Applied Data Science Capstone Course

January 2020

I. Introduction to Capstone Project.

Being a frequent traveler and often encountering delays one often looks for a bar to pass the time. The business challenge we are looking to address is which airport in the Toronto area would most benefit by having a new bar established in proximity to the gates.

This challenge of where to establish a new business appeals to a bar entrepreneur. He / She will want to determine which airports already have a saturation of similar establishments and therefore perhaps do not support the concept of opening a new bar. The underlying assumption is if there is a concentration of establishments, then competition would be greater and thus a higher entry barrier to the market (Porter's forces).

II. Data

The Data Source for the analysis is the FourSquare data. The locations of the two respective Neighborhood Airport locations, Toronto Pierson, and Billy Bishop Airport are used for their coordinates. Geographic coordinates for these two locations are also utilized as data derived from the Geopy function. The data of FourSquare is queried using the URL passing a request for "Bar" within a proximate radius. In particular we are looking for the name, longitude and altitude of the resultant locations and proximity to the two airport neighborhoods we are comparing.

III. Methodology

The results from the URL based data query of FourSquare are filtered to eliminate extra fields and then mapped to indicate the density of establishments. This is done by taking the json resultants from the FoureSquare query and manipulating them into a data frame. The date frame is cleaned up and the coordinate data is utilized to plot the respective locations.

The radius search for locations was increased from 450, 650, 850 and a count was made to determine the number of bar establishments are within proximity. This increasing radial search was done by passing a different radius parameter to the FourSquare URL query and noting the resultant number of returns in the converted data frame. The shape function was used to make the count of returned elements.

The Folium library is used to generate a plot of the 850 radius search for each of the respective locations in each of the respective neighborhoods.

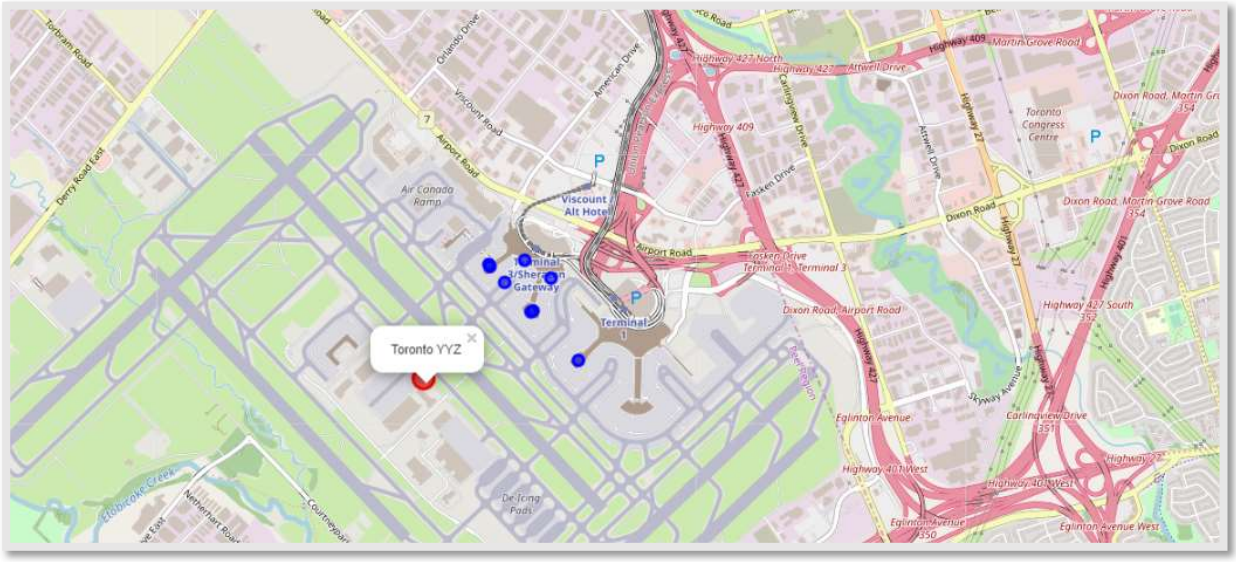
IV. Results

The results of the number of locations in proximity to each neighborhood is as follows using multiple FourSquare Query and shape function

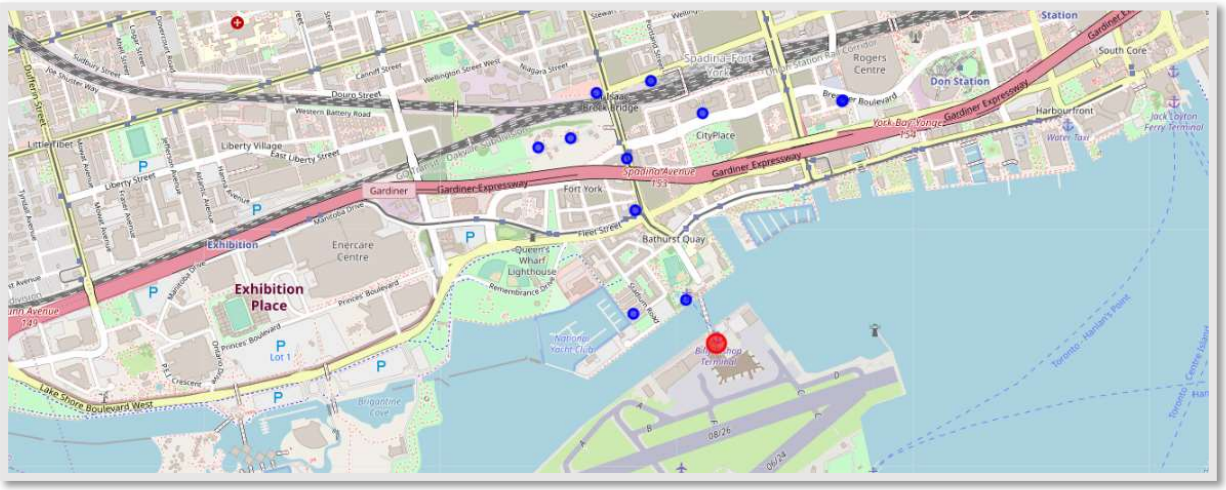
Number of Bars within Respective Neighborhoods		
Radius	**Toronto Pierson**	**Billy Bishop**
250	0	1
450	0	2
650	3	4
850	8	11

