
Exploratory Analysis and Clustering of Microdata on Lethal Violence Events in Ecuador: 2020-2024

Análisis Exploratorio y Clustering de microdatos de eventos de violencia letal en Ecuador: 2020-2024

Elio Edwin Sánchez Suárez*, Lester Xavier Rodríguez-Cruz*, Meiling Mishelle López Sotomayor*, Irma Ines Gaibor Garcia*

Abstract

Received: July 12, 2025

Approved: September 03, 2025

Sánchez, E., Rodríguez, L., López, M., Gaibor, I. (2025) Exploratory Analysis and Clustering of Microdata on Lethal Violence Events in Ecuador: 2020-2024. *Revista multidisciplinaria de investigación científica*, Vol. 9, No. 4. 48-64

The analysis of homicides between 2020 and 2024 reveals three territorial profiles. Using k-means, we can distinguish between cantons with very high levels of armed violence on the coast, moderate levels in the highlands and Amazon region, and low levels in rural areas. With DBSCAN, in addition to the main groups, atypical cantons with critical rates emerge, such as Olmedo, Pueblo Viejo, and La Troncal, considered extreme hotspots that require differentiated attention. Overall, the results show that violence is concentrated on the coast, associated with the intensive use of firearms, while other regions have lower incidence rates.

Keywords: clustering, criminology, risks, citizen security

* Magister En Finanzas Y Proyectos Corporativos
Instituto Superior Tecnológico Babahoyo,
Ecedwinsanchez@Gmail.Com, 0000-0002-0644-3238

* Economista, Universidad De Guayaquil,
Lester.Rodriguezc@Ug.Edu.Ec, 0000-0002-2931-0702,

* Ingeniera En Diseño Gráfico, Instituto Superior Tecnológico Babahoyo, 0009-0005-2928-5331

* Instituto Superior Tecnológico Babahoyo, 0000-0002-6262-9723, Igaibor@Istb.Edu.Ec, Magister En Administración De Negocios, Mención Negocios Internacionales

Resumen

El análisis de homicidios 2020-2024 revela tres perfiles territoriales. Con k-means, se distinguen cantones con violencia muy alta y armada en la Costa, niveles moderados en la Sierra y Amazonía, y bajos en zonas rurales. Con DBSCAN, además de los grupos principales, surgen cantones atípicos con tasas críticas como Olmedo, Pueblo Viejo y La Troncal, considerados focos extremos que demandan una atención diferenciada. En conjunto, los resultados muestran que la violencia se concentra en la Costa, asociada al uso intensivo de armas de fuego, mientras otras regiones presentan menor incidencia.

Palabras Clave: Clustering, criminología, riesgos, seguridad ciudadana

Introduction

The main objective of this research project is to conduct a detailed exploratory analysis of intentional homicides in Ecuador during the period 2020-2024, using the Ministry of Government's database. The phenomenon of intentional homicides is one of the most critical challenges to public safety, affecting quality of life and governance. The microdata recorded on the Open Data portal provides official, accessible, consistent, and timely statistical information on intentional homicides, representing an indispensable source for the development of studies that contribute to the design of effective prevention policies.

Exploratory data analysis (EDA) is essential for the study of intentional homicides in Ecuador, as it allows hidden patterns to be discovered and helps us understand how these events are distributed in both time and space. Through this approach, emerging trends in the analysis period can be identified and possible anomalies that reveal changes in criminal dynamics, the impact of public policies, or current social and economic situations can be recognized. In this way, EDA becomes an essential initial tool for the objective and rigorous interpretation of available information.

The application of clustering analysis techniques, such as k-means and DBSCAN, adds value to the analysis by allowing the segmentation of lethal violence into groups with common characteristics. These methods make it possible to identify "hot spots" and critical areas with a high concentration of homicides, as well as to pinpoint particularly sensitive time periods. The ability to group information objectively contributes to a clearer understanding of the underlying dynamics and facilitates the prioritization of territories and contexts that require immediate attention.

The study explores the distribution of crimes on an annual basis, their concentration in specific areas such as cantons, and the key characteristics of the victims and the incidents, such as gender, age, the weapon used, and the presumed motivation. Through data visualization and descriptive analysis, this

study will provide a clear and well-founded overview of crime dynamics in Ecuador.

