

Location Analysis for a new Smoothie and Juice Shop in Boston Area

Maheesha Tennakoon

IBM Coursera Applied Data Science Specialization

Capstone Project

July 2020

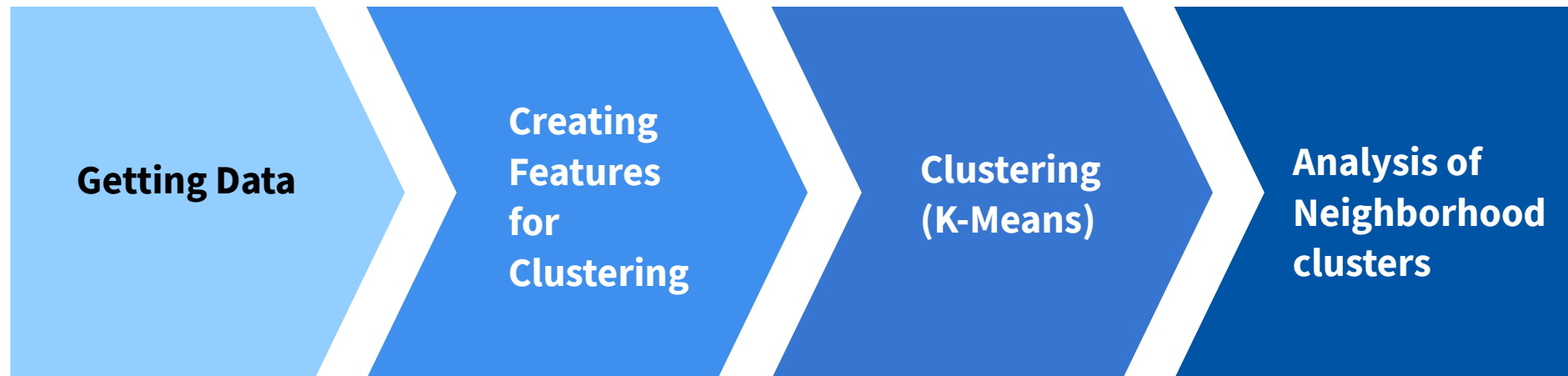
Business Problem

- When establishing a new business, the business owners can benefit higher sales and a good profit if they open their shop in a neighborhood having a higher demand for the product and less supply.
- This project is intended to solve this business problem for a smoothie and juice shop to provide recommendations on the locations with higher business potential.