The results of the folium map plots are as follows:

Toronto Pierson FourSquare Plotted locations of "Bars" 850 Radius

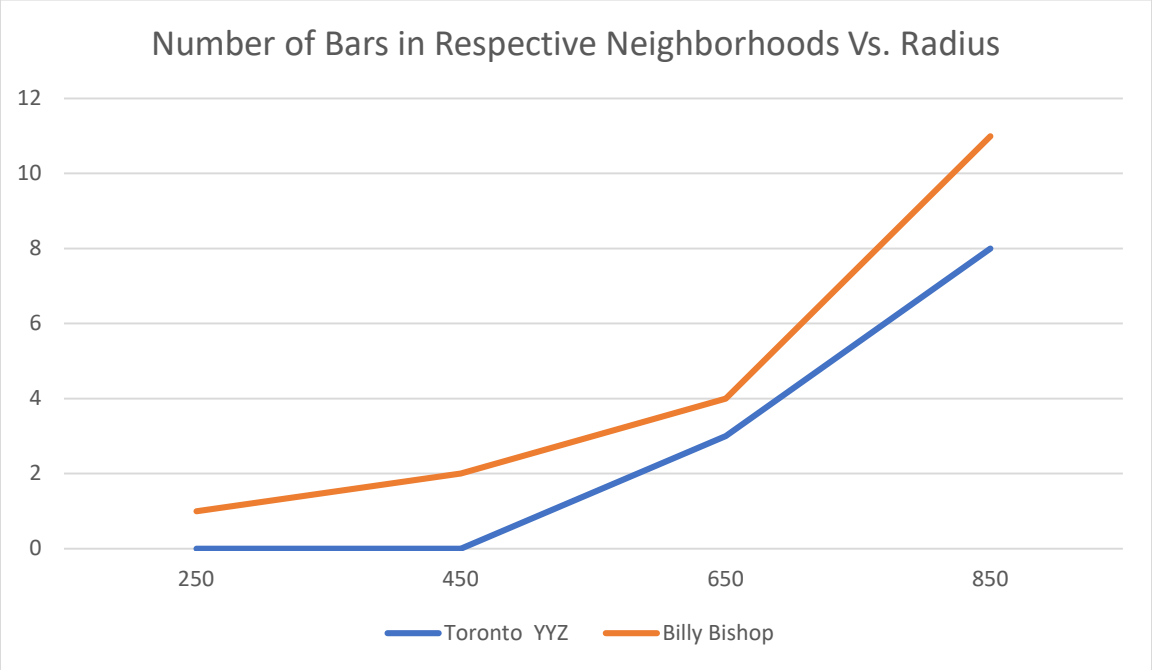


Billy Bishop FourSquare Plotted locations of "Bars" 850 Radius



V. Discussion

The increasing radial search is a powerful tool to determine the competition within proximity of a target neighborhood. The results can be visualized in this graph:



Toronto Pierson has a lower number of bars than Billy Bishop regardless of the distance from the center of the location. The map plots confirm this notion as we can see the concentration of bars is “on premise” at Toronto Pierson versus at Billy Bishop the bars are scattered through the adjacent neighborhood.

The ability to plot the locations within their neighborhood is necessary to making an informed decision about where to establish a new business. This will be discussed in the conclusion.

VI. Conclusion

The queries from FourSquare have helped indicate the potential competition to establishing a “bar” at either Toronto Airport or Billy Bishop. Typically when entering a new market with a business the goal is to avoid entering a saturated market where is the much competition. A increasing radial search using the location data provided by FourSquare can be a useful tool to determine what the market looks like.

In this particular case the table generated from radial search point to the conclusion to add a bar at Toronto YYZ as there are a lower number of bars in the neighborhood, and there are none within the 450 radius perimeter.

However we can not always rely purely on the “data”. The visual interpretation of the map indicates that all the bars in Toronto are within the Airport and clustered quite close. Adding a new business in this type of environment would encounter strong competition. On the other hand while there are more bars near Billy Bishop, there is only one at the 250 radius, therefore the suggestion to the entrepreneur who wants to open a bar in one of the metro area airports would be to add a bar on premise at Billy Bishop as there is less direct competition. Even though there are more bars near Billy Bishop, that particular airport is on an island, and transit back and forth to local establishments is not as convenient as the bars within the terminal at Toronto.