Peer-graded Assignment: Capstone Project - The Impacts of Neighborhoods on Restaurant Success or Failure

Finding the Best Restaurant Opportunity for Each Neighborhood

Abstract

A data science review the demographics of neighborhoods in Toronto, Canada, combined with the current restaurant offerings in that neighborhood, give key insights into what type of restaurant has the best chance of having a good customer base. The results of this report are intended go into a business decision process, as part of an overall selection process.

William Brent Lander
williambrentlander@gmail.com

Table of Contents

Table of Figures

Table of Tables

Introduction

The goal of this effort is to apply the tools and methodologies of data science in reducing the risk of starting a new restaurant in the Toronto area.

According to BPlans[i], a web site popular for promoting entrepreneurial business planning, the largest reason for restaurants failing is lack of planning. The research here can be applied early in the planning phase, providing location/competition market research. This early application allows for choices being made which reflect the impact of the choice of location (e.g., neighborhood), and choice of venue type, while avoiding areas where others have recently failed in their efforts.

This effort uses data science techniques, and available data, in an attempt to reduce risk in a potential new restaurant venture in the Toronto area. Restaurants tend to be neighborhood based – if they succeed, they will succeed in the surrounding neighborhood, and then expand their base. By using resources available on the Internet, a recommendation of the location and type of restaurant that is most likely to succeed is made. This recommendation would be used by the business developers, in combination with other business logic, for establishing a successful new venue.

In the course of this project, 205 neighborhoods were examined (with191 having existing restaurants), with 105 different venues of restaurants, and a total of 5209 individual restaurants. An additional 208 restaurants were considered, separately, which had failed (35 different categories of restaurants over 69 neighborhoods) in an effort to understand adverse impacts.

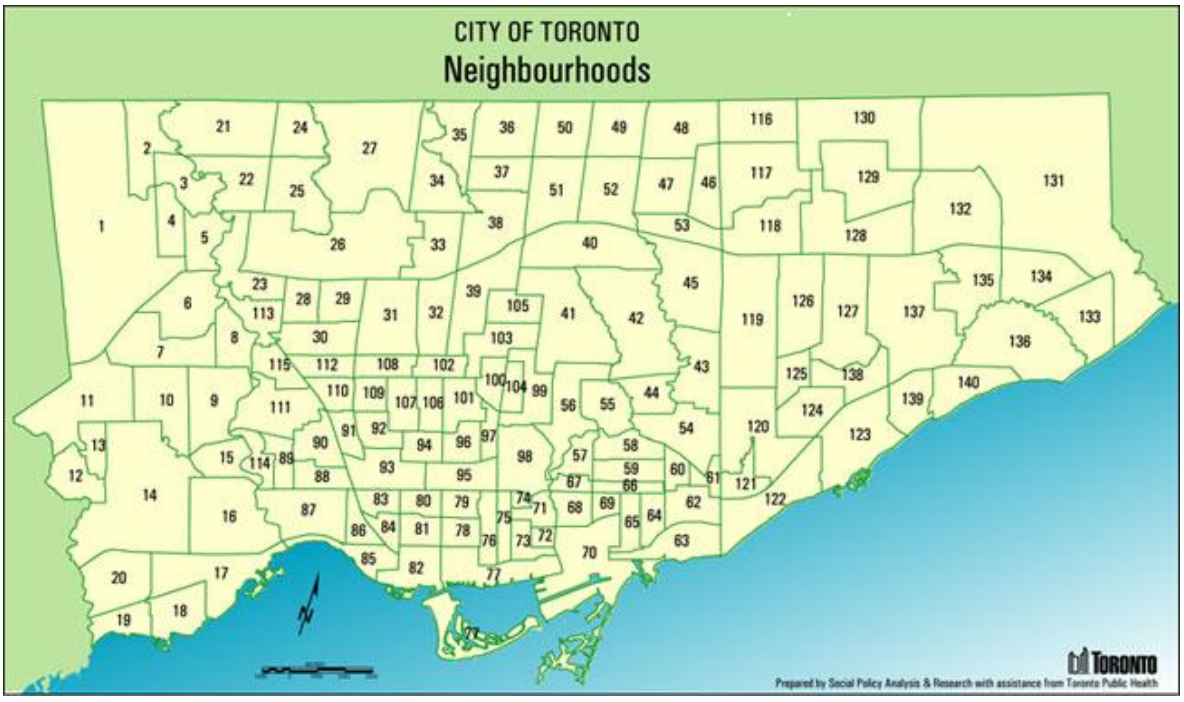


Figure 1- Overview of Neighborhoods in Toronto Canada

This work addresses the following analysis questions:

Q1. What are the most successful types of restaurants (e.g., coffee shops, national cuisine, snacks) in greater Toronto?
Q2. In what neighborhoods are successful restaurants currently located?
Q3. What types of successful restaurants are located in each neighborhood?
Q4. What types of restaurants failed in the recent past?
Q5. Where did those restaurants fail?

