Institutions and Economies
Vol. 14, No. 4, October 2022, pp. 25-54 https://doi.org/10.22452/IJIE.vol14no4.2

Characterising Crime Incidence in Mexico: A Hierarchical Linkage Clustering Analysis

Ana Lorena Jiménez Preciado[a] and Cesar Gurrola Ríos[b]

Abstract: *We aim to characterise the incidence and distribution of crime in 32 states of Mexico. A hierarchical grouping with average linkage is implemented with crime information from DataMéxico, segmented by state and type of crime from 2015 to 2020. Based on the proportional number of crimes over the population of each state, through the elbow method and the average linkage with a Cophenetic correlation coefficient, we validate the number of clusters. Subsequently, a principal component analysis (PCA) is performed to identify each state's contribution to the clusters proposed. The main results reveal criminal activity can be characterised by three groups. Drug trafficking is the crime that leads the first group, which in turn generates subgroups of interrelated crimes, such as crimes against the family, sexual abuse and harassment, and falsehood, to name a few. These crimes are committed homogeneously in most of the states of the country. Correspondingly, domestic violence and theft lead clusters two and three, and present significant concentration levels since four states accumulate 62% and 55% of crime incidence respectively. The results also provide an overview of how a particular crime can trigger the presence of others.*

Keywords: Crime incidence; Crime distribution; State crime in Mexico; hierarchical linkage clustering; Principal component analysis.
JEL Classification: C31, C49, E71

[a] Instituto Politécnico Nacional. Plan de Agua Prieta 66, Plutarco Elías Calles, Ciudad de México, México, CP 11350. *Email: ajimenezp@ipn.mx*. ORCID: https://orcid.org/0000-0001-9158-0685. Article derived from project SIP 20220920.

[b] Corresponding author. Universidad Juárez del Estado de Durango. Fanny Anitúa y Priv. Loza s/n, Durango, Dgo., México, CP 34000. *Email: cgurrola@ujed.mx.* ORCID: https://orcid.org/0000-0001-5806-4670. Work done during UJED sabbatical year.

1. Introduction

The economic and social dynamics of modern societies have been seriously affected in the last 30 years, among other aspects, due to financial and health disturbances; in the first case, the economic and financial crises of the end of the last century stand out.[1] The United Nations (UN) Office on Drugs and Crime concludes, in a broad research about the impact of the 2007 subprime crisis, that economic shocks cause significant increases in crime rates in the short term (UNODC, 2019).

The health crises caused by the SARS-CoV-2 virus spread rapidly worldwide due to its high rates of infection and mortality (Sahai et al., 2020). There are several undesirable effects due to the containment and restriction measures in economic activity and social mobility implemented by most countries in response to the COVID-19 pandemic decreed by the World Health Organization (WHO) authorities in March 2021. Besides more than five million deaths and 268 million infections confirmed, the COVID-19 crisis reduced confidence in financial markets due to the negative expectations of prolonged lockdowns and slow economic recovery (Daehler et al., 2021). The health crisis affected the global capital market's performance, causing losses of approximately 30% (Ali et al., 2020), with job losses, increased levels of poverty, damage to supply chains, even significant adverse effects on education and learning, among other manifestations of human behaviour (Ekinci, 2021). Although the COVID-19 pandemic affected all countries, its results were more significant in emerging economies (Cakmakli et al., 2020), especially in Latin America, because it was more susceptible to receiving shocks from abroad (De Salles, 2021).

There is no consensus on the effects of the COVID-19 pandemic on criminal activity. Empirical evidence is still scarce, and the discussion is under debate. Saucedo and Berry (2019) and Vilalta (2020) find a significant increase in crime rates as a direct consequence of the decrease in productive activity, increase in unemployment, social confinement, and the consequent levels of leisure, especially among the younger population. Abrams (2021) finds inconclusive results: he assures that crimes associated with drugs and home burglary show a decrease, but this is not the case for those related to non-residential robberies, homicides, or shootings. Likewise, the study highlights the importance of considering the size of the city analysed.

Estévez-Soto (2021) explores the behaviour pattern of criminal activity – grouped into seven categories – in Mexico's principal city. The results show a significant decrease in most categories in the months after the pandemic. However, the author suggests taking the results cautiously, since the reduction in criminal activity could be merely the effects of the decrease in social mobility and the consequent difficulties in reporting crimes. This may well explain that the incidence of some crimes has increased due to the COVID-19 pandemic, for example, those related to cyberattacks, fraud, and online extortion (Kemp et al., 2021). Although some previous studies show a significant reduction in the incidence of criminal activity due to social confinement policies, recent results warn that the coronavirus may increase crime levels in Latin America and the Caribbean region (Schargrodsky & Freira, 2021).

The increase in criminal activity triggers multiple costs held by a) governments, such as spending on judicial, police, and penitentiary systems; b) individuals – private sector – security services; and c) social costs, loss of income, and decreased quality of life. However, a full assessment of criminal activity costs should also consider the economic and social opportunity cost incurred when economic agents distract from their functions and reallocate resources to tackle crime. Some experts estimate, conservatively and leaving out some complementary aspects, that the costs of crime and violence exceed 3% of gross domestic product (GDP) in the Latin American and the Caribbean region (Jaitman & Torre, 2017). A more recent study of a sample of more than 160 countries reveals that the figure exceeds, on average, 12% of GDP; in the case of Mexico, it even exceeds 14% (Iqbal et al., 2021). Crime and violence have been positioned as one of the most significant aspects that threaten growth and development in modern societies, emphasising emerging countries and, notably, in the Latin American region, where crime rates are observed as the highest in the world[2] (Imbusch et al., 2011; UNODC, 2019; Cruz & Vorobyeva, 2021).

On the other hand, several studies point out the geographical heterogeneity in the incidence of criminal activity, representing significant differences between the states or regions in a given country (Cabral et al., 2018; Saucedo & Berry, 2019; Moreno & Saucedo, 2020; Mohammed & Baiee, 2020; Abrams, 2021; Bernardo et al., 2021).

However, although these studies identify the regions with the highest levels of criminal activity based on various local characteristics, they do not deal with how particular crimes can trigger the presence of others. This study aims to fill this gap by showing the incidence and spread of criminal activity in the states of Mexico.

Through a hierarchical cluster, principal component analysis (PCA), and the classification of the DataMéxico database of crimes between 2015 and 2020, our analysis offers an overview of the types of crime committed most frequently in Mexican states. The results show that criminal activity is grouped into drug dealing, domestic violence, and theft. Likewise, we use a phylogenetic tree to visualise the propagation sequence of the criminal activity. For example, in the first group, headed by drug dealing, subgroups of interrelated illegal activities are generated, such as crimes against the family, sexual freedom, abuse, sexual harassment, and falsehood, to name a few. The second group triggers other crimes, such as injuries, threats, and property damage, while the third group, headed by theft, generates its behaviour vector. Additionally, the study reveals significant differences in the concentration of crime – the first group shows more homogeneous behaviour, while groups two and three show more crime clustering in certain states.

The next section of the document provides an overview of the relevant literature that justifies the study. Sections three and four explain the methodological aspects and discuss the results obtained. Finally, the fifth section contains the main conclusions of the study.

