UNIVERSIDAD PANAMERICANA
Campus México

UNIVERSIDAD PANAMERICANA

Subject: Marketing Analysis and Design Management(COM145)

Professor: Sarahí Aguilar González

Delivery date: May 24

Cycle: 1222

Project name: Crime rate in the states of the Mexican Republic

Team Members		
ID	**Name**	**Career**
0215397	Mauricio Torres Minguez	Ing TI
0213360	Gilberto Guzmán González	Ing TI
0220791	Carlos Manuel Villalobos Santiago	Ing TI

Headings				
ID	**2-social**		**7-knowledge**	
	D	**C**	**A**	**JI**

Abstract

Over the past few decades, crime has been increasing in Mexican territory, reaching worrying values and becoming news not only in the country, but throughout the world. The project aims to be able to analyze the historical data of the crimes that have been registered in the country, as well as to offer in detail the relationships that the variables have with the crime rate, finally, it seeks to obtain a prediction of said rate to the first month of the year 2019, using the historical data of the year 2018.

Introduction

Among the events that damage, dissociate or violate laws and/or regulations in communities, there is a special class called criminality, which is usually conceived as that which has more serious consequences. (Quiroga, 89)

Because of the reports and the greater diffusion that the media have due to social networks, it is notorious that crime in Mexican territory is sadly an everyday occurrence and has become normalized in practically the entire country. Year after year figures and statistics have been released that are worse than the previous year, so it is practically a fact that violence and crime have been on the rise, at least so far this century.

Thanks to this information and the figures that are public, it was decided to look for social patterns in the states, to find out if there is a correlation between the crimes in the region with respect to deficiencies in the human composition, the data of the 2019 referring to the values of 2018, with the aim of being able to train the model for future predictions, this will make it easier to find key points to reduce the crime rate in principle to later carry out social programs to make the minimum crime in the states.

Development

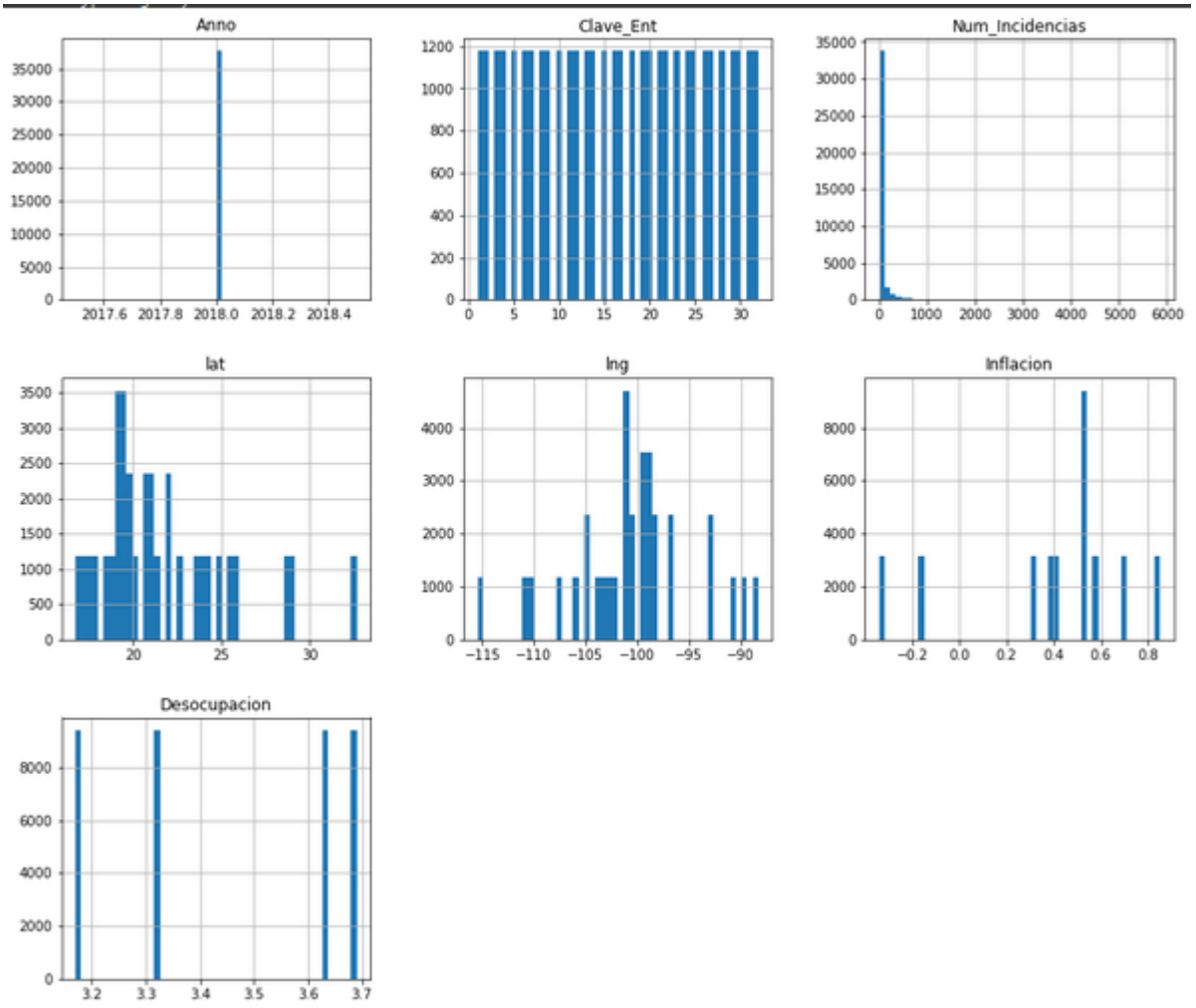
For the development of the project, the following x open data sets were used:

