Understanding crime in Los Angeles

Felipe Lambach, Joana Martins, Lucas Fischer

Abstract

Motivated by the influx of crime in the city of Los Angeles, in this paper we study the main questions arisen from our research. This work is based on the crime and arrest data provided by the Los Angeles Police Department available in [?] and [?]. These datasets contain information regarding crime incidents and arrests since 2010 to the present.

To provide a social context to this work, we employ complementary datasets on demographics, educational and economical indicators relative to the city of Los Angeles obtained from the US Census Bureau through American Fact Finder [?].

Among the questions we aim to explore are: what is the historical evolution of the most prevalent types of crime, at the city and at a local level; how the different types of crime are distributed geographically and if they are correlated with social and economical indicators such as unemployment rate, education level and percentage of population below poverty level. Additionally, we will examine how are the types of crimes distributed according to the age brackets, as well as the seasonality of crimes, and how historical events (e.g. crisis periods) impacted the crime incidence.

1 Introduction

Being one of the largest cities in the United States, Los Angeles is inherently amongst the cities with the highest crime rate. Crime is always a subject worth studying in order to improve the quality of life within a city, and the dataset used in this work provides us with the tools needed to perform this study. Upon an initial investigation concerning this subject the team was motivated to further our research in this topic and try to answer some important questions with the help of data visualization.

The main datasets are provided by the city of Los Angeles and the US Census Bureau. The datasets relative to the crime and arrests contain data regarding the date and location in which they occurred, the type and characteristics of the crime/arrest and the traits of the victims or perpetrators.

The complementary datasets from US Census extracted from [?] contain historical demographic information such as population number and distribution by age bracket and gender, educational data (e.g. educational attainment and school enrolment rates per age bracket and gender) and economical data (e.g. employment status in the past 12 months and population below and above poverty level).

The software used for the development of our research was Tableau to create all the data visualizations in order to communicate our findings, and Python to enables us to pre-process and enrich our dataset.

In section 2 we describe in detail all the datasets used, specifying all the fields present in each dataset, as well as the size of each dataset. To further contextualize the problem, we describe some important related works in section 3. Section 4 is the main body of this paper. In this section all the details from our proposed solution are specified accompanied with the data visualizations created answer our questions.

2 Datasets

In order to provide reliable information through the visualizations created, a good dataset is crucial. After a thorough search, the team found two datasets provided by the Los Angeles Police Department with information regarding events of crime incidents [?] and arrests [?]. These datasets enable us to build effective visualizations but in order to correlate these events with social and economic metrics complementary datasets provided by the US census were needed and obtained through the website American Fact Finder [?]. This website enables advanced search of Census data, including the selection of the geographical granularity. This section is dedicated to describing in detail these datasets.

2.1 Crime in Los Angeles Dataset

Both this dataset and the one described in the next section are provided by the Los Angeles Police Department. This reason and the fact that the dataset is updated weekly were crucial factors for the decision to use these datasets, as the reliability of a dataset is crucial to provide accurate information.

This dataset reflects incidents of crime that occurred in a specific time and place in the metropolitan area of the city of Los Angeles, since 2010. Each row in this dataset maps to a crime event that was transcribed by original crime reports that are typed on paper and as such, some information may be inaccurate. There are also some location fields with missing values denoted by (0^o, 0^o).

The dataset has a total of 1.967.976 rows and 28 fields describing each crime event such as the Location, the date, the age and gender of the victim and other useful information used to characterize a crime. A full description of every single variable of the 28 variables can be found in annex A in table 1.

2.2 Arrests in Los Angeles Dataset

Similar to the dataset described in section 2.1 this dataset is also provided by the Los Angeles Police Department and its records are also transcribed from original police reports, but each row reflects the booking of an arrestee.

There are a total of 1.250.047 arrest events described by 19 different fields similar to the ones described in section 2.1. A full description of every field can be found in the annex section B in table 2.

2.3 Census

In order to provide complementary information to the main datasets, some datasets provided by United States Census were used, namely a dataset on the population below and above poverty level in the past 12 months by sex and age (dataset B17001), school enrolment by sex and age (dataset S1401), distribution of population by sex and age (dataset S0101), educational attainment for the population with 18 years or more (dataset B15001) and employment status for the population with 16 or more years of age (dataset B23001), by sex and age. All these datasets were obtained at both city and zip code level to allow a geographical distribution analysis. At the zip code level, only data from 2011 to 2017 was available so this was the range used throughout the local level study. Furthermore zip code level data is only available in terms of estimates based on 5 years.

The fields available for these datasets are specified in detail in the appendix C.

2.4 Geocoding

Since the data obtained from the US census is available aggregated to a zip-code level, an additional processing was necessary to enable us to correlate crime/arrest incidents with a given zip-code, in order to create visualizations at geographical level of granularity.

In the [?] dataset there is a Location field which provides information on the latitude and longitude of a given arrest incident. With this information the python library `uszipcode` [?] was used to obtain the zip-code of a given latitude and longitude. It's important to note that a zip-code area is subject to change throughout the years, and since the team only used zip-codes corresponding to the year of 2019 some minor inaccuracies in the area of each zip-code may be present for past years, although these inaccuracies might not distort the credibility of our visualizations. With this library a new csv file was created, where each line maps the location variable present in [?] to a zip code. Upon retrieving this information, the Geocode capabilities of Tableau can be used to create the a visualization of different measures per zip-code area. The Tableau visualization of the zip code areas mapped is shown in Figure 1.

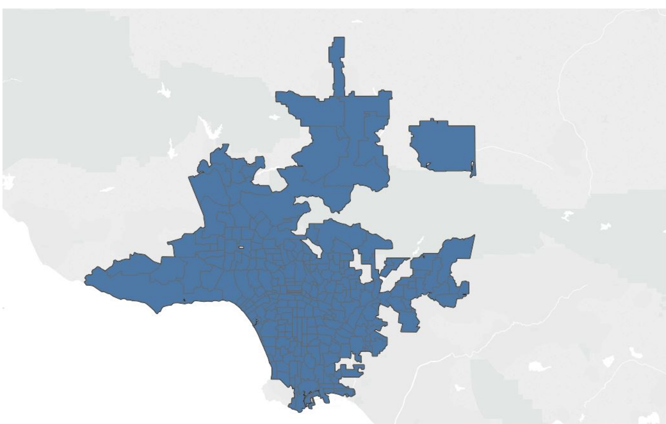


Figure 1: *Zip code areas obtained for the arrest dataset*

3 Related Work

The study of crime in a major metropolitan area such as Los Angeles is not a novel study. Various different authors have published studies regarding this subject. There are more formal studies such as [?] who investigated the impact of immigration and its relation to gang activity, and also some articles describing crime in Los Angeles in the last years. Aizman [?] lists the main crime incidents in recent history, [?] presents a project by students and researchers of the USC's spatial science institute who explored the effect that variables such as unemployment rate or proximity to a street light have on the location of crime incidents.

Although these works motivate us to investigate the questions described in the abstract, they lack the aid that data visualization brings and so our main body of work is to present a study where data visualization is used to help research crime in Los Angeles. Since the dataset is publicly available on Kaggle, there are notebooks who implement some generic data visualizations of our data, most notably the notebook of [?], who visually explores the variables in our dataset such as crime distribution per gender, age, time of day, etc. His work is the most similar to ours but it does not attempt to study the same questions our work does. Throughout this research we shall make the most comparisons with his work as it uses the same dataset as ours.

4 Proposal

This section refers to all the processes made in order to create the visualizations needed to accompany our study of crime in Los Angeles. It will be segmented into two main categories: Exploratory Data Analysis (section 4.1) and Main Visualizations (section 4.2)

4.1 Exploratory Data Analysis

A good rule of thumb in every project that handles large amounts of structured data is to do an exploratory data analysis, i.e. explore the data researching some generic indicators such as means and modes with no specific objective in mind.

Before trying to visualize anything, the group identified how many missing values the dataset contained and being that it is not a relevant portion of the data, the decision to discard these values was made.

The first simple questions the group sought to answer were: If males were committing more crimes than females; What was the prevalent gender of the victim and the offender; What was the distribution of the criminals and victims age and what were the most common crimes in our dataset.

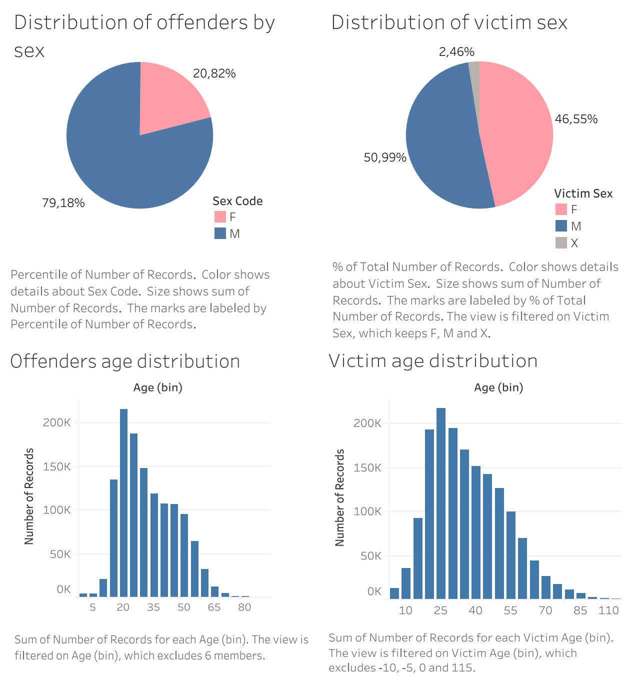


Figure 2: *Distribution of crimes by victim and offenders gender and age*

To help answer these questions two pie charts and two histograms were created. The pie charts visualize the distribution of offenders and victims gender and the histograms depict the distribution of offenders and victims by age with an interval of 5 years. These graphs are displayed in a matrix with two columns (one column for offenders and one for victims) and two rows. Figure 2 shows this matrix where can be observed that the majority of crimes are committed by males and males are also the majority of victims, and that distribution of both the victim and the criminal is largely concentrated around 20 to 30 years old.

To determine the most common types of crime, a bar plot of the their frequency was built using the charge description group field values as categories. This bar plot is shown in Figure 3, where it can be seen that simple assault dominates the reported crimes with over 8% of crimes reported as simple assault.

To extend this subject, the team chose to study in greater depth three types of crime: Vehicle Theft, Aggravated Assault with deadly weapons and Robbery. To do so, in the next visualizations a filter according to the charge group description field was applied to the visualizations, where the values *Aggravated Assault with deadly weapons*, *Robbery* and *Vehicle theft* were selected.

Figure 4 shows how much do these types of crimes amount relative to the total number of reported crimes. It is possible to see in all cases that men are the most common perpetrator in these types of crimes and that Aggravated assault accounts for roughly 6.5% of all reported crimes since 2010. It should be noted that to show the percentage of records for each of these crimes relative to the total number of records and not simply to the records belonging to these three crime types, a level of detail expression was used where the total number of records was fixed.

To further analyze these crimes, figure 5 maps the perpetrators age and gender relative to each one of these crimes. This image shows that for the three types of crime analysed, men are the most common perpetrators and that the age distribution in Robbery tends to be higher around 18 years all and falls down quickly after that, while in the other types of crime the distribution of the perpetrator age seems to take more time to fall down. During the development of this study, the team focused on these types of crime and so section 4.2 mainly focuses on them.

4.2 Main Visualizations

Having explored how the data is distributed and having a new sense of knowledge over our data, the team could start attempting to answer the questions it proposed. First the distinction was made between city and local level, where city level hopes to explain data relative to all the city of Los Angeles and local level tries to explore the same information but at a zip-code granularity level.

