The best place for a restaurant in Toronto city

Farhad Kiavash
Jun 2020

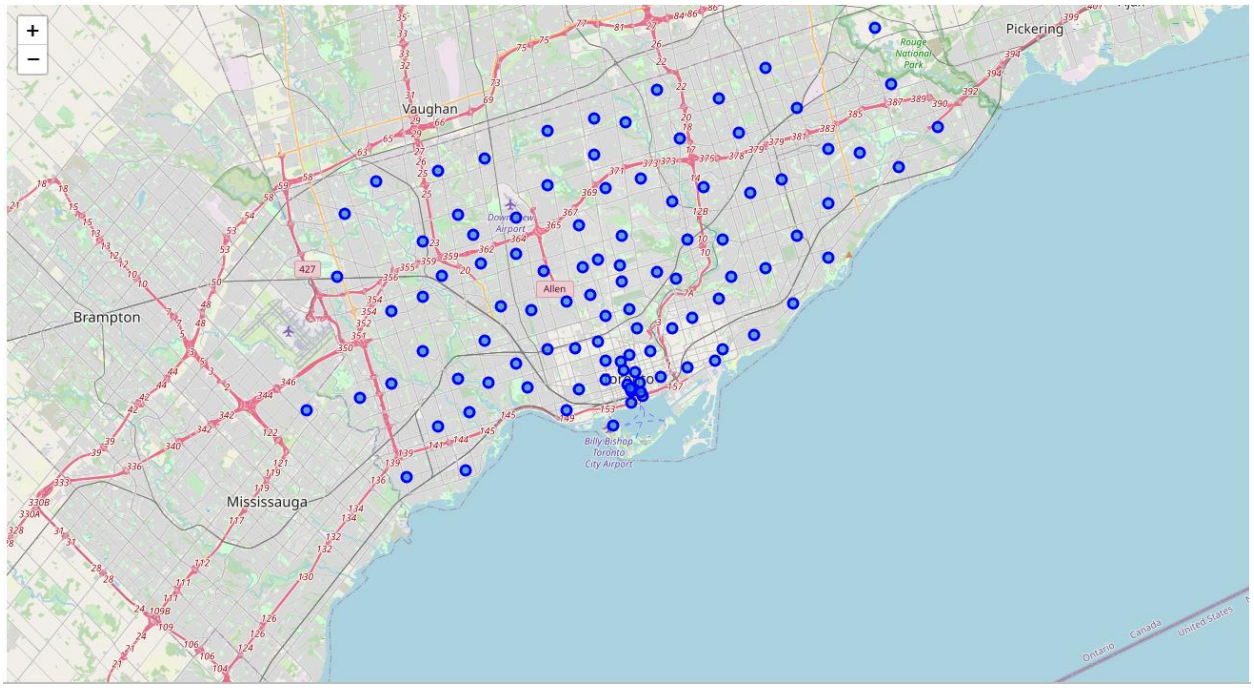
1. **Introduction - Business Problem:**

   The aim of this project is to find an optimal location for a restaurant in **Toronto**, Canada. The same approach can be taken for other cities and other businesses as well. The proper location is the key questions for every new businesses especially in dense cities.
   There are lots of restaurants in Toronto. It is clear that the locations that are already crowded with restaurants are not desirable. Also the preferred location should be as close as possible to the center of the city since it is the place people go to restaurants more often.

2. Data:
   Based on definition of our problem, factors that influence the decission are:
   - Number of existing restaurants in the neighborhood
   - Distance of neighborhood from the city center

   Following data sources are needed to extract/generate the required information:
   - List of boroughs and neighborhoods of the city of Toronto and their postal codes can be obtained from the following Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
   - The geographical coordinates of each postal code can be found using the following URL : http://cocl.us/Geospatial_data
   - Coordinate of Toronto center can be obtained from geopy library
   - Foursquare needs to be utilized in order to find the number of restaurants in each neighborhood

