

Coursera Capstone Project

The Battle of Neighborhoods

Rishabh Karmakar

28th April 2020

Introduction

- Helsinki is the capital and most populous city in Finland and has not only its own unique vibe of a charming Scandinavian city but also a contemporary feel.
- In 2017, according to , the number of overnight stays in Helsinki was about 5.3 million for both domestics and international visitors. Helsinki's main attractions are urban culture with the diversity in events and foods. In this final project, we are going to focus on exploring Helsinki in different categories in all areas to observe what people have been enjoyed when they come to Helsinki using Foursquare.

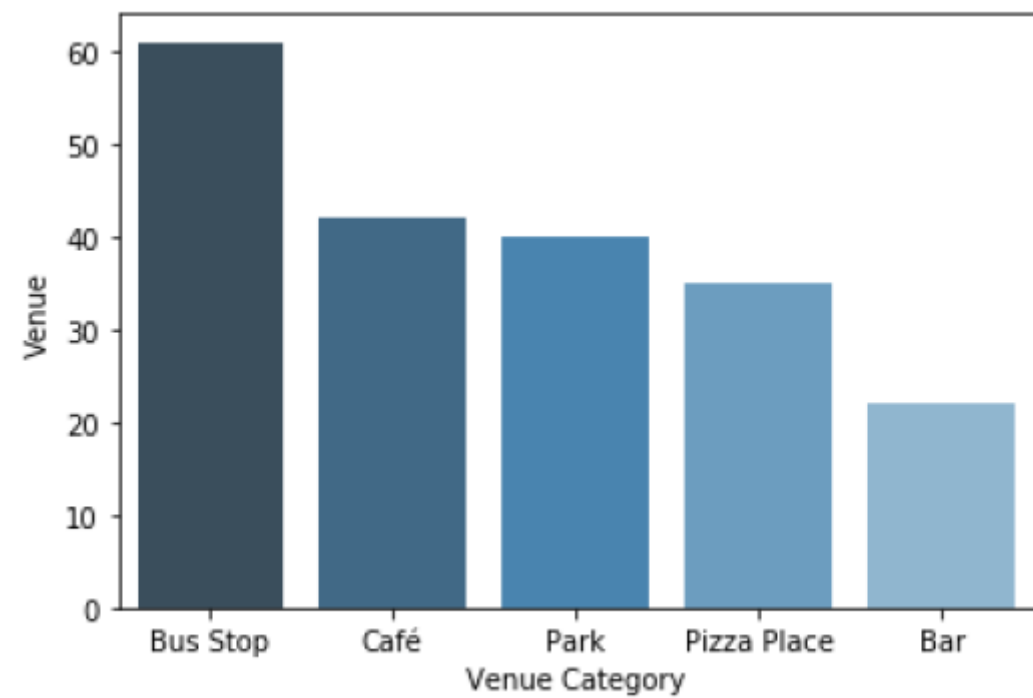
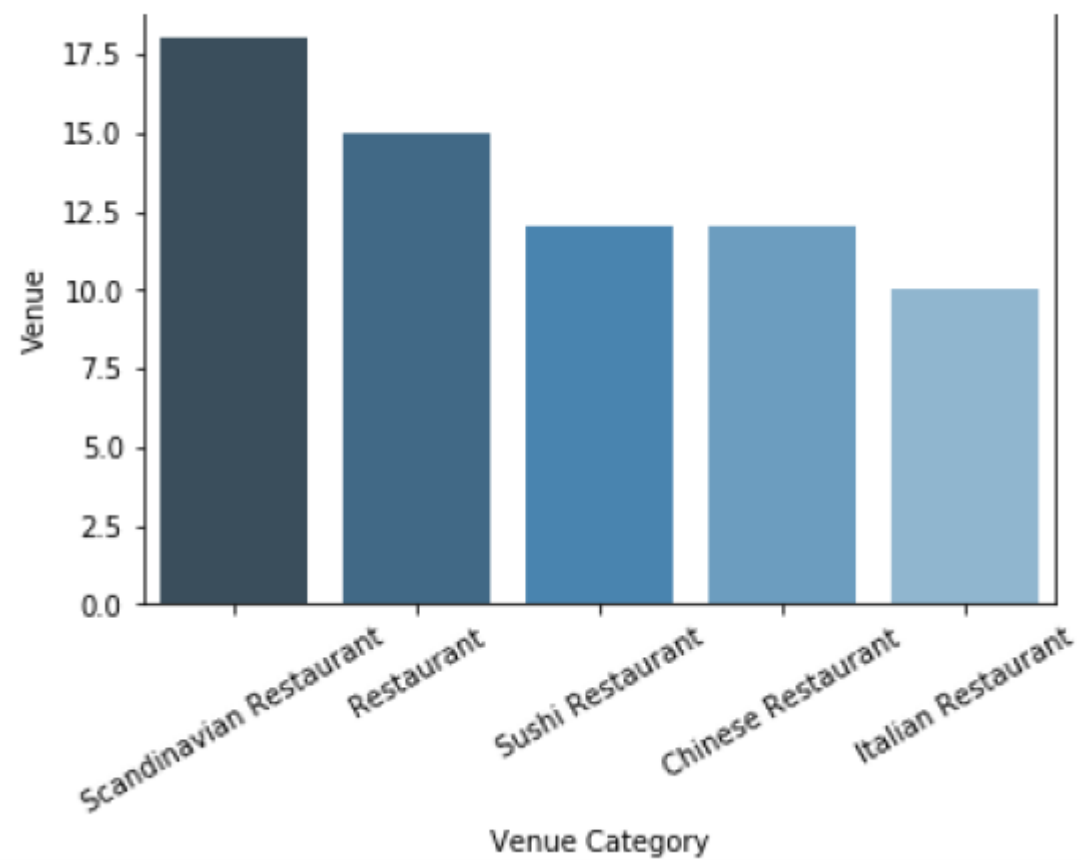
Data

