# Data Visualization on U.S Crime Data through 1960-2014

P.Anand Sagar

11197418

## Introduction:

Crime is a major issue in the United states since decades. Every day we hear about them. To get to know about them is not what that is important, we need to use modern technology and data science tools to wisely act against this problem. Crime classifications are based upon preliminary information supplied to the police department by the reporting parties or agencies and maybe changed later based upon investigation. Applying visualization to these data bring us valuable information that can be used to increase the safety of our society and lower the crime rate. Data visualization is an efficient means to represent distribution and structures of huge datasets and find out hidden structures in the data.

## Motivation:

This data set has been chosen with a motive that analyzing this dataset and visualizing them to draw some insights can lead to public safety. This data is selected with an intention to find out the answer whether the crime rate in U.S is rising or falling. It can also help Police officials justify the allocation of resources needed to develop a crime analysis function in their departments. This project is carried out with an aim to get a better understanding of the security status in the United states. Using this data, one can get awareness about the crime incidents happening at a particular location, so that effective measures can be taken to ensure public safety.

## Methods:

Here I have used the Tableau software system. The visualizations created are in the form of dashboards and worksheets. It converts raw data into easily understandable format, and It is particularly suited to the users without having any coding knowledge and it is best while working with large amounts of data. Using Tableau, we get an overall idea about trends in a dataset. Here the complex ideas hidden in the dataset are communicated statistically with clarity and approach has been made to be precise.

Various techniques have been used for visualization of the data. They are:

Tree maps - It provides hierarchical structures which may be helpful in displaying numerous variables. The colorful presentations in tree maps can accommodate hundreds or thousands of items in a meaningfully organized display that allows patterns and exceptions to be spotted in seconds.

Bar charts – It illustrates comparisons among different items. It maps the data using horizontal or vertical rectangular bars and this chart is best suited when data is arranged in column and row wise manner on a worksheet.

Bubble charts – These charts are extremely useful to represent data as 3-d and to compare the relationship between items. 3-dimensions are X-axis Y-axis and data represented as bubbles. Bubbles are usually differentiated by color, pattern, number or name labels, or a combination of these.

Pie chart - Pie Charts help show proportions and percentages between categories, by dividing a circle into proportional segments. You can use pie charts to show relative sizes of many things.

## Data description:

The data describes the crime information recorded in US from 1960 to 2014. This data is according to FBI's crime reporting statistics. This data consists of following columns.

**State** in which crime occurred, Year( from 1960 to 2014), **Population** of that state

**Violent crime total** – It represents number of violent crimes occurred. This column represents the sum of following four columns.

- **Murder and nonnegligent Manslaughter**
- **Legacy rape /1**
- **Robbery**
- **Aggravated assault**

**Property crime total** – It represents the incidents which has affected the property by some means such as theft. This column represents the sum of following three columns.

- **Burglary**
- **Larceny-theft**
- **Motor vehicle theft**

This data not only represents the number of incidents occurred, but also it provides the crime rate for the above mentioned columns. i.e., we also have columns like **Violent Crime rate, Murder and nonnegligent manslaughter rate, Legacy rape rate /1, Robbery rate, Aggravated assault rate, Property crime rate, Burglary rate, Larceny-theft rate, Motor vehicle theft rate.**

The crime rate is calculated as crime figures in each category/ Total population of that state per 1,00,000 people. It is simply average rate per 100000 people.

The data consists of 2800 rows and 21 columns. This data has been collected by FBI from many agencies

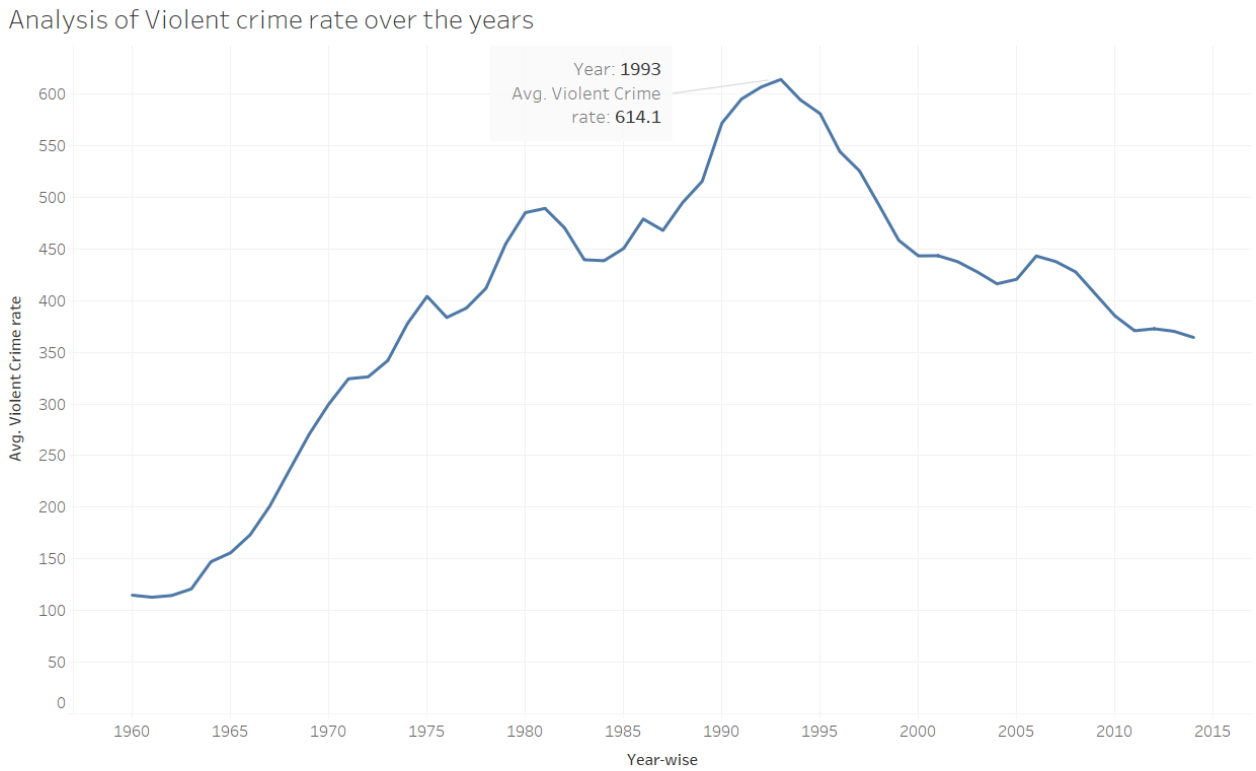
## Problem definition:

