

BATTLE OF THE NEIGHBORHOODS

DATA SCIENCE CAPSTONE PROJECT



INTRODUCTION

- Relocating firm helping people relocate from Toronto, Ontario, Canada to Manhattan, New York, USA.
- Help customers ease into their new environment
 - Find neighborhoods that match well with their existing neighborhood in Toronto

“Which neighborhoods in Manhattan most closely matches with a given neighborhood in Toronto, in terms of venues and facilities?”



DATA SOURCES

- Manhattan neighborhood information
 - Source https://geo.nyu.edu/catalog/nyu_2451_34572
 - Cached https://cocl.us/new_york_dataset
- Toronto neighborhood information
 - Web scraping https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
 - Cached coordinates https://cocl.us/Geospatial_data
- Four Square API
 - Explore and Search (venues, categories)

EXPLORING TORONTO NEIGHBORHOODS

	Toronto Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	1st Venue Weight	2nd Venue Weight	3rd Venue Weight	4th Venue Weight	5th Venue Weight	Total Wt
29	St. James Town	Café	Coffee Shop	Cocktail Bar	Restaurant	American Restaurant	0.263158	0.263158	0.157895	0.157895	0.157895	0.237500
0	Berczy Park	Coffee Shop	Cocktail Bar	Seafood Restaurant	Bakery	Restaurant	0.357143	0.214286	0.142857	0.142857	0.142857	0.241379
31	Stn A PO Boxes	Coffee Shop	Café	Seafood Restaurant	Cocktail Bar	Restaurant	0.416667	0.166667	0.166667	0.125	0.125	0.247423
13	Garden District, Ryerson	Clothing Store	Coffee Shop	Bubble Tea Shop	Middle Eastern Restaurant	Café	0.36	0.28	0.12	0.12	0.12	0.250000
30	St. James Town, Cabbagetown	Coffee Shop	Pizza Place	Pub	Italian Restaurant	Bakery	0.25	0.25	0.166667	0.166667	0.166667	0.272727
6	Church and Wellesley	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Restaurant	Gay Bar	0.285714	0.238095	0.190476	0.142857	0.142857	0.272727
25	Richmond, Adelaide, King	Coffee Shop	Restaurant	Café								
17	Kensington Market, Chinatown, Grange Park	Café	Bakery	Mexican Restaurant								
14	Harbourfront East, Union Station, Toronto Islands	Coffee Shop	Aquarium	Hotel								
2	Business reply mail Processing Centre, South C...	Yoga Studio	Skate Park	Auto Workshop								

