Capstone Project

Denver Crime Data Analysis

Introduction

- Crime Rates are increasing. Therefore, the demand on police officers is getting higher.
- There is a huge need to improve the performance of crime monitoring and detection.
- An understanding of the criminal behavior and the areas where crime occurs will ease the police officers' job and save time and effort.

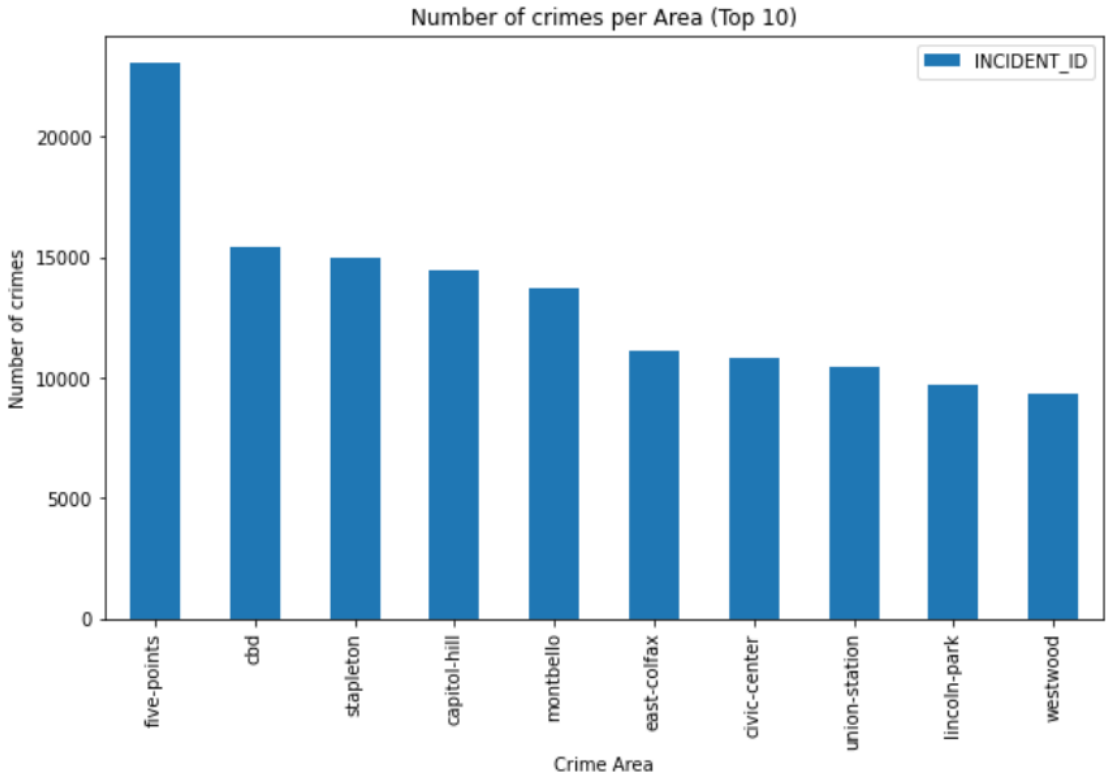
Data Accusation

- Two sets of data will be used :

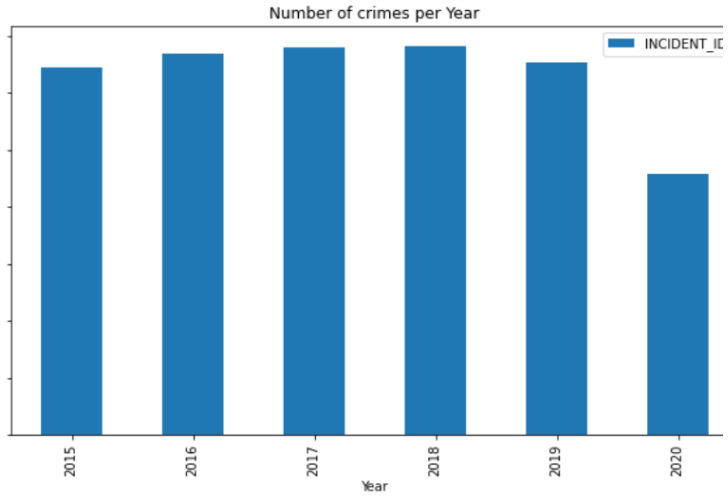
- A dataset from Kaggle that contains historical data of Denver's city crimes (offense type, offense category, first occurrence, incident address, longitude, latitude, Neighborhood)