To understand the crime context in the united states, here are the questions which I have answered through visualization.

1) What are the insights on violent crime?

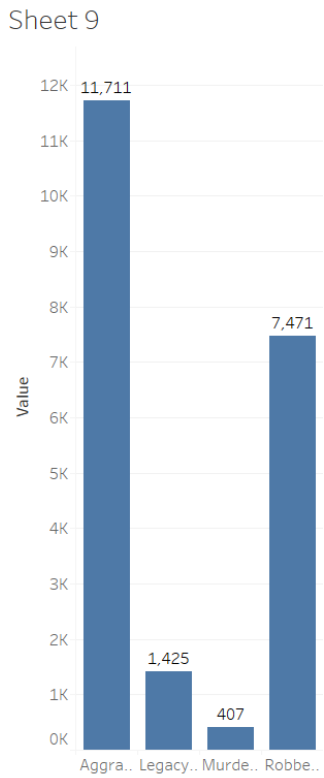
Sol: This question has been answered by creating three scenarios. They are:

a) Average violent crime rate over the years:




For this visualization, Avg. violent crime rate Is taken on Y-axis i.e., as rows and Years are taken on X-axis. This curve shows the trends in total Violent crime rate beginning from 1960 to 2014. Here, Violent crime is referred in the sense that it includes the Murder and nonnegligent Manslaughter, Legacy rape, Robbery, and Aggravated assault. From the curve, we can say that the violent crimes have occurred more in the year 1993 with a rate of 614.1. The crime rate was very low in years of 1960's. We can say that violent crime rate took a higher pace in years between 1985-1995.

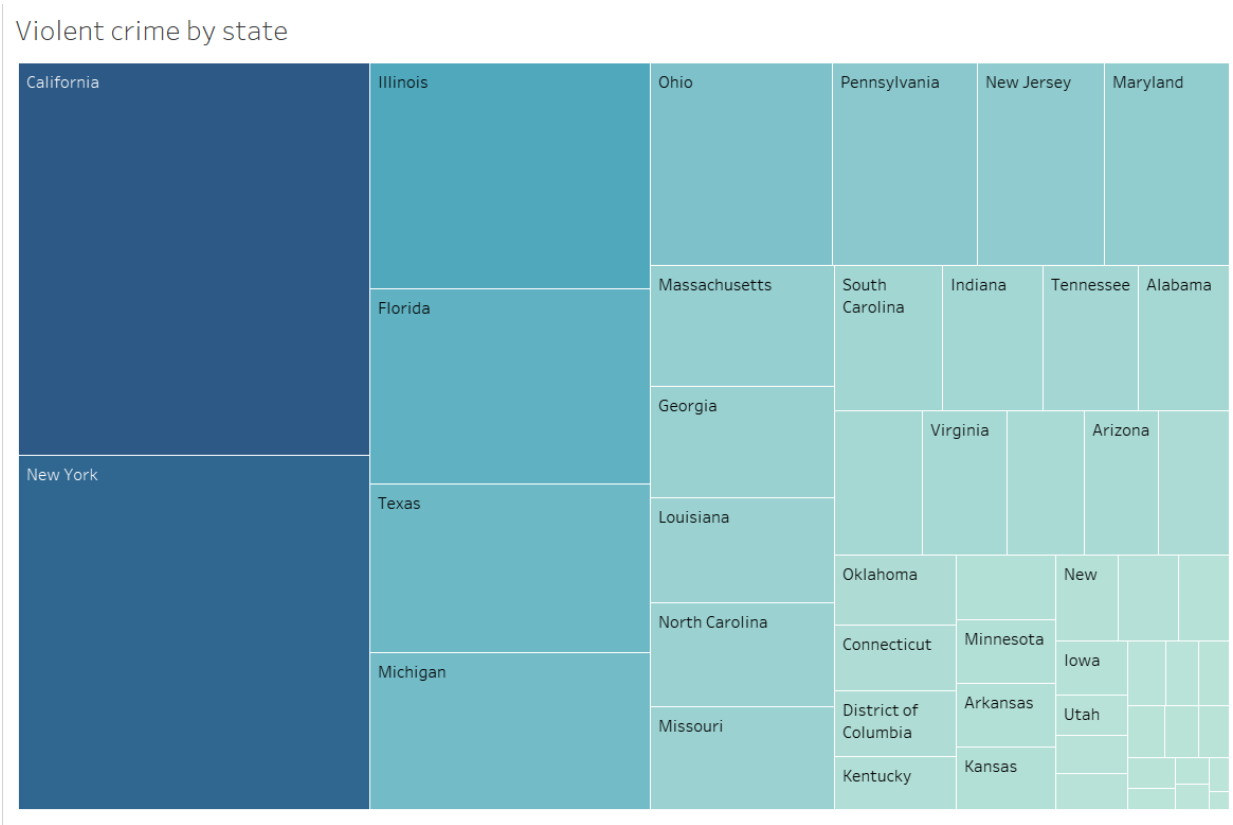
b) <u>Breakdown of Violent crime by crime type:</u>




