

# Identification of crime prone areas in Detroit

---

# Introduction

---

Detroit has the fourth highest murder rate among major cities in the United States (As of 2016)

It has the 42nd highest murder rate in the world.

This analysis will help local government agencies as well as tourists to identify geographical areas of interests.

Government agencies will be able to make informed and focused decisions to reap out desired outcomes more efficiently.

The tourists will be aware of the areas that should be avoided for safe travel and stay.

Entrepreneurs can also understand demographics of various areas around city to make better investments for growth and profits.

# Data acquisition and cleaning

---

The initial data required can be broken down into two separate data sets:

- The FourSquare Venues to Visit in Detroit
- The Detroit Police Department Crime Data from 1920 to June, 2019 (<https://data.detroitmi.gov/api/views/6gdg-y3kf/rows.csv?accessType=DOWNLOAD>)

Features to keep from crime database

- Crime ID
- Incident Date & Time
- Offense Category
- Neighborhood
- Latitude
- Longitude

Add new columns for:

- Hour
- Day
- Month
- Year

Predictive modelling

- Prepare data to include only numerical data and by removing unneeded columns

# Data Visualisation and Analysis

---

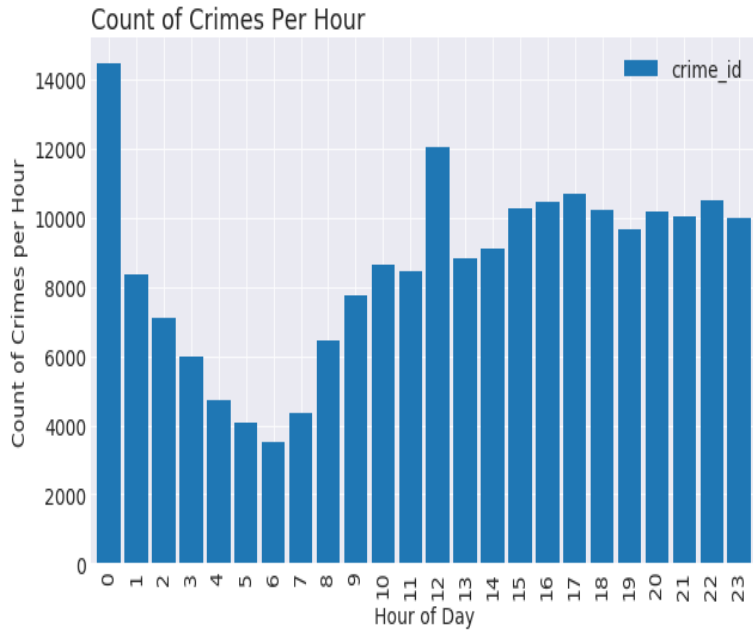
Bar chart

Area plots

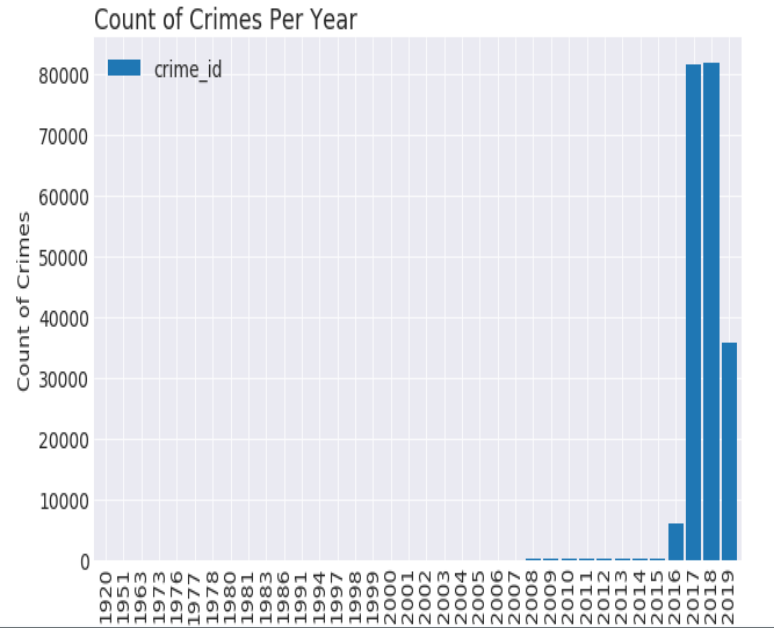
Scatter plots

Folium Maps with marker

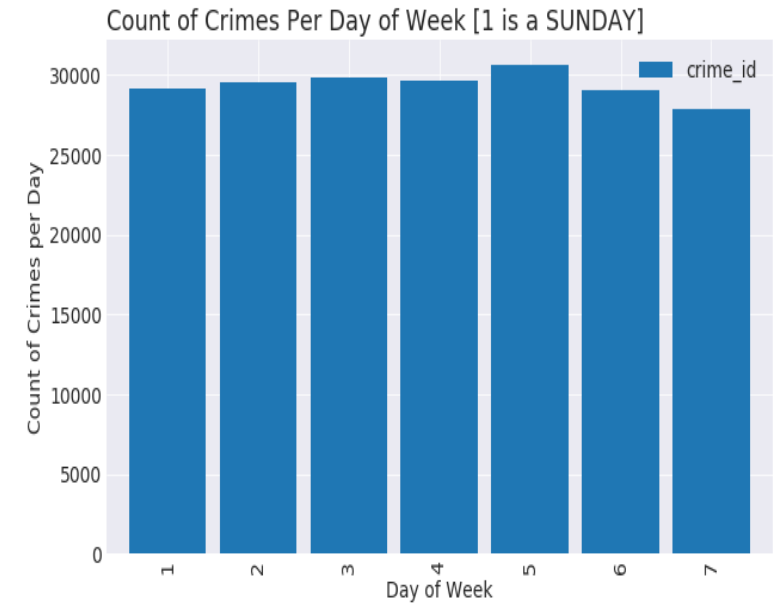
# Bar charts



- Most crimes happen around midnight and noon
- Least around 5am morning.



