Data Science Capstone

Clustering Montreal Neighbourhoods to Improve Real Estate Agent Efficiency

A Presentation by Vishnu Kumaraswami

For Coursera 24.06.2021

Contents

Overview

- In the current economic climate, with many individuals resources stretched, time truly is money
- For real estate agents with limited resources, knowing exactly where to take a client without running in circles around the city would boost efficiency tremendously

- With this project, we aim to:
  - Cluster the neighbourhoods in a city based on the facilities and amenities available in each location.
  - Cluster an existing client database to fit multiple profiles.
  - Assign client profiles to neighbourhood clusters based on assumptions regarding popular activities.
  - For the purposes of this project, we will use Montreal as an example.

Data Used

- A list of postal codes along with their geographic coordinates for the city of Montreal, Quebec, Canada.

- Foursquare API data to pinpoint venues and their categories within a given radius of a location

- Pre-existing database of customer profiles, including age, income and education

Methodology

- Data was successfully scraped from the list of postal codes available for the city of Montreal
- The coordinates for each postal code were mapped into the dataframe as well, to ensure the Foursquare API would function as intended
- Venues were mapped to each of the postal codes, and we analyzed which were the most common venues within a specific neighbourhood.

	Postal Code	Neighborhood
0	H1A	Pointe-aux-Trembles
1	H2A	Saint-Michel,East
2	H3A	Downtown Montreal North
3	H4A	Notre-Dame-de-GrâceNortheast
4	H5A	Place Bonaventure
5	H7A	Duvernay-Est
6	H9A	Dollard-des-OrmeauxNorthwest
7	H1B	Montreal East
8	H2B	AhuntsicNorth
9	H3B	Downtown MontrealEast
10	H4B	Notre-Dame-de-GrâceSouthwest

```
In [17]: df_data_1.head()
```

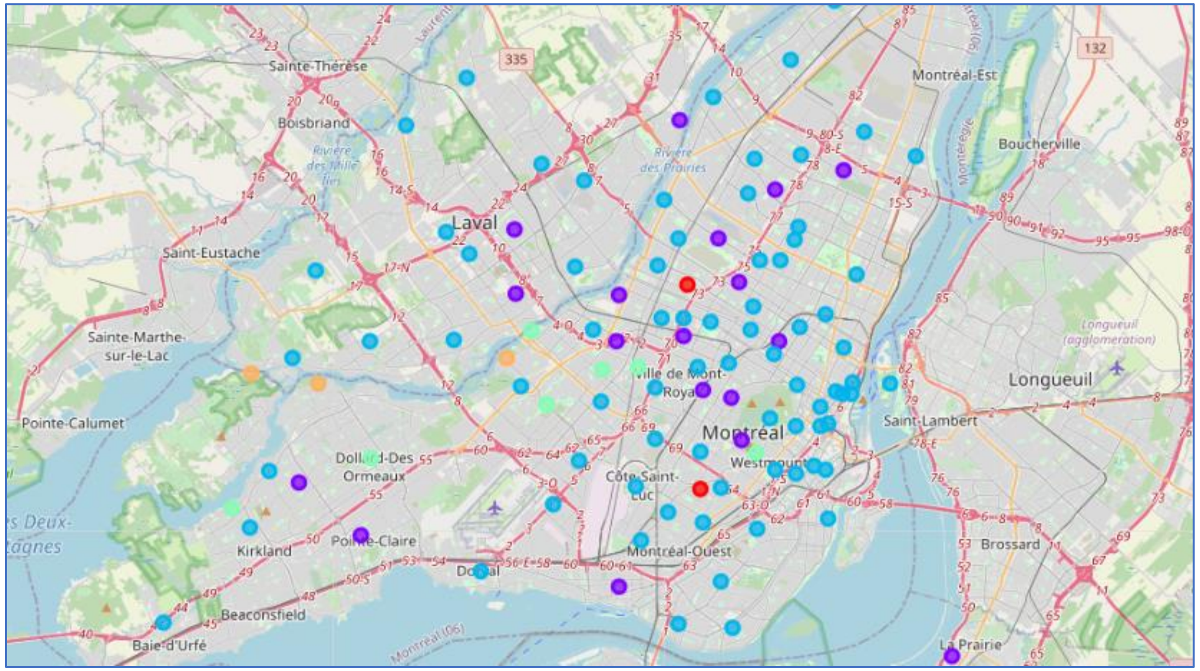
Out[17]:

	Unnamed: 0	Postal Code	Neighborhood	Latitude	Longitude
0	0	H1A	Pointe-aux-Trembles	45.454300	-73.483774
1	1	H2A	Saint-Michel East	45.561567	-73.601288
2	2	H3A	Downtown Montreal North	45.506992	-73.568941
3	3	H4A	Notre-Dame-de-Grâce Northeast	45.476496	-73.622168
4	4	H5A	Place Bonaventure	45.500795	-73.565230

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Ahuntsic Central	Breakfast Spot	Ice Cream Shop	Coffee Shop	Pharmacy	Restaurant
1	Ahuntsic East	Playground	Athletics & Sports	Yoga Studio	Discount Store	Farm
2	Ahuntsic North	Soccer Field	Yoga Studio	College Science Building	Farm	Factory
3	Ahuntsic Southeast	Italian Restaurant	Women's Store	Clothing Store	Sandwich Place	Bank
4	Ahuntsic Southwest	Home Service	Ice Cream Shop	Pizza Place	Bakery	Park

Methodology (contd)

- All the neighbourhoods in the city were mapped into 5 clusters, depicted in the adjoining map.
- Clustering was on the basis of commonality of venues adjacent to a neighbourhood



- The client database was also clustered, to give three distinct profiles, which were then assigned neighbourhood clusters based on the venues available

	Customer Id	Age	Edu	Years Employed	Income
Clus_km					
0	402.295082	41.333333	1.956284	15.256831	83.928962
1	432.468413	32.964561	1.614792	6.374422	31.164869
2	410.166667	45.388889	2.666667	19.555556	227.166667

Results and Observations

- **For Neighbourhoods**: Assumed profiles based on most common venues

Cluster	Most Common Venues	Assumed Client Profile
0	Playground, Stores, Yoga Studios	Young families with children below 5
1	Parks, Gym/Yoga, Convenience Stores	Middle Aged, Upper middle income, families
2	Pubs, Restaurants, Entertainment	Young adults, disposable income
3	Park, Yoga Studio, Farm	Older, High Income, Possibly retired
4	Construction, Yoga, Fast food	Non-residential area

- **For Clients:** From the clusters above, it would be possible to classify the customers as follows:
    - **Type A (Cluster 0)**: Middle Aged, Upper Middle Income
    - **Type B (Cluster 1)**: Young, Upwardly Mobile
    - **Type C (Cluster 2)**: Older, High Income

- Accordingly, we allocated Client Types to Clusters:

Client Profile	Neighbourhood Cluster(s)
Type A	1,0
Type B	2,1
Type C	3

Conclusion

- The clustering of neighbourhoods in Montreal can act as a guide for real estate agents when assessing potential clients, helping save time and improve efficiency.

- A prospective client’s details would need to be updated in the database and then the program would need to be run. Accordingly, they would be assigned a cluster and provided with a suitable set of real estate options (either to sell or to rent depending on their requirement).

THANK YOU