TABLE OF CONTENTS

1. Introduction

1.1 Background

Toronto is the fifth-largest city in all of North America and the most populous city in Canada, with a population of over 2 million people, as of 2018 [1]. It is known for its extremely diverse population, with over 250 ethnicities and 180 different languages represented [3]. It is the home to many popular attractions; for example, the CN Tower, Casa Loma — the only real castle in North America. It consists of over 1,500 parks, 150 outdoor skating rinks and is the host of the largest single-day parade in North America.

Considering all these elements, it is no wonder that Toronto attracts about 43 million visitors annually [1]. As a result of all these factors, Toronto is an ideal place to start up a new business, with over 800,000 successful businesses running already. One of the strongest business sectors in Toronto is the food-service industry, consisting of over 8,000 restaurants [3].

1.2 Problem

Individuals who might want to open a restaurant in Toronto might be reluctant to do so considering all the competition. This project will be a source for feature business owners to refer to when they are planning to open a restaurant. The key problem this project aims to solve is: which neighbourhood is the best for business owner to open a restaurant.

To provide the best solution for this problem, several factors will be taken into consideration. The first being which neighbourhoods of Toronto have the most restaurants. This will indicate to the business owner where the competition is high.

The second factor was out of the restaurants in each neighbourhood, which cultural cuisine do they belong to. This will help the business owner decide the best location for their restaurant. For example, if an individual is looking to open a Chinese restaurant, then it is a better idea for them to open it in a neighbourhood with few or no Chinese restaurants.

The third factor was: what amenities are in the neighbourhood? Neighbourhoods with existing commercial, educational or recreational venues more likely to attract customers, making them an ideal location.

The fourth factor was the population distribution in Toronto. Neighbourhoods with a higher population are more likely to attract more customers.

The fifth and final factor is the income distribution of the neighbourhoods. This factor helps the restaurant owner choose the appropriate neighbourhood that fits with their price range. Furthermore, if a restaurant owner has already chosen a neighbourhood of their liking, they can change their price range to reflect the average income of the residents.

After considering all these factors, the business owner should have a good idea of which neighbourhood is the most ideal for their restaurant.

1.3 Audience

The key audience for this report is anyone who would like to open a new restaurant in Toronto and would like to know where the ideal spot for their new restaurant is.

2. Relevant Data

2.1 Data Sources

The list of neighbourhoods in Toronto can be found from the Wikipedia webpage: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. This page was scraped and the information was placed into a pandas data frame as shown in **Figure 1**.

	Postcode	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned\n
1	M2A	Not assigned	Not assigned\n
2	M3A	North York	Parkwoods\n
3	M4A	North York	Victoria Village\n
4	M5A	Downtown Toronto	Harbourfront\n

Figure 1: Data frame scraped from Wikipedia Page

The population and income distribution were obtained from the Toronto Open Data Portal website: https://open.toronto.ca/dataset/neighbourhood-profiles/ and the data were downloaded as a CSV file. This link provides the data set corresponding to the 2016 census.

The JSON file used to generate the maps was obtained from http://adamw523.com/toronto-geojson/.

The Foursquare API was used to acquire the rest of the data. The Foursquare API is a location-based service, which provides developers with data on locations around the world. Available data includes information about venues, users, photos and much more.

For this project, the API is used to find out which neighbourhood has the most restaurants. This is done by running a search of all the restaurants in Toronto and getting information about where they are located. Furthermore, the API was used to collect data about all the restaurants in each neighbourhood. The data was sorted by restaurant type, to see the prominent cuisines in each neighbourhood. Finally, the API was used to explore other venues in the neighbourhood. Information such as venue type and frequency of each venue was used to determine if the venues nearby are more likely to attract customers.

2.2 Data Cleaning

Data that was scraped from the Wikipedia page needed to be wrangled before it was ready to use. After storing the data into a Pandas data frame, several steps were taken to make the data cleaner. The first step was to remove any neighbourhoods which did not have an associated borough. The unwanted character (shown in **Figure 2**) in the 'Neighbourhood' column was also removed.

	Postcode	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Harbourfront
3	M6A	North York	Lawrence Heights
4	M6A	North York	Lawrence Manor

Figure 2: data frame with '\n' removed from the 'Neighbourhood' column

Then the data was grouped by the postal code. If there was a postal code with more than 1 neighbourhood, the rows were combined, with a comma separating the neighbourhood names, as seen in **Figure 3**.

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Malvern,Rouge
1	M1C	Scarborough	Port Union,Rouge Hill,Highland Creek
2	M1E	Scarborough	Guildwood,Morningside,West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

Figure 3: Data frame which is grouped by postal code

If there was a cell with a borough name, but no neighbourhood name, the neighbourhood was assigned to be the same as the borough name. The final step was downloading the longitude and latitude data for each neighbourhood and appending it to the existing data frame, producing the data frame in **Figure 4**.

