

Battle of the Toronto Neighborhoods: The Mexican Restaurant

by Jose Calucag

(To view the code, please go to the following link [here](#))

I. Introduction/Business Problem

Introduction

As of 2016, Wikipedia states that Toronto is Canada's most populous city and the fourth most in North America. It's diverse, it's multicultural, it's a global hub of commerce and entertainment.. and its profile is always growing. To focus upon Toronto's restaurant industry, many have come and gone but there have been a rare and resilient few who have come to Toronto and actually thrive in the market. Which is why any hopes of a restaurant business to remain afloat and be successful needs to be analyzed and planned carefully. Even before any notion of looking at a piece of real estate or buying your first piece of cutlery, it is important for stakeholders to be able find all data possible and leverage it to make a sound business decision.

So, I hope that this project can illustrate how restaurant stakeholders strategize launching and breaking ground for a new restaurant in a metropolitan city like Toronto.

Business Problem

My client, a wealthy and successful restaurateur is eager to expand his business operations and brand in the city of Toronto. They're hoping to create an authentic Mexican eating experience rich in culture and cuisine. Since the city of Toronto is a very competitive market, my client needs insightful data in order to decide if it's good to establish this restaurant in the city and in which neighborhood.

Interested Audience

Albeit that this is a project about trying to break ground for a Mexican restaurant, I believe that the analyzed data can be relevant for any restaurateur trying to make it in the city of Toronto. As well, with the data science practices and techniques used throughout this project, this can be a useful example and resource for any data science practitioner.

II. Data

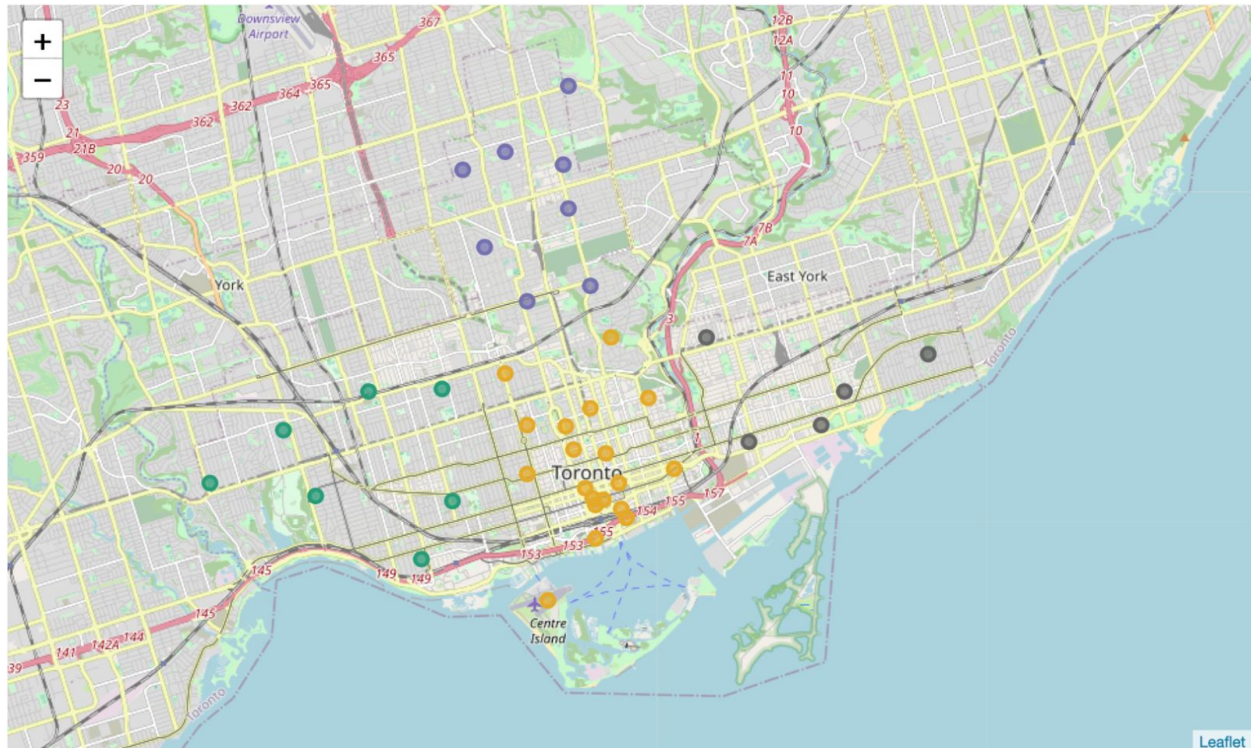
Data Sources

These are the data sources used for this report:

- From Wikipedia, I will pull postal code data of Toronto's neighborhoods. With this data, I will create data frameworks and geological mappings. Here's the link:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- I will need the latitude and longitude coordinates for each postal code, and that is pulled from this csv file: https://cocl.us/Geospatial_data
- The Foursquare API will be used to obtain location and restaurant data. The pulled data will provide categorical data that can give further analysis that will determine the location for the restaurant.

