



Definition of geographic profiles of car theft. Case applied in Cartagena de Indias

Enrique Jose Delahoz Dominguez

Universidad Tecnológica de Bolívar, Colombia

edelahoz@utb.edu.co

<https://orcid.org/0000-0003-2531-6389>

Tomás José Fontalvo Herrera

Universidad de Cartagena, Colombia

tfontalvoh@unicartagena.edu.co

<https://orcid.org/0000-0003-4642-9251>

Adel Alfonso Mendoza Mendoza

Universidad del Atlántico, Colombia

adelmendoza@uniatlantico.edu.co

<https://orcid.org/0000-0002-4278-1226>

Received: November 7, 2019 / Accepted: February 11, 2020

<https://doi.org/10.17081/just.25.37.4215>

Abstract

This study develops a geospatial analysis of car theft that occurred in the city of Cartagena de Indias, Colombia. In defining crime profiles, theft was analyzed in the five modalities considered to have a high impact on the generation of citizens' perception of security, namely, motor vehicles, motorcycles, residences, commercial establishments, and financial institutions, in 2015 and 2016. The data used are from the annual crime report of the Colombian National Police. As a result, seven geographic profiles of car theft were identified, generating characterized crime patterns visualized geographically.

Keywords: machine learning, management, decision-making, SMEs.

How to quote:

Delahoz-Dominguez, E. J., Fontalvo-Herrera, T. J., & Mendoza-Mendoza, A. A. (2020). Definición de Perfiles Geográficos de hurto de automóviles. Caso Aplicado en Cartagena. *Justicia*, 25(37), 99-108. <https://doi.org/10.17081/just.25.37.4215>

I. Introduction

This study defines geographic profiles of vehicle theft in Cartagena using the data mining (DM) technique. The stated objective is to analyze and describe the geospatial patterns and trends identified in car thefts that occurred in the city in 2015 and 2016, considering the records obtained from the Colombian government website for the publication of public data, www.datos.gov.co. The study is focused on the perception of Cartagena as a highly recognized tourist city, which means that the increase of these types of criminal acts not only affect the city's good image as a world-class tourist destination but also affect the well-being of the native population, who are exposed daily to the risks associated with citizen insecurity.

This study proposes a replicable and comparable model, which serves as a source of information and is the basis for future research to analyze security in other cities. In addition, the study seeks to generate knowledge among citizens through the geographical presentation of the results. Accordingly, the above will establish a visual approach as an implementation of experience, considering concepts and their representation as indivisible (Fontalvo-Herrera et al., 2018).

The first part of the article provides a general overview of theft at the global and local levels to contextualize the city of Cartagena with respect to other cities in the world. Subsequently, the problem is stated with the respective research objectives. Then, the methodology is explained, detailing each of the steps carried out during the study. Next, different studies that have used the DM technique to analyze large databases on criminality are described, serving as sources of information for the present study's development (Fajardo Pascagaza, 2016; Ripoll, 2015; Soto, 2017). The approach to the research methodology is then stated, which refers to implementing the DM technique as a viable solution to analyze the behavior of car thefts. Finally, the analysis of results presents the conclusions observed and obtained through the geospatial map and the clusters derived from the research.

II. Global context

Thefts constitute a major security issue worldwide due to their high frequency and social impact, reaching the point of becoming one of the greatest social concerns since thefts impact not only on the population's assets but also the social capital, well-being, and trust of citizens. In this regard, violence in Latin America is positioned among the top five causes of death of the population, being the leading cause of death in Brazil, Colombia, Venezuela, El Salvador, and Mexico (Güitron & Guerrero, 2017).

The increase in theft worldwide is significant, and according to official indicators, 28 million families in Latin America are subject to theft or robbery every year, equivalent to 54 families being robbed per minute (Londoño & Guerrero, 1999). From this, the following findings regarding Latin America stand out: the problem of robberies has spread, and the use of violence has increased significantly.