The following table provides guidance to where in the content the various questions are addressed.

Table 1 - Reference Matrix Showing Where Questions are Addressed in Analysis and Results

Question	Analysis	Results
Q1	A.3	R.3
Q2	A.4, A.2	R.4
Q3	A.5	R.5
Q4	A.10	R.7
Q5	A.8 A.9	R.7, R8

Data Sources

Demographic data came from Statistics Canada, the national statistical office, as published in 2016. Direct access to this information is no longer available outside Canada, but is reported in Wikipedia. This information included the list of neighborhoods, size (square kilometers), population, and average income.

Geolocation information for the various neighbors came from the Nominatim geolocator service (https://nominatim.openstreetmap.org/), and was used to generate the location (latitude/longitude) for the center of each neighborhoods.

Information about existing restaurants came from the Foursquare venue API,
https://api.foursquare.com/v2/venues
This information included the location, and categoryID (or venue category) for each specific restaurant. Foursquare describes itself as "a technology platform that powers leading business solutions and consumer products through a deep understanding of location."

Additional information on the venues' categoryID came from the Foursquare API documentation found at https://developer.foursquare.com/docs/resources/categories . This information linked the numeric categoryID to a verbal description (e.g., "Afghan Restaurant", "Indonesian Restaurant", "Korean Restaurant"). The Foursquare documentation also included a taxonomy of restaurants, allowing for a categorial roll up of some of the more specific categories. For example, "Korean Restaurant" can be either a "Korean Restaurant", "Bossam/Jokbal Restaurant", "Bunsik Restaurant", "Gukbap Restaurant", "Janguh Restaurant", or a "Samgyetang Restaurant".

Information on restaurant closures in 2017 and 2018 came from

"The RIP LIST: 150 Toronto restaurants that closed in 2017 " https://dailyhive.com/toronto/rip-list-toronto-restaurants-closed-2017
"The RIP LIST: 147 Toronto restaurants that closed in 2018" https://dailyhive.com/toronto/rip-toronto-restaurants-closed-2018

Methodology

section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, and what machine learnings were used and why.

The following approach was taken:

M1. Gather the list of neighborhoods. This includes the Statistics Canada information, and additional lists of smaller more ethnic-centered neighborhoods. This gave a list of 205 neighborhoods for potential examination.
M2. Geo-locate (latitude, longitude) each of the neighborhoods. If no geolocation information is available, then remove those neighborhoods from further consideration. This resulted in the list of neighborhoods reduced to 191 neighborhoods.
M3. Transcribe the restaurant taxonomy from the Foursquare documentation into a machine usable dictionary. This preparatory step allows later consolidation of restaurant categories where there is not a meaningful number of those restaurants.
Early analysis showed, for example, that the individual numbers for the sub-categories of Korean Restaurants only obscured the success of Korean Restaurants by diluting the numbers while not providing any additional insights.
M4. For each neighborhood, query Foursquare about the available restaurants. Consolidate the restaurant venue breakdown by categories into the neighborhood data frame. Because Foursquare will only return information about 50 or less restaurants in a given query, the neighborhoods were divided into a series of circular sub-regions each of which was queried in turn. The Foursquare results were then compiled into a list, so that duplicate responses (e.g., the same restaurant being reported twice) were eliminated. In order to keep the number of responses from each sub-region less than 50, the size of each sub-region was reduced to a circle with a radius of 100 meters. The centers of the sub-regions were selected so that there were no gaps in coverage.
M5. Examine the marginals over the data frame from step 4. Summing across rows tells the number of each category of restaurant in the greater Toronto area (this is an indicator of the popularity of various categories, the most popular is “Coffee Shop”, followed closely by “Pizza Place”). Summing down each column gives the count of restaurants per neighborhood (This gives an indication of how favorable a neighborhood is to restaurants, the most popular neighborhoods being “Willowdale” and “Little Italy”
M6. Normalize the number of each category of restaurant with the population of each neighborhood. One of the important indicators of restaurant success is the size of the potential customer base. For example, if there is only one Italian Restaurant in a neighborhood of 1,000 people, the number of potential customers is larger than in one where there are 20 existing Italian Restaurants.
M7. Cross check the results so far by looking at a map showing for each neighborhood what category of restaurant is most popular.

Begin looking at counter indicating evidence – restaurants closing in the past two years by neighborhood.