2. Literature Review

From a psychological perspective, criminal activity is considered a deviation from expected behaviour. However, crime has also been explained from an economic perspective. The pioneering studies (Becker, 1974; Ehrlich, 1973; Lin & Loeb, 1980) argue that committing a crime is a rational function between the possibility of being discovered versus the monetary benefit generated. In other words, the criminal's motivation is to maximise his expected utility. Besides, if individuals prone to crime consider the high probability of not being discovered, apprehended, and punished, they will have significant incentives to commit or continue committing a crime (Cortés et al., 2018). The decision to commit a crime is a balance between aversion versus tolerance to risk, so those who commit a crime are lovers of risk and seek to maximise utility. In contrast, the intensity of the punishment determines the level of risk aversion (Chalfin & McCrary, 2017).

The literature offers numerous efforts to characterise and quantify the effects of criminal activity. However, the evidence provides some critical figures; it remains a pending task, especially its impact on economic activity, transmission mechanisms, and the presence of spatial externalities, among other aspects (Saucedo & Berry, 2019). The quantifying process has been hampered by the wide range of variables disturbed by criminal activity and the impossibility of measuring some of its effects: the loss of confidence in the government institutions in charge of prosecuting the crime and imparting justice.

In the Central American region, multiple effects of crime and violence – social, physical, emotional, and economic – can be observed in the lives of individuals, companies, and governments. They are forced to allocate significant resources to police and judicial labour, instead of promoting economic activity, which undermines the credibility and legitimacy of institutions. The former explains the increase in corruption, impunity, and other vices of a similar nature with the consequent impact, on individuals who have been prey to criminals, of distrust in criminal systems and processes, even considering that the only way to achieve justice is on their own (World Bank, 2011).

Crime and violence have been increasing at alarming rates in both the type and number of illegal activities committed (Cortés et al., 2018; Cabral et al., 2018; Bansal et al., 2022). This trend has become, in recent years, a determining factor in explaining low levels of development and growth and is a direct consequence of high levels of inequality and poverty (Mohammed & Baiee, 2020; Mojsoska et al., 2021), or increase in unemployment (Soemarsono et al., 2021), especially in emerging regions like Latin America and the Caribbean (UNODC, 2019; Cruz & Vorobyeva, 2021).

The Department of Sustainable Development and the Unit for Poverty Reduction and Economic Management of the World Bank assure that a 10% reduction in violence and homicides could cause a rebound of up to 1% in economic growth per capita in the Latin America region (World Bank, 2011; Pan et al., 2012). The most recent report of the UNODC, with analyses of 202 countries, documents a strong inverse link between development and criminal activity. The countries with wide disparities in population income are, on average, four times more violent (UNODC, 2019).

There is ample evidence at a global level that there is an inverse relationship between crime and economic growth; for example, Mocetti and Rizzica (2021) and Bernardo et al. (2021) provide evidence for the Italian economy, Santos-Marquez (2021) offers evidence in the Colombian case. Addy et al. (2021) also document an inverse relationship between crime and economic growth and development, using panel data modelling for a sample of Sub-Saharan African countries. Nayebyazdi (2017) studies the relationship between criminal activity and economic growth in 27 member countries of the European Union. The author demonstrates a bidirectional causal relationship between variables through a panel vector autoregressive (VAR) model. According to the Kuznets curve, this relationship varies according to the level of growth: the greater income inequality is, the more criminal activity will have a more significant effect, indicating less growth.

In the Latin American and Caribbean region, the ways in which economic dynamics respond to criminal activity has also been studied. Mohan (2021) demonstrates, in a cross-sectional study covering 13 Caribbean countries, an inverse relationship between crime incidence and the financial performance – income – of private companies, which affects economic growth. Schargrodsky and Freira (2021) suggest a positive relationship between economic activity – inequality – and crime in the Latin American and Caribbean regions.

2.1 Effects of crime and violence in Mexico

Several studies have documented how criminal activities have decimated Mexico's economic life through different mechanisms. For example, the decrease in productivity and competitiveness of companies (González, 2014; Soria, 2017; Saucedo & Berry, 2019), effects on the labour plane – a contraction of employment and increase in inequality (Cortés et al., 2018; Moreno & Saucedo, 2020), a decrease in salary (Altindag, 2012; Velázquez & Lozano, 2019), increase in absenteeism (González, 2014). The reduction of incentives for private investment also has been identified because of spreading crimes in Mexico (Torres et al., 2015; Pan et al., 2012). Distractions of governments from their substantive functions – promoting economic activity – to prosecute crime is another negative consequence of the proliferation of crime in national life (World Bank, 2011). Verdugo-Yepes et al. (2015) confirm that a crime shock induces a fall of 0.5% in different GDP per capita Mexican states.

Empirically, the Mexican economy has documented a negative relationship between criminal activity and foreign private investment, an essential variable for the domestic economy.[4] (Ashby & Ramos, 2013). Cabral et al. (2018) confirm inverse relationships and find that the effect on foreign direct investment (FDI) in Mexico is even more significant in the most violent states in the country. This relationship is based on a deterioration in the perception of security in businesses and negatively influences investment flows, affecting relocation decisions (Torres et al., 2015).

Beyond the direct costs that crime and violence cause in the different economic agents, it is essential to recognise that criminal activity also undermines scenarios that favour investment, employment, trust, and social welfare, among other aspects. Either because of the economic cost or because of the perception of insecurity, crime has damaged the investment climate by preventing the creation of new companies and increasing the number of those that close or decide to relocate to a less violent region (Pan et al., 2012; González, 2014).

More recently, Saavedra et al. (2021) demonstrated the negative influence of crime – through business robberies in businesses and violence – on business entrepreneurship, which drives the growth and development of the economies of Mexico's states. These arguments suggest a long-term growth trend of criminal activity in social and economic dynamics, with its consequent adverse effects on economic agents. Also, reducing illegal activities would achieve various social and economic benefits. Finally, the literature highlights the need to recognise the heterogeneity in crime incidence related to the level of growth or geographic location, among other aspects.

Notwithstanding the statistical significance and applicability of the results discussed, the specialised literature still lacks studies addressing how crimes are related to each other, identifying how a specific type of crime can trigger the appearance of crime. This study aims to fill this gap by showing the incidence and spread of criminal activity in Mexican states.

3. Methodological Aspects

To analyse the categorisation, incidence, and distribution of crime in Mexican states, we use hierarchical clustering and PCA. Since there are multiple crime categories, there is a possibility of correlated crimes. In addition, it is essential to find which crimes could show higher incidences, meaning that even though all crimes threaten economic agents, these need to be classified first. Hierarchical clustering allows the grouping of similar objects into clusters, while PCA permits reducing variables by an orthogonal transformation (linear combinations of original data) into a set of linearly uncorrelated values.

3.1 Data and variables

We use DataMéxico information supported by the Mexican Economics Department[5] and Datawheel. Crime data, from 2015 to 2020, is segmented by state and crime type (including 39 categories). For a better approach, the total number of crimes is proportional to the population of each state in Mexico (32 states total). Population data is taken from the latest Population and Housing Census 2020 conducted by Mexico's Economics and Statistics Office.[6] Table 1 shows the crime categories for this study.

