was not used to filter a subset of data for a specific model, but rather to choose the variables that would be relevant for exploration. visual.

2. **Box Plot:** The "Box Plot " widget or scatter diagram visually represents the relationship between two variables, offering the reader key information immediately. These graphics allow observe the shape of the relationship, the size of the sample analyzed and the impact that outliers may have on the analysis (Sainani , 2016). In this study, HE used for explore the distribution of the selected variables (Type of death by type of weapon – Guy of death by age range), using a color scheme that differentiates each variable used.

3. **Distribution:** This widget was used to visualize the distribution of each variable individually. It presented histograms and probability density plots, providing an understanding of the distribution. detailed how the data is distributed for each feature.

4. **Scatter Plot:** It was used to visualize the relationship between pairs of variables numerical. Scatter diagrams were generated that allowed us to explore whether there were correlations or patterns between the selected variables.

Flow 5: The workflow implemented for the clustering analysis and classification included the following stages:

1. **Data loading and preprocessing:** The dataset was loaded using the "File" widget on the platform Orange. Due to the Limitations of the algorithm K- means for process further of 5,000 records, random sampling was applied with the "Data Sampler " widget. This procedure allowed the size of the data set to be reduced to a representative sample of 5,000 cases, ensuring that the diversity of observations was maintained (Fig. 8).

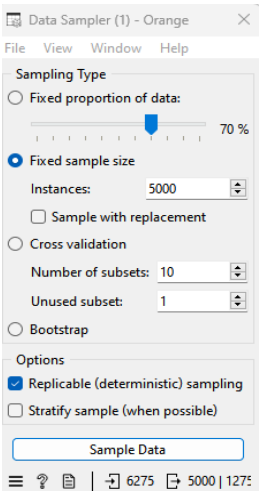

Figure 8. Widget Data Sampler – Flow 5

2. **Clustering with K- means:** A clustering analysis was performed using the K -means algorithm, setting the model for explore a range of cluster values between 2 and 18. The evaluation of the quality of the clusters was based on the Silhouette Score, which measures the cohesion and separation of the clusters. According to Artus (2024), a high score in this metrics indicates a well-defined structure in the data. In this case, the model with two clusters got the highest score with a silhouette coefficient of 0.624, it that suggests a clear separation and cohesion of the data. Therefore, the two-cluster model was selected for further analysis. In addition, a column normalization process was applied to ensure that the variables contributed equally to the analysis, avoiding the predominance of those with a larger scale (Fig. 9)

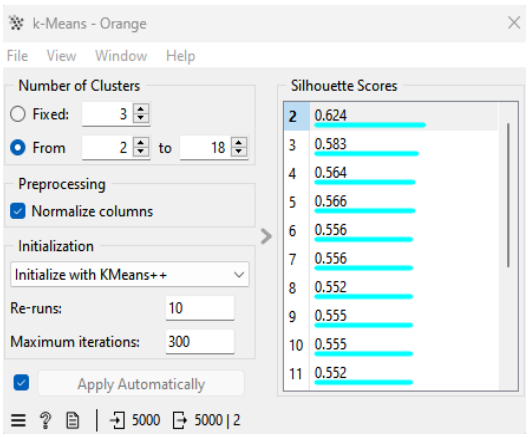

Figure 9. Configuration of the Widget K- Means – Flow 5

3. **Variable selection: To identify the most relevant features in the characterization of each cluster, the** " Select " widget was used. Columns ". This process allowed discrimination between the predictor variables and the target variable. The selected variables included province, area of the event, gender, type of death, presumed observed motivation, weapon, location and ethnicity, assigning the generated cluster as the target.

4. **Classification with Naive Bayes:** To understand the relationship between the selected variables and the clusters, the widget was used Naive Bayes (NB). This probabilistic model, based on Bayes ' theorem, assigns observations to the most probable class considering the independent variables (Romano et al., 2024, p. 237). The NB allowed to analyze the importance of each variable in the classification, showing how each attribute influences the assignment to a cluster specific.

5. **Display of results:** Finally, to visualize the relative relevance of each variable in the classification model, is used he widget Nomogram. This widget is configured for show the variables in order of absolute importance for both clusters.

Flow 6: As mentioned above, A flow based on text analysis was used, the purpose of which was to create a word cloud based on the theme "Security in Ecuador". Thus, the flow included the following widgets:

1. **Wikipedia:** This allowed obtaining information directly from Wikipedia on a specific topic. That is, it takes a concept, searches for the corresponding article on the platform and returns a summary or fragment of the text. In the case of this study, the term "security in Ecuador" was queried, which must be searched in Spanish. The retrieved text included the title, content and relevant summaries of the article.