During the period 2020-2024, Ecuador has faced significant increases in intentional homicides, with a persistent concentration in urban areas such as Guayaquil and Durán. The Ecuadorian Observatory on Organized Crime (OECO) documented a high homicide rate, placing the country among the most violent in Latin America, despite a reduction in 2024 compared to 2023. The predominance of firearms in these crimes and the geographical grouping of homicides highlight the need for studies that identify territorial clusters for better institutional and social targeting (OECO, 2025).

In line with this, Basantes et al. (2025) applied spatial analysis and clustering techniques in Region 1 of Ecuador to study homicidal violence, finding a high concentration of violent acts in public spaces and specific urban areas. Their study highlights the usefulness of clustering to geographically segment risk areas, allowing for the addressing of fundamental structural factors such as poverty and inequality that underlie these homicides.

Complementarily, InvestigaGeográfica (2025) analyzed the spatiotemporal evolution of homicides from 2015 to 2022, locating persistent clusters in provinces on the Ecuadorian coast. The combined use of statistical analysis and territorial clustering in their research facilitated the understanding of spatial trends in violence, crucial information for the formulation of selective and targeted public policies. At the micro-local level, Simbaña Collaguazo (2024) studied intentional homicides in the La Delicia district of Quito, combining quantitative analysis with clustering techniques to identify critical areas affected by disputes between criminal groups. This approach also allowed for the integration of social and cultural dimensions into the analysis, broadening the understanding of the phenomenon in complex urban contexts.

Finally, the official annual reports of the Ministry of the Interior and the National Police reinforce the evidence on the concentration of homicides in specific territorial clusters. The systematic analysis they apply incorporates these methodologies to improve resource allocation and design targeted interventions that respond to the heterogeneity of the violent phenomenon in Ecuador (Ministry of the Interior and National Police, 2024; OECO, 2025).

Materials and methods

This study uses a database of intentional homicides in Ecuador, provided by the Directorate of Violent Deaths, Disappearances, Extortion, and Kidnappings (DINASED) of the Ministry of Government. The dataset covers the period from 2014 to 2024, with a focus on the years 2020 to 2024 for the main analysis.

The database, with a total of 30,508 records, includes detailed information on the victims, the type of crime, the characteristics of the events, and the geographic location (provinces, cantons, and coordinates). The key variables used in this study are: date of offense, canton, province, sex, age, weapon, and presumed motivation.

The methodology of this research is structured in two main phases: exploratory data analysis (EDA) and cluster analysis. This sequential approach allows for an in-depth and systematic exploration of the crime phenomenon, moving from a general understanding of trends to the identification of more specific patterns and clusters in the homicide data.

The first phase focuses on Exploratory Data Analysis (EDA). Using descriptive statistics and visualization techniques, the aim is to obtain an initial understanding of the distribution of intentional homicides in Ecuador. Temporal trends will be examined over the period 2020-2024, analyzing the variation in crimes on an annual basis. The geographical distribution will also be explored to identify cantons with the highest incidence. In addition, victims and criminal acts will be characterized through variables such as sex, age, type of weapon used, and presumed motivation, providing a detailed overview of lethal violence in the country. This analysis is supported by the capabilities of R libraries such as tidyverse, lubridate, and ggplot2 (Wickham et al., 2023; R Core Team, 2024).

The second phase of the study focuses on cluster analysis, an unsupervised learning technique. The purpose is to go beyond simple description to identify natural groupings of cantons that share similar crime profiles.

To achieve this, two complementary algorithms will be used. The first algorithm, K-means, will partition the cantons into a predefined number of groups based on the similarity of their characteristics, such as the homicide rate per 100,000 inhabitants and other key variables. This method is effective for grouping spherical and well-separated data (James et al., 2013). In this analysis, five (5) key variables were used for the period, such as homicide rate; and the proportions of crimes involving firearms, on public roads, due to criminal violence, and with male victims.

Formally, k-means is a participatory method that seeks to minimize cluster inertia, defined as the sum of the squared distances of each point to its centroid.

Inertia Function:

Where:

- K: The number of clusters.
- i : The i -th cluster.
- x : A data point in cluster i .

· : The centroid of cluster , which is the average of all points in

The algorithm operates iteratively: (1) assigning points to their nearest cluster and (2) updating centroids to the average of the assigned points. This process continues until convergence.