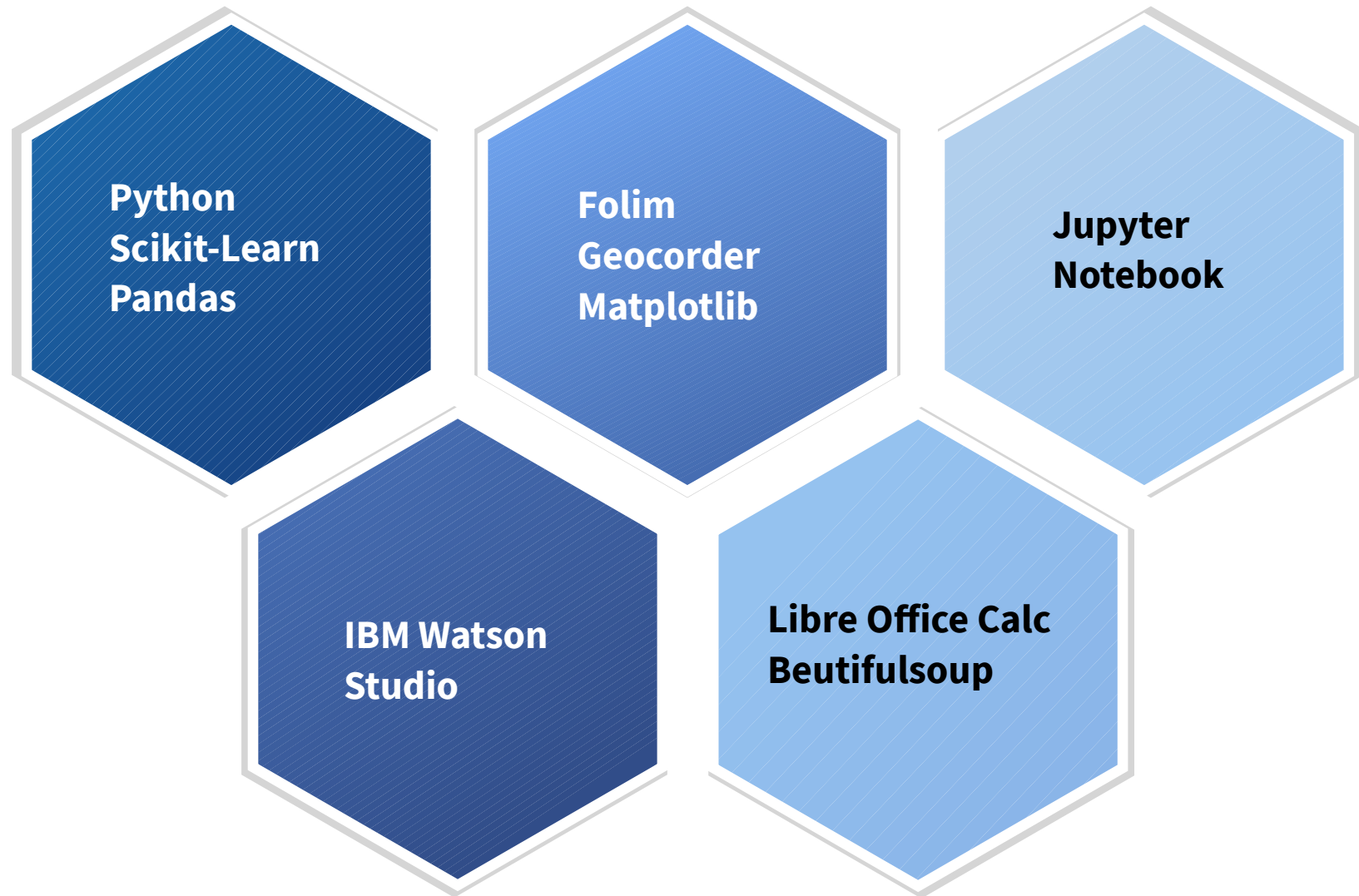
Data

- **List of neighborhoods in Boston, MA, USA.**
 - Wikipedia page: https://en.wikipedia.org/wiki/Neighborhoods_in_Boston
- **Latitude and longitude of each neighborhoods.**
 - Python geocoder package
 - Project Link: <https://github.com/DenisCarriere/geocoder>
- **Demographic Data for Boston's Neighborhoods**
 - Boston Neighborhood Demographics, 2013-2017 American Community
 - Data Link: <https://data.boston.gov/dataset/neighborhood-demographics>
- **Venue data about other business in each neighborhood.**
 - Foursquare API
 - API Link: <https://api.foursquare.com/v2/venues/explore>

Methodology



Software and Tools Used



Results

Cluster 4 Neighborhoods

[37]:

| Neighborhood_name | Allston | Back Bay | Brighton | Chinatown | East Boston | Fenway Kenmore | Hyde Park | South End | West End |
|--------------------------------|----------|----------|----------|-----------|-------------|----------------|-----------|-----------|----------|
| Median Income | 0.757530 | 1.645742 | 1.000326 | 1.086198 | 0.853507 | 0.637685 | 1.141712 | 1.402667 | 1.560554 |
| Total Households | 0.024530 | 0.037321 | 0.082077 | 0.008031 | 0.061870 | 0.041508 | 0.048973 | 0.061517 | 0.011906 |
| Total Population | 0.028936 | 0.027162 | 0.077388 | 0.006641 | 0.069722 | 0.048715 | 0.055434 | 0.047881 | 0.009225 |
| Median Age | 0.812500 | 1.031250 | 0.906250 | 1.031250 | 1.062500 | 0.718750 | 1.218750 | 1.093750 | 1.062500 |
| Smoothie_Juice | 0.000000 | 0.037037 | 0.000000 | 0.000000 | 0.100000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| Gym_Fitness | 0.750000 | 0.111111 | 0.428571 | 0.454545 | 0.500000 | 0.333333 | 1.000000 | 0.090909 | 0.571429 |
| Dessert_Drink | 0.000000 | 0.111111 | 0.285714 | 0.181818 | 0.300000 | 0.000000 | 0.000000 | 0.090909 | 0.000000 |
| Park | 0.000000 | 0.000000 | 0.000000 | 0.181818 | 0.100000 | 0.166667 | 0.000000 | 0.363636 | 0.142857 |
| Transportation | 0.000000 | 0.037037 | 0.142857 | 0.000000 | 0.000000 | 0.166667 | 0.000000 | 0.000000 | 0.000000 |
| Shopping | 0.250000 | 0.555556 | 0.142857 | 0.090909 | 0.000000 | 0.333333 | 0.000000 | 0.181818 | 0.142857 |
| Books_Gifts | 0.000000 | 0.148148 | 0.000000 | 0.090909 | 0.000000 | 0.000000 | 0.000000 | 0.272727 | 0.142857 |
| Smoothie_Juice_per_Gym_Fitness | 0.000000 | 0.333333 | 0.000000 | 0.000000 | 0.200000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| Smoothie_Juice_per_Shopping | 0.000000 | 0.066667 | 0.000000 | 0.000000 | -1.000000 | 0.000000 | -1.000000 | 0.000000 | 0.000000 |

