Applied Data Science Capstone Project Report

The Battle of Neighborhoods

Coursera - IBM Data Science Professional Certificate

Finding an optimal location for a Supermarket in Brooklyn, New York City

Submitted by:

Somraj Chowdhury

Table of contents

1. Introduction

The New York City (NYC) is the most populous city in the United States and provides a lot of business opportunities. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, advertising, legal services, accountancy, insurance, theater, fashion and the arts in the United States. In the 21st century, New York has emerged as a global node of creativity and entrepreneurship, social tolerance and environment sustainability.

New York City consists of five boroughs, each of which is a separate county of the State of New York.

The five boroughs are:

1. Brooklyn
2. Queens
3. Manhattan
4. The Bronx
5. Staten Island

New York is a highly developed city and so the cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully.

2. Business Problem

2.1 Problem

This project will aim at finding a suitable and optimal location for a Supermarket in Brooklyn, in the City of New York, United States.

2.2 Problem Description:

A supermarket is a self-service shop offering a wide variety of food, beverages and household products, organized into sections and shelves. With the increase in the population, the demand and availability of essential resources like food and other

households increase and supermarkets serve this very purpose of providing the resources under one roof.

2.3 Target Audience

This project is aimed at those individuals and business groups who are interested to invest in opening a supermarket in Brooklyn, New York.

3. Data

In this project we will be exploring the New York City dataset and analyzing the neighborhoods of Brooklyn borough.

3.1 New York City Data

This dataset contains data of the boroughs and the neighborhoods that exist in each borough along with the geographical coordinates of each neighborhood.

This dataset is available for free online.

Link: https://geo.nyu.edu/catalog/nyu_2451_34572

This dataset holds data of a total of 5 boroughs and 306 neighborhoods.

Borough	Number of Neighborhoods
Bronx	52
Brooklyn	70
Manhattan	40
Queens	81
Staten Island	63
Total Neighborhoods	306

The New York dataset has the following features and format.

	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

3.2 Supermarket Data from Foursquare API

Information about the existing supermarkets in Brooklyn, New York City like the name of supermarkets in a neighborhood and their location along with geographical coordinates etc. will be obtained from the Foursquare API.

We request the required data by writing an URL along with the client_id, client_secret, version, latitude, longitude, radius, limit and category id for the type of venue you would like to retrieve information in this case it supermarket.

The following screenshot shows a part of the supermarket data retrieved from the Foursquare API.

Supermarket Name	Supermarket Latitude	Supermarket Longitude
Met Fresh Supermarket	40.616528	-74.034003
Metropolitan CityMarket	40.617375	-74.030735
CTown Supermarkets	40.629234	-74.022803
Jmart 新世界超市	40.610080	-74.001221
Foodtown	40.619927	-74.032301
Scaturro Supermarkets	40.629409	-74.005051
Food Dynasty	40.611275	-74.008544

3.2 Population data for each Brooklyn neighborhood

Population dataset for each Brooklyn neighborhood was created by finding the population of each neighborhood individually from the web.