GEOVISUALIZATION ANALYSIS OF GENDER-BASED VIOLENCE IN MEXICO CITY USING 3D MAPPING APPROACH

M. Saldana-Perez 1, C. Palma 1,*, Y. Z. Contreras 1, N. Carrillo 1, M. Moreno-Ibarra 1
1 Instituto Politécnico Nacional (IPN), Centro de Investigación en Computación (CIC), Av. Juan de Dios Batiz, s/n, 07320, Mexico City, Mexico - (amagdasaldana, cpalma2020, ycontreras2020, ncarrillog2020, marcomoreno)@cic.ipn.mx

KEY WORDS: Crime analysis, Crimes against women, GIS, Heatmap, Open data, Open source.

ABSTRACT:

The study of criminal acts is a topic of great interest to society and with the increase of violence against women in the world makes working with this type of data an important subject, where it can be used to detect areas with high crime rates to help in decision making. This work uses open data sets available from government entities of Mexico City, where three sectors of interest were studied: economic, education, and transportation around the criminal investigation reports made by the government. For the processing and analysis of spatial data, QGIS was used with the help of open-source libraries, the 2D and 3D modelling was carried out to create heatmaps and detect hotspots. It was found that the Historic center of Mexico City is the most insecure and that the Cuauhtémoc district stands out as the most dangerous zone, for transport stations, the ones that had the most intersections with other stations are the ones that presented greater problems of gender violence, in the economic sector there was a high incidence in the limits of each district, contrary to the results obtained on the educational services.

1. INTRODUCTION

In Mexico, violence and crimes are relevant subjects given the fact that just in 2020 were reported 27.6 million crimes into several categories (INEGI, 2021). One of these categories, and the one this paper is based on, is gender crimes. Gender crimes can be defined as all crimes of violence against women just for the fact that they are women (Sanchez, 2020). Another definition is all the crimes and aggressions, including physical violence and psychological violence that are committed against the women, several manifestations could be considered violence like sexual, economic, and social violence (Torres, 2007).

One of the most dangerous cities in Mexico is Mexico City. In the 2020 Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) was reported the prevalence of a crime with a rate of 35,238 for every 100,000 women. For that reason, Mexico City is where more gender crimes are reported (ENVIPE, 2020).

The Mexican government provides a repository with open data, where this data is available to anyone who needs it. Open data is categorized into 12 general sectors related to the city, for example, health, economy, education, infrastructure, security, justice, etc. In the portal of Mexico City, there exists data of investigations about gender crimes reported since 2016. This data is georeferenced data, and for that reason, it is possible to do an analysis using a Geographic Information System (GIS).

The geographic information system (GIS) are systems or tools computed-based that are used to collect, store, process, analyze and visualize geographic information in an efficient way. The geographic information Science is the study of the science that investigates geographic information, and seeks ways to represent phenomena in the real world. (Longley et al., 2015). These systems allow making geographic representations or map-view representations, where these representations provide a realistic view of the events that occurred.

In this case, we propose the use of free software tools such as QGIS for doing the data analysis, modeling and geovisualization with the aim that this study can later be replicated.

This paper presents the study of violent acts made against women obtain from the public repository of Mexico City in a period of two years from 2019 to 2020 around three sectors of interest, such as: economic (convenience stores), transportation (subway and subway bus) and education (private and public schools) in Mexico City. The analysis applied implements 3D geospatial analysis techniques to see the incidence of gender-based violence, in order to help the population and organizations in decision-making to prevent crime.

This work is made up of five sections. The first section is the introduction, the second one presents the related works. In the third section, we propose the methodology, which considers both 2D and 3D geo-visualization. The fourth section shows the obtained results and finally in the last section conclusions are made.

2. RELATED WORK

In the following section, some studies related to the research proposal are described. The works are related to 3D mapping, crimes against women, and geo-visualization in 2D and 3D.

