Applied Data Science Capstone

Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods

Now that you have been equipped with the skills and the tools to use location data to explore a geographical location, over two weeks, you will have the opportunity to be as creative as you want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of your choice or to come up with a problem that you can use the Foursquare location data to solve.

1) Introduction/Business Problem

Clearly define a problem or an idea of your choice, where you would need to leverage the Foursquare location data to solve or execute. Remember that data science problems always target an audience and are meant to help a group of stakeholders solves a problem, so make sure that you explicitly describe your audience and why they would care about your situation.

This study aims to help people plan to open a new restaurant in Toronto to choose the right location by providing data about the income and population of each neighborhood and the competitors already present in the same regions.

2) Downloading and Prepping Data

Describe the Data that you will be used to solve the problem or execute your idea. Remember that you will need to use the Foursquare location data to solve the problem or execute your idea. You can use other datasets in combination with the Foursquare location data. Ensure that you provide adequate explanation and discussion, with examples, of the Data you will be using, even if it is only Foursquare location data.

To provide the stakeholders with the necessary information, I'll be combining Toronto's 2016 Census that contains Population, Average income per Neighborhood with Toronto's Neighborhoods shapefile, and Foursquare API to collect competitors on the same neighborhoods.

Toronto's Census data is publicly available at this website: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#8c732154-5012-9afe-d0cd-ba3ffc813d5a

Toronto Neighborhoods' shapefile is publicly available at this website: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#a45bd45a-ede8-730e-1abc-93105b2c439f

3) Methodology

The methodology section, which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, and what machine learnings were used and why.

For this report, I used a few different maps to help a new investor decide the best neighborhood to open a restaurant in Toronto based on its income, population, and available competitors. To do that I've used the 2016 Census information combined with choropleth maps to visually display the wealthier and more populational neighborhoods and Foursquare data to show the current restaurants in each region.

4) Results

Results section where you discuss the results.

Comparing the maps, we can notice most of the restaurants grouped on main streets and the south of the city, although some of the wealthiest neighborhoods are up to the north. Also, the areas with a dense population don't reflect on the number of restaurants.

5) Discussion

Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

When I first decided to create this study, I was expecting to find clusters of restaurants in specific regions and the final result didn't meet that expectation.

6) Conclusion

Conclusion section where you conclude the report.

This report may help someone plan to open a restaurant in Toronto by comparing the current offers and neighborhood profiles. However, it may not cover all variables such as access to public transportation or even the restaurant profiles, so it shall not be used as a single decidion making tool.