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Malvern,Rouge	43.806686	-79.194353
1	M1C	Scarborough	Port Union,Rouge Hill,Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

Figure 4: Final data frame with longitude and latitude coordinates

3. Methodology

3.1 Initial Data Visualization

To get an idea of what the original data looked like, each neighbourhood was plotted on a map. Each blue marker, shown in **Figure 5**, represents a neighbourhood of Toronto.

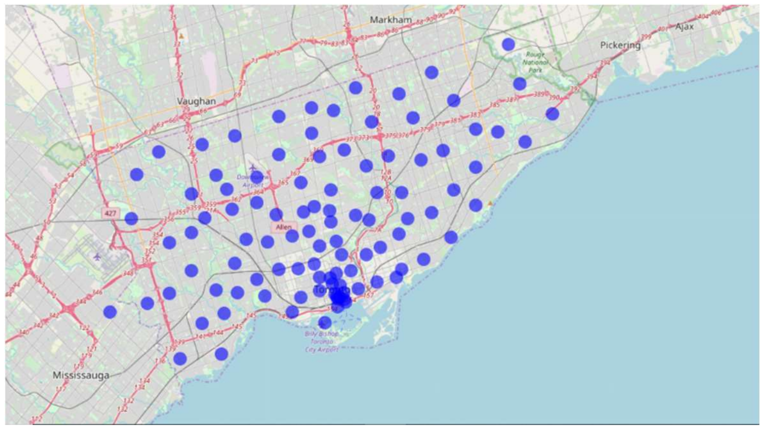

Figure 5: Neighbourhoods of Toronto

3.2 Neighbourhoods with Most Restaurants

Next, the neighbourhoods with the most number of restaurants were identified. This is valuable information because the more restaurants a neighbourhood has, the more competition there is. To obtain this information the Foursquare API was used to search for all venues with the category Id 4d4b7105d754a06374d81259, which is the ID corresponding to the food category. The information about the restaurants and their corresponding neighbourhood were placed in a Pandas data frame as shown in **Figure 6**.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	restaurant	restaurant Latitude	restaurant Longitude	restaurant Category
0	Malvern,Rouge	43.806686	-79.194353	Wendy's	43.802008	-79.198080	Fast Food Restaurant
1	Malvern,Rouge	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
2	Malvern,Rouge	43.806686	-79.194353	Caribbean Wave	43.798558	-79.195777	Caribbean Restaurant
3	Malvern,Rouge	43.806686	-79.194353	Harvey's	43.800106	-79.198258	Fast Food Restaurant
4	Malvern,Rouge	43.806686	-79.194353	Mr Jerk	43.801262	-79.199758	African Restaurant

Figure 6: Restaurants in Toronto and their Corresponding Neighbourhoods

Then the above data frame was stored based on the neighbourhoods to find out which neighbourhood had the highest frequency of restaurants. The data frame was then sorted, in descending order, and a new data frame, containing only the top 20 entries, was made. This data frame was used to plot the bar graph, as shown in **Figure 7**, which shows the top 20 neighbourhoods with the highest number of restaurants.

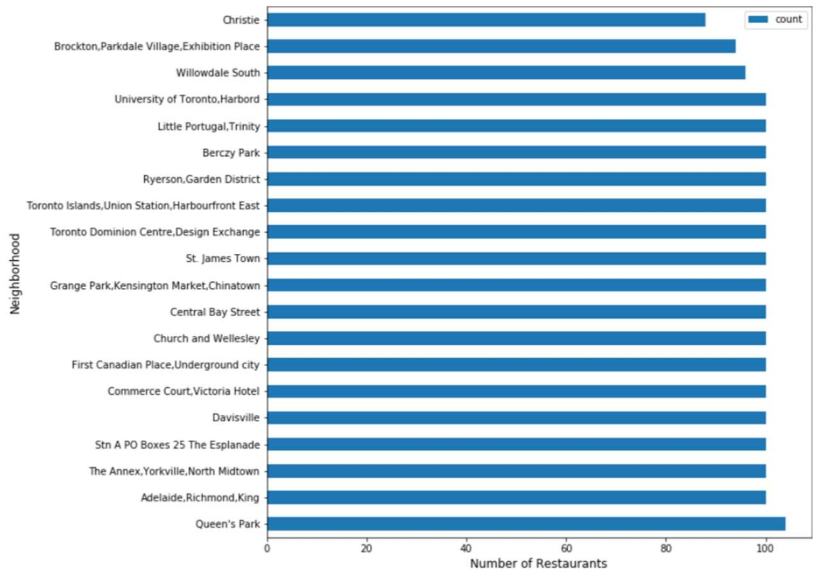

Figure 7: Neighborhoods with Most Restaurants in Toronto

3.3 Clustering Neighbourhoods Based on Restaurant Types

3.3.1 Data Preparation

Figuring out which type of restaurants already exist in the neighbourhood was the next step in the data analysis. With this information, a future restaurant owner can choose the best location for their restaurant. For example, let's say a person would like to open an Italian restaurant. They would, ideally, want to open this restaurant in an area with no or very few existing Italian restaurants. So this information is crucial for determining an ideal location.