Since the incidence of crime is a topic of interest to the population, different approaches have been developed, which include the implementation of 3D mapping techniques. In a study conducted by the German city of Cologne (Wolff & Asche, 2009), the authors introduce 3D geovisualization...
techniques to produce crime maps in order to identify hotspots using a kernel density estimation (KDE). They create a GIS plug-in that calculates the minimum distance between a building and the closest crime that has occurred. Likewise, Hashim et al. (2019) applied 3D visualization to identify criminal patterns of urban crime (13 types) through hotspots by applying Emerging Hot Spot Analysis (Space-Time), they also applied Ordinary Least Squares (OLS) Regression to determine the factors that influence the patterns, the study used ArcGIS Pro 2.4 tool to perform the analysis.

Crime incidence can be seen using heatmaps to identify areas prone to violence against women, Garfias Royo et al. (2020) present a proposal that shows heatmaps derived from surveys carried out in the Corregidora in Querétaro with the purpose of locating places susceptible to violence against women and determine the causes of the incidence of gender-based violence in this area of Mexico.

The proposed method is described as low cost since it collects data through surveys carried out in the place of study with the aim to help the prevention of gender crime. The result of this work was a geospatial analysis of areas of violence in Querétaro City and concludes that there is a relation between social norms, gender structure (such as appearance, sexuality, capabilities, and religion), and gender violence. It mentions the importance of generating georeferenced data relating to gender violence since that information is important to decrease the problem.

Gender violence is not a problem that happens just in Mexico as seen in Crime against women in Chandigarh: A GIS analysis (Bhattacharyya, 2016). A study about crime against women was done with the help of surveys of gender violence victims. The surveys were applied in some cities such as Agartala, Kohima, Imphal, Shillong and Guwahati. The purpose of this work is to find the common factors of violence and use the results to make better decisions. The authors identified that some of the elements that determine gender violence are: being a woman, belonging to a certain religion, age, having a disability, being from another country, and speaking a different language. Furthermore, it was found that some of the places where women felt less safe are: public transport, streets, bus stops, and taxis.

In previous investigations, the importance of specifying the geographical areas with the highest incidence of gender violence is established, in the same way, Kahlon (2014) carry out a GIS analysis in order to identify the most susceptible spaces and reasons for gender crimes. This study makes use of the records of eleven police stations in the city of Chandigarh in India, a geospatial and spatio-temporal analysis of gender crimes was developed, resulting in that the highest incidence of gender crimes occurred in informal neighborhoods and in areas of student influx, which suggests that this criminal incidence is highly related to the socioeconomic and demographic profile of the locality.

Moreover, the use of open data is crucial for geospatial analysis, in the following works the use of open data and its importance are mentioned. In the study of Belesiotis et al. (2018) they describe the use of crowdsourced and open data as crucial for geospatial analysis, they used several heterogeneous open data sources in order to create predictions of spatial distributions of crimes and identify hotspots of crimes. Likewise, Groff and La Vigne (2002) use an open data set made by the Vancouver Police Department to predict crime. They develop a statistical model to make the predictions (break and entries for residential and commercial locations.), whereby with the help of mobile GIS predictive maps were created in order to assist patrol units to make decisions.

Similarly, a geospatial crime analysis study that employs the United Kingdom crime statistics and analyses crime trends was conducted to develop a web-based system that helps visualized heatmaps and is capable of displaying crime trends in the UK (Bonatos et al, 2013).

Finally, a study that assesses the walkability of women based on location and use of open data in New York City was performed by Gorrini et al. (2021), this analysis concentrates on the security issues that women face, the objective was to recognize the most insecure geographical areas.

3. METHODOLOGY

The present work proposes a way to analyze data from different open sources in a social, economic, and educational context to face gender-based violence. The methodology takes into consideration 2D and 3D modeling with the aim of showing a visual and realistic way the results obtained. Figure 1, shows the four main steps of the methodology: data acquisition, data preprocessing, geoprocessing, and geovisualization.

Each stage contains a series of steps to follow, for example, the first stage, which consists of data acquisition, refers to identifying the important data for the analysis to be carried out. For its part, the preprocessing stage applies data cleaning, and identification of relevant variables, among others. The third stage has two sub-stages: two-dimensional and three-dimensional modeling. Lastly, stage four presents the visualization of the maps.