- Since we are going to clustering the data, having the neighborhood information helps us in grouping the data into each district in Helsinki. Using Google and BeautifulSoup4 library in Python, a data-frame contains the name of 51 subdivisions in the capital of Finland is created.
- Using Geopy Client, a python library that helps us getting the latitudes and longitudes of all districts. Combining with the list of subdivisions above, we can have a completed data frame as below:
- Using Foursquare API, we can get a very large database of nearby venues for each neighborhood in JSON format. Here in this table, we are able to choose 100 popular locations maximum for each district within a 1 km radius.

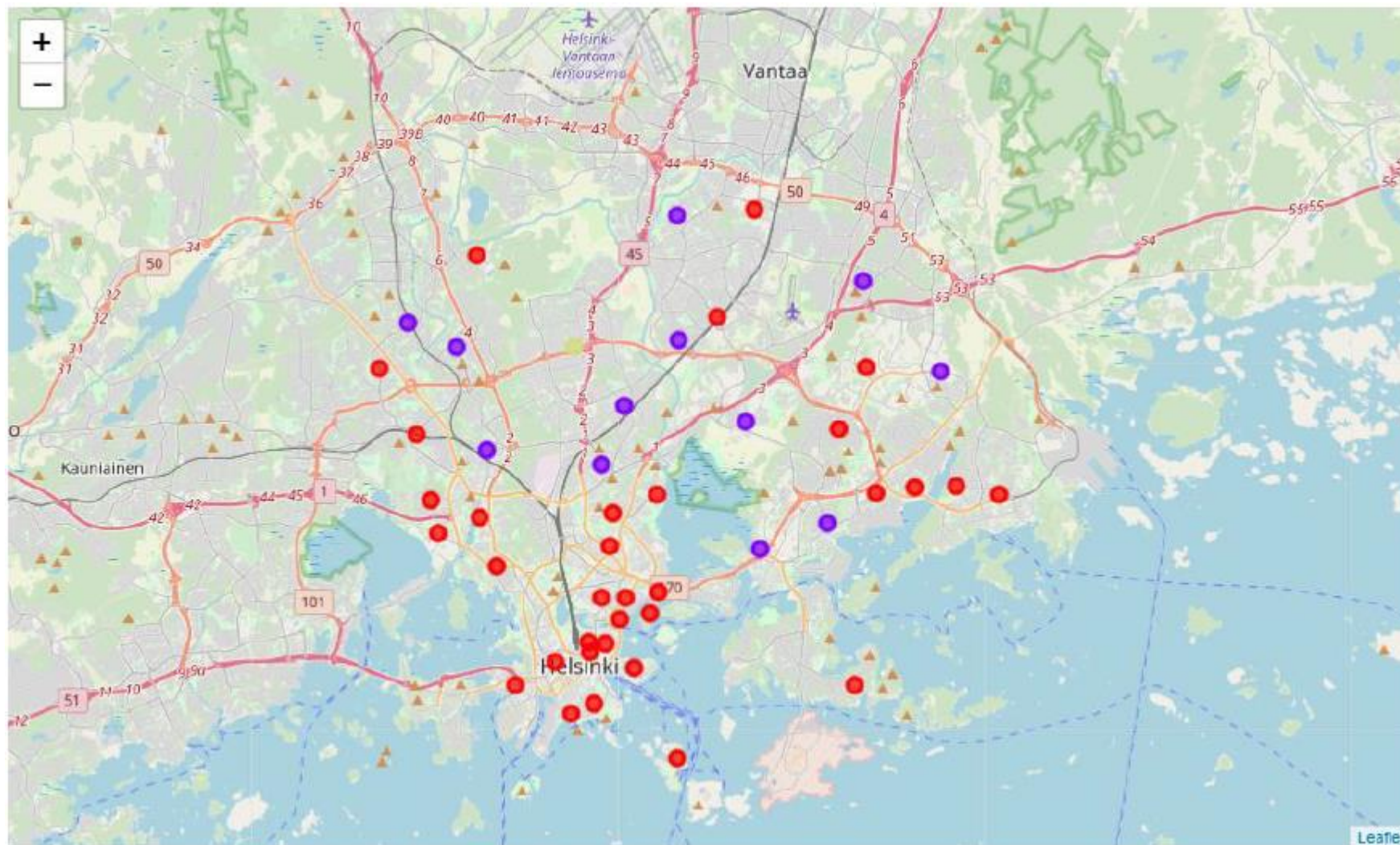
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude		Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kamppi	60.168535	24.930494		Cafe Rouge	60.168711	24.933027	Middle Eastern Restaurant
3	Kamppi	60.168535	24.930494		Pobre	60.169500	24.933484	Filipino Restaurant
12	Kamppi	60.168535	24.930494		Singapore Hot Wok	60.169163	24.933890	Malay Restaurant
14	Kamppi	60.168535	24.930494		Sushibar + Wine	60.167187	24.932170	Sushi Restaurant
15	Kamppi	60.168535	24.930494		Hoku	60.169490	24.933543	Hawaiian Restaurant
...
1162	Kaisaniemi	60.173182	24.946400		Momotoko	60.169834	24.946957	Ramen Restaurant
1169	Kaisaniemi	60.173182	24.946400		Döner Harju City	60.170082	24.941644	Doner Restaurant
1173	Kaisaniemi	60.173182	24.946400	Kiinalainen Ravintola Family		60.174245	24.950121	Chinese Restaurant