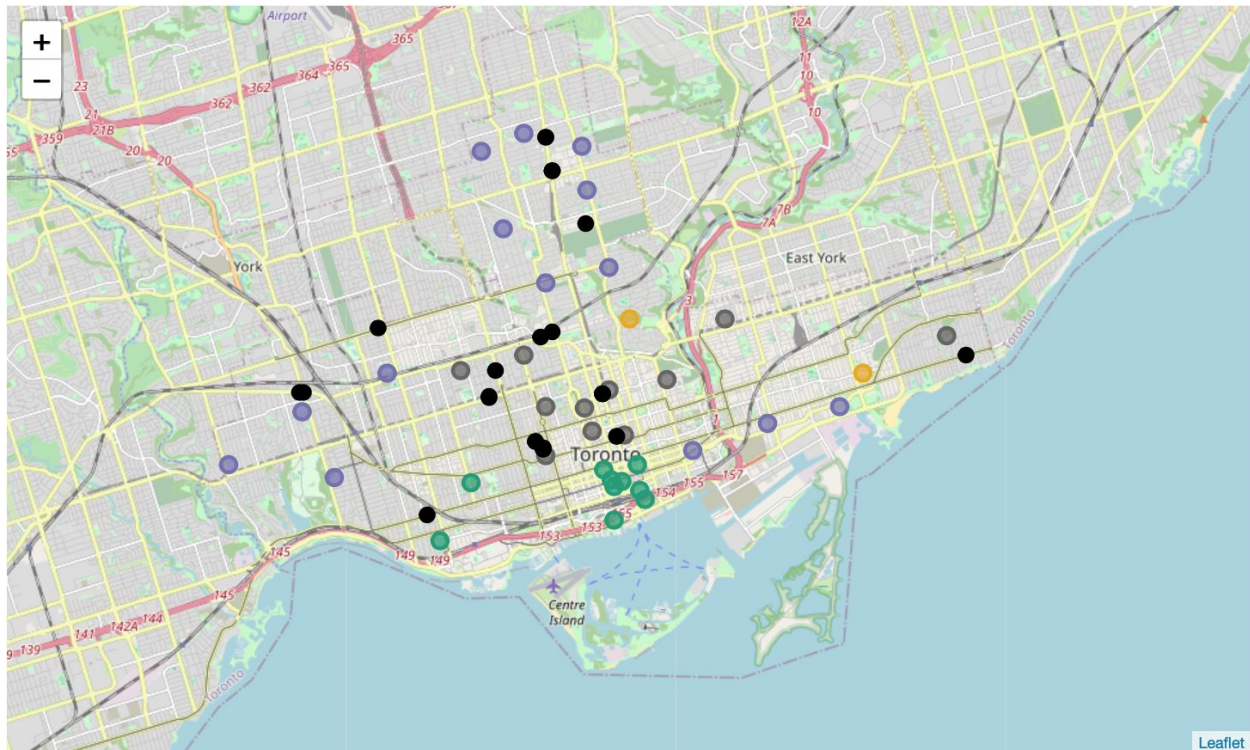
III. Methodology

Our main goal is to find an optimum location for a new Mexican restaurant in the city of Toronto. To do that, I needed to create a recommendation model based upon our data sources. So, to start, I pulled the data from the aforementioned geological sources, cleaned them of unusable data (i.e. 'Not Assigned' data points and non-'city of Toronto' boroughs) to create usable dataframes. For analysis, we used a kmeans clustering method to create clusters of the Toronto neighborhoods; to which I also used to create visualizations using the Folium package.



With these datasets, I used the Foursquare API to find restaurant data. Here, I was able to see the frequency of restaurant cuisines for each neighborhood and also locate nearby Mexican restaurants around the city. With this information, we use K-means clustering again to re-establish the Toronto neighborhoods (based upon the Foursquare information) and some machine learning to explore the common food cuisines of each neighborhood. With this information, we can explore the areas of the city that have an established food scene where a new restaurant can tether it's success onto and to not be built around an already established Mexican restaurant.

IV. Results and Discussion



Our analysis dictates that even though Toronto has an abundant amount of restaurants (789, according to Foursquare), there are pockets of low density restaurant data for a lot of the areas. Albeit a new restaurant can vitalize a neighborhood and create a new food scene, this is still a gamble. There can be various reasons why there aren't that many restaurants there (i.e. low income residential where people don't eat out much, there's no infrastructure to build a new restaurant, lack of Foursquare data in those areas). So, with a lack of data, it's a gamble that my client would like to avoid. As well, if you look at the above map, we can see established neighborhoods; but if you notice the black points, those are Mexican restaurants. So, we would also want to avoid the surrounding areas as well.

Furthermore, if you look at the breakdown of each cluster and their common restaurants for each, you can see that Mexican restaurants aren't that popular in the city; compare to the most common in asian cuisines (Japanese/Sushi, Korean, Vietnamese, Thai), Middle-Eastern cuisines (Indian, Halal, Pakistani) and Italian looking like the most common. So, adding a Mexican restaurant might give some variety to dining in any of these neighborhoods.

However, with the 2 criterias mentioned prior to an established food scene and a lack of Mexican restaurants in a cluster, I can recommend that the green cluster is the most optimal to break ground for a new Mexican restaurant.

V. Conclusion

The purpose of this project was to identify areas in Toronto that would facilitate a new Mexican restaurant for my client. By looking at clean datasets and cross referencing them with Foursquare data, albeit a limited collection of data points, I was able to create some insights that my client should consider moving forward. And with the recommendation of the green cluster, it's not a bad choice in any way. South downtown Toronto is an area that filled with various food options, shopping districts and tourist attractions all placed along the subway line.

If only they had a Mexican restaurant....