1. Data set with the number of historical incidents in the Mexican Republic, provided by INEGI.
2. Geographic data set of the Mexican Republic, provided by INEGI
3. Data set with inflation rates by state per month
4. Data set with unemployment rates by state per quarter

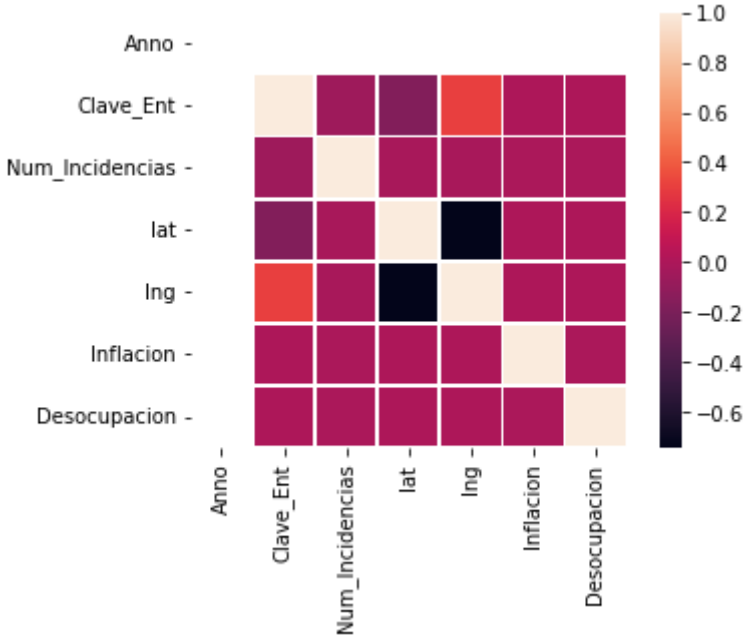
After obtaining all the necessary data to find the solution, it was loaded all this information to the development environment that was initially a google colab notebook.

Some type of processing or cleaning was done to all the data sets, this to improve the quality of the results that were going to be obtained and to be able to continue with the development of the model, the operations that were carried out were melts, merge, drops of columns and joins.