Table 1: Crime Type Based on DataMéxico Classification

ID	Crime type	ID	Crime type	ID	Crime type
1	Abortion	14	Falsehood	27	Other crimes against property
2	Against the environment	15	Falsification	28	Other crimes against sexual freedom and security
3	Breaking and entering	16	Femicide	29	Other crimes against society
4	Child trafficking	17	Fraud	30	Other crimes of the common law
5	Corruption of minors	18	Gender-based violence in all its different modalities to family violence	31	Other offenses against the family
6	Crimes committed by public servants	19	Homicide	32	Rapture
7	Damage to property	20	Incest	33	Sexual abuse
8	Dispossession	21	Injury	34	Sexual harassment
9	Domestic violence	22	Kidnapping	35	Simple violation
10	Electoral	23	Letting prisoners escape	36	Theft
11	Equated violation	24	Drug dealing	37	Threats
12	Extortion	25	Other crimes against life and bodily integrity	38	Trafficking
13	Failure to comply with obligations of family assistance	26	Other crimes against personal freedom	39	Trust abuse

Source: Authors' elaboration.

Even if crimes are proportional to populations, the differences between crime types may lead to clustering bias. In that sense, the variables are scaled, and the normalisation process is defined as:

$$scale = \frac{x_i - \mu}{\sigma} \tag{1}$$

Where μ is the mean of each observation, and σ is the standard deviation. Normalisation allows crime variables to be scaled and centred with $\mu = 0$ and $\sigma = 1$, allowing the exact weight of the variables when clustering the distance matrix calculation. Since all crime proportions values are continuous, we use the Euclidean distance.

$$D_{ij} = \sqrt{\sum_{k=1}^{n} \left(x_{ki} - x_{kj}\right)^2} \tag{2}$$

Where D_{ij} reflects the distance measured between pairs of crime types. To evaluate the suitability of clustering crime incidences, the Hopkins statistic is implemented. The Hopkins test finds the closest neighbour for each crime and computes their distance. Besides, a data set is simulated extracted from a uniform distribution with the same variation as the original data. Finally, the Hopkins test finds the closest neighbouring observation for each simulated observation and calculates the distance between pairs.

$$H = \frac{\sum_{k=1}^{n} y_k}{\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k} \tag{3}$$

Equation (3) represents the distance between crime pairs x_k and the simulated observations from the uniform distribution y_k. If the H value is below or near 0.5, then $\sum_{k=1}^{n} x_k$ and $\sum_{k=1}^{n} y_k$ are closer to each other, meaning that data is uniformly distributed, and the clustering process is not suitable. For instance, there is data clustering if the H statistic approaches 1 (see Table 2).

Table 2: Hopkins Statistic for Crime Data

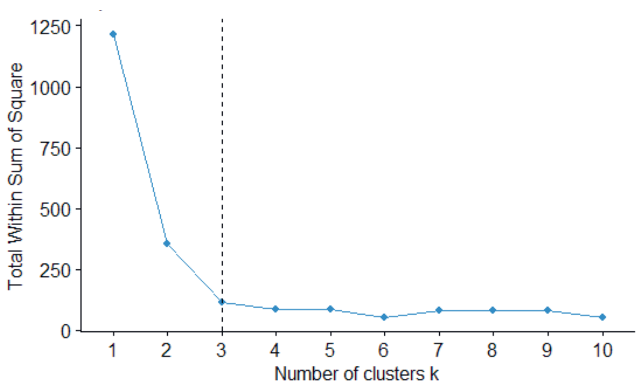
Null hypothesis	Threshold	H score
Data set follows a uniform distribution	0.5	0.8295

Source: Authors' elaboration.

3.2 Clustering selection criteria and average linking

We use the elbow method to select the number of clusters to get the optimal cluster number for crimes. The elbow method computes the sum of square error (SSE) differences among centroids (middle sets). Since data is scaled and follows a Euclidean distance, the elbow visualisation allows us to see if the crimes are similar while the test simulates random centroid iterations. Figure 1 presents the outcome.

Figure 1: Optimal Number of Clusters



Source: Authors' elaboration based on (Kassambara & Mundt, 2020) in R programming language.

Figure 1 exhibits the total within SSE and the number of clusters. When the lengths among centroids are distant from the data set, the greater the number of groups is, forming an elbow-shaped silhouette (Umargono et al., 2019). Once the distances are close, incorporating a more significant number of clusters loses relevance. In Figure 1, clusters $k = 3$ and $k = 4$ are close, compared to the SSE distance between cluster $k = 3$ and $k = 2$. For that reason, we choose three clusters for data segmentation.

Now that the number of clusters is selected, we visualise the agglomeration of crime with hierarchical trees or hierarchical clustering, also known as dendrograms, representing a two-dimensional diagram. Hierarchical clustering separates the crime data set into narrower groupings (in this case, three). In addition, we use the nearest neighbour method or average link for hierarchical tree representation, meaning that each cluster is merged through the mean distance between each pairwise crime vector:

$$D(C_i, C_j) = \frac{1}{|C_i|}\frac{1}{|C_j|}\sum_{x_1 \in C_i}\sum_{x_2 \in C_j} D(x_1, x_2) \tag{4}$$

Where D represents the distance between the number of clusters (C_i, C_j), and (x_1, x_2) are the elements contained in each cluster. Once the dendrogram has been created, it is necessary to evaluate if the average linking reflects the actual distances between observations. For instance, we use the Cophenetic coefficient, a linear correlation measure for dendrograms, and assess the height of the nodes from the actual distance matrix. The closer the value is to 1, the better the dendrogram reflects the fundamental similarity between the observations (Saraçli et al., 2013).

Following Kumar and Toshniwal (2016), Equation (5) represents the Cophenetic correlation, where d_{ij} refers to the Euclidean distance between i and j crimes, t_{ij} represents the distance between the nodes of the hierarchical cluster, and, finally, it is assumed that $\bar{d}_{ij}$ and $\bar{t}_{ij}$ are the mean values of d_{ij} and t_{ij} respectively.

$$Cophenetic = \frac{\sum_{i<j}(d_{ij} - \bar{d}_{ij})(t_{ij} - \bar{t}_{ij})}{\sqrt{\left(\sum_{i<j}(d_{ij} - \bar{d}_{ij})^2\right)\left(\sum_{i<j}(t_{ij} - \bar{t}_{ij})^2\right)}} = 0.9884 \tag{5}$$

The Cophenetic coefficient with average linking is 0.9884 representing high goodness of fit similarity between observations.

