The Battle of Neighborhoods in Vancouver, BC, Canada

Zhanerke Shulenbayeva

Business problem

- For the past 10 years Canada has been the number one choice for immigrants all over the world. The stable economy, great educational system and the beautiful nature have been attracted people to apply for Canadian permanent residency. However, for anyone who decides to immigrate or just travel to Great White North, there is a lack of information about certain cities, specifically about their neighborhoods.

Questions to be addressed

Potential new immigrants could face a problems:

- Where to live in a certain town?
- Which neighborhoods have more parks nearby
- Is there any coffee shops and yoga studious nearby?

Data acquisition

The data required for this project are the information about neighborhoods in Vancouver and the data about venues in the city.

- The data about venues was taken via the Foursquare.com API, which returns the data about venues in the specified location. It is required to pass the user credential information, longitude and latitude of the interested area and the interested number of venues to the API. The data from the API is obtained in json file with the information about venue, address, category, price, users' comments and likes. For this project the data about venues categories was collected and analyzed.

- The required information about city's districts boundaries was taken from the website *City Of Vancouver Open Data Portal (www.opendata.vancouver.ca)*. The database *Local area boundary* contains the data about Postal codes, names of neighborhoods and its geo-coordinates.

Data transformation

Collected data was transformed into the table using pandas library

[2]:

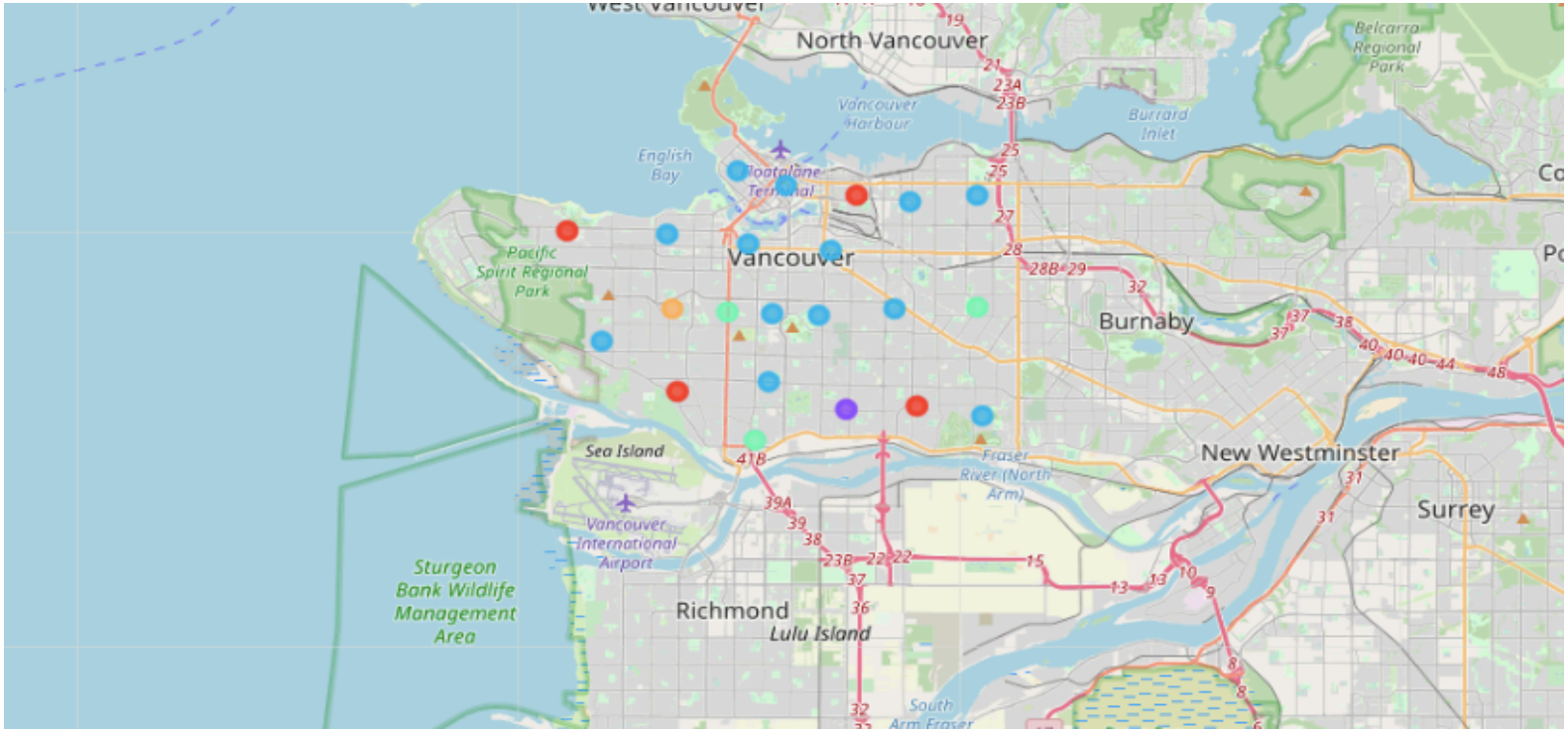
	PostalCode	Neighborhood	Latitude	Longitude
0	AR	Arbutus-Ridge	49.246805	-123.161669
1	CBD	Downtown	49.280747	-123.116567
2	FAIR	Fairview	49.264540	-123.131049
3	GW	Grandview-Woodland	49.276440	-123.066728
4	HS	Hastings-Sunrise	49.277934	-123.040270
5	MARP	Marpole	49.210207	-123.128382
6	RP	Riley Park	49.244766	-123.103147
7	SHAU	Shaughnessy	49.245681	-123.139760
8	STR	Strathcona	49.278220	-123.088235

