Coursera | IBM Data Science

Battle of Denver Neighborhoods

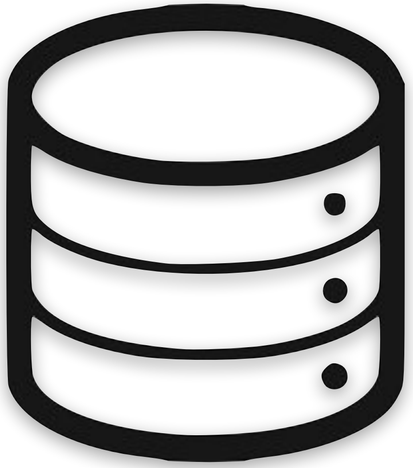
Roman Kandirovskiy | June 2021

Table of Contents

Introduction

- Moving to a new city can be a challenging task (at least it was for me in the past)
- You need to find an apartment/house for rent in a neighborhood that fits your criteria for transportation, safety, and nearby availability of specific venues such as grocery stores, hospitals, restaurants, entertainment, schools, and parks.
- As a resident of Denver, CO, I would like to make that process easier for new-comers by segmenting the city into neighbrohoods and clustering them based on socioeconomic and location characteristics. Ideally, I'm looking to find a location similar to my current residence that is in close proximity to a grocery store, park, gym, and restaurants, public transport is a plus.

DATA



- **Foursquare API** will be used to access most up to date venue/facility data in Denver, CO. Exact venues that we are looking for:
  - Residential Building/Apartment Complex
  - Grocery Store
  - Gym/Fitness Studio
  - Park
  - Restaurants
- **City and County of Denver (later CCD)** offers over 200 open data sets in csv and other formats and for this project I selected Crime Data which includes criminal offenses in the City and County of Denver for the previous five calendar years plus the current year to date. The data is based on the National Incident Based Reporting System (NIBRS) which includes all victims of person crimes and all crimes within an incident. We will use to find a safe location with low amounts of crime and theft.
- **Rent Data** that was sourced at the beginning of this project from RentCafe and Zumper websites using BeautifulSoup. However, the webpages used for web-scraping have changed and thus I have had to use a backup copy of that portion of the project that is saved locally.

Methodology

1 Install and Import Libraries to Jupiter Notebook

In [1]:

```
#!pip install html5lib
#!pip install sklearn
#!pip install seaborn
import pandas as pd
import numpy as np
from folium.plugins import HeatMap
import requests

import matplotlib.pyplot as plt
from geopy.geocoders import Nominatim
import matplotlib.cm as cm
import matplotlib.colors as colors
from sklearn.cluster import KMeans
import folium
import seaborn as sns
```

2 Process Denver Crime Data Using Pandas

```
Would you like to load local or external data? Enter "local" or "external" :  local
Loading Local Version
```

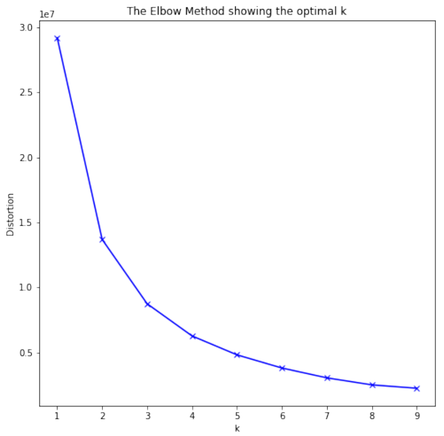
	INCIDENT_ID	OFFENSE_ID	OFFENSE_CODE	OFFENSE_CODE_EXTENSION	OFFENSE_TYPE_ID	OFFENSE_
0	2021224206	2021224206220200	2202	0	burglary-residence-by-force	
1	2021225308	2021225308240400	2404	0	theft-of-motor-vehicle	
2	20216009452	20216009452239900	2399	0	theft-other	
3	20216009439	20216009439230500	2305	0	theft-items-from-vehicle	theft-fro
4	20218017976	20218017976240400	2404	0	theft-of-motor-vehicle	

