RECOMMENDING WHERE TO OPEN A COFFEE SHOP IN MANHATTAN

IBM & Coursera – Data Science Capstone:
Battle of the Neighborhoods

Introduction

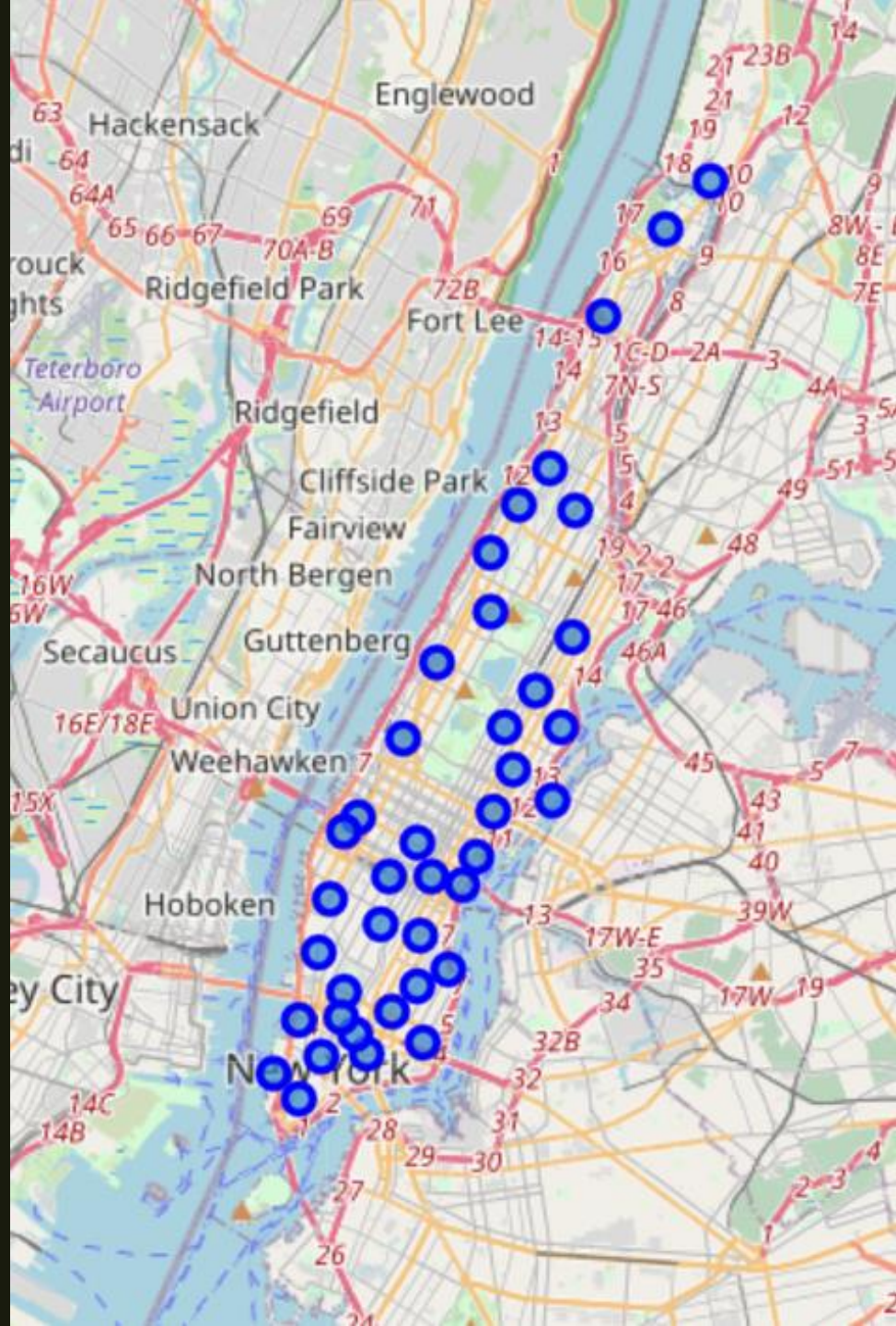
- Manhattan – A Borough within NYC. World known location for food & culture
- In this fictitious scenario, a team of investors (**our target audience**) want to open a coffee shop (“*24/7 Caffeine Queen*”) in Manhattan.
- However, they do not know where to open the coffee shop.
- Initial research shows that coffee shops usually tend to do favorably when they are located near bookstores.

Data

- For analysis, k-means was utilized to determine which neighborhoods in Manhattan were similar by venue type utilizing Foursquare API information.
  - *The ideal number of clusters were determined using the elbow method (5).*
- Python programming in combination with Folium were incorporated into the analysis to analyze for each cluster of neighborhood the following:
  - *Number of Coffeshops*
  - *Number of Bookstores (new & used)*
- Through normalization, the clusters were rated and the final results yielded two neighborhoods for the investors (target audience) to consider in further detail.

