

Coffee shop expansion in Downtown Toronto

CAPSTONE DATA SCIENCE PROJECT

LUYING PEI

MAY 22, 2019

Introduction to business problem

- Business owner of a small coffee shop is looking to expand his business to new locations in Downtown Toronto
- Current location of the coffee shop is on Central Bay Street, a neighbourhood of Downtown Toronto
- Due to the complex and varying characteristics of different neighbourhoods in Downtown Toronto, the owner hired a team of data scientists for determining the best locations to open up new coffee shops
- List of criteria for selecting desired neighbourhoods:
 - Popular/high-demand area: comparable number of venues to current location
 - Less potential competition: lower number of existing coffee shops
 - Similar neighbourhood in characteristics as current neighbourhood

Project approach and workflow

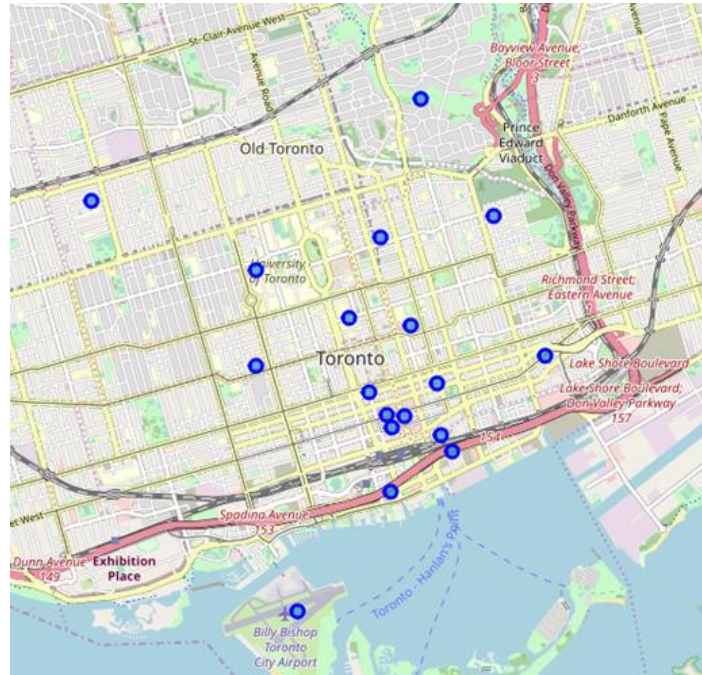
Data acquisition, cleaning, analysis and visualization

- Reliable data sources to conduct data exploration and data analysis:
 - Toronto neighborhood data from Wikipedia ('https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M')
 - Geographical coordinates matching postal codes of Toronto in csv file (provided in Week 3)
 - Foursquare location data (venues) will be used to cluster, segment, target, and position to craft recommendations for determining locations for business expansion
- Retrieve and clean up data (i.e. remove redundancies and irrelevant data)
- Merge and reorganize data (i.e. geographical coordinates and neighbourhood data)
- Analyze data using statistical testing and machine learning methods
- Visualize data using tables, maps and graphs
- Draw conclusions based on results and provide future recommendations

Exploring Downtown Toronto

Select popular neighbourhoods with high venue counts

- We have compiled a list of 17 neighbourhoods based on postal codes in Downtown Toronto
- After retrieving a list of venues in each neighbourhood, we ranked the neighbourhoods based on the number of venues
- We have selected 11 neighbourhoods (including Central Bay Street) that has a venue count of 87 or higher