To make the decision easier for the restaurant owner, the neighbourhoods were grouped based on the types of restaurants. This was done by applying the machine learning technique called clustering. Clustering is a technique that involves grouping a set of data points, based on similar features. The algorithm which is used in this project is the k-means method. The algorithm sorts the data points into *k* groups. To begin the grouping, *k* centroids are chosen, one for each group. Then the distance between each data point and the centroids is calculated and the point is classified to the group who it is closest to. Based on the points in each group, the center is recalculated. The process is repeated until the centroids do not move as much.

So, to apply the k-means algorithm the data set has to be transformed. The first step in the transformation was applying one-hot encoding. One hot encoding is the process of assigning a binary value (0 or 1) to represent categorical variables. This data set is shown in **Figure 8**.

	Afghan Restaurant	African Restaurant	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Belgian Restaurant	...	Tex-Mex Restaurant	Thai Restaurant	Theme Restaurant	Tibetan Restaurant	Turkish Restaurant	Udon Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Wings Joint	Neighborhood
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	Malvern,Rouge
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	Malvern,Rouge
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	Malvern,Rouge
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	Malvern,Rouge
4	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	Malvern,Rouge

Figure 8: Data set after one-hot encoding has been applied

As you can see in the above figure, each categorical value has a value of 0, unless the row is of that category. For example, the 4th entry is an African restaurant, so the "African Restaurant" column has a value of 1, while the rest have a value of 0.

Next, the data frame was sorted by the "Neighbourhood" column and the mean frequency for the occurrence of each category was calculated. This is shown in **Figure 9**.

	Neighborhood	Afghan Restaurant	African Restaurant	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	...	Tapas Restaurant	Tex-Mex Restaurant	Thai Restaurant	Theme Restaurant	Tibetan Restaurant	Turkish Restaurant	Udon Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Wings Joint
0	Adelaide,Richmond,King	0.0	0.0	0.04000	0.0	0.0	0.030000	0.00000	0.00	0.030000	...	0.0	0.0	0.03	0.0	0.0	0.0	0.0	0.03	0.00000	0.00
1	Agincourt	0.0	0.0	0.02381	0.0	0.0	0.047619	0.02381	0.00	0.047619	...	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.00	0.02381	0.00
2	Alderwood,Long Branch	0.0	0.0	0.00000	0.0	0.0	0.000000	0.00000	0.00	0.000000	...	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.00	0.00000	0.00
3	Bayview Village	0.0	0.0	0.00000	0.0	0.0	0.000000	0.00000	0.00	0.000000	...	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.00	0.00000	0.00
4	Bedford Park,Lawrence Manor East	0.0	0.0	0.04000	0.0	0.0	0.000000	0.00000	0.04	0.040000	...	0.0	0.0	0.08	0.0	0.0	0.0	0.0	0.00	0.00000	0.04

Figure 9: Data frame sorted by Neighbourhood column, with mean frequency

3.3.2 Applying Algorithm

Now the algorithm can be applied. For this project, the value for *k* was chosen to be 15. This is so the individual who wishes to open the restaurant has enough groups to choose from. However, if the value was too high, let's say 70, it might be overwhelming to choose a cluster and there might not be enough neighbourhoods in each cluster, therefore the number 15 was chosen. After the clustering was done the results were shown on a map. The map can be seen in **Figure 10**.