Results (Clusters)

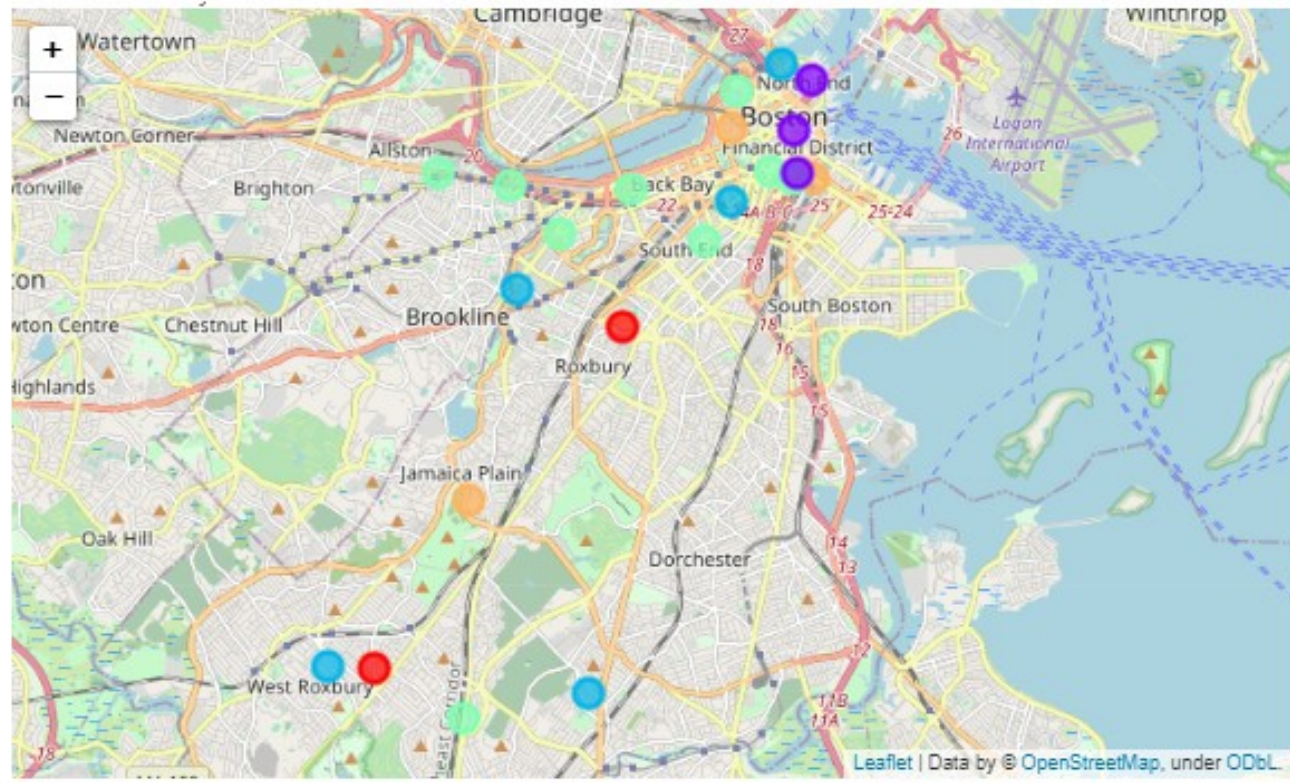
