

Finding an optimal location for opening an exclusive gym in Manhattan

Data:

Following dataset were used in this project:

1. NYC census data to determine neighborhoods with high incomes. This dataset was used from NYC census portal and contain many attributes such as: county, total population, gender, race, income, employment status and more. In order to get neighborhood information, we will have to combine with census block dataset which contains latitude and longitude and from there we will be extracting neighborhoods.
2. NYPD arrest dataset to determine safest neighborhoods. This dataset was obtained from NYPD current data that contains information about arrest date, type of offence, perpetrator gender, race, age. In order to obtain information which neighborhoods in Manhattan are consider to be safe will be focused on occurrences of crimes per each neighborhood.
3. NYC borough and neighborhood data with Latitude and Longitude which we will need in order to find venues.
4. Venues data obtained from Foursquare API to determine most common venues in neighborhoods. Purpose of this dataset is to match venues in each neighborhood and find 10 most frequent venues which will help us to determine location for opening new gym in area.