

# **IBM Applied Data Science Capstone Project:**

## **The Battle of Seattle Neighborhoods**

Ellen Zhang  
May 15, 2020

### 1. Introduction

Graduating from college is a big deal! My parents are just as excited as I am, and they were planning to fly to Seattle to attend my graduation. Unfortunately, the plan is cancelled due to the COVID-19 outbreak... But if they were to come here in Seattle to attend my college graduation, where would be the best location for them to stay? The most ideal location should be close to local attractions, nearby restaurants, public transportation, and more importantly, it can't be too far from my school! In this project, I will be using what we learned in this course to try to find this ideal location.

Besides my parents, all parents who wish to come attend their children's college graduation in Seattle can refer to this project as their guide to find a place to stay, if they also consider having a variety of restaurant choices is considered the most important factor for and "ideal" location to stay.

### 2. Data

In this project, I have searched for as well as generated the following data to solve the problem:

- 1) Geographical coordinates of my school (University of Washington) using Google Maps API geocoding
- 2) Centers of candidate areas and their approximate addresses using Google Maps API geocoding

3) Number of restaurants and their ratings in every candidate neighborhood using Foursquare API

4) Number of attractions in every candidate neighborhood using Foursquare API

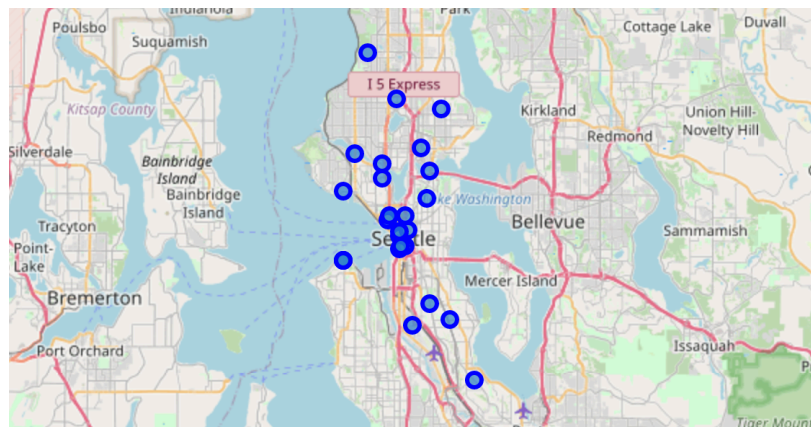
### 3. Methodology

In this project we will mainly search for areas of Seattle that have high restaurant density. We will limit our analysis to areas about 6 miles around university of Washington.

In the first step, we have collected a list of neighborhoods in the Great Seattle area by scraping and transforming data from Wikipedia. Then, we generated geographical coordinates for all 27 of these neighborhoods.

	Neighborhood	Latitude	Longitude
0	► Ballard, Seattle (1 C, 16 P)	47.657813	-122.376231
1	► Beacon Hill, Seattle (10 P)	47.562010	-122.305145
2	► Belltown, Seattle (20 P)	47.615760	-122.344640
3	► Broadview, Seattle (3 P)	47.722380	-122.364980
4	► Capitol Hill, Seattle (46 P)	47.629700	-122.308610
5	► Cascade, Seattle (4 P)	47.589357	-122.388168
6	► Central District, Seattle (12 P)	47.605530	-122.334320
7	► Central Waterfront, Seattle (13 P)	39.952224	-75.141729
8	► Chinatown-International District, Seattle ...	47.599170	-122.327960
9	► Denny Triangle, Seattle (13 P)	47.596800	-122.334230
10	► Downtown Seattle (3 C, 64 P)	47.589357	-122.388168
11	► Eastlake, Seattle (6 P)	47.618157	-122.329181
12	► First Hill, Seattle (1 C, 10 P)	47.608780	-122.326430

Then, we created a map of Seattle with all the name of the neighborhoods.