The first question was to determine if the most prevalent types of crime remain the same throughout the years. To obtain the answer, a simple line plot was chosen to visualize this concept as it gives us a good mapping when the x axis is a continuous value as it is the case with time. To build it, the number of records was plotted against the date of occurrence, using the year function to extract the year information from the date of occurrence dimension.

Figure 6 shows the temporal evolution of the number of occurrences of the most common types crimes from 2010 to 2018. In this image it can be observed that Simple Assault is the most common type of crime since 2010 except in 2016 when the number of Vehicle Thefts incidents surpassed that of Simple Assaults. In this period Vehicle Theft saw a rapid increase from 2014 to 2016 along with Burglary from Vehicle since these two crime types are largely correlated. It is also interesting to point out that Petty Theft saw a large increase from 2010 to 2012 and stabilizes after that, a period that coincides with the aftermath of the financial crisis of 2008 which suggests this might have been a factor that contributed to this rise in theft.

To answer the second question, on how the different crime types are distributed geographically, the team focused on the three specified types of crime (Aggravated assault with deadly weapons, Vehicle theft and Robbery). To visualize the geographic distribution of these crimes, we took advantage of the zip-code obtained through the process described in section 2.4. To build the maps that will depict the geographic distribution of each of the crimes per zip-code, the zip code dimension (which has a geographic role known to Tableau) was used. This results in a map where each zip code area is colored according to the number of crimes occurred in it.

Figures 7, 8 and 9 are the result of this process and represent the geographic distribution on these types of crime. In these figures it is clear that these crimes are concentrated in the area of South Los Angeles, a zone known to be problematic, whereas in Figure 9 in addition to this finding, it is possible to distinguish a more concentrated number of crimes in the far south (Long Beach), northeast of downtown Los Angeles (Passadena) and northwest of downtown Los Angeles (North Hollywood and Burbank), which are typically richer areas of the city of Los Angeles with a higher number of expensive vehicles on the street.

In order to determine possible factors that contribute to the higher concentration of these crimes in these areas, a map of percentage of people below poverty line per zip-code was created. Our goal was to understand if poorer areas in the city of Los Angeles have higher crime rates. The plot, present in Figure 10, colors each zip code with the percentage of people below the poverty level. This is useful to highlight areas of extreme poverty. It is possible to see that the areas with higher percentage of population below poverty level are near the zones of Southeast Los Angeles. This almost perfectly coincides with the areas where the above crimes happen more frequently, and Figure 12 plots the correlation between the number of crimes and the percentage of people below the poverty line which confirms this finding.

The same process was used to plot the geographical distribution of the percentage of unemployed people in the labor force, which is shown in Figure 11.

To better verify if poverty is really a factor in the increase of these types of crime the plots present in Figures 12, 13 were created. These scatter plots depict the correlation between the number of these crimes committed versus the percentage of people below the poverty level and the percentage of people unemployed respectively. Each circle in the plots represents a specific zip code and its position marks the corresponding sum of records for this crime and zip code and the percentage of population below poverty level or the percentage of unemployed civilian population (averaged over the years) for the same zip code. To produce this graphs, it was necessary to connect the data sources associated with the crime dataset and the Employment status and Poverty level datasets. To do so, we established custom relationships between the different data sources, linking them by zip code and year, the dimensions common to all the data sources.

It is possible to see that most zip-code areas present a strong correlation between number of committed crimes and these poverty metrics. In the previous plots, it is possible to distinguish some zip-code areas with a high number of reported crimes. We extended our search for correlation in these outliers to better understand them. Figures 14, 15, 16, 17, 18, 19 display the correlation between this social-economic factors with the different types of crime, separating by color the zip-code areas that have more than 2% of the reported crime type records. To create these plots, histograms of the distribution of each of these crimes per zip code were done and from them sets of zip codes with a percentage of records higher than 2% of records for each crime were created. These plots suggest that in these outlier areas, although there is a correlation between poverty and unemployment, it is not enough to justify the high number of crime incidents.

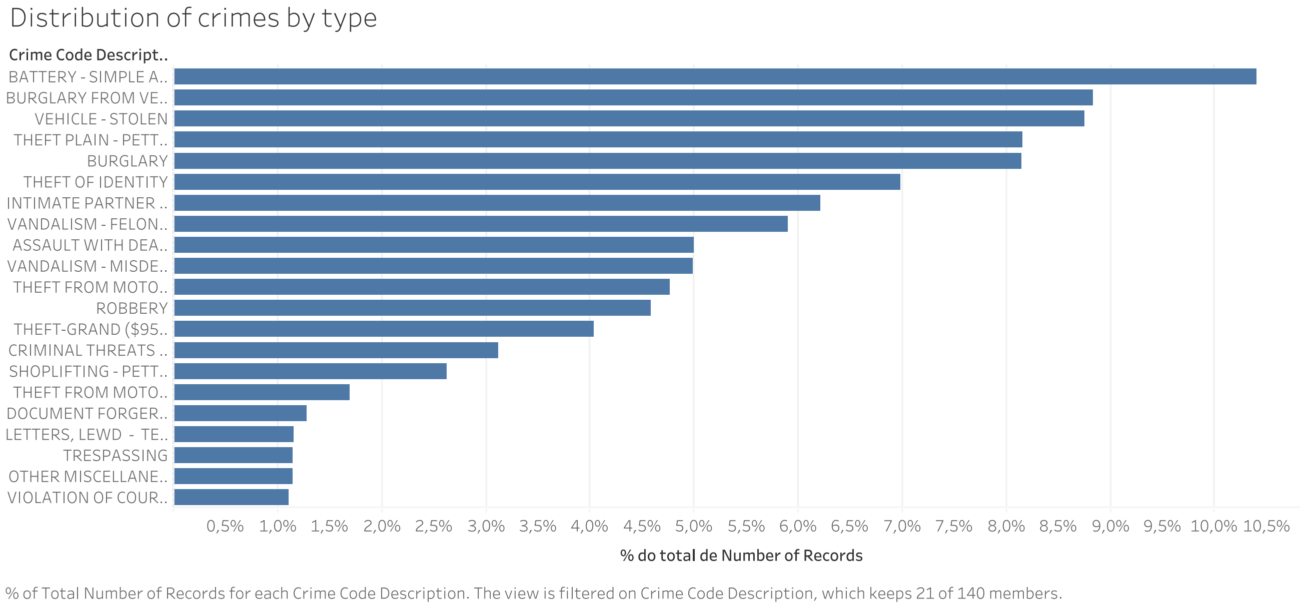


Figure 3: *Distribution of crimes by its type*

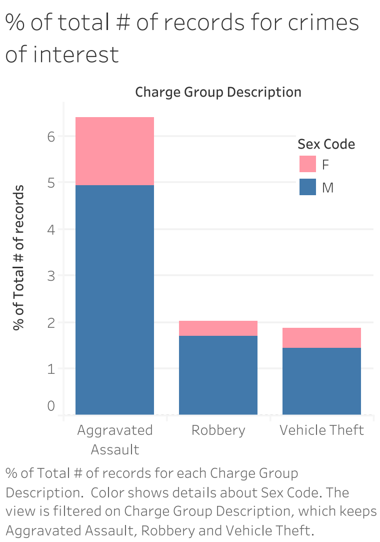


Figure 4: *Distribution of how much the crimes of interest amount to the total number of reported crimes, colored by portion of incidents committed by males and females*

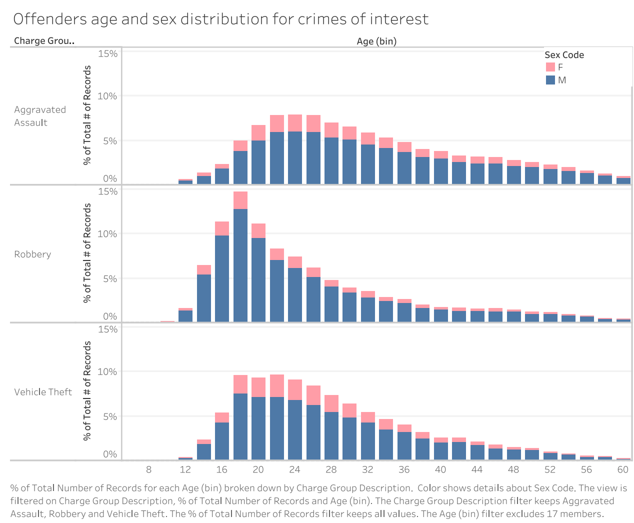


Figure 5: *Distribution of offenders age and gender in the three crime types studied*

Another question of relevant interest is to determine how are these three types of crimes are distributed in terms of the offenders age groups. It is interesting to understand if, for example, youths commit more robbery than aggravated assaults, if the predominant age group in a certain time evolves over the years suggesting that the same persons that committed a crime years before are the same ones committing the crimes now. Although Figure 5 already visualizes the predominant age group of offenders for each crime type, it does not inform on how these distributions evolve over time. To tackle this question a heat map was built where the columns are the years ranging from 2010 to 2018 (2019 was excluded with a filter since it is only a partial year as of the time of this writing) and the rows are the different age bins the team discretized. The heat maps for each crime type are present in Figures 20, 21, 22. Figure 21, which corresponds with the heat map of robbery crimes, indicates that the predominant age group that commits this offenses is densely concentrated around the ages of 16-22, a fact that can also be observed in Figure 5. Perhaps the most interesting all the three heat maps is in Figure 22 that refers to vehicle theft. The interest in this heat map is in verifying that the darker cells in the heat map tend to move down as the x axis increases. This finding suggests that the predominant age of vehicle thieves tends to increase throughout these 9 years, which may suggest that the same persons that were committing the crimes before are still the ones committing the crimes as the years

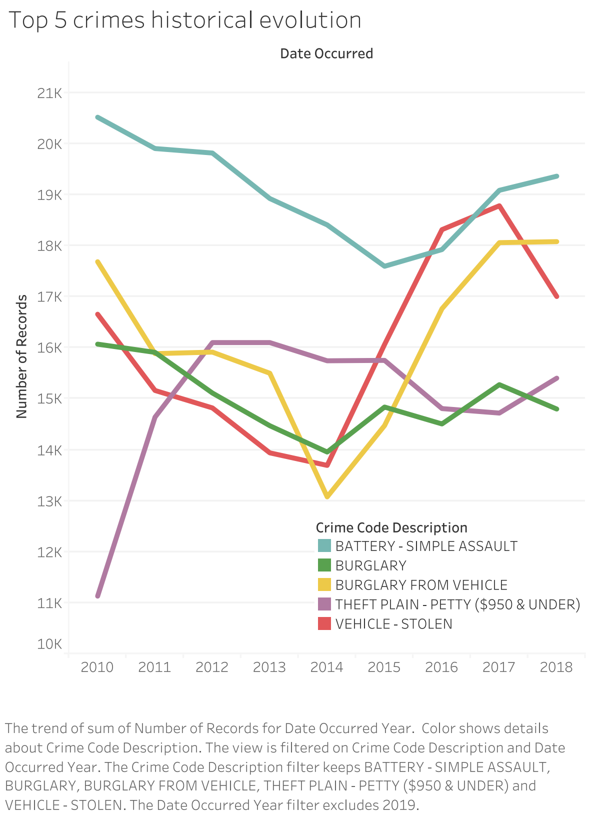


Figure 6: *Most common types of crime throughout the years*

go by. This finding is also supported by Figure 23 that shows that the average age in the offenders for these crime types are rising as the years pass.

The fourth and final question the team sought out to answer was to determine if there is some seasonality factor in the three types of crime studied. Motivated with the notion that tourism increases in Los Angeles in the summer, it is easy to extrapolate and think that there might be more robberies or other types of crime during this period. To factually answer this question, a line chart was built where the x axis represents the several years in study correctly divided by the month in order to determine in which months do crimes happen the most. The y axis represents the number of reported crimes for each month in each year. This line chart is displayed in Figure 24. To achieve this in Tableau, both the year and month were extracted from the date of occurrence dimension and placed in the columns shelf and the number of records was put in the row shelf (it was aggregated by summing the number of records). A filter was applied to select the crimes of interest from the charge code description field and the year of 2019 was also excluded from the view by filtering it out. The crime code description field was also applied to the color section of the Marks card to color each of the crimes of interest distinctly.