Aggravated assault rate, Legacy rape rate /1, Murder and nonnegligent manslaughter rate and Robbery rate. The data is filtered on Action (Year), which keeps 1 member.

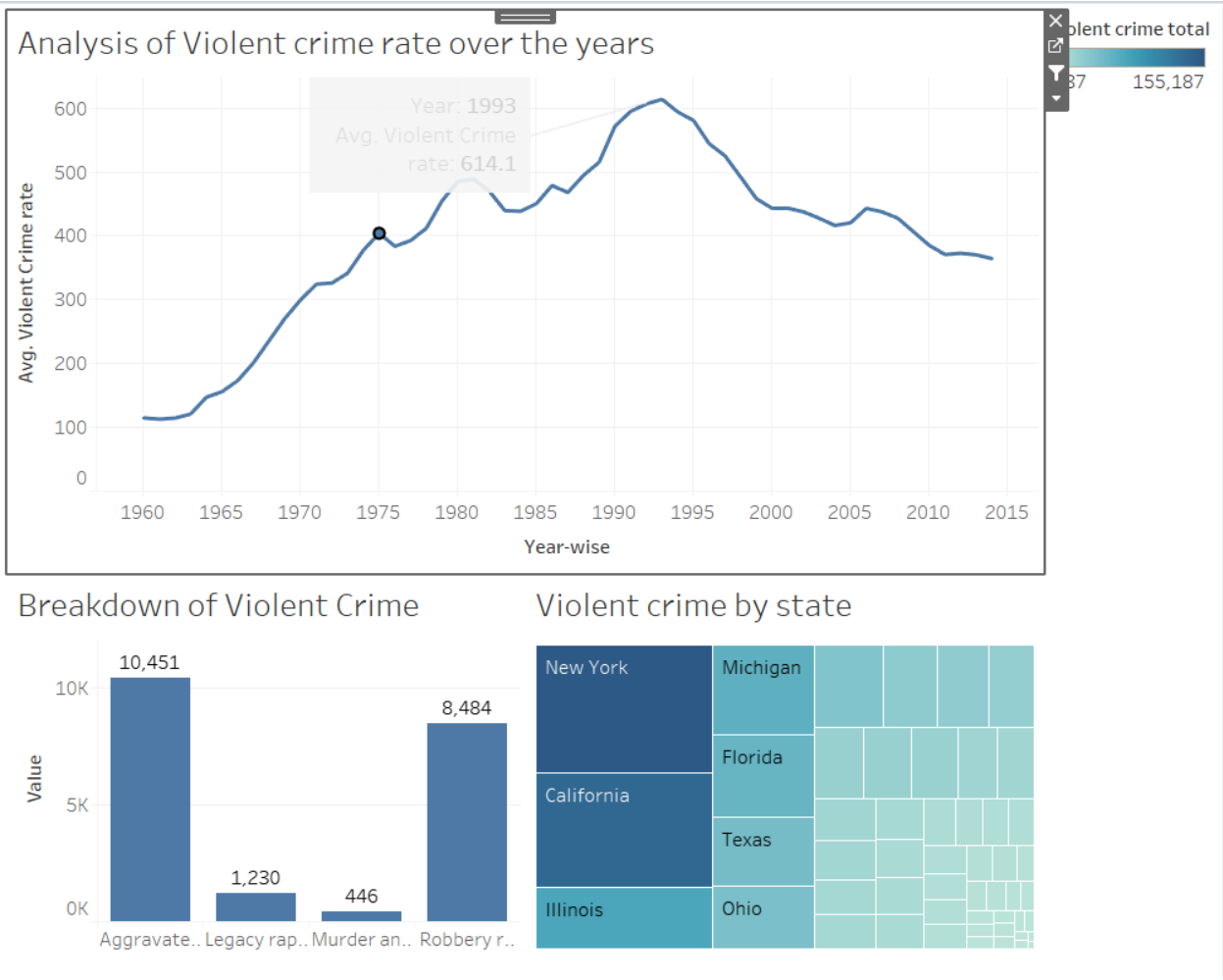
As discussed, Aggravated assaults, Legacy rape, Murders and Robbery crimes come under the category of violent crime., the above bar chart shows the figures of these crimes independently over the years. We can say that Aggravated assault rate is very high (11,711), followed by Robbery(7471) . The least violent crime occurred in US is Murders.