Thus, markets for stolen goods have developed, and a percentage of thefts are not reported by the victims (Félix, Ponce, Hernández, & Llamas, 2015). As a result, theft has become an everyday occurrence globally owing to its high frequency. An alarming aspect is the use of firearms or bladed weapons to carry out robberies, which, in some situations, can lead to injuries and, in the worst circumstances, homicides. In other circumstances, Argentina has a high rate of firearm intimidation in robberies, above 70%, whereas countries such as Chile and El Salvador do not exceed 40%. Half of the known robberies committed in Argentina were carried out with violence, 39% in Mexico, 37% in Chile, 30% in Peru and Brazil, and 12% in El Salvador (Cubel, 2016). However, in Nicaragua, according to a study conducted by the Institute for Strategic Studies and Public Policy (IEEPP), robberies top the list of security problems, at a rate of 63.7%, making robberies to homes and businesses a primary problem. According to the Ministry of the Interior, along with the United Nations, the participation of young people between 18 and 25 years old in criminal activities is notorious: they are arrested for 54% of robberies with force and 67.5% of robberies with intimidation (Gutiérrez, 2013).

Spain has the third highest number of robberies in commercial establishments in the European Union, with figures higher than 10% of the EU average in robberies of commercial establishments and individuals. Losses caused by theft in this country exceed 2.4 billion euros (Ripollés, 2006). For its part, in the year of study, Guatemala had a high level of crime; in addition to being classified as an extreme risk region along with Mexico and Central America, crime cost it some 200 billion USD due to thefts and people killed because of the lack of security (Abullarade, 2016).

Another country worth highlighting is Brazil, where, according to the Public Safety Institute of Rio de Janeiro, there has been an increase in the number of street robberies and assaults, which rose by 23.7%, up

to 38,461 cases between January and April 2016. In addition, a growth of 19.7% was recorded in the number of vehicle thefts, up to 13,074 cases (Dammert & Arias, 2007). These statistics lead to the interpretation that the analyses related to the theft records have not been effective in creating policies and plans to prevent this global phenomenon because the indicators show significant upward trends; therefore, tools are required to make operational decisions for improving citizen security.

III. Local context

Cartagena is a historic city characterized by its colonial streets, full of magic and color. This has positioned it as a privileged destination for tourism, culture, world-class business events, and foreign investment.

According to the Ministry of Commerce, Industry, and Tourism, from January to May 2016, the tourism sector in the city of Cartagena recorded the arrival of 161,156 passengers owing to the entry of 109 cruise ships, 12.37% higher than the same period in the previous year. In addition, the number of travelers who reported Bolivar as their main destination grew from 131,008 to 162,423, making it the department with the second highest tourism demand in the country by non-residents after the city of Bogotá. Nevertheless, the city continues to struggle for progress and sustainable development, which has been obstructed by the daily criminal acts committed; such acts do not generate an adequate perception of security for the establishment and prosperity of local and foreign investment.

This is why reducing these acts should be a priority for local authorities, particularly considering that thefts sometimes not only generate material damage but also escalate to homicides, which are caused mostly by the resistance that is put up by the inhabitants to safeguard their personal property.

IV. Methodology

This research article presents a method for characterizing theft-related crimes in the city of Cartagena, supported by the development of DM techniques. This study employs primary information provided by the government. The central development of the study is based on an unsupervised data learning model, implementing principal component analysis techniques and the clustering method using the data obtained. The study focused on developing two stages: reducing the dimensions of the variables studied and creating theft scenarios. The dimension reduction process creates a set of new uncorrelated variables, allowing for a visual interpretation of the results. Moreover, as a further contribution, a method for characterizing thefts was provided to identify characteristics associated with the respective weights of each variable by the group found.

This study has a quantitative approach and uses data collection to test a hypothesis based on numerical measurement and statistical data analysis to establish behavior patterns and test theories (Oates, 2005). We seek to use this type of research to objectively explain and predict the social problems related to car thefts in Cartagena in 2015 and 2016, searching for consistency and causal relationships within the data.

In addition, the purpose of the present research is to analyze a social problem from a quantitative point of view, implementing the inductive method for the study and conducting an analysis of the database provided by the government. Therefore, the starting point is a theory or approach to the problem based on the observation of data on the criminal acts that have occurred in the city to subsequently establish proposals through the inferences obtained from the data, which will ultimately be validated. To this end, the DM technique will be implemented, allowing us to extract substantial information from the dataset to transform it into an understandable structure.

Crime characterization methodology

Problem analysis: This phase aims to comprehensively examine the problem in question, which refers to various types of thefts in Cartagena in 2015 and 2016. Thus, the initial procedure was to obtain information on the problem, identify the occurrences in detail, and determine their possible causes. For this reason, a study of the global and local contexts and the city's current situation in terms of theft-related crime levels was carried out.