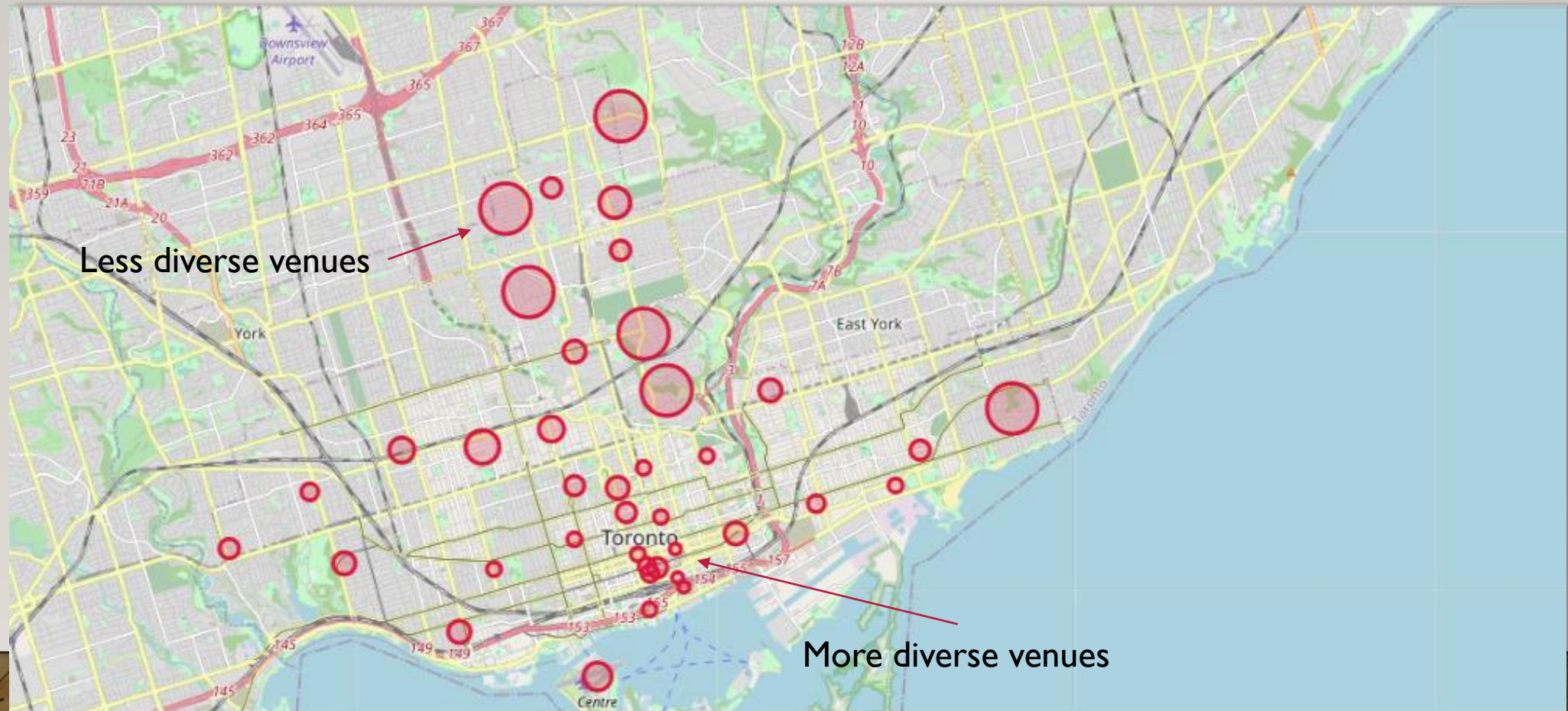
Most diverse venues

Least diverse venues

	Toronto Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	1st Venue Weight	2nd Venue Weight	3rd Venue Weight	4th Venue Weight	5th Venue Weight	Total Wt
10	Dufferin, Dovercourt Village	Bakery	Pharmacy	Bank	Supermarket	Bar	0.285714	0.285714	0.142857	0.142857	0.142857	0.500000
3	CN Tower, King and Spadina, Railway Lands, Har...	Airport Service	Airport Lounge	Airport Terminal	Boutique	Coffee Shop	0.333333	0.222222	0.222222	0.111111	0.111111	0.529412
9	Davisville North	Gym	Food & Drink Shop	Sandwich Place	Hotel	Dog Run	0.2	0.2	0.2	0.2	0.2	0.625000
5	Christie	Grocery Store	Café	Park	Baby Store	Nightclub	0.363636	0.272727	0.181818	0.090909	0.090909	0.687500
26	Rosedale	Park	Playground	Trail	None	None	0.5	0.25	0.25	0	0	1.000000
27	Roselawn	Music Venue	Garden	None	None	None	0.5	0.5	0	0	0	1.000000
12	Forest Hill North & West, Forest Hill Road Park	Park	Trail	Jewelry Store	Sushi Restaurant	None	0.25	0.25	0.25	0.25	0	1.000000
35	The Beaches	Trail	Neighborhood	Health Food Store	Pub	None	0.25	0.25	0.25	0.25	0	1.000000
20	Moore Park, Summerhill East	Gym	Trail	None	None	None	0.5	0.5	0	0	0	1.000000
18	Lawrence Park	Park	Swim School	Bus Line	None	None	0.333333	0.333333	0.333333	0	0	1.000000

Normalised weightage
(each row sum to 1)

TORONTO NEIGHBORHOODS



SEARCHING MANHATTAN NEIGHBORHOODS

	Manhattan Neighborhood	Café	Coffee Shop	Clothing Store	Yoga Studio	Bar	Thai Restaurant	Sandwich Place	Park	Greek Restaurant	...	Baby Store	Gay Bar	Scenic Lookout	Burrito Place	Men's Store	Salad Place	Pet Store	Dog Run	Indian Restaurant	Nightclub
0	Marble Hill	3	6	16	1	3	1	7	3	0	...	3	0	2	0	0	0	3	3	1	0
1	Chinatown	49	49	48	4	50	16	18	18	5	...	1	1	7	2	21	0	15	3	6	28
2	Washington Heights	9	11	37	1	13	1	5	5	0	...	2	0	5	0	6	2	8	1	2	2
3	Inwood	8	7	17	2	10	2	4	7	0	...	0	0	3	0	3	2	5	1	0	6
4	Hamilton Heights	9	14	14	5	12	4	10	8	1	...	0	2	2	0	1	0	2	1	2	1
5	Manhattanville	8	8	7	3	7	2	4	7	0	...	0	0	2	0	0	1	0	1	2	4
6	Central Harlem	4	6	9	0	13	0	5	2	0	...	0	1	0	0	2	0	0	3	1	0
7	East Harlem	10	8	19	3	8	5	10	5	1	...	0	1	6	0	2	0	6	3	3	0
8	Upper East Side	27	34	48	11	17	6	6	9	0	...	3	0	10	3	5	5	6	4	5	3
9	Yorkville	15	21	15	5	37	11	11	8	0	...	1	1	12	0	0	0	17	3	7	7