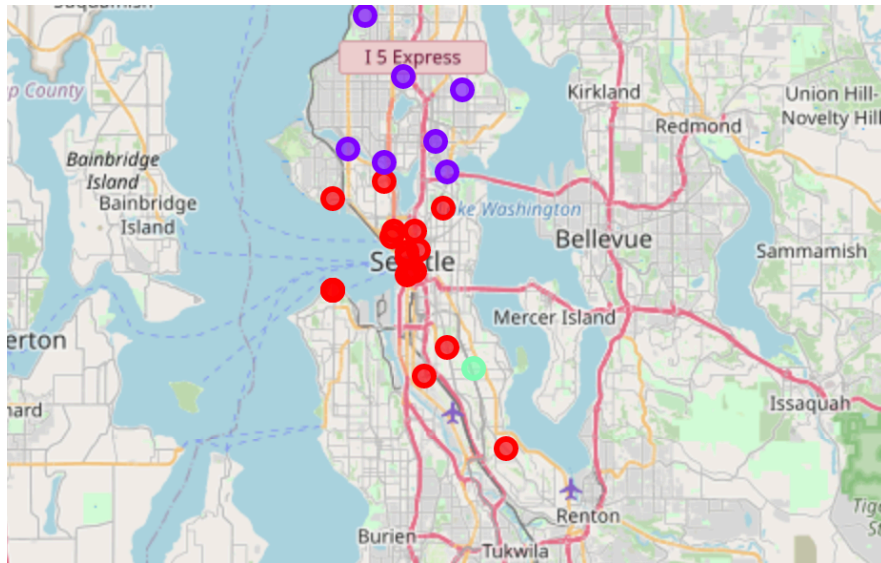
In the next step, we use Foursquare data to analyze these neighborhoods. We found all types of venues as well as the number of them in these neighborhoods such as grocery stores, fish market, ice cream shop, restaurants, etc.

	Neighborhoods	African Restaurant	Airport	American Restaurant	Antique Shop	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	...	Video Game Store	Video Store	Vietnamese Restaurant	Waterfront	Wine Bar	Wine Shop
0	► Ballard, Seattle (1 C, 16 P)	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1	► Ballard, Seattle (1 C, 16 P)	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2	► Ballard, Seattle (1 C, 16 P)	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
3	► Ballard, Seattle (1 C, 16 P)	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4	► Ballard, Seattle (1 C, 16 P)	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0

We then filtered this raw data and collected data only on restaurant densities on those qualifying areas, areas within 6 miles radius from my school.

	Neighborhoods	Restaurant
0	► Ballard, Seattle (1 C, 16 P)	0.00
1	► Beacon Hill, Seattle (10 P)	0.01
2	► Belltown, Seattle (20 P)	0.01
3	► Broadview, Seattle (3 P)	0.00
4	► Capitol Hill, Seattle (46 P)	0.01

In the final step, we presented a map of all such locations as well as clusters (using k-means clustering) of those locations to identify general neighborhoods that could be considered as an optimal zone for my parents to stay during their visit for my graduation.



#### 4. Results and Discussion

Most of the restaurants are concentrated in rainier valley, with the highest number in cluster 2 and moderate number in cluster 1. On the other hand, cluster 0 has a relatively small number of restaurants in the neighborhoods. This means that there are not a lot of choices for restaurants in that neighborhood and this area might not be the most ideal location for my parents to stay. Meanwhile, number of restaurants in cluster 2 ranks top in all neighborhoods and might be a great choice to stay. Therefore, this project indicates that if my parents were here to attend graduation and were looking for somewhere with a lot of restaurants to stay, cluster 2 might be their best choice, where cluster 0 will be the next option and cluster 1 for the last option.

#### 5. Conclusion

In this project, through thorough analysis with the help from various sources, we can conclude that if my parents were here to attend graduation and were looking for somewhere with a lot of restaurants to stay, rainier valley would be the best neighborhood for them to stay.