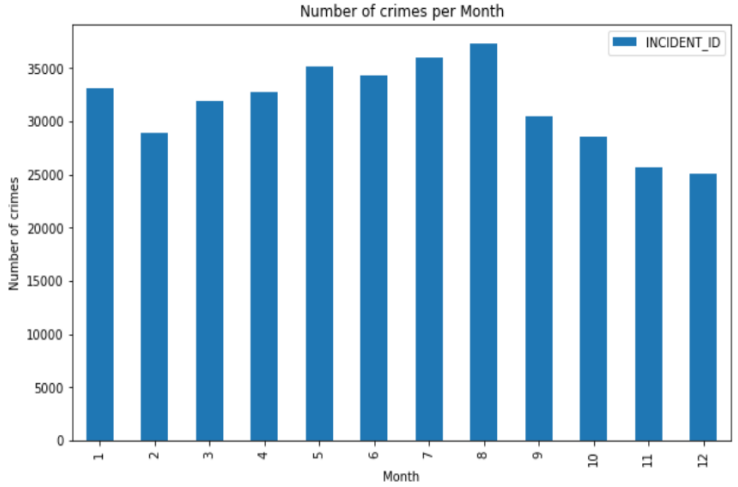
- Dataset extracted using the Foursquare API. Using the neighborhood from the crime dataset, a set of venues will be extracted (Arts and Entertainment, Fitness Center, Food, Medical Center, Nightlife Spot, and Shop and Service)

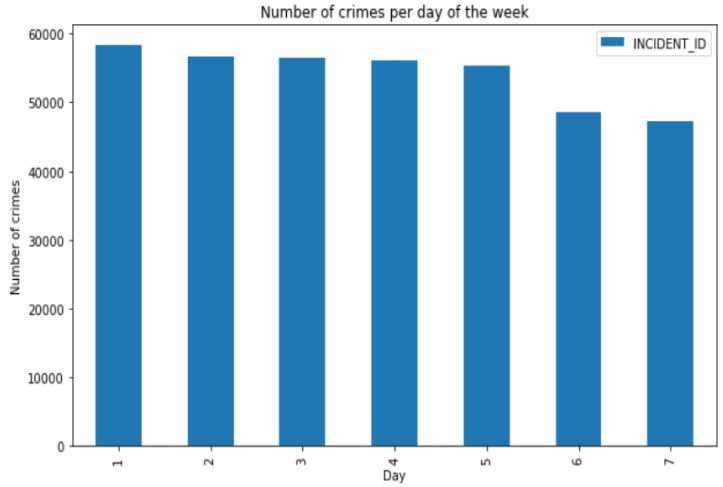
Identifying the Crime Rate Relationship with TimAreae



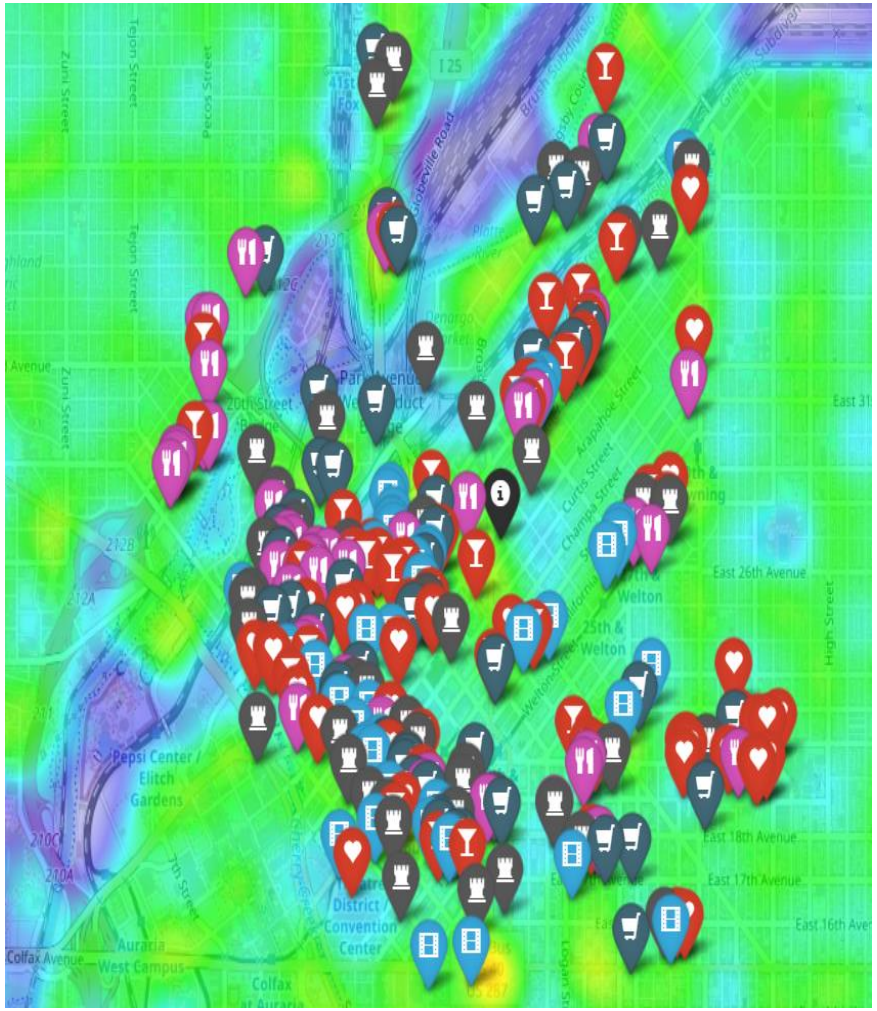
Identifying the Crime Rate Relationship with Time