M8. Collect the list of failed restaurants for 2017 and 2018, including name, venue, year of closure, and neighborhood where the restaurant was closed.
M9. Examine the number of restaurant closings by neighborhood, for estimating how impenetrable a neighborhood is to a new restaurant.

M10. Examine the number of closings by restaurant venue, determining which categories of restaurants are most likely for failing.

Analysis

A1. The 20 most popular restaurant venues in Toronto are shown in Figure 2.The most popular venue is Coffee Shop. While Coffee Shops were most popular in the business neighborhoods, they were pervasive, and found in most neighborhoods. There were 103 different types of venue found, most with only one restaurant featuring that venue.

A2. During the geolocation phase, the location and size of the following neighborhoods in Toronto could not be determined, so were removed from further consideration:

Table 2 - Neighborhoods Defined by Statistics Canada but no Geospatial Information Available

Clairville
Cricket Club
Distillery District/West Don Lands
Fort York/Liberty Village
Governor's Bridge/Bennington Heights
Humber Bay Shores
Humbermede
Humewood–Cedarvale
Playter Estates
Princess Gardens
Regal Heights
Regent Park/Trefann Court
Rockcliffe–Smythe

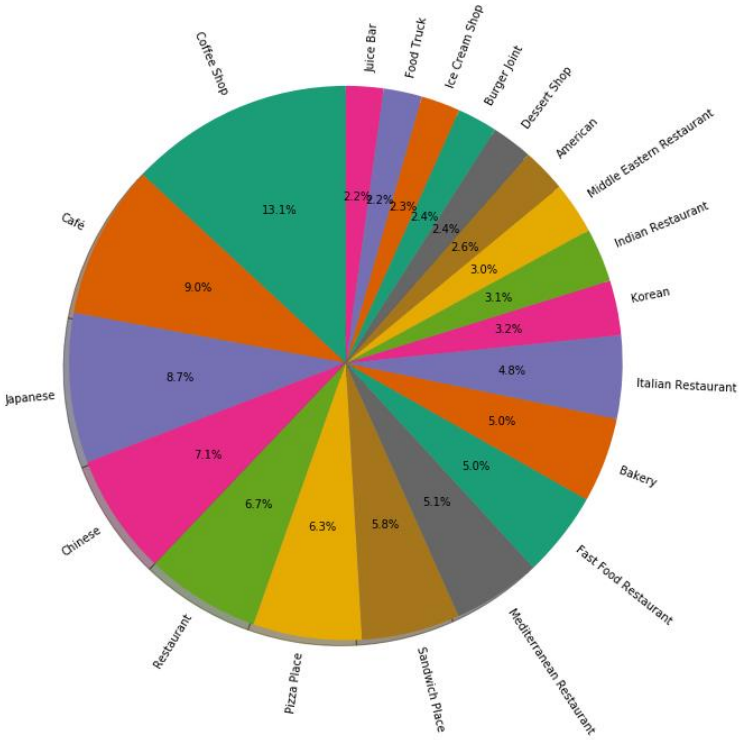


Figure 2- Distribution of Top 20 Restaurant Venues in Toronto Canada

A3. The histogram in Figure 3 shows the number of restaurants (across all venues) by neighborhood. Due to the number of neighborhoods, only those neighborhoods with 10 or more restaurants are shown. From this chart, it appears that the neighborhoods which are most