3 Collect Venue Data Using FourSquare API and Analyze It Using Pandas

```
denver_venues.groupby("Venue Category").count()
```

Venue Category	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
Coffee Shop	327	327	327	327	327	327
Grocery Store	140	140	140	140	140	140
Gym / Fitness Center	277	277	277	277	277	277
Park	257	257	257	257	257	257
Residential Building (Apartment / Condo)	384	384	384	384	384	384

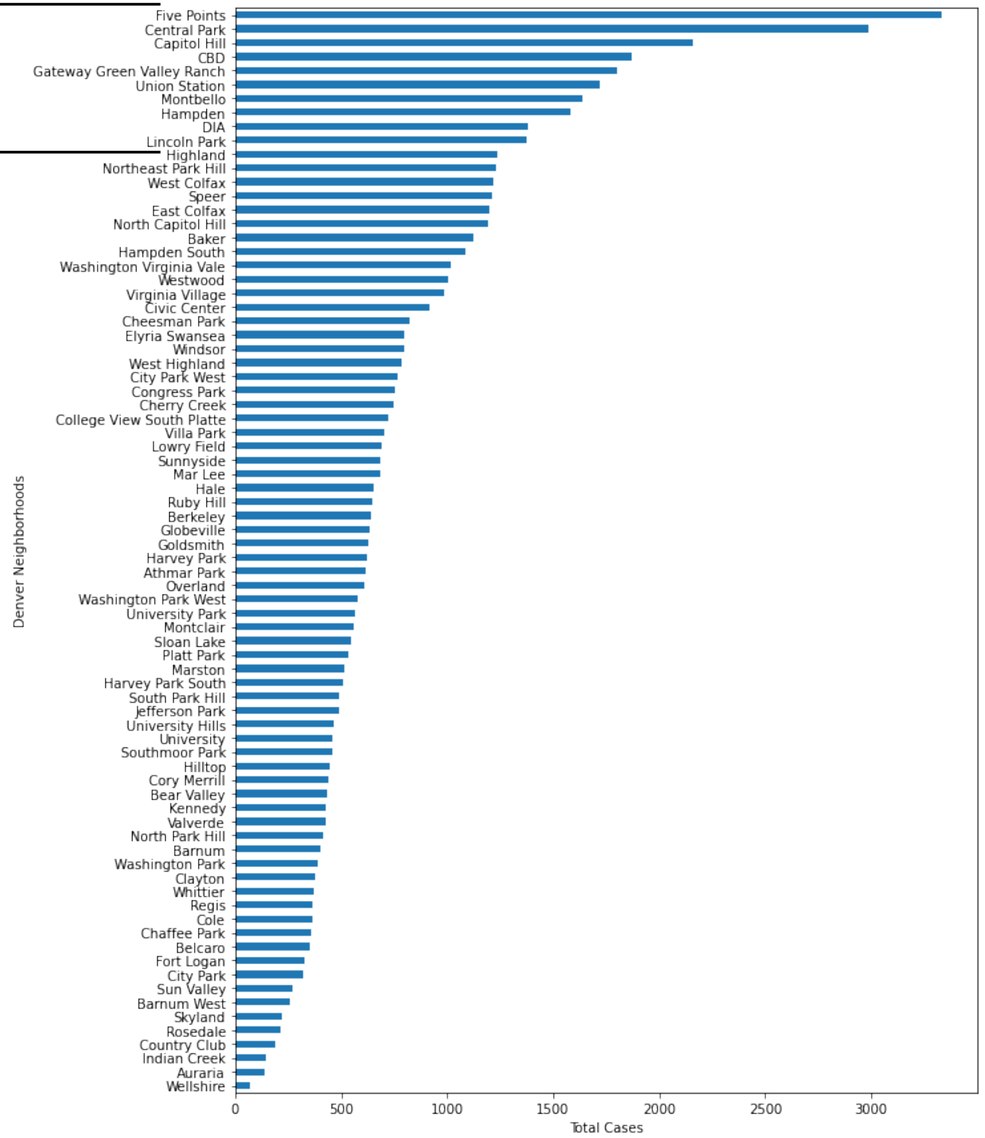
4 Combine All Data for K-Means Clustering
Pick the Optimal K Using the Elbow Method
Cluster Neighborhoods into 5 Groups Using K-Means Method