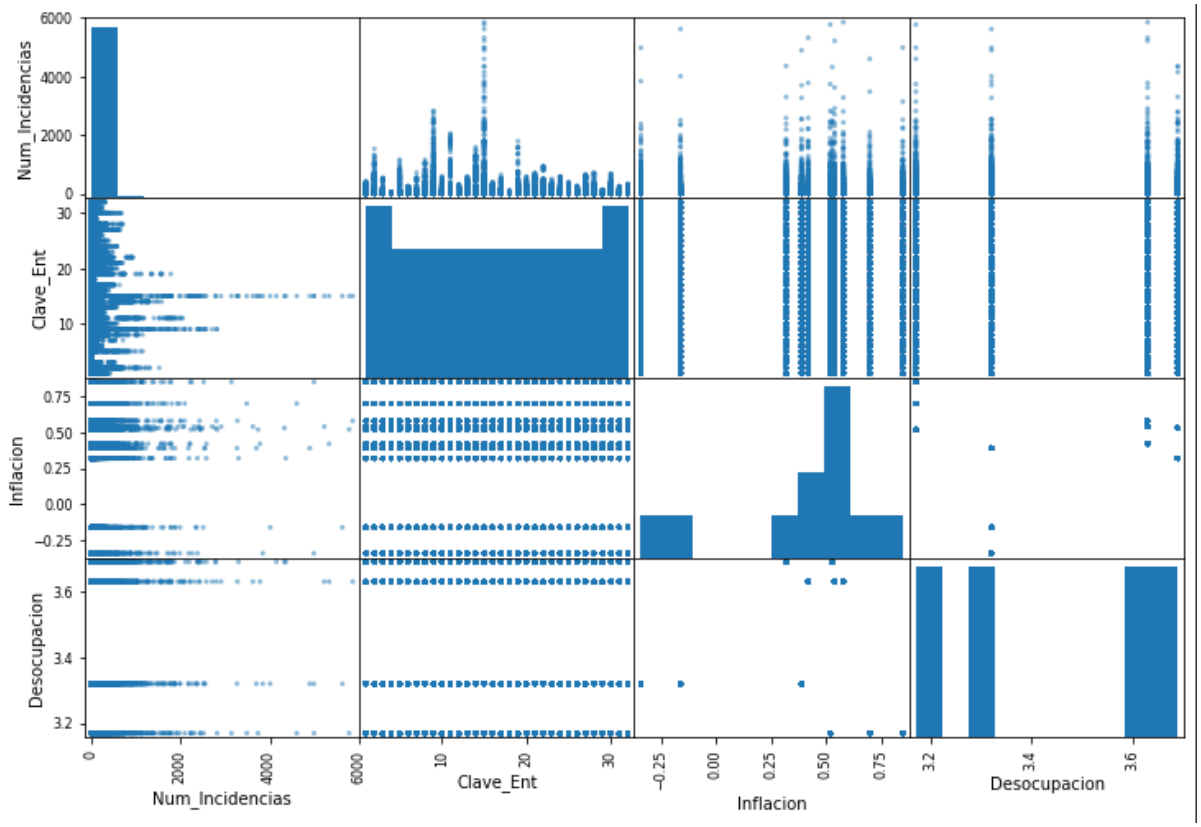
Once all this information was ready to be displayed and analyzed, the following graphs and tables were generated:

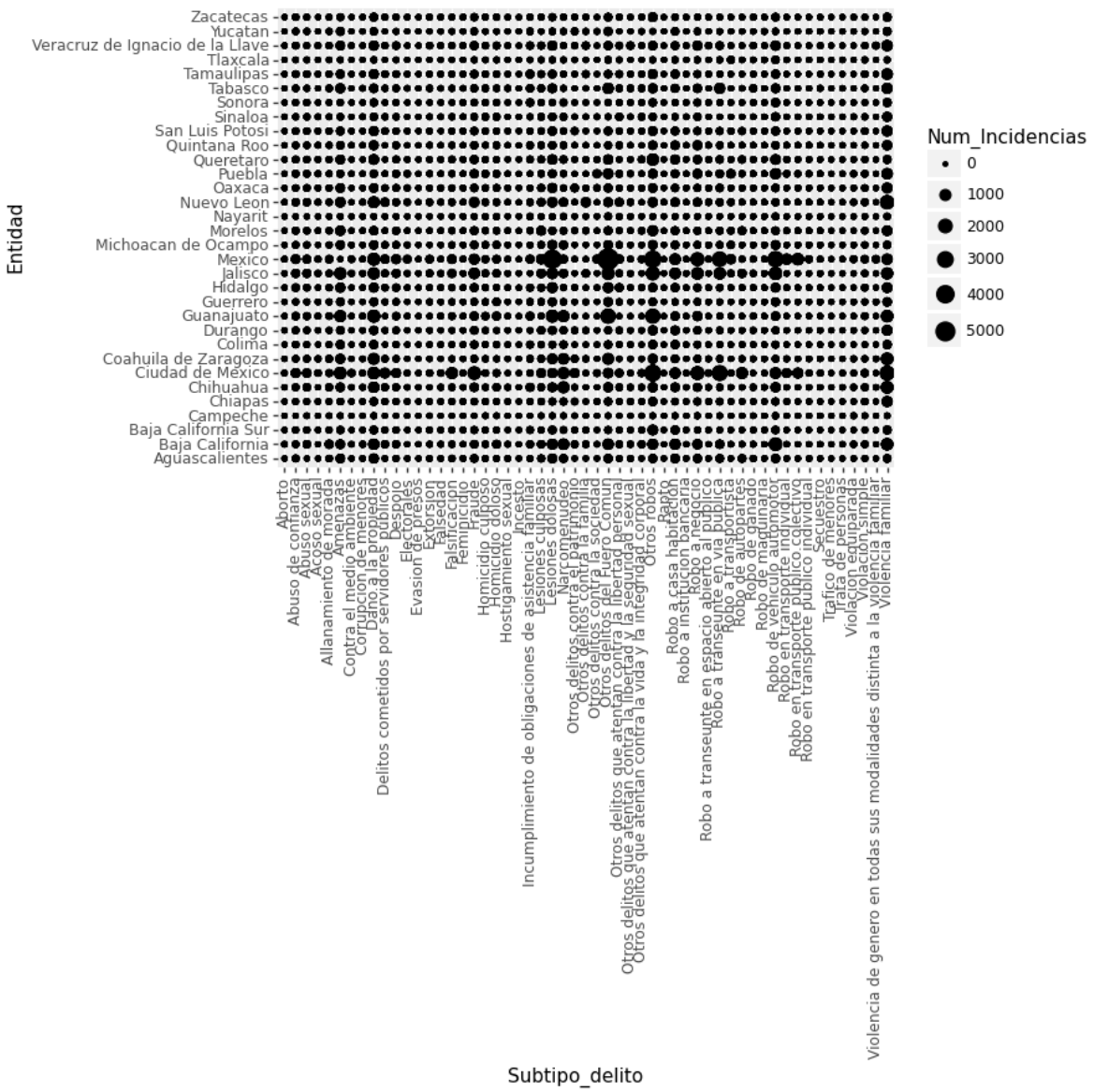


For the initial exploratory analysis, a histogram of our main table was generated to be able to visualize the initial values that are available.