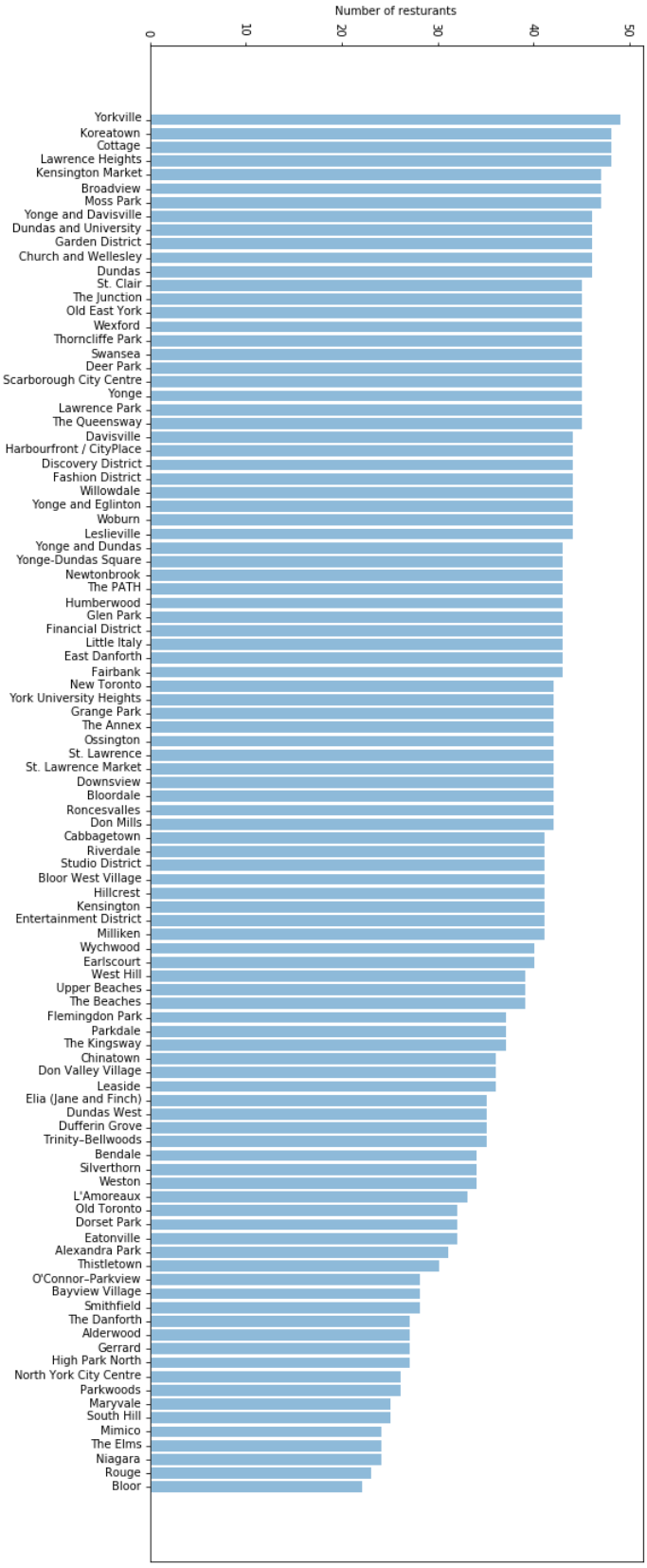


Figure 3- Count of Restaurants by Neighborhood, with 10 or more Restaurants

A4. Table 3 shows the number of restaurants for the most popular venues. From this, it is very clear that Toronto's citizens really like quick service restaurants

Table 3- Number of Restaurants in Each Venue in Toronto

Restaurant Venue	Count
Coffee Shop	406
Pizza Place	256
Café	205
Chinese	169
Fast Food Restaurant	166
Sandwich Place	163
Bakery	162
Japanese	144
Generic Restaurant	144
Mediterranean Restaurant	110
Italian Restaurant	103
Caribbean Restaurant	92
Indian Restaurant	90
Middle Eastern Restaurant	76
Burger Joint	75

Restaurant Venue	Count
American	60
Breakfast Spot	58
Ice Cream Shop	56
Deli / Bodega	53
Dessert Shop	53
BBQ Joint	51
Juice Bar	49
Korean	43
Bubble Tea Shop	39
Latin American Restaurant	39
Diner	35
Greek Restaurant	32
Vegetarian / Vegan Restaurant	29
African	26
Hot Dog Joint	25

A5. The following table scales the population of a neighborhood with the number of restaurants of a given venue. Only neighborhoods with three or more of a given restaurant venue are shown. This gives an estimate of the size of the customer base.