3.3 PCA analysis for state crimes contribution

Once the crime clustering and their nodes are recognised, we represent the results of the hierarchical clustering by combining them with a reduction in dimensionality with PCA, which represents a linear combination of pn-dimensional vectors with the maximum variance:

$$PCA = \sum_{j=1}^{p} a_j x_j = Xa \qquad (6)$$
$$var(Xa) = a^T Sa$$

Where X is a nxp matrix with $x_1, \ldots, x_p$ pn-dimensional vector with j-columns (variables). Linear combinations are represented by the maximum value of $var(Xa) = a^T\, Sa$ where S is the covariance matrix of crimes multiplied by $a_1, a_2, \ldots, a_p$ vector of constants. We use the traditional solution for solving the linear combination of PCA by introducing a Lagrange multiplier (λ) with unit-norm vectors, this is:

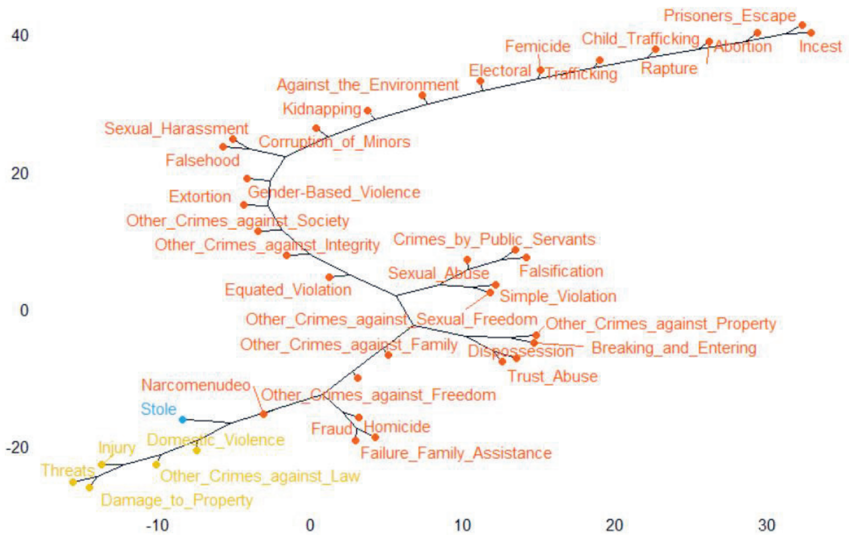
$$a^T\, Sa - \lambda\, (a^T\, a - 1) \qquad (7)$$
$$a{:}\ var\,(Xa) = a^T\, Sa = \lambda a^T\, a = \lambda$$

By differentiating Equation (7), we are looking to find $Sa - \lambda a = 0$, particularly the largest eigenvalue λ corresponding to its eigenvector a, creating new uncorrelated variables (Jolliffe & Cadima, 2016). Once the eigenvectors are obtained and ordered, we make a dimensional reduction, meaning that X_a matrix is a new dimensional subspace: the first principal component will have the most considerable variance, which explains more variance from crimes incidence.

4. Results

Figure 2 displays the hierarchical clustering using the Euclidean distance and average linkage to measure similarity (Equation 4).

Figure 2: Hierarchical Clustering (phylogenetic tree)
Euclidean distance and average linkage, k=3



Source: Authors' elaboration based on (Csárdi, 2020) in R programming language.

The results presented in Figure 2 show the interrelationships of crimes in Mexican states and groups them. As seen in the lower section of Figure 2, domestic violence heads the first group from the bottom up in the phylogenetic tree. It triggers other crimes such as injuries, which groups the crimes of threats and property damage.

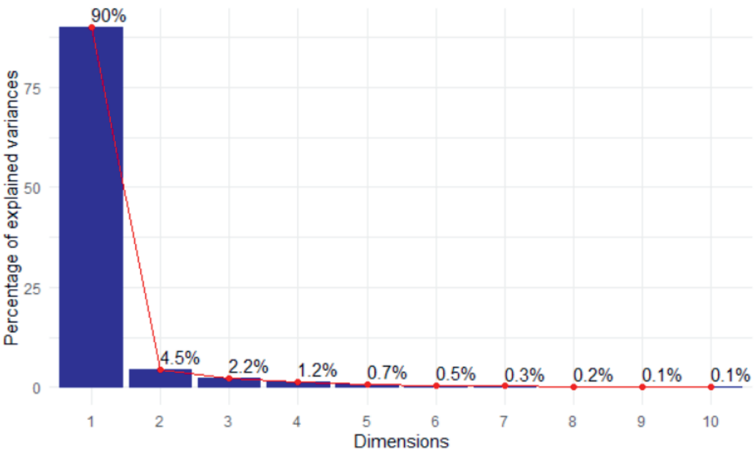
The third cluster in Figure 2 is headed by the crime of drug dealing, branching into other criminal activities, such as crimes against freedom, which has repercussions in homicides, fraud, and lack of family care. It is important to note that drug dealing has the most ramifications or subgroups. For example, crimes against the family and sexual freedom trigger two mini-groups: 1) dispossession and trust abuse; and 2) crimes against property and breaking and entering. Within this conglomerate there is another subclassification: sexual abuse and rape; they share the same ramification and crimes committed by public servants and falsification.

Although rape, crimes against integrity, society, and extortion are a subset of the large group led by drug dealing, between gender violence and corruption of minors, there is a ramification composed of falsehood and sexual harassment. In the last part of the phylogenetic tree, we find kidnapping, crimes against the environment, electoral fraud, femicide, trafficking, and trafficking of children. Finally, the last leaf on the tree is abortion, which contains two crimes: prisoner breaking and incest.

The first group refers to the one headed by drug dealing, while groups two and three are led by domestic violence and stealing, respectively. Figure 3 shows the eigenvalue variance; the first dimension is considerably more significant than the following dimensions: the highest eigenvalue identifies our principal component.

Figure 3: Eigenvalue Variance

Percentage of explained variances

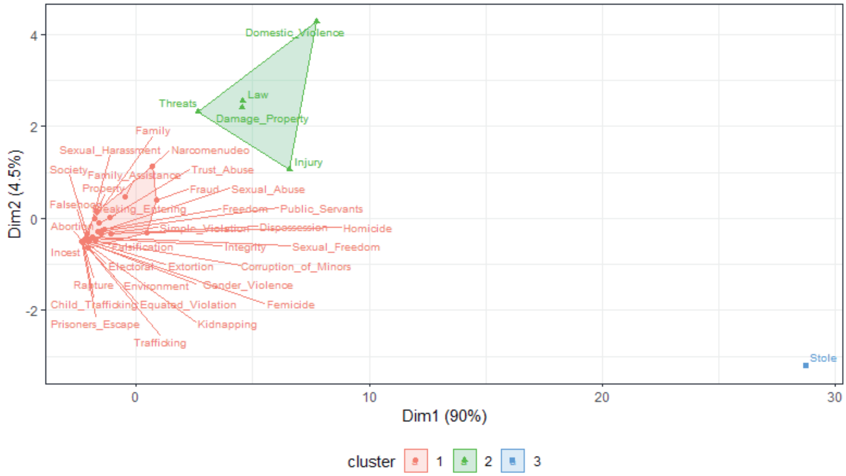


Source: Authors' elaboration based on (Kassambara and Mundt, 2020) in R programming language.

In Figure 3, the first principal component exhibits the largest possible variance in data crime, containing the 90.03% eigenvalue percentage of variance (Dimension 1), while the second component denotes the 4.5% eigenvalue percentage of variance (Dimension 2); the following components provide little information making them less significant. Recall the elbow method for optimal cluster selection; Figure 4 displays the hierarchical clustering combined with PCA solution.

Figure 4: Hierarchical Clustering + PCA Projection

Euclidean distance and average linkage, k=3



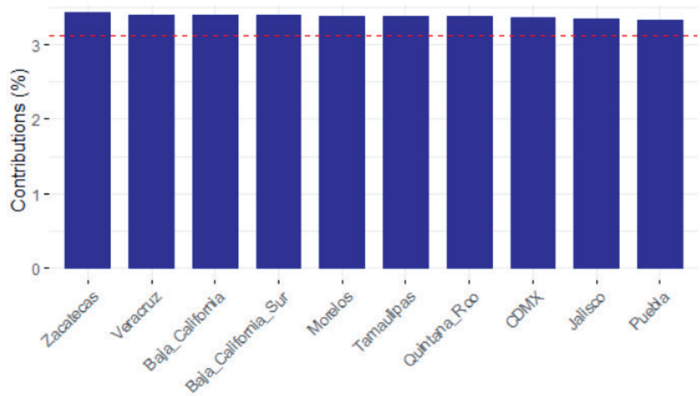
Source: Authors' elaboration based on (Kassambara & Mundt, 2020) in R programming language.