MANHATTAN NEIGHBORHOODS – NORMALIZED SCORING

	Manhattan Neighborhood	Café	Coffee Shop	Clothing Store	Yoga Studio	Bar	Thai Restaurant	Sandwich Place	Park	Greek Restaurant	...	Baby Store	Gay Bar	Scenic Lookout	Burrito Place	Men's Store	Salad Place	Pet Store	Dog Run	Indian Restaurant	Nightclub
0	Marble Hill	0.06	0.063830	0.32	0.020833	0.06	0.037037	0.104167	0.030303	0.000000	...	0.6	0.000000	0.06250	0.000000	0.00	0.000000	0.176471	0.3	0.029412	0.000000
1	Chinatown	0.98	0.978723	0.96	0.083333	1.00	0.592593	0.333333	0.484848	0.555556	...	0.2	0.052632	0.21875	0.222222	0.42	0.000000	0.882353	0.3	0.176471	0.595745
2	Washington Heights	0.18	0.170213	0.74	0.020833	0.26	0.037037	0.062500	0.090909	0.000000	...	0.4	0.000000	0.15625	0.000000	0.12	0.052632	0.470588	0.1	0.058824	0.042553
3	Inwood	0.16	0.085106	0.34	0.041667	0.20	0.074074	0.041667	0.151515	0.000000	...	0.0	0.000000	0.09375	0.000000	0.06	0.052632	0.294118	0.1	0.000000	0.127660
4	Hamilton Heights	0.18	0.234043	0.28	0.104167	0.24	0.148148	0.166667	0.181818	0.111111	...	0.0	0.105263	0.06250	0.000000	0.02	0.000000	0.117647	0.1	0.058824	0.021277

$$\frac{N - N_{min}}{N_{max} - N_{min}}$$

← Watch out for divide-by-zero!

RECOMMENDATION SYSTEM

Matching Score

$$S_{m,t} = \sum_{i=1}^5 \tilde{N}_{m,i} \times W_{t,i}$$

	St. James Town	Berczy Park	Stn A PO Boxes	Garden District, Ryerson	St. James Town, Cabbagetown	Church and Wellesley	Richmond, Adelaide, King	Kensington Market, Chinatown, Grange Park	Harbourfront East, Union Station, Toronto Islands	Business reply mail Processing Centre, South Central Letter Processing Plant Toronto	Little Portugal, Trinity	First Canadian Place, Underground city	Toronto Dominion Centre, Design Exchange
count	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000	40.000000
mean	0.516720	0.450157	0.489284	0.544259	0.441317	0.425036	0.536833	0.441330	0.453077	0.228229	0.445943	0.535344	0.492545
std	0.322717	0.287701	0.295978	0.277428	0.268894	0.277667	0.318203	0.270373	0.267347	0.140817	0.269655	0.319190	0.306030
min	0.022654	0.006211	0.015435	0.016849	0.000000	0.010021	0.041473	0.017647	0.023491	0.000000	0.000000	0.038117	0.028560
25%	0.205385	0.169978	0.201258	0.276036	0.208079	0.164605	0.250582	0.217140	0.205808	0.112500	0.188064	0.256897	0.232923
50%	0.482120	0.376558	0.441772	0.612425	0.421005	0.402246	0.538225	0.406612	0.434832	0.234028	0.446883	0.537877	0.463388
75%	0.825258	0.688776	0.762926	0.786125	0.679695	0.659434	0.858806	0.699255	0.710632	0.307986	0.658032	0.857430	0.786003
max	0.986819	0.912799	0.905833	0.902647	0.910474	0.890927	0.987770	0.884447	0.799440	0.562500	0.956044	0.986556	0.989565

RESULTS

...AND THE WINNERS ARE ...

	1st Matching Neighborhood	1st Matching Score	2nd Matching Neighborhood	2nd Matching Score	3rd Matching Neighborhood	3rd Matching Score	4th Matching Neighborhood	4th Matching Score
Toronto Neighborhood								
Commerce Court, Victoria Hotel	Midtown South	0.99051	Midtown	0.988213	Financial District	0.937204	Murray Hill	0.929568
Toronto Dominion Centre, Design Exchange	Midtown	0.989565	Midtown South	0.96102	Murray Hill	0.917313	Financial District	0.890235
Richmond, Adelaide, King	Midtown South	0.98777	Midtown	0.98027	Financial District	0.939819	Murray Hill	0.937939
Studio District	Midtown South	0.987567	Soho	0.95375	Greenwich Village	0.946224	Midtown	0.916797
St. James Town	Midtown	0.986819	Midtown South	0.982041	Noho	0.945272	Flatiron	0.941845
First Canadian Place, Underground city	Midtown South	0.986556	Midtown	0.982103	Soho	0.940827	Financial District	0.939222
The Annex, North Midtown, Yorkville	Midtown	0.981283	Midtown South	0.951889	Financial District	0.914043	Murray Hill	0.890407
Moore Park, Summerhill East	Financial District	0.968085	Greenwich Village	0.878723	Flatiron	0.689362	Civic Center	0.678723
Little Portugal, Trinity	Little Italy	0.956044	Noho	0.859302	Greenwich Village	0.853575	Chinatown	0.84031
Central Bay Street	Financial District	0.955029	Midtown	0.934248	Noho	0.910914	Murray Hill	0.892248

THANK YOU

