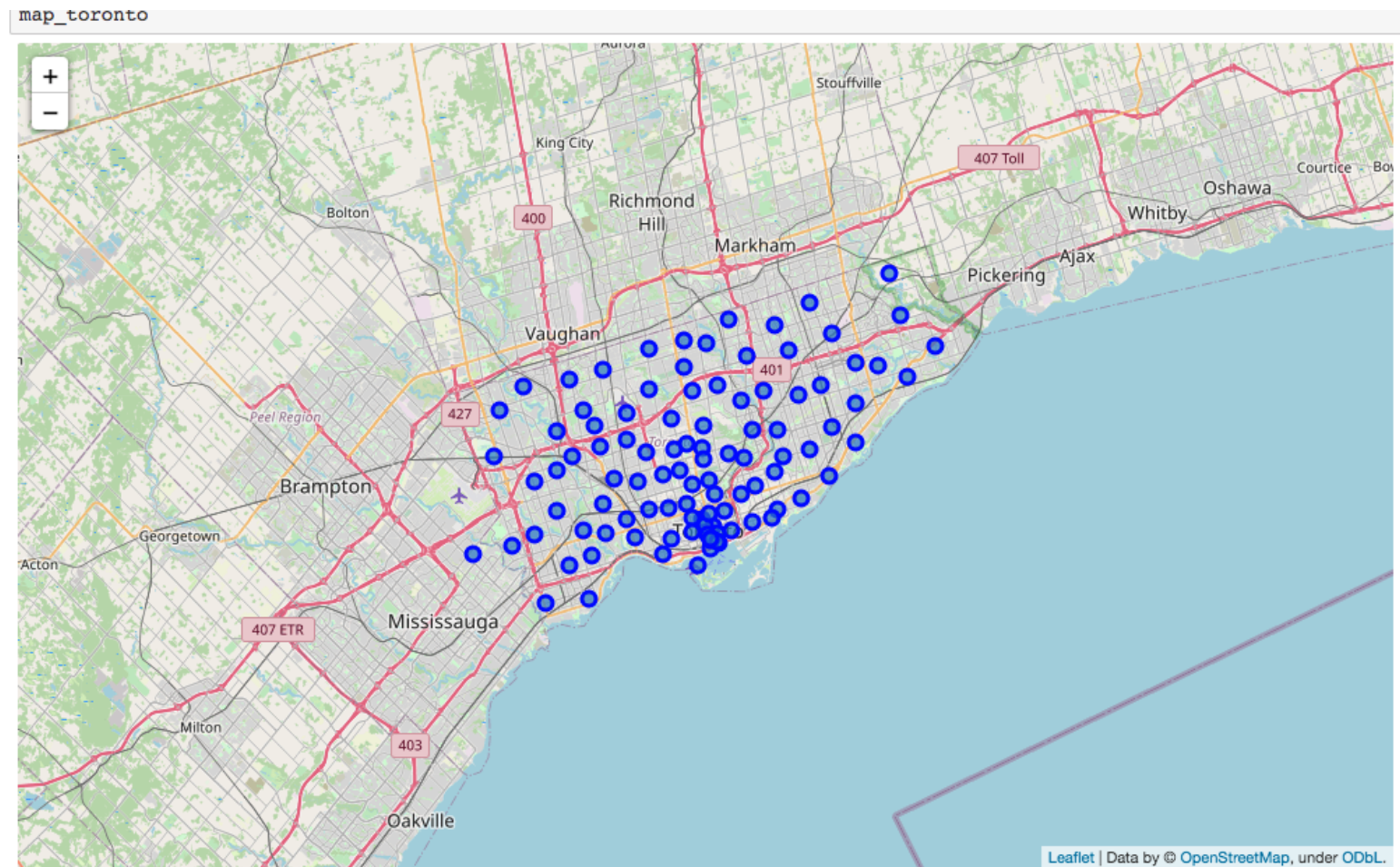# Prediction of Ideal Neighborhoods to Open New Coffee Shops

- Coffee shop is the most popular venue in Toronto
- Law of supply and demand
- An ideal location to open a new coffee shop is in a neighborhood have lower than expect number of shop

Distribution of 103 postal code neighbor hood in Toronto

FourSquare API found 2123 venues in Tronoto

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	TTC stop #8380	43.752672	-79.326351	Bus Stop
2	Parkwoods	43.753259	-79.329656	Sun Life	43.754760	-79.332783	Construction & Landscaping
3	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
4	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena

Number of venus vary greatly in different neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Agincourt	4	4	4	4	4	4
Alderwood, Long Branch	8	8	8	8	8	8
Bathurst Manor, Wilson Heights, Downsview North	24	24	24	24	24	24
Bayview Village	4	4	4	4	4	4
Bedford Park, Lawrence Manor East	27	27	27	27	27	27
Berczy Park	57	57	57	57	57	57
Birch Cliff, Cliffside West	4	4	4	4	4	4