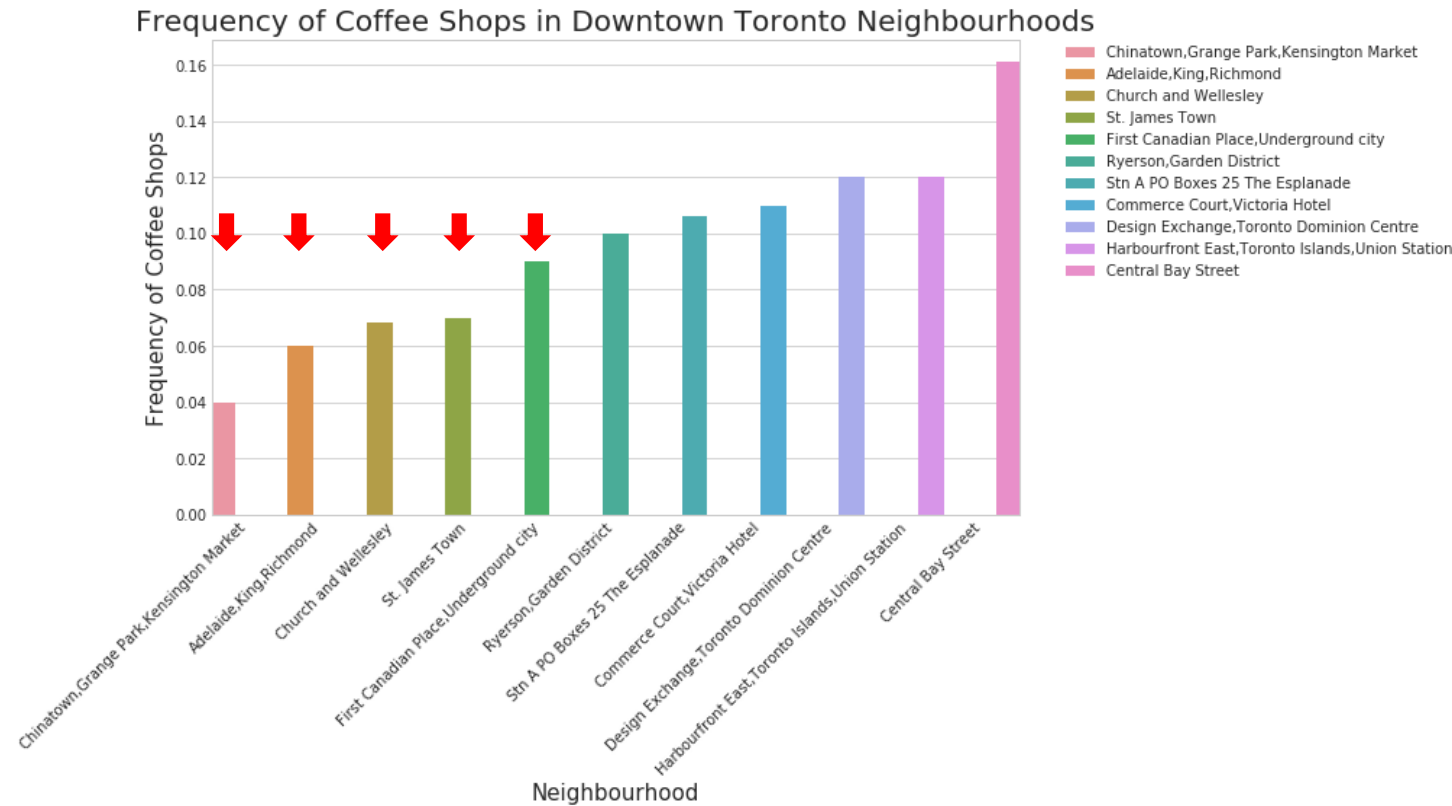


	Neighbourhood	Venue
0	Adelaide, King, Richmond	100
8	Commerce Court, Victoria Hotel	100
16	St. James Town	100
15	Ryerson, Garden District	100
12	Harbourfront East, Toronto Islands, Union Station	100
10	First Canadian Place, Underground city	100
9	Design Exchange, Toronto Dominion Centre	100
5	Chinatown, Grange Park, Kensington Market	100
17	Stn A PO Boxes 25 The Esplanade	94
7	Church and Wellesley	88
4	Central Bay Street	87
1	Berczy Park	55
13	Harbourfront, Regent Park	48
3	Cabbagetown, St. James Town	44
11	Harbord, University of Toronto	34
6	Christie	16
2	CN Tower, Bathurst Quay, Island airport, Harbourf...	14
14	Rosedale	4

