**Data Analysis**

**COMP – 8157 Advanced Database Topics**

**Submitted By: Aakarsha Mahajan**

**Student ID: 105103000**

1. **Title and Author of the chosen dataset:**

The title of the dataset I chose is: **Arrest Data – Reported Number of Arrests by Crime**

The author of the chosen dataset: **Crime Data Explorer for Federal Bureau of Investigation (USA)**

1. **Research Questions:**

* **What has been the general trend of the arrests of the various crimes conducted across the country from 1995 -2016?**

During my research, I found out the general trend of arrests over the past 22 years. This data is definitely the most important as this helps the government to analyze the various loopholes which might have occurred in reporting arrests of various crimes. The more number of arrests either imply the arrests reported for various crimes have tremendously varied.

* The period of 1995-1997 observed the maximum number of arrest. As per reports, approx. 15,284,300 arrests were made in the year 1997, which is the highest is the span of 22 years. The years 1996 and 1995 took the second and third spot respectively
* The number of arrests started descending from the year 1998 onwards and the year 1998 and 1999 saw a massive decrease.
* The 21st century however maintained a steady pace with the decrease in the number of arrests till the year 2003.
* But the years 2005-2006 saw a significant increase in the number of arrests and showed that more and more agencies were reporting the arrests.
* From 2007 to 2016, the count kept on going down.
* **Explain the relationship between population and total number of arrests for a particular year.**

In this particular question, I am trying to analyze and draw a comparison between the population and the number of arrests is a particular year.

* Keeping a track of the relationship between the population and the number of arrests, help us in computing the ‘Arrest Ratio.’
* The arrest ratio is nothing but the number of arrests made by the law enforcement agencies per 100,000 population or per the 100,000 population considered to be at risk for arrest..
* The arrest ratio is calculated by dividing the total number of reported arrests by the respective population.
* **How can you compare the total number of arrests with the arrests of a particular crime?**

Here we are trying to analyze the crime for which the maximum arrests have been made for a particular year.

* It seems that disorderly conduct and aggravated assaults have been very common and thus the arrests for them are among the highest.
* The state and local law enforcement agencies have been making estimates from 1995 to 2016 for each of the mentioned crimes in the data set and they make sure to determine the sex, the age group and the race of the people arrested for each of the crimes.
* This helps in making sure that a particular section of society is not being targeted.