c) Breakdown of Violent crime by states: Keeping in view to convey the readers, in which state the violent crime was more prevalent, the following heat map has been created to understand clearly. It is clear that Violent crimes were high in California, followed by New York. And also, these crimes occurred almost same in Illinois, Florida, Texas and Michigan




Interactive visualization: For the above graphs, I have created the interactive visualization so that users can have more access and interact with the environment. For instance, if a user wants to see the violent crime data that took place in 1975, he can easily grasp the context he wants by looking at interactive visualized data.

One can understand that, in 1975, violent crime incidents are high in New York and Aggravated Assaults occurred mostly than any other Violent crime types.

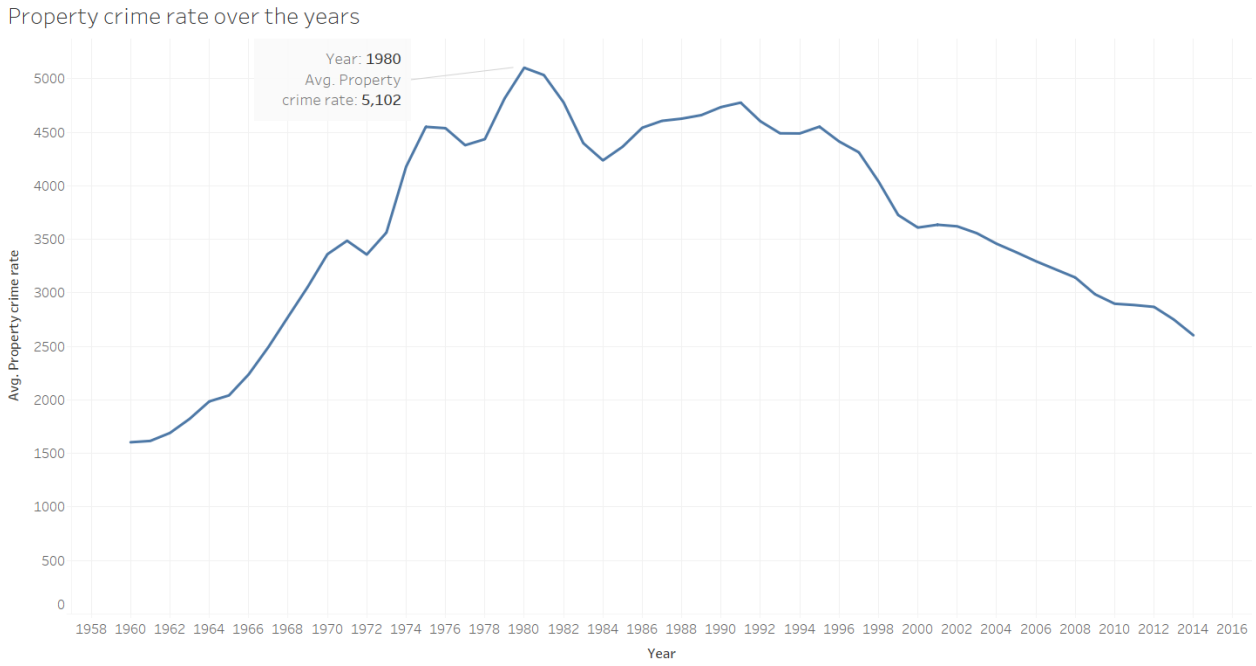



2) What are the insights on Property crime?

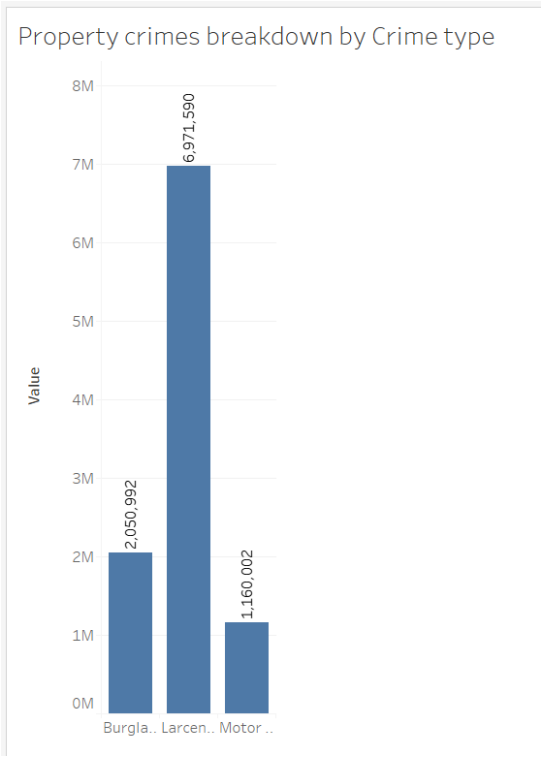
Sol: Property crimes usually involve private property at the incident such as Burglary, Vehicle theft and Larceny theft etc. Let us see some visualizations with respect to property crime.

a) <u>Average violent crime rate over the years:</u>

Similar to Violent crime rate, visualization has been crated to represent the trends of property crime over the years. We can say that Property crime rate was high in between 1975-1995. In the recent times, the curve has seen a downfall which indicates that the safety measures have been increasing.