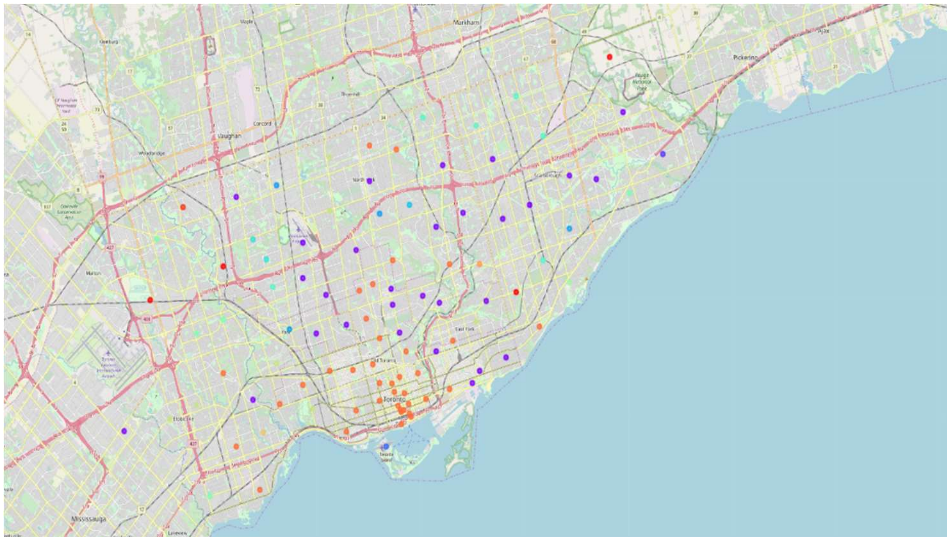

Figure 10: Map depicting the clusters of Toronto, based on the restaurant type

Each cluster can be examined to see what the most common types of restaurants are. For example in **Figure 11**, it can be seen that in cluster 1, “Pizza Place” is very common.

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Restaurant Type	2nd Most Common Restaurant Type	3rd Most Common Restaurant Type	4th Most Common Resturant Type	5th Most Common Resturant Type
0	M1B	Scarborough	Malvern,Rouge	43.806686	-79.194353	1	Fast Food Restaurant	Greek Restaurant	African Restaurant	Caribbean Restaurant	Chinese Restaurant
3	M1G	Scarborough	Woburn	43.770992	-79.216917	1	Fish & Chips Shop	Indian Restaurant	Pizza Place	Restaurant	Chinese Restaurant
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	1	Bakery	Indian Restaurant	Chinese Restaurant	Wings Joint	Caribbean Restaurant
10	M1P	Scarborough	Dorset Park,Wexford Heights,Scarborough Town C...	43.757410	-79.273304	1	Fast Food Restaurant	Asian Restaurant	Bakery	Chinese Restaurant	Indian Restaurant
11	M1R	Scarborough	Maryvale,Wexford	43.750072	-79.295849	1	Middle Eastern Restaurant	Pizza Place	Burger Joint	Bakery	Seafood Restaurant
13	M1T	Scarborough	Clarks Corners,Tam O'Shanter,Sullivan	43.781638	-79.304302	1	Pizza Place	Sandwich Place	Fast Food Restaurant	Deli / Bodega	Taiwanese Restaurant
18	M2J	North York	Fairview,Henry Farm,Oriole	43.778517	-79.346556	1	Fast Food Restaurant	Sandwich Place	Food Court	Pizza Place	Asian Restaurant
22	M2N	North York	Willowdale South	43.770120	-79.408493	1	Pizza Place	Korean Restaurant	Japanese Restaurant	Café	Sushi Restaurant
25	M3A	North York	Parkwoods	43.753259	-79.329656	1	Fish & Chips Shop	Pizza Place	Chinese Restaurant	Caribbean Restaurant	Café
26	M3B	North York	Don Mills North	43.745906	-79.352188	1	Japanese Restaurant	Pizza Place	Restaurant	Burger Joint	Fast Food Restaurant
30	M3K	North York	CFB Toronto,Downsview East	43.737473	-79.464763	1	Turkish Restaurant	Italian Restaurant	Café	Pizza Place	Latin American Restaurant
33	M3N	North York	Downsview Northwest	43.761631	-79.520999	1	Pizza Place	Chinese Restaurant	Fast Food Restaurant	Snack Place	Caribbean Restaurant
35	M4B	East York	Woodbine Gardens,Parkview Hill	43.706397	-79.309937	1	Pizza Place	Bakery	Fast Food Restaurant	Breakfast Spot	Gastropub
37	M4E	East Toronto	The Beaches	43.676357	-79.293031	1	Pizza Place	Bakery	Japanese Restaurant	Breakfast Spot	Sandwich Place
38	M4G	East York	Leaside	43.709060	-79.363452	1	Pizza Place	Sandwich Place	Sushi Restaurant	Burger Joint	Restaurant
39	M4H	East York	Thorncliffe Park	43.705369	-79.349372	1	Indian Restaurant	Burger Joint	Pizza Place	Sandwich Place	Fast Food Restaurant
41	M4K	East Toronto	Riverdale,The Danforth West	43.679557	-79.352188	1	Greek Restaurant	Pizza Place	Café	Restaurant	Sushi Restaurant
42	M4L	East Toronto	India Bazaar,The Beaches West	43.668999	-79.315572	1	Indian Restaurant	Café	Pizza Place	Sandwich Place	Restaurant
45	M4P	Central Toronto	Davisville North	43.712751	-79.390197	1	Pizza Place	Italian Restaurant	Fast Food Restaurant	Café	Sushi Restaurant
47	M4S	Central Toronto	Davisville	43.704324	-79.388790	1	Pizza Place	Italian Restaurant	Fast Food Restaurant	Sushi Restaurant	Sandwich Place

Figure 11: Cluster 1 with its corresponding neighbourhoods

Doing a similar analysis in cluster 7, we can see the most common type of restaurant is a Chinese restaurant, as shown in **Figure 12**.