3.1 Data selection

The data used was obtained from different open data sources, the gender crime investigations from Fiscalía General de Justicia (FGJ) of the portal Abierto de Datos de la Ciudad de México. Due to the fact that this data set incorporates records from 2016, only those that occurred in 2019 and 2020 were selected, since they use the same classification method and reflect the current situation, in this period 498,863 reports were registered (Datos abiertos CDMX, 2021).

The data crime repository has all the investigations of the prosecution that occurred in the period between 2019 - 2020, because of this a subset that contains only gender crimes was generated. First, we created a subset with all the crimes committed against women but, because not all violent acts against women are considered gender-based crimes they were separated according to their classification (Ramos, D., 2020). A total of 65,343 reports were set apart.

The second dataset used in this work was collected from the Directorio Estadístico Nacional de Unidades Económicas (DENUE). From this set, two subsets were selected: economic activities (convenience stores) and one focused on education that includes any educational service, either public or private (INEGI, 2022).

For the subsets of the economic and educational sectors, the codes of the activities that belong to each one were identified. The DENUE implements the North American Industry Classification System (NAICS), when obtaining the code of the sector the necessary data can be separated, for the economic sector code 46 was used, due to the large number of business
that are registered in the DENUE it was necessary to use a small portion of the data, for this purpose the subsector retail and minisupe with code 462112 was selected obtaining 3,135 records, and for the educational sector with code 61, a total of 11,249 schools were detected.

The third dataset contains the 12 subway lines and 195 subway stations of the transportation of the collective metro system of Mexico City; this dataset was also obtained from Portal de Datos Abiertos de la Ciudad de Mexico (Datos abiertos CDMX, 2021). Within the public transport system is the Public Passenger Transport Corridor System of Mexico City, also known as Metrobus, it has 7 lines and 283 stations. This set of data is also found on the portal.

For the 3D visualization and modeling of the city, it was necessary to incorporate information from the Sistema Abierto de Informacion Geografica de la Ciudad de Mexico (SIGCMX), which provides cadastral data and ground records (Datos abiertos CDMX, 2020).

3.2 Data preprocessing

Given the fact that the data collected belongs to different sources and each one has a different format, the datasets require preprocessing, this includes the cleaning and integration of the data.

In this step, we consider the most important attributes that are: investigation folder identifier, date, classification and type of crime, district, latitude and longitude. Because the files handle similar attributes such as the date of issue and start date, one of these were dropped. Additionally, void and incomplete values were removed.

3.3 Geoprocessing

To analyze and compile the geographic information the QGIS software was used, which has an open source license, this system was utilized because of its different functions and available plugins that allow the application of multiple algorithms to perform the required analyses useful for this investigation. This phase has two main components: 2D and 3D modeling. To create the modeling, the first layer a real-world representation was added, the next one was the gender-based crime layer, and lastly a layer of one of the three sectors: transportation, economic or educational.

3.3.1 2D Geospatial crime model

As a first step two-dimensional geoprocessing was carried out since the layers obtained during this process are later necessary to conduct the 3D modeling. The first layer that is added is the real-world representation, then a delimitation layer for the districts in Mexico City, next the crime layer, and from the previously selected sectors (transport, economic and educational) areas of influence (buffer) are delimited to study the incidence of gender crimes that have occurred.

The processing tools to perform crime analysis and modeling that happened in Mexico City are part of the plugins available in QGIS, first the data acquired from FGJ is loaded, then the appropriate sector of interest is identified: educational centers, economic stores, stations, and lines of the collective transport system. Taking into consideration the selected points from the area of interest we made a 300 meters (about 984.25 ft) buffer for the first two sectors and a 1-kilometer buffer for transportation (Figure 2).

In the second step, the count point algorithm is applied to the buffer created in the last step to determine how many crimes occurred around a place (in the buffer area). This point counting is later used to apply a categorization with a color ramp, where pink represents the lowest value and purple represents the highest (number of crimes).
In a similar way, the analysis carried out on the crimes that occurred around a building was carried out, firstly with the centroids algorithm, the midpoint of each construction were identified, then a buffer of 50 meters was created from the centroid obtained, like the analysis carried out for the sectors of interest, the crimes committed in the delimited area were counted, then the Join attributes by field process was applied, later this value will be employed to identify the number of crimes that occurred around a place along with the use of identifying colors according to the quantity.