In this algorithm, the elbow method is used to obtain the optimal number of clusters (k). This method searches for the point where the addition of a new cluster does not significantly reduce the variance within the groups. Once the optimal k is determined, the k-means algorithm groups the cantons into clusters, minimizing the distance of each canton to the center of its cluster.

The second algorithm, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), will be used to complement the K-means analysis. Unlike the latter, DBSCAN does not require the number of clusters to be specified in advance and is capable of identifying groups arbitrarily, as well as detecting outliers (considered "noise"). This feature is particularly useful in spatial analysis, where crime concentration can vary dramatically from one area to another (Ester et al., 1996).

DBSCAN is a density-based algorithm that defines clusters as dense regions of points, separated by regions of lower density. The formalization is based on the concept of density accessibility.

All points within a distance of a point: -neighborhood

A point is a core point if its: -neighborhood contains at least minPts points

minPts

A point is reachable by density from a point if there is a sequence of core points that connects it to .

Clusters are formed by expanding the core points. Points that are not reachable by density from any core point are considered noise (outliers). The combination of these two approaches to clustering will allow us to discover distinctive risk profiles and understand the underlying structure of crime, which can serve as a basis for the design of more precise and targeted public safety policies.

Results

Data on intentional homicides in Ecuador between 2020 and 2024 show an explosive growth trend, with a massive increase in criminal violence, as shown in Table 1. Throughout this period, the number of homicides peaked in 2023, with the most significant increase occurring between 2021 and 2022, when the number of cases nearly doubled, rising from 2,495 to 4,886, representing an increase of 95.83%. Although the data for 2024 shows a decrease of 14.73% compared to the previous year, the total number of homicides (7,033) remains

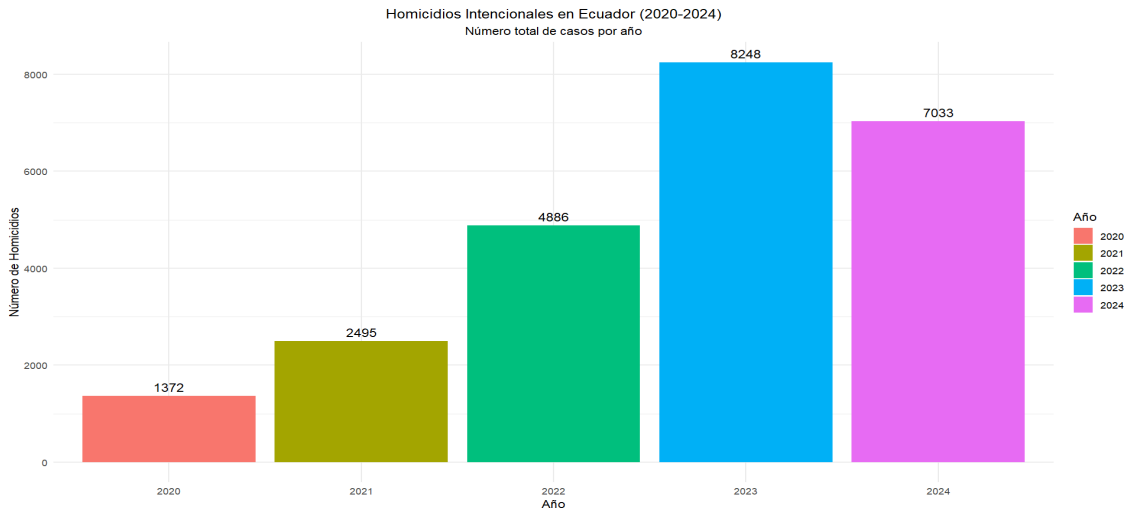
high. This figure is more than triple that recorded in 2020 (81,372), underscoring the persistence of a serious security crisis.

Table 1. *Intentional homicides per year in Ecuador: 2020-2024*

| Year | Total Homicides |
|-------|-----------------|
| 2020 | 1,372 |
| 2021 | 2,495 |
| 2022 | 4,886 |
| 2023 | 8,248 |
| 2024 | 7,033 |
| Total | 24,034 |

The general trend for the period 2020-2023 indicates a rapid and sustained deterioration in public safety, as can be seen in Figure 1; and, despite a slight improvement in 2024, Ecuador continues to face a significant challenge in terms of violence.

