

IBM Data Science Capstone – Setting Up a Coffee Shop in Toronto



A Business owner is looking to set up a coffee shop in Toronto:

He is looking for a venue that is relatively affluent and populous...



... Which does not have a lot of existing competition (Eg. Other cafes).



From this, I understood that I needed to:

1) Find out population statistics for each of Toronto's districts



2) Get geodata for all of Toronto's districts and coffee shops



3) Combine these 2 data points to help the business owner decide



Some of the data sources I used include:



City of Toronto's Open
Data sets



Wikipedia on Burroughs
and Neighborhoods in
Toronto

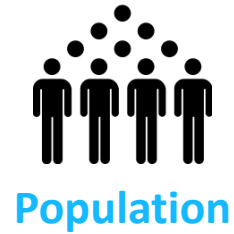


Foursquare Data on
Businesses in Toronto

These are the following that went into the tables:



From:



Geodata

From:



I scored each district based on Census Data, and combined the scores to rank each district:

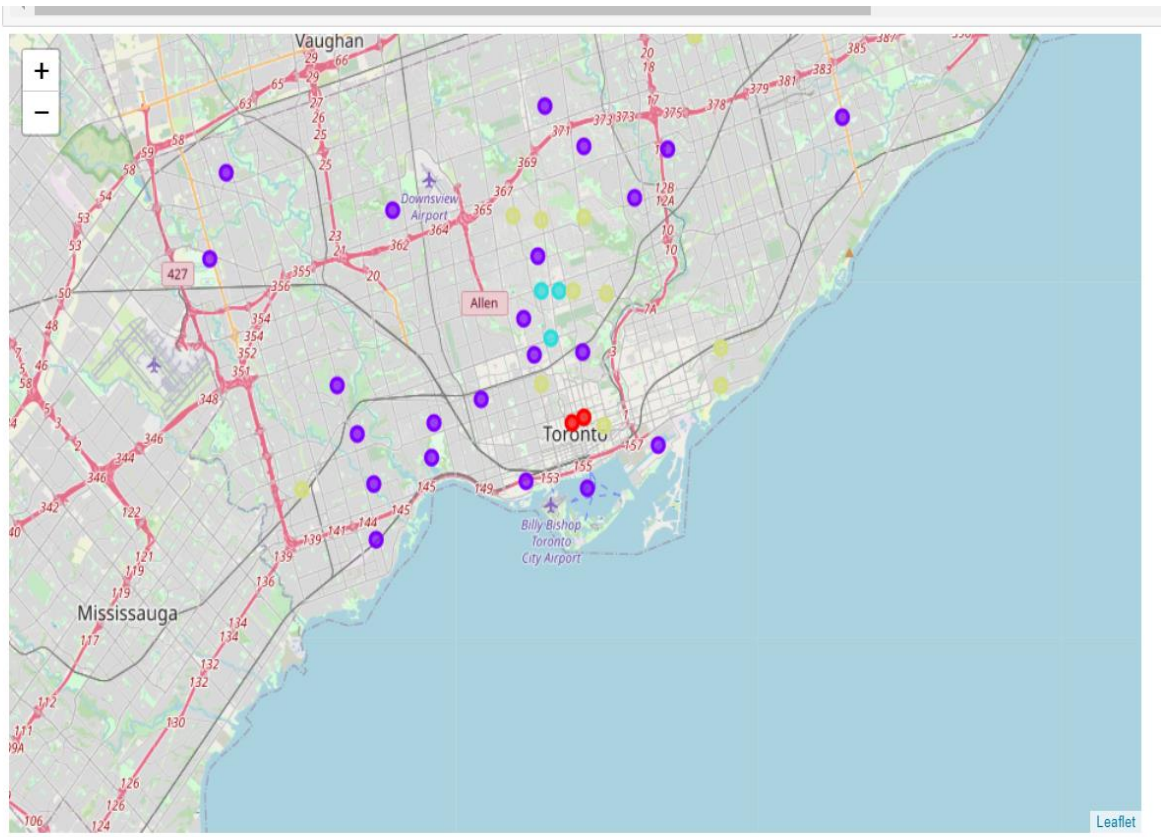
$$\text{Total Score} = \frac{1}{3} \text{ Population Score*} + \frac{1}{3} \% \text{ Working Population Score*} + \frac{1}{3} \text{ Income Score*}$$


The diagram illustrates the calculation of a Total Score for each district. It is presented as an equation: Total Score equals the sum of three components, each weighted by 1/3. The components are: Population Score* (represented by an icon of four people with dots above them), % Working Population Score* (represented by an icon of three people inside a circle with a percentage sign), and Income Score* (represented by an icon of a hand holding three dollar signs). The weights 1/3 are shown in blue, and the plus signs and the 'Total Score' label are also in blue.

