1. Introduction

1.1 Problem

A foreign entrepreneur wants to open his first restaurant in Toronto. The objective is to find which area in Toronto most suited to support an Afghan restaurant based on market demand, competitors, demographic and other economic indicators.

2. Data

2.1 Data Sources

Foursquare's API is used to provide venue data, especially 'Afghan Restaurant' categories. Wikipedia's data on Toronto Post Codes was used for the location data in this assignment, and also to locate what neighborhoods may host Afghan restaurants presently. Data from the Open Data Toronto portal was used to collect income and ethnic group information for the city by neighborhood. These neighborhoods are not defined by post-code, and are instead defined geographically by the Canadian Census, through StatsCanada. All of this data is the most up-to-date information available from their respective sources.

2.2 Data Collection and Cleaning

Data scraped from Wikipedia was cleaned to remove N/A variables, and then grouped by post code. These were then geocoded, with the location data being used to eliminate neighborhoods that were not in Toronto proper. This data, from Wikipedia, was then combined with the Foursquare API data into one dataframe. This data was one-hot-encoded to be able to find the Afghan restaurants in Toronto.

After this I imported the Canadian Census data from a CSV file produced through the Toronto Open Data Portal. This information contained income data grouped by deciles, as well as the numbers of Afghan-Canadians living in each neighborhood.

3. Methodology

Using the aforementioned data, I transposed the dataframe so neighborhoods were all in one column with the demographic attributes and values in corresponding columns. The titles of these categories served as the index along axis 0. I then dropped the index as there were 2 categories and I only needed one to analyze the data. Following this I sorted the data to find where the highest concentrations of Afghans were. As this would be the community to support the restaurant, I then compared the neighborhoods with the highest concentrations of Afghans: Willowdale East, Newtonbrook West, Humber Summit, and Willowridge-Martingrove-Richview. These 4 neighborhoods each had 95 Afghans living in them – not a great population base to support a restaurant. In comparing their decile income information, I found that the first 3 had more Canadians living in the poorest 5 deciles rather than the wealthiest 5. Only Willowridge-Martingrove-Richview had the inverse, so I selected this one as the neighborhood to site the restaurant.

The numbers of Afghan-Canadians by neighborhood never topped 95. In fact, in the top 4 neighborhoods by population, each had 95 such persons. As these had the highest concentrations of the Afghan population in Toronto, I selected these for analysis. This led me to then look at their decile income indicators.

4. Results

As stated above, the analysis I did of the data led me to choose 4 neighborhoods to look at for potentially siting a restaurant. As 3 of these neighborhoods had larger population in the bottom half of income deciles, I decided to choose the only neighborhood that had the inverse to site my restaurant.

5. Discussion

In analyzing the data, I did not encounter any problems for what I wanted to do. Luckily, python was able to handle all of the analysis within my notebook. However, my results concern me and bring me back to a point I made in my research question within the notebook. I questioned there what neighborhood would best support the restaurant, which I did. However, I have also determined that Toronto is not a good market for such a restaurant. As the Afghan population is spread among the neighborhoods of Toronto and may not be concentrated in Toronto proper, there is not enough of a population base among this ethnic group to support a restaurant within their neighborhood. Further, Afghans tend to live in poorer or middle-class neighborhoods that would not have the economic or social capital to support a restaurant. It may be that Toronto only has 1 Afghan restaurant now because that is all it can support and there is no market there for it, which my data suggests.

6. Conclusion

Based on the foregoing results and discussion, I found that Willowridge-Martingrove-Richview would be the best neighborhood to site an Afghan restaurant. However, I would generally caution against doing so, due to the low population concentration of Afghans and the low economic capital in this neighborhood that may support a business.