



IBM Applied Data Science Capstone Project

**Evaluating Restaurant Locations in Toronto
By Using Major Crime Indicators (MCI)**

Richard Xue, MBA, MS

April 2, 2020



Introduction/Business Problem

- This project uses Toronto Major Crime Indicators (MCI) data to identify neighborhoods in Toronto that may be good locations to build new restaurants.
- Demand for new restaurants with more menu options is growing in Toronto. Healthy restaurants are emerging to attempt to meet the demand for meatless cuisine, vegetarian or vegan tasting menus.¹
- Selecting new venue locations is always a challenge for restaurants owners and investors. Few factors such as health-conscious and safe neighborhoods need to be carefully considered before making the final decision of new venue locations.
- This project explores data insights specifically to identify restaurant locations in safe neighborhoods in Toronto that may be under-served by current venues. Detail data analysis is performed to explain the factors that lead to such data insights on selecting of new restaurant locations.



Audience & Stakeholders

- The audience for this project is restaurants owners, investors and anyone servicing the fast-food and sit-in food segments. These stakeholders have a vested interest in responding to market demand by making investment and operations decisions based on data insights. This project explores some of those data insights and provides guidance specific to venue location.

Data Sources

1. Major Crime Indicators (MCI)

- Toronto Policy Service website
- Download MCI data as CSV file

premisetype	ucr_code	ucr_ext	offence	...	occurrencedayofyear	occurrencedayofweek	occurrencehour	MCI	Division
Outside	1610	220	Robbery - Other	...	283	Saturday	1	Robbery	D53
Outside	1610	200	Robbery - Mugging	...	283	Saturday	1	Robbery	D53
House	2120	200	B&E	...	84	Wednesday	10	Break and Enter	D53
House	2120	200	B&E	...	280	Wednesday	17	Break and Enter	D53
House	2120	200	B&E	...	14	Wednesday	14	Break and Enter	D53

Data Sources

2. Foursquare Venue Data

- Foursquare provides useful venues data.
- Get Venues via Foursquare REST GET API

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park / Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park / Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park / Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
3	Regent Park / Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa
4	Regent Park / Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot

Data Sources

3. *Neighborhood Geocodes Data*

- Wiki webpage contains Neighborhood Geocodes data
- Use Python BeautifulSoup package to get Postal Code, Borough, and Neighborhood information.

	Postal Code	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park / Harbourfront
3	M6A	North York	Lawrence Manor / Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park / Ontario Provincial Government



Data Gathering & Cleansing

- Load Toronto MCI data from Toronto Policy Service
- Retrieve Foursquare Food Venues Data
- Load Neighborhood Geocodes