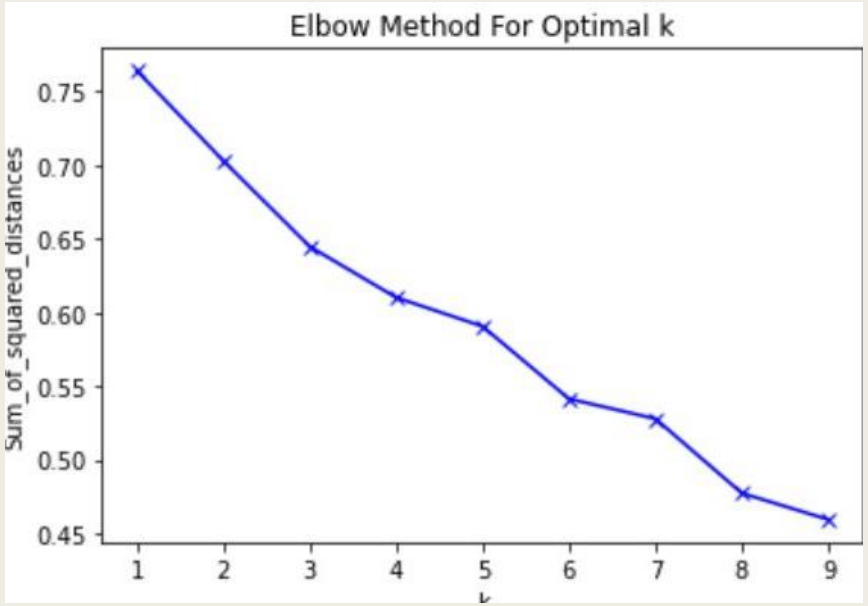
Methodology

- First, we imported gedodata created a visual map of Manhattan using Folium within Jupyter notebook utilizing Python language to depict the neighborhoods (visual shown).
- Three dataframes were created: Manhattan bookstores , Manhattan coffee shops, and Manhattan neighborhoods.
- k-Means was utilized to segment the neighborhoods based on similarity.
- Data normalization was used on the coffee shops & bookstores data to calculate a rating per each neighbhorhood.