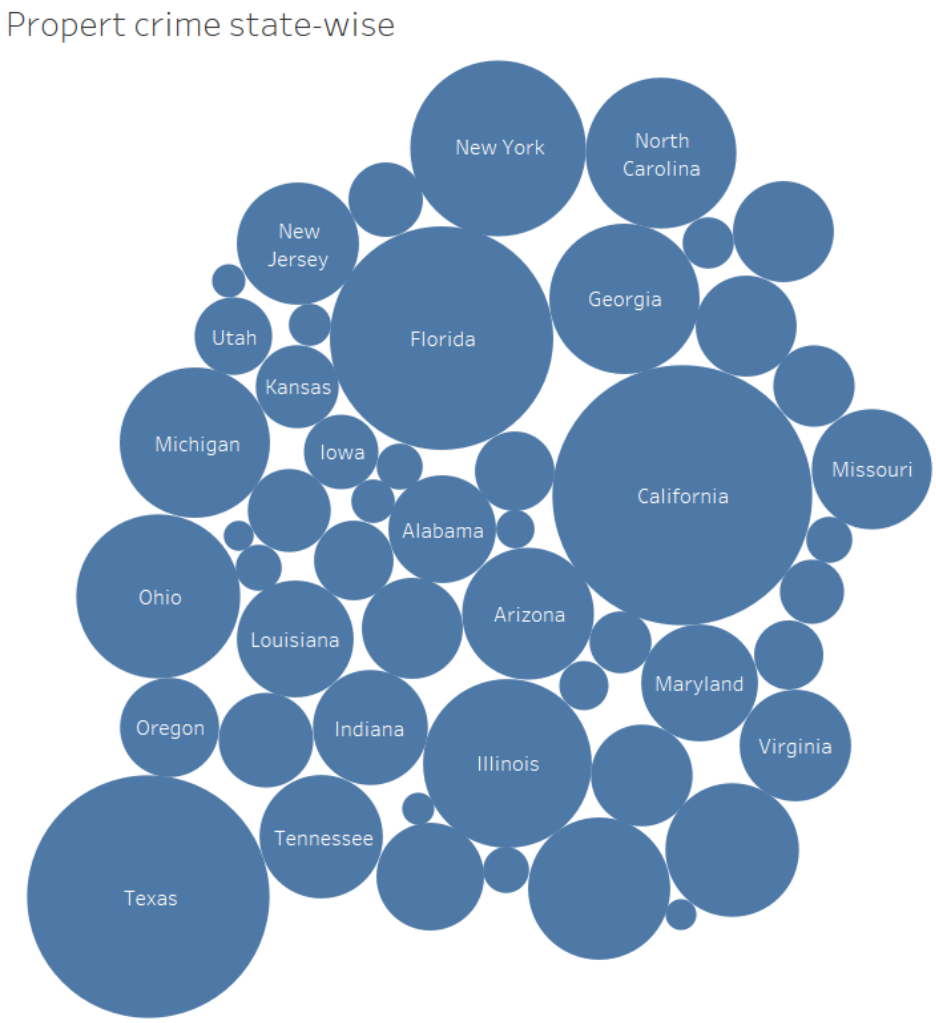



b) <u>Breakdown of Property crime:</u>




Property crime can be categorized into Burglary, Larceny theft and Motor Vehicle theft. We can see that Larceny theft is highest(69,71,500) compared to Burglary and Motor vehicle theft incidents.

