

Coursera Capstone

IBM Applied Data Science
Capstone

**Living in DC, planning for
NYC**

Nam Le

June 12, 2020



Business Problem

- ◆ Moving from one city to another can be a stressful process, given your personality and ideal situation
- ◆ I am currently living in DC, specifically in the Dupont Circle area, but have been offered an irresistible opportunity in Manhattan. I love fitness, variety of food, and parks
- ◆ Objective: To analyze Manhattan neighborhoods and selected apartment units based on given criteria, preferences, and similarity to current residence, and recommend a specific unit that matches these specification
- ◆ Business Question: Given geographic data, nearby venues information, and rent price distribution, can this process recommend an apartment unit that fits my needs?

Data

- ◆ Data Required:

- ◆ List of Manhattan neighborhoods (name, borough, and location data)
- ◆ List of potential apartment residences (address, size, bedroom, and price)
- ◆ Venue data, particularly related to gyms, food, parks, etc.

- ◆ Sources of Data:

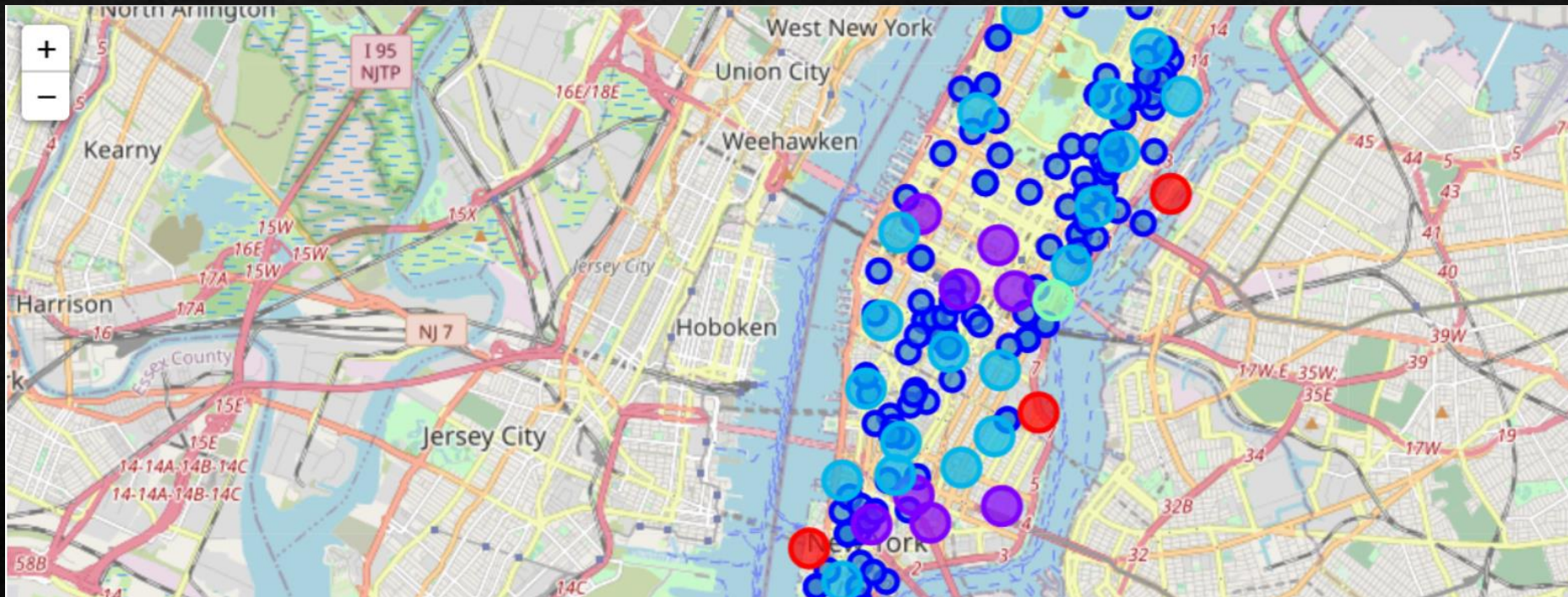
- ◆ NYU data for neighborhood names
- ◆ Apartment listing websites such as Rent Manhattan, Nestpick, and Realtor
- ◆ Geocoder package for latitude and longitude
- ◆ Foursquare API for venue data

Methodology

- ◆ Web-scraping off websites
- ◆ Integrating latitude and longitude coordinates using Geocoder
- ◆ Leveraging Foursquare API for venue data
- ◆ Grouping data by neighborhood and performing K-means clustering
- ◆ Visualize cluster algorithm using Folium mapping
- ◆ Make recommendation

Results

- ◆ Categorized neighborhoods into 5 clusters based on rent prices:
- ◆ Cluster 0 is orange, Cluster 1 is purple, Cluster 2 is blue, Cluster 3 is green, and Cluster 4 is yellow (not seen because it is at the northern tip of Manhattan)
- ◆ Cluster 1 is generally most expensive followed by Cluster 2, due to proximity to commercial centers



Discussion and Recommendation

- ◆ After cluster analysis, I narrowed down my search to two apartments, one in Cluster 1 and one in Cluster 2
- ◆ Apartment unit located in W 23rd St in Flatiron District is in Cluster 2, while Apartment unit in 52 Spring St in Little Italy is in Cluster 1
- ◆ The first apartment costs \$4250 per month, and its top venues are gyms, restaurants, and parks, while the second apartment costs \$6000 per month, and its top venues are restaurants
- ◆ From weighing my preferences for fitness, food options, and outdoor recreation, the first apartment makes most sense for me and matches my expectations

Conclusion

- ◆ Answer to Business Question: Neighborhoods in cluster 2 are the best suited for me and the apartment unit located at W 23rd St in Flatiron District is the most ideal location for me coming from Dupont Circle
- ◆ The results of this study will hopefully be replicated and built on to create the go-to recommender system for moving from one metropolitan to another