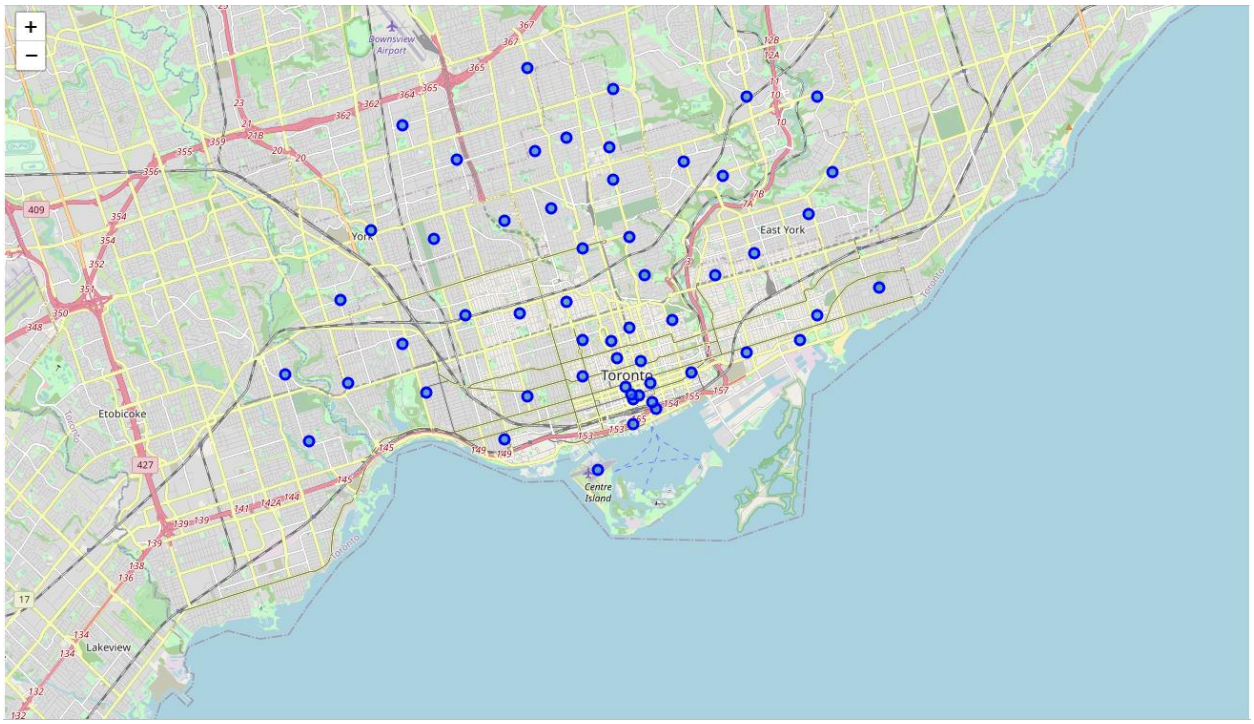
3. Exploring the data:
   Before doing anything lets create a map of Toronto with neighborhoods superimposed on top and see how our data is visualized.



4. Methodology:
   In order to find the optimum locations for our restaurant, we need to get rid of neighborhoods that are very far from the center of the city. So here are the steps:

   1) Find the distance of each neighborhood to the center of the city
   2) Remove all neighborhoods whose distances to the center of the city is more than 10 KM
   3) Find the number of restaurants in each neighborhood
   4) Find the locations that have the low number of the restaurant and are also close to the center of the city.
   5) Cluster the results to show the 5 best area for our restaurant

5. Analysis:
   Exploring the data show us many restaurants that are very far from the center of the city. They are not definitely in our interest. Let's filter out the neighborhoods whose distances are more than 10 KM from the center of the city.
   Let's visualize the remaining neighborhoods one more time and see how they look like:

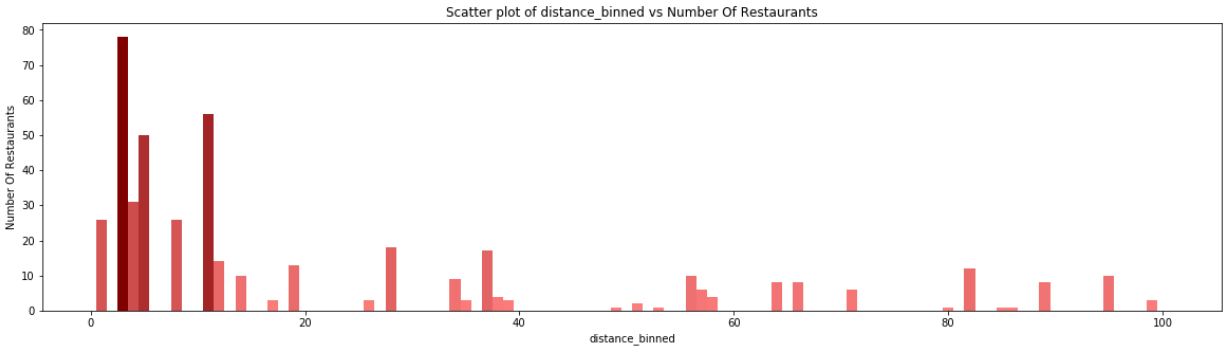


We are now interested in knowing the number of restaurants for each neighborhood. Foursquare is used to fetch the top 100 venues for each neighborhood. We are not interested in all venues so anything except restaurant are filtered out from our data set.