1. **DataSet: The following data set has been used for the analysis.**

**URL:** <https://crime-data-explorer.fr.cloud.gov/downloads-and-docs>

**Arrest Data – Reported number of arrests by crime.**

**This particular dataset describes:**

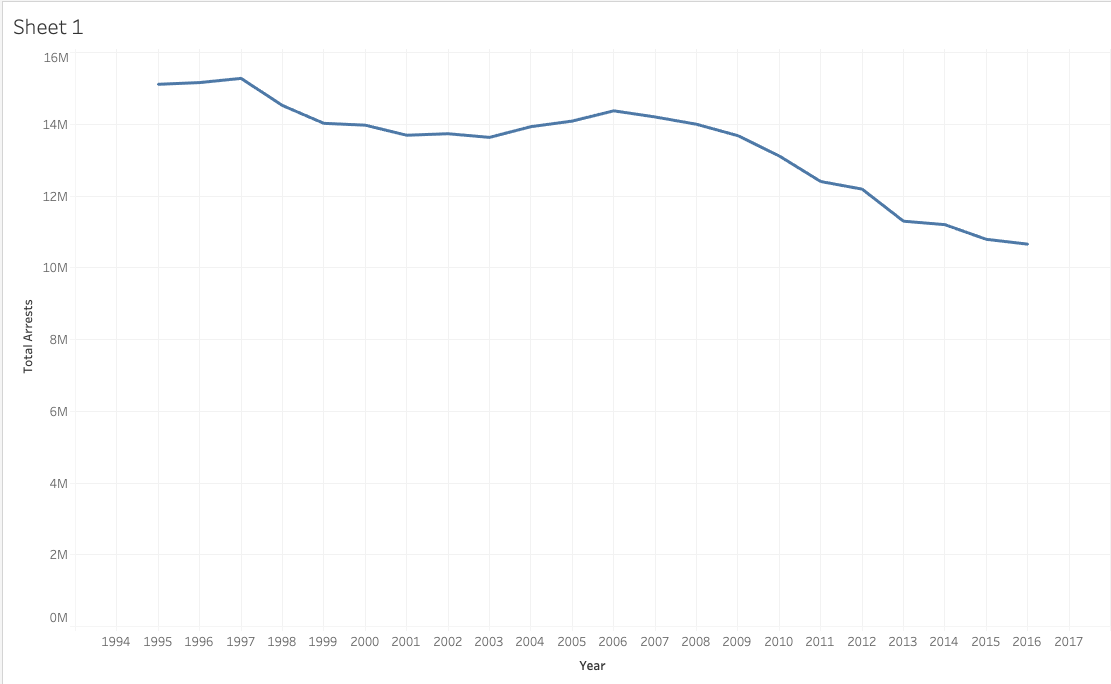
This dataset contains the monthly number of reported arrests for various offenses reported by participating law enforcement agencies from 1995.

The FBI’s Uniform Crime Reporting (UCR) Program collects counts of arrests (including citations and summonses) in several offense and demographic categories from participating state and local law enforcement agencies

This report expands the FBI’s set of published national arrest estimates to include estimates of arrests in various age, sex, and race groups within offense categories.