Cluster Labels	1st Most Common Restaurant Type	2nd Most Common Restaurant Type	3rd Most Common Restaurant Type	4th Most Common Resturant Type	5th Most Common Resturant Type
7	Chinese Restaurant	Pizza Place	Asian Restaurant	Sandwich Place	Fast Food Restaurant
7	Chinese Restaurant	Asian Restaurant	Sandwich Place	Restaurant	Caribbean Restaurant
7	Chinese Restaurant	Pizza Place	Bakery	Noodle House	Fast Food Restaurant
7	Chinese Restaurant	Bakery	Fast Food Restaurant	Indian Restaurant	Noodle House
7	Pizza Place	Diner	Sandwich Place	Chinese Restaurant	Bakery
7	Pizza Place	Sandwich Place	Chinese Restaurant	Bakery	Mediterranean Restaurant

Figure 12: Cluster 7

A similar analysis can be done for each cluster but is omitted to ensure the report is not too long.

3.4 Exploring Other Venues

In this section, we will explore other venues in the neighbourhoods. We will explore venues under the following categories: “Entertainment”, “Colleges/Universities”, “Outdoors and Recreations” and “Professional and Other Places”. If any venues of these categories are present more customers will be attracted to the restaurant. Therefore if they are present in high volumes, the restaurant has a better chance at business. For each category, we will find the top 20 neighbourhoods, with the highest count for that category.

3.4.1 Entertainment Venues

To find the neighbourhoods with the most venues in these categories, the Foursquare API was used to fetch all the venues with the category ID 4d4b7104d754a06370d81259. Then the results were placed in a data frame. Which was then grouped by neighbourhood and ordered in descending order.
The results were then plotted in a bar graph, as shown in **Figure 13**.

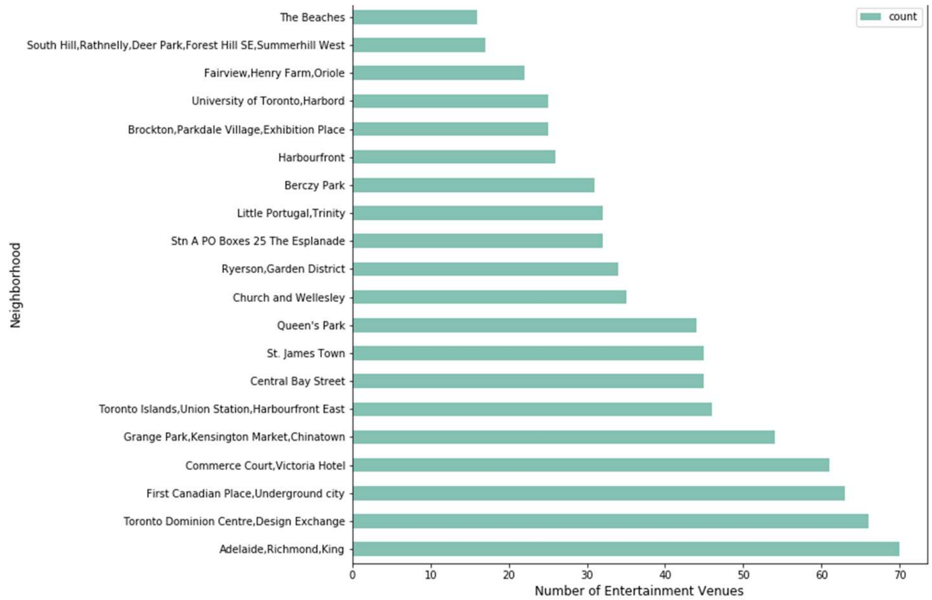

Figure 13: Neighbourhoods with most Entertainment Venues in Toronto

3.4.2 College/University Venues

Next, it was time to determine which neighbourhoods had the most colleges/universities nearby. The students of these institutions will be attracted to the new restaurant if it is close by and this increases the number of customers. The business owner can also use this information to decide what price range they would like to keep their restaurant in, as students are more likely to be attracted to lower prices.

The same steps as in **Section 3.4.1** were taken to obtain the neighbourhoods with the most college/university institutions. However, this time the category ID 4d4b7104d754a06370d81259. The graph in **Figure 14** depicts this information.