Figure 4 shows the concentration of crimes between the main component (Dimension 1) and the second component (Dimension 2) and the clusters obtained in the phylogenetic tree.[7] In this representation, the crimes that suit each cluster are exhibited again.

4.1 Cluster crimes contribution

From the PCA, it is also possible to know which states are most exposed to crimes for each component. Figure 5 shows the ten most representative states with the highest percentage of crimes in the first cluster. This is important since the crimes in Cluster 1 are held homogeneously in the Mexican states with the most significant participation. Far from being a positive aspect, it indicates that the crimes incidence of Cluster 1 spread similarly in Mexico.

Figure 5: Mexican States' Contribution to Cluster 1 (Top 10)



Source: Authors' elaboration based on (Kassambara & Mundt, 2020) in R programming language.

Figure 5 reveals the top ten states that contribute the highest percentage to Cluster 1, being Zacatecas (3.44%), Veracruz (3.41%), Baja California (3.40%), Baja California Sur (3.40%), Morelos (3.39 %), Tamaulipas (3.38%), Quintana Roo (3.38%), CDMX (3.38%), Jalisco (3.35%) and Puebla (3.34%), which have the highest incidence of crimes in the first component. It should be noted that, although Figure 5 shows only the top 10 states, most of them contribute in similar proportions in Cluster 1, indicating that crimes are scattered in most parts of the country, as shown by Table 3.

Table 3: Contribution of States to Cluster 1

Contributions (%) ordered from highest to lowest

State	Contribution	State	Contribution
Zacatecas	3.44%	Oaxaca	3.28%
Veracruz	3.41%	Queretaro	3.26%
Baja California	3.40%	Tabasco	3.25%
Baja California Sur	3.40%	Campeche	3.24%
Morelos	3.39%	Sinaloa	3.24%
Tamaulipas	3.38%	Michoacan	3.20%
Quintana Roo	3.38%	Guerrero	3.18%
CDMX	3.38%	Hidalgo	3.15%

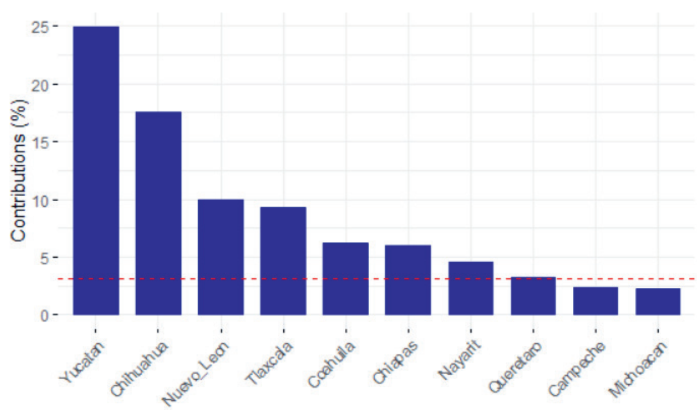
State	Contribution	State	Contribution
Jalisco	3.35%	Coahuila	3.02%
Puebla	3.34%	Guanajuato	3.00%
EDOMEX	3.33%	Tlaxcala	2.97%
Sonora	3.31%	Chiapas	2.87%
San Luis Potosi	3.31%	Nuevo León	2.51%
Durango	3.30%	Nayarit	2.49%
Aguascalientes	3.29%	Chihuahua	2.34%
Colima	3.28%	Yucatán	1.32%

Source: Authors' elaboration.

Table 3 shows that the crimes in Cluster 1 exhibit a slightly lower proportion in Tlaxcala, Chiapas, Nuevo León, Nayarit, and Chihuahua (less than 3% in all cases). In comparison, Yucatán is the only one that remains at 1.32%.

The second component contains five crimes: domestic violence, other crimes of the common law, injury, damage to property, and threats. Yucatán has the highest contribution (24.89%) (Figure 6), which is interesting since, for the first cluster, Yucatán had the lowest representation.

Figure 6: Mexican States' Contribution to Cluster 2 (Top 10)



Source: Authors' elaboration based on (Kassambara & Mundt, 2020) in R programming language.

After Yucatán, the state of Chihuahua ranks second in terms of the highest contribution of crimes of the second component with 17.55%;

Nuevo León and Tlaxcala present matching contributions (9.97% and 9.30%, respectively), Coahuila, Chiapas, Nayarit, while Queretaro, Campeche, and Michoacan are below 3%. Table 4 presents the contribution in percentage terms of all the states, highlighting Tabasco, Zacatecas, and Sonora having the lowest contributions. The CDMX (Mexico City) results stand out despite being the country's second largest state in terms of population.

Table 4: Contribution of States to Cluster 2

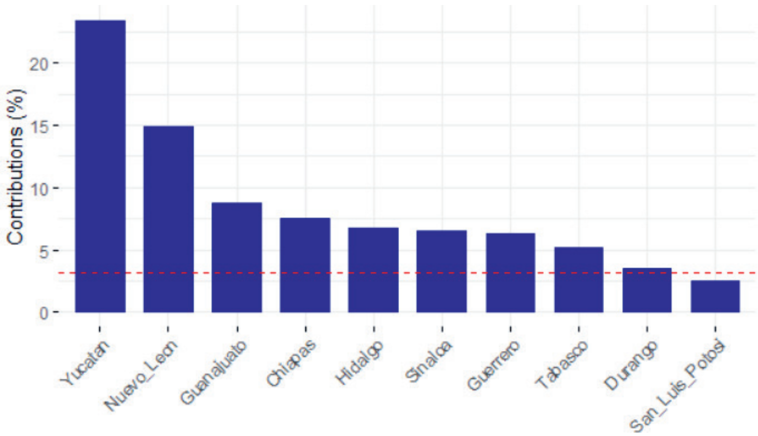
Contributions (%) ordered from highest to lowest

State	Contribution	State	Contribution
Yucatán	24.89%	Aguascalientes	0.79%
Chihuahua	17.55%	Quintana Roo	0.64%
Nuevo León	9.97%	Durango	0.59%
Tlaxcala	9.30%	Baja California	0.55%
Coahuila	6.23%	Oaxaca	0.50%
Chiapas	5.99%	Morelos	0.39%
Nayarit	4.55%	Hidalgo	0.34%
Queretaro	3.21%	Jalisco	0.26%
Campeche	2.36%	Sinaloa	0.17%
Michoacan	2.26%	Veracruz	0.17%
Guanajuato	2.04%	Colima	0.17%
EDOMEX	1.94%	Tamaulipas	0.11%
Puebla	1.78%	CDMX	0.09%
San Luis Potosi	1.09%	Tabasco	0.09%
Guerrero	1.01%	Zacatecas	0.07%
Baja California Sur	0.87%	Sonora	0.05%

Source: Authors' elaboration.