Eliminating potential competition

Select neighbourhoods with less existing coffee shops

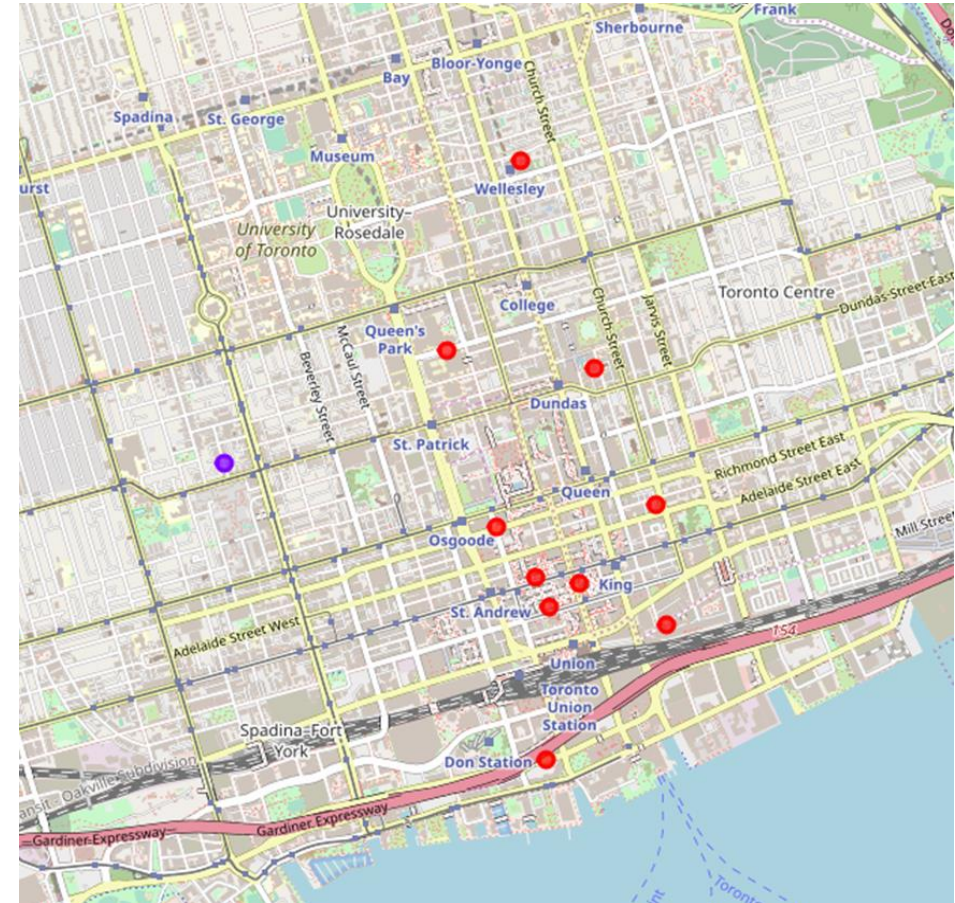
- Using one-hot-encoding, we have retrieved a summary table of venues sorted by category in the selected neighbourhoods
- We have selected the frequency of coffee shops in each neighbourhood
- We have ranked the selected neighbourhoods from the lowest to highest frequency of coffee shops
- We can select the 5 neighbourhoods with the lowest frequencies of coffee shops



Clustering neighbourhoods via k-means

Create clusters of neighbourhoods based on common venues

- Using k-means clustering method, we have categorized the 11 selected neighbourhoods of Downtown Toronto into 2 different clusters based on the most common venues occurring in each
 - Cluster 1: 10 neighbourhoods (red)
 - Cluster 2: 1 neighbourhood (purple)
- Neighbourhoods in Cluster 1 (including Central Bay Street) share similarities with each other



Examining cluster with Central Bay Street

Observe neighbourhoods similar to current business location

- The first cluster contains 10 neighbourhoods including Central Bay Street
- The 9 other neighbourhoods all share similarities with Central Bay Street in terms of the types of venues and how common they are in each neighbourhood
- Interestingly, the 5 neighbourhoods with the lowest frequencies of coffee shops that we selected previously are all included in this cluster

	Neighbourhood	Cluster Labels	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	Church and Wellesley	0	Japanese Restaurant	Coffee Shop	Sushi Restaurant	Restaurant	Gay Bar	Gastropub	Men's Store	Mediterranean Restaurant	Hotel	Gym
1	Ryerson, Garden District	0	Coffee Shop	Clothing Store	Cosmetics Shop	Café	Middle Eastern Restaurant	Ramen Restaurant	Sporting Goods Shop	Bubble Tea Shop	Pizza Place	Italian Restaurant
2	St. James Town	0	Coffee Shop	Café	Hotel	Restaurant	Cosmetics Shop	Bakery	Breakfast Spot	Gastropub	Clothing Store	Cocktail Bar
3	Central Bay Street	0	Coffee Shop	Café	Italian Restaurant	Burger Joint	Middle Eastern Restaurant	Bar	Sushi Restaurant	Bakery	Chinese Restaurant	Japanese Restaurant
4	Adelaide, King, Richmond	0	Coffee Shop	Café	Thai Restaurant	Bar	Steakhouse	American Restaurant	Burger Joint	Cosmetics Shop	Bakery	Hotel
5	Harbourfront East, Toronto Islands, Union Station	0	Coffee Shop	Hotel	Aquarium	Italian Restaurant	Café	Fried Chicken Joint	Scenic Lookout	Pizza Place	Bakery	Restaurant
6	Design Exchange, Toronto Dominion Centre	0	Coffee Shop	Hotel	Café	Restaurant	Deli / Bodega	Steakhouse	Bakery	Seafood Restaurant	Gastropub	Italian Restaurant
7	Commerce Court, Victoria Hotel	0	Coffee Shop	Café	Hotel	Restaurant	American Restaurant	Bakery	Italian Restaurant	Deli / Bodega	Seafood Restaurant	Steakhouse
9	Stn A PO Boxes 25 The Esplanade	0	Coffee Shop	Restaurant	Café	Italian Restaurant	Cocktail Bar	Hotel	Seafood Restaurant	Beer Bar	Fast Food Restaurant	Bakery
10	First Canadian Place, Underground city	0	Coffee Shop	Café	Hotel	Bakery	Deli / Bodega	Burger Joint	Restaurant	Seafood Restaurant	Bar	Steakhouse