Data Feature Selections

1. MCI Data

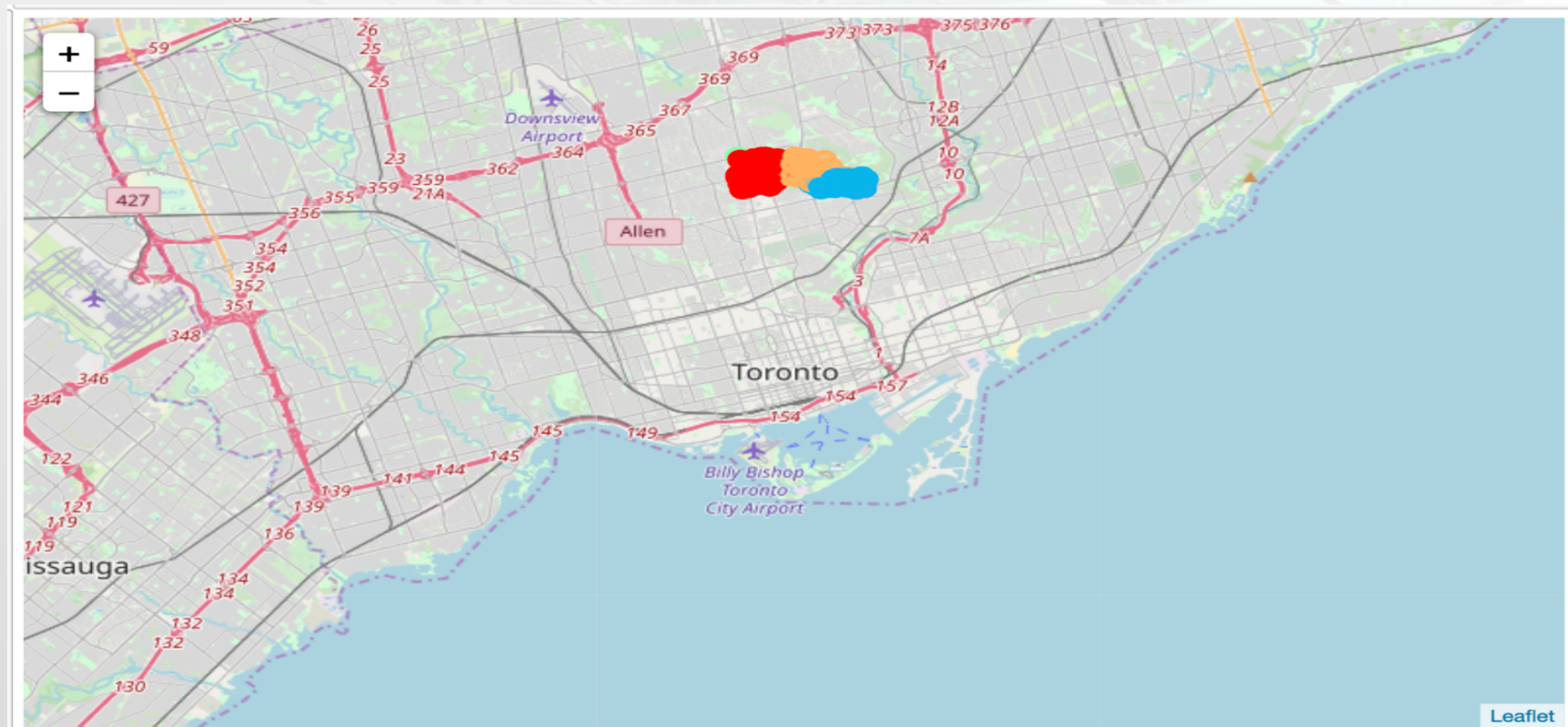
- Neighborhood Name
- MCI indicator

2. Foursquare Venues Data

- Neighborhood Name
- Category

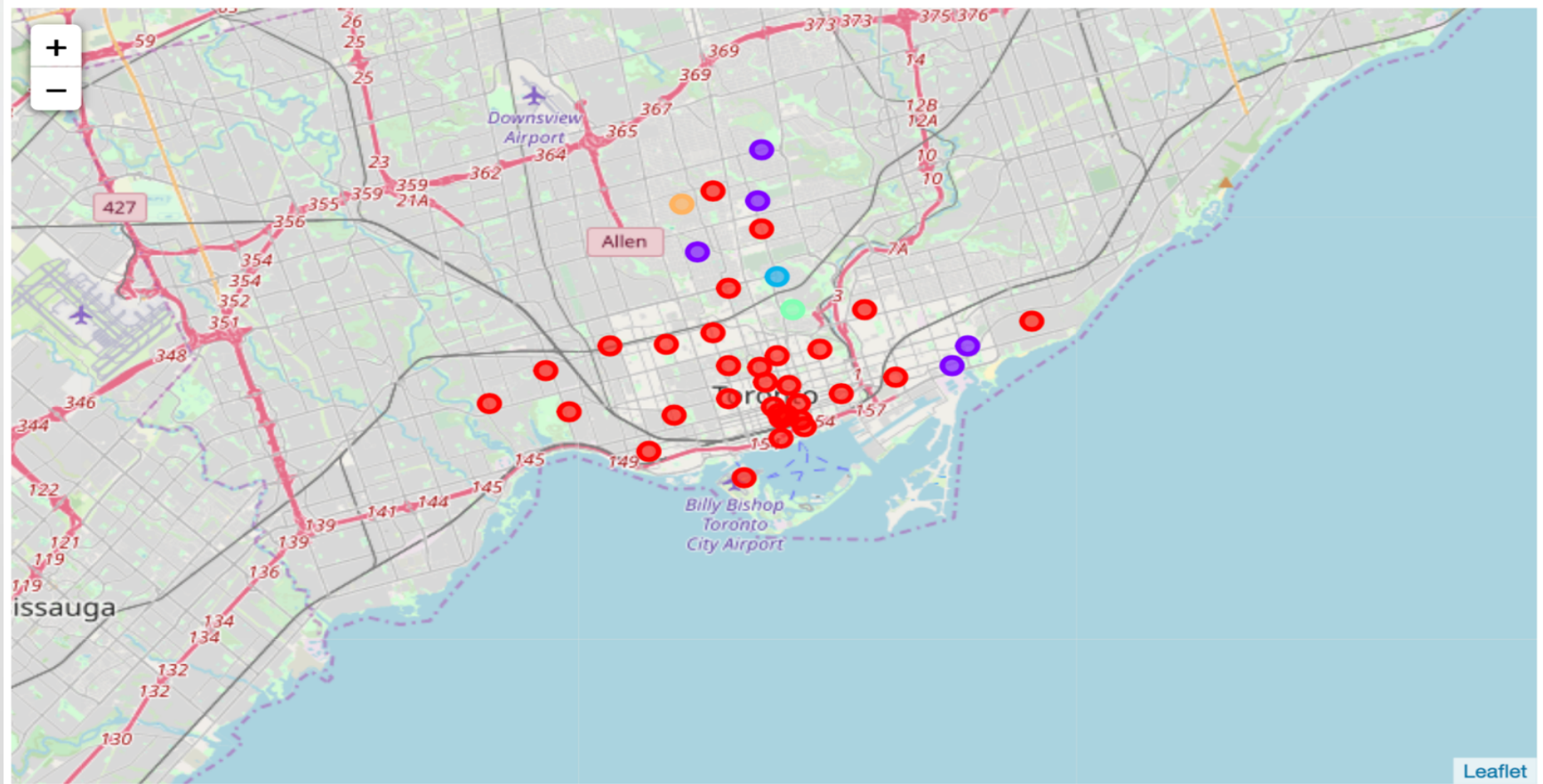
- Data Preparation
- Data Exploration
- Data Categorization
- Clustering of Neighborhoods
- Choropleth Crime Neighborhood Map
- Choropleth Venue Neighborhood Map
- Data Comparison
- Examine the Clusters