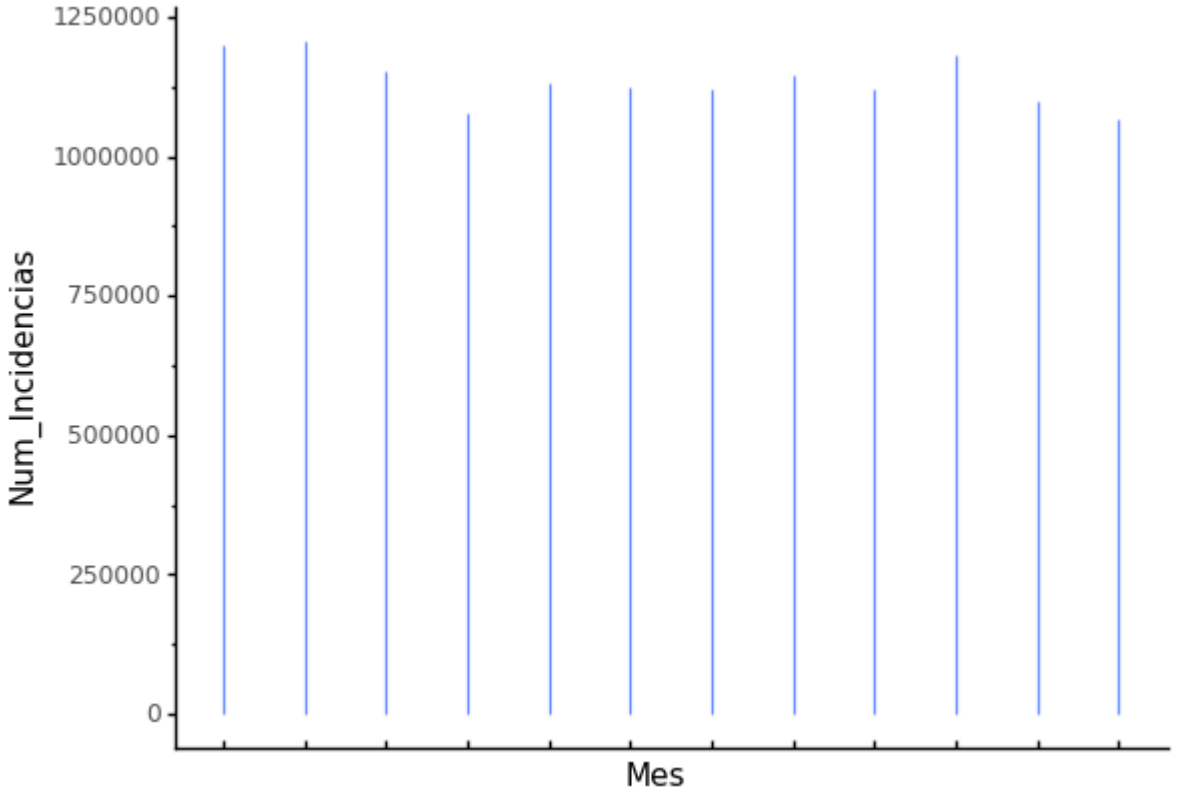


Subsequently, a heat map was generated in which we can observe the relationships that some variables in our table have with other variables. Remembering that what we want is to have dark colors in the graph, we cannot obtain great insight about the effect of some variables on others.



