Presentation for Applied Data Science Capstone

By Qasim Nazir

The Battle of Neighbourhoods

Problem Statement

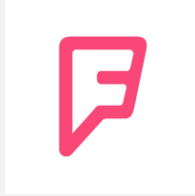Use foursquare data to get information about venues in a neighborhood

Compare and segment neighbourhoods of Toronto

Give useful insights to targeted audience and help them in decision making

Data Sources

Forsquare Data

Toronto Open Data

Wikipedia

	Neighbourhood	Population	Land area (km2)	Density (people/km2)	Average Income
0	Agincourt	44577	12.45	3580	25750
1	Agincourt North, L'Amoreaux East, Milliken, St...	96830	18.86	5177	26092
2	Albion Gardens, Beaumond Heights, Humbergate, ...	16790	3.97	4229	28955
3	Alderwood, Long Branch	21281	7.16	3348	36263
4	Bathurst Manor, Downsview North, Wilson Heights	65290	24.96	2924	32966

	Neighbourhood	Venue Count	Top 3 Venues	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	9th Most Common Venue	10th Most Common Venue	Access
0	Adelaide, King, Richmond	100	Coffee Shop, Café, Bar	Coffee Shop	Café	Bar	Steakhouse	Thai Restaurant	Hotel	Sushi Restaurant	Bakery	American Restaurant	Asian Restaurant	
1	Agincourt	4	Sandwich Place, Lounge, Breakfast Spot	Sandwich Place	Lounge	Breakfast Spot	Clothing Store	Drugstore	Discount Store	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	
2	Agincourt North, L'Amoreaux East, Milliken, St...	2	Playground, Park, Drugstore	Playground	Park	Drugstore	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Dumpling Restaurant	
3	Albion Gardens, Beaumond Heights, Humbergate, ...	11	Grocery Store, Pharmacy, Coffee Shop	Grocery Store	Pharmacy	Coffee Shop	Sandwich Place	Discount Store	Fried Chicken Joint	Beer Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	
4	Alderwood, Long Branch	9	Pizza Place, Pharmacy, Gym	Pizza Place	Pharmacy	Gym	Sandwich Place	Coffee Shop	Athletics & Sports	Skating Rink	Pub	Dog Run	Dim Sum Restaurant	

Results

- Discovered 3 clusters based on Venue Count

	Cluster Label	Neighbourhood Count	Venue Count Mean	Description
0	0	14	50.571429	Mid Venue Count
1	1	9	97.555556	High Venue Count
2	2	77	8.389610	Low Venue Count

- Discovered 5 clusters based on demographics

	Cluster Label	Density (people/km2)	Average Income	Description
0	0	5566.310345	48996.275862	Mid Dense Mid Income
1	1	5717.666667	98826.666667	Mid Dense High Income
2	2	2324.500000	214025.500000	Low Dense Very High Income
3	3	5234.571429	30101.542857	Mid Dense Low Income
4	4	30106.000000	37011.000000	High Dense Low Income

- Scatter plots
- Map Visualizations

Discussion & Conclusion

- Made sense from raw data
- Discovered clusters in neighbourhoods
- Useful insights for entrepreneurs
- Useful insights for marketing campaigns
- Useful insights for comparing residence options
- Correlations among different parameters such as distance from city center, population density and venue count

Questions