Choropleth Crime Neighborhood Map



- Three top clusters highlighted in red, orange and blue are located outside of the downtown area of Toronto.
- The downtown area of Toronto are relative safer comparing to these three clusters.

Choropleth Venue Neighborhood Map



- Top venue cluster is highlighted in red which is the downtown area of Toronto.
- Many restaurants locate in the downtown area of Toronto.

Common Venues of Top Venue Cluster

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	M5A	Downtown Toronto	Regent Park / Harbourfront	43.654260	-79.360636	0	Coffee Shop	Pub	Park	Theater	Mexican Restaurant
1	M7A	Downtown Toronto	Queen's Park / Ontario Provincial Government	43.662301	-79.389494	0	Coffee Shop	Gym	Diner	Music Venue	Mexican Restaurant
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Clothing Store	Coffee Shop	Café	Bubble Tea Shop	Japanese Restaurant
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0	Coffee Shop	Italian Restaurant	Café	Restaurant	Clotier's
4	M4E	East Toronto	The Beaches	43.676357	-79.293031	0	Trail	Pub	Health Food Store	Donut Shop	Diner
5	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Coffee Shop	Cocktail Bar	Bakery	Cheese Shop	Café
6	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	0	Coffee Shop	Italian Restaurant	Sandwich Place	Japanese Restaurant	Restaurant
7	M6G	Downtown Toronto	Christie	43.669542	-79.422564	0	Grocery Store	Café	Park	Gas Station	Baby's
8	M5H	Downtown Toronto	Richmond / Adelaide / King	43.650571	-79.384568	0	Coffee Shop	Restaurant	Café	Bakery	
9	M6H	West Toronto	Dufferin / Dovercourt Village	43.669005	-79.442259	0	Bakery	Pharmacy	Bar	Pool	Bar
10	M5J	Downtown Toronto	Harbourfront East / Union Station /	43.640816	-79.381752	0	Coffee Shop	Aquarium	Italian Restaurant	Café	Hospital

- Most common venues in top venue cluster include: *Coffee Shop, Pub, Italian Restaurant, Mexican Restaurant, Japanese Restaurant, etc..*

Comparison of Choropleth Crime and Venue Neighborhood Maps

Downtown Neighborhoods:

- Have higher business density
- Have many existing restaurant venues
- Attract many consumers to dine in or have a drink
- Have much lower crimes reported in the past five years (2014 – 2019)
- MCI indicator is another key factor to seek venue locations for new businesses



Discussion

- Although, number of venues and safety are among the key reasons for restaurant owner and investors to consider when they select a location to open a restaurant. There are other factors we may want to consider in our data analytic and model methodology.
- For example, should we consider Neighborhood Venue Average Health Ratings (VAHR) when selecting a location beside the safety reason?
- Should we consider the existing competitors with the same food type?
- Population to Venue Ratio, Population Density and Population Growth Rate may also be good indicators for restaurant owners or investors to look at details on the restaurant, fast-food, or dine-in food market opportunities.



Recommendations

- The in-depth data analysis and machine learning of this project provide several decision factors and the insights gained by the data analysis, visualizations and machine learning.
- These factors and insights together highlight a framework in considering locations for restaurant owners, investors and stakeholders seeking to open new venues in neighborhoods of Toronto.
- The data analyses, visualizations, and machine learning in this project suggest how the MCI data and venue data can be considered as part of a decision framework for locating new venues.
- One key contribution is the data comparison of Toronto crime data with venue data.



Conclusion

- This project explores data from various data sources, composites data, and create Choropleth maps for neighborhood venues and crime data.
- The data analysis, visualization of choropleth maps along with insights contribute to the identification of restaurant locations in safe neighborhoods in Toronto.
- These data analysis and insights may be useful for restaurant owners, investor and stakeholders as considering factors for new restaurant locations.
- These data analysis, insights and factors could also be applied for evaluating existing venues to rebrand or renovate their venues with respect to neighborhood food category.

Questions?