From this, it became clear that although there are records on the different types of crimes in Cartagena that entail criminality, the population does not always effectively report situations that affect them and represent a threat to their security to the authorities, creating operational restrictions for the proper functioning of the police force to ensure security and a high quality of life in the city. In addition, after a

document review, it was found that several types of theft, such as motorcycle and cell phone theft, have increased year over year.

Data: The data used for this study were provided by a Colombian government database dedicated exclusively to providing this information for research, developing applications, and creating visualizations and stories. Thus, we focused on searching for records related to crimes, specifically in four modalities: thefts, theft of automobiles, theft from residences, theft of motorcycles, and theft from businesses, focusing on 2015 and 2016. It should be noted that this study is limited to the data obtained from the government, which only represent reported crimes without considering cases with no formal record of occurrence.

Data purging: At this stage, a systematic process is implemented to clean the data to be processed for the research to ensure that there are no errors and that the data are accurate.

It should be noted that all the above steps are focused on ensuring appropriate management and data analysis, with the objective of representing reality with accurate and consistent information by detecting and correcting incorrect data in a timely manner in order to characterize the patterns of car thefts more precisely.

Clustering: This stage is characterized by grouping a set of data without using predefined classes based on the similarities in the attribute values of the different data items. Unlike classification, this association is performed unsupervised since the classes of the training dataset are not known beforehand.

Modeling: To create the code, the R Studio programming language is used as a tool for processing and analyzing the data in this study by generating reports supported by graphs.

Knowledge generation: Based on the analysis and conclusions drawn from the research, geographic profiles of vehicle thefts are created, representing crime trends in a structured manner.

Literature review

Several studies have proposed concepts related to the classification, analysis, and strategies for analyzing problems associated with crime in urban environments. In the works analyzed in this study, the different lines of thought can be categorized according to the methodological approach used. These conceptual lines are as follows.

Table 1 shows conceptual developments based on geospatial analysis techniques: these studies analyze the determining factors in the occurrence of criminal acts using spatial variables that are associated with the place of occurrence of the acts.

TABLE 1. SUMMARY OF GEOSPATIAL CRIME INVESTIGATIONS

Investigation	What did it study?	Technique
Geographical analysis of crime and crime maps (Vázquez González & Soto Urpina, 2013)	A tool to analyze urban crime based on digital mapping was presented, analyzing the geographical aspects of crime and identifying configurations and trends in data.	Digital cartography
Vehicle theft and its spatial relationship with the sociodemographic context in three central delegations of Mexico City (Sánchez Salinas & Fuentes Flores, 2016)	The spatial patterns of vehicle theft crime in Mexico City were analyzed, finding four variables, namely, nonresidential land use, population density, female-headed households, and concentration of young population, as factors influencing vehicle theft.	Multiple linear regression model and exploratory spatial data analysis, specifically the Moran index
Modeling and prediction of crime hotspots based on probabilistic models (Flores Arias, 2014)	An empirical method to automatically generate spatial probabilistic risk functions and a mechanism for characterizing the temporal evolution of crime hotspots	Probabilistic models

Model to characterize crime in Bogotá, applying spatial data mining techniques (Peña Suárez, 2017)	A model to characterize crime behavior by applying spatial data mining techniques to find trends and behavioral variations in the data	Spatial data mining
Crime prediction using hidden Markov models (Obrecht Ihl, 2014)	How to prevent and analyze crime based on situational forecasting using the characteristics of the place and circumstances surrounding the occurrence of a crime to explain the observed patterns	Markov models
Generation of the crime map of the municipality of Zipaquirá using geostatistical and GIS tools (Rodríguez Rojas, 2017)	Crime in the municipality of Zipaquirá was analyzed by applying geostatistical and spatial analysis techniques, generating crime maps, and showing the spatial relationships of the variables.	Geostatistical tools