Figure 24 helps to illustrate that although there is no strong seasonality in the studied crime types, they do tend to have a slight increase in number of incidents in the summer months in some particular years. This is observed for vehicle thefts in 2011 and for aggravated assaults with deadly weapon and robberies in 2010 to 2014 and in 2017 and 2018.

The possible correlation between the crimes of interest and the percentage of children, teenagers and young adults enrolled in school or the educational attainment will also be studied through visualizations analog to those presented above (geographical distribution and correlation plots), since we possess the necessary datasets to do so.

4.3 Interactive Visualizations

In our datasets we have both spatial and temporal information regarding crimes, arrests and social variables such as employment status of population in labor force and population below poverty level. In static visualizations it is hard to describe the historical evolution simultaneously with the geographical data. Therefore, we will implement interactive data visualizations that allow the user to dynamically explore the data in both time and space.

4.3.1 Historical information of measures at zip code level

An interactive visualization will be implemented where a map of Los Angeles zip code areas colored according to a measure M1 will be displayed. The user can then click on a zip code and visualize in a line plot presented in an extra navigation panel, the historical evolution of this and other measures (M2, M3,...) filtered for this zip code. In this manner, the user can navigate both geographically and in time. A sketch of how this interactive visualization will work is depicted in Figure 25.

4.3.2 Evolution of measures at zip code level for two selected years

In the plots with the historical evolution of a given measure M at year level appearing in the extra navigation panel described in 4.3.1, the user will be able to select two years (Y1 and Y2) and get a new view where the map of a measure M at years Y1 and Y2 will appear side by side along with a new map colored according to the difference of the value of measure M between the two years. This will allow the observation of how measures such as the number of crimes or the percentage of people below poverty level in each of the zip code areas changed from one year to another. The sketch of how this interactive visualization will work is shown in Figure 26.

5 Conclusion

After the analysis made in section 4.2, it is possible to take some key findings out of this study.

In the first question, an important point made was how the financial crisis may affect the behaviour of crimes. In 2010 up to 2012, a rapid increase in petty theft can be identified, which stabilizes after this period. This period also coincides with the aftermath of the housing crisis of 2008 experienced in the United States.

The second question answered identifies the south zone of Los Angeles, an historically problematic zone, as having a higher number of reported crimes. Beyond this, it was also possible to uncover that vehicle theft also seems to happen with some frequency in richer areas of Los Angeles such as North Hollywood, Burbank, Pasadena and Long Beach.

For the third question, an analysis was made to the profile of the offenders age group in the three crime types studied. After viewing Figure 23 an increase in the average age of perpetrators is noticeable, which suggests that the persons that were committing these crimes in the past are still committing them over the years.

In the fourth and final question, the group tried to analyze if these crimes have a seasonality factor to them. Upon visualizing the

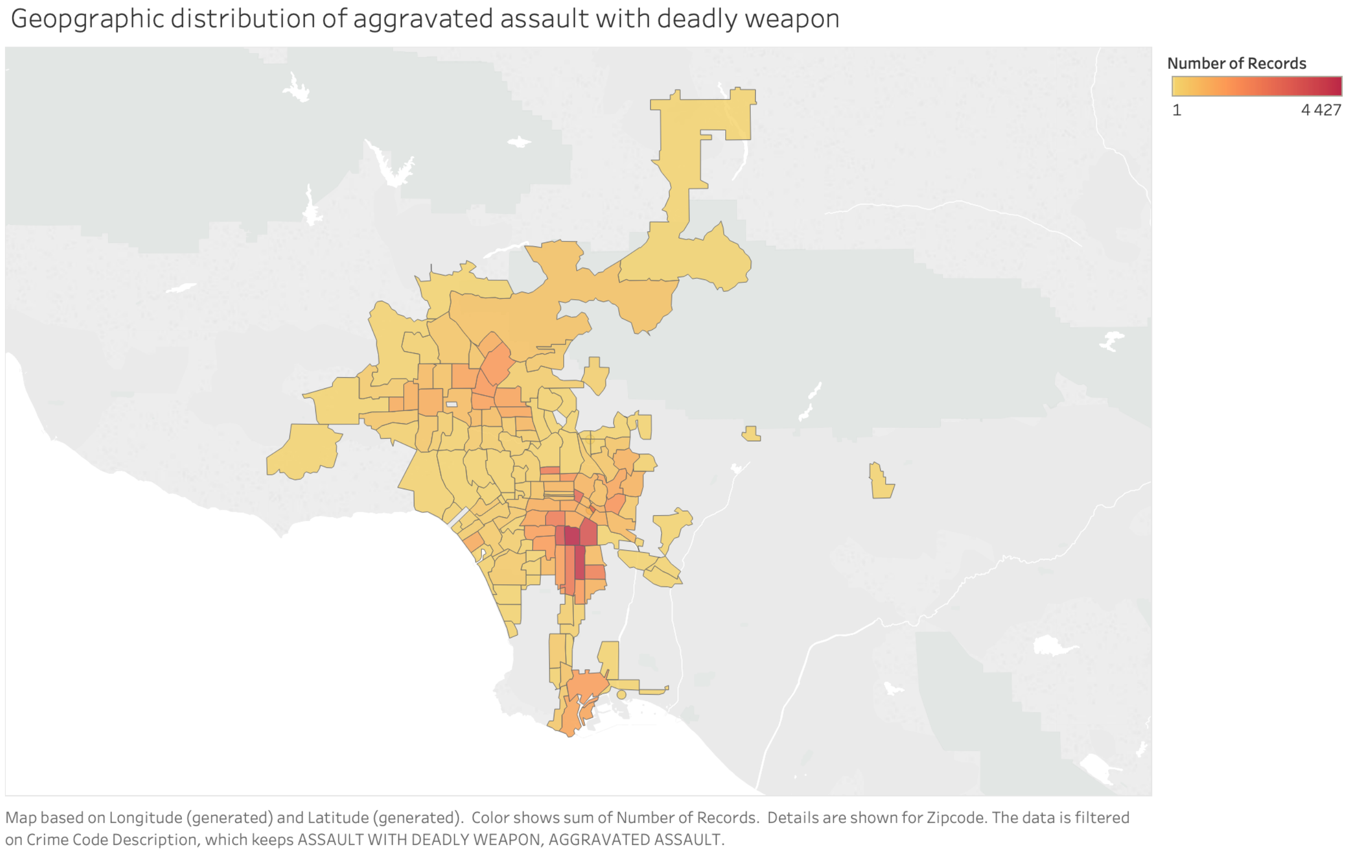


Figure 7: *Geographic distribution of aggravated assaults in the city of Los Angeles*

data to answer this question, present in Figure 24, there seems to be no significant seasonality influence in the number of these reported crimes, although a slight increase is noticeable in the summer months for some years, which can suggest a correlation with a higher influx of tourists at these times.

Although some of the proposed questions have already been answered in this report, mostly at a city level, it is intended that this research will be deepened with a more thorough analysis at zipcode level with the aid of interactivity.

A Crime Data from 2010 to Present dataset description

The first of the two main datasets used in this study refers to the incidents of crimes in the city of Los Angeles. It has a total of 1.967.976 rows and 27 columns. As described in section 2.1 this dataset has spatial and time information about each crime record as well as other variables. Table 1 presents a detailed description of the fields in this dataset that were used to create visualizations. This dataset consists of transcripts from the original criminal reports typed in paper, and as such, some missing values do occur. Since the number of rows with missing values tends to be very small, these rows were simply removed from our analysis with Tableau as they did not impact much the visualizations.

B Arrest Data from 2010 to Present dataset description

The second of the two main datasets refers to reported events of arrests in the city of Los Angeles. This dataset is slightly smaller, consisting of 1.250.047 rows and 18 columns. Table 2 describes in detail all used fields for this research. As with the crime dataset, rows with missing values present in this dataset were removed from the analysis made with Tableau, since the number of lines with missing values is insignificant.

C US Census Bureau dataset descriptions

In this section, a detailed description of the datasets taken from the US Census Bureau through American Fact Finder is provided, divided per dataset.

For a given dataset code, the data was provided in separate dataset for each year, which we combined for ease of use, using a python script. To be able to correctly identify the data belonging to each year we added an additional field called year.

C.1 Datasets B17001: Poverty status in the past 12 months by sex and age

Each of the yearly B17001 datasets is comprised of 121 fields and 127 records. In addition to the estimates for the different quantities, margins of error are also provided which we did not use. We have also extracted only a part of the available fields. Therefore, the

Table 1: *Crime Data from 2010 to Present*

Field	Descriptions
DR Number	Division of Records Number: Official file number made up of a 2 digit year, area ID, and 5 digits
Date Reported	Date of the crime report
Date Occurred	Date of the crime
Time Occurred	Hour of the crime
Area ID	Station where the occurrence was recorded. The Geographic Areas are sequentially numbered from 1-21.
Area Name	The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for
Reporting District	A four-digit code that represents a sub-area within a Geographic Area.
Crime Code	Indicates the crime code committed
Crime Code Description	Describes the Crime Code provided.
MO Codes	Activities associated with the suspect in commission of the crime.
Victim Age	Victim Age
Victim Sex	Gender of the Victim
Victim Descent	Descent of the Victim
Premise Code	The type of structure, vehicle, or location where the crime took place.
Premise Description	Defines the Premise Code provided.
Weapon Description	Defines the Weapon Used Code provided.
Location	The location where the crime incident occurred. Actual address is omitted for confidentiality. X,Y coordinates reflect the nearest 100 block.

Table 2: *Arrest Data from 2010 to Present*

Field	Descriptions
Report ID	ID for the arrest
Time	In 24 hour military time
Area ID	Station where the occurrence was recorded. The Geographic Areas are sequentially numbered from 1-21.
Area Name	The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for
Reporting District	A four-digit code that represents a sub-area within a Geographic Area.
Age	Offender age
Sex Code	Gender of the offender
Descent Code	Descent of the offender.
Charge Group Code	Category of arrest charge.
Charge Group Description	Defines the Charge Group Code provided.
Arrest Type Code	A code to indicate the type of charge the individual was arrested for.
Charge	The charge the individual was arrested for.
Charge Description	Describes the Charge provided
Location	The location where the crime incident occurred. Actual address is omitted for confidentiality. XY coordinates reflect the nearest 100 block.

number of fields in the combined dataset is much smaller.

The combined dataset has a total of 8 fields and 889 records. The following table describes these fields.

C.2 Datasets B23001: Employment status of population with 16 years of age or more by sex and age

Each of the yearly B23001 datasets is comprised of 349 fields and 127 records. In addition to the estimates for the different quantities, margins of error were also provided which we did not use.

The combined dataset has a total of 175 fields and 889 records. The following table describes these fields. In the description below we omit the data relative to the population in the armed forces, which we did not use.