c) Property crime by state -wise

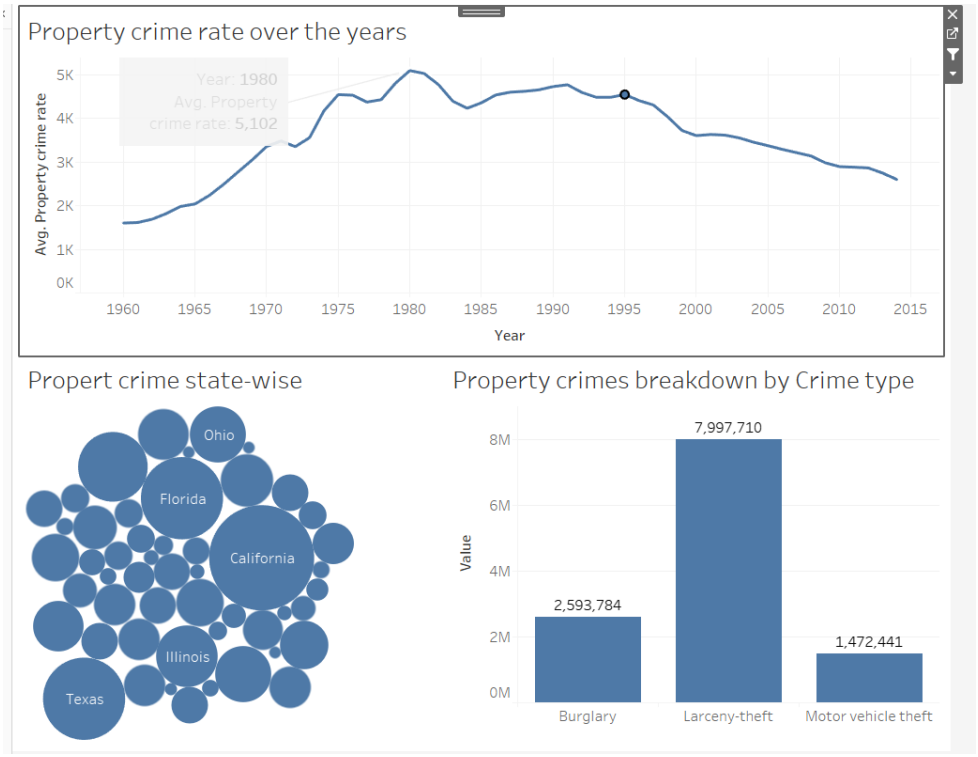



State. Size shows sum of Property crime total. The marks are labeled by State. The data is filtered on Action (Year), which keeps 1 member.

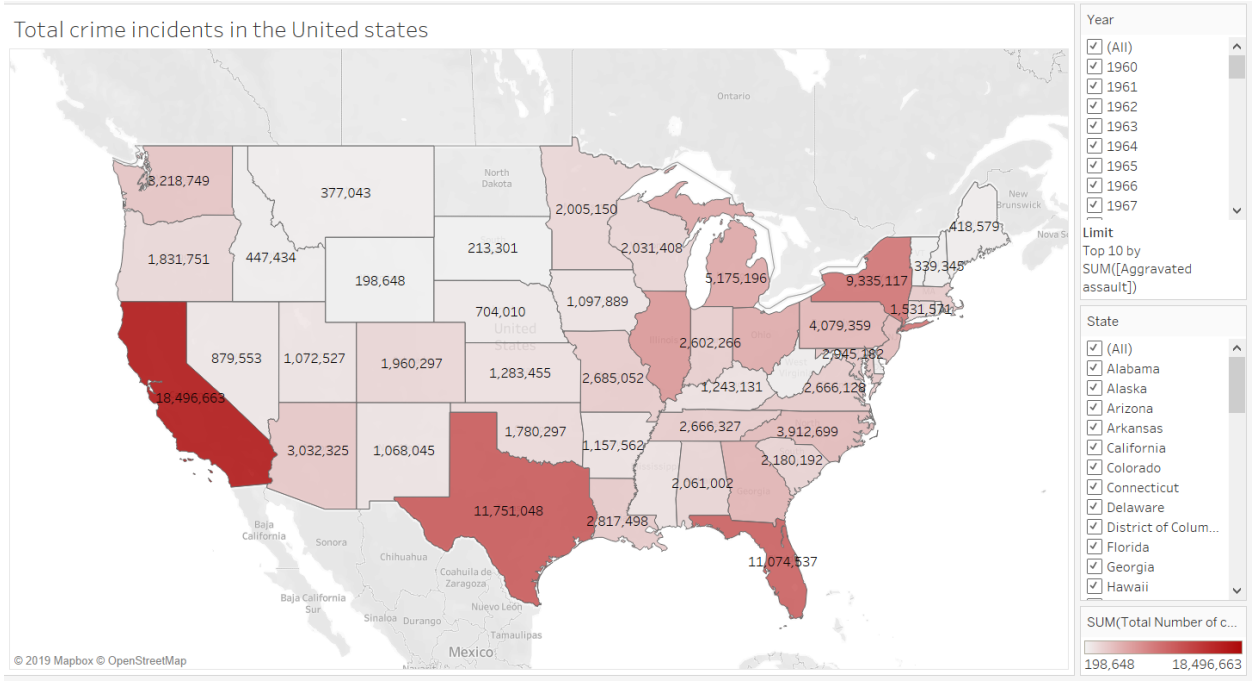
A bubble chart has been created to view the property crimes occurring in each state. It is clear that California, Texas and Florida ranks as the top 3 states in the context of Property crime.

Interactive visualization for property crime:

The user can play on this data by hovering over the graphs and finding some useful information.




3) In which state, more safety measures have to be taken to prevent crime?