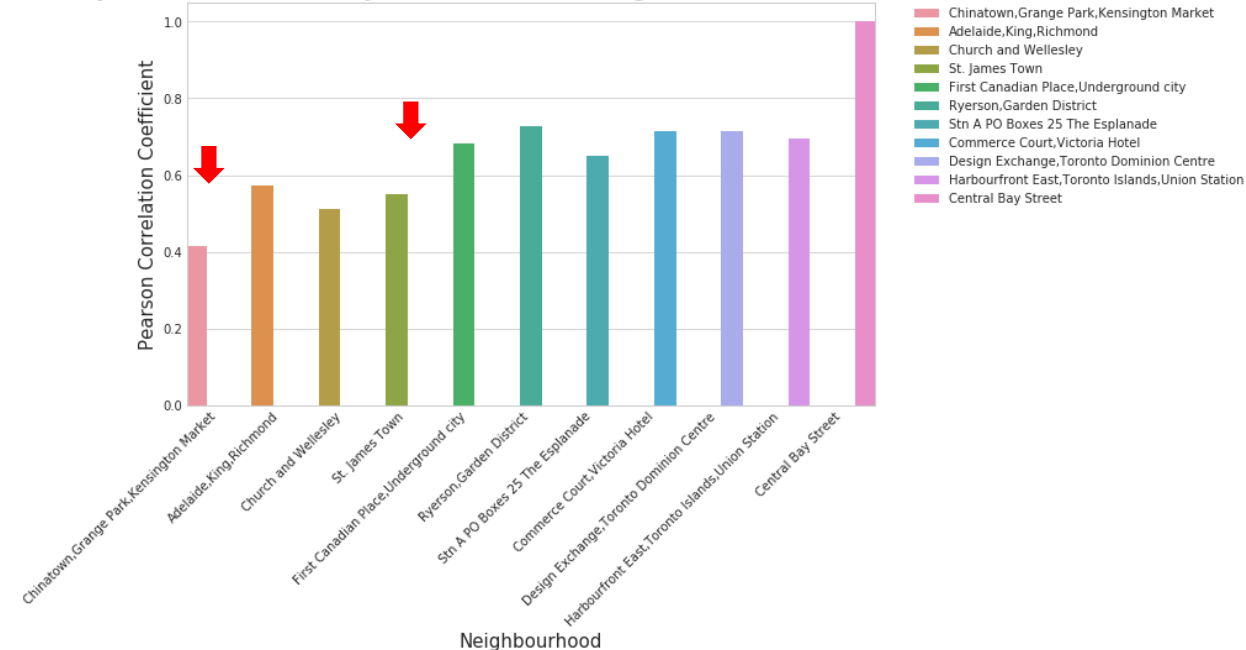
Statistical testing for correlation

Evaluate similarities via Pearson correlation coefficient and p-values

- We have tested the correlation between each neighbourhood vs. Central Bay Street by calculating Pearson correlation coefficients and p-values
- From the 5 neighbourhoods with the lowest frequencies of coffee shops, we select the 2 with the highest Pearson correlation coefficients (and lowest p-values)
 - Adelaide/King/Richmond
 - First Canadian Place/Underground City

	Neighbourhood	Coffee Shop	Pearson Correlation Coefficient	P-value
0	Chinatown,Grange Park,Kensington Market	0.040000	0.414713	7.124575e-09
1	Adelaide,King,Richmond	0.060000	0.571578	5.304815e-17
2	Church and Wellesley	0.068182	0.512197	1.990909e-13
3	St. James Town	0.070000	0.549027	1.464666e-15
4	First Canadian Place,Underground city	0.090000	0.682063	5.596237e-26
5	Ryerson,Garden District	0.100000	0.726425	8.293996e-31
6	Stn A PO Boxes 25 The Esplanade	0.106383	0.649456	6.105092e-23
7	Commerce Court,Victoria Hotel	0.110000	0.715710	1.477158e-29
8	Design Exchange,Toronto Dominion Centre	0.120000	0.714348	2.110117e-29
9	Harbourfront East,Toronto Islands,Union Station	0.120000	0.696131	2.050364e-27
10	Central Bay Street	0.160920	1.000000	0.000000e+00

Similarity between Central Bay Street vs. Other Neighbourhoods in Downtown Toronto



Conclusions and future directions

- After conducting a detailed data analysis of the neighbourhoods in Downtown Toronto, we have found the most suitable locations for opening up a new coffee shop based on the business owner's criteria
- We recommend the business owner to open new coffee shops in these two particular neighbourhoods in Downtown Toronto: Adelaide/King/Richmond and First Canadian Place/Underground City
- Other neighbourhoods can also be considered if the business owner decides to change the scale of expansion or modify the criteria for expansion