The heatmap was obtained from the heatmap algorithm (kernel density estimation) the values used were radius: 0.005, 2000 rows, and 1750 columns. At the end of the process, it is necessary to prepare the layer for 3D modeling in such a way that the raster layer is converted into a vectorial layer with the polygonizer algorithm (raster to vector) in order to implement the height variable according to the density of crimes when using 3D visualization (for the general gender-based crime map).

### 3.3.2 3D Geospatial crime model

The 3D implementation allows volumetric information to be added, such as mountains, buildings, routes, and elevations. In this particular case, we can use the number of crimes as height. One of the most important advantages of representing the information in this type of cartography is the intuitive symbology it offers since it facilitates the representation and compression of the results.

In the same way as the 2D modeling, 3D modelling needs a surface layer, the Google Map Terrain Hybrid was used to get a digital elevation model (DEM) which is the base layer. This version contains both geographical relief as well as the streets and points of interest, this is important for the representation of crime zone incidence, since this form allows the consultation of various results of a specific area.

Subsequently, a three-dimensional environment of Mexico City is created with a dataset obtained from SIGDMX, to determine the areas and buildings affected by gender violence. The dataset contains 2,373,869 registers which have an attribute denominated level range, with this variable it was possible to appoint the height for each building and form the model of the city (Figure 3).

For the study of criminal cases in the transportation sector, arbitrary height was given to the subway and subway bus routes so that when the geovisualization stage was carried out, they would stand out more and the routes could be seen better on the map. The official colors of the Mexican subway were assigned to distinguish between each transport line.

The results derived from the two-dimensional analysis done in the 2D modeling step were used to create the 3D modeling patterns of the crimes against the woman, with the help of Qgis2threejs plugging.

### 3.4 Geovisualization

For the implementation, integration, and visualization of the data crimes in a three-dimensional format, we use the Qgis2threejs plugin which allows us to visualize the study area depending on the sector. With this plugin, through the generation of files, it is possible to see the maps in a web browser and it can also support file formats like gTIF. In this stage the layers made in the previous steps are used to build interactive maps where the public can consult the results obtained from the 3D modeling, in a quick and easy way, they can identify hotspots, and thereby identify the areas with more crimes in Mexico City.

### 4. RESULTS

As a part of the results, heatmaps were generated where the hotspots of gender crimes can be seen. The created maps contemplate both an overview map of Mexico City and certain areas of Mexico City to show specific cases in more detail.

In Figure 4, a general heatmap of Mexico City is presented, since this overview uses all the data as a base, it is later used to compare with the ones obtained from sectors of interest. It is shown that the district where most gender crimes happened is Cuauhtemoc as it has cases between 150-180 and 120-150. The next district is Coyoacan followed by Gustavo A. Madero and Venustiano Carranza.

The general view shown in Figure 4, uses a similar analysis presented in Figure 5 since it makes use of all the records of the crimes that occurred in a delimited range but does not apply extra analysis by sector.

![Figure 3. 3D model of Mexico City.](image-url)

![Figure 4. Gender crime heatmap of Mexico City.](image-url)
The results obtained from the study of gender crimes in the different sectors of interest can be seen in the next maps. Figure 6 corresponds to educational services that include both public and private schools that range from elementary to universities, trades schools, and others (music, art, among others), similarly to the presented 3D heatmaps, it is shown that the districts Cuauhtemoc and Coyoacan have the highest crime index, where is detect that more crimes occur in the center of each district. To identify the most dangerous places a radius of 300 meters was implemented (the width of the cylinders on the map), the height of the cylinders represents the number of crimes that occurred.

Although there are few educational services that have had more than 120 criminal reports around them, a pattern can be noticed, where when a peak appears on the map (represents the place with the most cases), around this there are other peaks, although smaller between the 60-100 crimes, which indicates that the area around these schools is unsafe.

Finally, in the result of the 3D visualization of the gender crimes in Mexico City around subway routes, the analysis presents that there exists a higher criminal incidence when there is a greater number of subway station transfers, and the influx is higher, likewise for the downtown area of Mexico City (Figure 8).