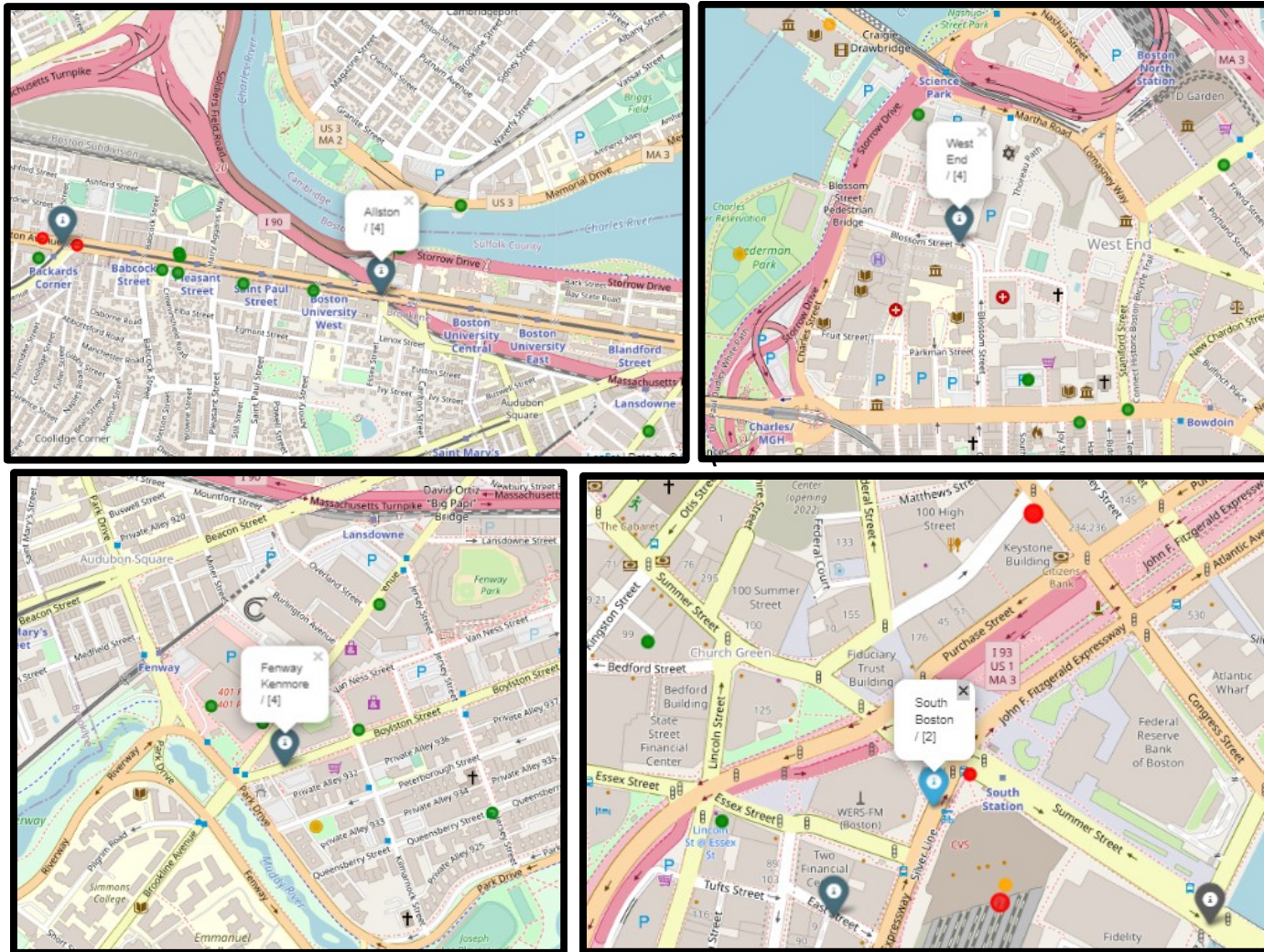
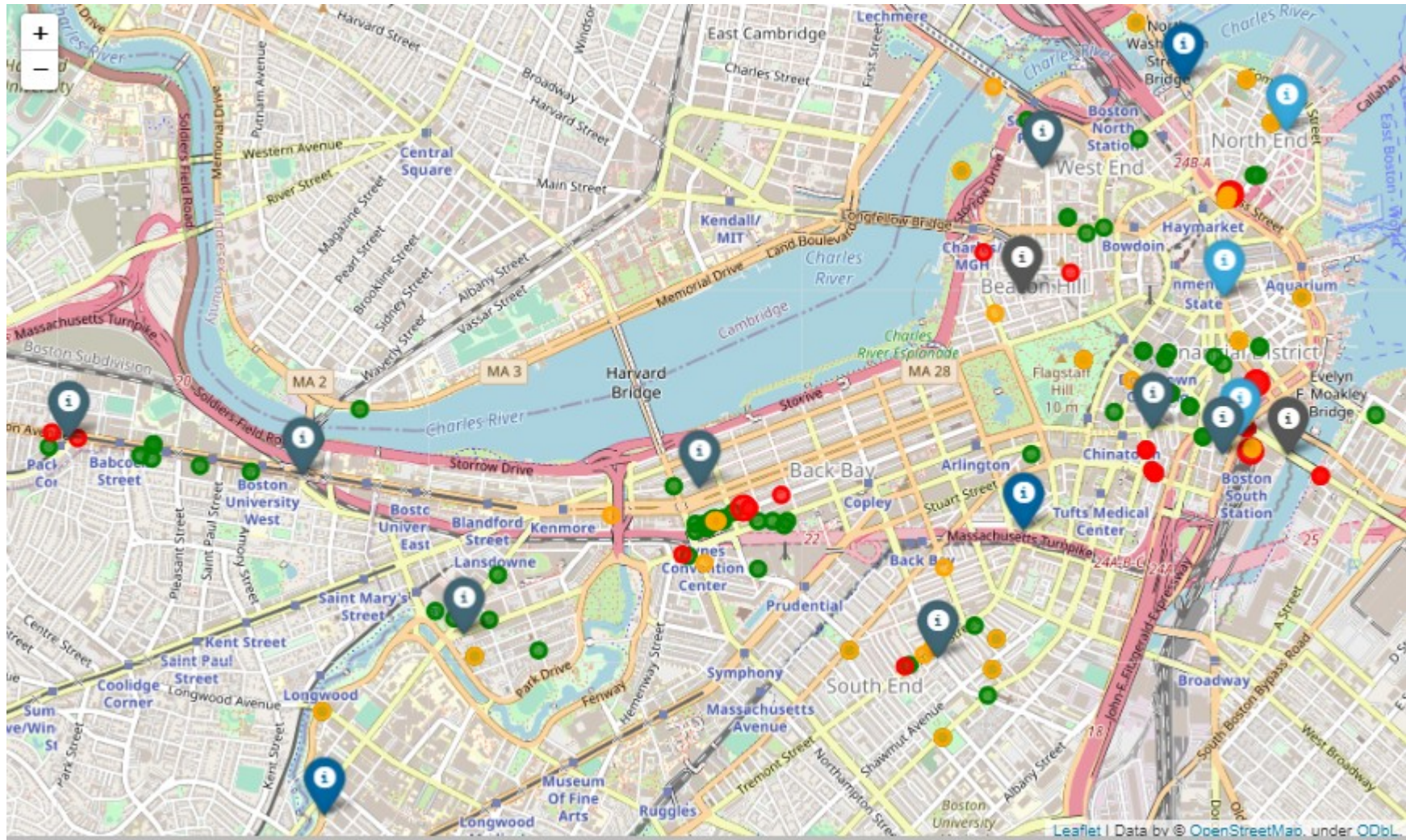


