**San Francisco Crime Analysis**

Group 3: Arnav Gaur, Yutao Zheng, Deepanker

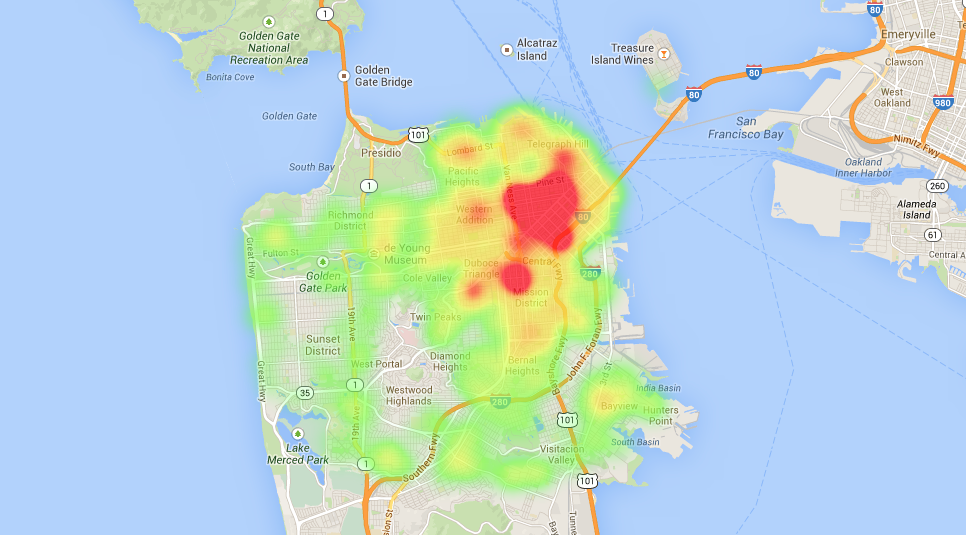
Subject: Urban Big Data Analytics

UNIVERSITY OF BRITISH COLUMBIA

* Introduction

San Francisco officially called the City and County of San Francisco, is a city in, and the cultural, commercial, and financial center of, Northern California.

San Francisco is the 13th-most populous city in the United States, and the fourth-most populous in California, with 883,305 residents as of 2018.It covers an area of about 46.89 square miles (121.4 km2).



* Problem of Interest

From 1934 to 1963, San Francisco was infamous for housing some of the world's most notorious criminals. Today, the city is known more for its tech scene than its criminal past. But, with rising wealth inequality, housing shortages, and a proliferation of expensive digital toys riding BART to work, there is no scarcity of crime in the city by the bay

This project examines the San Francisco crime records and the trends of major crimes in the city across time and locations.

Rstudio and Gapminder Tools have been used for data processing and visualizations.