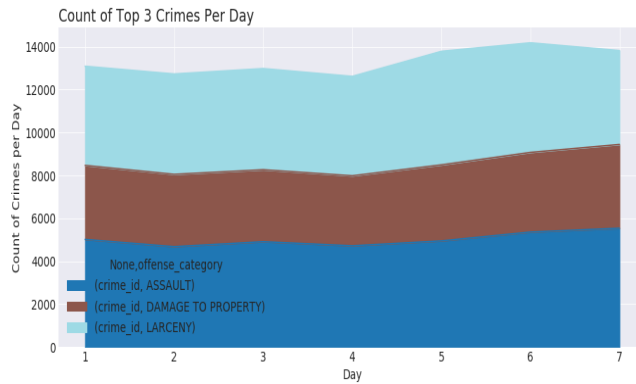
Sharp increase in 2017



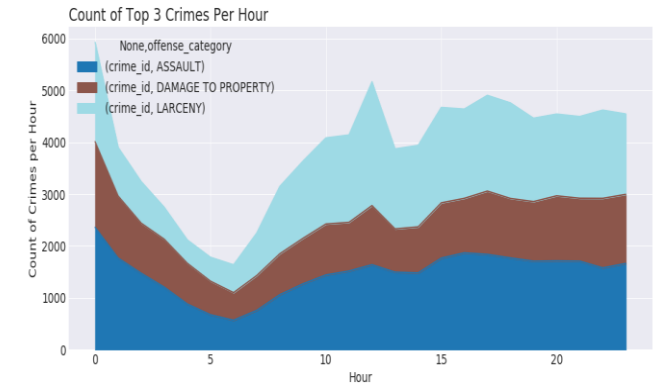
More crimes happen on Thursday

# Area plots

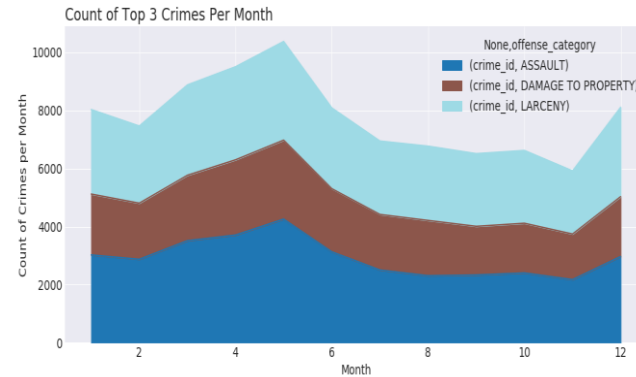
Second half of week is more dangerous



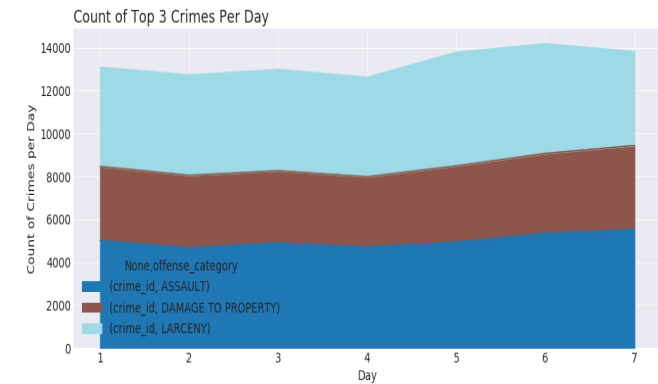