Source: created by the authors

TABLE 2. SUMMARY OF STATISTICAL RESEARCH ON CRIME

Article	What did it study?	Technique
Analysis of the perception of theft from people between 2012 and 2015 in Bogotá (Rojas López & Ramírez, 2017)	A quantitative analysis of citizens' perception of theft from persons in Bogotá was conducted in 2012 and 2015, analyzing the number of times that an individual was a victim of theft.	Logit-probit model
Crime and economic activity in Mexico (Félix et al., 2015)	The relationship between crime and economic activity in Mexico was explored by calculating cointegration and common cycle equations to assess whether the series shares a trend.	Econometrics and time series
Security/insecurity assessment. An exploratory study in an urban community (Sani & Nunes, 2016)	An exploratory study was developed for a community with high crime rates in Porto using a questionnaire.	Questionnaire
Determination of the places with the highest incidence of crime and violence in the Metropolitan District of Quito based on spatial statistical techniques (Vizuet Galeas, 2013)	Delimitation of the criminal phenomenon in time, space, and crime typologies that afflict the city of Quito to determine the so-called "hot spots" of crime	Kriging, a form of Bayesian inference

Source: made by the authors

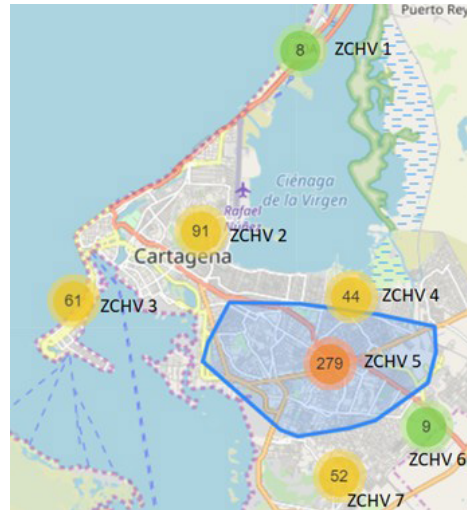
By contrast, Table 2 shows conceptual developments based on statistical analysis: these studies integrate case analysis methodologies and surveys to provide insights on the factors associated with determining causal relationships for the occurrence of criminal events.

V. Results

Car theft profiles

In the hierarchical cluster analysis, the partitioning around medoids or PAM algorithm was implemented, where a central observation of the data is related to the object that minimizes the sum of the average dissimilarity between members of the same group (De-La-Hoz, De-La-Hoz, & Fontalvo, 2019). The representative members of each cluster were obtained based on the medoids. Consequently, the analysis of the geospatial plane in two dimensions made it possible to determine crime profiles through characteristic vehicle theft zones (ZCHV) (see Figure 1), as described below.

Figure 1. Geographic cluster of automobile theft



ZCHV 1 represents the access zone to the city through the road that connects the city of Barranquilla. The analysis shows that during the study period, there were eight vehicle thefts, all using firearms and carried out between 10 pm and 6 am. Regarding the type of stolen goods, five were pickup trucks and three were vehicles.

ZCHV 2 represents the area of the city's historical center, where some of the most visited emblematic sites by tourists are located. There were 91 thefts, where master keys (79%) comprised the main tool used for committing the crime, and alternative tools such as scopolamine (9%) were also dominant. No specific pattern was identified in the time slot variable; however, 61% of the thefts occurred between 6 pm and 5 am. It is important to clarify that motorcycles are prohibited in this zone, which explains why 93% of the thefts involve motor vehicles.

ZCHV 3 represents the hotel zone in the coastal sector of the city, where large hotel chains are located, and is one of the areas with the highest value per square meter in the city. There were 61 thefts in this zone, and master keys (88%) were again the main tool used for committing the crime. The analysis of the days of occurrence shows that 73% of the thefts took place on weekends and 69% of the stolen vehicles corresponded to models older than 2013. This zone has a 24-hour restriction on motorcycles.

ZCHV 4 represents the area that is adjacent to the Ciénaga de la Virgen sector, which is characterized mainly as residential area of strata 2 and 3. There were 44 vehicle thefts in this area during the study period, with firearms being the most used theft tool (65%), followed by bladed weapons (20%). Of the 44 thefts committed, 40 involved motorcycles.

ZCHV 5 represents the public market area of the city, which is characterized by the intensive presence of businesses in general, congested traffic due to trespassing on public roads, and the logistical work of cargo vehicles, including neighborhoods such as Santa Lucia, La Esperanza, El Prado, Calamares, Zaragocilla, Escallón Villa, Chinquirá, and Las Gaviotas, among other neighboring sectors. It is the area with the highest incidence of vehicle theft: 284 in total. Of these, 60% corresponded to motorcycles, 23% to automobiles, and 17% to trucks. The area's specific characteristics include the high use of firearms (61%) as well as blunt and sharp objects (27%). An important fact to note is the high frequency of events on weekdays.