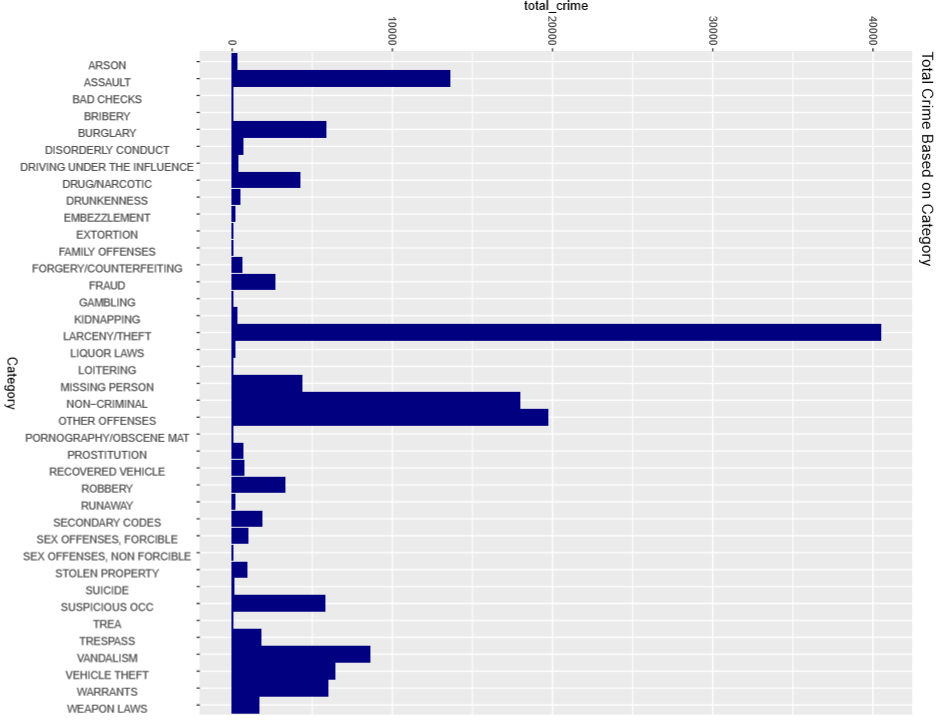
* Data prepossessing

There were 940,668 observations in the original dataset. Duplicates that shared the same incident numbers were removed, leaving 724,426 observations in the data frame. Data which was cause inconsistency was filtered and curated accurately for better visualization.

* Data Analysis

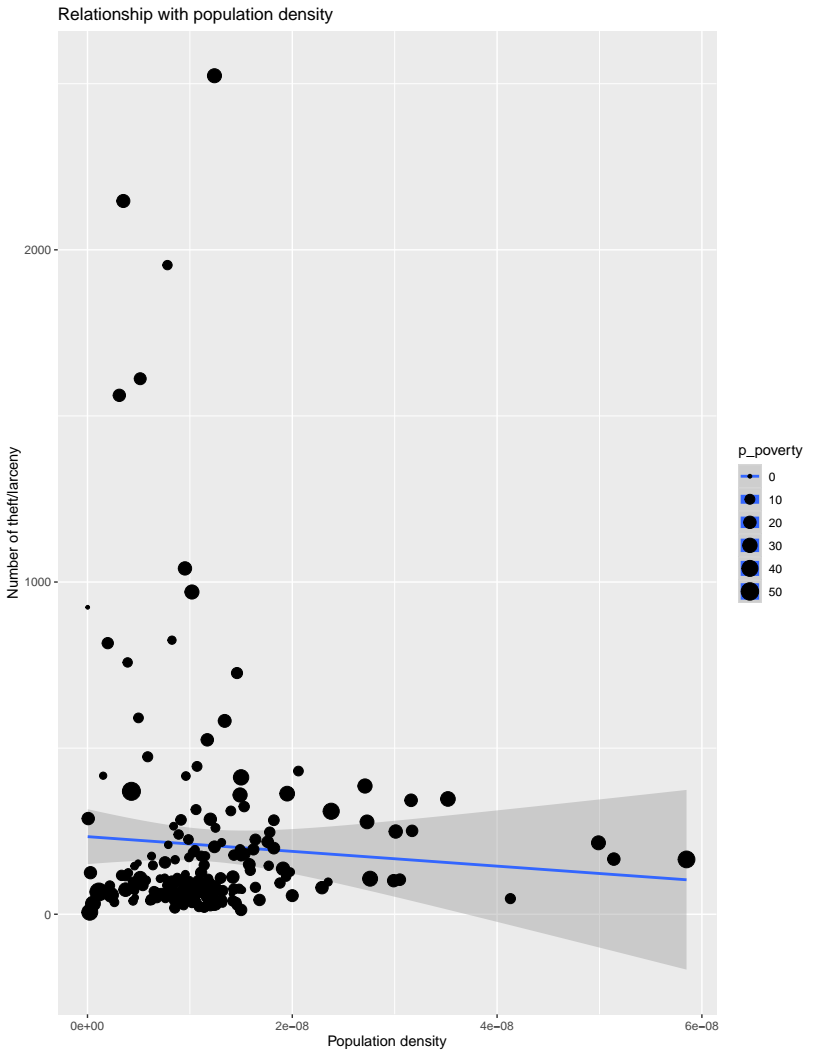
To find various trends and relationships, the crime data was associated with various factors like population density, poverty, unemployment rate, Time and age group. This helped us to conceive a better understanding for the analysis and recommendation.

During the analysis, main focus went about various types of crime and the number of its occurrences, a plot for which is depicted below.



* Population density

It seems very astonishing to say that as the population density increases, there is no significant growth or depreciation seen in the number of crimes. However, in when population becomes larger, the confident interval also increases, therefore, there is more uncertainty of our result.



* Poverty

Its is very evident that poverty does not stand behind to fuel the increase in crime rates.

As the poverty increases, we see that the crime rates follow a positive growth.

