Exploring the Neighbourhoods of Seattle

Thato Kagiso Malema

March 3, 2020

1 Introduction and Business Problem

The city of Seattle also known as *The Emerald city* is the largest city in the state of Washington and is the inspiration behind Grey's Anatomy's hospital, Seattle Grace / Greys Sloan Memorial Hospital which is set in the city. Seattle, with a Mediterranean climate is home to the headquarters of the two tech giants, Microsoft and Amazon. The city has a rich culture and up to this day still maintain's its status as a regional centre of the performing arts.

Our project is defined by the following problem. A very famous and fast growing vegan cakes and ice cream restaurant franchise wants to open its doors in the city of Seattle after establishing itself in Europe and some of the vibrant cities in the United States of America. The restaurant is to be opened in a "popular" district with a lot of neighbourhoods that are already well established with lots of restaurants, pubs, movie theaters and etc. The proposed district should be the most popular in the whole city with more than five neighbourhoods. The franchise is hoping to be provided with a list of these neighbourhoods so that they can make their final decision based on site visits to the various neighbourhoods in one of the districts of Seattle.

Figure 1: The city scape of Seattle[?]

2 Data : Acquisition and Cleaning

2.1 Acquisition

The city of Seattle is split into neighbourhoods and those neighbourhoods are further organized into larger districts. We scraped our data from a Wikipedia and the data is mostly based on the maps provided by the Seattle City Clerk's Neighborhood Atlas. This was the only data in a tabulated format that we could acquire and its form made our data wrangling much easier. The data included both the neighbourhoods and their corresponding districts and that was left was to clean the data and include the coordinates (latitudes and longitudes) of the different districts. We used Geopy to add coordinates to our different data frames and the Foursquare API to explore the different venues.

2.2 Cleaning

The scraped data had a lot of reference values on most of the values in the Neighbourhood column and so we corrected the names manually after dropping columns such as *Street map and Notes* which were not of use for this project. Some of the data included multiple values, like a district that overlapped with another and so we assigned those particular districts to the most common in values or to the first district in the entry. Our clean data contained 27 larger districts. For the Foursquare acquired data, the json file format was flattened into a data frame and further exploited to get relevant

values to process cleaning.

3 Methodology

Our target variable in our case is the most *popular* district and the exploration of the venues of the different neighbourhoods in this district. Our project is not as numerically motivated as the standard data analysis case and we relied a lot on the Foursquare API data calls that we made in our notebook. The client required us to ensure that the categories of the different venues in these neighbourhoods are varying and that the neighbourhoods are well established in terms of uniqueness of the different categories of venues. We made an API call for each of the different districts to find the popular venues in each of these districts. The top three districts were Downtown which returned 122 venues, followed by South End and Central Area which both returned 109 venues each.
Since the district of Downtown was the most *popular*, we concentrated our analysis on the 10 neighbourhoods of Downtown. After adding coordinates of the different neighbourhoods, we made API calls for each neighbourhood and combined all the results into one data frame with over 800 venues using the features : Neighbourhood, Venue Name, Categories (Thai Restaurant, movie theater etc.) and the coordinates of each venue. We grouped the different categories together to determine the total number of venues in each category and used our grouped data to make bar charts of the different categories.

4 Results

The API calls returned 867 venues with 7 out of the 10 neighbourhoods returning top 100 venues each with the Denny Triangle neighbourhood returning 95 venues. The category with the most venues was Coffee Shop with the second most category being hotels.
The dessert venues account for 5.882% of the total venues and of the Ice Cream Shop category, 4 of the 10 Neighbourhoods have Ice Cream Shops, namely; Denny Triangle, Central Business District, International District and West Edge.

Our analysis concludes with the number of dessert restaurants and shops

	venues	number of venues for each category
7	Bakery	17
11	Donut Shop	15
42	Ice Cream Shop	5
52	Chocolate Shop	4
65	Dessert Shop	4
71	Vegetarian / Vegan Restaurant	3
101	Frozen Yogurt Shop	2
125	Snack Place	1

Figure 2: Categories of Dessert Venues.

in each Neighbourhood. Belltown and West Edge both have 7 dessert shops in their respective neighbourhoods followed by Central Business District with 6 shops and International District with 5 shops. Pike-Market, Pioneer Square and Central Waterfront each have 4 dessert shops and the Neighbourhood with the least amount of dessert venues is First Hill with 2 venues and Yesler Terrace with no dessert venues.

5 Discussion

From our analysis, the most popular district is the Downtown district and the neighbourhood with the most dessert shops is Belltown. Yesler Terrace either has no dessert shops or has no recorded data about the amount of dessert shops in the neighbourhood. In our analysis, there are no vegan ice cream and cake shops in any of the neighbourhoods. It seems that it would be a profitable idea to exploit the 100% market of the dessert shops in Yesler Terrace and establish the vegan ice cream shop in that neighbourhood. There is of course the issue of Yesler Terrace only having 27 venues that are frequently visited and that could mean that people do not visit the neighbourhood as much as all the other neighbourhoods in the district.