Table 3: *Datasets B17001: Poverty status in the past 12 months by sex and age*

Field #	Field	Description
1	Id2	5-digit LA zip code
2	Estimate; Total:	Estimate of population per zip code (for a given year)
3	Estimate; Income in the past 12 months below poverty level:	Estimate of population per zip code with income below poverty level in the past 12 months
4	Estimate; Income in the past 12 months below poverty level: - Male:	Estimate of male population per zip code with income below poverty level in the past 12 months
5	Estimate; Income in the past 12 months below poverty level: - Female:	Estimate of female population per zip code with income below poverty level in the past 12 months
6	Estimate; Income in the past 12 months at or above poverty level:	Estimate of population per zip code with income above poverty level in the past 12 months
7	Estimate; Income in the past 12 months at or above poverty level: - Male:	Estimate of male population per zip code with income above poverty level in the past 12 months
8	Estimate; Income in the past 12 months at or above poverty level: - Female:	Estimate of female population per zip code with income above poverty level in the past 12 months

Table 4: *Datasets B23001: Employment status in the past 12 months by sex and age of civilian population*

Field #	Field	Description	age brackets
1	Id2	5-digit LA zip code	-
2	Estimate; Total:	Estimate of total population per zip code	-
3	Estimate; Male:	Estimate of male population per zip code	-
4	Estimate; Male: - X to Y years: - In labor force: - Civilian:	Estimate of civilian male population per zip code in labor force	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64
5	Estimate; Male: - X to Y years: - In labor force: - Civilian: - Employed	Estimate of the civilian male population in labor force employed	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64
6	Estimate; Male: - X to Y years: - In labor force: - Civilian: - Unemployed Estimate of the civilian male population in labor force unemployed	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64	
7	Estimate; Male: - X to Y years: - In labor force:	Estimate of male population per zip code in labor force	65-69, 70-74, ¿75
8	Estimate; Male: - X to Y years: - In labor force: - Employed	Estimate of male population per zip code in labor force employed	65-69, 70-74, ¿75
9	Estimate; Male: - 65 to 69 years: - In labor force: - Unemployed	Estimate of male population per zip code in labor force unemployed	65-69, 70-74, ¿75
10	Estimate; Female: Estimate of male population per zip code	-	
11	Estimate; Female: - X to Y years: - In labor force: - Civilian:	Estimate of civilian female population per zip code in labor force	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64
12	Estimate; Female: - X to Y years: - In labor force: - Civilian: - Employed	Estimate of civilian female population per zip code in labor force employed	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64
13	Estimate; Female: - X to Y years: - In labor force: - Civilian: - Unemployed	Estimate of civilian female population per zip code in labor force unemployed	16-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-61, 62-64
14	Estimate; Female: - 65 to 69 years: - In labor force:	Estimate of female population per zip code in labor force	65-69, 70-74, ¿75
15	Estimate; Female: - 65 to 69 years: - In labor force: - Employed	Estimate of female population per zip code in labor force employed	65-69, 70-74, ¿75
16	Estimate; Female: - 65 to 69 years: - In labor force: - Unemployed	Estimate of female population per zip code in labor force unemployed	65-69, 70-74, ¿75