Figure 1. *Trend in intentional homicides in Ecuador: 2020-2024*



Prepared by: Authors

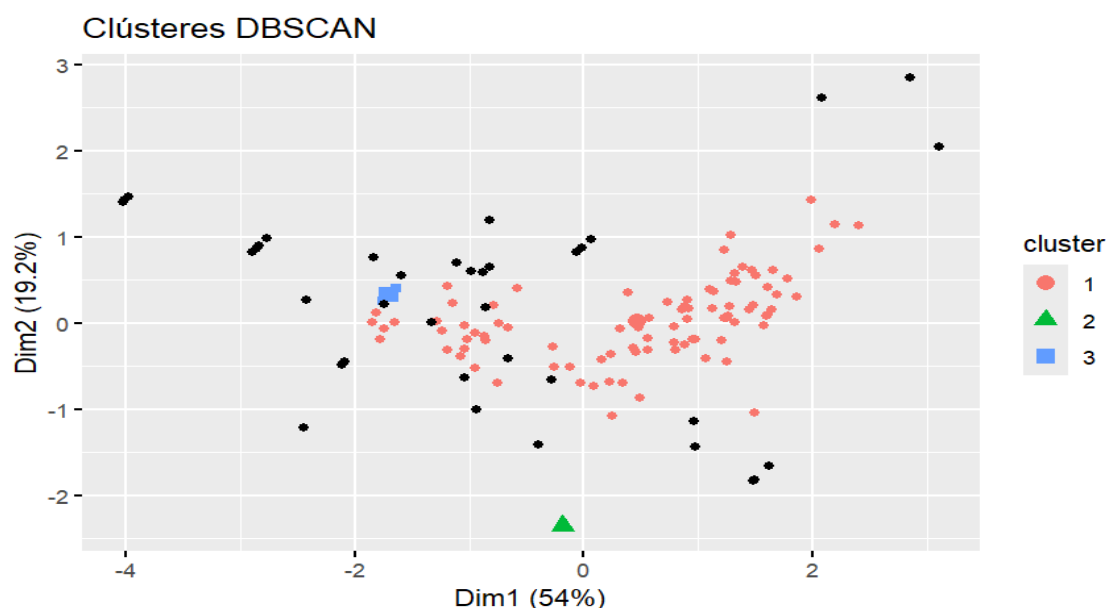
Cluster analysis applied to homicides in Ecuador between 2020 and 2024, using the DBSCAN and K-means algorithms, revealed significant patterns in the distribution and concentration of crime in the country's cantons.

Cluster 0 represents noise (outliers); the 38 cantons in this group are statistical anomalies. Although most of them have a very low total number of homicides, their homicide rates are extremely high due to their low population. This group includes cantons such as Olmedo, Pueblo Viejo, and La Troncal, which stand out with rates of 696, 619, and 575 per 100,000 inhabitants, respectively, making them critical areas or hotspots of lethal violence. However, it also includes cantons with only one or two homicides but with such a small population that

their rate skyrockets. These anomalous cantons do not fit the profiles of the main clusters.

The rest of the cantons are grouped into denser clusters. Cluster 1 (89) cantons is the largest group. Clusters 2 and 3 are very small (5 cantons each), indicating that they group together cantons with very particular profiles, but which are not as extreme as those in the noise group.

Figure 2 DBSCAN clusters of crime rates by canton in Ecuador: 2020-2024



Prepared by: Authors

Figure 2 of the DBSCAN cluster shows how cantons are grouped according to the density of their homicide characteristics. The black dots represent noise, i.e., cantons without a dense grouping, while the other colors and symbols indicate clusters with similar characteristics within each group. The distribution reflects the spatial heterogeneity of crime, highlighting isolated critical areas and groups with homogeneous patterns. This pattern suggests that violence is not distributed evenly, but is concentrated in certain cantons with particular dynamics, which are important for targeting interventions.