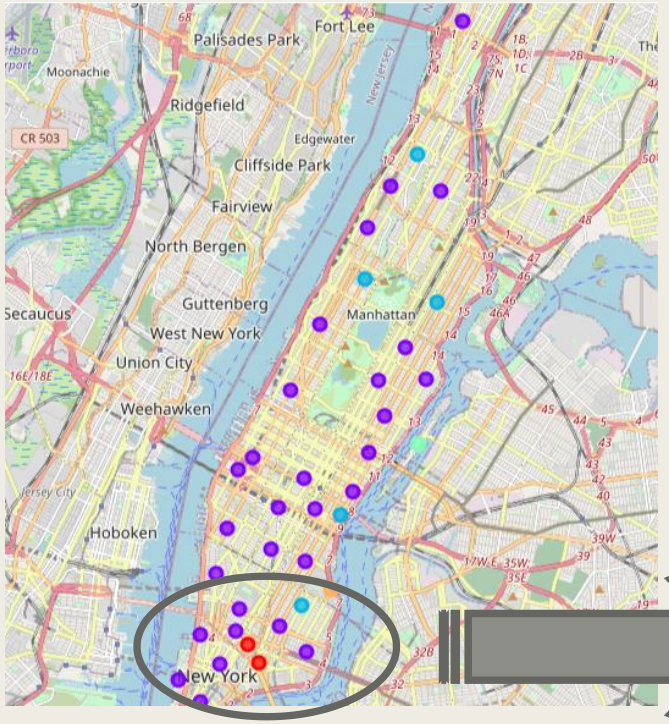
Results

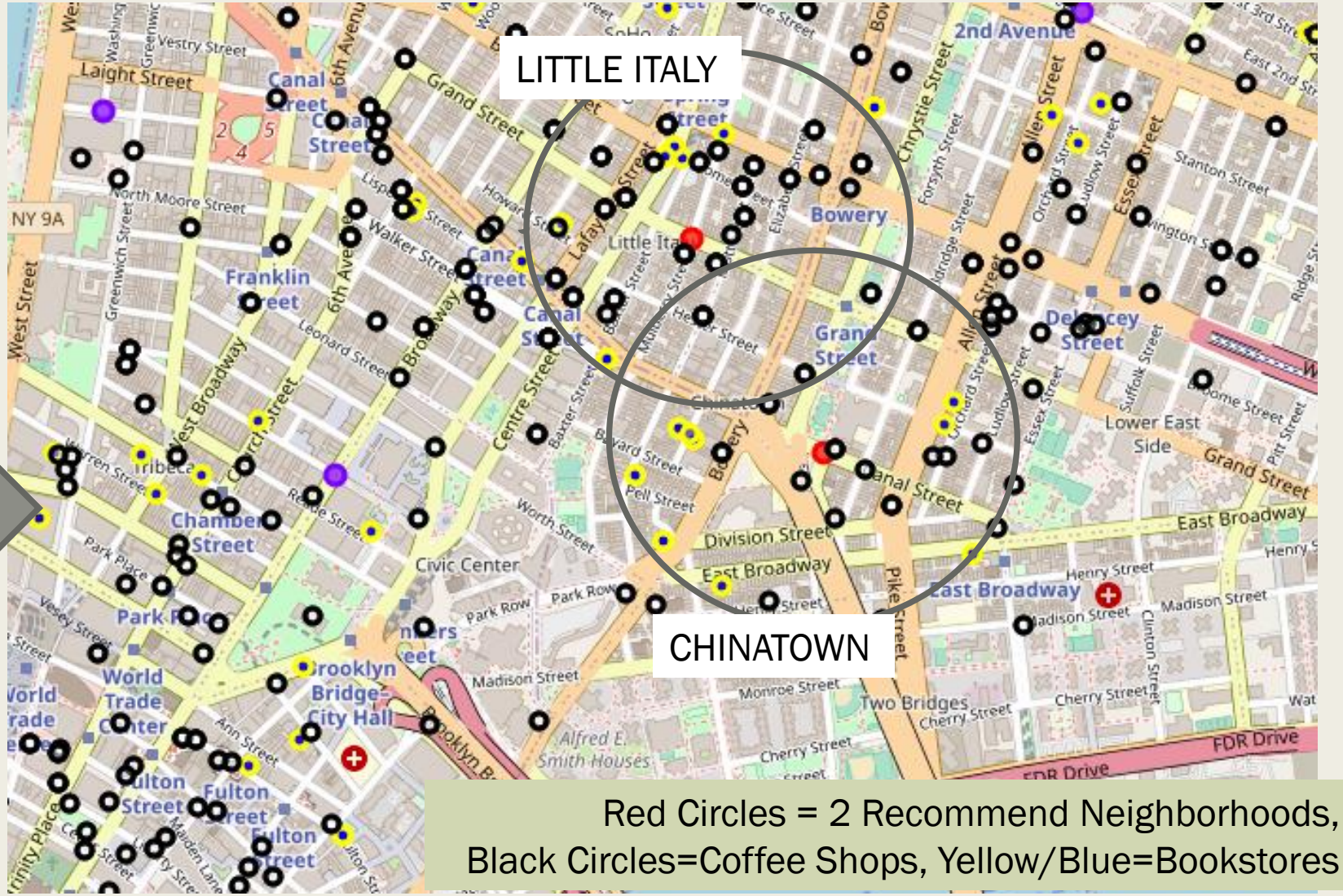
- There are 1174 coffee shops and 287 bookstores in Manhattan.
- Manhattan neighborhoods can be segmented into 5 neighborhoods based on similarity (elbow method, shown).
- It was discovered that one cluster containing two neighborhoods was the ideal location to search real estate listings for opening the coffee shop.
  - *Little Italy & Chinatown*



Cluster Labels	Count of Bookstores	Count of Coffee Shops	Count of Neighborhoods	Normalized Bookstores	Normalized Coffee Shops	Rating	Percentage B to C
0	30.0	96.0	2	0.129870	0.099792	0.012960	31.250000
1	231.0	962.0	30	1.000000	1.000000	1.000000	24.012474
2	25.0	113.0	6	0.108225	0.117464	0.012713	22.123894
3	1.0	3.0	1	0.004329	0.003119	0.000014	33.333333
4	0.0	0.0	1	0.000000	0.000000	0.000000	NaN

Results – 2 Manhattan Neighborhoods





Red Circles = 2 Recommend Neighborhoods,
Black Circles=Coffee Shops, Yellow/Blue=Bookstores

Discussion

- Observations:
  - *Manhattan has a lot of coffee shops – not much of a surprise!*
  - *The recommended area has quite a few bookstores close to each other and less coffee shops. This is ideal for our business venture!*
    - Bayard St. in particular looks very appealing based on number of book stores.
  - *If we had Premium API access with Foursquare, we could have evaluated these bookstores and competing coffee shops in these two neighborhoods to review user check-ins (frequency) to get a sense of typical volumes of customers.*
    - Would recommend this for deeper analysis
  - *If we had real estate commercial listings for this area, we could calculate down to the city block where within these two neighborhoods our business could have the best opportunity.*
    - Combining this with average numbers of customers in the area from check-in data could allow us to calculate high level ROIs for real estate considerations.

Red Circles = 2 Recommend Neighborhoods,
Black Circles=Coffee Shops, Yellow/Blue=Bookstores

Conclusion

- Recommended coffee shop location would be in Little Italy or Chinatown.
- This was determined using k-Means to segment neighborhoods.
- Based off this segmentation, we use Foursquare API data with normalization to visually depict ideal locations within these two neighborhoods.
- Next steps would be to review real estate options in these two neighborhoods and evaluate.