2. **ChatGPT Constructor: Allowed the integration of the** ChatGPT language model to perform text processing and content generation. In the current work, a search parameter focused on "Keywords on security in Ecuador" was generated with a "gpt-3.5-turbo" model, obtaining a unique code to identify and authenticate the user (API Key).

3. **Preprocess Text:** This widget allowed you to clean and prepare the text before analyzing it. This The process was essential since raw texts often contain errors and unnecessary elements. which can negatively affect the results of the analysis. Having well-processed data significantly improves the quality of the results. In this case study, the text was processed by transforming it into lowercase, tokenization was performed using expressions regular (\w+) and stop words and numbers were excluded (Fig. 10).

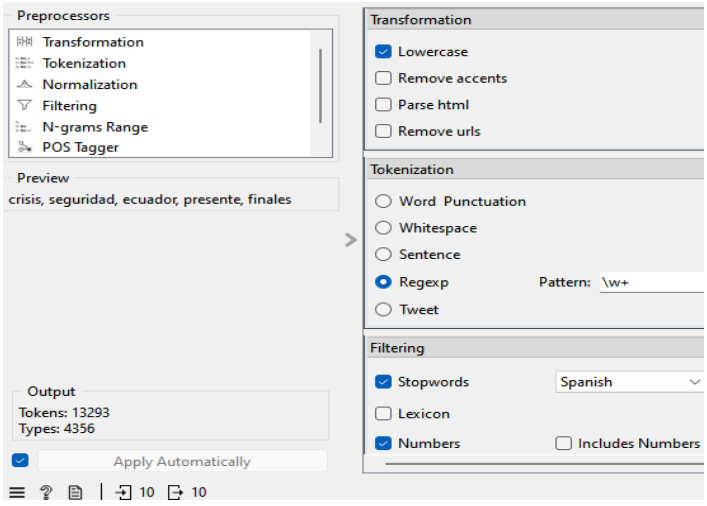

Figure 10. Configuration of the Widget PreprocessText – Flow 6

4. **Word Cloud:** Allowed the generation of a visual word cloud from of a set of text, in In this case, on keywords related to security in Ecuador previously requested. The most frequent and important words are shown in a larger size, making it easier to identification of the most frequently occurring terms.

3. Results and Discussion

Flow 1: Exploration of data through Box Plot

The distribution of the data can be observed in the box diagram according to the t ipo of death occurred By Province of Ecuador between January and November 2024:

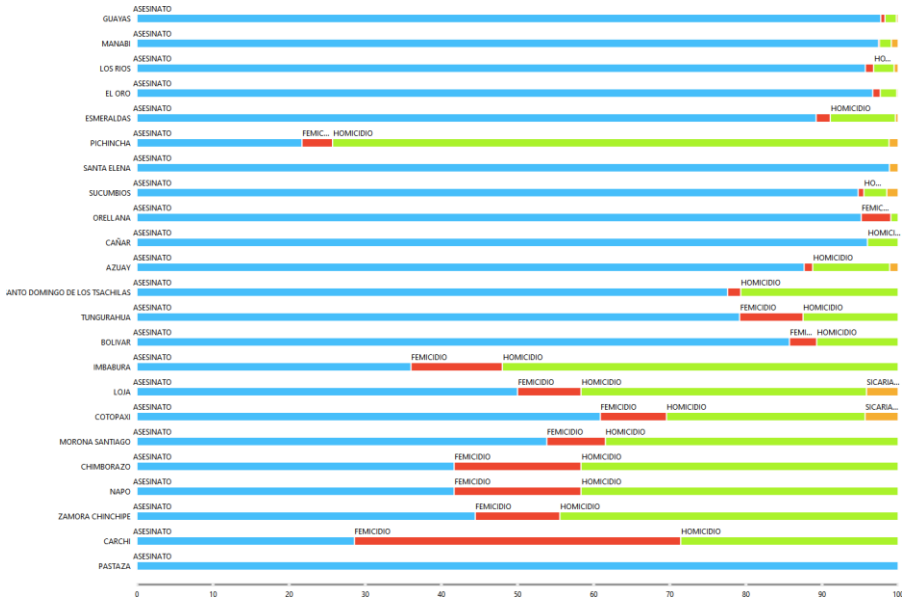

Figure 11. Guy of death by Province of the Ecuador

The graph shows the distribution of murders, femicides, homicides and contract killings in various provinces of Ecuador, showing that murder predominates in most of them, although with significant variations. Provinces such as Carchi, Imbabura, Cotopaxi, Morona Santiago, Chimborazo, Napo and Zamora Chinchipe show a combination of these crimes, including notable proportions of femicide and homicide, and in some cases, contract killing, while Guayas, Manabí, Los Ríos, El Oro, Esmeraldas, Santa Elena and Pastaza present a high incidence of murder as the main crime. The diversity in the proportions of each type of crime between provinces suggests different and complex dynamics of violence, highlighting the need of more detailed studies to understand these particularities and to be able to implement prevention and control strategies of the effective crime in each region (Fig. 11).