In Figure 9 the central zone of Mexico City, as well as some of the subway stations near the Historic center of Mexico City, are presented. It is evident that in this zone exists a higher index of gender violence, to be more specific, in the station called Guerrero, which has 2 intersections one with line 3 and another one with line B, there were between 150-200 crimes. At the intersection of the stations Balderas and Niños Heroes there were around 100-150 crimes.
In the north zone of the city, the index of crime rate is 0-50 (an example of this is the Politécnico station). Moreover, in Chabacano station where there is an intersection of 3 stations, there are between 100-150 reports. These stations are lines 3, 8, and 9. It can also be seen that, on the west side of the city, the last stations of lines 3, 9, 6, and 7 (Universidad, Tacubaya y El Rosario) stand out compared to previous stations, they have between 50-100 crimes compared to 0-50 reported crimes presented in the nearby stations.

Figure 9. Heatmap of gender-based crimes on subway lines in CDMX.

Figure 10 shows the crimes committed around the subway bus routes, a buffer of 1 kilometer was used. In a similar way to the results presented above, the pattern that expresses the concentration of gender crimes happening in the Historic center of Mexico City is repeated. The stations that have intersections with other subway bus routes also presented more insecurity.

Figure 10. Heatmap of gender-based crimes on subway bus lines in CDMX.

A detailed view of the downtown area of CDMX that shows as mentioned before a more detailed map presents the specific stations that have the greatest problems. The crimes in subway bus routes between Garibaldi in line 7 and Guerrero in line 1 have about 150-200 gender crimes, (Figure 11).

Figure 11. Heatmap of gender-based crimes on subway bus lines in CDMX.

5. CONCLUSION AND DISCUSSION

The use of both 2D and 3D geovisualization techniques allowed the creation of heatmaps to help specifically detect points or trends that occur in the data. By applying the study in different sectors (economic, educational, and transportation), the need to generate knowledge from the open data obtained can be verified.

The areas with the highest crime rate were successfully detected, during the investigation, it was found that, although the three sectors present the Cuauhtémoc district as the place with the most reported cases, certain patterns change according to the analysis of the crimes from a chosen sector, in the economic sector crimes tend to occur in the territorial limits of the districts, however for educational services the opposite occurs since more crimes occur in the center, on the other hand in the transport sector the results are very similar between the metro and Metrobus due to their concentration in the part of the Historic center of Mexico City.

The use of open-source tools and resources and free data allowed us to find crime patterns in Mexico City from the 3D geoprocessing and geovisualization, which can then be used for decision-making to prevent crime. The hotspot detection helped determine the areas with the highest crime rates, so the government can use this type of analysis to allocate resources to these areas in order to face and prevent gender-based crimes. Likewise, having a geovisualization of the data allows us to observe certain patterns of the events that have occurred, which would allow the government to deploy strategies to minimize such events.

6. FUTURE WORK

Although a relationship was found between the sectors of interest and the number of crimes against women that occurred according to the industry, other areas could be used to find other types of relationships or implement a multi-criteria space analysis that takes several sectors of interest simultaneously. Likewise, applying a study using the population density of women by area to identify if there is a correlation between the number of reports that have occurred and the number of people who live in a place could allow us to associate if more crimes occur in a place due to its insecurity or if the higher amounts are related to the number of inhabitants that each delegation has.
Due to the fact that the data present date and time variables, they can be later used to carry out a space-time analysis. Finally, this study applied the use of the 30 types of criminal acts against women, research according to the type of crime to generate more specific results than those presented can be conducted.

ACKNOWLEDGEMENTS

The work was done with partial support from the Mexican Government through the grants 1083730, 1083728, 1084083 of CONACYT. The authors also are grateful to Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, through the projects 20221894, 20221469 and grants B200432, B200500, B200501.

We would like to thank the reviewers of this work for taking the time and effort to read this paper, we appreciate the feedback and opinions made to improve the work.

REFERENCES


This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper. https://doi.org/10.5194/isprs-annals-X-4-W2-2022-241-2022 | © Author(s) 2022. CC BY 4.0 License.