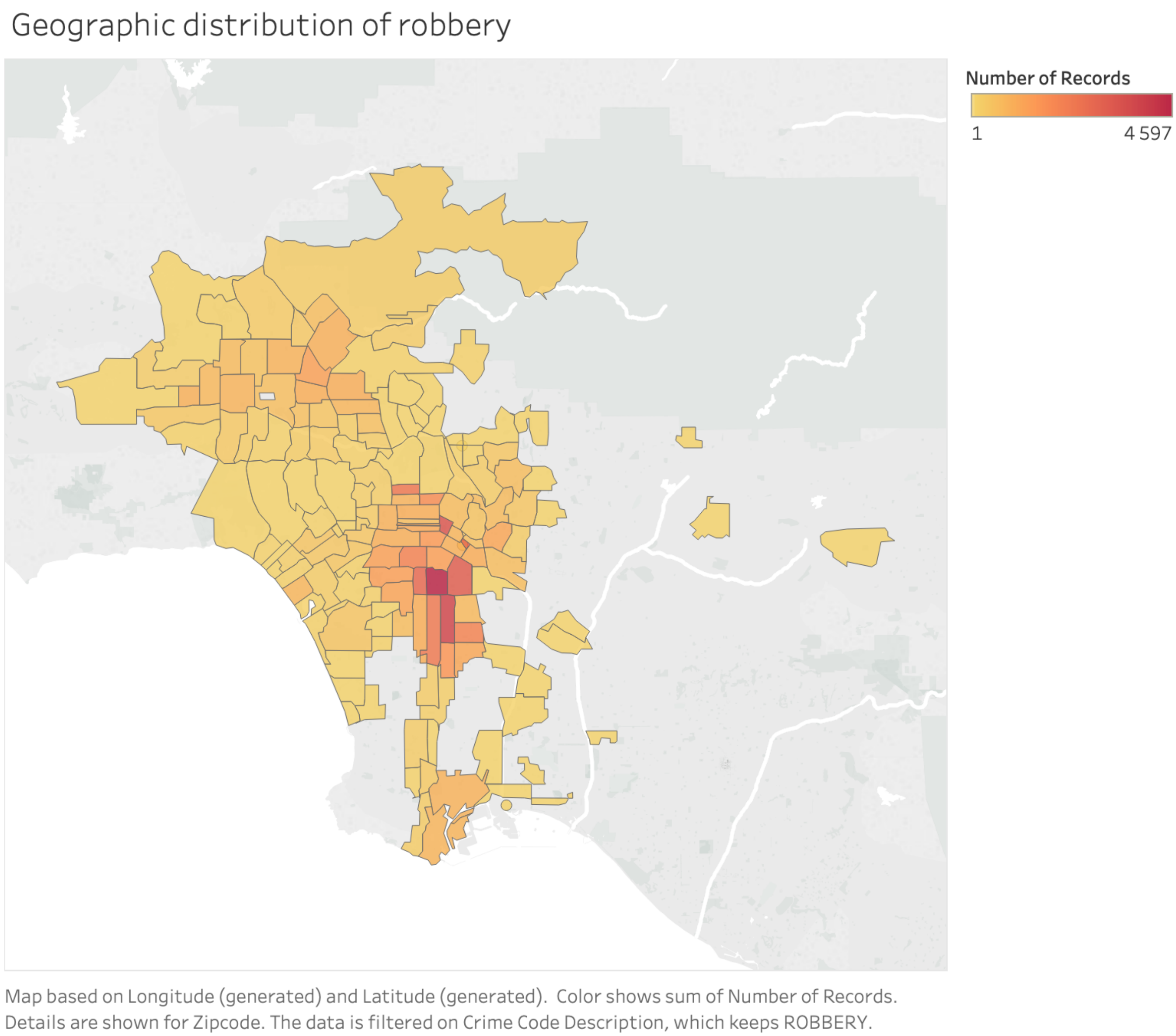


Figure 8: *Geographic distribution of robbery in the city of Los Angeles*

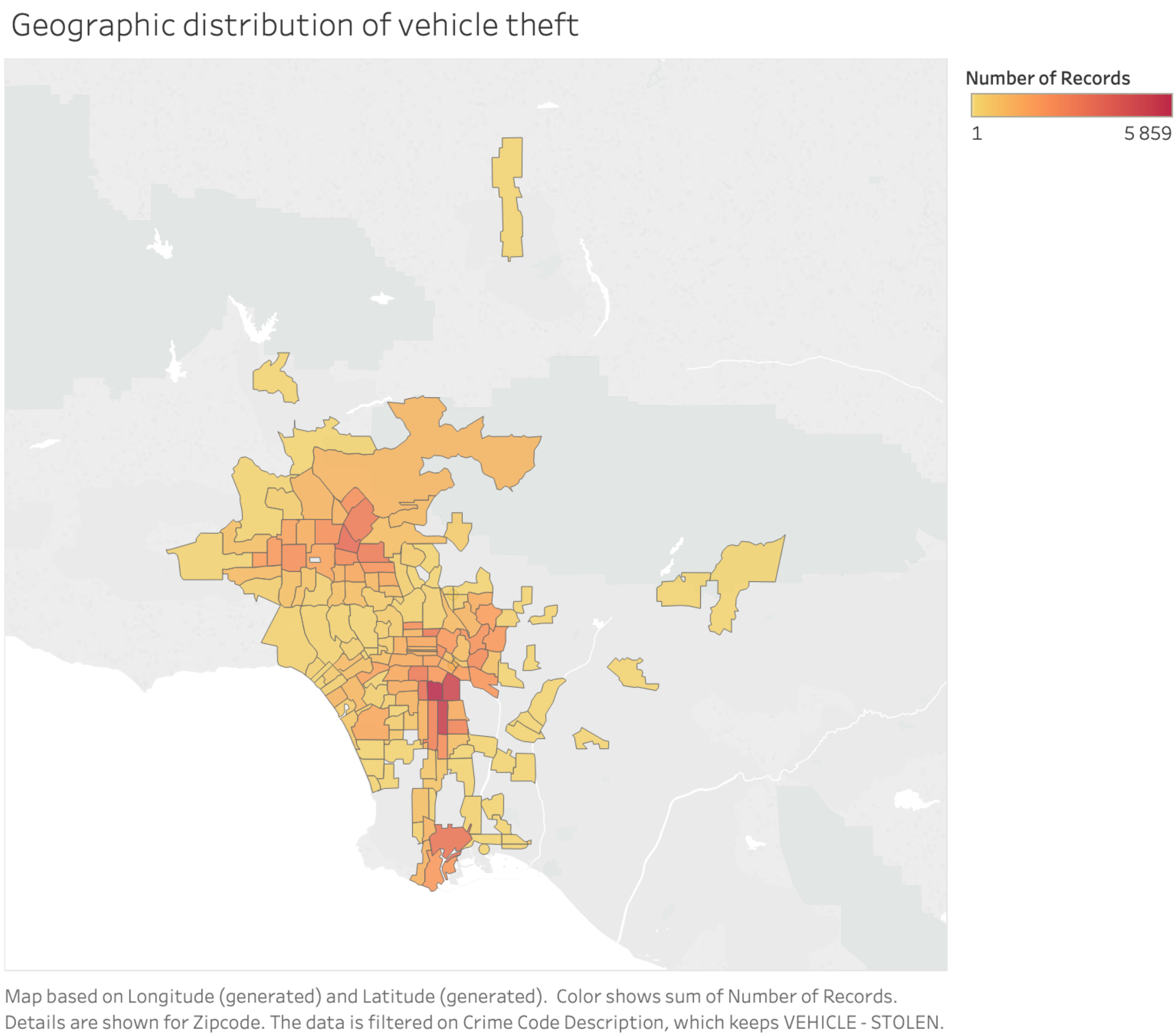


Figure 9: *Geographic distribution of vehicle theft in the city of Los Angeles*

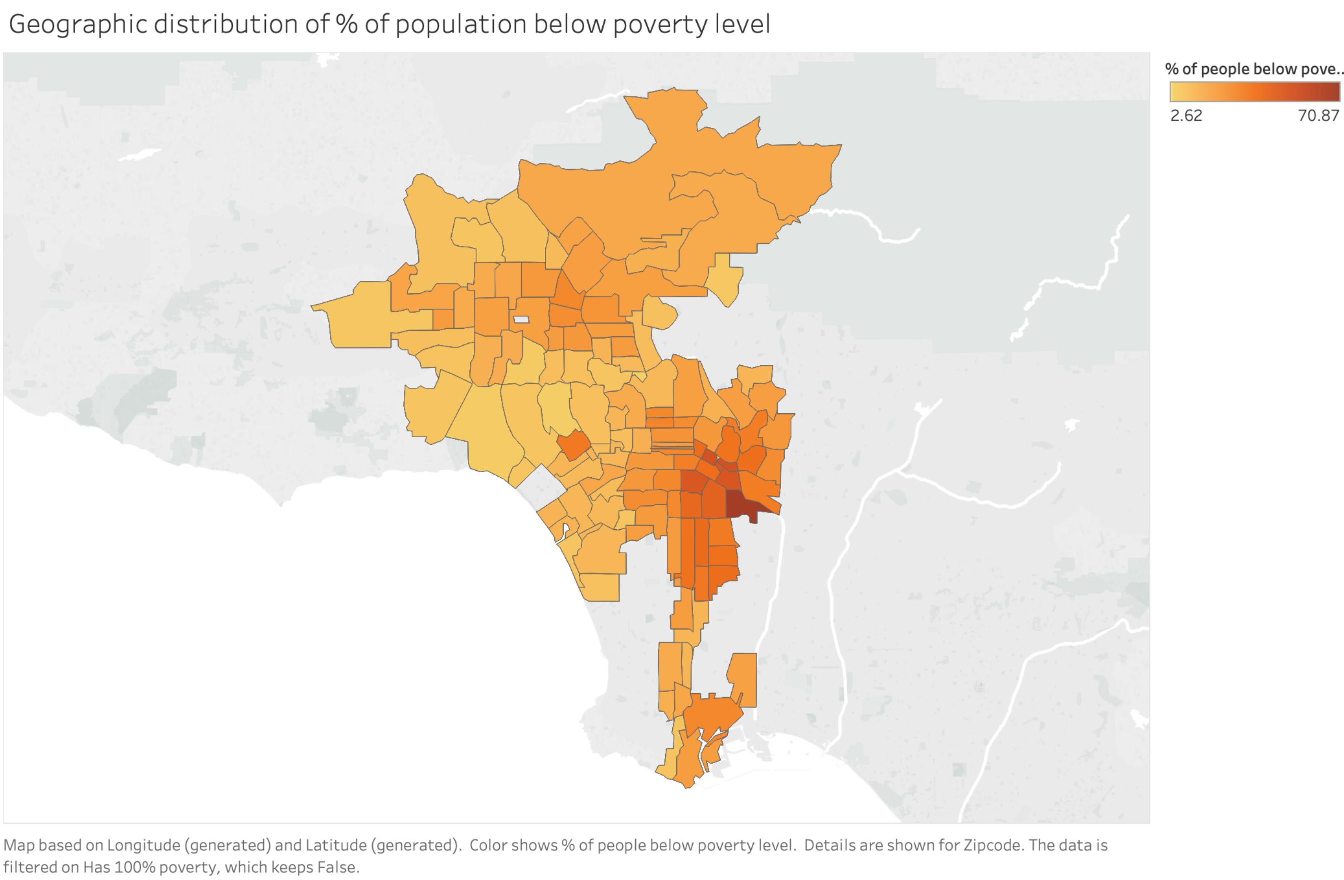


Figure 10: *Geographic distribution of percentage of people below the poverty line in the city of Los Angeles*

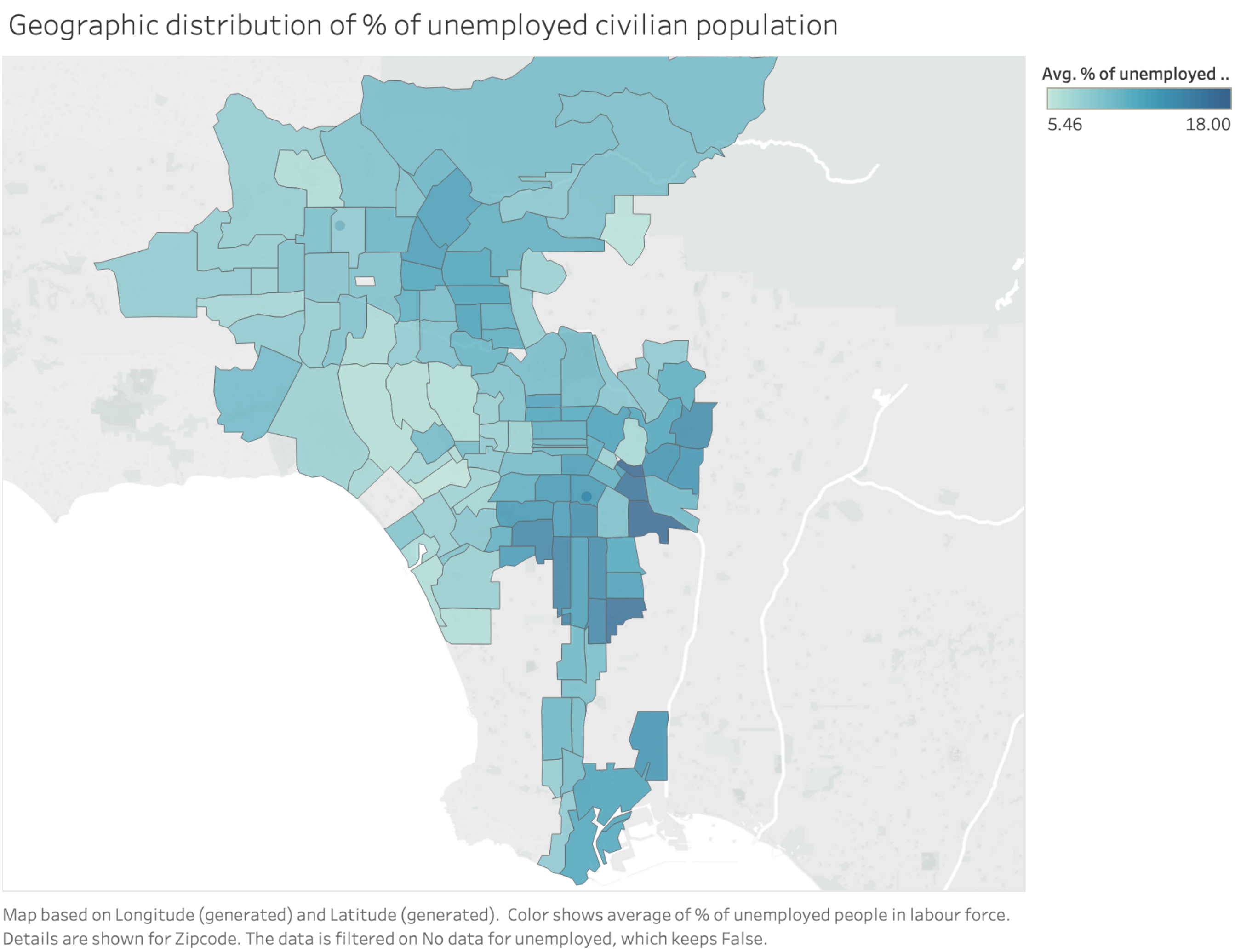
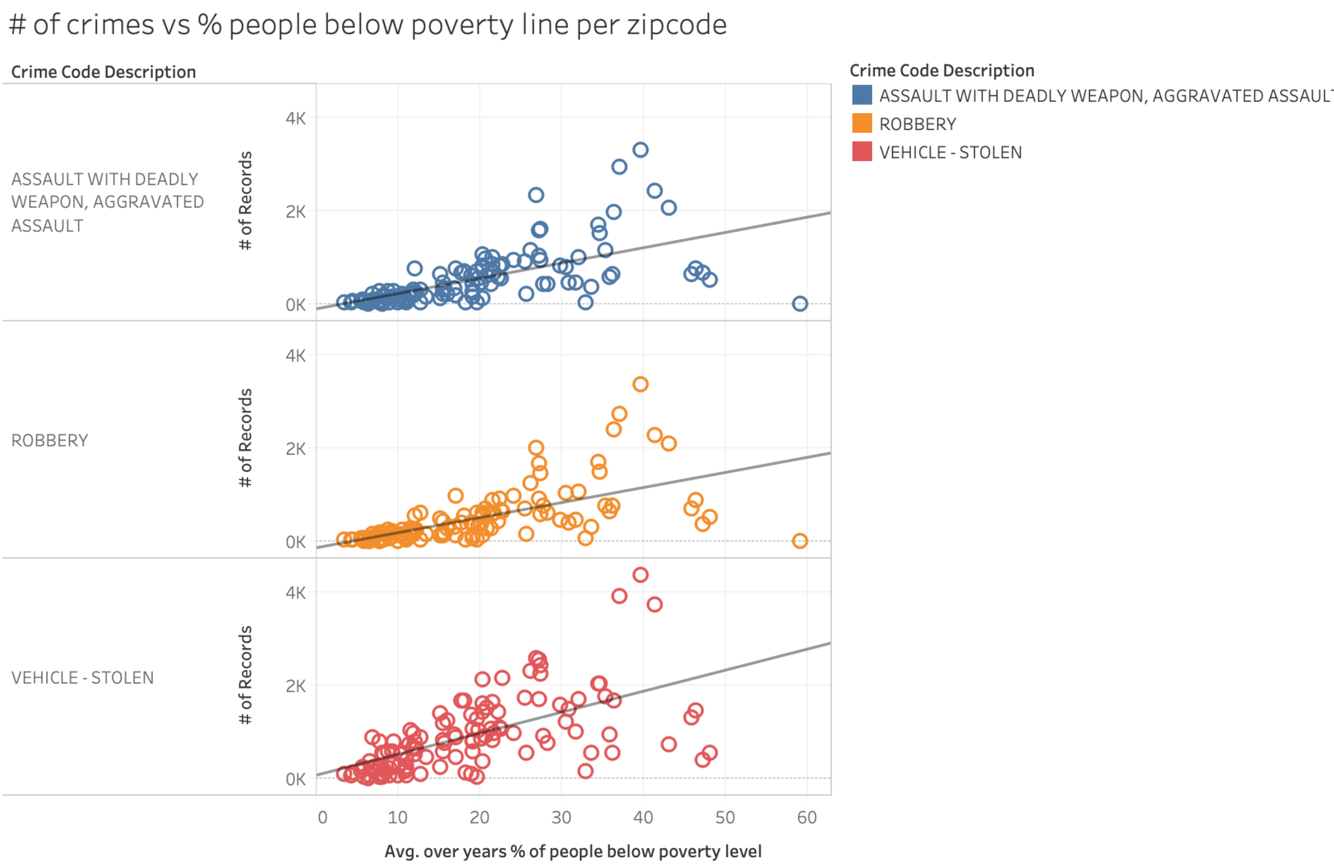


Figure 11: *Geographic distribution of percentage of unemployed people in the labor force in the city of Los Angeles*



Average of % of people below poverty level (Poverty_status_all_5YR_B17001_with_ann) vs. sum of Number of Records broken down by Crime Code Description. Color shows details about Crime Code Description. Details are shown for Zipcode (Poverty_status_all_5YR_B17001_with_ann). The data is filtered on Date Occurred Year, which excludes 2010, 2018 and 2019. The view is filtered on Crime Code Description, which keeps ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT, ROBBERY and VEHICLE - STOLEN.

