



Applied Data Science Capstone

Capstone Project - The Battle of Neighborhoods

Wency Yan

Introduction

► Problem

- My friend wants to move to Vancouver, Canada, and open a gym there
- Not sure where he should open the gym
- Wants to make sure not too many other gyms/indoor sports facilities are located close by already

► Target audience

- My friend - looking to invest in opening a gym



Data

► Sources:

- Vancouver neighborhood - from Wikipedia (borough and neighborhood data by postal code):
 - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V
- Latitude & Longitude coordinates by postal code - from GeoNames.org:
 - <http://download.geonames.org/export/zip/> (CA.zip file)
- Nearby venues data - from Foursquare

► Steps:

- Downloaded and scraped the neighborhood data and coordinates data
- Combined the neighborhood data with coordinates data
- Narrowed data to contain only postal codes in the Vancouver borough
- Used Foursquare to get nearby venues data (Top 100 venues, radius=500m)

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	V6A	Vancouver	Strathcona, Chinatown, Downtown Eastside	49.2779	-123.0908
1	V6B	Vancouver	NE Downtown, Gastown, Harbour Centre, Internat...	49.2788	-123.1139
2	V6C	Vancouver	Waterfront, Coal Harbour, Canada Place	49.2866	-123.1158
3	V6E	Vancouver	SE West End, Davie Village	49.2833	-123.1298
4	V6G	Vancouver	NW West End, Stanley Park	49.2990	-123.1408
5	V6H	Vancouver	West Fairview, Granville Island, NE Shaughnessy	49.2559	-123.1322
6	V6J	Vancouver	NW Shaughnessy, East Kitsilano, Quilchena	49.2603	-123.1460
7	V5K	Vancouver	North Hastings-Sunrise	49.2807	-123.0397
8	V6K	Vancouver	Central Kitsilano, Greektown	49.2646	-123.1648
9	V5L	Vancouver	North Grandview-Woodland	49.2795	-123.0667
10	V6L	Vancouver	NW Arbutus Ridge, NE Dunbar-Southlands	49.2497	-123.1660
11	V5M	Vancouver	South Hastings-Sunrise, North Renfrew-Collingwood	49.2600	-123.0398
12	V6M	Vancouver	South Shaughnessy, NW Oakridge, NE Kerrisdale,...	49.2344	-123.1451
13	V5N	Vancouver	South Grandview-Woodland, NE Kensington-Cedar ...	49.2551	-123.0667
14	V6N	Vancouver	West Kerrisdale, South Dunbar-Southlands, Musq...	49.2302	-123.1890
15	V5P	Vancouver	SE Kensington-Cedar Cottage, Victoria-Fraserview	49.2220	-123.0683
16	V6P	Vancouver	SE Kerrisdale, SW Oakridge, West Marpole	49.2151	-123.1396
17	V5R	Vancouver	South Renfrew-Collingwood	49.2397	-123.0407
18	V6R	Vancouver	West Kitsilano, West Point Grey, Jericho	49.2686	-123.1976
19	V5S	Vancouver	Killarney	49.2175	-123.0380
20	V6S	Vancouver	NW Dunbar-Southlands, Chaldecutt, South Univer...	49.2491	-123.2088
21	V5T	Vancouver	East Mount Pleasant	49.2620	-123.0923
22	V6T	Vancouver	UBC	49.4635	-122.8220
23	V5V	Vancouver	West Kensington-Cedar Cottage, NE Riley Park-L...	49.2480	-123.0913
24	V5W	Vancouver	SE Riley Park-Little Mountain, SW Kensington-C...	49.2327	-123.0917
25	V5X	Vancouver	SE Oakridge, East Marpole, South Sunset	49.2156	-123.0979
26	V7X	Vancouver	Bentall Centre	49.2935	-123.1162
27	V5Y	Vancouver	West Mount Pleasant, West Riley Park-Little Mo...	49.2492	-123.1104
28	V7Y	Vancouver	Pacific Centre	49.2819	-123.1190
29	V5Z	Vancouver	East Fairview, South Cambie	49.2475	-123.1210
30	V6Z	Vancouver	SW Downtown	49.2767	-123.1300