1178	Kaisaniemi	60.173182	24.946400	Fazer Food & Co Oodi		60.173648	24.937799	Restaurant
1195	Kaisaniemi	60.173182	24.946400	Ravintola Kaisaniemi		60.176665	24.941657	Restaurant

Methodology

- There are 228 unique venue categories in total and surprisingly as the surprising result, **Bus stop** is on top of the list. People might find it beautiful and attractive. Comes in the top 5 are **Cafe, Park, Pizza Place, and Bar**. Additionally, when it comes to food categories, it's clear that **Scandinavian** restaurant takes the win alongside **Chinese, Sushi, and Italian** restaurants.
- Overall, A data frame for venue contributions grouped by each neighborhood is created. (noted that each one can query 100 venues maximum)



Result - Clustering Neighbourhoods



Conclusion and Summary

- After using k-means clustering, we can specify that **Cluster 1** is a group of neighborhoods that has the common known for good restaurants and pubs while **Cluster 2** is mainly reviewed by local people. Meanwhile, **Cluster 3** and **Cluster 4** are classified when contains some unexpected venues categories such as **Gas Station**, **Garden**, etc.
- By using data collected in Foursquare, we have been able to have a small project about analyzing Helsinki venue information. Even though the amount of available data from Foursquare for the Helsinki area is still quite limited, A small guidance for tourists is made in order to provide solutions for future individuals and businesses to improve.