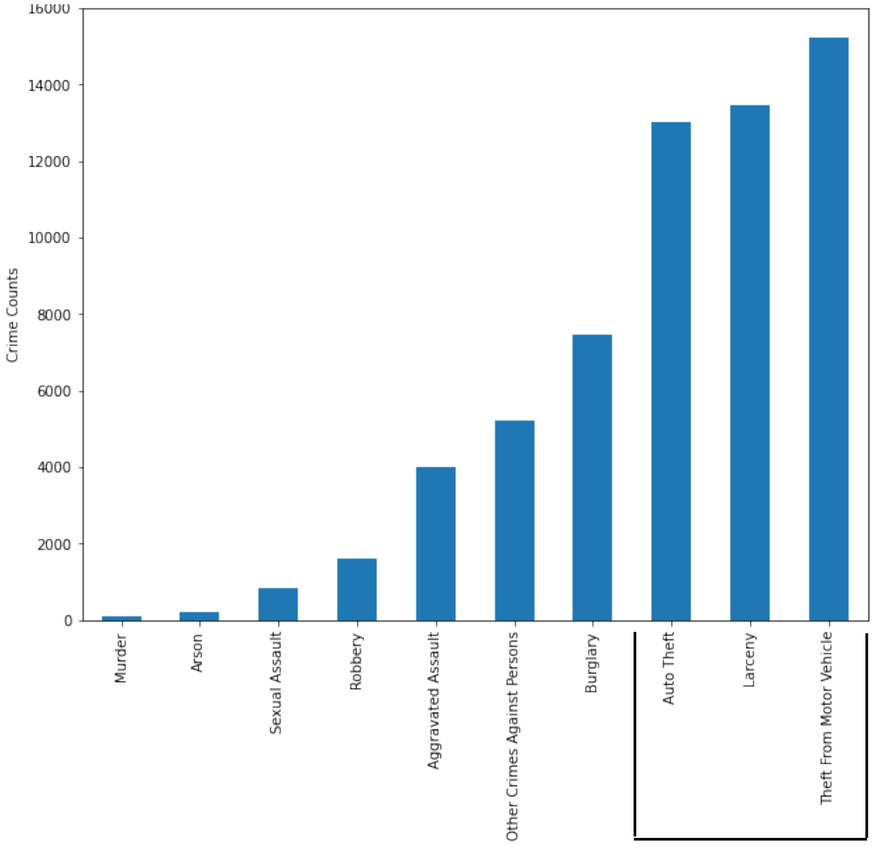


Results

Total Crimes in Denver by Neighborhood and Category

Top 10 Neighborhoods By Crime Count

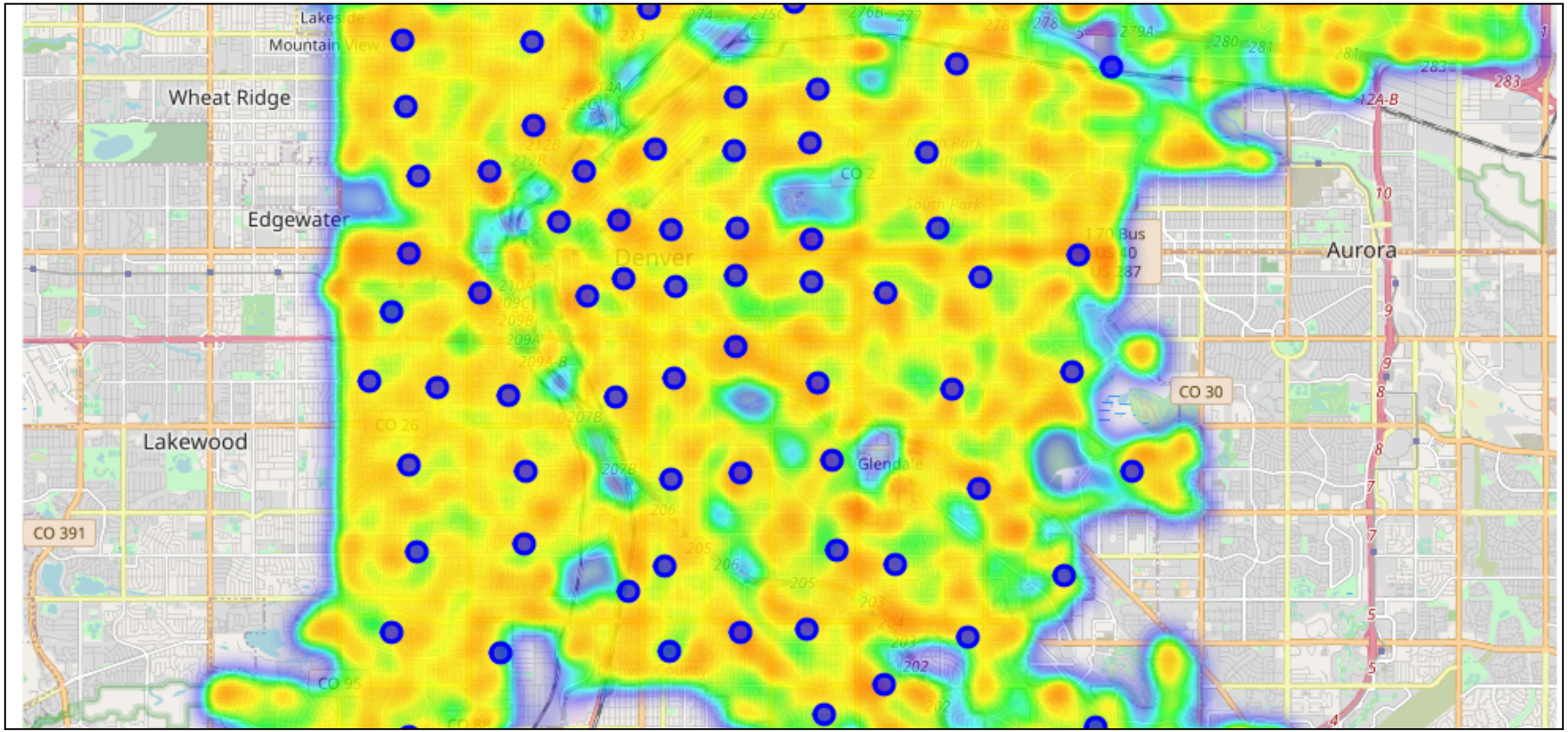




Top 3 Crime Types

Results

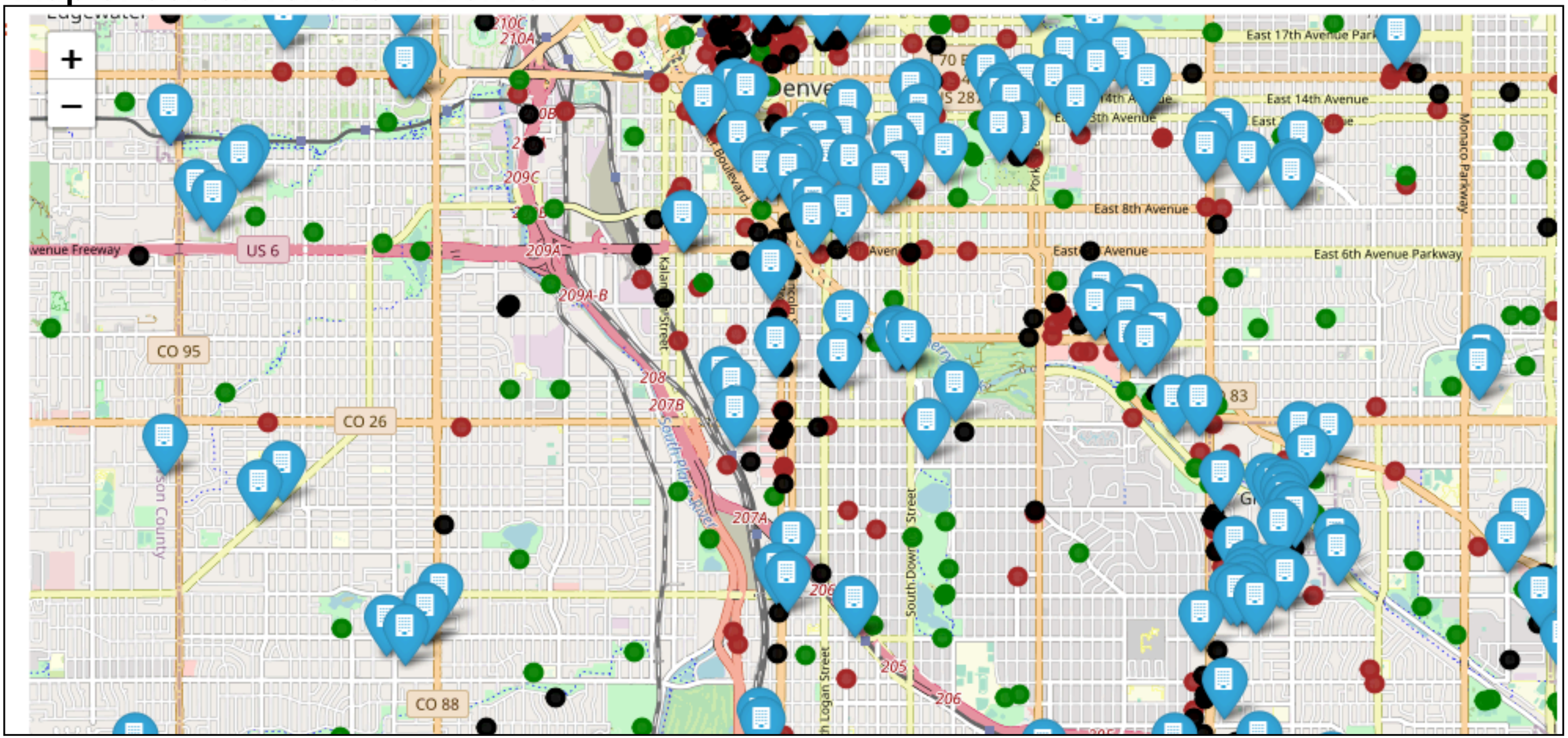
Heatmap of Crime in Denver*



* Interactive version is available via notebook **hosted on IBM**

Results

Map of Select Venues in Denver*

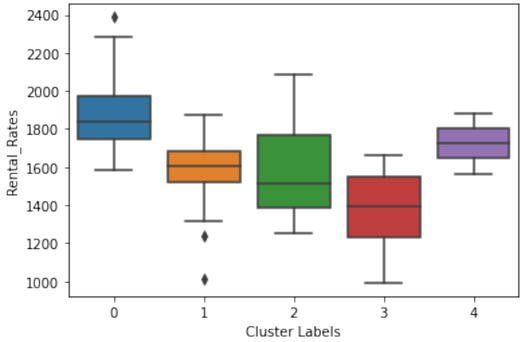


