Predicting Prime Italian Restaurant Locations

Ryan Hull

March 1, 2020

1. Introduction

1.1. Background

According to Lorri Mealey from The Balance Small Business, a bad location is one of the biggest, if not *the biggest*, reason a restaurant fails. Reasons can include poor visibility, poor parking, and low foot traffic. These problems make it nearly impossible for a restaurant to earn a profit. To mitigate this problem, it's all about location, location, location.

1.2. Problem

How can we pick a location that is popular and would maximize foot traffic? Can we compare that information to existing restaurants in those locations to aid in informing our decision-making process when choosing a location for a new restaurant?

1.3. Interest

An entrepreneur has expressed an interest in opening an Italian restaurant in Toronto, CA. They want to know where they could place the restaurant and maximize customers by minimizing the competition. They are specifically interested in popular locations with high traffic volume that don't have an established Italian restaurant.

2. Data Acquisition and Selection

2.1. Data Sources

A list of Toronto, CA Postal Codes can be found [here](#). Latitude and Longitude data can be obtained from Geopy Library. Finally, venue data (including location) can be pulled using a Foursquare API.

2.2. Data Cleaning

Data was put into json format and cleaned on venue categories. This allowed us to properly categorize the information retrieved from Foursquare prior to further exploration and visualization.

2.3. Feature Selection

After cleaning the data, it was important to find out how many different venues there were in the surrounding area of Toronto, CA. The resulting dataframe included 2,239 venues and 7 features which is visible in Table 1.

Table 1 – Toronto Venues from Foursquare

```
print(toronto_venues.shape)
toronto_venues.head()
```

(2239, 7)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Downtown Toronto	43.650571	-79.384568	Four Seasons Centre for the Performing Arts	43.650592	-79.385806	Concert Hall
1	Downtown Toronto	43.650571	-79.384568	Rosalinda	43.650252	-79.385156	Vegetarian / Vegan Restaurant
2	Downtown Toronto	43.650571	-79.384568	The Keg Steakhouse + Bar - York Street	43.649987	-79.384103	Restaurant
3	Downtown Toronto	43.650571	-79.384568	Nathan Phillips Square	43.652270	-79.383516	Plaza
4	Downtown Toronto	43.650571	-79.384568	Shangri-La Toronto	43.649129	-79.386557	Hotel

3. Methodology and Exploratory Data Analysis

3.1. Toronto Venue Information

Once we collected the Foursquare venue data on Toronto, our next step pertained to grouping those venues by their respective neighborhoods for analysis. Those results can be seen as depicted in Table 2.

Table 2 – Toronto Venues Count by Neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Central Toronto	116	116	116	116	116	116
Downtown Toronto	1314	1314	1314	1314	1314	1314
East Toronto	126	126	126	126	126	126
East York	74	74	74	74	74	74
Etobicoke	70	70	70	70	70	70
Mississauga	13	13	13	13	13	13
North York	245	245	245	245	245	245
Scarborough	88	88	88	88	88	88
West Toronto	172	172	172	172	172	172
York	21	21	21	21	21	21

3.2. Toronto Italian Restaurant Information

Our entrepreneur customer wants to know where the Italian restaurants are in relation to this information. The next step in our process was to find the number of Italian restaurants and group them into neighborhoods, which can be seen in Table 3 and Table 4 respectively. Table 3 shows us that there are 49 Italian restaurants in the city of Toronto, CA according to Foursquare.

Table 3 – Toronto Italian Restaurants from Foursquare

```
toronto_italian = toronto_venues[toronto_venues['Venue Category'] == 'Italian Restaurant']
print(toronto_italian.shape)
toronto_italian.head()
```

(49, 7)

45]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
83	Downtown Toronto	43.650571	-79.384568	Mercatto	43.650243	-79.380820	Italian Restaurant
160	North York	43.733283	-79.419750	Francobollo	43.734557	-79.419549	Italian Restaurant
169	North York	43.733283	-79.419750	Il Fornaro	43.734073	-79.419870	Italian Restaurant
202	Downtown Toronto	43.644771	-79.373306	The Old Spaghetti Factory	43.646964	-79.374403	Italian Restaurant
245	West Toronto	43.636847	-79.428191	Caffino	43.639021	-79.425289	Italian Restaurant

Table 4 – Toronto Italian Restaurant Counts by Neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Central Toronto	2	2	2	2	2	2
Downtown Toronto	31	31	31	31	31	31
East Toronto	6	6	6	6	6	6
North York	3	3	3	3	3	3
Scarborough	1	1	1	1	1	1
West Toronto	6	6	6	6	6	6