Table 2. DBSCAN Cluster Counties

| No | Clúster | Número de Cantones | Cantones | Observación |
|----|---------|--------------------|--|--|
| 1 | 0 | 38 | ATAHUALPA, CALUMA, CALVAS, CELICA, CEVALLOS, CHAGUARPAMBA, CHAMBO, CHIMBO, | Estos puntos representan a los cantones que no pudieron ser asignados a ningún |

| | | | | |
|---|---|----|---|--|
| | | | CHUNCHI, COLTA, EL CHACO, EL TAMBO, GUALACEO, HUAMBOYA, LA TRONCAL, LAS LAJAS, MOCHA, OLMEDO, PABLO SEXTO, PALLATANGA, PATATE, PAUTE, PEDRO VICENTE MALDONADO, PIMAMPIRO, PINDAL, PUEBLOVIEJO, PUERTO QUITO, PUYANGO, ROCAFUERTE, SALCEDO, SAN JUAN BOSCO, SAN MIGUEL DE LOS BANCOS, SANTA ISABEL, SANTIAGO, SARAGURO, SEVILLA DE ORO, TIWINTZA, ZAPOTILLO | clúster. Según la lógica de DBSCAN, estos cantones son valores atípicos con perfiles de homicidios únicos o extremos, que los separan de los grupos densos. Los puntos de ruido son de particular interés para la investigación, ya que representan los focos de violencia más inusuales y a menudo más graves del país. |
| 2 | 1 | 89 | 24 DE MAYO, AMBATO, ANTONIO ANTE, ARCHIDONA, ARENILLAS, ATACAMES, AZOGUES, BABA, BABAHoyo, BALAO, BALZAR, BUENA FE, CARLOS JULIO AROSEMENA TOLA, CATAMAYO, CAYAMBE, CHILLANES, CHONE, COLIMES, COTACACHI, CUENCA, DAULE, EL CARMEN, EL GUABO, EL TRIUNFO, ELOY ALFARO, ESMERALDAS, FLAVIO ALFARO, GUALAQUIZA, GUANO, GUARANDA, GUAYAQUIL, HUAQUILLAS, IBARRA, ISIDRO AYORA, JAMA, JIPIJAPA, LAS NAVES, LATACUNGA, LOJA, LOMAS DE SARGENTILLO, | Grupo más grande y más denso, representa a los cantones con una problemática de criminalidad promedio o moderada en el país. |

| | | | | |
|---|---|---|---|---|
| | | | <p>MACHALA, MANTA, MERA, MILAGRO, MIRA, MOCACHE, MONTALVO, MONTECRISTI, MORONA, MUISNE, NARANJAL, NARANJITO, NOBOL, OLMEDO, OTAVALO, PALENQUE, PALESTINA, PALORA, PANGUA, PASAJE, PASTAZA, PEDERNALES, PEDRO CARBO, PEDRO MONCAYO, PICHINCHA, PLAYAS, PORTOVELO, PORTOVIEJO, QUEVEDO, QUINSALOMA, RIOBAMBA, RIOVERDE, SALITRE, SAN JACINTO DE YAGUACHI, SAN LORENZO, SAN MIGUEL, SAN PEDRO DE PELILEO, SAN VICENTE, SANTA ANA, SANTA ROSA, SUCRE, TAISHA, TENA, TOSAGUA, URDANETA, VALENCIA, VENTANAS, VINCES, ZARUMA</p> | |
| 3 | 2 | 5 | <p>ARAJUNO, GUAMOTE, PALTAS, QUIJOS, TISALEO</p> | <p>Este clúster es extremadamente pequeño, con solo unos pocos puntos, e incluso podría considerarse un único punto en el espacio. Su aislamiento indica que estos cantones tienen características de homicidios muy específicas que los separan del resto.</p> |
| 4 | 3 | 5 | <p>BALSAS, EL PAN, QUERO, SANTA CLARA, SIGCHOS</p> | <p>Similar al Clúster 2, este grupo es</p> |

pequeño y compacto. Su ubicación también lo diferencia de la gran mayoría de los cantones. Es probable que este clúster contenga cantones con un perfil de criminalidad ligeramente diferente al de los otros grupos, pero no tan extremo como para ser considerado ruido.

For its part, the K-means analysis separated the total number of cantons into three groups with clear differences in the average homicide rate and the proportion associated with the use of firearms. Cluster 2, with 65 cantons, exhibits a high crime rate (207 homicides per 100,000 inhabitants) and a high incidence of armed violence, indicating areas where violence is a particularly serious problem.

Clusters 1 and 3 show lower crime rates and less association with firearms, indicating the existence of areas with varying levels of risk, from moderate to relatively low.