So same HE obtained the distribution of data in the box plot according to the t Type of death occurring due to presumed motivation between January and November 2024:

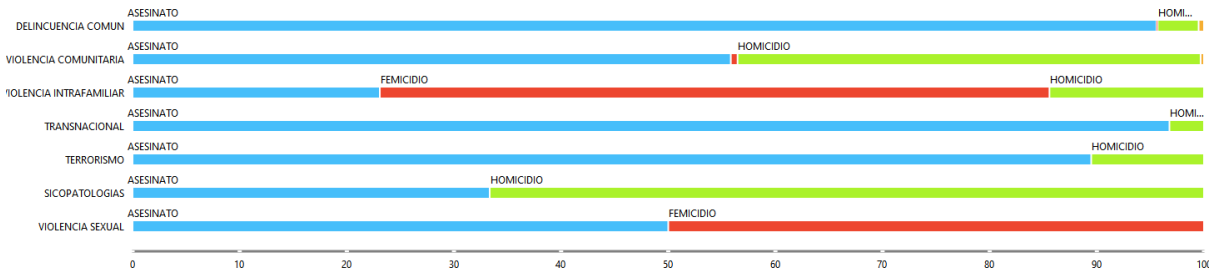

Figure 12. Guy of death by Alleged Motivation

The graph shows the distribution of murders, femicides and homicides according to different causes or contexts, where murder predominates in the most of the categories, although with notables variations in proportions. Common crime, transnational causes and terrorism show a high incidence of murders with a minimal representation of other crimes, while community violence, domestic violence and psychopathologies reflect a

combination of murders and homicides. By his part, the violence sexual stands out by a proportion significant of femicides, even surpassing murders in this category. These differences in proportions suggest that each cause or context has particular dynamics of violence that require deeper analysis and implementation. of specific prevention and control strategies crime for each case (Fig. 12).

By on the other hand, it obtained the distribution of the data in he diagram of boxes according to het ipo of Death occurring due to the sex of the victim, which includes the categories male, female, and undetermined, between January and November 2024:

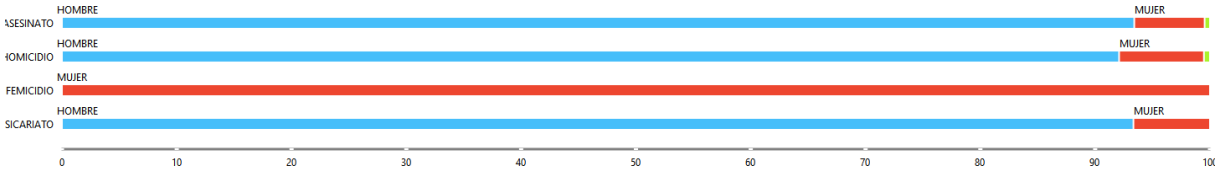

Figure 13. Guy of Death by Sex of the victim

The graph shows the distribution of murders, homicides, femicides and contract killings according to the gender of the victim (Fig. 13), where it can be seen that murders, homicides and contract killings mostly affect men, although with a small proportion of women and undetermined cases, while femicide is almost entirely presented as a crime committed against women. The presence, although minimal, of undetermined cases in the crimes analyzed, suggests the need for a record further necessary to obtain a complete overview of the situation. This marked differentiation in the proportion of victims by gender highlights the urgency of implementing prevention and care strategies specific to each type of crime and each gender.

Flow 2: Geographical Analysis

The distribution map showed the following geographic results for the variable “Presumed motivation”:

