Analysis for opening Whole Sale Store for Food Products

Mahesh Venkat Reesu

August 18, 2020

1. INTRODUCTION

1.1 Background

Instead of letting stores buy regionally, most purchasing is now handled through wholesale headquarters, giving them greater leverage with suppliers and potentially streamlining the entire delivery process. Centralized buying is likely to eliminate many regional distributors in favour of large distributors that can deal in large volume and accept a lower price.

Hence, this project deals with the analysis of the neighbourhoods of Brooklyn and finding out the clusters based on customers & venues that is best suitable for constructing and Finally finding the mean position in a particular cluster by taking travelling and other conditions into account.

Brooklyn got diverse food items, because of its diversity in culture. There are many restaurants in Brooklyn, each belonging to different categories like Chinese, Indian, French etc. Who are main customers for Wholesale retailers.

1.2 Problem Statement

Find the best location for constructing(opening) a Wholesale Store for raw materials used in preparation of Food Stuffs in Brooklyn.

This project aims to predict location of establishing a store in Brooklyn based on the dataset and visualise the clusters and analyse the customers.

1.3 Interest

Obviously, all famous Restaurants, Factories and Stores are interested in doing business with whole sale retailers.

2. DATA AQUSITION AND CLEANING

2.1 Data Sources

1. New York City data that contains list Boroughs, Neighbourhoods along with their latitude and longitude. From this data set all neighbourhoods of Brooklyn are collected.

source: *https://cocl.us/new_york_dataset*

2. By using the below API we will get all the venues in each neighbourhood. We can filter these venues to get only those do business with warehouse.

source: *Fousquare API*

3. By using the below link geo space data we will get the New york Borough boundaries that will help us visualize map.

source: *https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm*

2.2 Data Cleaning

Data downloaded or scraped from multiple sources were combined into one table. All the unwanted data from the downloaded json file was removed. Only Borough, Neighbourhood and it's Latitude along with Longitude were kept into a table.

As this project analyses the data of Brooklyn, entries in Borough containing only Brooklyn were taken into a new data set (Figure 2.1:). This is the final data set for processing and visualising.

```
In [65]: Brooklyn_data = neighborhoods[neighborhoods['Borough'] == 'Brooklyn'].reset_index(drop=True)
         Brooklyn_data.head()
```

Out[65]:

	Borough	Neighborhood	Latitude	Longitude
0	Brooklyn	Bay Ridge	40.625801	-74.030621
1	Brooklyn	Bensonhurst	40.611009	-73.995180
2	Brooklyn	Sunset Park	40.645103	-74.010316
3	Brooklyn	Greenpoint	40.730201	-73.954241
4	Brooklyn	Gravesend	40.595260	-73.973471

Figure 2.1: Boroughs Coordinates

3.DATA ANALYSIS AND VISUALISING

3.1 Analysing and Mapping the Venues

Geographical co-ordinates were acquired from geolocator using which a Follium map was obtained to observe the neighbourhood of New York. Geographical co-ordinates were acquired from geolocator using which a Follium map was obtained to observe the neighbourhood of Brooklyn.

Foursquare Client ID and Client Secret were loaded to get the neighbour hoods and venues in Brooklyn. After getting the neighbour, one of them which was nearer to New York and located in Brooklyn was selected and its Geographical coordinates were noted.

	name	categories	lat	lng
0	Pilo Arts Day Spa and Salon	Spa	40.624748	-74.030591
1	Bagel Boy	Bagel Shop	40.627896	-74.029335
2	Pegasus Cafe	Breakfast Spot	40.623168	-74.031186
3	Leo's Casa Calamari	Pizza Place	40.624200	-74.030931
4	Cocoa Grinder	Juice Bar	40.623967	-74.030863
5	Ho' Brah Taco Joint	Taco Place	40.622960	-74.031371
6	The Bookmark Shoppe	Bookstore	40.624577	-74.030562
7	Brooklyn Market	Grocery Store	40.626939	-74.029948
8	Blue Door Souvlakia	Greek Restaurant	40.624567	-74.030311
9	Mimi Nails	Spa	40.622571	-74.031477
10	Windy City Ale House	Sports Bar	40.628117	-74.029128
11	Georgian Dream Cafe and Bakery	Caucasian Restaurant	40.625586	-74.030196

Figure 3.1: Neighbourhood Coordinates and categories

Using Foursquare API and json normalisation all the venues in selected Neighbourhood(Bay ridge) were Adopted into a table with name, category, and location Coordinates. Total count venues in that neighbourhood obtained is 78.