Figure 3. Elbow method for K-means clustering

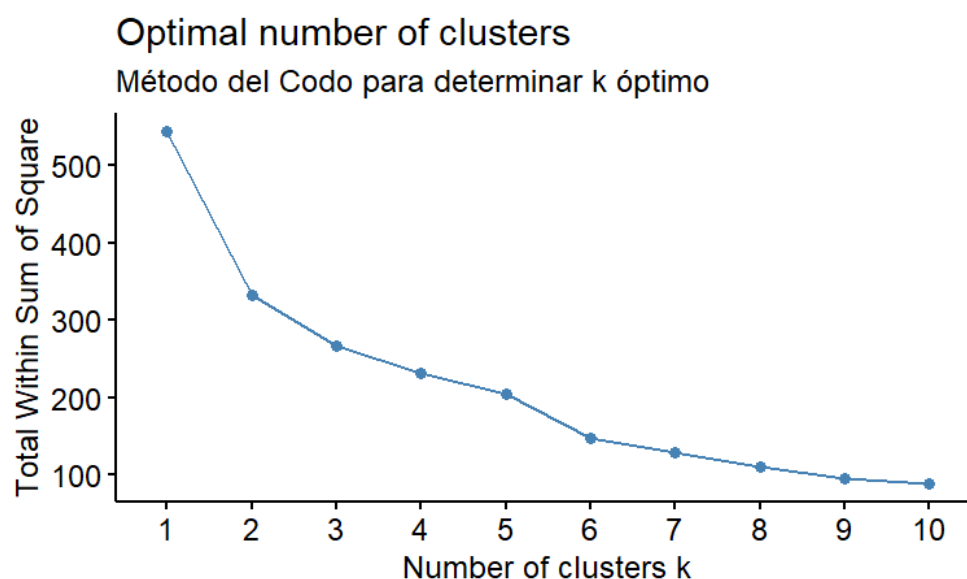


Table 4. *K-means Cluster Results*

| Clúster K-means | Cantones | Tasa Promedio Homicidios | Proporción arma de fuego promedio |
|-----------------|----------|--------------------------|-----------------------------------|
| 1 | 56 | 30.33 | 0.36 |
| 2 | 65 | 206.68 | 0.87 |
| 3 | 16 | 21.05 | 0.12 |

Figure 4. *K-means clusters of crime rates by canton in Ecuador: 2020–2024*

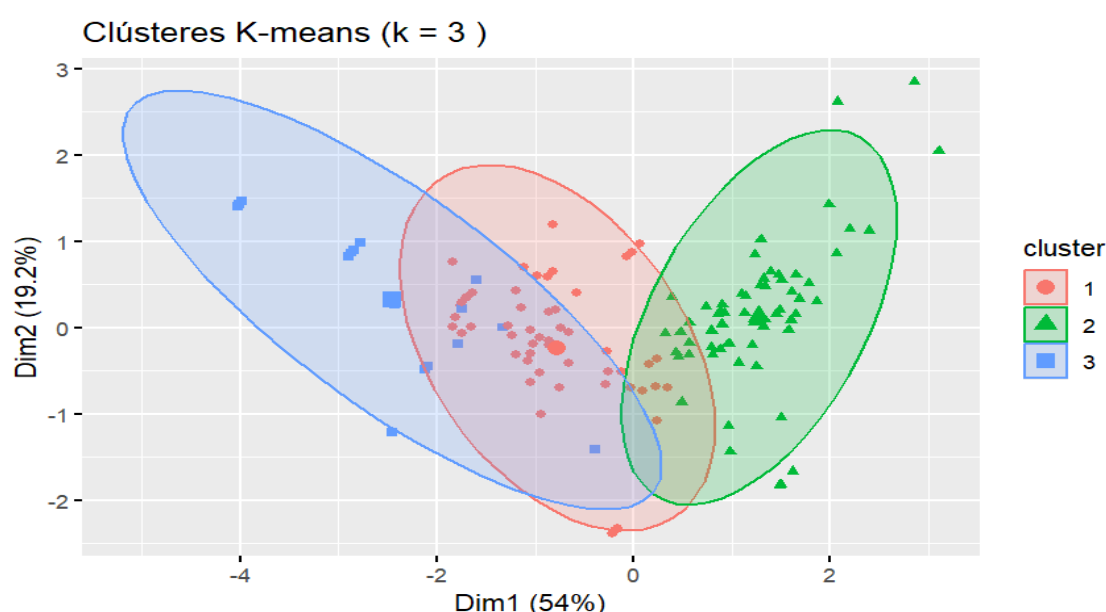


Figure 4 shows the three clusters identified: Cluster 1 (red dots) in the center, Cluster 2 (green dots) on the right, and Cluster 3 (blue dots) on the left. The colored ellipses surrounding each group do not define the exact boundary, but rather illustrate the dispersion of points within each cluster. The overlap of the red and blue ellipses in the central area suggests that the boundaries between these groups are not rigid, and there are points in that area that could be assigned to either of the two, indicating a degree of overlap in their characteristics.