Midnight and noon are more dangerous



May and June are more dangerous

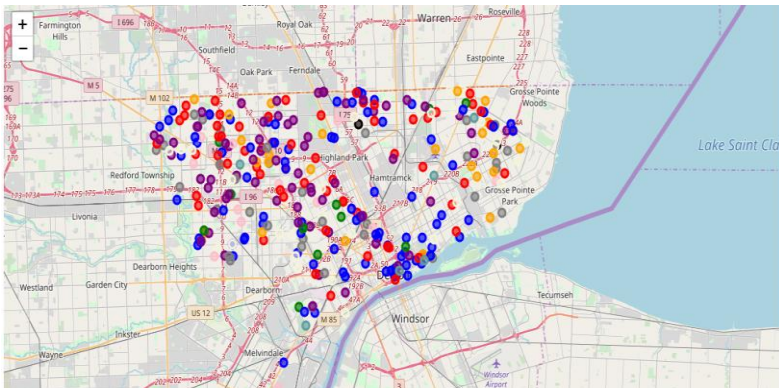


Weekends are more dangerous

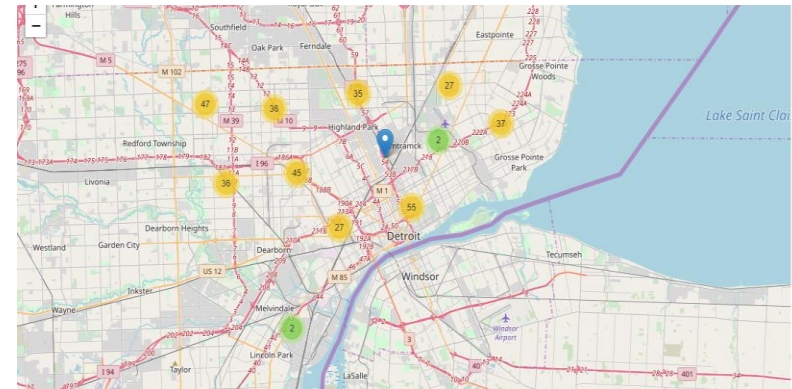


# Folium maps

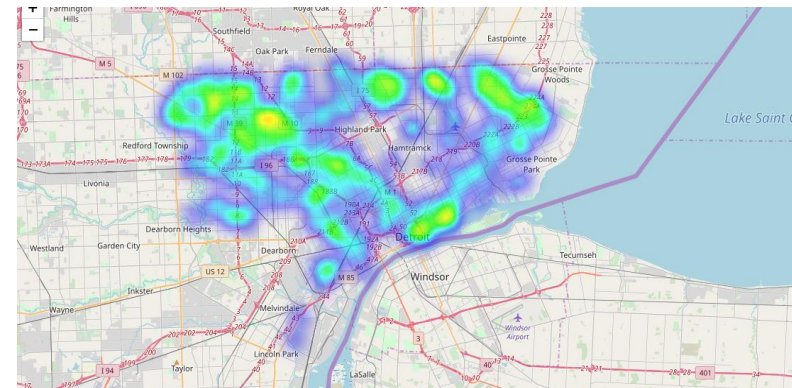
Markers of crime incidents on Detroit map