Figure 2: Cluster Map: cluster 1 (red), cluster 2 (purple), cluster 3 (blue), cluster 4 (green), cluster 5 (orange)

Results (Neighborhoods and Venues)



Dashboard Map



Conclusion

- Neighborhoods in Boston, MA were clustered using k- means clustering algorithm providing demographic and venues data.
- Neighborhoods and the clusters were analyzed to get insights into business opportunity to open and successfully run a Smoothies and Juice shop.
- The neighborhoods Allston, Brighton, Fenway Kenmore, South End and West End are clustered together and found those neighborhoods have venues favorable to a nearby Smoothies and Juice shops.
- The data also showed the absence of competitive businesses in the area.
- The demographic data also provided insights favorable to opening a Smoothies and Juice shop in any of these neighborhoods.

Reference

1. "Segmenting and Clustering Neighborhoods in New York City" by Alex Aklson and Polong Lin, Cognitive Class.
2. Neighborhoods in Boston, https://en.wikipedia.org/wiki/Neighborhoods_in_Boston
3. Geocoder, <https://github.com/DenisCarriere/geocoder>
4. Boston Neighborhood Demographics, 2013-2017 American Community, <https://data.boston.gov/dataset/neighborhood-demographics>
5. Foursquare API Documentation, <https://developer.foursquare.com/docs/api-reference/venues/search/>
6. Scikit-Learn Kmean, <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>
7. Folium, <https://python-visualization.github.io/folium/>
8. BeautifulSoup, <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>