ZCHV 6 represents the border area with the municipality of Turbaco, a sector dominated by industrial and logistic warehouses in the Ternera sector, with a high concentration of second- and third-class housing. In this area, the incidence of robberies is low and 100% of the robberies involve motorcycles under the intimidation of firearms.

ZCHV 7 represents a densely populated sector of the city, characterized by subnormal neighborhoods with little infrastructure for public roads, transportation, and urban furniture. This is the area with the second highest number of thefts, 65% of which involve motorcycles, with knife and knife intimidation being the most frequent mode of theft.

VI. Conclusion

As the study was being conducted, we were able to identify key variables that affect the perception of security in the city, such as the time of day, specific area of the robbery, day of the week, and type of weapon used.

Through a variety of approaches, we provide the scientific community with a framework for studying a city's vehicle theft problem using DM. In the operational aspects, the results obtained support for a sophisticated interpretation of the theft problem in Cartagena by presenting the crime clusters in the form of characteristic vehicle theft zones or ZCHVs, based on which strategic and operational planning processes can be developed in consideration of the specific characteristics of each zone, the stolen object, the weapon used to execute the theft, and other aspects.

The above framework is proposed in the first instance for the generation of informative reports to optimize the use of security resources, recognizing that such resources are scarce in Colombian cities.

VII. References

- ACalderón Abullarade, J. (2016). *Criminalidad y emprendimiento en Guatemala: 2016* [Crime and entrepreneurship in Guatemala: 2016]. https://gem.ufm.edu/wp-content/uploads/2020/11/2016_Criminalidad-y-Emprendimiento.pdf
- Cubel, F. M. (2016). *El estado de la seguridad en América Latina 2015* [The state of security in Latin America 2015]. https://www.ieee.es/Galerias/fichero/docs_opinion/2016/DIEEO01-2016_Seguridad_AmericaLatina_FernandoMartinCubel.pdf
- Dammert, L., & Arias, P. (2007). *El desafío de la delincuencia en América Latina: diagnóstico y respuestas de política* [The challenge of crime in Latin America: diagnosis and policy responses]. In L. Dammert & L. Zúñiga (Eds.), *Seguridad y violencia: desafíos para la ciudadanía* (1a ed., Vol. 1, pp. 21–66). FLACSO.
- De-La-Hoz, E. J., De-La-Hoz, E. J., & Fontalvo, T. J. (2019). *Metodología de aprendizaje automático para la clasificación y predicción de usuarios en ambientes virtuales de educación* [Machine learning methodology for the classification and prediction of users in virtual education environments]. *Información tecnológica*, 30(1), 247–254. <https://doi.org/10.4067/S0718-07642019000100247>
- Fajardo Pascagaza, E. (2016). *Hacia la caracterización de los valores democráticos y ciudadanos de las niñas y niños escolares: una mirada desde la filosofía para niños* [Towards the characterization of democratic and citizenship values of school children: a look from philosophy for children]. *Amauta*, 14(27), 71–86. <https://doi.org/10.15648/am.27.2016.7>
- Félix, J. Q., Ponce, R. A. C., Hernández, J. M. O., & Llamas, R. V. (2015). *Delincuencia y actividad económica en México* [Crime and economic activity in Mexico]. *Norteamérica*, 10(2), 187–209.
- Flores Arias, P. (2014). *Modelación y predicción de focos de criminalidad basado en modelos probabilísticos* [Modeling and prediction of crime hotspots based on probabilistic models] [Universidad de Chile]. <https://repositorio.uchile.cl/handle/2250/129832>
- Fontalvo-Herrera, T. J., Delahoz, E. J., Mendoza-Mendoza, A. A., Fontalvo-Herrera, T. J., Delahoz, E. J., & Mendoza-Mendoza, A. A. (2018). *Aplicación de minería de datos para la clasificación de programas universitarios de ingeniería industrial acreditados en alta calidad en Colombia* [Application of data mining for the classification of high quality accredited industrial engineering university programs in Colombia]. *Información tecnológica*, 29(3), 89–96. <https://doi.org/10.4067/S0718-07642018000300089>
- Güitron, A. C. M., & Guerrero, L. R. V. (2017). *Inseguridad pública y miedo al delito, un análisis de las principales perspectivas teóricas y metodológicas para su estudio* [Public insecurity and fear of crime, an analysis of the main theoretical and methodological perspectives for its study]. *Letras Jurídicas*, 6(6). <http://letrasjuridicas.cuci.udg.mx/index.php/letrasjuridicas/article/view/63>
- Gutiérrez, J. (2013). *Violencia juvenil urbana: Estrategias de políticas públicas en Nicaragua y El Salvador* [Urban youth violence: public policy strategies in Nicaragua and El Salvador.]. <https://repositorio.utdt.edu/handle/utdt/1466>
- Londoño, J. L., & Guerrero, R. (1999). *Violencia en América Latina: epidemiología y costos* (Núm. 3074) [Violence in Latin America: epidemiology and costs (No. 3074)]. Inter-American Development Bank, Research Department website: <https://ideas.repec.org/p/adb/wpaper/3074.html>
- Oates, B. J. (2005). *Researching information systems and computing*. Sage.