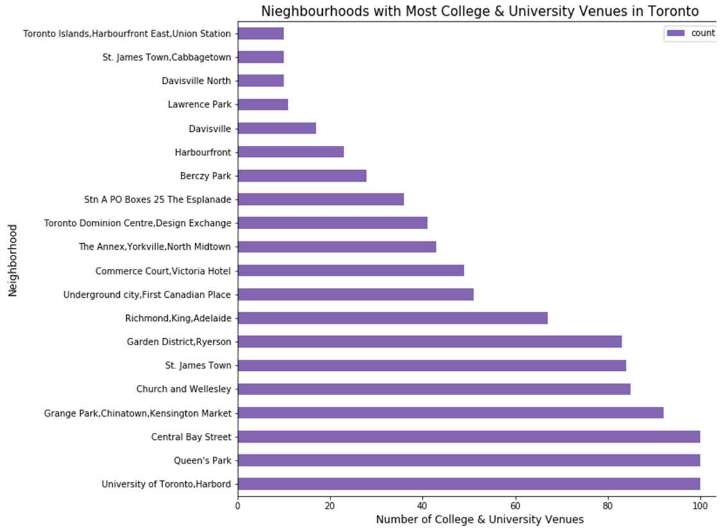

Figure 14: Neighbourhoods with the most College/University Institutions

3.4.3 Outdoors and Recreation Venues

Now the neighbourhoods with the most outdoor and recreational venues were identified. The presence of these venues will attract their users to the restaurant, therefore increasing the chance of business.

The same steps as in **Section 3.4.1** were taken to obtain the neighbourhoods. However, this time the category ID 4d4b7105d754a06377d81259. The graph in **Figure 14** depicts this information.

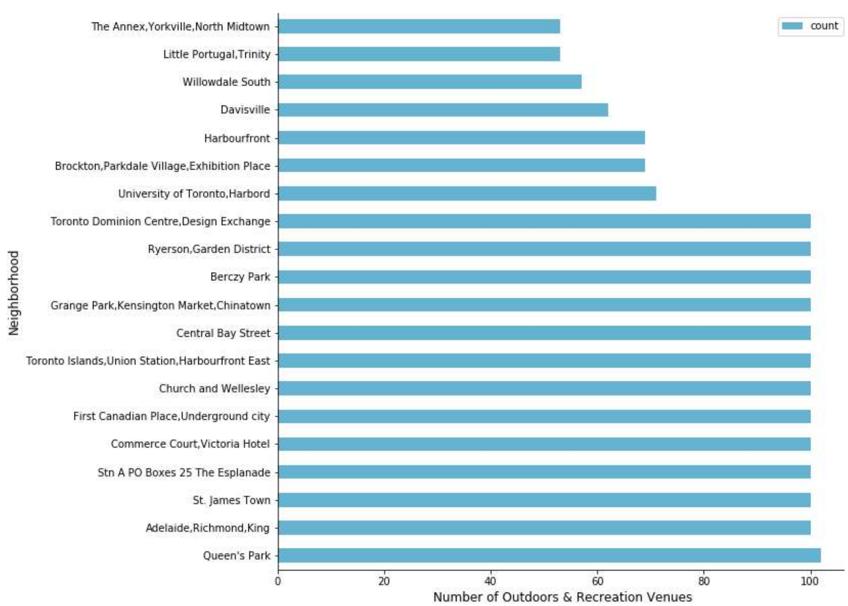

Figure 15: Neighbourhoods with the most Outdoor/Recreation Venues

3.4.4 Professional and Other Places

Finally, the neighbourhoods with the most professional/other places were determined. The ID number 4d4b7105d754a06375d81259 was used to do this.

Figure 16: Neighbourhoods with the most Professional and Other Venues

3.5 Exploring Population/Income

3.5.1 Data Preparation

To explore the population and income data, the data which was downloaded from the open data portal needed to be cleaned. The original data set can be seen in **Figure 17**. This is the data set from the 2016 census, for Toronto.

	_id	Category	Topic	Data Source	Characteristic	City of Toronto	Agincourt North	Agincourt South-Malvern West	Alderwood	Annex	...	Willowdale West	Willowridge-Martingrove-Richview	Woburn	Woodbine Corridor	Woodbine-Lumsden	Wychwood	Yonge-Eglinton	Yonge-St.Clair	York University Heights	Yorkdale-Glen Park
0	1	Neighbourhood Information	Neighbourhood Information	City of Toronto	Neighbourhood Number	NaN	129	128	20	95	...	37	7	137	64	60	94	100	97	27	31
1	2	Neighbourhood Information	Neighbourhood Information	City of Toronto	TSNS2020 Designation	NaN	No Designation	No Designation	No Designation	No Designation	...	No Designation	No Designation	NIA	No Designation	No Designation	No Designation	No Designation	No Designation	NIA	Emerging Neighbourhood
2	3	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2016	2,731,571	29,113	23,757	12,054	30,526	...	16,936	22,156	53,485	12,541	7,865	14,349	11,817	12,528	27,593	14,804
3	4	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2011	2,615,060	30,279	21,988	11,904	29,177	...	15,004	21,343	53,350	11,703	7,826	13,986	10,578	11,652	27,713	14,687
4	5	Population	Population and dwellings	Census Profile 98-316-X2016001	Population Change 2011-2016	4.50%	-3.90%	8.00%	1.30%	4.60%	...	12.90%	3.80%	0.30%	7.20%	0.50%	2.60%	11.70%	7.50%	-0.40%	0.80%

