Battle of Neighborhoods

New York – Smart Neighborhood Finder

Detomas Simone

Project objective

- Classify New York neighborhoods through crime data downloaded from NYPD and venues data collected form foursquare API.
- To recommend the best neighborhood to live for a young couple when moving to New York

Python packages and Dependencies:

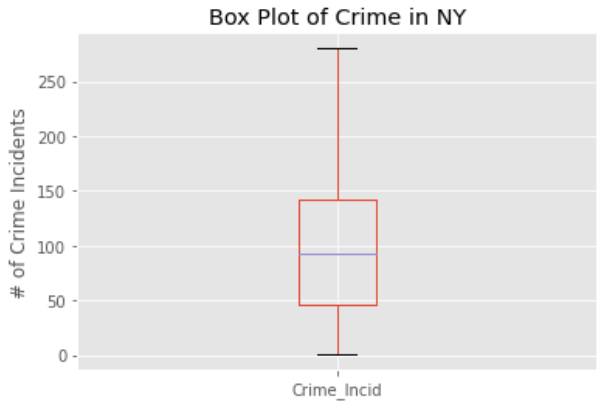
- Pandas - Library for Data Analysis
- NumPy – Library to handle data in a vectorized manner
- JSON – Library to handle JSON files
- Geopy – To retrieve Location Data
- Requests – Library to handle http requests
- Matplotlib – Python Plotting Module
- Sklearn – Python machine learning Library
- Folium – Map rendering Library

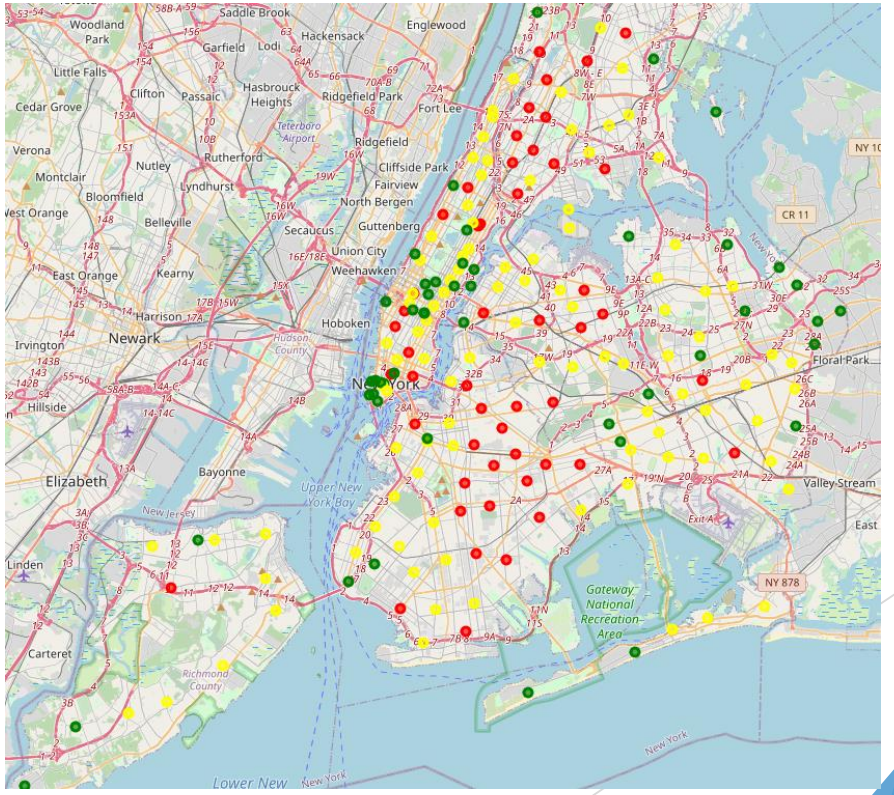
Data collection

- Web Scraping and Data Wrangling for NYPD crime data
- Venues collection from Foursquare API

Exploratory data analysis

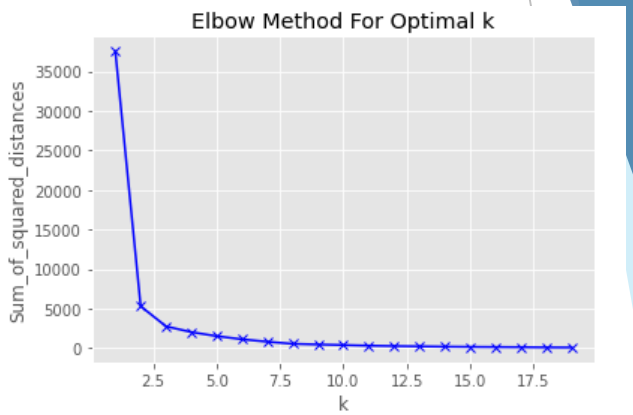
- Neighborhoods are filtered based on safest areas





Clustering

- K-Means clustering to classify the neighborhoods



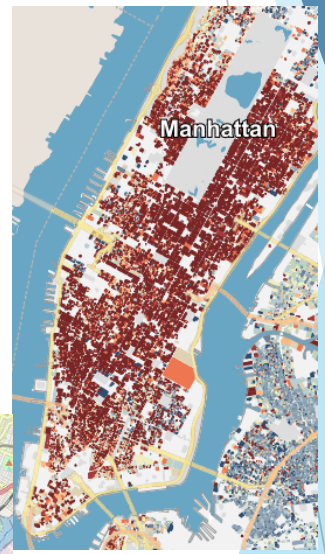
- Analysis of the cluster ot verify the cluster that is coherent with my desires

- Filtering of neighborhoods based on clustering results

Conclusion

- From the huge area of NY we were able to narrow down the potential neighborhhods to 13 Postal Codes

- Let’s then check the average cost per square meter in NY



So…. let’s start to look for some ads in Lower Manhattan!