The k-means analysis classifies the cantons into three distinct clusters. In this case, Cluster 1 represents cantons with a homicide rate of 30 per 100,000 inhabitants and an incidence of criminal violence involving firearms in approximately 36% of cases. These cantons are mainly inland, mountainous, and Amazonian, with average levels of violence during the period analyzed.

In contrast, Cluster 2 has a significantly higher homicide rate than Cluster 1 and is characterized by a higher proportion of homicides related to criminal violence and the use of firearms, suggesting a greater presence of organized crime and violent disputes. Table 4 shows that the average rate for this cluster is 207 homicides per 100,000 inhabitants, with firearms used in approximately 87% of cases. These are the most critical cantons, especially on the coast, including Guayaquil, Esmeraldas, Machala, Quevedo, and Manta. In other words, they represent the hotbed of armed violence.

Cluster 3 is the smallest, including cantons with an average homicide rate of 21 per 100,000 inhabitants and the use of firearms in approximately 12% of cases. These are cantons with very low and scattered violence, in rural or mountainous areas with lower population density.

Table 3. *K-MEANS Cluster Cantons*

| No | Clúster | Número de Cantones | Cantones | Observación |
|----|---------|--------------------------|--|--|
| 1 | 1 | 56 | AMBATO, ANTONIO ANTE, ARAJUNO, ARCHIDONA, ATAHUALPA, AZOGUES, BALSAS, CALVAS, CARLOS JULIO AROSEMENA TOLA, CATAMAYO, CAYAMBE, CEVALLOS, CHAGUARPAMBA, CHILLANES, CHIMBO, COTACACHI, CUENCA, EL PAN, FLAVIO ALFARO, GUALACEO, GUALAQUIZA, GUAMOTE, GUARANDA, HUAMBOYA, IBARRA, JIPIJAPA, LATACUNGA, LOJA, MERA, MIRA, MORONA, MUISNE, OTAVALO, PALLATANGA, PALORA, PALTAS, PANGUA, PASTAZA, PAUTE, PEDRO MONCAYO, PUYANGO, QUERO, QUIJOS, RIOBAMBA, SAN MIGUEL, | Grupo de criminalidad baja a moderada, presenta una tasa de homicidios promedio de 30.3 por cada 100,000 habitantes. |

| | | | | |
|---|---|----|---|--|
| | | | SAN PEDRO DE PELILEO, SANTA CLARA, SANTA ISABEL, SANTIAGO, SEVILLA DE ORO, SIGCHOS, TAISHA, TENA, TISALEO, ZAPOTILLO, ZARUMA | |
| 2 | 2 | 65 | 24 DE MAYO, ARENILLAS, ATACAMES, BABA, BABAHOYO, BALAO, BALZAR, BUENA FE, CHONE, COLIMES, DAULE, EL CARMEN, EL CHACO, EL GUABO, EL TRIUNFO, ELOY ALFARO, ESMERALDAS, GUAYAQUIL, HUAQUILLAS, ISIDRO AYORA, JAMA, LA TRONCAL, LAS NAVES, LOMAS DE SARGENTILLO, MACHALA, MANTA, MILAGRO, MOCACHE, MONTALVO, MONTECRISTI, NARANJAL, NARANJITO, NOBOL, OLMEDO, PABLO SEXTO, PALENQUE, PALESTINA, PASAJE, PATATE, PEDERNALES, PEDRO CARBO, PICHINCHA, PINDAL, PLAYAS, PORTOVELO, PORTOVIEJO, PUEBLOVIEJO, PUERTO QUITO, QUEVEDO, QUINSALOMA, RIOVERDE, ROCAFUERTE, SALITRE, SAN JACINTO DE YAGUACHI, SAN LORENZO, SAN VICENTE, SANTA ANA, SANTA ROSA, | Grupo con criminalidad alta y violenta, clúster más grande con una tasa de homicidio promedio de 207 por cada 100,000 habitantes. |