Different venue preference vary by neighborhood

	Neighborhood	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
0	Berczy Park	Coffee Shop	Cocktail Bar	Bakery	Cheese Shop	Restaurant	Seafood Restaurant	Pharmacy	Beer Bar	Farmers Market	Hotel
1	CN Tower, King and Spadina, Railway Lands, Har...	Airport Lounge	Airport Service	Airport Terminal	Coffee Shop	Rental Car Location	Boat or Ferry	Airport	Airport Food Court	Airport Gate	Sculpture Garden
2	Central Bay Street	Coffee Shop	Café	Sandwich Place	Italian Restaurant	Salad Place	Bubble Tea Shop	Burger Joint	Japanese Restaurant	Korean Restaurant	Indian Restaurant
3	Christie	Grocery Store	Café	Park	Baby Store	Italian Restaurant	Nightclub	Candy Store	Restaurant	Coffee Shop	American Restaurant
4	Church and Wellesley	Coffee Shop	Japanese Restaurant	Sushi Restaurant	Gay Bar	Restaurant	Yoga Studio	Pub	Hotel	Fast Food Restaurant	Mediterranean Restaurant

Top 20 venue categories in Toronto

```
: Coffee Shop               183
  Café                       94
  Restaurant                 62
  Park                       54
  Pizza Place                49
  Italian Restaurant         45
  Hotel                      41
  Japanese Restaurant        40
  Sandwich Place             40
  Bakery                     40
  Clothing Store             36
  Gym                        35
  Fast Food Restaurant       30
  Grocery Store              29
  Sushi Restaurant           28
  Bank                       28
  Bar                        27
  Breakfast Spot             23
  Seafood Restaurant         23
  Pharmacy                   22
```

- One-hot dataframe of each categories
- K-fold validation for splitting the train/test set due to small sample size
- *Multi-linear regression (scikit-learn)*
- Metrics: *Coefficient of determination,(R2) Mean squared error and coefficeincy to determine if multiple linear regression is suitable for predicting the number of a specific category.*

Use multi-linear regression for predicting different categories number

	name	r2	mse	coef
0	Coffee Shop	0.631204	3.228614	3.228614
18	Seafood Restaurant	0.419422	0.266531	0.266531
1	Café	0.269827	1.332593	1.332593
9	Bakery	0.107140	0.467728	0.467728
8	Sandwich Place	-0.058658	0.439785	0.439785
12	Fast Food Restaurant	-0.160709	0.358247	0.358247
2	Restaurant	-0.260870	0.968311	0.968311
6	Hotel	-0.349577	1.136891	1.136891
5	Italian Restaurant	-0.454025	1.226595	1.226595
7	Japanese Restaurant	-0.485812	0.673359	0.673359
19	Pharmacy	-0.545532	0.288691	0.288691
3	Park	-0.614441	0.708853	0.708853
14	Sushi Restaurant	-0.712241	0.762295	0.762295
4	Pizza Place	-0.745298	1.023304	1.023304
15	Bank	-1.067827	0.551912	0.551912
11	Gym	-1.141358	0.827415	0.827415
17	Breakfast Spot	-1.347859	0.603516	0.603516
16	Bar	-1.605440	0.656598	0.656598
13	Grocery Store	-1.807140	0.822999	0.822999
10	Clothing Store	-12.434890	2.082187	2.082187

Training model to predict coffee shop number in each neighborhood

- R2 score: 0.8719796647644302

- coefficiency: [-0.05698575 0.36527995 0.49928682 0.06049385 0.6038981 0.61981172 0.43625695 0.82013551 0.44346885 0.34628185 -0.05717205 -0.55038192 -0.10135574 0.18406391 0.71333691 0.15898316 0.49716479 0.61669379 -0.61514679]

- intercept: -0.1697859631662182

Rank most favorable place for opening a coffee shop (predict number – existing number)

	Neighborhood	Coffee Shop	predict	diff
89	University of Toronto, Harbord	0	4.252432	4.252432
18	Davisville	2	4.685824	2.685824
19	Davisville North	0	2.269935	2.269935
88	Toronto Dominion Centre, Design Exchange	11	12.575825	1.575825
27	Downsview West	0	1.459938	1.459938
39	Guildwood, Morningside, West Hill	0	1.405996	1.405996
38	Golden Mile, Clairlea, Oakridge	0	1.216439	1.216439
37	Glencairn	0	1.209227	1.209227
46	India Bazaar, The Beaches West	0	1.205436	1.205436

Conclusion

- University of Toronto, Harbord, Davisville and Davisville North are neighborhood are neighborhoods that ideal for open new coffee shop.