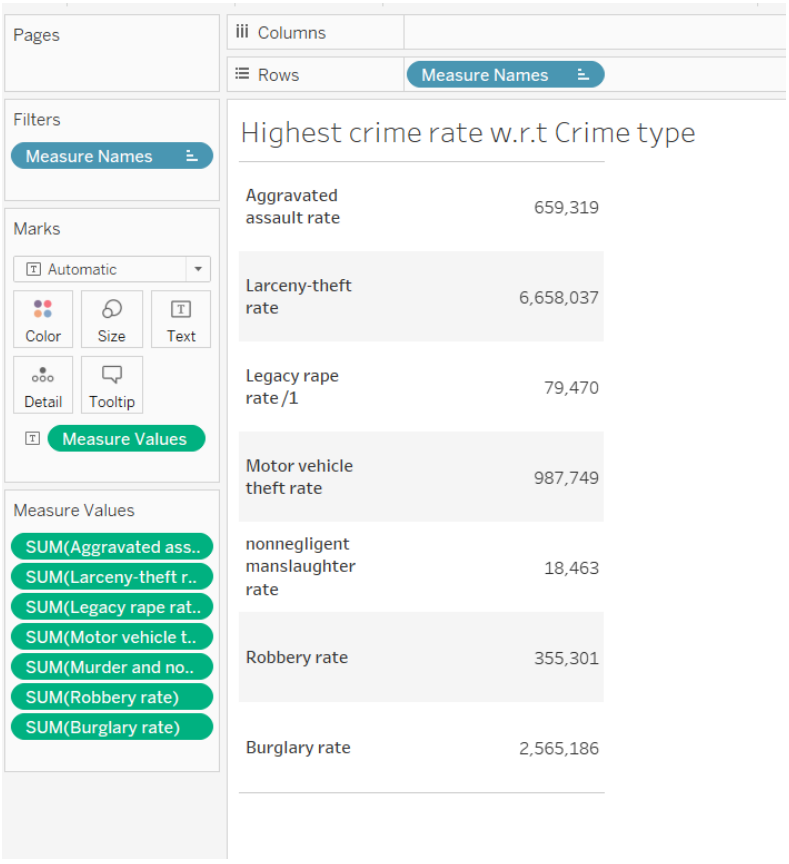
To visualize this, a separate calculated field has been created.

Total number of crimes = [Aggravated assault] + [Burglary] + [Larceny-theft] + [Legacy rape /1] + [Motor vehicle theft] + [Murder and nonnegligent Manslaughter] + [Robbery] .

Using this calculated field, map has been created to represent the total crimes in the United states through 1960-2014. Red color has been chosen to replicate the crime rate. Light color for the state with lower crime and dark color for the state with the greatest number of crimes. We can say that, over the years, California has experienced highest crime scenes and we can draw an insight that safety measures have to be increased here keeping in view for the public safety. The numbers here are really high IN California with crime incidents crossing over 18 M, and then followed by Texas with 11M incidents.

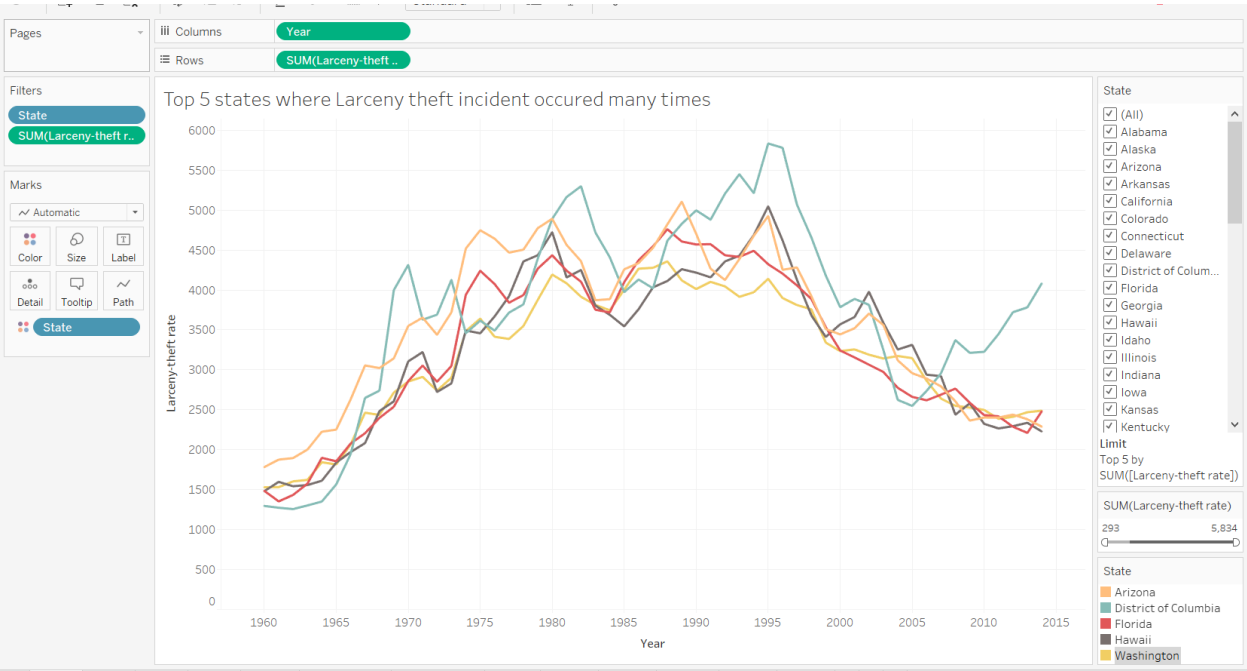
4) Which type of crime has the highest crime rate among all, and how are the trends varying for the top 5 states of that crime?

Sol: To know what the highest crime rate is, a table chart has been created with types of crime as rows. They are Aggravated Assault rate, Larceny-theft rate, Legacy rape rate, Motor-vehicle theft rate, Murder and nonnegligent manslaughter rate, Robbery rate and Burglary rate.