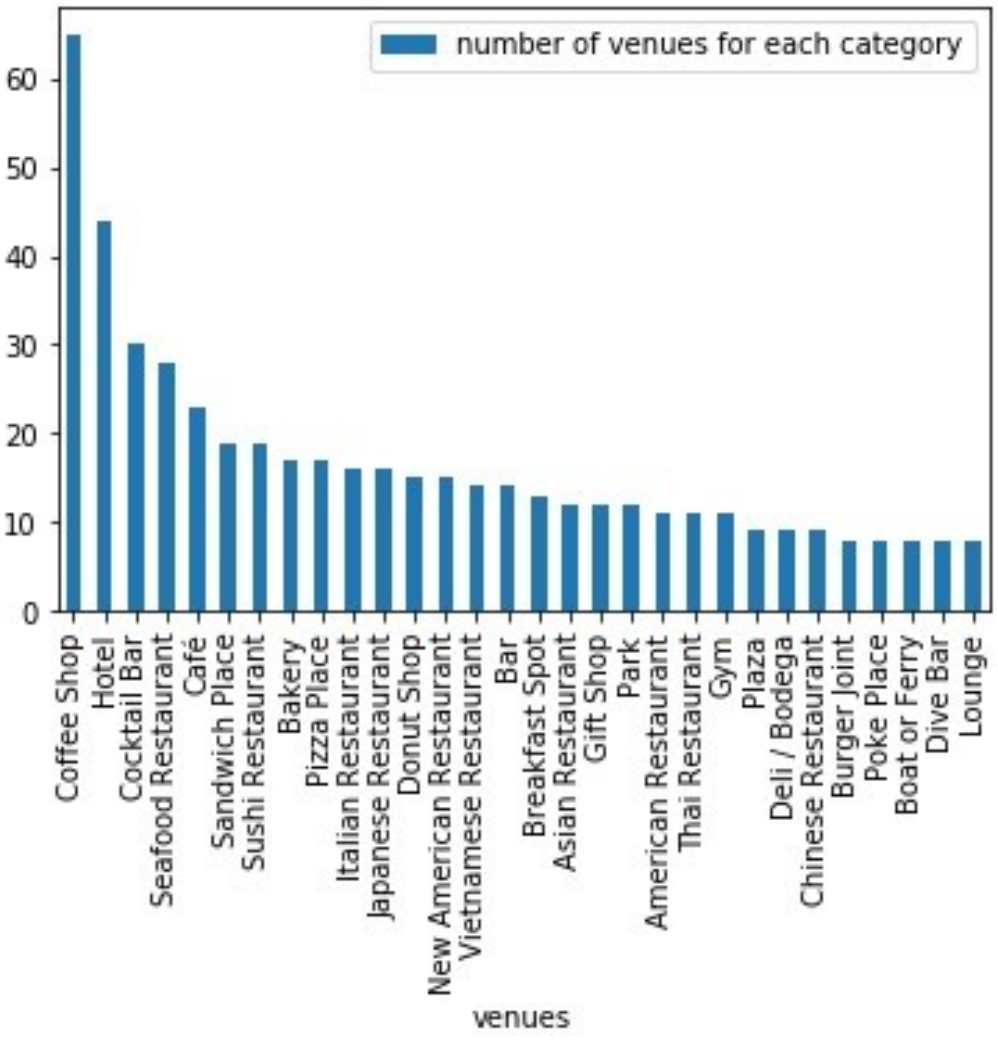

Figure 3: The top 30 Categories.

neighborhood	venue name	categories	latitude	longitude
Belltown	7	7	7	7
Central Business District	6	6	6	6
Central Waterfront	4	4	4	4
Denny Triangle	5	5	5	5
First Hill	3	3	3	3
International District	5	5	5	5
Pike-Market	4	4	4	4
Pioneer Square	4	4	4	4
West Edge	7	7	7	7

Figure 4: Dessert venues in every Neighbourhood.

The other option would be to establish the vegan ice cream shop in the most popular neighbourhood (using the metric of both the amount of popular venues and the number of dessert shops), Belltown. Belltown does not have any ice cream shops and so the competition in that regard is much less than in a neighbourhood like Central Business District which is one of the popular neighbourhoods with 2 ice cream venues and out of the 6 dessert venues. There is of course the cake aspect of the franchise that we need to consider as there are quite a number of bakeries in this district. The Belltown neighbourhood might not have any ice cream shops but does have the most bakeries out of all the neighbourhoods in the district and that may prove to be problematic for the franchise with regards to competition.

6 Conclusion

An exploration of at least 3 of the most popular districts would have provided an even more comprehensive analysis and is of course out the scope of this project. For future analysis, it would create a better picture of the venues of

Seattle if we explored more neighbourhoods in different districts. Exploring the most popular neighbourhood in each district and drawing a comparison between them would have increased the completeness our study but this is also beyond the scope of our defined project.
The franchise can use the data provided in the notebook and report to help with the site seeing of the possible neighbourhoods for the establishment of the vegan ice cream shop.