Table 4- Neighborhoods, Highest Count Venue, and Estimated Customer Base

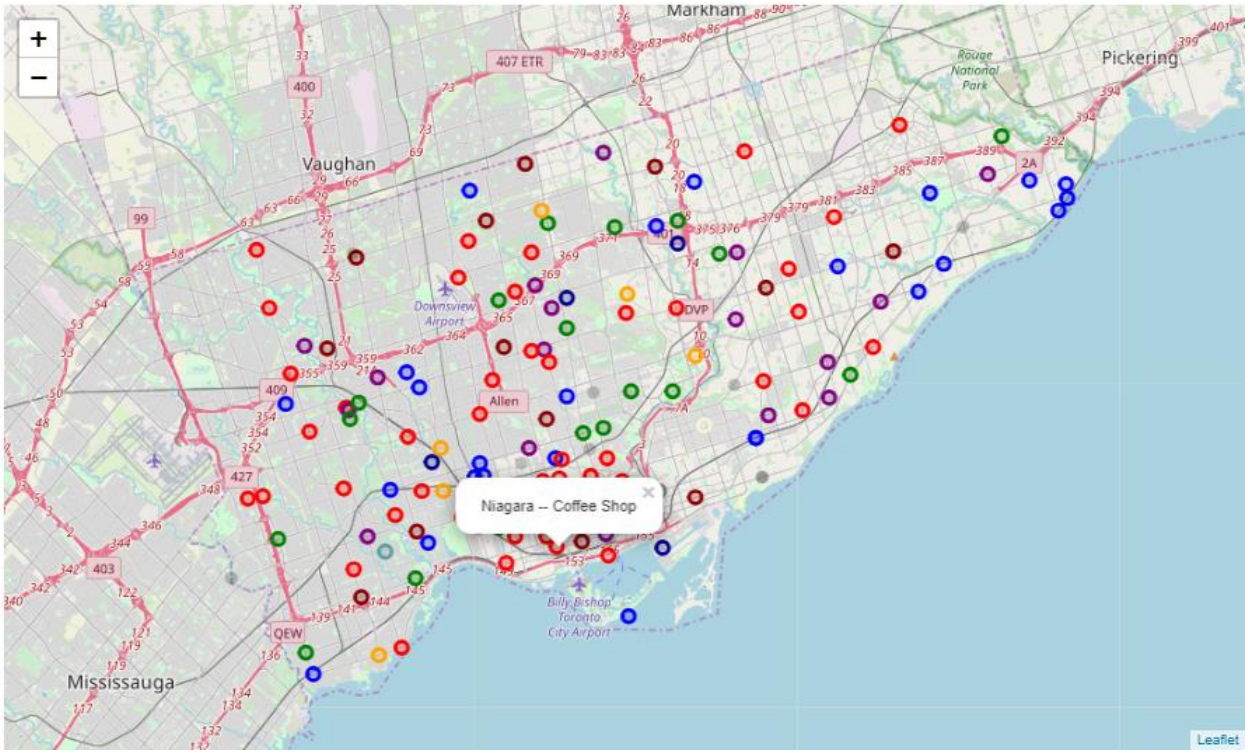
Neighborhood	Venue	#	Estimated Customer Base
Old East York	Sandwich Place	4	13,055
Woburn	Generic Restaurant	4	12,126
L'Amoreaux	Coffee Shop	4	11,465
Downsview	Café	4	9,153
Newtonbrook	Café	4	9,011
Smithfield	Japanese	4	8,749
Willowdale	Korean	5	8,628
Fairbank	Bakery	4	8,530
Parkdale	Pizza Place	4	7,091
Parkwoods	Coffee Shop	4	6,633
York University Heights	Pizza Place	4	6,535
Riverdale	Bakery	5	6,201

Neighborhood	Venue	#	Estimated Customer Base
Don Valley Village	Chinese	5	5,948
Davisville	Japanese	4	5,931
Leslieville	Coffee Shop	4	5,891
Bendale	Coffee Shop	5	5,789
High Park North	Pizza Place	4	5,686
East Danforth	Café	4	5,360
Don Mills	Japanese	4	5,343
Elia (Jane and Finch)	Caribbean Restaurant	9	5,333
Flemingdon Park	Sandwich Place	4	5,321
Mount Dennis	Caribbean Restaurant	4	5,321

Neighbor-hood	Venue	#	Estimated Customer Base
West Hill	Breakfast Spot	5	5,126
The Beaches	Café	4	5,104
Upper Beaches	Coffee Shop	4	4,957
Glen Park	Café	4	4,606
Hillcrest	American	4	4,581
Rouge	Coffee Shop	5	4,544
Thorncliffe Park	Coffee Shop	4	4,487
Silverthorn	Bakery	4	4,439
O'Connor–Parkview	Coffee Shop	4	4,435
Amesbury	Coffee Shop	4	4,329
Earlscourt	Coffee Shop	4	4,310
Thistletown	Caribbean Restaurant	4	4,197
Weston	Coffee Shop	4	4,119
Scarborough City Centre	Fast Food Restaurant	4	4,100
Roncesvalles	Pizza Place	4	3,999
The Annex	Japanese	4	3,900
Deer Park	Café	4	3,791
Mimico	Café	4	3,549
Dorset Park	Indian Restaurant	4	3,547
Leaside	Coffee Shop	4	3,469
Wilson Heights	Bakery	4	3,433
Church and Wellesley	Bubble Tea Shop	4	3,349
Eatonville	Coffee Shop	6	3,188
The Junction	Bakery	4	2,847
Swansea	Pizza Place	4	2,783
Cabbage – town	Coffee Shop	4	2,780
Steeles	Chinese	9	2,744
North York City Centre	Bubble Tea Shop	4	2,606
Dufferin Grove	Coffee Shop	4	2,468