Now we have the required elements in our dataset, i.e. "Distance" and "Number of Restaurants". These two elements are continuous numbers. We prefer the lower distance and the lower number of restaurants.

Let's created two bins for each of these two elements and grade them from 1 to 10. Here is the plot for the sum of restaurants for each distance bin:

As we can see in the above chart the number of restaurants are much higher in the first two bins and it is decreased sharply. It seems that the number of bins are nor selected correctly. Let's change the number of bins to 100 and see how it works.

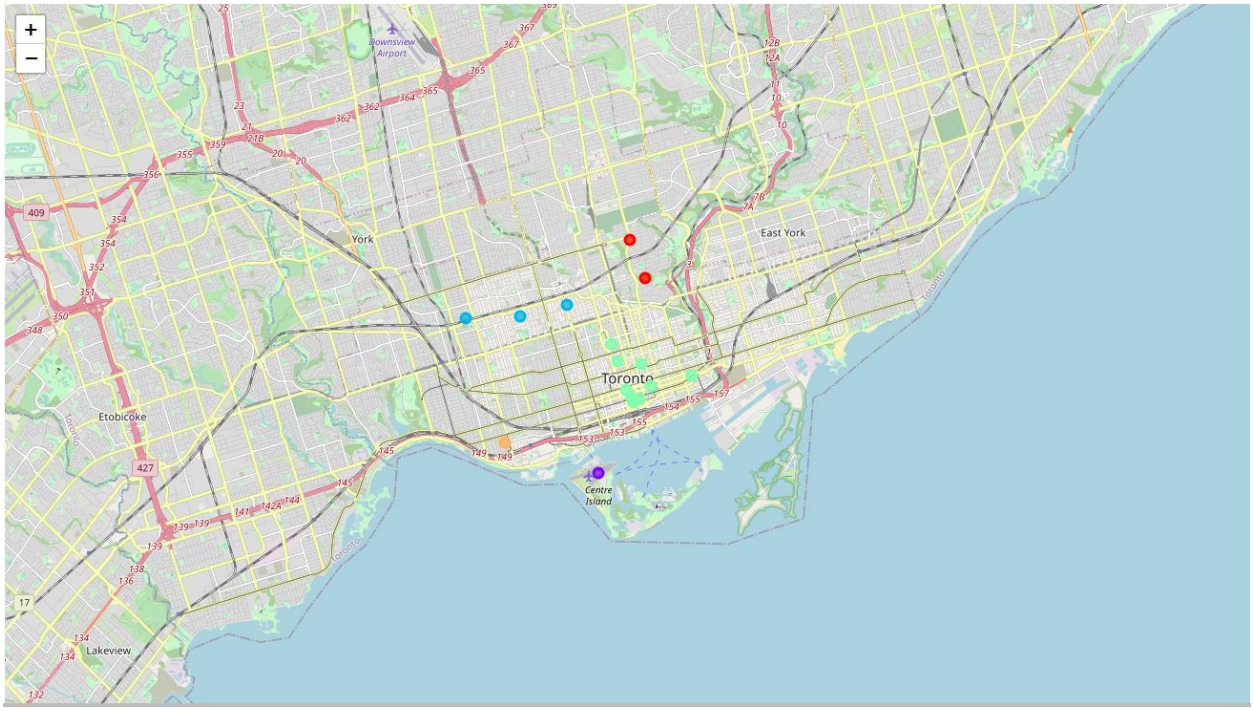


Much better result now. Let's continue with 100 bins for distance and 10 bins for number of restaurants.

Here is the considerations for finding the optimal locations:

- Since we prefer the lower grades for both elements, let's multiply them. The result is between 1 to 1000.
- 1 is the most desirable one. It has the lowest number of restaurants and the least distance.
- 1000 is the worst one. It has the highest number of restaurants and the most distance.
- In order to choose the optimal neighborhoods, only the ones that are selected as the proper locations whose results are 50 and lower

6. Cluster Neighborhoods:
   In order to find the best areas, the proper locations need to be clustered:



7. Results and Discussion:
   The result of our analysis shows 5 different areas that are the optimal locations for our restaurant:

   1) East of university avenue in downtown
   2) The area around Dupont street between Spadina and Dufferin
   3) Toronto Island
   4) Dufferin and King west
   5) Mount Pleasant Rd, between St Clair and Bloor East

   It is entirely possible that there is a very good reason for smaller number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

8. Conclusion:

Purpose of this project was to identify Toronto areas close to center with low number of restaurants in order to aid stakeholders in narrowing down the search for optimal location for a new restaurant. By calculating number of restaurants in each neighborhood and considering their distance from the center of the city, first the best neighborhoods are nominated and then clustering of those locations was performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.