Figure 12: *Correlation plot between the number of reported crimes and the percentage of people bellow the poverty line*

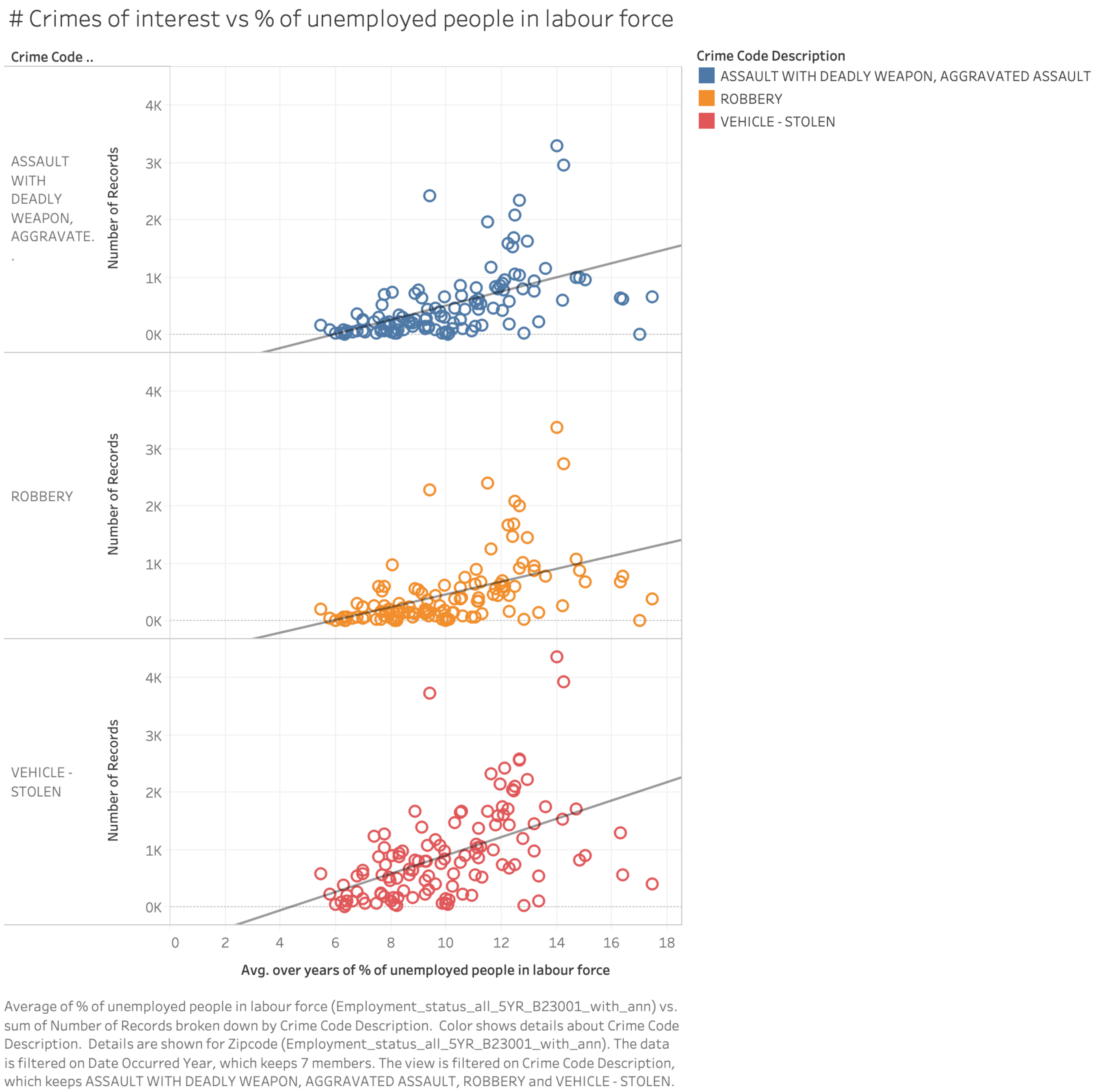


Figure 13: *Correlation plot between the number of reported crimes and the percentage of people unemployed*

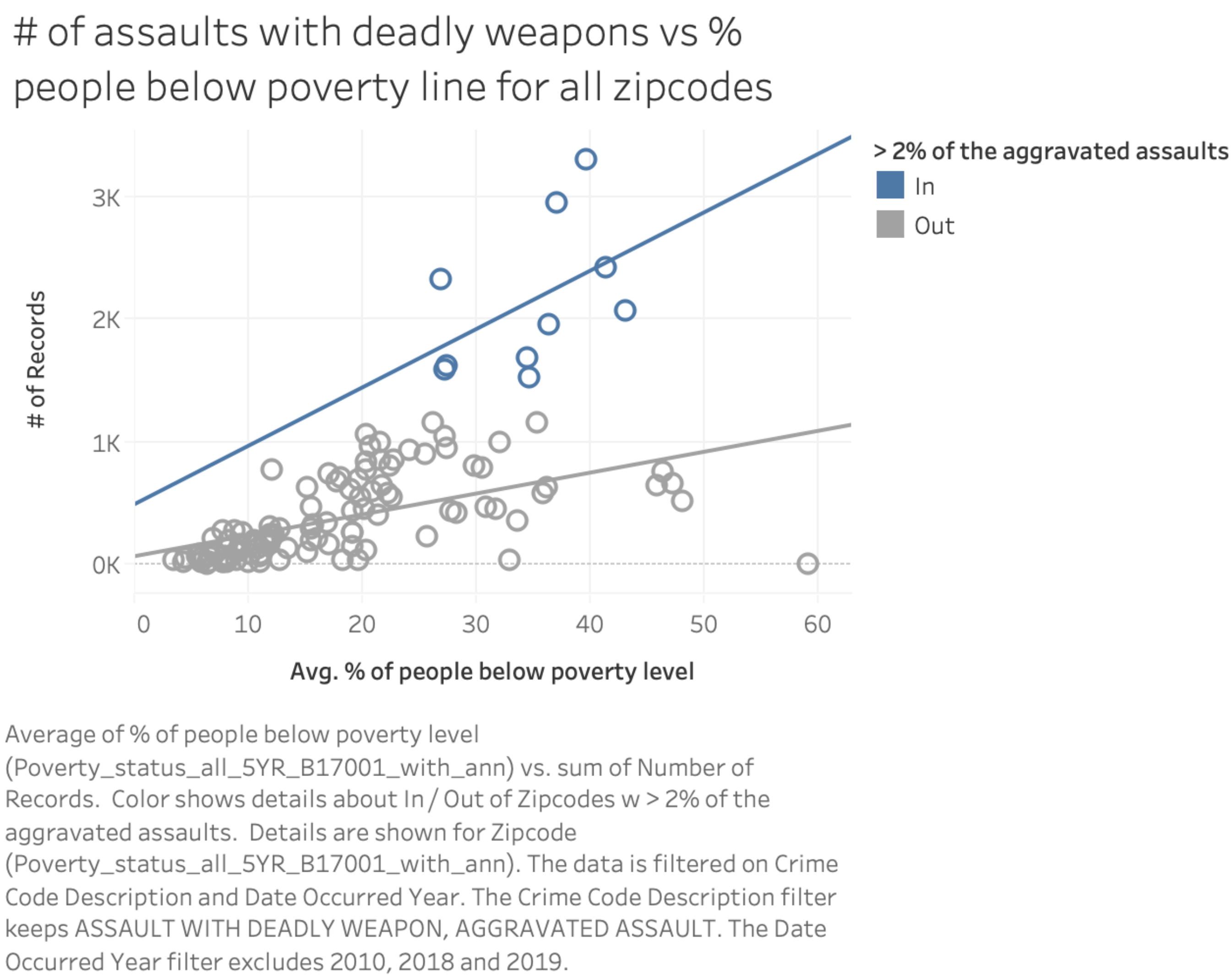
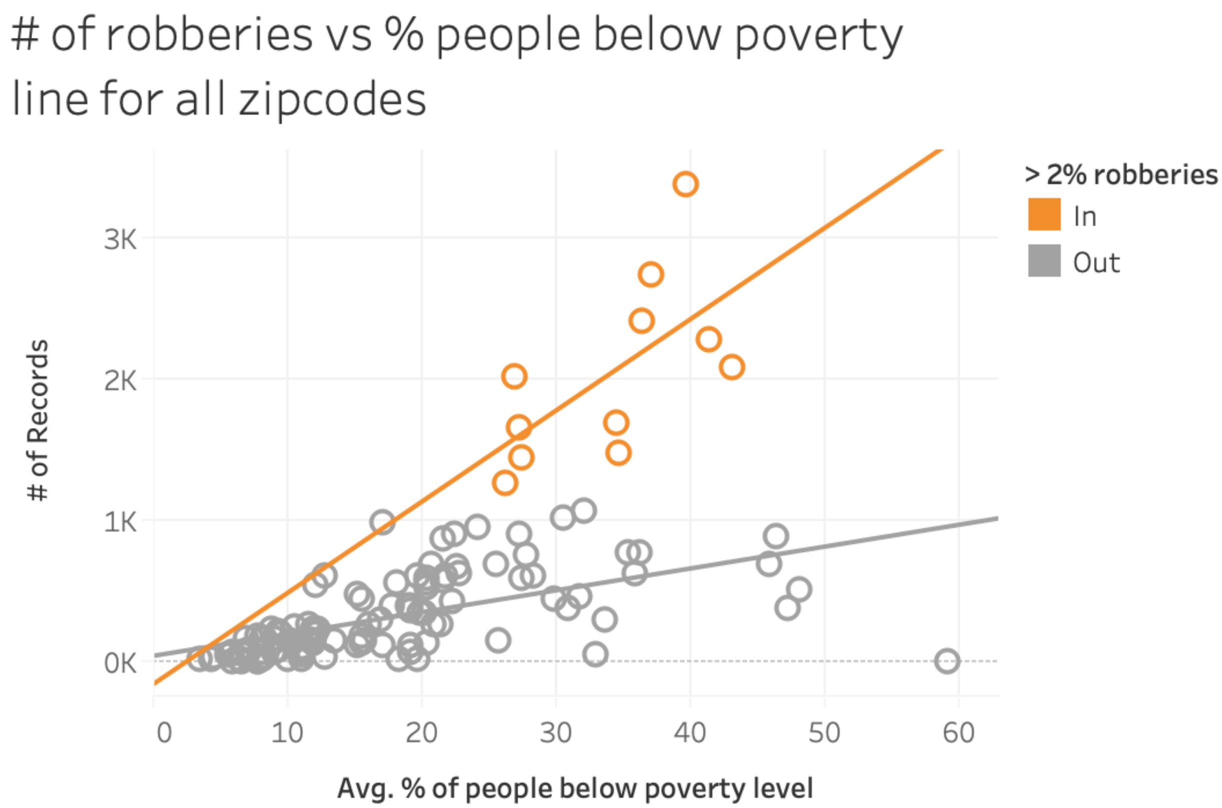


Figure 14: *Correlation plot between number of reported crimes and the percentage of people bellow the poverty line to zip-code areas with over 2% of all the reported aggravated assaults*

Average of % of unemployed people in labour force (Employment_status_all_5YR_B23001_with_ann) vs. sum of Number of Records. Color shows details about In / Out of Zipcodes w > 2% of the aggravated assaults. Details are shown for Zipcode (Employment_status_all_5YR_B23001_with_ann). The data is filtered on Crime Code Description and Date Occurred Year. The Crime Code Description filter keeps ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT. The Date Occurred Year filter keeps 7 members.

Figure 15: *Correlation plot between number of reported crimes and the percentage of people unemployed to zip-code areas with over 2% of all the reported aggravated assaults*



Average of % of people below poverty level (Poverty_status_all_5YR_B17001_with_ann) vs. sum of Number of Records. Color shows details about In / Out of Zipcodes w > 2% robberies. Details are shown for Zipcode (Poverty_status_all_5YR_B17001_with_ann). The data is filtered on Crime Code Description and Date Occurred Year. The Crime Code Description filter keeps ROBBERY. The Date Occurred Year filter excludes 2010, 2018 and 2019.