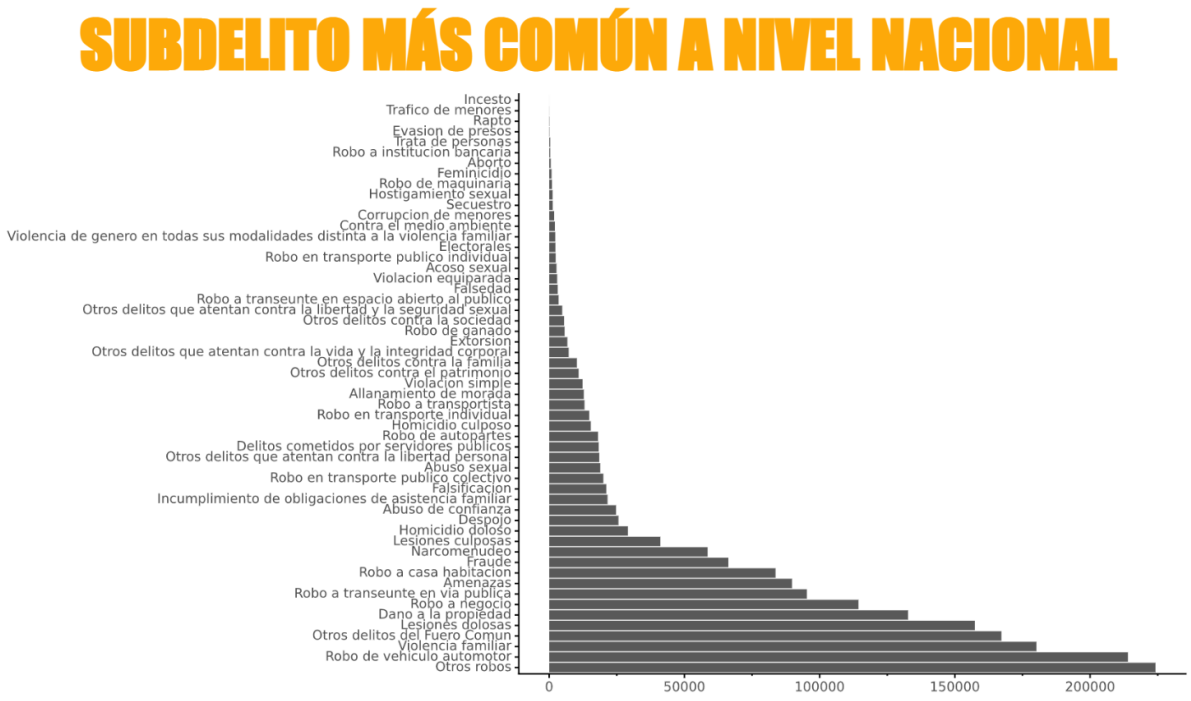


In this visualization it can be seen that the states of Mexico and Mexico City are the ones with the most accumulated crimes, states such as Jalisco and Baja California also have higher values than the rest of the country. This does not mean that the other states are safer, the insight that can be drawn from this graph is that the states with more incidents make the crime rate seem lower in the rest of the states, but in reality in all of them there is a high criminality index in all of them.



In this visualization we can see the accumulated incidences in each of the months of the year 2018, we can take as an important note the fact that the incidences have never been below one million units.

The first analysis that we can carry out on this graph is the number of incidents by state since those that are more inhabited are where more crimes take place, finding some exceptions where this does not happen in particular, because the values are not normalized, An objective result is not found since states with a small population show a low rate but it does not necessarily mean that they have a low crime rate.



As a second analysis, the most common sub-crimes can be observed, the most prominent being the category of other robberies, this is very general and does not tell us much, the next two positions are vehicle theft and family violence.

For training, the set was divided into two parts, 70% were the train set while 30% were test, the categorical variables for the model had to be transformed to numeric with onehotencoder, then the original numeric variables are added to form a same table, this table is converted to ndarray to put it into the training model which is a polynomial regression model. This resulted in a vector with thetas which are the coefficients by which we will arrive at the dependent variable (number of occurrences). Then the root mean square error is calculated, which was 4.4.

Social implications

A correlation was found between social conditions and crime increases, unemployment and inflation increase with respect to crime, in our point of view more variables that are involved were missing in this problem, in order to have a completely detailed analysis in its reciprocity, to know if there are more influential aspects in which insecurity in Mexico is increasing

Conclusions

Thanks to the model, it was possible to obtain the value of the squared error, which has a value of 4.4, this means that the model can predict the number of normalized incidences +- 4.4. This can be considered a positive result, since we believe that this value is low and does not represent a large margin of error, so it can be concluded that the model was well trained

and gives possible results that can be taken into account in order to generate public policies. to improve security in those states.

For crimes with many incidents, the mean square error can represent a very small error, but for less common crimes or with fewer incidents, it can be considered a very large error, having only a couple of incidents at most each month.

Bibliography

Quiroga, Héctor. *Delinquency as a social fact*. Mexican journal of sociology, 1960.