Neighbor-hood	Venue	#	Estimated Customer Base
Bayview Village	Coffee Shop	5	2,456
Cliffside	Pizza Place	4	2,346
Brockton	Pizza Place	4	2,259
Grange Park	Japanese	4	2,251
The Kingsway	Bakery	4	2,195
New Toronto	Coffee Shop	5	2,091
Garden District	Middle Eastern Restaurant	4	2,060
Little Italy	Café	4	1,979
The Danforth	Café	4	1,962
Milliken	Chinese	14	1,876
Humberwood	Restaurant	4	1,82
Discovery	Japanese	4	1,815
Trinity–Bellwoods	Café	5	1,737
Lawrence Park	Coffee Shop	4	1,663
The Queensway	Coffee Shop	4	1,551
Yorkville	Restaurant	4	1,511
Wexford	Middle Eastern Restaurant	13	1,372
The Elms	Pizza Place	7	1,348
Harbourfront / CityPlace	Coffee Shop	11	1,306
South Hill	Café	5	1,243
Fashion District	Generic Restaurant	4	1,160
Bloor West Village	Japanese	5	1,035
Caribou Park	Pizza Place	4	1,024
Lawrence Heights	American	4	942
Kensington Market	Bakery	4	935
Humber Heights	Pizza Place	5	934
Hoggs Hollow	Coffee Shop	4	780
Financial District	Coffee Shop	12	45

A6. Geographically, the distribution of most popular types of restaurant venues looks as follows:



The venue legend for the map is:

Venue	Color
Coffee Shop	Red
Café	Blue
Pizza Place	Green
Japanese	Purple
Restaurant	Orange
Bakery	Dark Red
Chinese	Light Red
Fast Food Restaurant	Beige
Italian Restaurant	Dark Blue
Sandwich Place	Dark Green

Venue	Color
Mediterranean Restaurant	Cadet Blue
Caribbean Restaurant	Dark Purple
Indian Restaurant	White
Middle Eastern Restaurant	Pink
Burger Joint	Light Blue
American	Light Green
Breakfast Spot	Gray
BBQ Joint	Black
Dessert Shop	Light Gray

A7. Based on reports of restaurants closing in 2017 and 2018, the following list of number of failed restaurants by Neighborhood was produced.

Table 5- Count of Restaurant Failures by Neighborhood 2017-2018

Neighborhood	Restaurant Failures
Yorkville	21
Chinatown	18
Little Italy	17
Alexandra Park	14

Neighborhood	Restaurant Failures
The Annex	8
The Danforth	7
The Beaches	7
Cabbagetown	7

Neighborhood	Restaurant Failures
Riverdale	6
Leslieville	6
Little Portugal	5
Roncesvalles	4
Fort York/Liberty Village	4
Church and Wellesley	4
The Junction	3
Ossington	3
Leaside	3
Hillcrest	3
High Park North	3
Dundas West	3
Bloor	3
Yonge and Eglinton	2
Yonge and Dundas	2
St. Clair West	2
Seaton Village	2
Parkdale	2
Old Toronto	2
New Toronto	2
Kingsview Village	2
Financial District	2
St. Lawrence Market	1
Seato Village	1
Rosedale	1
Riverside	1

Neighborhood	Restaurant Failures
Old Mill/Baby Point	1
Oakwood Village	1
North York City Centre	1
North York	1
Moss Park	1
Liberty Village	1
Koreatown	1
King West Village	1
Kensington	1
Harbor Village	1
Harbord Village	1
Gerrard	1
Garden District	1
Fashion District	1
Eglinton West	1
East York	1
Dundas and University	1
Distillery District/West Don Lands	1
Cottage	1
Cliffcrest	1
Church	1
Christie Pits	1
Broadview	1
Bloordale	1
Bloor West Village	1
Allenby	1

A8. The first four neighborhoods contain the only double-digit numbers of closings. Each of those areas featured the following venues:

Table 6- Most Common Venue Closings for Neighborhoods with Double Digit Restaurant Closings

Neighborhood	Most Common Closing Venue
Yorkville	Seafood Restaurant
Chinatown	Chinese
Little Italy	Italian Restaurant
Alexandra Park	Italian Restaurant

A review of the existing restaurants returned by Foursquare showed that each of these neighborhoods are popular venues for the restaurants closed.