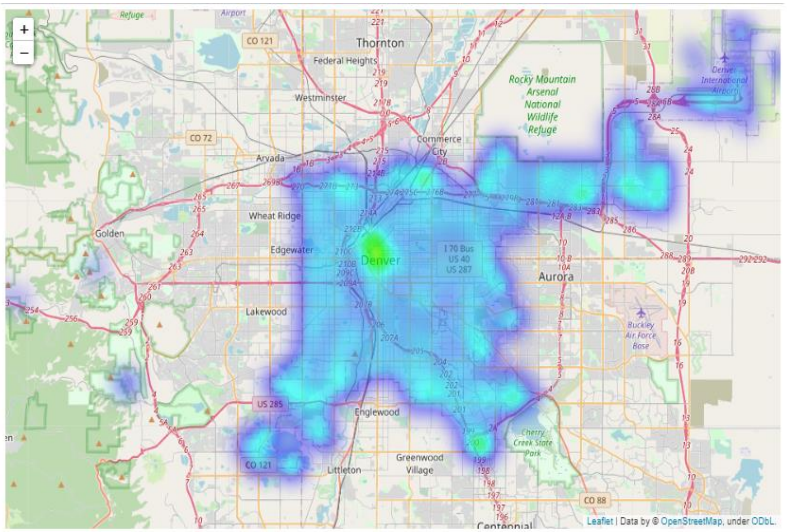


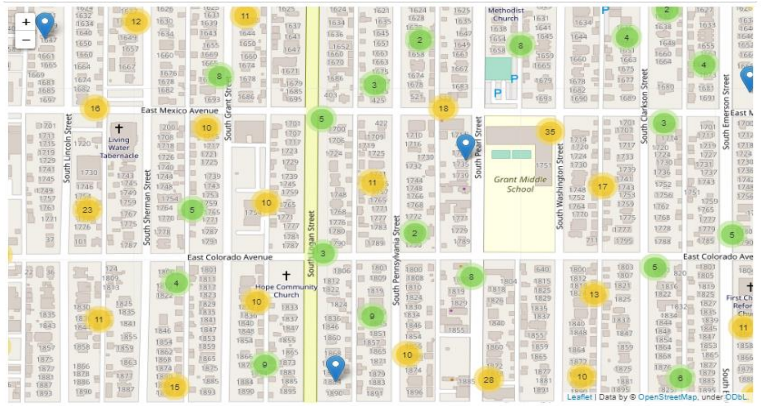




Mapping of all Crimes with Location and venues

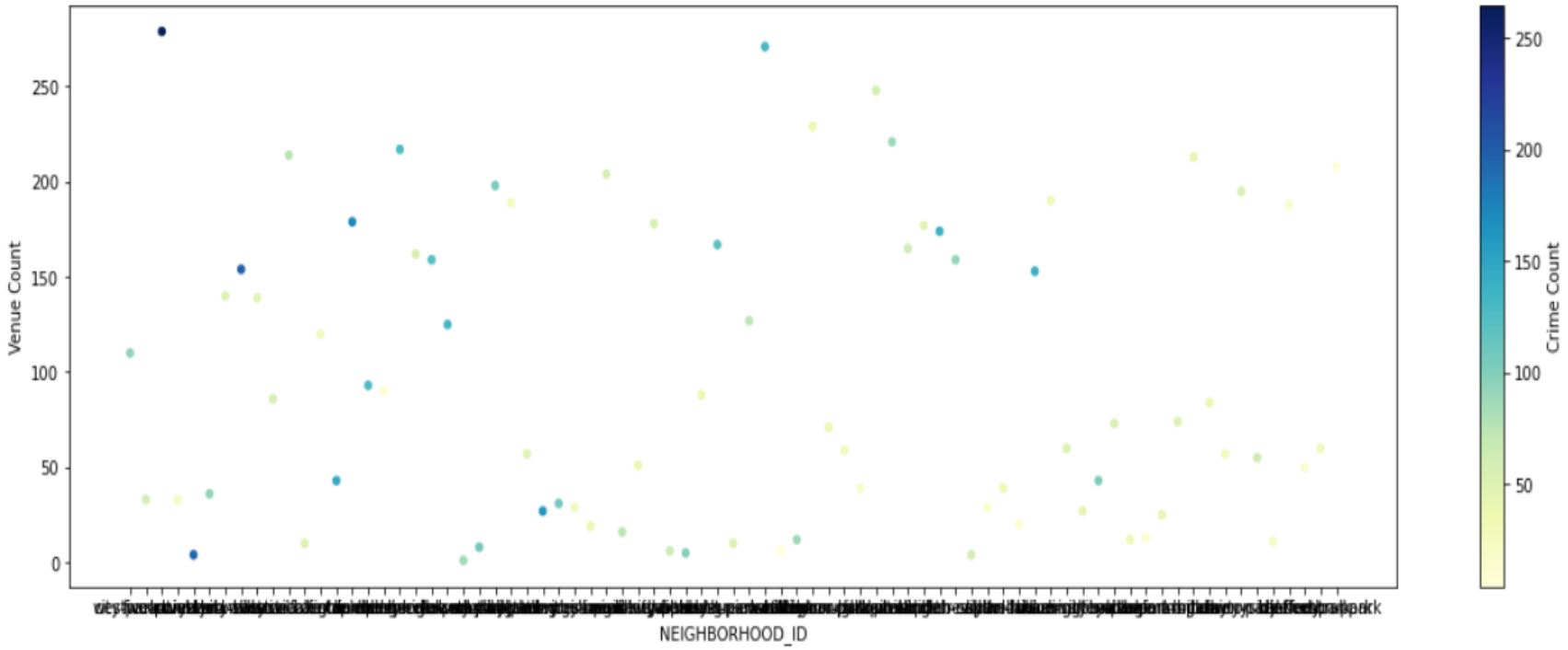




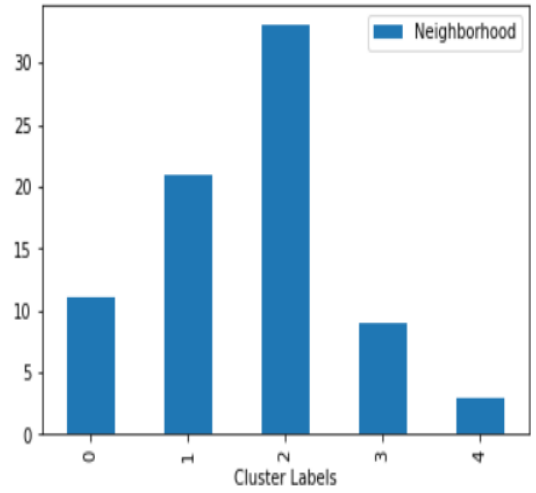


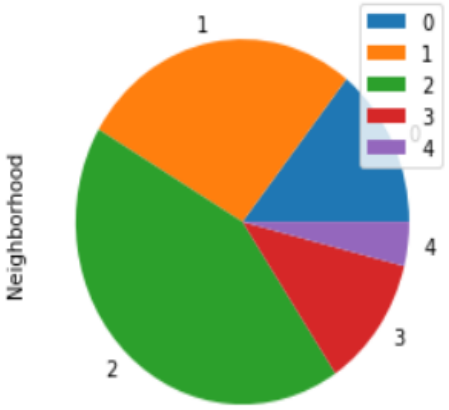
relationship between the number of venues and crimes by neighborhood

- Considering only business burglary crimes



Clustering Neighborhoods by Venues and Crime rate





- Using K- means algorithm
- K=5

Results and Discussion

- Each cluster have a special characteristic that requires the attention of police officers
- Cluster 0: Contains a very high burglary crime count even though the number if venues is small. Thus, the neighborhoods in this cluster requires a more extensive surveillance.
- Cluster 1: have a very low burglary crime rate even though it's populated with venues. Police officers may try to understand why the crime rate is lower in this neighborhoods comparing to others
- Cluster 2 and 3: the number of crimes and venues is relatively proportional.
- Cluster 4: have a very high burglary crime rate and a high venues count. And a further investigation of the area is needed to understand the causes.