* Interactive version is available via notebook **hosted on IBM**

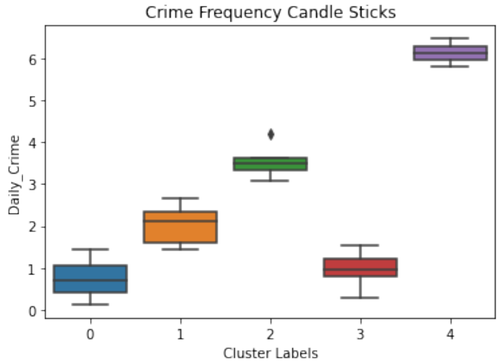
Results

Total Crimes in Denver by Neighborhood and Category

Spread of Average Rental Rates for Different Neighborhood Clusters

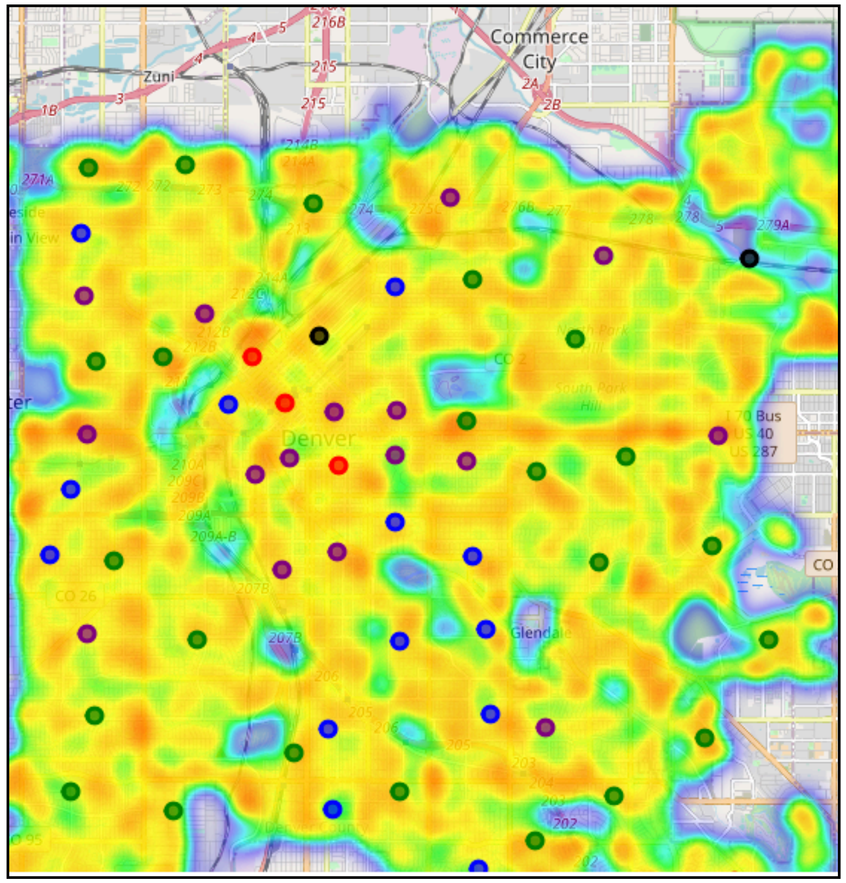


Spread of Daily Crime Cases for Different Neighborhood Clusters



Denver Crime HeatMap*

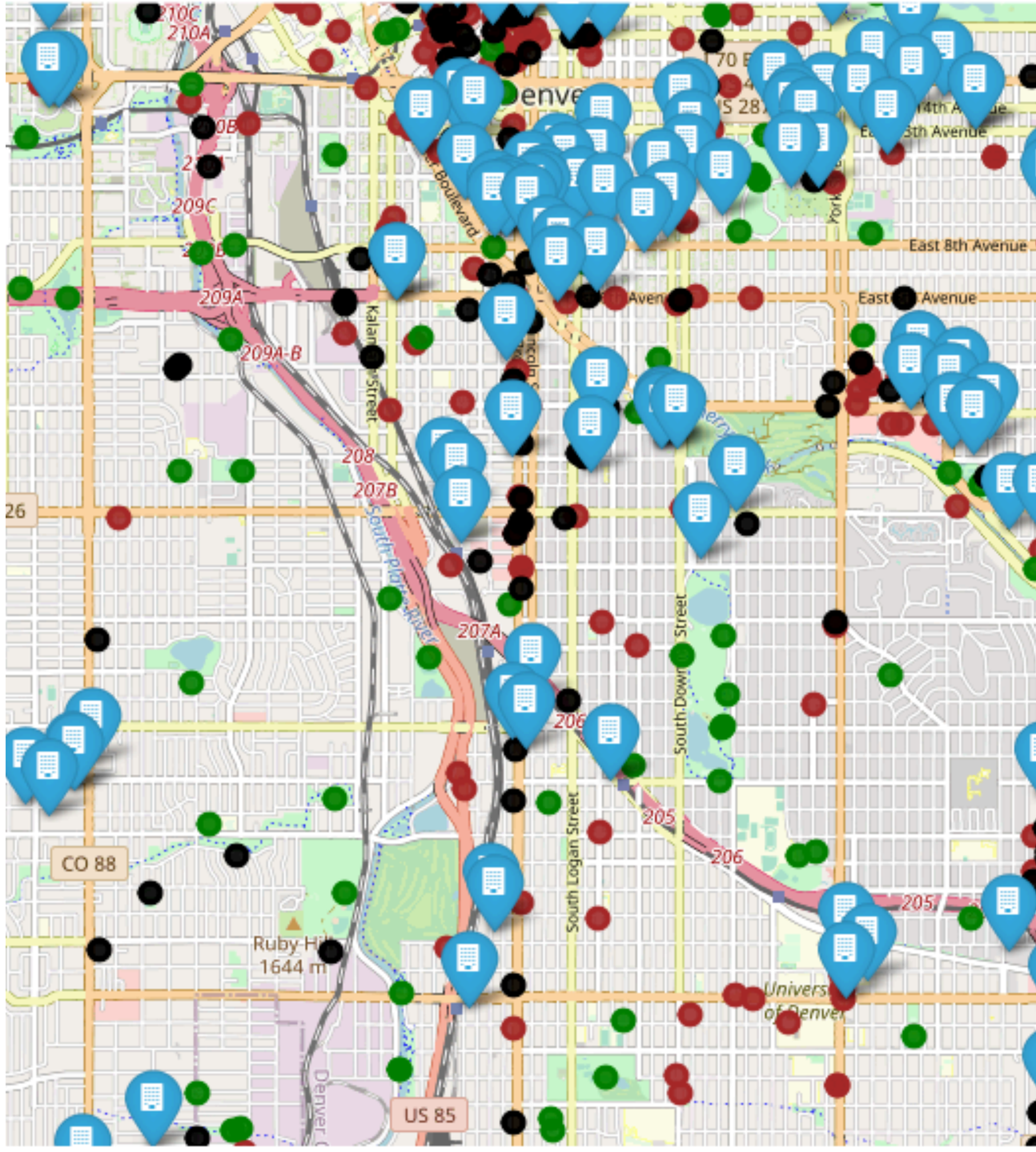
Markers Represent Neighborhoods and Colors Represent Cluster