Figure 16: *Correlation plot between number of reported crimes and the percentage of people bellow the poverty line to zip-code areas with over 2% of all the reported robberies*

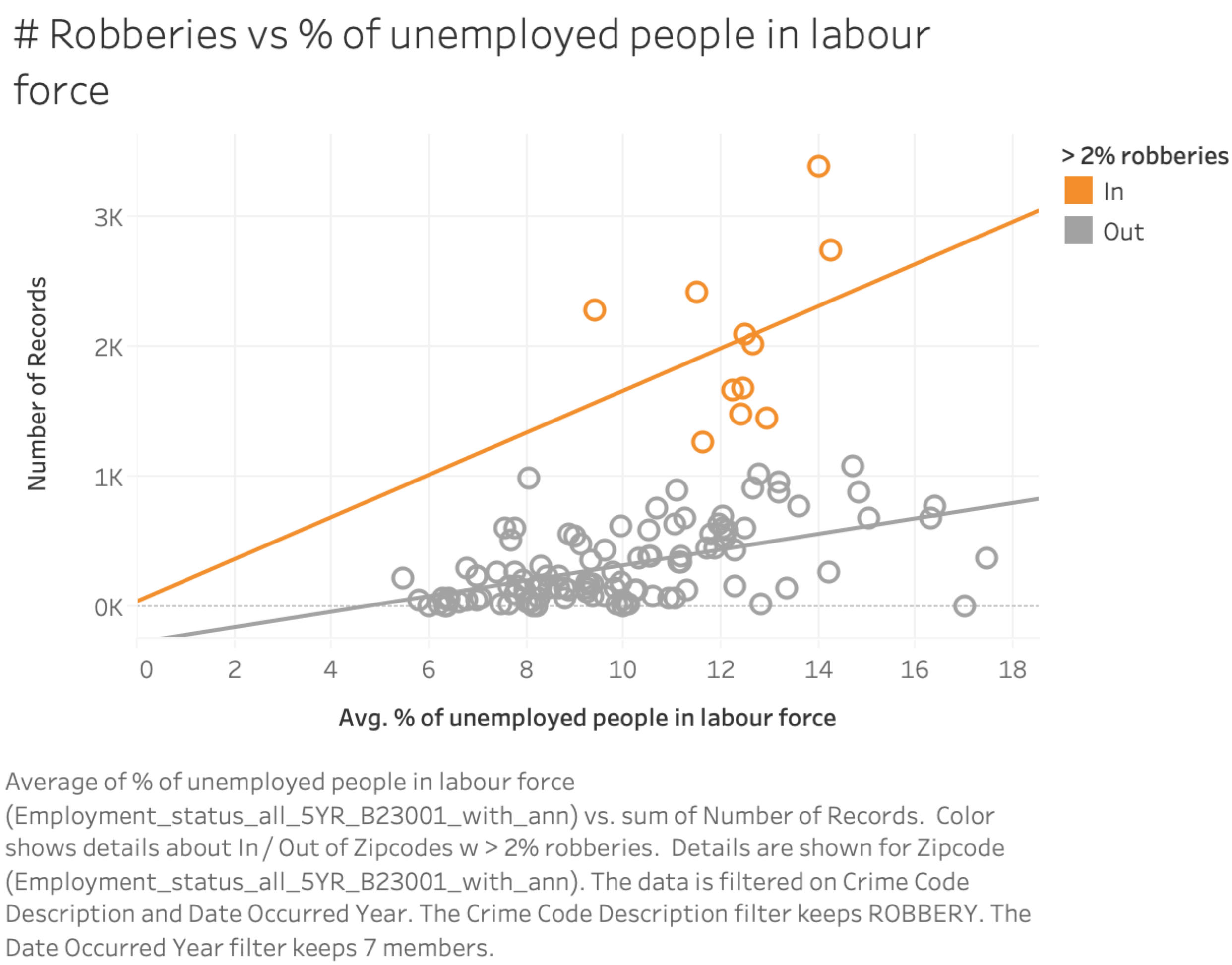
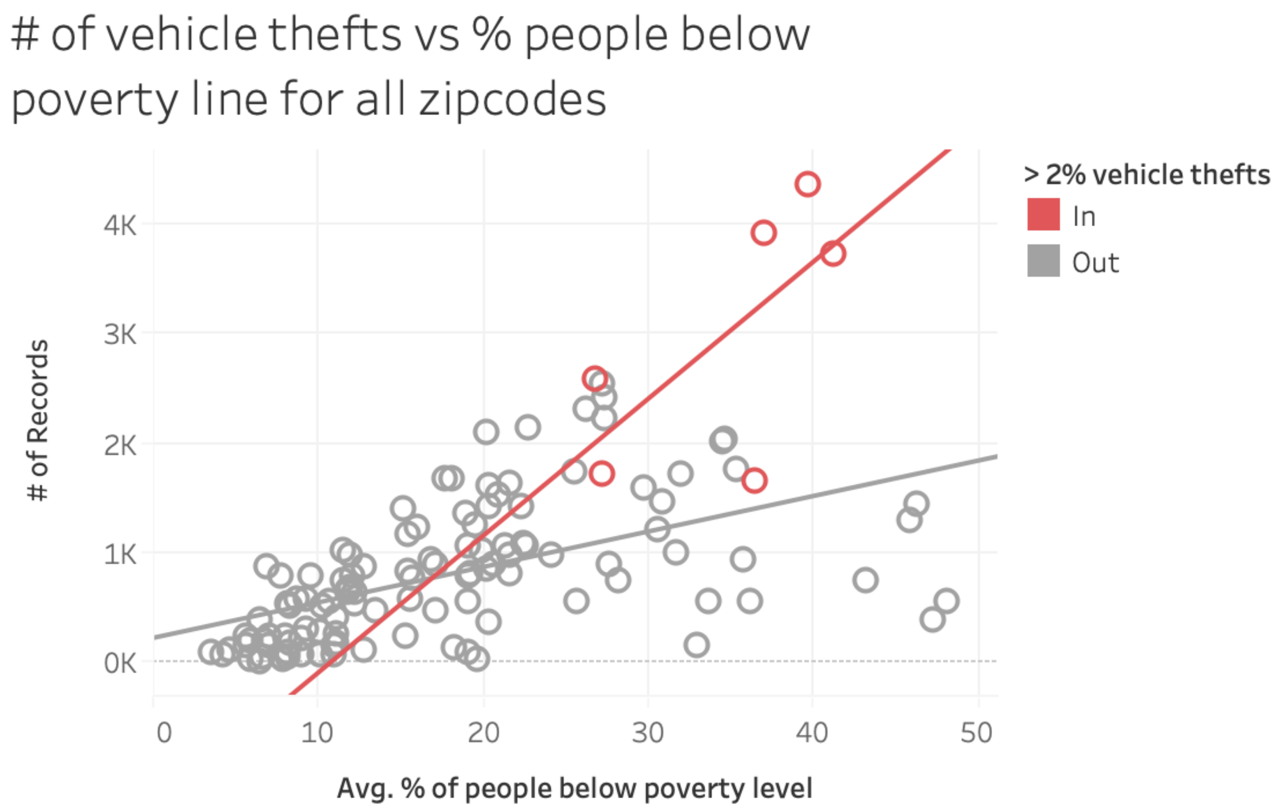
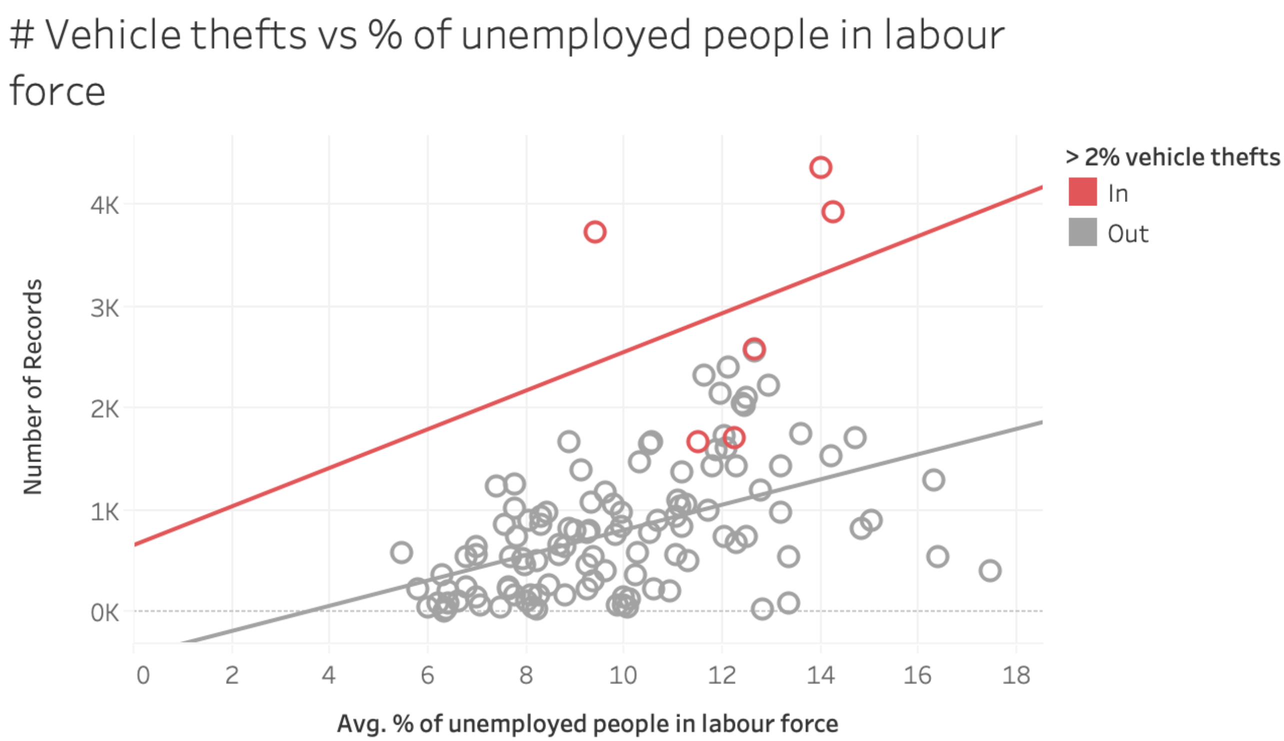


Figure 17: *Correlation plot between number of reported crimes and the percentage of people unemployed to zip-code areas with over 2% of all the reported robberies*



Average of % of people below poverty level (Poverty_status_all_5YR_B17001_with_ann) vs. sum of Number of Records. Color shows details about In / Out of Zipcodes w > 2% vehicle thefts. Details are shown for Zipcode (Poverty_status_all_5YR_B17001_with_ann). The data is filtered on Crime Code Description and Date Occurred Year. The Crime Code Description filter keeps VEHICLE - STOLEN. The Date Occurred Year filter excludes 2010, 2018 and 2019.

Figure 18: *Correlation plot between number of reported crimes and the percentage of people bellow the poverty line to zip-code areas with over 2% of all the reported vehicle thefts*



Average of % of unemployed people in labour force (Employment_status_all_5YR_B23001_with_ann) vs. sum of Number of Records. Color shows details about In / Out of Zipcodes w > 2% vehicle thefts. Details are shown for Zipcode (Employment_status_all_5YR_B23001_with_ann). The data is filtered on Crime Code Description and Date Occurred Year. The Crime Code Description filter keeps VEHICLE - STOLEN. The Date Occurred Year filter keeps 7 members.