Finally, Figure 7 shows the states that present the most significant contribution to Cluster 3. Recall that this component is a single crime cluster (theft). In that sense, Yucatán once again leads the component by contributing 23.41%. Likewise, Nuevo León presents the second-highest contribution with 14.89%, and Guanajuato is in third place with 8.75% incidence. The states of Chiapas, Hidalgo, Sinaloa, Guerrero, and Tabasco contribute between 7% and 5% in Cluster 3, while Durango and San Luis Potosi do so in 3.48% and 2.48% respectively.

Figure 7: Mexican States' Contribution to Cluster 3 (Top 10)



Source: Authors' elaboration based on (Kassambara & Mundt, 2020) in R programming language.

The data referring to all contributions are shown in Table 5, including the states with the lowest incidence of Cluster 3 (stealing) being: Chihuahua, Baja California, Coahuila, Zacatecas, and Morelos (less than 0.05%).

Table 5: Contribution of States to Cluster 3

Contributions (%) ordered from highest to lowest

State	Contribution	State	Contribution
Yucatán	23.41%	Queretaro	0.97%
Nuevo León	14.89%	Sonora	0.95%
Guanajuato	8.75%	Oaxaca	0.84%
Chiapas	7.52%	Campeche	0.57%
Hidalgo	6.75%	Tlaxcala	0.52%
Sinaloa	6.47%	Jalisco	0.49%
Guerrero	6.26%	Puebla	0.34%
Tabasco	5.19%	EDOMEX	0.32%
Durango	3.48%	Veracruz	0.27%
San Luis Potosi	2.48%	CDMX	0.12%
Nayarit	2.22%	Baja California Sur	0.08%
Colima	1.88%	Chihuahua	0.05%
Michoacan	1.62%	Baja California	0.04%
Tamaulipas	1.38%	Coahuila	0.02%
Aguascalientes	1.11%	Zacatecas	0.01%
Quintana Roo	0.98%	Morelos	0.01%

Source: Authors' elaboration

The information presented in Tables 3, 4, and 5 reveals essential differences in the incidence of criminal activity among Mexico's states in each cluster. As can be seen, this information shows that the crimes contained in the first cluster are spread homogeneously. In contrast, the behaviour of the crimes in clusters two and three show significant levels of concentration. Table 4 reveals that four states – Yucatán, Chihuahua, Nuevo León, and Tlaxcala – concentrate 62% of the crimes in the cluster headed by domestic violence. In the group represented by stealing, the concentration level of crimes reaches 55% on average in Yucatán, Nuevo León, Guanajuato, and Chiapas.

4.2 Discussion of results

The result of the study identifies the most relevant crimes due to their level of incidence and the way they trigger the appearance of others in Mexican states. Likewise, the outcomes also suggest significant differences between states in terms of the presence of criminal activity, attributable to local characteristics and factors.

Our results agree with previous studies in several respects. Some studies have pointed out that theft is a determining factor in economic dynamics; for example, Saavedra et al. (2021) find significant evidence of this as an explanatory factor of the economy through the financial performance of companies. Our results highlight the importance of theft since it represents a particular vector in the national criminal configuration.

The importance in terms of the relative weight of homicides shown in our analysis corresponds to Torres et al. (2015), who relate them to the negative effects on private investment; Cabral et al. (2018) document an inverse relationship with FDI. Meanwhile, Moreno and Saucedo (2020) show that homicide and kidnapping threaten employment. In Mohammed and Baiee (2020), there is also evidence in favour of the relative importance of homicides, assault by threat, and shooting as major crimes.

Our results coincide with Saucedo and Berry (2019) regarding the relative importance of drug-related crimes. They find that crimes associated with drug trafficking, including extortion, homicide, and kidnapping, reflect a decrease in labour productivity. Our results show that extortion, homicide, and kidnapping appear because of drug trafficking (see Figure 2). The results of Bernardo et al. (2021) also coincide with our findings, by pointing out the relative importance of drug trafficking, prostitution and money laundering, extortions, and kidnapping.

Another exciting result of our analysis is the geographic heterogeneity of criminal activity; in other words, the incidence of illegal activity varies according to differences and characteristics attributable to the states. Such statement coincides with the findings of Torres et al. (2015), Cabral et al., 2018; Saavedra et al. (2021); and Bernardo et al. (2021). In particular, the contrast of our results with those of Abrams (2021) is striking. In a study carried out in 25 large cities in the United States (US), the incidence of crime differs depending on population density of localities.

5. Conclusions

Criminal activity represents a detriment to the different economic agents of society through aspects such as disbursements to protect an individual's inheritance, the deterioration of investment and employment conditions, even the loss of confidence in democratic processes and the governance of a country.

Using hierarchical clusters with average linkage, principal component analysis, and the DataMéxico crime database, the study characterised the incidence of crime in Mexican states between 2015 and 2020. The modelling results reveal that Mexico's criminal activity can be characterised into three large groups headed by drug dealing, domestic violence, and theft respectively. The first cluster trigger subgroups of interrelated felonies, such as crimes against the family and sexual freedom, sexual abuse, rape or falsehood, and sexual harassment, to name a few.

These crimes are committed proportionally in most states. In the second cluster, the crime of domestic violence has ramifications in injuries, threats, and property damage. The study shows that stealing is consolidated as a single cluster in Mexican crime activity. Another exciting result is the significant differences of crime concentration levels among Mexican states. The group represented by domestic violence shows more homogeneous behaviour. In contrast, groups two and three, characterised by drug dealing and stealing, respectively, show significant concentration levels.

We highlight that the first cluster contains the highest concentration of crimes. However, drug dealing is the main crime that triggers others: drug dealing is a spotlight for mitigating other crimes. Cluster 2, led by domestic violence, coincides with its increase during Mexico's confinement peak in 2020. Finally, the single cluster represented by theft also includes reported crimes that affect the business environment. In that sense, the major crimes in Mexico are exposed.

The study contributes to the existing literature in two aspects. First, an overview of the geographic incidence of criminal activity in Mexico is offered, recognising different regional heterogeneity. On the other hand, the analysis shows the interrelationships among the different crimes committed, allowing us to identify how a particular crime generates subgroups, triggering the presence of other crimes. Our results have potential implications for several economic agents (policymakers and the business sector) in designing strategies to confront crime and its undesirable economic consequences; likewise, for preventive measures and warnings for families and individuals.

There are plenty of studies that analyse the levels of criminal incidence based on various local characteristics; nevertheless, they do not point out how particular crimes can trigger others. This study fills this gap by showing the incidence and spread of criminal activity in Mexican states.

Further research could consider the incidence of criminal activity resulting from different crises – financial, economic, health-testing the strategies adopted to confront illegal activity based on allocated resources. Also, it allows the exploration of possible relationships with macroeconomic indicators, such as confidence indices, production, and unemployment, among other variables.

Notes

1 Among the most recent: the crisis in emerging markets – headed by Mexico in 1994 – the subprime crisis of 2007 originated in the subprime mortgage sector in the US, or the debt crisis in European countries of 2010. Calvo and Reinhart (1996), Rigobon (2002), Mendoza et al. (2011), and Santillán (2015) offer a detailed review of the financial crises of the late 20th century.

2 Considering the homicide rate and ruling out places with active warfare, the region contains the ten most dangerous cities on the planet.

[3] For example, Spain and Central America have approximately the same population, but while the European country registered 336 murders in 2006, the American region accounted for 14,257 cases (World Bank, 2011).