Number of Venues in each neighbourhood around Brooklyn that are expected to do business

With this warehouse are plotted as a Bar Graph (Figure 3.2).

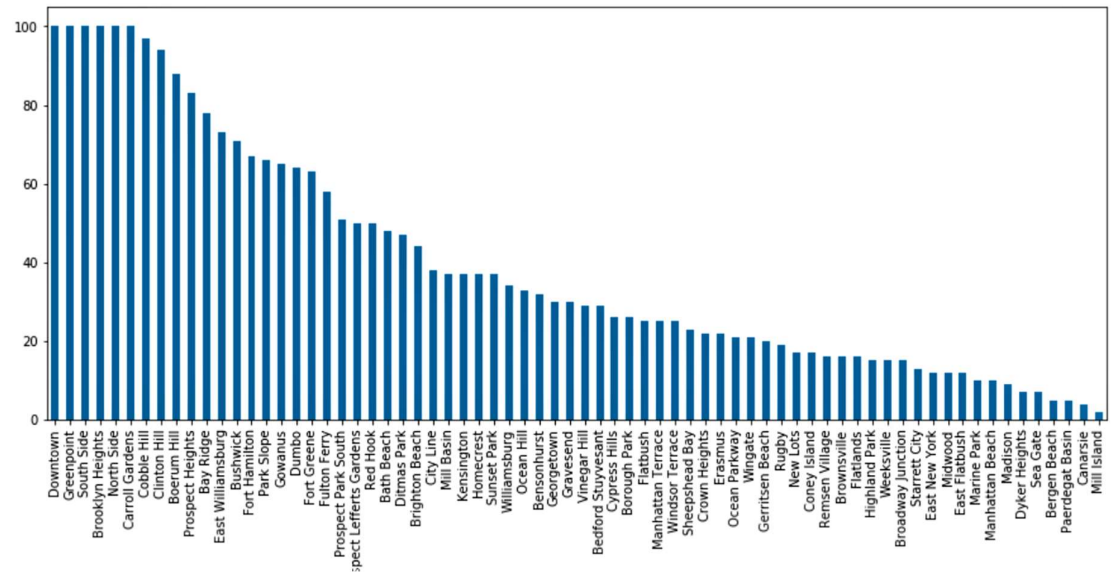

Figure 3.2: Number of Venues in each neighbourhood

All those Venues are shown on the map by using markers and location coordinates as shown below