Data Analysis / Methodology section

- **One Hot Encode -**methodology is used in machine learning to analyze the categorical data
- **K-means clustering** is one of the simplest and popular methodologies in machine learning to structure data. Clustering is the process of grouping similar data into subgroups (clusters) and separate different data into other subgroups. K-means clustering algorithm partitions data into non-overlapping datasets with the pre-defined K-number centroids. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum.

Data examination

The K-mean clustering algorithm with the number of centroids equal 5, partitioned the data about venues in each neighborhood of Vancouver into five clusters.



Discussion section / Recommendations

From the Results section, one could observe that the neighborhoods in Vancouver mostly similar to each other. Mostly half of the neighborhoods fell into the third cluster, while the rest of the neighborhoods were classified into four different clusters.

- **The cluster 1** includes neighborhoods with the recreational venues, such as parks and pools. Thus, these neighborhoods could by considered by the families who are looking for the active sport life or just want to have nice parks nearby.
- **The cluster 2** has the only one neighborhood with the Indian restaurant as the most favorable places. Thus, the Sunset neighborhood in Vancouver could be populated with Asian people. It could be helpful for people who want to live in a Asian populated area.
- **The cluster 3** is the biggest one and includes mostly 60% of all neighborhoods. Those neighborhoods have a lot of cafes, restaurants and coffee shops. So, this area could be considered by people looking to have places to go nearby. These districts could be good for students or young professionals.
- **The cluster 4** includes neighborhoods that don't have a lot of places to go, thus it could be just suburbs with houses. These places could be considered by families with kids or retired people.
- **The cluster 5** has only one neighborhood which is probably the business center

Results

	PostalCode	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	AR	Arbutus-Ridge	49.246805	-123.161669	4	Park	Business Service	Fish & Chips Shop	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
1	CBD	Downtown	49.280747	-123.116567	2	Hotel	Coffee Shop	Restaurant	Bar	Sandwich Place	Concert Hall	Vegetarian / Vegan Restaurant	Steakhouse
2	FAIR	Fairview	49.264540	-123.131049	2	Coffee Shop	Park	Asian Restaurant	Breakfast Spot	Japanese Restaurant	Furniture / Home Store	Camera Store	Spa
3	GW	Grandview-Woodland	49.276440	-123.066728	2	Coffee Shop	Pizza Place	Deli / Bodega	Vegetarian / Vegan Restaurant	Theater	Indian Restaurant	Cajun / Creole Restaurant	Café
4	HS	Hastings-Sunrise	49.277934	-123.040270	2	Bridal Shop	Theme Park Ride / Attraction	Café	Pizza Place	Beer Garden	Park	Portuguese Restaurant	Theater
5	MARP	Marpole	49.210207	-123.128382	3	Bus Stop	Pizza Place	Plaza	Bus Station	Japanese Restaurant	Taiwanese Restaurant	Yoga Studio	Electronics Store
6	RP	Riley Park	49.244766	-123.103147	2	Farmers Market	Vietnamese Restaurant	Japanese Restaurant	Coffee Shop	Sporting Goods Shop	Grocery Store	Lounge	Café
7	SHAU	Shaughnessy	49.245681	-123.139760	3	Bus Stop	Print Shop	Chocolate Shop	Park	Yoga Studio	Dive Bar	Falafel Restaurant	Event Space
8	STR	Strathcona	49.278220	-123.088235	0	Park	Food Truck	Arts & Crafts Store	Coffee Shop	Deli / Bodega	Pub	Cheese Shop	Soup Place
9	WE	West End	49.285011	-123.135438	2	Café	Farmers Market	Gay Bar	Sushi Restaurant	Sandwich Place	Falafel Restaurant	Restaurant	Spanish Restaurant
10	DS	Dunbar-Southlands	49.237962	-123.189547	2	Grocery Store	Bus Stop	Coffee Shop	Japanese Restaurant	Liquor Store	Yoga Studio	Electronics Store	Farmers Market
11	KERR	Kerrisdale	49.223655	-123.159576	0	Pool	Park	Café	Bar	Deli /	Dessert	Diner	Disc Golf

Support

Conclusion

To sum up, in this project, neighborhoods of Vancouver were analyzed based on the types of venues located at these areas and recommendations were made. The analysis included the data preparation with the pandas library, data analysis with the one hot encoding and K-means clustering methodologies. The recommendations were made based on the results from the clustering process. Hopefully, this analysis could help new or potential immigrants to find a suitable place to live or stay in Vancouver, BC, Canada.