[4] According to the World Investment Report 2020 of the United Nations Conference on Trade and Development (UNCTAD), in 2018, Mexico was the 14th largest FDI recipient with 35 billion dollars.

[5] Secretaría de Economía (SE).

[6] Instituto Nacional de Estadística y Geografía (INEGI). Information about the population states of the Mexican Republic is shown in Appendix A.

[7] Appendix B contains the hierarchical clustering and PCA crimes classification.

References

Abrams, D. S. (2021). Covid and crime: An early empirical look. *Journal of Public Economics*, *194*(104344). https://doi.org/10.1016/j.jpubeco.2020.104344

Addy, W.O., HongXing, Y., Otchere, S.K., & Beraud, J.J. (2021). Economic development, political and socio-political violence crisis in the Sub-Saharan Africa: An economic panel study. *European Journal of Business and Management Research*, *6*(4), 34-40. https://doi.org/10.24018/ejbmr.2021.6.4.763

Ali, M., Alam, N., & Rizvi, S. (2020). Coronavirus (Covid-19) - An epidemic or pandemic for financial markets. *Journal of Behavioral and Experimental Finance*, *27*, 1-6. https://doi.org/10.1016/j.jbef.2020.100341

Altindag, D. (2012). Crime and unemployment: Evidence from Europe. *International Review of Law and Economics*, *32*, 145-157. https://doi.org/10.1016/j.irle.2011.10.003

Ashby, N., & Ramos, M. (2013). Foreign direct investment and industry response to organized crime: The Mexican case. *European Journal of Political Economy*, *30*, 80-91. https://doi.org/10.1016/j.ejpoleco.2013.01.006

Bansal, K., Arora, S., Pritam, K.S., Mathur, T., & Agarwal, S. (2022). Dynamics of crime transmission using fractional-order differential equations. *Fractals*, *31*(1), 1-16. https://doi.org/10.1142/S0218348X22500128

Becker, G. (1974). Crime and punishment: An economic approach. In G. Becker, & W. Landes (Eds.), *Essays in the Economics of Crime and Punishment* (pp. 1-54). Cambridge: National Bureau of Economic Research.

Bernardo, G., Brunetti, I., Pinar, M., & Stengos, T. (2021). Measuring the presence of organized crime across Italian provinces: A sensitivity analysis. *European Journal of Law and Economics*, *51*(1), 31-95. https://doi.org/10.1007/s10657-020-09676-0

Cabral, R., Mollick, A.V., & Saucedo, E. (2018). The impact of crime and other economic forces on Mexico's foreign direct investment inflows. *Working Paper No. 2018-24*, Mexico City: Banco de México.

Cakmakli, C., Demiralp, S., Kalemli-Ozcan, S., Yesiltas, S., & Yildirim, M. (2020). Covid-19 and emerging markets: An epidemiological model with international production networks and capital flows. *International Monetary Fund IMF Working Paper Series* (20/133), 1-75. https://ssrn.com/abstract=3670613

Calvo, S., & Reinhart, C. (1996). Capital flows to Latin America: Is there evidence of contagion effects? *Policy Research Working Paper Series 1619*. Washington: World Bank.

Chalfin, A., & McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, *55*(1), 5-48. https://doi.org/10.1257/jel.20141147

Cortés, R., Castellanos, L., & Cortés, M. (2018). Causas sociales y económicas de los delitos en México. *Boletín criminológico*, *24*(176), 1-11. https://doi.org/10.24310/Boletin-criminologico.2018.v24i0.5308

Cruz, J.M., & Vorobyeva, Y. (2021). State presence, armed actors, and criminal violence in Central America. *The Sociological Quarterly*. https://doi.org/10.1080/00380253.2021.1940349

Csárdi, G. (2020). Igraph: Network analysis and visualization. https://cran.r-project.org/web/packages/igraph/index.html

Daehler, T., Aizenman, J., & Jinjarak, Y. (2021). Emerging markets sovereign CDS spreads during Covid-19: Economics versus epidemiology news. *Economic Modelling*, *100*(105504). https://doi.org/10.1016/j.econmod.2021.105504

De Salles, A. (2021). Covid-19 pandemic initial effects on the idiosyncratic risk in Latin America. *The Mexican Journal of Economics and Finance*, *16*(3), 1-21. https://doi.org/10.21919/remef.v16i3.632

Ehrlich, I. (1973). Participation in illegitimate activities: Theoretical and empirical investigation. *Journal of Political Economy*, *81*(3), 521-565. https://doi.org/10.1086/260058

Ekinci, A. (2021). Modelling and forecasting of growth rate of new Covid-19 cases in top nine affected countries: Considering conditional variance and asymmetric effect. *Chaos, Solitons & Fractals*, *151*(111227). https://doi.org/10.1016/j.chaos.2021.111227

Estévez-Soto, P. (2021). Crime and Covid-19: Effect of changes in routine activities in Mexico City. *Crime Science*, *10*(15). https://doi.org/10.1186/s40163-021-00151

González, S. (2014). Criminalidad y crecimiento económico regional en México. *Frontera norte*, *26*(51), 75-111. https://doi.org/10.17428/rfn.v26i51.557

Imbusch, P., Misse, M., & Carrión, F. (2011). Violence research in Latin America and the Caribbean: A literature review. *International Journal of Conflict and Violence*, *5*(1), 87-154. https://doi.org/10.4119/ijcv-2851

Iqbal, M., Bardwell, H., & Hammond, D. (2021). Estimating the global economic cost of violence: Methodology improvement and estimate updates. *Defence and Peace Economics*, *32*(4), 403-426. https://doi.org/10.1080/10242694.2019.1689485

Jaitman, L., & Torre , I. (2017). Estimación de los costos directos del crimen y de la violenvia. In L. Jaitman (Ed.), Los costos del crimen y de la violencia Nueva evidencia y hallazgos en América Latina y el Caribe (p. 132). Washington, D.C.: Banco Interamericano de Desarrollo.

Jolliffe, I.T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 1-16. https://doi.org/10.1098/rsta.2015.0202

Kassambara, A., & Mundt, F. (2020). Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. https://cran.r-project.org/web/packages/factoextra/index.html

Kemp, S., Buil-Gil, D., Moneva, A., Miró-Llinares, F., & Díaz-Castaño, N. (2021). Empty streets, busy internet: A time-series analysis of cybercrime and fraud trends during Covid-19. *Journal of Contemporary Criminal Justice*, *37*(4), 480-501. https://doi.org/10.1177%2F10439862211027986

Kumar, S., & Toshniwal, D. (2016). Analysis of hourly road accident counts using hierarchical clustering and cophenetic correlation coefficient (CPCC). *Journal of Big Data*, *3*(13), 1-11. https://doi.org/10.1186/s40537-016-0046-3

Lin, V., & Loeb, P. (1980). An economic analysis of criminal activities in Mexico. *Journal of Behavioral Economics*, *9*(2), 25-39. https://doi.org/10.1016/S0090-5720(80)80002

Mendoza, A., López, F., & Watkins, K. (2011). *Reflexiones sobre las crisis financieras*. Mexico City: CIIEUPAEP-UNAM.

Mocetti, S., & Rizzica, L. (2021). La criminalità organizzata in Italia: Un'analisi economica. *Bank of Italy Occasional Paper* (661). Rome: Banca d'Italia.