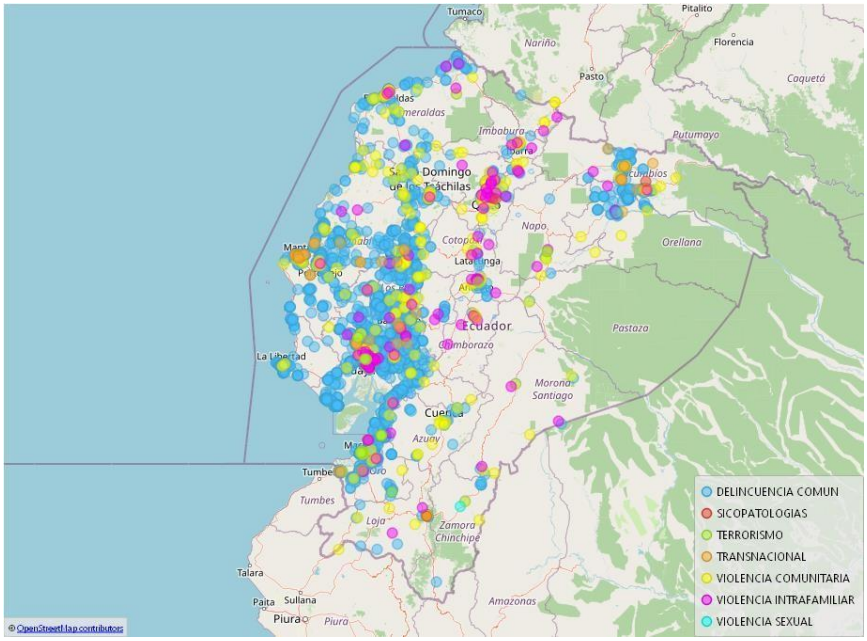

Figure 14. Map of distribution by alleged motivation

"Common crime" (represented in blue) is the most prevalent category on the map, showing a wide distribution across various regions. Specific geographic patterns can be identified for certain categories of violence. Domestic violence (represented in magenta) tends to be more dispersed and present in several areas of the country, while other types of violence such as "psychopathologies" (red) or "terrorism" (green) are more concentrated in certain regions. Community violence (yellow) is another type with a good distribution on the map. The points corresponding to "sexual violence" (cyan) show a more dispersed distribution, but with some notable foci, particularly in urban and coastal areas. The presence of transnational violence (orange) is scarce compared to

other categories, although some specific foci can be identified (Fig. 14).
The distribution map also shows the geographic results for the variable “Actual area”:

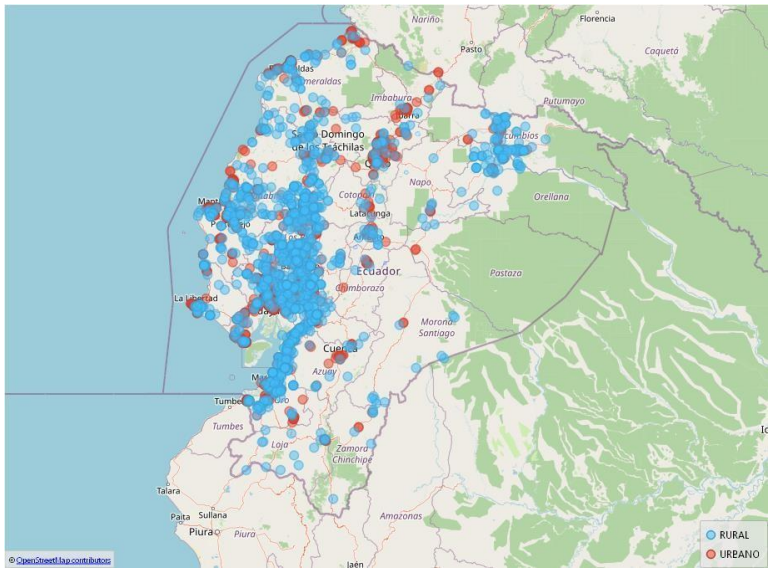

Figure 15. Map of distribution by Area of made

The map shows the geographical distribution of incidents in Ecuador by area, differentiating between urban (red) and rural (blue) areas. Incidents in rural areas predominate, although Urban incidents show concentration in major cities. The visualization highlights the disparity between urban and rural areas, and the importance of considering differentiated approaches for each type of area (Fig. 15).

Flow 3: Analysis of the model of classification (Naive Bayes)

The results showed through Features Importance that variables as he place of the crime and the The age of the victims was the most influential, which could indicate specific patterns in the areas with higher homicide rates or in certain demographic groups (Fig. 16).

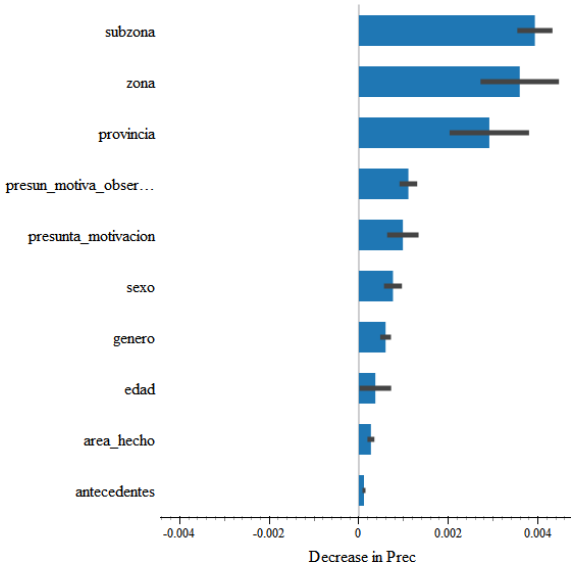

Figure 16. Result Feature Importance - Flow 3

On the other hand, through Test and Score the results showed that the model of Naive Bayes had reasonable accuracy, suggesting that, despite its simplicity, it is suitable for predicting homicide probabilities under certain conditions (Fig. 17)