**Where Score = $\frac{x}{x_{median}}$*

This was used with the number of coffee shops in each district to do K-means clustering – 4 Clusters

Out[262]:



Clusters	L	neighborhood	CDN	borough	LONGITUDE	LATITUDE	pop_score	Income_score	work_score	total_score	No. of Coffee Shops
39	1	Waterfront Communities-The Island	77	Old City of Toronto	-79.377202	43.633880	1.312427	0.497174	0.533259	2.34	0.0
28	1	Rosedale-Moore Park	98	Old City of Toronto	-79.379669	43.682820	0.418608	1.235014	0.295162	1.95	0.0
30	1	Edenbridge-Humber Valley	9	Ettobicoke	-79.522458	43.670886	0.309325	0.680762	0.297961	1.27	0.0
31	1	Forest Hill South	101	Old City of Toronto	-79.414318	43.694526	0.213690	1.306094	0.307509	1.83	0.0
32	1	Woburn	137	Scarborough	-79.228586	43.766740	1.064967	0.250373	0.319734	1.64	0.0
29	1	Stonegate-Queensway	16	Ettobicoke	-79.501128	43.635518	0.498803	0.448731	0.331602	1.28	0.0
34	1	StAndrew-Windfields	40	North York	-79.379037	43.756248	0.354664	0.649123	0.297496	1.30	0.0
35	1	Downsview-Roding-CFB	26	North York	-79.490497	43.733292	0.897938	0.272653	0.334475	1.31	0.0
36	1	Parkwoods-Donalda	45	North York	-79.330180	43.755033	0.893020	0.324255	0.330916	1.35	0.0
37	1	Kingsway South	15	Ettobicoke	-79.510577	43.635520	0.184600	0.895924	0.272755	1.35	0.0
33	1	Willowdale East	51	North York	-79.401484	43.770602	1.004217	0.336443	0.399413	1.74	0.0
38	1	South Riverdale	70	Old City of Toronto	-79.335651	43.649292	0.555053	0.396488	0.420019	1.37	0.0
27	1	High Park North	88	Old City of Toronto	-79.466302	43.657565	0.441279	0.420281	0.409991	1.27	1.0
26	1	Mount Olive-Silverstone-Jamestown	2	Ettobicoke	-79.587259	43.748688	0.856164	0.220895	0.321954	1.20	1.0
25	1	High Park-Swansea	87	Old City of Toronto	-79.467872	43.645065	0.476383	0.502485	0.376848	1.36	1.0
24	1	Lawrence Park South	103	Old City of Toronto	-79.406039	43.717212	0.302237	1.021839	0.301356	1.63	1.0
23	1	Mimico (includes Humber Bay Shores)	17	Ettobicoke	-79.500137	43.615924	0.678274	0.402660	0.405991	1.48	1.0
22	1	West Humber-Clairville	1	Ettobicoke	-79.596356	43.716180	0.663292	0.257012	0.323875	1.24	1.0
21	1	Dovercourt-Wallace Emerson-Junction	93	Old City of Toronto	-79.438541	43.665677	0.729259	0.309310	0.421069	1.46	1.0
20	1	Banbury-Don Mills	42	North York	-79.349718	43.737657	0.551449	0.475031	0.304165	1.33	2.0
19	1	Niagara	82	Old City of Toronto	-79.412420	43.636681	0.620841	0.501853	0.582824	1.71	2.0
17	1	Malvern	132	Scarborough	-79.222517	43.803658	0.872004	0.242717	0.317887	1.43	2.0
16	1	Rouge	131	Scarborough	-79.186343	43.821201	0.925805	0.308934	0.310224	1.54	2.0
18	1	Casa Loma	96	Old City of Toronto	-79.408007	43.681852	0.218389	1.053404	0.323628	1.60	2.0
15	3	Islington City Centre West	14	Ettobicoke	-79.543317	43.633483	0.875409	0.389648	0.365837	1.63	3.0
14	3	East End-Donforth	62	Old City of Toronto	-79.299359	43.684174	0.425728	0.389861	0.364101	1.19	3.0
13	3	L'Amoreaux	117	Scarborough	-79.314084	43.795716	0.875967	0.255895	0.304847	1.44	3.0
12	3	Leaside-Bennington	56	East York	-79.368072	43.703797	0.335071	0.782922	0.298916	1.42	3.0
11	3	Bridle Path-Sunnybrook-York Mills	41	North York	-79.378904	43.731013	0.184500	1.771537	0.259867	2.22	3.0
8	3	Moss Park	73	Old City of Toronto	-79.367297	43.665518	0.406305	0.420114	0.476162	1.30	4.0
10	3	Bedford Park-Norhtown	39	North York	-79.420227	43.731486	0.462664	0.784588	0.282046	1.53	4.0
9	3	Lawrence Park North	105	Old City of Toronto	-79.403978	43.730060	0.290847	0.724351	0.319025	1.33	4.0
7	3	Annex	95	Old City of Toronto	-79.404001	43.671585	0.607819	0.733856	0.383939	1.73	6.0
6	3	Mount Pleasant East	99	Old City of Toronto	-79.384924	43.704852	0.334015	0.584810	0.353979	1.27	6.0
5	3	The Beaches	63	Old City of Toronto	-79.299601	43.671050	0.429431	0.618033	0.346508	1.39	8.0
4	2	Yonge-Eglinton	100	Old City of Toronto	-79.403590	43.704689	0.235294	0.598492	0.386434	1.22	11.0
3	2	Yonge-St.Clair	97	Old City of Toronto	-79.397871	43.687859	0.249451	0.737675	0.370723	1.36	15.0
2	2	Mount Pleasant West	104	Old City of Toronto	-79.393360	43.704435	0.590535	0.416122	0.449303	1.46	16.0
1	0	Church-Yonge Corridor	75	Old City of Toronto	-79.379017	43.659649	0.624027	0.394126	0.466962	1.49	30.0
0	0	Bay Street Corridor	76	Old City of Toronto	-79.385721	43.657511	0.513657	0.397679	0.394662	1.31	30.0

From this, it was clear that Waterfront-Communities, the Island was the best venue!

Relatively large population & in the workforce → Likely to come to cafes



No coffee shops set up as of today! (16 July 2020)





Thank You!

Eric Tay Jing Long