4. Results

4.1. Clustering Popular Toronto Locations

Clustering our data into popular areas with the highest amounts of foot traffic results in Figure 1.
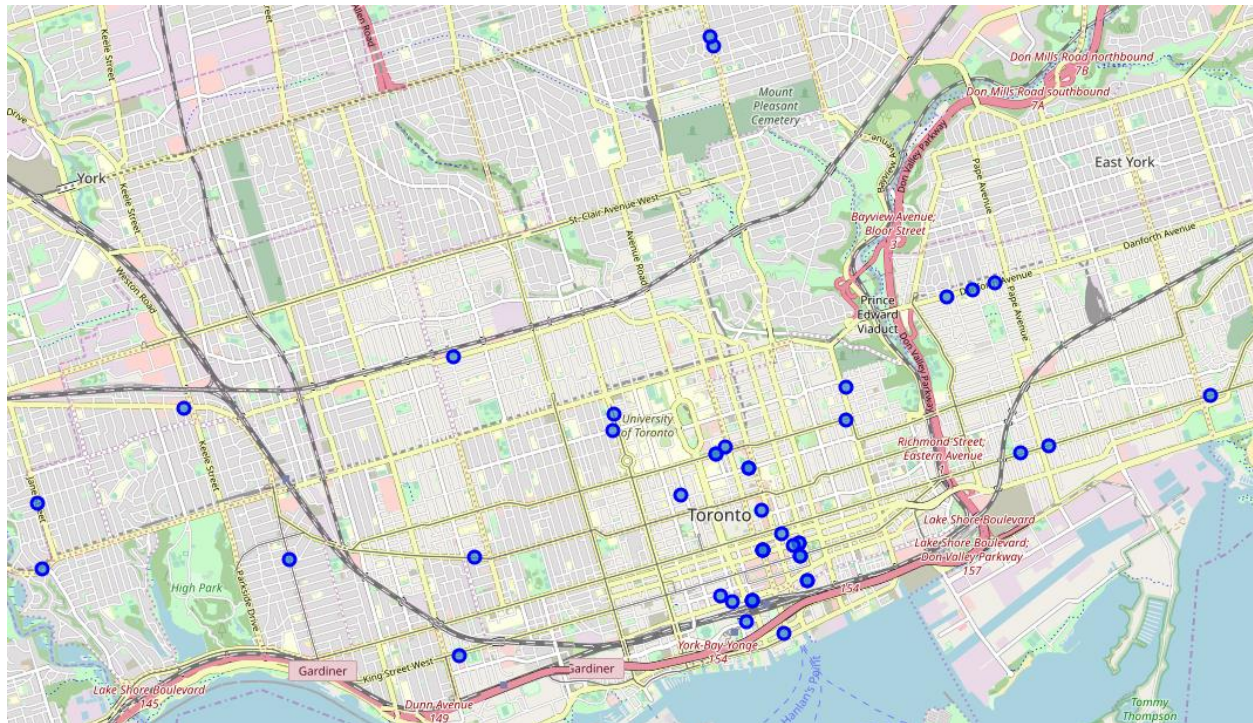
Figure 1 – Toronto Hot Spot Locations from Foursquare
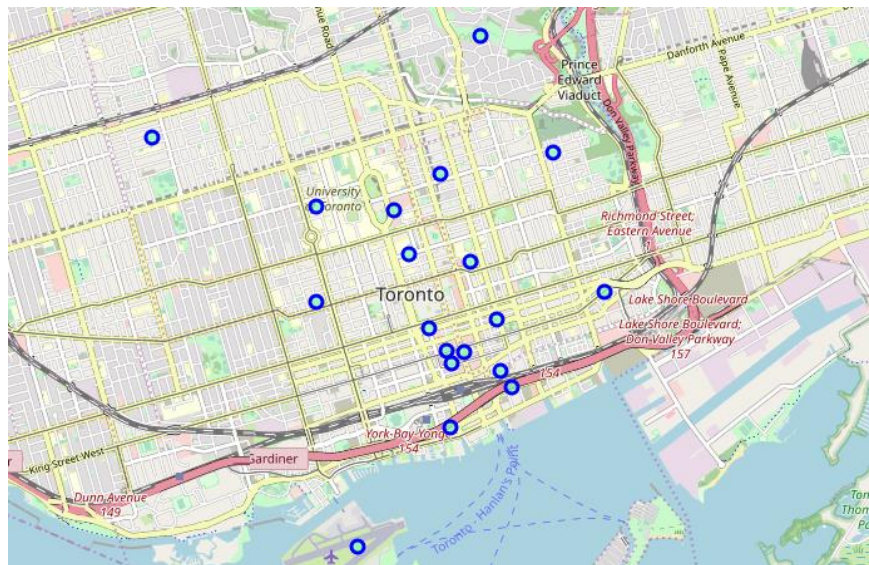
4.2. Mapping Toronto Italian Restaurant Locations

Finally, we mapped the locations of all 49 Italian restaurants located in Toronto (Figure 2). This information can be used to compare to the areas with the highest amount of foot traffic.

Figure 2 – Toronto Italian Restaurants from Foursquare

5. Discussion

By mapping out the popular areas in Toronto with high levels of foot traffic, we are able to mitigate the risk of selecting a poor location for our entrepreneurial customer's Italian restaurant. Combining that information with the location of existing Italian restaurants in Toronto will further improve the entrepreneur's chance of success.

6. Conclusion

The decision to open a business without doing your homework up front can prove fatal from the start. Our customer came to us with a problem. By using data, we were able to find a solution for our customer and use exploratory data analysis to exceed their needs.