| | | | | | |
|---|---|----|-----------------------|---------------|---------------------|
| | | | SUCRE, | TOSAGUA, | |
| | | | URDANETA, | VALENCIA, | |
| | | | VENTANAS, | VINCES | |
| 3 | 3 | 16 | CALUMA, | CELICA, | Grupo con |
| | | | CHAMBO, | CHUNCHI, | criminalidad baja y |
| | | | COLTA, | EL TAMBO, | menos violenta. Con |
| | | | GUANO, | LAS LAJAS, | una tasa de |
| | | | MOCHA, | PEDRO VICENTE | homicidios promedio |
| | | | MALDONADO, | | de 21 por cada |
| | | | PIMAMPIRO, | SALCEDO, | 100,000 habitantes. |
| | | | SAN JUAN BOSCO, | SAN | |
| | | | MIGUEL DE LOS BANCOS, | | |
| | | | SARAGURO, | TIWINTZA | |

Conclusions

The k-means algorithm allows for clear segmentation, with high levels of armed violence (cluster 2) compared to moderate violence (cluster 1) and minimal levels of armed violence (cluster 3). The K-means algorithm with a fixed number of three groups ($k = 3$), unlike methods that identify clusters based on density, divides the dataset into three equal parts, assigning each point to the group whose centroid is closest. The Dim1 and Dim2 axes represent the two dimensions that explain most of the variation in the original data, allowing for a simplified visualization of the structure.

The DBSCAN algorithm highlights atypical cantons with disproportionate rates, which should be treated as critical cases or anomalies. The application of the DBSCAN algorithm yielded three clusters plus one "noise" group. Cluster 1, with 89 cantons, represents the main core and can be specified as moderate violence. Clusters 2 and 3, with five cantons each, are small groupings with particular patterns. With regard to the "noise" cluster, they do not fit into any group, many with extreme homicide rates. In this particular case, DBSCAN is useful because it identifies outliers that k-means dilutes within a larger cluster.

The combination of both methods provides a richer and more multifaceted view of lethal violence in the nation. While K-means offers a clear and numerically balanced segmentation of crime, DBSCAN allows for the capture of irregularities and extreme cases that require priority attention, such as high-risk areas that emerge as noise. Together, these techniques make it possible to draw up a detailed map of territorial crime profiles and structure a more accurate early warning system focused on the most affected areas.

In conclusion, the results support the need to design differentiated public policies that consider the particularities of each territorial cluster and, especially,

special treatment for cantons in noise that show extreme risks. The methodologies applied support evidence-based security strategies aimed at intervening with special emphasis on the critical points detected, thus improving the efficiency and effectiveness of reducing homicidal violence in Ecuador.

Referencias

- Basantes, C., Barahona, A., Barrionuevo, D., & Quespás, S. (2025). Análisis de tendencias y factores determinantes de la violencia en la zona 1 de Ecuador. *Dominio de las Ciencias*.
<https://www.dominiodelasciencias.com/ojs/index.php/es/article/view/4199>
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*.
- InvestigaGeográfica. (2025). Evolución espacio-temporal de los homicidios en Ecuador de 2015 a 2022. *Investigaciones Geográficas*.
<https://www.investigacionesgeograficas.com/article/view/27758>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer.

Ministerio del Interior y Policía Nacional del Ecuador. (2024). *Boletín anual de homicidios intencionales en Ecuador 2023*. <https://oeco.padf.org/wp-content/uploads/2024/04/OECO.-BOLETIN-ANUAL-DE-HOMICIDIOS-2023.pdf>

Observatorio Ecuatoriano de Crimen Organizado (OECO). (2025). *Boletín anual de homicidios intencionales en Ecuador ajustado 2025*. https://oeco.padf.org/wp-content/uploads/2025/06/Boletin-anual-de-homicidios-intencionales-en-Ecuador-ajustado_compressed.pdf

R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Simbaña Collaguazo, M. (2024). Análisis descriptivo de los homicidios intencionales en el distrito La Delicia, Quito. *Innovación y Saber*. <https://innovacionysaber.isupol.edu.ec/index.php/innovacion/article/view/283/586>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., ... & Yutani, H. (2023). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.