Figure 19: *Correlation plot between number of reported crimes and the percentage of people unemployed to zip-code areas with over 2% of all the reported vehicle thefts*

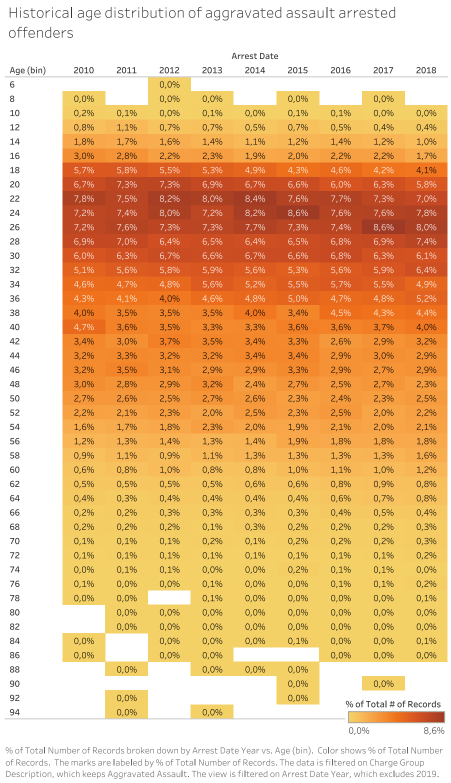


Figure 20: *Historical evolution of the age distribution of aggravated assault arrestees*

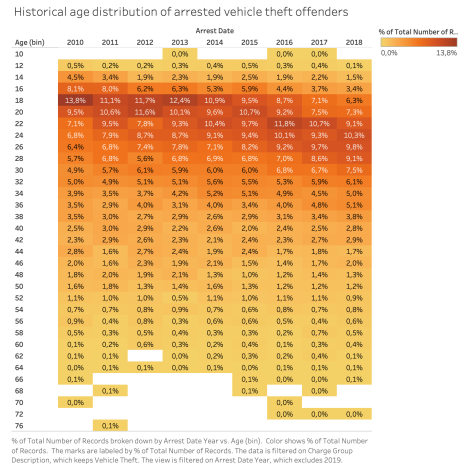


Figure 22: *Historical evolution of the age distribution of vehicle theft arrestees*

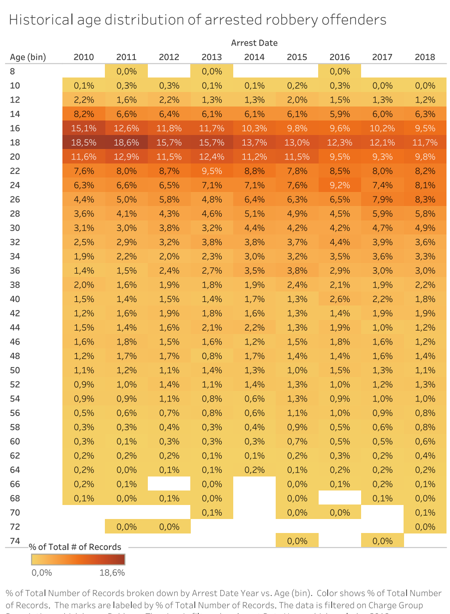


Figure 21: *Historical evolution of the age distribution of robbery arrestees*

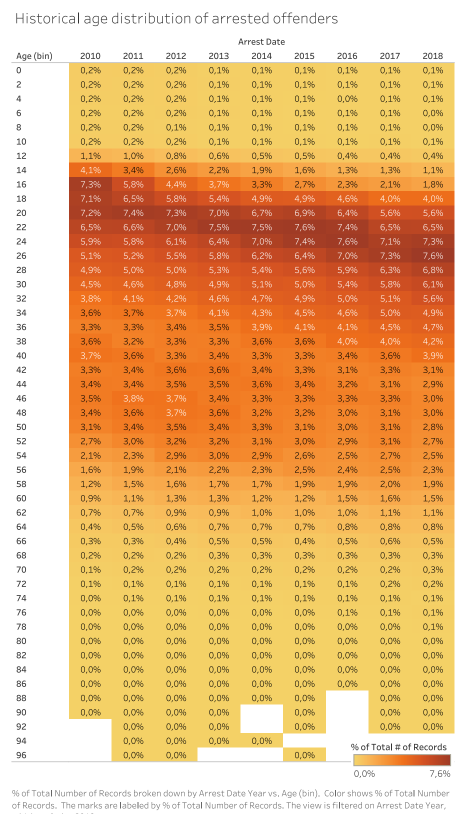


Figure 23: *Historical evolution of the average age distribution on the three types of crime studied*

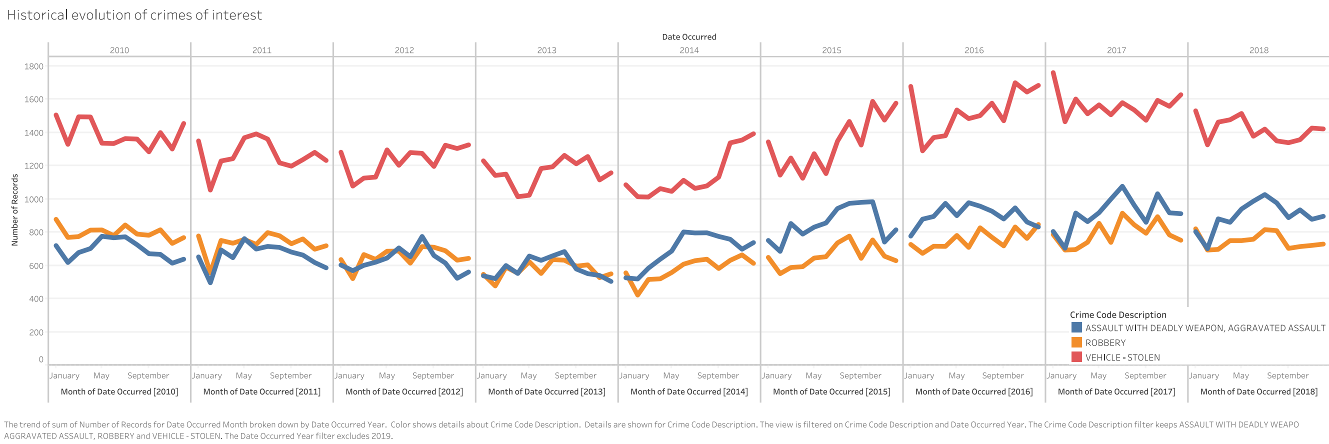


Figure 24: *Historical evolution of the seasonality*

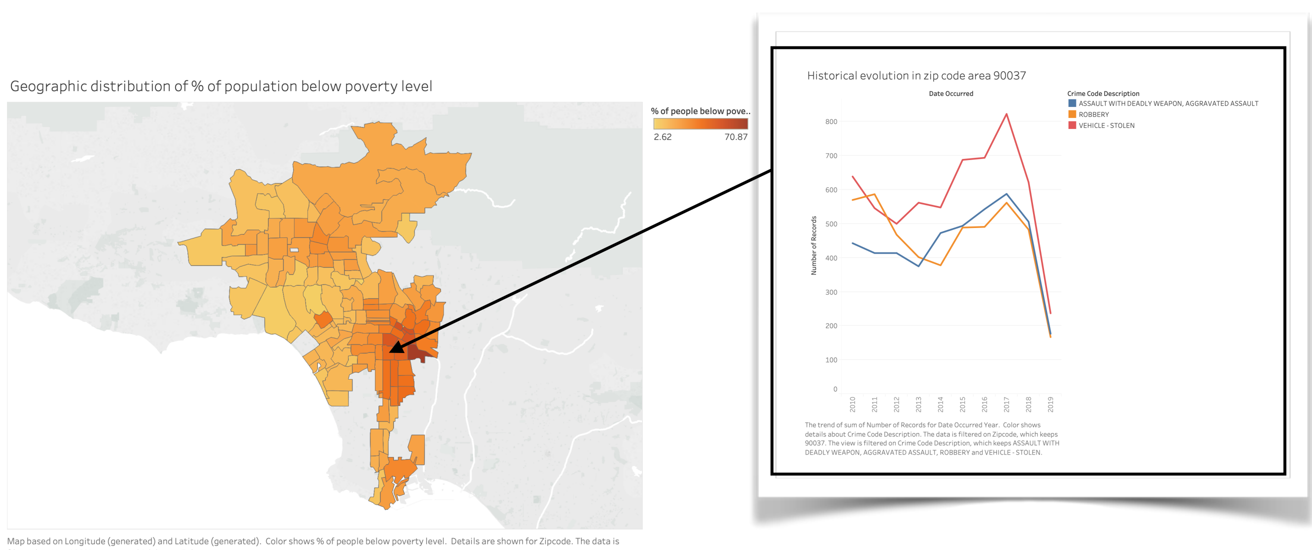


Figure 25: *Sketch of interactive visualization. User clicking on a zip code area leads to the appearance of a new navigation window where time evolution of different measures filtered by selected zip code can be seen.*

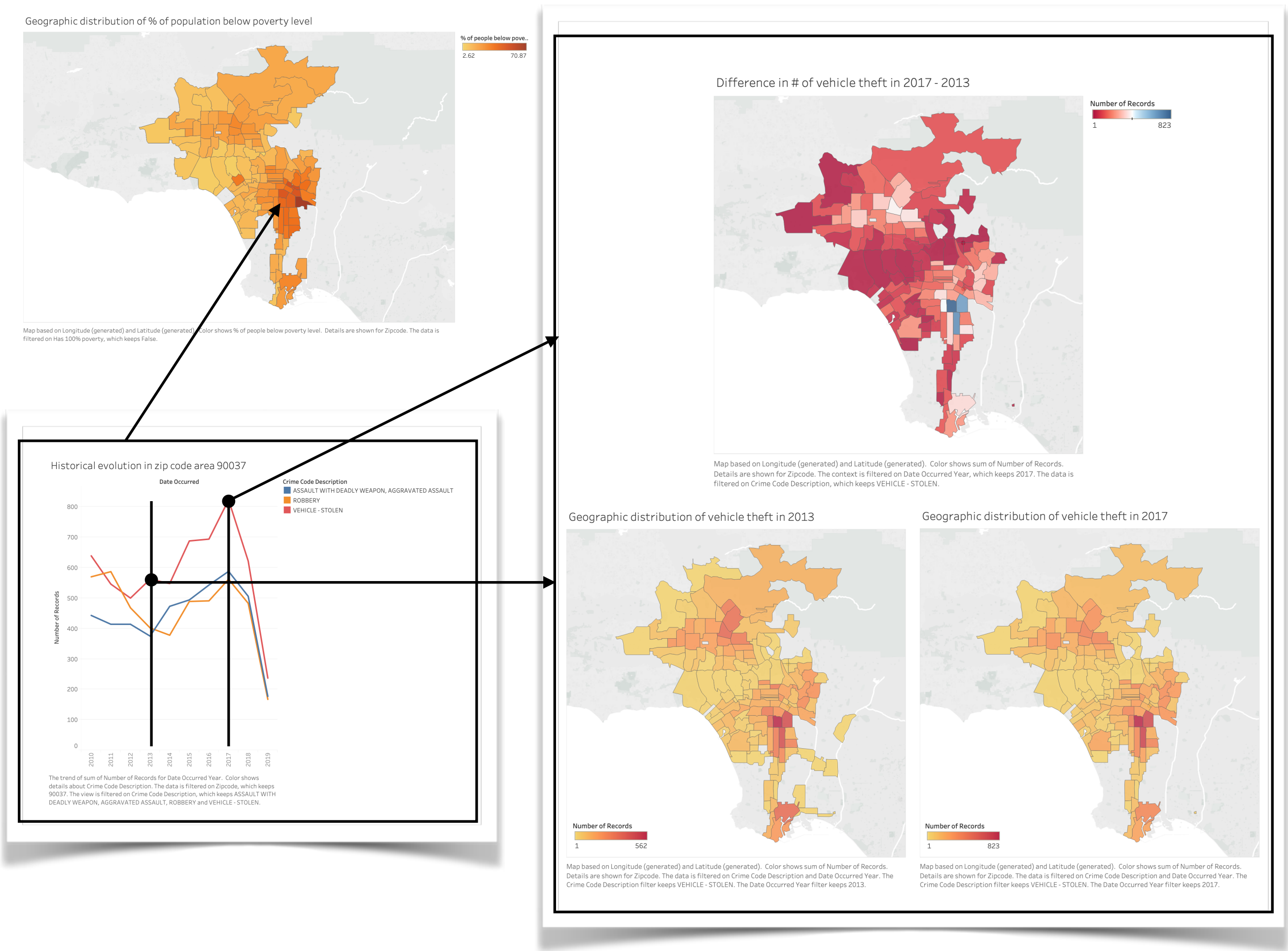


Figure 26: *Sketch of interactive visualization with selection of two years (Y1 and Y2). When the user selects two years in the extra navigation panel for a specific measure, a new view appears in which this measure is presented in two maps, one for each of the selected years. In addition, a third map in the same view shows the difference of this measure between year Y2 and Y1.*