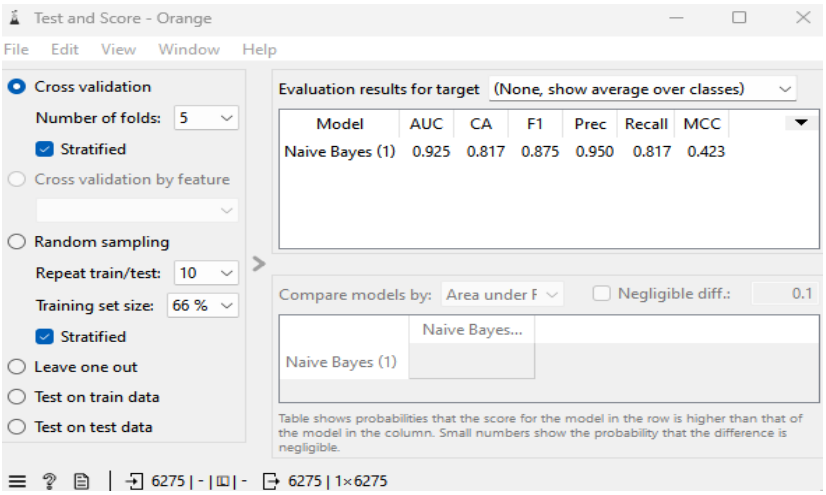

Figure 17. Results Test and Score - Flow 3

Analysis with the Naive classifier Bayes showed an accuracy of 76.6%, revealing that the variables Most influential in predicting homicide were geographic location and age of the victim.

The confusion matrix showed greater accuracy in the classification of intentional homicides in urban areas, while the most frequent errors occurred in cases from rural areas with lower data density.

Table 1. Results Matrix of Confusion

Actual \ Predicted	MURDER	FEMICIDE	HOMICIDE	CONTRACT KILLING
MURDER	76.6%	0.6%	3.8%	19.0%
FEMICIDE	4.1%	93.2%	1.4%	1.4%
HOMICIDE	12.9%	3.2%	71.8%	12.1%
CONTRACT KILLING	60.0%	0.0%	10.0%	30.0%
Σ	4509	112	496	1158

Flow 4: Analysis exploratory of data

Once the columns considered relevant for the analysis have been selected using the widget Select Comumns was connected to a Box Plot to analyze the categorization of homicides by type of death and type of weapon, thus obtaining.

The graph illustrates the distribution of the types of weapons used in different types of deaths: murders, homicides, femicides and contract killings. In the In the case of murders, a combination of weapons is observed, although with a marked predominance of firearms (green), followed by a significant proportion of knives (blue) and a lower incidence of blunt weapons (red). Homicides, on the other hand, show a higher proportion of use of knives (blue), followed by blunt weapons (red) and firearms (green). A portion of homicides caused by construction or similar objects (orange) and other types of objects (yellow) is also observed. In the femicides, the picture is similar to that of homicides, with a predominance of bladed weapons (blue), followed by blunt weapons (red) and firearms (green). In contrast, hitmen They are characterized by an almost exclusive use of firearms (green), which suggests intentionality and planning in this type of crime. general, the graph sample that the Murders, homicides and femicides are usually carried out with a greater variety of weapons, while contract killings are almost exclusively associated with the use of firearms. These differences may indicate variations in the planning, context and motivation of each type of crime.

On the other hand, when making the comparison with the ages of the victims, HE noticed that in all the guys

of death the mean and the median are close, that is to say that there is no major dispersion in the data, in addition to the fact that the ages in all cases are similar, so it can be said that the population group mainly Those affected by this problem are adults between 22 and 43 years old.

Widget was used Distributions to analyze how the data is distributed within this sample and it was observed that Guayas is by a lot the province with elderly index of homicides, followed of Manabi and Los Rios.

In order to have a clearer vision of the patterns that could influence to the Commitment of the homicides, was related by Scatter Plot the variables type of deaths and presumed motivation associated to the sex of the victims. The distribution of the data shows that in general terms homicides are mainly associated with common crime, however, it can be notice that out of this cause, The cases where the victims are men are distributed in different groups, that is, they are motivated by different reasons, unlike the homicides of women, which are concentrated in domestic violence

Flow 5: Analysis of the nomogram and characterization of clusters

The results obtained reflect how the applied clustering and classification techniques allow identifying significant patterns in the data and understanding the relevance of the variables selected in the segmentation. This section describes the main observations derived from the analysis of the Cluster 1 and 2 (C1 and C2) where the "dot" scale at the top of the graph indicates how each variable influences the probability of a case belonging to the analyzed cluster, while the "total" values at the bottom show the total contribution of the selected variables.

