Finding Pharmacy Location in NYC

Coursera: Data Science Capstone Project for Final Assignmen

Introduction

- Due to spread of Covid-19 healthcare infrastructure is getting more and more crucial to public health
- location of a new pharmacy in Manhattan will be analysed
- turnover of a pharmacy and the supply of medicines to the population depends not only on the number of pharmacies but also on the population density of a neighbourhood

Data Description

- Foursquare API is used for this purpose, which uses queries to output the venues in each neighbourhood within a given radius
- combined with public data on the demographics of each neighborhood. This data is publicly available from the City of New York and can be accessed at https://data.cityofnewyork.us/api/views/8m6s-esnp/rows.csv.

Extract from NYC neighborhood population data

	Borough	Year	FIPS County Code	NTA Code	NTA Name	Population
0	Manhattan	2010	61	MN01	Marble Hill-Inwood	46746
1	Manhattan	2010	61	MN03	Central Harlem North-Polo Grounds	75282
2	Manhattan	2010	61	MN04	Hamilton Heights	48520
3	Manhattan	2010	61	MN06	Manhattanville	22950
4	Manhattan	2010	61	MN09	Morningside Heights	55929

https://data.cityofnewyork.us/api/views/8m6s-esnp/rows.csv.

Methodology

- From the population data, a so-called population score (Pop Score) was then derived, which represents the relative population of Manhattan in the respective neighborhood. This score is to serve as one of the bases for decision-making at a later stage of the analysis.
- use the Foursquare API to request and store specific information on the individual venues in the respective neighborhoods
- relative frequency of pharmacies in the respective neighbourhood was calculated and a so-called Pharmacy Score was derived
- To combine both scores, their relative weight is set at 50% each, as this weight is considered realistic by the authors of the study. An overall score is then calculated from the difference between the Pop Score times relative weight minus the Pharmacy Score times relative weight.
- a K-Means cluster analysis was performed to identify possible similarities between individual neighborhoods

Results

- Upper West Side is the best neighborhood for a new pharmacy (high population and low competition)
- Marble Hill closed as the worst neighbourhood for opening, not least because the relatively lower population was not able to compensate for this either.
- If the average overall score is calculated for each cluster, clusters 0, 2 and 3 are particularly attractive for a particular branch concept
- Cluster 0 represents the densely populated lively neighborhoods such as Upper West Side, Lower East Side and Yorkville.
- Cluster 3 represents the affluent neighborhoods such as Roosevelt Island, Hamilton Heights and Manhattanvile, where a different branch concept that focuses on more affluent customers, for example, could be promising.

Neighborhood	Latitude	Longitude	Population	Pop Score	Pharmacy Score	Overall Score
Upper West Side	40.787658	-73.977059	132378.0	12.487813	0.000000	6.243907
Yorkville	40.775930	-73.947118	77942.0	7.352620	1.000000	3.176310
West Village	40.734434	-74.006180	66880.0	6.309092	0.000000	3.154546
Lincoln Square	40.773529	-73.985338	61489.0	5.800534	1.041667	2.379434
Hamilton Heights	40.823604	-73.949688	48520.0	4.577110	0.000000	2.288555

#1 Upper West Side

#2 Yorkville

#3 West Village

Discussion

- further analyses with the help of additional data are certainly useful
- costs for a branch, such as rent, personnel costs and incidental expenses, as these can vary from quarter to quarter
- turnover potential also depends not only on the population size but also on the average income in the respective neighbourhoods