Exploring Crime Occurrences and Neighbourhood Venues in Toronto

Robert McMillan

Crime Occurrences and Neighbourhood Venues in Toronto

Can we predict the safety of a neighbourhood by the type of venues it has?

- If we could, it would be useful for:
  - People moving to a new city
  - Tourists going to a city for the first time
  - City planners, to create safer neighbourhoods

Crime Occurrences and Neighbourhood Venues in Toronto

Data Acquisition

- Crime data from 2014 - 2018 in a csv file downloaded from the Toronto Police Service Public Safety Data Portal site (containing information licensed under the Open Government Licence – Ontario), with occurrences of:
  - assault
  - auto theft
  - break and enter
  - robbery
  - theft over $5,000
  - homicide by neighbourhood for the years 2014 to 2018.

- Venue types in each of the Toronto neighbourhoods obtained by API calls to Foursquare API

Crime Occurrences and Neighbourhood Venues in Toronto

Crime Data Cleaning and Processing

- serious crime occurrences extracted for 2018
- row with missing value dropped
- number entries set to Python type int

- serious crime occurrences summed into one column
- data sorted in descending order of crime occurrences

- neighbourhood coordinates added with geopy geocoder
- latitude, longitude and crime ranking added to columns

- resulting data size:
  - 105 rows (neighbourhoods) and 10 columns

Crime Occurrences and Neighbourhood Venues in Toronto

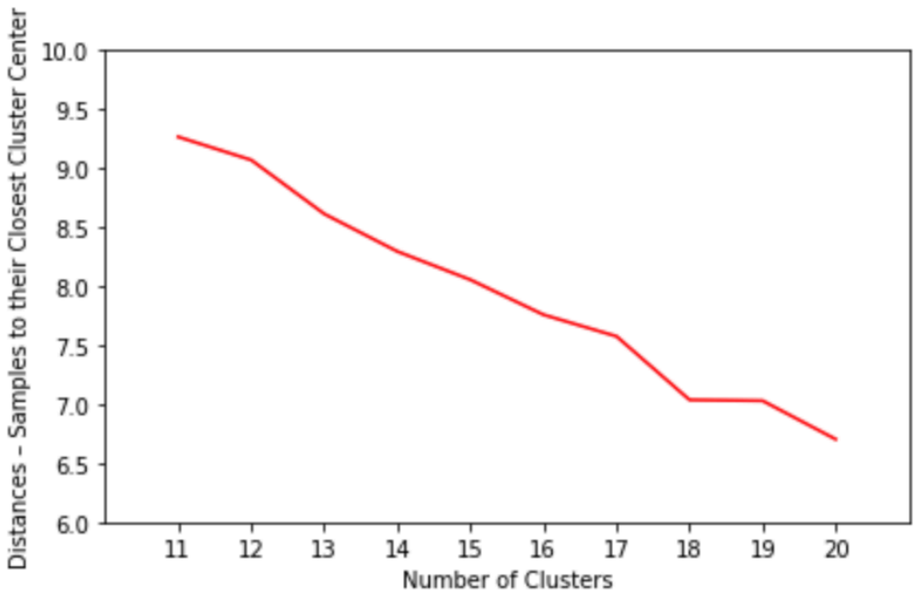
Venue Data Cleaning and Processing

- maximum 30 venues added to each neighbourhood
- venues one-hot encoded by neighbourhood
- mean frequency of occurrence of each venue category calculated
- resulting data size:
  - 105 rows (neighbourhoods) x 236 columns (venues)

Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis I

- k-means clustering performed for cluster (k) sizes 11 - 20
- elbow point found at k=18
- k-means run for k=18
- the 18 cluster labels added to the crime dataframe for visual analysis



Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis II

- 63.8% of neighbourhood fall into two cluster groups (1 and 11)

Cluster Labels	Neighbourhood
0	1
1	25
2	1
3	1
4	1
5	1
6	1
7	7
8	6
9	5
10	2
11	42
12	3
13	1
14	4
15	1
16	1
17	2

Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis III

- neighbourhoods from cluster groups 1 and 11 dominate the 10 neighbourhoods with the highest crime occurrences

Crime_Rank	Cluster Labels
1	1
2	8
3	11
4	12
5	14
6	11
7	1
8	5
9	1
10	4

Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis IV

- neighbourhoods from cluster groups 1 and 11 also dominate the 10 neighbourhoods with the lowest crime occurrences

Crime_Rank	Cluster Labels
96	1
97	1
98	11
99	1
100	1
101	1
102	11
103	11
104	1
105	1

Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis V

- clusters 8, 12, 14, 5 and 4 appear in the highest 10 crime neighbourhoods
- clusters 4 and 5 have only one neighbourhood each and can be ruled out

hbourhood	Assault_2018	Auto_Theft_2018	BreakandEnter_2018	Robbery_2018	Theft_Over_2018	Homicide_2018	Crime_Occurrence	Latitude	Longitude	Crime_Rank	Cluster Labels
Black Creek	1005	79	221	224	46	4	1579	43.734634	-79.505355	1	1
Cliffcrest	787	58	314	93	50	1	1303	43.721939	-79.236232	2	8
Ionview	284	495	154	69	50	0	1052	43.735990	-79.276515	3	11
Palmerston-Little Italy	547	40	145	159	37	0	928	43.655854	-79.410116	4	12
he Beaches	457	22	236	78	30	0	823	43.671024	-79.296712	5	14
hurst Manor	404	109	83	64	14	0	674	43.763893	-79.456367	6	11
Centennial carborough	385	60	98	72	13	0	628	43.787491	-79.150768	7	1
dale-Kipling	295	34	189	54	53	0	625	43.721362	-79.565513	8	5
carborough Village	340	77	96	85	22	0	620	43.743742	-79.211632	9	1
nnedy Park	411	26	103	51	5	2	598	43.724878	-79.253969	10	4

Crime Occurrences and Neighbourhood Venues in Toronto

Data Analysis VI

- clusters 8, 12 and 14 are grouped by the similarity of their venues, but do not show similar crime rankings

Crime_Rank	Cluster Labels
2	8
39	8
41	8
44	8
56	8
91	8

Crime_Rank	Cluster Labels
4	12
36	12
77	12

Crime_Rank	Cluster Labels
5	14
32	14
33	14
82	14

Crime Occurrences and Neighbourhood Venues in Toronto

Conclusion

- This study explored the relationship between crime and venue types in neighbourhoods in Toronto
- If a correlation exists, it would be useful to people moving or travelling to a new city, as well as city planners
- However, no correlation was found between crime occurrences and neighbourhood venues

Crime Occurrences and Neighbourhood Venues in Toronto

Caveats

- Geopy's geocoder used for coordinates may not be so accurate
- Some neighbourhoods do not have enough venues to make accurate comparisons

Crime Occurrences and Neighbourhood Venues in Toronto

Future Direction

- Explore crime occurrences and neighbourhood venues with more accurate latitude and longitude coordinates
- Explore an intra-city comparison of neighbourhoods with more venues to better cluster the neighbourhoods