- **Cluster C1:**

The categories inside of the variables that influence in the belonging to the cluster C 1 (Fig. 23) are the following:

1. **Place:** Homicides in places such as "public transport", "warehouses", "parking lots", "beauty salons", "hotels", " night clubs", "industries", and "bus stops/bars" tend to have a negative impact on membership in cluster C1. In contrast, places such as "river", "coffee shops", "beach", "ports", "apartment in a house or building", "farm", "country houses", "public companies" and "shopping centres" have a negative impact on membership in cluster C2. a greater probability belonging to cluster C1. That is, the place is a strong differentiator. Homicides that occur in places more associated with the life everyday life has more to do with cluster C1.

2. **Alleged motivation:** Motivations such as "debts", "terrorism", "property litigation" and "car theft" tend to decrease the probability of belonging to C1. On the other hand, motivations such as "threat", "lynching", "cattle rustling" and "smuggling" show a higher probability of belonging. to the cluster C1. HE can infer that he guy of motivation of homicide play a crucial role. Conflicts or criminal motivations have a greater association with cluster C1, unlike homicides motivated by economic issues. or organized crime.

3. **Ethnicity:** The "Asian" ethnicity has a negative correlation with cluster C1. The "indigenous" and "Afro" ethnicities have a negative correlation with cluster C2. positive with cluster C1.

4. **Gender:** Homicides involving " transfeminine " or " trans " gender persons have a high correlation with cluster C1. In contrast, homicides involving "transfeminine" or "trans" gender persons are highly correlated with cluster C1. "male" have a lower probability of belonging to cluster C1. Thus, it can be observed that gender is a strong predictor, cases of trans or transfeminine people have a great positive influence, while the presence of a man in homicide has a effect negative.

5. **Province:** Provinces such as Pastaza, Imbabura, Los Ríos, Chimborazo, Napo and Zamora Chinchipe have a positive relationship with cluster C1. Provinces such as Azuay, Loja and El Oro decrease the probability of membership.

6. **Type of death:** The category " femicide " has a strong correlation with cluster C1. "Homicide" or "Murder" have a strong correlation negative.

7. **Weapon:** The type of weapon plays a role, although it is minor. The presence of a "blade weapon" slightly increases the probability of belonging to cluster C1, unlike "blunt weapon" or "construction weapon".

8. **Area of the event:** Cases occurring in rural areas have a slight positive correlation with cluster C1.

- **Cluster C2:**

The categories inside of the variables that influence in the belonging to the cluster C2 (Fig. 24) are the following:

1. **Location:** Location is a distinctive factor. Places associated with commerce and leisure, such as shopping malls or gyms, have less of a relationship with the cluster C2. Places associated with the crime, urban life or public transport are more likely to belong to this cluster.

2. **Alleged motivation:** The type of motivation is a differentiator. homicides partners to Robberies, sexual crimes or mental disorders are less likely to belong to C2, while homicides associated with criminality, conflicts or emotional issues are more likely to belong to this cluster.

3. **Ethnicity:** Ethnicity is a relevant factor. Homicides involving people of mixed-race, indigenous or montubia ethnicity have a higher correlation with cluster C2, while people of white or Asian ethnicity have a lower correlation.

4. **Gender:** Gender is a strong predictor. Homicides in which the victim is male are more strongly associated with cluster C2, while the presence of trans or transfeminine people has the opposite effect.

5. **Province:** Provinces such as El Oro, Sucumbíos, Loja, Azuay, Tungurahua and Guayas have a negative correlation with cluster C2. On the other hand, provinces such as Bolívar, Pichincha, Cotopaxi, Manabí, Napo and Pastaza present a correlation positive.

6. **Type of death:** The type of death is a distinguishing factor. Hitman and murder HE associate more with cluster C2, and femicides have less relation with this cluster.

7. **Weapon:** The use of "Firearm" has a positive correlation with the C2 cluster, while "Blunt Weapon" or "Constructor" have a negative correlation.

8. **Area of the event:** Homicides that occur in "urban areas" have a elderly correlation with Cluster C2. Those that occur in rural areas present correlation negative.

Between the patterns further significant in each cluster HE find:

- **Cluster C1:** Nomogram analysis revealed that the place, the alleged motivation and the ethnicity are very influential variables for cluster C1. Within this cluster, a significant pattern emerges in homicides that occur in places associated with daily life, such as apartments in houses or buildings or farms, where the motivation is linked to interpersonal conflicts or criminality such as threats or lynching, and where the victim is of indigenous or Afro ethnicity. This pattern suggests that cluster C1 is associated with homicides of a local and conflictive nature, which occur in a domestic or everyday context where indigenous or Afro-descendant populations are more vulnerable.
- **Cluster C2:** The results of the Nomogram indicate that the place, the alleged motivation and the gender are key components of cluster C2. A significant pattern in this cluster is homicides that occur in places associated with crime or urban life, such as public transport or public roads, motivated by acts of hate or threats, and where the victims are of gender male. This pattern of cluster C2 point out toward a typology of homicides with a strong burden of interpersonal violence occurring in urban areas and with male victims.

