The Battle of Neighborhoods

Finding the best location to set up a new boutique café in Manhattan,NY

Applied Data Science Capstone Project

Danielle O'Sullivan

1. Introduction

1.1   Background

New York City is one of the most diverse cities in the world bursting with attractions. With over 3000 coffee shops in New York it can be said that New Yorkers love their coffee. Recently, MassiveHealth found that New Yorkers drink 6.7 times the amount of coffee consumed by the average denizen of any other US city. -https://medium.com/topos-ai/the-next-wave-predicting-the-future-of-coffee-in-new-york-city-23a0c5d62000.

1.2   Business Problem

Selecting a location in New York to open a new boutique cafe is crucial to a businesses' success.

There are many factors to consider when choosing the location of this new business:

1. Nearby Venues: Other cafes, restaurants, gyms, etc.
2. Crime: Will employees be safe coming to/leaving work, will there be people visiting this area
3. Nearby subway stations: a coffee shop near a subway station might attract more visitors

In this project we will analyse the above factors to determine which area would be the best to open this new café. We will be looking at neighbourhoods in the Manhattan borough only.

1.3   Target Audience:

The target audience for this analysis would be those looking to open a new café/small business in Manhattan.

2   Data

For this analysis, the following datasets were used:

2.1   New York City Data

New York City Data which was downloaded as a json file from https://cocl.us/new_york_dataset. This includes each neighborhood, borough, latitude and longitude.

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Figure 1 - New York City Data

2.2 Crime Data

The crime data was downloaded as csv file from NYC Open Data
https://data.cityofnewyork.us/Public-Safety/Crime-Map/5jvd-shfj. The data was filtered to
Manhattan borough. Violation was removed from the law category as a lot of these crimes will not
be relevant. We will only be looking at felony and misdemeanour crimes. The dataset was also
filtered to only include crimes reported in 2019. After filtering there is 98,693 records.

	CMPLNT_FR_DT	LAW_CAT_CD	BORO_NM	Latitude	Longitude
1830027	01/02/2019	FELONY	MANHATTAN	40.773332	-73.961074
1830636	01/03/2019	FELONY	MANHATTAN	40.787567	-73.943132
1832787	01/06/2019	MISDEMEANOR	MANHATTAN	40.765024	-73.984836
1832815	01/09/2019	MISDEMEANOR	MANHATTAN	40.815732	-73.945420
1832851	01/14/2019	FELONY	MANHATTAN	40.794515	-73.966324
1832873	01/18/2019	MISDEMEANOR	MANHATTAN	40.723659	-73.991022
1832897	01/27/2019	MISDEMEANOR	MANHATTAN	40.732356	-73.984941
1832910	01/31/2019	MISDEMEANOR	MANHATTAN	40.749780	-73.987781
1832917	01/29/2019	MISDEMEANOR	MANHATTAN	40.823575	-73.937675
1832941	02/04/2019	MISDEMEANOR	MANHATTAN	40.710783	-73.996632

Figure 2- New York City Crime Data

2.3 Foursquare API

Foursquare data will be used to explore each neighborhood in Manhattan and venues nearby each
of the coordinates in the New York City data.

2.4 Nearby Subway Stations

Subway Entrances dataset was downloaded as a geojson file from NYC open Data
https://data.cityofnewyork.us/Transportation/Subway-Entrances/drex-xx56.

The json file was loaded and read into a new Dataframe. The dataset consists of 98 records.

	Object_ID	Name	URL	Latitude	Longitude
0	1734	Birchall Ave & Sagamore St at NW corner	http://web.mta.info/nyct/service/	40.849169	-73.868356
1	1735	Birchall Ave & Sagamore St at NE corner	http://web.mta.info/nyct/service/	40.849128	-73.868213
2	1736	Morris Park Ave & 180th St at NW corner	http://web.mta.info/nyct/service/	40.841223	-73.873499
3	1737	Morris Park Ave & 180th St at NW corner	http://web.mta.info/nyct/service/	40.841453	-73.872892
4	1738	Boston Rd & 178th St at SW corner	http://web.mta.info/nyct/service/	40.840815	-73.879623

Figure 3 - New York City Subway Data