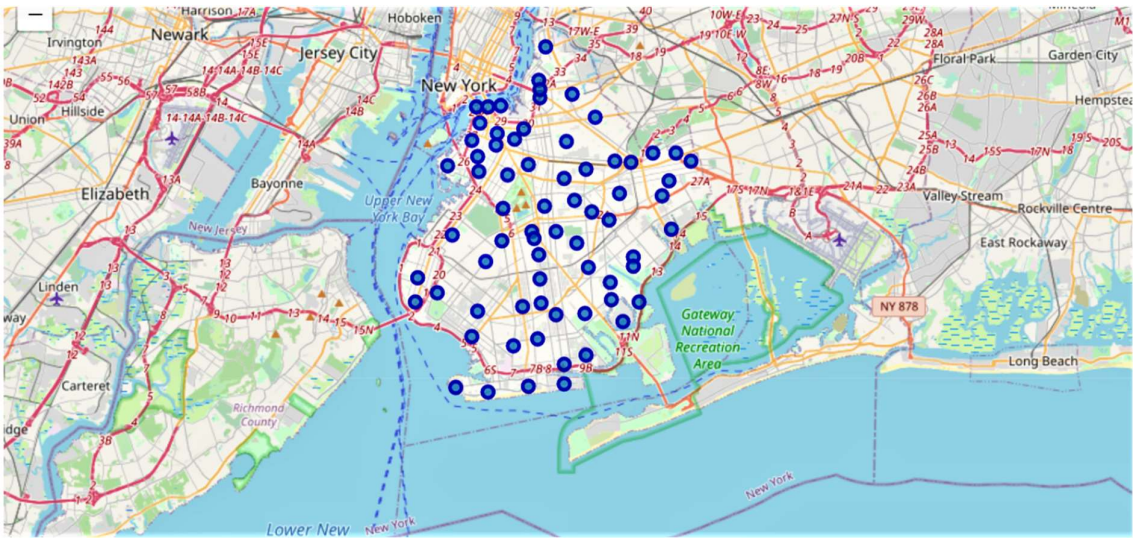

Figure 3.1.2: Neighbourhoods

3.2 Clustering the venues

Firstly, all the venues are grouped according to the neighbourhood and each neighbourhood was analysed by using “one hot encoding”. Top ten most common venues are appended into a new data frame for each neighbourhood. From SciKit learn Kmeans was imported at the beginning. Using, Kmeans the entire neighbourhoods were divided into 5 clusters.

Visually all the five clusters are shown by using markers on a map. This map was plotted on by using Follium and foursquare API.

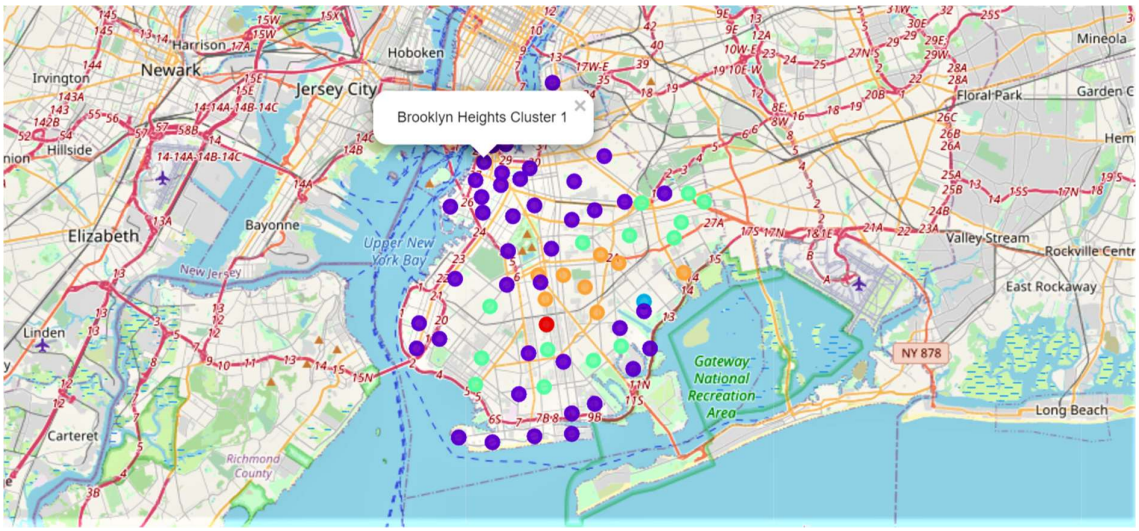

Figure 3.2: Visualizing the Cluster of Neighbourhoods

Cluster 0: Red

Cluster 1: Violet

Cluster 2: Blue

Cluster 3: Green

Cluster 4: Orange

4. ANALYSIS AND RESULTS

4.1 Analysis

Each cluster was analysed separately by creating a separate table for each cluster. From the cluster map we can see that the Cluster 1 has more relevant customers to do business with the Whole Sale Ware House.

From the Bar Graph of the Venues the among the first half of the venues most of them are located in upper part of the Cluster 1.

4.2 Results: Hence, the mean location for upper part of the Cluster 1 can be selected for setting up the ware house. If we want the location that is nearer to New York and the Sea Shore according to the Analysis Brooklyn Heights or Down Town are best locations for construction.