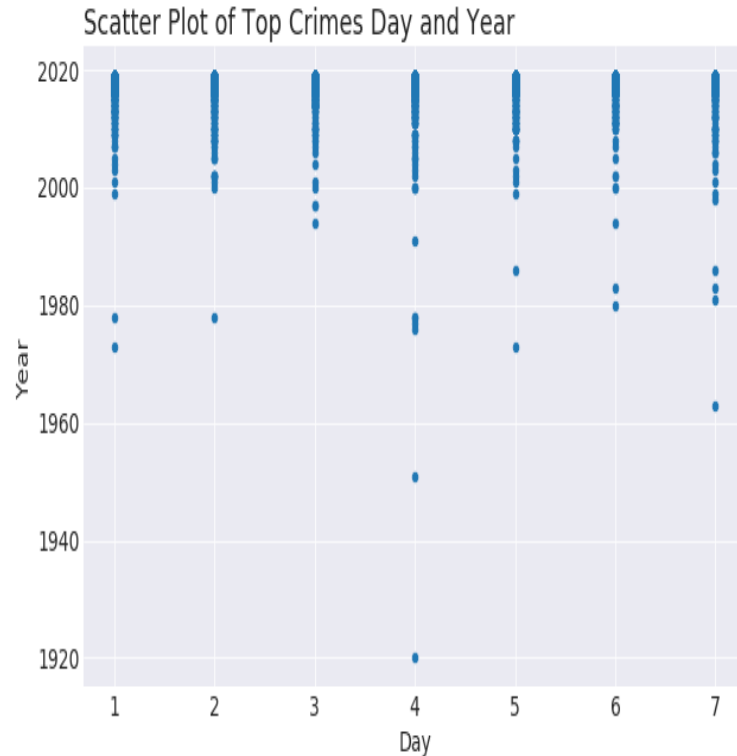
Cluster of crime incidents on Detroit map concentrated around the periphery of Detroit



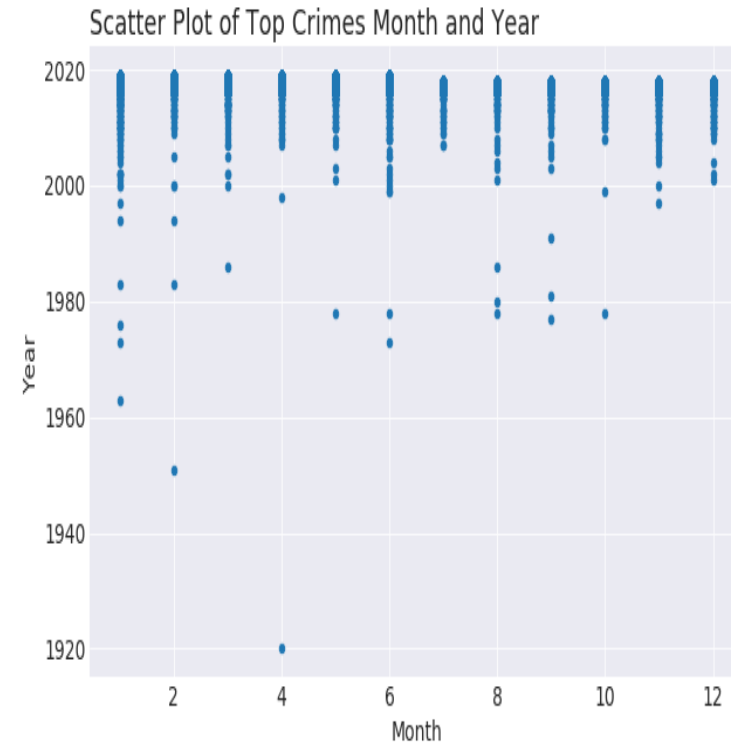
Heat Map showing **Greektown** and **Macomb street** have a high crime rate occurrence



