Oportunistic Opening of a new Gym in the Center of NY City

Capstone Project – Mar 30, 2020

Introduction

I am a freelance consultant, advising small businesses. One customer, active in the fitness sector, contacted me with the following idea: Due to the current Corona (Covid-19) crisis he expects that several restaurants will go bankrupt; especially bad restaurants. He wants to take advantage of this situation and cheaply take over one restaurant and convert it into a gym. He already owns several gyms, but wants to expand into New York with a first gym, in the center of Manhattan - where cheap locations are hard to find. Of course the new gym has to be at a location, where no other gym is located in the close area. My client asks to provide him a list with bad restaurants in the center of Manhattan, where no gym is located in the close area.

Data Source

The problem will be solved as follows: 1 - Define the center of Manhattan (e.g. time square) and get the location data of the center from Foursquare. 2 - Find all restaurants in the wider area of the center (e.g. within 1000 meters); get the rating of all restaurants, and only select restaurants with a bad rating (e.g. rating below mean). Venues (Restaurants) and Ratings are available in Foursquare. 3 - Drop all restaurants that have a gym in the close area (e.g. within 100 meters). Gyms and their geographical locations are available in Foursquare.
There is a limitation to access rating data with a free Foursquare account. Therefore, the analysis is only done for a limited number of restaurants.

Methodology

For this project the key relevant piece is location data, i.e. data describing places and venues incl. their geographical locations. In this project Foursquare will be used as data provider, providing a massive set of location data. With 'calls' venues (specifically 'Restaurants' and 'Gyms') in the database will be searched and loaded. A Foursquare account has to be created to get access.

When the data from Foursquare is downloaded, the data is further manipulated and analysed in table format (dataframes) with the open source tool 'pandas'. Lists of the restaurants, meeting the search criteria, are exported to csv files, that could be provided to the customer.

For the visualization of the locations of the restaurants, the location data from Foursquare is used, and shown on interactive maps by using the package 'folium'.

Results

The results are provided in three csv-files and in three maps.

1 – List/Map with all restaurants in the close area of the center
17 restaurants were identified in the close area of the center (Limited entries due to free Foursquare account)

Preview:

2 – List/Map with all 'bad' restaurants in the close area of the center
9 of the 17 Restaurants were classified as 'bad'. The mean rating of the restaurants is 7.16. Restaurants with a rating below the mean (rounded 7) or with no rating, were defined as 'bad'

Preview

3 – List/Map 'Bad' Restaurants in the Center – not next to existing Gym On average the next Gym is within 57 meters of the restaurants. We defined the minimal required distance to an existing Gym has to be at least 100 meters. There is only one restaurant left fulfilling this condition!

Preview

Discussion

As only 1 Restaurant was identified some rework is required. As a first step a paid Foursquare Account is required to avoid limitation regarding rating information.

As a next step the Gyms in the close area of the restaurants could be classified; there might be Gyms that must not be considered as competitors. The categories of the Gyms should be analysed to understand the customer segment.

Furthermore, the radius (currently defined as 1000 meters) of the center could be expanded to identify additional restaurants.

Conclusion

Further discussions with the customers are required (Could the definition of the 'Center' be changed? How could the existing 'Gyms' be categorised?). Additional iteration in the analysis are needed; especially given the high density on existing Gyms in the center of New York; and the fact that only one single restaurant was identified.