Figure 17: Data set from the 2016 Census

The first step to clean the data was to remove the unwanted columns. The next step was to take the transpose, so the neighbourhood names would come up in the rows, instead of the columns. After this, a few columns were renamed so they represented the values they displayed. Finally, the columns corresponding to the total population and total income were changed from type object to type float. The final data set, shown in Figure 18, was used to complete the analysis.

Characteristic	Neighbourhood	Neighbourhood Number	TSNS2020 Designation	Population, 2016	Population, 2011	Population Change 2011-2016	Total private dwellings	Private dwellings occupied by usual residents	Population density per square kilometre	Land area in square kilometres	...	External migrants	Total - Mobility status 5 years ago - 25% sample data	Non-movers	Movers	Non-migrants	Migrants	Internal migrants	Intraprovincial migrants	Interprovincial migrants	External migrants
0	Agincourt North	129	No Designation	29113.0	30,279	-3.90%	9,371	9,120	3,929	7.41	...	605	27,490	18,865	8,610	5,445	3,170	880	735	135	2,280
1	Agincourt South-Malvern West	128	No Designation	23757.0	21,988	8.00%	8,535	8,136	3,034	7.83	...	490	22,325	13,565	8,775	5,610	3,145	980	760	220	2,170
2	Alderwood	20	No Designation	12054.0	11,904	1.30%	4,732	4,616	2,435	4.95	...	70	11,370	8,235	3,130	2,200	925	680	615	70	245
3	Annex	95	No Designation	30526.0	29,177	4.60%	18,109	15,934	10,863	2.81	...	835	27,715	12,980	14,735	8,340	6,390	3,930	2,630	1,310	2,460
4	Banbury-Don Mills	42	No Designation	27695.0	26,918	2.90%	12,473	12,124	2,775	9.98	...	380	25,925	16,300	9,625	6,480	3,140	1,405	1,190	220	1,735

Figure 16: Final data set after data cleaning

3.5.2 Population Analysis

The population intensity for each neighbourhood was mapped using a choropleth map, as shown in **Figure 17**. The intensity of the colour varies with the population, with a darker colour representing a higher population.

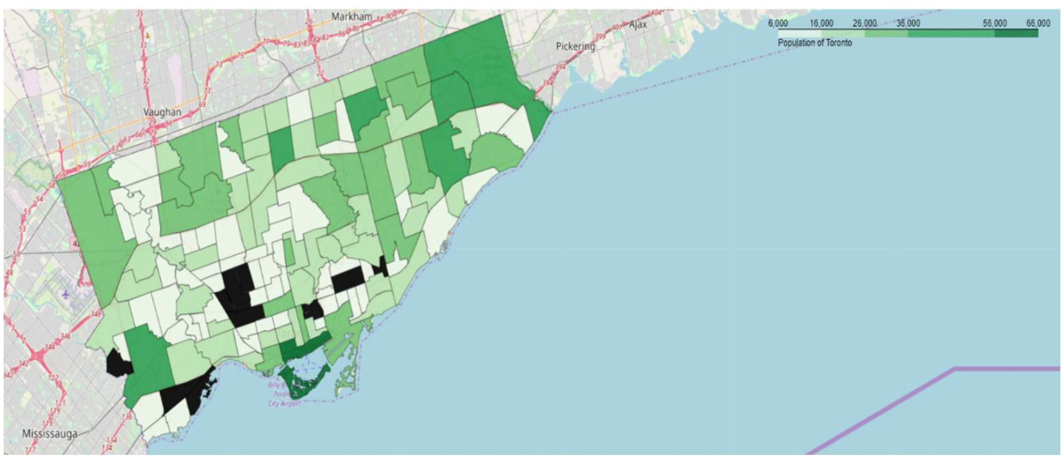

Figure 17: Population Intensity in Toronto

3.5.3 Income Analysis

The income intensity for each neighbourhood was also mapped using a choropleth map. The map is shown in **Figure 18**.