We can see that Larceny theft rate is 66,58,037, which is the highest. The rate is calculated per 100000 people.

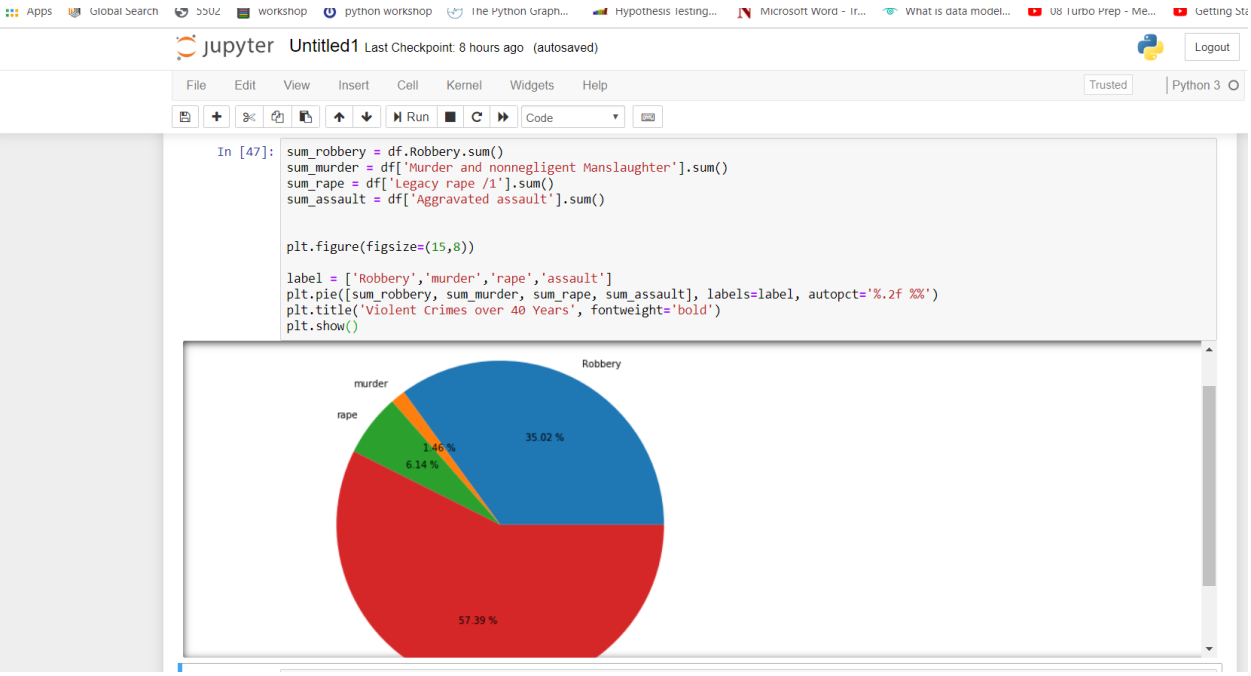
Now, I have created a visualization of the top 5 states with this crime type. i.e., Top 5 states in which Larceny theft was frequently occurred.




The insights from this graph are that Arizona, District of Columbia, Florida, Hawaii and Washington are the states where Larceny theft rate occurred frequently through 1960-2014. For all the states, the peak rate was in years between 1990-2000.

<u>Bonus – Visualization in Python</u>

Pie chart for visualizing violent crimes over 40 years:



The pie chart shows the proportionate of violent crimes in U.s and it is clear that Aggravated Assaults are most common.

## Findings from the data:

The audience can get to learn the following facts from the information.

Crime rates have varied over time, reaching the peak in between 1970's and 1990's and then the crime has declined majorly. The highest crime rate was in mid-1990's.

The audience can be able to differentiate between the type of crime incidents such as which violent crime type was highest or lowest, and how were the trends of violent crime types and property crime types.

From this visualization, Aggravated assaults were higher in case of Violent crimes and Larceny theft ranks the top among Property crimes. One can say that over the years property crimes were prevalent in United states compared to Violent crimes.

Also, the user can go through interactive dashboard to get the crime information at a particular state and respective year. They can have a glance at state-wise Crime information. This visualization tells that California, New York and Texas are the states where crime incidents are occurring mostly.

Related work – There is a huge amount of crime data and lot of projects are doing a research in this area due to a lot of insecurities rising and public safety is a more concern. Similar projects include 911 police system in Dallas city area, Creating tableau visualizations for Crime in Cincinnati, Chicago crime rate analysis etc.

## Future work:

This dataset can be updated later by including the incident reports city-wise and also adding crime incident information such as incident number, exact location of the incident with Latitude and Longitude, details of Accused and victim information and to collect more meaningful data to standardize the crime reporting system and such that the user must be satisfied with the good usability and interactivity

## References:


[1] J. A. Shaffer, "Creating a Tableau Data Visualization on Cincinnati Crime".

[2] s. Tafzoli, "My notes on Chicago Crime data analysis," 2018. [Online]. Available: https://medium.com/@stafa002/my-notes-on-chicago-crime-data-analysis-ed66915dbb20.

[3] A. M. R. Enrico G. Caldarola, "Big Data Visualization Tools: A Survey".