Finding the Optimal Neighborhood in Manhattan for Café Business

Ruby Wu

December 22, 2020

1 Introduction

1.1 Background

Manhattan, the most densely populated borough of the New York City, is a well-known area which serves as the economic and administrative center of the United States. It covers 33.58 square miles' land and has approximately 1,628,706 residents (by 2019) in total. People all around the world are attracted to Manhattan because of its leading economy, diverse culture, historical tourism and, more importantly, huge potential for success. Therefore, it is advantageous for us to explore deeper into Manhattan in terms of business opportunities.

1.2 Business Problem

There are lots of coffee shops locating in different neighborhoods of Manhattan. Their business performances highly depend on various factors including location, residence and other business categories. Therefore, based on these considerations, this project aims to find the optimal neighborhood for opening a coffee shop/café. Specifically, we will explore 40 neighborhoods in Manhattan and generate a best solution for stakeholders interested in coffee business.

2 Data

We will be using three types of data in this project: Manhattan neighborhood data, Manhattan population data and Foursquare venue data. In this section, we will explain our data sets in terms of their sources and content.

2.1 Manhattan Neighborhood Data

Manhattan neighborhood data can be found in *IBM Skills Network*. The data is in JSON format containing features including boroughs, neighborhoods, latitude and longitude across all 5 boroughs in the New York City. The dataset has 306 neighborhoods in total along with their coordinates. We will only use borough Manhattan data from this data set.

2.2 Manhattan Population Data

Manhattan population data can be scraped from the Neighborhoods in New York City page of Wikipedia. It contains area, population and population density for each community board which is the appointed advisory group of the community districts of Manhattan. Manhattan has 12 community boards (CB) in total. Based on this dataset, we can see which neighborhoods are assigned to the same community board and the population distribution across all community areas.

2.3 Foursquare Location Data

We will also use the Foursquare API to pull the venue information based on each Manhattan neighborhood's coordinates and to explore what kinds of venue categories are in the

neighborhood and their frequencies. We will also pay closer attention to coffee related categories specifically. According to the Foursquare documentation, venue categories related with coffee are named to Coffee Shop, Cafeteria, or Café. Therefore, when we explore coffee business in Manhattan, we will extract those venue data based on the three category names.