# Scatter Plots of crimes from 1920 to 2019



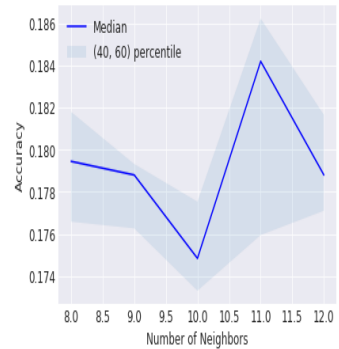
Wednesday is more dangerous



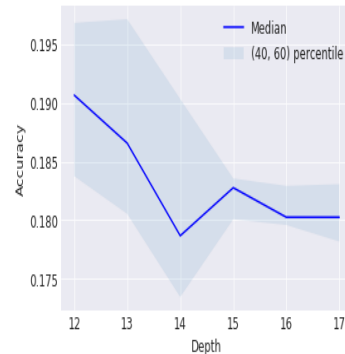
January has more occurrence of crimes



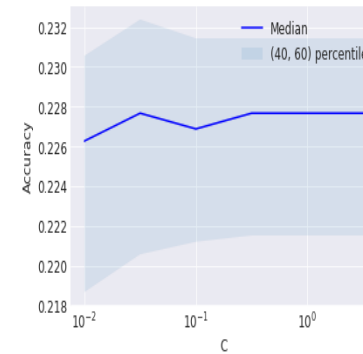
# Predictive modelling



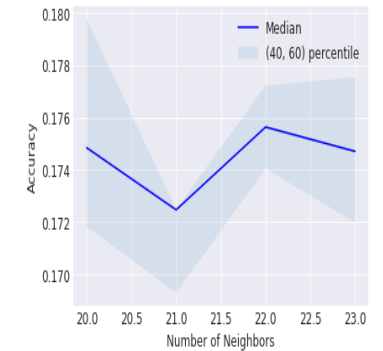
KNN(K=9)



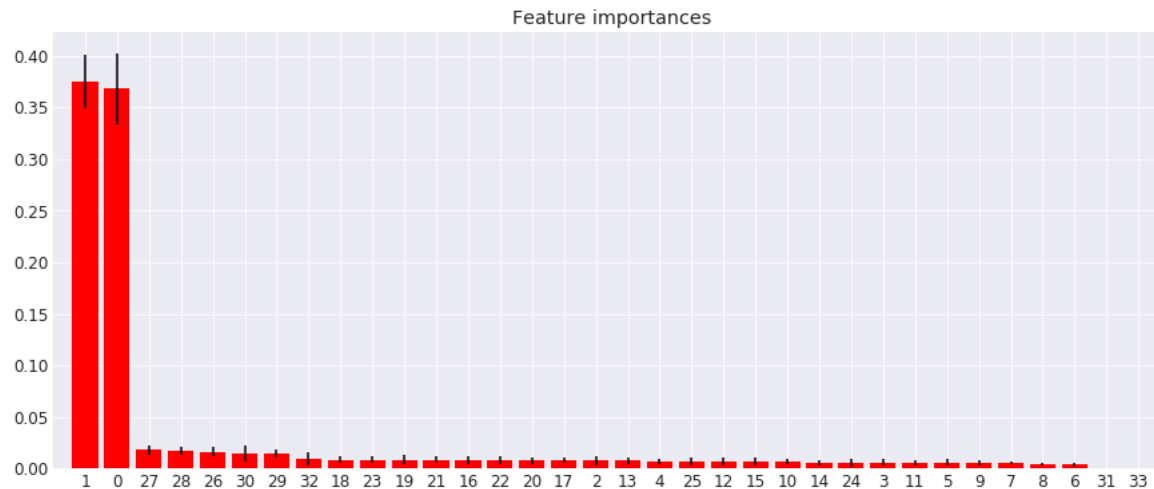
Decision Tree



Logistics Regression



Random Forest



- Due to computational constraints, filtered and processed dataset to crimes that happened at 6<sup>th</sup> hour in June of 2018
- Best Model – Random Forest based on F1-Score, Jaccard and Log Loss
- Latitude and Longitude influence the model most

# Conclusion and Results

---

Based on the analysis, this project offers a way for travelers to analyses venues to travel from foursqaure api quickly. The crime data analysis suggests the following:

Avoid travelling to Detroit in January

Take extra precautions on Wednesday if planning to travel

Avoid travelling to Detroit city peripheral areas as they are more prone to crime

Avoid being outdoors during noon

# Future Work

---

Link venue data from Foursquare to crime data and make a combined visualization for on the spot recommendations based on crime level of locality.

Include more dependent variables and collect more dataset to create more accurate predictive model.