Mohammed, A.F., & Baiee, W.R. (2020). Analysis of criminal spatial events in GIS for predicting hotspots. *IOP Conference Series: Materials Science and Engineering*, *928*. https://iopscience.iop.org/article/10.1088/1757-899X/928/3/032071

Mohan, P.S. (2021). Violent crime and firm performance: Evidence from the Caribbean. *International Journal of the Economics of Business*, *28*(2), 309-327. https://doi.org/10.1080/13571516.2021.1896926

Mojsoska, S., Nikolovska-Vrateovska, D., & Vrteovski, S. (2021). The economic crime, social costs and economic growth. *Thematic Conference Proceedings Of International Significance*, *11*. http://eskup.kpu.edu.rs/dar/article/view/282

Moreno, O.J., & Saucedo, E. (2020). Crime and employment destruction in Mexico: Do firms' size and location matter. *Análisis económico*, *35*(88), 9-41.

Nayebyazdi, A. (2017). The relationship between occurrence of crime and economic growth in the European Union: Panel VAR approach. *International Journal of Contemporary Economics & Administrative Sciences*, *7*(1-2), 57-80. http://ijceas.2020staging.com/index.php/ijceas/citationstylelanguage/get/ieee?submissionId=150

Pan, M., Widner, B., & Enomoto, C.E. (2012). Growth and crime in contiguous states of Mexico. *Review of Urban & Regional*

Development Studies, *24*(1-2), 51-64. https://doi.org/10.1111/j.1467-940X.2012.00188.x

Rigobon, R. (2002). *International financial contagion: Theory and evidence in evolution*. Charlottesville: The Research Foundation Publications of AIMR. https://www.cfainstitute.org/-/media/documents/book/rf-publication/2002/rf-v2002-n2-3918-pdf.ashx

Saavedra, R., Morones, A., & Martínez, G. (2021). El robo como obstáculo para el emprendimiento en México, 2005-2018. *Análisis económico*, *36*(92), 145-163. https://doi.org/10.24275/uam/azc/dcsh/ae/2021v36n92

Sahai, A., Rath, N., Sood, V., & Singh, M. (2020). ARIMA modelling & forecasting of Covid-19 in top five affected countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(5), 1419-1427. https://doi.org/10.1016/j.dsx.2020.07.042

Santillán, R. (2015). *La gran recesión (2007-2012): Lecciones y oportunidades para México*. México: EGADE Business School-Instituto Mexicano de Ejecutivos de Finanzas-PwC.

Santos-Marquez, F. (2021). Spatial beta-convergence forecasting models: Evidence from municipal homicide rates in Colombia. *Journal of Forecasting*, *41*, 294-302. https://doi.org/10.1002/for.2816

Saraçli, S., Doğan, N., & Doğan, İ. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, *203*, 1-8. https://doi.org/10.1186/1029-242X-2013-203

Saucedo, D., & Berry, B. (2019). Efecto de la violencia en la productividad laboral en México: Análisis de un modelo de regresión espacial de datos panel. *Investigaciones Geográfica*, *100*, 1-15. https://doi.org/10.14350/rig.60021

Schargrodsky, E., & Freira, L. (2021). Inequality and crime in Latin America and the Caribbean: New data for and old question. *CAF Working Paper 2021/08*. http://scioteca.caf.com/handle/123456789/1786

Soria, R. (2017). Impacto de la violencia e inseguridad en la competitividad de los estados Mexicanos. *Economía, Sociedad y Territorio*, *17*(54), 279-307. https://doi.org/10.22136/est002017802

Torres, V.H., Polanco, M., & Venegas, F. (2015). Actividad criminal e inversión privada en México: Una perspectiva espacial, 1997-2010. *Papeles de población*, *21*(83), 203-251. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-74252015000100008

Umargono, E., Suseno, J., & Gunawan, S. (2019). K-Means clustering optimization using the elbow method and early centroid determination based-on mean and median. *Proceedings of the International Conferences on Information System and Technology*, (pp. 234-240). https://doi.org/10.5220/0009908402340240

UNODC. (2019). *Global study on homicide*. Vienna: United Nations Office on Drugs and Crime.

Velázquez, D., & Lozano, R.L. (2019). El aumento del crimen y la disminución de la participación laboral en el ingreso en México. *Apuntes del Cenes*, *38*(68), 38-68. https://doi.org/10.19053/01203053.v38.n68.2019.8573

Verdugo-Yepes, C., Pedroni, P., & Hu, X. (2015). Crime and the economy in Mexican states: Heterogeneous panel estimates (1993-2012). *IMF Working Paper*. https://www.imf.org/external/pubs/ft/wp/2015/wp15121.pdf

Vilalta, C. (2020). Violence in Latin America: An overview of research and issues. *Annual Review of Sociology*, *46*(1), 693-706. https://doi.org/10.1146/annurev-soc-073018-022657

World Bank. (2011). *Crime and Violence in Central America: A Development Challenge*. Washington: World Bank.

Appendix A: Mexico's Population

State	Population	State	Population
Aguascalientes	1,425,607	Morelos	1,971,520
Baja California	3,769,020	Nayarit	1,235,456
Baja California Sur	798,447	Nuevo León	5,784,442
Campeche	928,363	Oaxaca	4,132,148
Coahuila de Zaragoza	3,146,771	Puebla	6,583,278
Colima	731,391	Queretaro	2,368,467
Chiapas	5,543,828	Quintana Roo	1,857,985
Chihuahua	3,741,869	San Luis Potosi	2,822,255
CDMX (Mexico City)	9,209,944	Sinaloa	3,026,943
Durango	1,832,650	Sonora	2,944,840
Guanajuato	6,166,934	Tabasco	2,402,598
Guerrero	3,540,685	Tamaulipas	3,527,735
Hidalgo	3,082,841	Tlaxcala	1,342,977
Jalisco	8,348,151	Veracruz	8,062,579
EDOMEX (Estado de México)	16,992,418	Yucatán	2,320,898
Michoacan	4,748,846	Zacatecas	1,622,138

Source: Authors' elaboration based on Population and Housing Census 2020.

Appendix B: Hierarchical Clustering and PCA Crimes Classification

D	Cluster 1	ID	Cluster 1
1	Abortion	25	Other crimes against life and bodily integrity
2	Against the environment	26	Other crimes against personal freedom
3	Breaking and entering	27	Other crimes against property
4	Child trafficking	28	Other crimes against sexual freedom and security
5	Corruption of minors	29	Other crimes against society
6	Crimes committed by public servants	31	Other offenses against the family
8	Dispossession	32	Rape
10	Electoral	33	Sexual abuse
11	Equated violation	34	Sexual harassment
12	Extortion	35	Simple violation
13	Failure to comply with obligations of family assistance	38	Trafficking
14	Falsehood	39	Trust abuse
15	Falsification	**ID**	**Cluster 2**
16	Femicide	9	Domestic violence
17	Fraud	30	Other crimes of the common law
18	Gender-based violence in all its different modalities to family violence	21	Injury
19	Homicide	7	Damage to property
20	Incest	37	Threats
22	Kidnapping	**ID**	**Cluster 3**
23	Prison escape	36	Theft
24	Drug dealing		

Source: Authors' elaboration based on elbow method and hierarchical clustering with average linking.