CAPSTONE PROJECT: THE BATTLE OF NEIGHBORHOODS

PRESENTATION

ZHONGYU ZHAO

1.Introduction

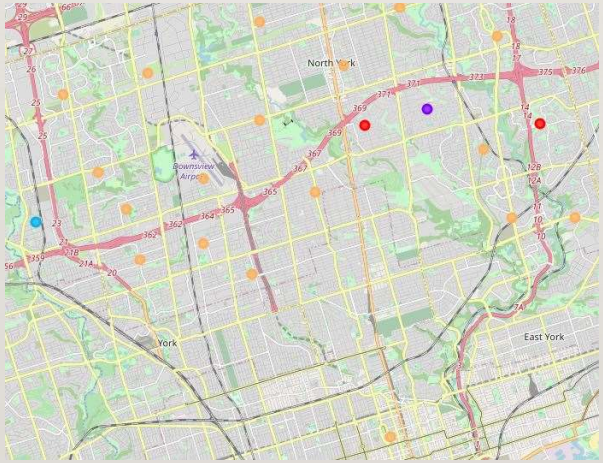
- **Background:** Lots of people are migrating to Canada. There's no doubt that they need to do a lot of research for good housing prices and living enviroments. This project is for those people who are looking for better neighborhoods. Precisely speaking, this project focus on those who wants to move to North York and willing to find a neighborhood that meets his demand.
- **Goal:** In this project we will try to find an optimal location in North York for Torontonians to move in. Specifically, this report will be targeted to find a safe and convenient neighborhood that is similar to the previous one. Since there are lots of neighborhoods in North York we will try to do clustering to filter out some of them. We will also use crime rate data and traffic data to secure the safety and convenience of the neighborhood. We would prefer neighborhood as safe as possible. We also want the neighborhood could offer venues that users love and not too far from the public transit.

2.Data Description

- Three datasets are used in this project.
- **1.Foursquare API Data:** we connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.
- **2.Toronto Crime Data:**
  In order to select the best neighborhood, We will use crime rate data to find out if a neighborhood is safe enough. We will focus on differ ent kinds of crime rating for each neighborhood and rank them.
- **3.Toronto Subway Station Data:** It's scraped from wikipedia, showing all the subway stations' name in Toronto.

3.Methodology

- The problem we are going to solve is to find a new neighborhood that is similar to the previous one for a person who wants to move to North York.
- We do clustering for all neighborhoods in order to find the feasible one.



3.Methodology

- After clustering we do key word searching to meet customer's demand.

```
[ ] NY_chosen = NY_merged_plus.iloc[[2,3,8,9,10,17,19]]
    NY_chosen.reset_index(drop = True,inplace = True)
    NY_chosen
```

	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	4.0	Clothing Store	Furniture / Home Store	Boutique	Event Space	Miscellaneous Shop
1	M3B	North York	Don Mills	43.745906	-79.352188	4.0	Gym	Beer Store	Coffee Shop	Japanese Restaurant	Restaurant
2	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	4.0	Clothing Store	Coffee Shop	Fast Food Restaurant	Japanese Restaurant	Cosmetics Shop
3	M3J	North York	Northwood Park, York University	43.767980	-79.487262	4.0	Coffee Shop	Bar	Vietnamese Restaurant	Massage Studio	Metro Station
4	M2K	North York	Bayview Village	43.786947	-79.385975	4.0	Chinese Restaurant	Café	Bank	Japanese Restaurant	Yoga Studio
5	M5M	North York	Bedford Park, Lawrence Manor East	43.733283	-79.419750	4.0	Sandwich Place	Coffee Shop	Italian Restaurant	Liquor Store	Pub
6	M2N	North York	Willowdale, Willowdale East	43.770120	-79.408493	4.0	Ramen Restaurant	Sandwich Place	Coffee Shop	Café	Pizza Place

3.Methodology

- We do crime rate analysis to find out which neighborhood is safe and which is not.

	Neighbourhood	Assault_Rate_2019	AutoTheft_Rate_2019	BreakandEnter_Rate_2019	Homicide_Rate_2019	Robbery_Rate_2019	TheftOver_Rate_2019	crime_rate_avg
1	York University Heights	1340.9	521.9	391.4	0.0	286.3	101.5	4.157715
3	Yorkdale-Glen Park	1411.8	412.1	567.4	6.8	283.7	195.9	5.675270
9	Danforth East York	483.1	52.4	139.7	0.0	34.9	40.7	1.018460
45	Bayview Woods-Steeles	342.1	136.8	152.0	7.6	38.0	15.2	1.982641
49	Old East York	498.2	43.3	140.8	0.0	86.6	32.5	1.152379
81	Englemount-Lawrence	518.5	102.8	147.5	0.0	71.5	31.3	1.233705
84	Banbury-Don Mills	267.2	151.7	292.5	0.0	36.1	50.6	1.379827
95	Willowdale East	333.1	113.0	174.5	0.0	79.3	47.6	1.279963
96	Willowdale West	661.3	183.0	124.0	0.0	165.3	59.0	1.916132
107	Parkwoods-Donalda	454.0	91.9	195.4	2.9	103.4	14.4	1.658286
113	Bayview Village	308.5	172.9	172.9	0.0	65.4	60.8	1.393322
116	Bridle Path-Sunnybrook-York Mills	474.9	118.7	550.4	0.0	140.3	32.4	2.189292
117	Don Valley Village	547.1	77.6	110.9	0.0	125.7	51.8	1.435130
119	Lawrence Park South	197.6	184.5	382.1	0.0	85.6	52.7	1.734904
127	Lawrence Park North	260.1	123.2	150.6	0.0	34.2	27.4	0.945037

3.Methodology

- We calculate the distance between neighborhoods and subway stations in order to choose a neighborhood that can reach public transit easily.

	Neighborhood	Latitude	Longitude	Closest	Distance
0	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	Yorkdale	0.962
1	Don Mills	43.745906	-79.352188	Leslie	1.863
2	Fairview, Henry Farm, Oriole	43.778517	-79.346556	Don Mills	0.221
3	Northwood Park, York University	43.767980	-79.487262	Finch West	0.305
4	Bayview Village	43.786947	-79.385975	Bessarion	1.317
5	Bedford Park, Lawrence Manor East	43.733283	-79.419750	Lawrence	0.989
6	Willowdale, Willowdale East	43.770120	-79.408493	North York Centre	0.230

4.Result and Conclusion

According to the analysis, York University neighborhood is not safe so we will not consider this place. For the rest candidates, Fairview neighborhood is close to subway station and it's safe according to the record. Willwodale is the second choice, it is also close to the subway station, but the west part of Willowdale is not that safe. Don Mills, Bayview Village and Lawrence are too fart from subway station so they are not that good. In general, Fairview neighborhood is the best choice.

This project helps a person get a better understanding of the neighborhoods with respect to venues, crime rates and other data in that neighborhood. This project shows how data analysis helps people live better. Further discussion may include taking other factors such as quality of education, supermarkets, highways, cost of living, house price and so on.