Project Proposal

Capstone 1: NYC Crime Rate 2015-2019

Sherlock Holmes and John Watson just recently moved to New York city. Upon hearing the news, New York Police Department hired the detectives as their consultants. Since the detectives are new in the city, they want to know the crime statistics of various areas in NYC based in the date of the crime rate from the past five years. They want to know what time of crimes is prominent in certain cities, the density of crimes and the demographic of individuals that committed the crime. NYPD wants to be data-driven in establishing appropriate measure on how to lessen prominent crimes in certain areas and what programs they can do to help out the community.

Data will be extracted from the NYC OpenData https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Map-Year-to-Date-/2fra-mtpn. This Dataset contains approximately 500,000 data points over 5 predictor variables. Some of the datasets are categorized by codes but this is also provided in the website. These codes will not be used, rather, state its real label names.

A comprehensive report will be submitted, which includes background statement, problem statement, method, data analysis, result, and discussion. The approach is to established regression and compilation learning models. We will be using different type of data visualization and through that, we can establish and predict the appropriate conclusions.

In deliverance of the dataset, we will be using Jupyter notebooks and its methods and codes for data visualization. We will be using slide decks to present the step-by-step process of this project. A final report will also be provided and submit the codes, report and slide deck in Github.