4. Discussion

This study reveals patterns differentiated in the homicides in Ecuador, evidencing two big guys violence with its own characteristics. The first is associated with homicides resulting from interpersonal conflicts and everyday contexts, while the second is more closely linked to organized crime and urban violence. As for the location of the crime, homicides tend to occur in private spaces such as homes or buildings, followed by public spaces such as beaches, rivers and transportation, which reflects different dynamics of violence in the country.

Regarding the motivations for homicides, they are mainly related to interpersonal conflicts and criminal motivations such as drug trafficking. This pattern is consistent with what has been reported by Ochoa (2023), who pointed out that sentimental motivations predominate in femicides, while organized crime contributes to the increase in violence, as Pierre highlights. (2013).

In how much to the ethnicity, HE observe a elderly incidence of homicides in people of ethnic groups mixed race,

indigenous and Afro-descendants, which could be related to the practices of racial discrimination prevalent in Ecuador, as pointed out by Benavides (2013). On the other hand, the relationship between mestizo, indigenous and Afro-descendant ethnic groups and montubias with urban violence also highlights the intersection between discrimination, crime and violence, which is presented as an important variable for the study of the criminality in the country.

He gender is other factor clue in the patterns observed. The homicides of people trans and

Transfeminine women associate with interpersonal conflicts, while men are more linked to organized crime. This finding reinforces the need to address violence from a gender perspective. It coincides with the studies of Cisne et al. (2020), who reported a higher incidence of violent deaths in men, which is linked to homicides related to the organized crime. Finally, the type of death also reflects this division, already that he femicide HE associated with interpersonal violence, while contract killings and murders with firearms are more closely linked to criminal violence. This is compared with the report of the Ecuadorian Observatory of Organized Crime (OECO, 2024), which links the increase in homicides with the increase in crimes committed with firearms.

A time preprocessed he text related with the theme "Security in he Ecuador" and using the Word Cloud tool, the following word cloud was obtained.

The word cloud generated from the theme "Security in Ecuador" reveals central terms such as "police," "security," "social," "Ecuador," and "national," which are the key concepts in this topic. In addition, "forces," "armed," "government," and "ministry" stand out, indicating the importance of institutions and entities in managing security. The presence of words such as "crisis," "crime," and "murdered" points out problems and challenges in the matter. The legal and political context is reflected in words such as "system," "law," and "constitution." So Same, aspects are mentioned social and economic ones like "banks," "work," "health", "well-being" and "education", demonstrating that security is intertwined with various aspects of society. The word cloud, taken together, offers a vision general overview of the key components that influence the security issue in Ecuador (Fig. 25).

5. Conclusions

This study offers a detailed overview of intentional homicides in Ecuador, analyzing the types of death, the provinces with the highest incidence, and the motivational factors behind these crimes. Murder was identified as the most frequent type, while contract killing represents a lower percentage. In addition, the provinces of Guayas, Manabí, and Los Ríos have the highest homicide rates, suggesting that geographic, economic, and social factors may be influencing this problem. As for motivations, common crime is the main cause of homicides, with the majority of victims being mestizo men. However, in cases where the victims are women, a strong relationship with domestic violence was evident, which highlights the persistence of gender violence as an urgent social problem. It was determined that the age of the victims ranges between 22 and 43 years, with no significant variations. The integration of C1 and C2 cluster analysis allowed for a more holistic understanding of the

phenomenon, identifying distinct patterns in the occurrence and motivation of homicides. This data segmentation provides a solid basis for the design of more effective prevention and mitigation strategies. In this context, the Orange tool proved to be fundamental for data analysis, offering a visual interface that facilitates the exploration of trends and patterns. Its ability to integrate data mining, preprocessing and visualization processes allows for efficient and accessible analysis for data-driven decision making.

From a neutrosophic perspective, this research can be strengthened by applying neutrosophic logic and its methods, which allow for handling the uncertainty, contradiction, and indeterminacy present in the analysis of complex phenomena such as intentional homicides. Neutrosophic theory, by extending classical and fuzzy logic, allows for modeling factors that cannot be classified in a binary manner (such as "common crime" vs. "gender violence"), but rather in degrees of belonging that simultaneously consider the certainty, falsity, and indeterminacy of the data. The use of fuzzy neutrosophic sets could help to better interpret the interrelationship between variables such as geography, socioeconomic level, motivations, and context of homicides, offering a more nuanced view of causal relationships. Likewise, neutrosophic decision-making could be applied to evaluate multiple scenarios in the design of crime prevention policies, considering the variability and uncertainty inherent to the phenomenon of violence. In conclusion, the results of this study provide key information for future research and homicide prevention strategies in Ecuador. The integration of advanced analysis tools and innovative approaches such as neutrosophy will allow for the development of more accurate predictive models and more effective public policies, thus contributing to the reduction of violence and the improvement of citizen security.