* Interactive version is available via notebook **hosted on IBM**

Conclusion

- Are you looking to move into a safe neighborhood in Denver that offers a lot of coffee shops, gyms, grocery stores, and parks next to your residence? Clusters #0 and #3 are for you!
  - Examples of Cluster #0 (Higher than Average Rent) Neighborhoods:
    - Platt Park
    - Washington Park
    - Berkely
    - Barnum West
  - Examples of Cluster #3 (Lower than Average Rent) Neighborhoods:
    - Sloan Lake
    - Jefferson Park
    - Athmar Park
- If you're looking to live in the middle of everything, cluster #1 and #2 will offer you that but make sure to stay clear of cluster #4 as it appears to have a much higher crime rate than the rest of the neighborhoods.



References

- FourSquare Places API Docs: https://developer.foursquare.com/docs/places-api/
- Denver Open Data Catalogue: https://www.denvergov.org/opendata/
- Folium Docs: https://python-visualization.github.io/folium/index.html
- Pandas Docs: https://pandas.pydata.org/docs/
- K-Means Elbow Method: https://predictivehacks.com/k-means-elbow-method-code-for-python/
- Rent Data:
- Zumper: https://www.zumper.com/rent-research/denver-co
- Rent Cafe: https://www.rentcafe.com/average-rent-market-trends/us/co/denver/

Acknowledgments

- Coursera and IBM Data Science Instructors, TAs, and Fellow Students.

GitHub Link to Project

- https://github.com/roman-pk/python-capstone-project