- Obrecht Ihl, P. (2014). Predicción de crimen usando modelos de Márkov ocultos [Crime prediction using hidden Markov models] [Universidad de Chile]. <https://repositorio.uchile.cl/handle/2250/117357>
- Peña Suárez, A. (2017). Modelo para la caracterización del delito en la ciudad de Bogotá, aplicando técnicas de minería de datos espaciales [Model for the characterization of crime in the city of Bogota, applying spatial data mining techniques] [Universidad Distrital Francisco José de Caldas]. <https://repository.udistrital.edu.co/handle/11349/6519>
- Ripoll, C. A. (2015). Plenitud de la metafísica de la alteridad [Fullness of the metaphysics of alterity]. *Amauta*, 13(25), 197–206. Retrieved from <http://investigaciones.uniatlantico.edu.co/revistas/index.php/Amauta/article/view/1285>
- Ripollés, J. L. D. (2006). Algunos rasgos de la delincuencia en España a comienzos del Siglo XXI [Some features of crime in Spain at the beginning of the 21st Century]. *Revista Española de Investigación Criminológica: REIC*, (4), 1. <https://dialnet.unirioja.es/servlet/articulo?codigo=2083343>
- Rodríguez Rojas, L. A. (2017). Generación del mapa de crimen del municipio de Zipaquirá mediante herramientas geoestadísticas y SIG [Universidad Distrital Francisco José de Caldas] [Generation of the crime map of the municipality of Zipaquirá using geostatistical tools and GIS [Universidad Distrital Francisco José de Caldas]]. <https://repository.udistrital.edu.co/handle/11349/6430>
- Rojas López, D. A., & Ramírez, C. F. (2016). Análisis de la percepción del hurto a personas entre 2012 y 2015 en Bogotá [Fundación Universitaria los Fundadores] [Analysis of the perception of theft from persons between 2012 and 2015 in Bogota [Fundación Universitaria los Fundadores]]. <https://repository.libertadores.edu.co/handle/11371/1027>
- Sánchez Salinas, O. A., & Fuentes Flores, C. M. (2016). El robo de vehículos y su relación espacial con el contexto sociodemográfico en tres delegaciones centrales de la Ciudad de México (2010) [Vehicle theft and its spatial relationship with the sociodemographic context in three central delegations of Mexico City (2010)]. *Investigaciones geográficas*, (89), 107–120.
- Sani, A., & Nunes, L. M. (2016). Diagnóstico de seguridad/inseguridad. Un estudio exploratorio en una comunidad urbana [Security/insecurity assessment. An exploratory study in an urban community]. *Anuario de Psicología Jurídica*, 26(1), 102–106.
- Soto, J. (2017). Desde una ontología del lenguaje hacia una ética intercultural de la alteridad [From an ontology of language to an intercultural ethics of otherness]. *Amauta*, 15(30), 135–150. <https://doi.org/10.15648/am.30.2017.10>
- Vázquez González, C., & Soto Urpina, C. (2013). El análisis geográfico del delito y los mapas de la delincuencia [Geographical analysis of crime and crime mapping]. *Revista de Derecho Penal y Criminología*, 9, 419–448. <https://dialnet.unirioja.es/servlet/articulo?codigo=4369529>
- Vizuite Galeas, D. A. (2013). Determinación de los lugares de mayor incidencia de delitos y violencia en el Distrito Metropolitano de Quito con base en técnicas estadísticas espaciales [Determination of the places with the highest incidence of crime and violence in the Metropolitan District of Quito based on spatial statistical techniques]. *Escuela Politécnica Nacional*.