Methodology

► Data exploration

- Using Foursquare to get the Top 100 nearby venues data (within 500 meters radius) for each postal code in Vancouver neighborhood data

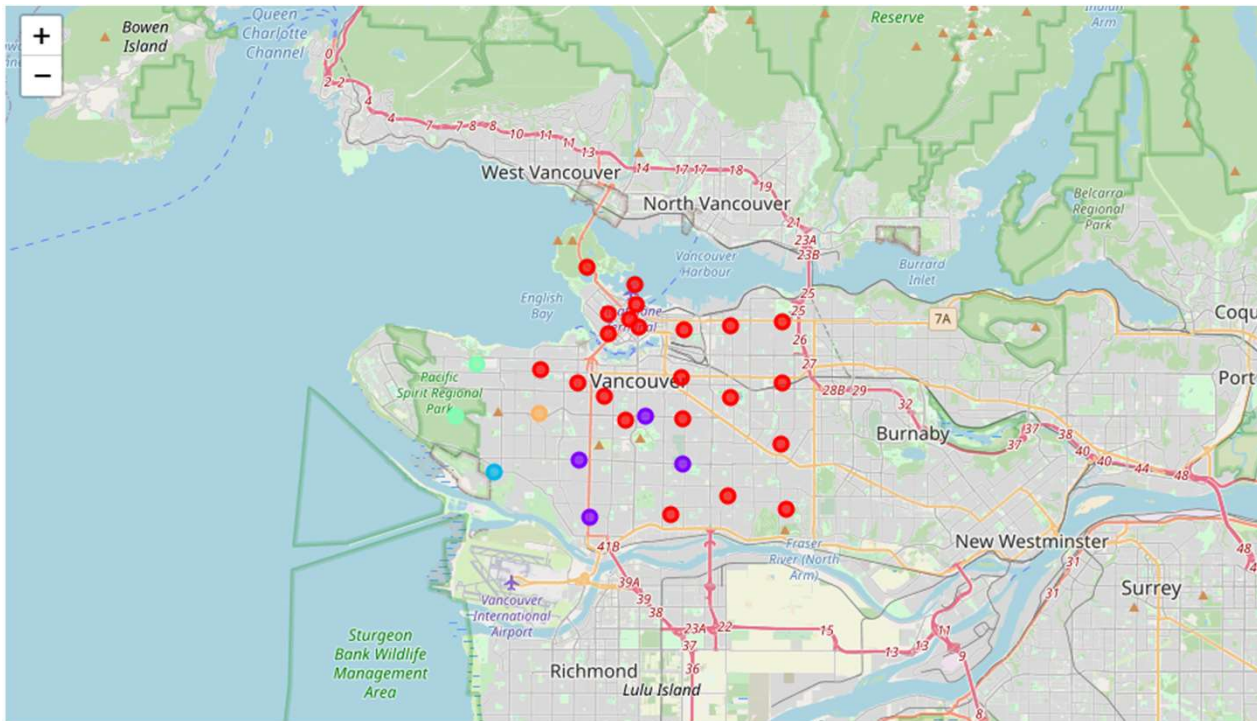
► Clustering

- Grouping results by neighborhood and by taking the mean of the frequency of the occurrence of each venue category
- Using the Top 10 most common venue categories in each neighborhood for clustering
- Using k-means model for neighborhood clustering ($k=5$)



Results

- 5 clusters were created from k-means model



Discussion

► Observations

- Cluster 0: includes most neighborhoods; lots of restaurants/coffee shops; sports opportunities quite common
- Cluster 1: 4 neighborhoods; restaurants most common venue; gyms not listed in Top 10 most common venues
- Cluster 2: 1 neighborhood; yoga studio (Top 3rd most common venue)
- Cluster 3: 2 neighborhoods; more sports facilities
- Cluster 4: 1 neighborhood; yoga studio (Top 4th most common venue)

► Recommendation

- Neighborhoods in Cluster 1 as potential location to open his gym

Conclusion

- ▶ Recommendation to my friend to look at neighborhoods in Cluster 1 in Vancouver as potential location to open his gym
 - ▶ Postal Codes: V6M, V6P, V5W and V5Y