References

[1] V. Artus (2024), " Clustering Metrics: Evaluate the complexity, make it simple," Medium, [Online]. Available: https://medium.com/@vladimir-artus/%D1%81lustering-metrics-evaluate-the-complex-make-it-simple-6ae70c0f164b.

[2] M. Benavides Benalcázar (2013), "The crime of hate and the protection of the right to equality and non-discrimination," in Criminological Profile. Hate crimes: A recognition of equality and dignity, Attorney General's Office of the Ecuadorian State, pp. 4–7, [Online]. Available: https://www.fiscalia.gob.ec/images/PerfilCriminologico/criminologico7.pdf.

[3] D. Hofmeyr, F. Kamper , and M. Melonas (2024), "Optimal Projections for Classification with Naive Bayes," arXiv [stat.ML], [Online]. Available: http://arxiv.org/abs/2409.05635.

[4] M. Lopez Bravo, R. Maria del Swan, E. Mejia Suarez, and L. Ordonez Mullo (2020), "Deaths from injuries caused by firearms and explosives, registered at Quito Institute of Forensic Sciences and Criminalistics from 2017 to 2019," Universidad Ciencia y Tecnología , vol. 24, no. 103, pp. 35–40, doi :10.47460/uct.v24i103.355.

[5] Ministry of the Interior EC (2025), "Intentional Homicides Ecuador 2024," Open Data Catalog of the Government of Ecuador, [Online]. Available: https://www.datosabiertos.gob.ec/dataset/homicidios-intencionales/resource/cb8f704e-2b27-4d7f-9431-d40c4e27fa48.

[6] Ecuadorian Observatory of Organized Crime (OECO) (2024), "Biannual bulletin of intentional homicides in Ecuador," [Online]. Available: https://oeco.padf.org/wp-content/uploads/2024/09/Boletin-semestral-de-homicidios-Primer-semestre-de-2024_compressed.pdf.

[7] Femicidal violence in Ecuador: analysis with a criminological perspective," Revista de Derecho, vol. 8, no. 1, pp. 80–102, doi : 10.47712/rd.2023.v8i1.236.

[8] M. Romano, G. Zammarchi , and C. Conversano (2024), "Iterative threshold-based Naïve Bayes classifier," Statistical Methods & Applications, vol. 33, no. 1, pp. 235–265, doi :10.1007/s10260-023-00721-1.

[9] KL Sainani (2016), "The value of scatter plots," PM & R: The Journal of Injury, Function, and Rehabilitation, vol. 8, no. 12, pp. 1213–1217, doi: 10.1016/j.pmrj.2016.10.018.

[10] P. Salama (2013), "Homicides: Is violence inevitable in Latin America?" Frontera Norte, vol. 25, no. 49, pp. 7–27, [Online]. Available: https://fronteranorte.colef.mx/index.php/fronteranorte/article/view/790/301.

[11] JV Neto, CB Dos Santos, É. M. de Torres, and C. Estrela (2017), " Boxplot: a graphic resource for the

analysis and interpretation of quantitative data," Revista Odontológica do Brasil Central, vol. 26, no. 76, pp. 1–6, doi :10.36065/robrac. v26i76.1132.

[12] HA Zapata Mayorga and A. Kumar Vivas (2024), "Analysis of Intentional Homicides in Ecuador Using Tree Regression and Random Forest Models," Scientific Journal.

[13] P. Liu and Y. Wang (2014), "Multiple attribute decision-making method based on single-valued neutrosophic normalized weighted Bonferroni mean," Neural Computing and Applications, vol. 25, no. 7-8, pp. 2001–2010, doi : 10.1007/s00521-014-1688-8.

[14] S. Broumi, M. Talea, A. Bakali , and F. Smarandache (2016), "Single valued neutrosophic graphs," Journal of New Theory, no. 10, pp. 86–101, [Online]. Available: http://www.newtheory.org/archive/volume-10-2016.

[15] R. Şahin (2017), "Neutrosophic hierarchical clustering algorithm and its application in decision making," Journal of Intelligent Systems, vol. 26, no. 3, pp. 515–526, doi: 10.1515/jisys-2015-0131.

[16] N. Batista-Hernández, M. Leyva-Vázquez, and F. Smarandache (2020), "A neutrosophic multicriteria decision-making approach for personnel selection," Neutrosophic Sets and Systems, vol. 34, pp. 76–85, doi :10.5281/zenodo.3892958.

Recibido: febrero 28, 2025. Aceptado: marzo 18, 2025