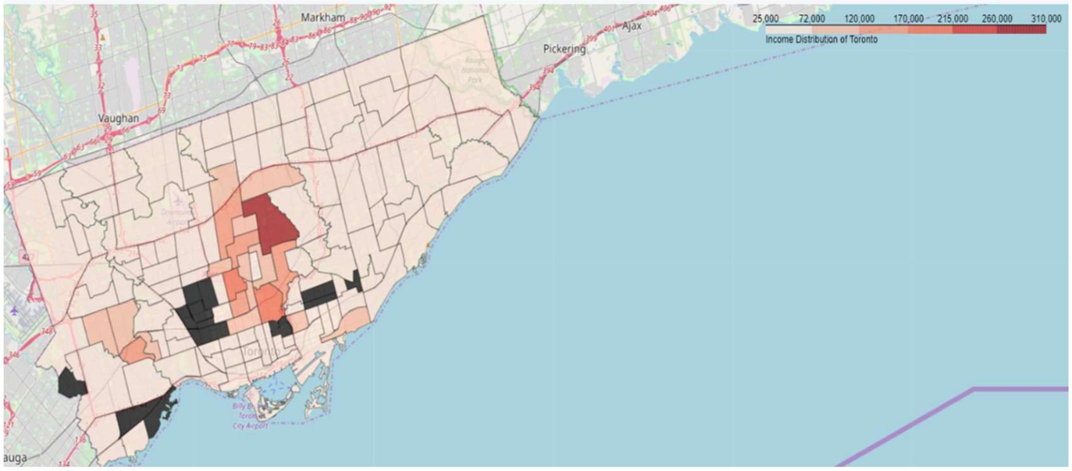

Figure 18: Income Intensity in Toronto

4. Results

The first section discussed which neighbourhoods had the most restaurants. This would help factor out a few neighbourhoods, especially the ones with too many existing restaurants. We see that neighbourhoods such as Queens Park, Richmond, and King all have a high number of restaurants.

The second part of the study focuses on the type of restaurant the individual wishes to open. By clustering the neighbourhoods based on their top 5 restaurant types, the neighbourhoods were divided into 15 different sections. Each of these 15 sections had similar characteristics when it came to their restaurant types. As observed in **Figure 12**, cluster 7 includes neighbourhoods who have a lot of existing Chinese restaurants. Say a person wants to open a new Chinese restaurant, then they would benefit from opening it in a neighbourhood which is not a part of cluster 7.

The third factor analyzed was the other venues in the neighbourhood. This information would give insight on which neighbourhoods will attract more customers. Neighbourhoods with any of the venues discussed in this part are more likely to attract customers. For example, as shown in **Figure 14** the University of Toronto neighbourhood has a high number of college/university venues. So the students of these places can be considered potential customers.

The final section focused on the population size and average income of the households. As you can see in **Figure 17** the main population resides in the downtown core. With this information, the restaurant owner can determine which areas are more likely to attract customers. **Figure 18** shows that most neighbourhoods have households whose incomes are between 25,000 and 75,000. This is valuable information for the new restaurant owner because it can help them determine the price range for their restaurants.

5. Discussion

After carefully considering all of the information provided in this report, a new restaurant owner can confidently choose the best location for their restaurant. This information can also help personalize the restaurant. For example, if an individual chooses to open a restaurant that is in a neighbourhood with a high number of college/university venues, they can choose to offer special promotions for the students to attract them to the restaurant.

The analysis in this study can be enhanced and personalized more in the future. Some enhancements include taking into consideration the average rent for each neighbourhood. This information can help the restaurant owner determine which area is within their budget. To personalize the study, we can take into consideration the type of restaurant the person is opening. This study was a general study meant as a resource for anyone. However, if a particular client needed information for their restaurant, the study can be replicated using specific information provided by the client. For example, using the cultural cuisine of their restaurant we can determine the neighbourhoods which have the most restaurants with such cuisine.

6. Conclusion

In conclusion, this study used data analysis, data visualization and machine learning to determine which neighbourhood would be the best one to open any given type of restaurant. An individual may use this information to find out which neighbourhood in Toronto is the best to open a restaurant in. This information is useful for not only individuals who are opening a restaurant, but also for consulting companies, operating in Toronto, who provide advice to such people.

7. Resources:

[1] City of Toronto. (2019, May 8). Toronto at a Glance. Retrieved from https://www.toronto.ca/city-government/data-research-maps/toronto-at-a-glance/

[2] Curiocity, Curiocity Toronto Staff, & Curiocity Toronto Staff. (2019, August 16). 20 fun facts about Toronto that prove it's the coolest Canadian city. Retrieved from https://curiocity.com/toronto/lifestyle/20-fun-facts-about-toronto-that-prove-its-the-coolest-canadian-city/

[3] Ryan, A. (n.d.). How multicultural is Toronto? Let us count the ways... Retrieved from https://torontoglobal.ca/TG-Blog/March-2019/How-multicultural-is-Toronto-Let-us-count-the-way

[4] Toronto Global - Toronto region quick facts. (n.d.). Retrieved from https://torontoglobal.ca/Discover-Toronto-region/Toronto-region-quick-facts