

Battle of the Neighbourhoods

Part 1

Business Problem

The problem I have chosen to analyse involves the decision regarding where to open a gym in New York City , USA. There are several factors which will play a role in determining this optimal location. The gym needs to be centrally located and needs to be easily accessible from all parts of the city.

It also needs to be closer to commercial buildings which will make it easier for customers to hit the gym for a workout before heading to work. It should also not exist in a place where several gyms already exist which will generate unnecessary competition. I will also attempt to determine the time at which gyms at most frequently visited.

I will attempt to generate a list of neighbourhoods most suitable which will result in maximum revenue for the stakeholder wishing to set up the gym.

Data

Based on the factors I discussed above, the following data will be required:

- The distance of each neighbourhood from the centre of the city.
- The presence of commercial office spaces in the neighbourhood.
- The number of gyms that exist in each neighbourhood.
- The time at which the maximum traffic will be observed for gyms.

The neighbourhoods have been identified from the source: https://cocl.us/new_york_dataset. All the features will be extracted from this data source.

The other data will be extracted using the Foursquare API. The location of each neighbourhood will be obtained from the GeoPy GeoCoder package in Python.