1. **Chart/Graph: The following charts have been prepared:**

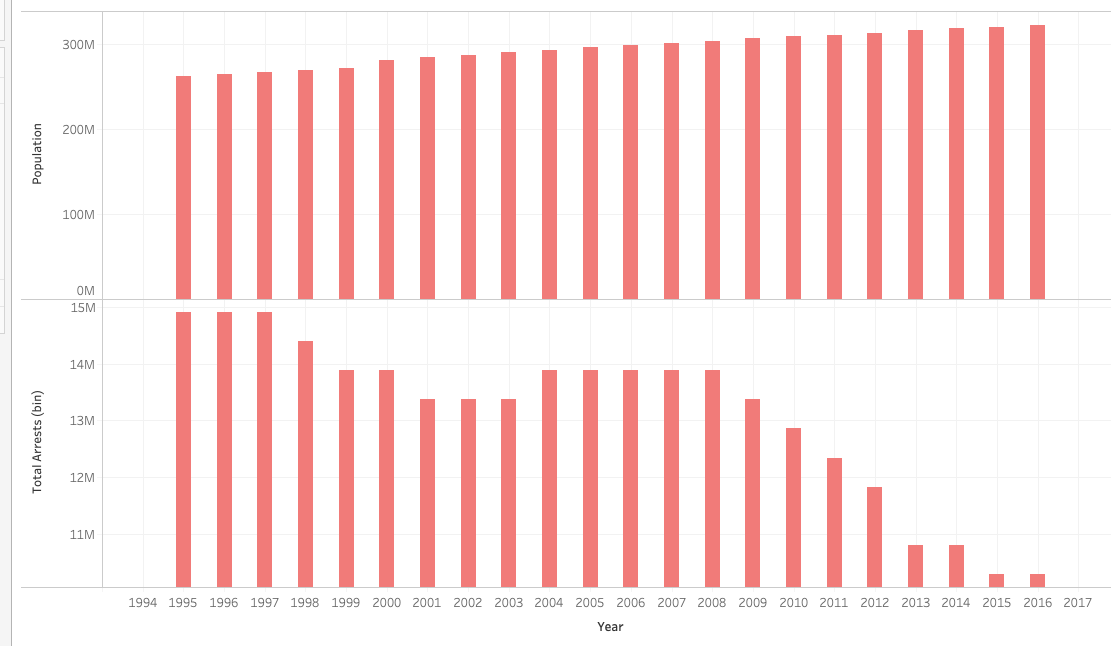
**Graph 1:**

****

**Description of Graph 1**:

* The above graph is used to show the total number of arrests made during each of the years for a span of 22 years. As the time began, the number of arrests was quite high in number showing that various law enforcement agencies were reporting the crimes that were conducted.
* As the time elapsed, the number of arrest gradually reduced and then rose again, during the mid of 21st century. And then the number of arrests was the lease during the 2014, 2015 and 2016.

**Graph 2:**

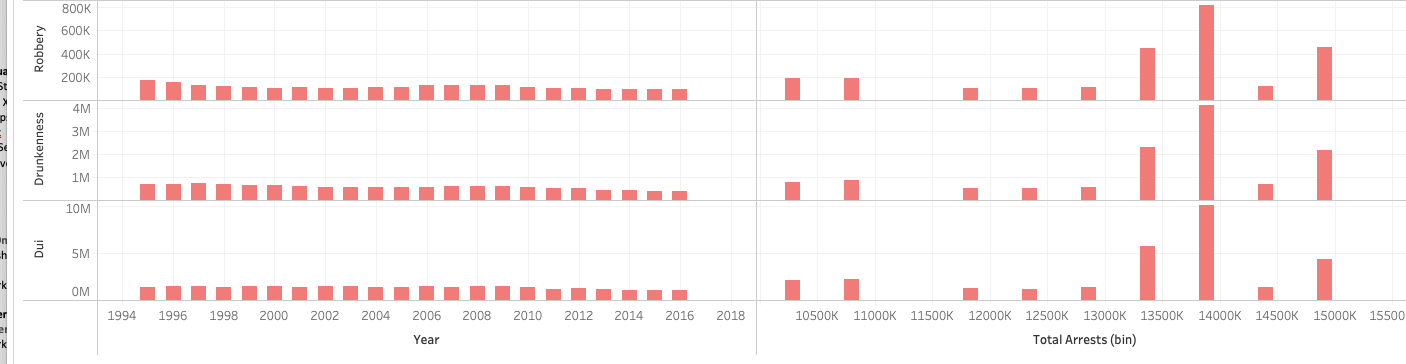


**Description of Graph 2:**

* The above graph depicts the relationship between population and total number of arrests made in a particular years.
* From the above graph, we can also calculate the arrest ratio, which is computed by dividing the total number of arrests by the population of the people.
* Each of the results is calculated for the given years across the span of 22 years.

**Graph 3:**





**Description of Graph 3:**

* The above graph shows the comparison between the total number of arrests to that of a particular crime.
* **There have been quite a few crimes, which have been of a serious concern and have seen the maximum number of arrests being made.**
* **Also the law enforcement agencies have made sure to estimate the above data based on the sex, race and the age group for which the maximum number of arrests have been done.**

1. **Conclusion:** I can conclude my case study with the following points below:

* We saw that how the uniform crime reporting program of the FBI, makes sure to analyze and report the data in a systematic fashion by collecting the results for the number of arrests in several offenses and various demographic locations.
* The various crime-reporting agencies make sure that the data provided is upto the mark and help the government in coming with various ways to curb it.
* The collected data is then depicted in the form of tables, which are very important for the analysis.
* Although, there can be multiple comparisons which can be made using the given data sets, but in order to ensure simplicity, I have only taken up analysis of the number of arrests made for a particular crime.
* The results were quite shocking, as we can clearly see, that even in a minimalistic population of a particular area, the number of arrests have been quite a lot. There are many crimes, which are still a cause of serious worry for the country’s government.

**As the analysis is huge, I believe that just one report might not be able to cover the complete analysis, as there are multiple other factors, which need to be considered.**

**Therefore, I have tried to highlight and explain one data set.**