A9. Graphically, the number of restaurant failures by neighborhood is:

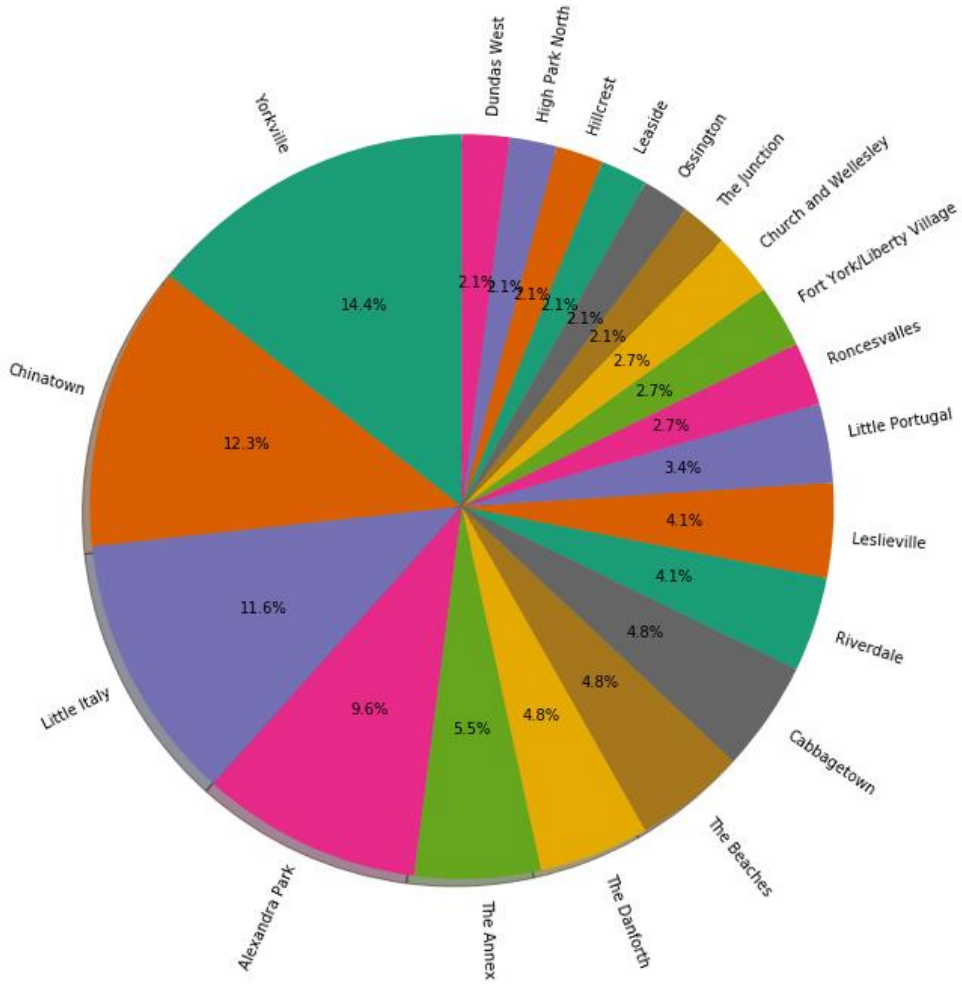


Figure 4- Restaurant Failures by Neighborhood 2017-2018

A10. The observed number of restaurant failures by venue is:

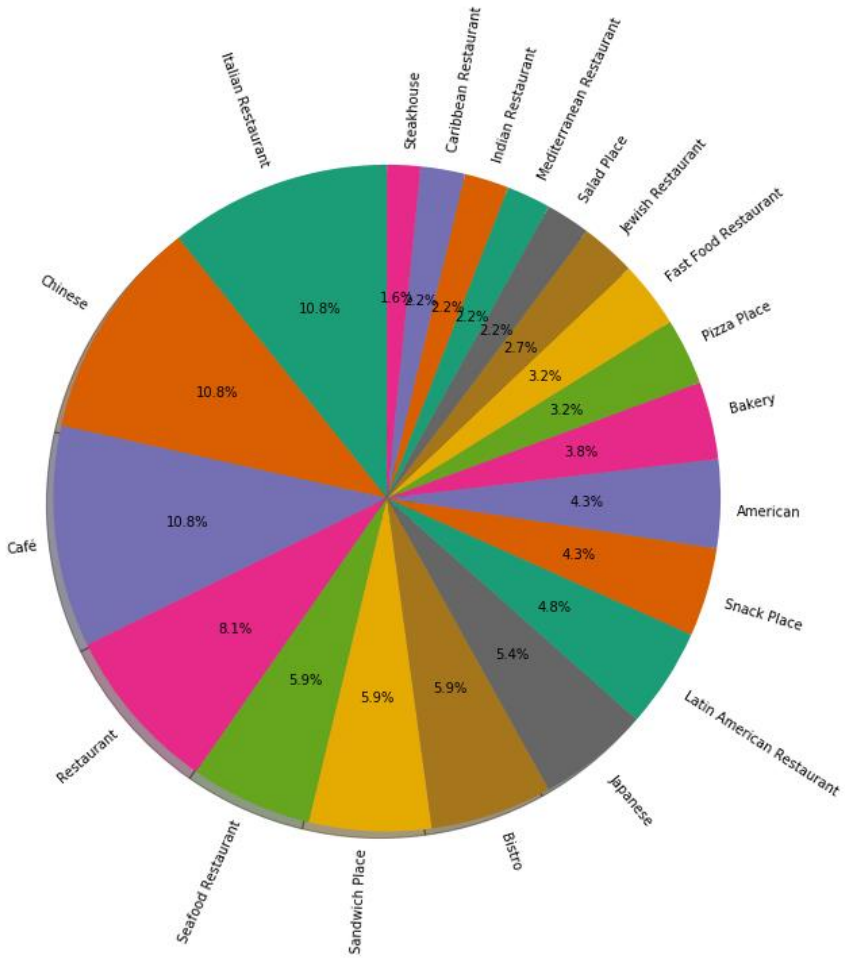


Figure 5- Restaurant Venue Failures in Toronto 2017-2018

Results

R1. The number, and diversity, of restaurant venues in Toronto is large, offering a number of potential opportunities for new restaurants.

R2. There were substantially more quick service venues (e.g., coffee shop, cafe', pizza shops) that full-service restaurants.

R3. Figure 2 shows the most popular restaurant venues, while Figure 5 shows the most frequent failures. This gives the most popular restaurant venue as Coffee Shop, with Japanese being a distant second.

R4. Several neighborhoods offered a large potential customer base, for very popular venue:

Table 7- Some Neighborhood with Potential for Restaurants and Large Customer Base

Neighborhood	Venue	Current #	Estimated Customer Base
Old East York	Sandwich Place	4	13,055
Woburn	Generic Restaurant	4	12,126
L'Amoreaux	Coffee Shop	4	11,465

R5. If neighborhood selection is driven by other considerations, `Table 4` shows the most popular venue.

R6. Only about 5% of restaurants closed in 2017, 3% in 2018. This is well below the expected where most new openings close in the first year, and 70% close in the next 3 to 5 years[1]. This is not an exact comparison, as no information was available about age of existing restaurants.

R7. `Figure 5` shows the most common venue failures in Toronto in 2017/2018. The most common failure venues should be avoided, especially in high-failure neighborhoods (`Figure 4`).

R8. The following neighborhoods had an unusually large number of failures/closings in 2017/2018: Yorkville, Chinatown, Little Italy, Alexandra Park. Each of those neighborhoods had an established preferred venue (e.g., Yorkville, seafood; Chinatown, Chinese; Little Italy and Alexandra Park, Italian). This confirms the difficulty of successfully opening a specific venue in an area already known for that venue. The exact reason for this was not studied, but could include reduced market share of customers, competition with proven performers, and customer habitation.

Discussion

Decision factors in starting a new restaurant, with selection of neighborhood and venue is a very complex, and risky, venture. Many of the decision factors are covered in this report. Many additional factors, such as chef, experience in a particular venue, financial backing, and experience in starting other restaurants must also be considered.

Toronto Canada numbers reflect a strong base of customers. The number of successful restaurants which have a unique venue (e.g., Hungarian Restaurants) shows a potential opening for niche restaurants. However, many popular venues are grouped into neighborhoods (e.g., Little Italy, Koreatown, Chinatown), where competition all but prohibits a new, competing, venue.

Incomplete data, and limited information on certain venue types, increases the risk inherit in recommendations. The number of neighborhoods, restaurants, and venues is smaller than those typically used in data science.

Conclusion

While the results determined in this report are consistent with common sense, these results are based on reduction of a significant amount of real-world observations. This consistent agreement

[1] "The Average Life Span of a Restaurant", https://yourbusiness.azcentral.com/average-life-span-restaurant-6024.html

may help the stakeholders accept the results, but should not minimize the importance that rigor gives.

Caveat Emptor

The author of this report is not a restaurateur. They have no knowledge about running a business, such as a restaurant. They do not have any business degree.

In an actual data science effort, the data scientist would work closely with the subject matter expert (SME) to determine what areas needed to be explored, and what questions need to be answered.

As such, the author has attempted on-line research (however briefly) in an attempt to come up with reasonable areas of research/questions.

This work, as it stands, should not be used in justifying decisions about starting a restaurant, either in Toronto, or anywhere else.

[i] "How to Start a Successful Restaurant", https://articles.bplans.com/start-successful-restaurant-guide/