Battle of Neighborhoods – How Health Conscious if your Neighborhood?

DEVDATTA KANTHE

MAR 30, 2020

Introduction

As we see the world change around us, one of the noticeable changes is how much health-conscious people have become. They are conscious about working out, eating healthy and staying fit. One of the facilities within a neighborhood that can indicate whether the inhabitants are taking their health seriously is the number of Gyms available AND whether Gyms are among the top 10 recommended venues in that neighborhood.

We are going to find a statistical method to evaluate every neighborhood within borough of Manhattan in New York City and use folium maps to visualize the density. Based on the results, we would try to speculate if we can find top 5 health-conscious neighborhoods within the borough of Manhattan in New York city.

So, let's find out How Health Conscious is your neighborhood?

Data Sources

Google Geocoding API for listing coordinates for Manhattan Neighborhood. - Google API is used for finding out coordinates based on addresses.

NYU Spatial Data Repository for listing neighborhoods in Manhattan - NYU Data contains list of Neighborhoods with coordinates that can be used to find recommended venues around Manhattan.

The venue details are scraped from **Foursquare** API - We would be using the explore API to figure out recommended venues based on inputs

Methodology

Ranking

Finding Top 10 Venues in all Neighborhoods and checking if Gym Categories are among them

Visualization

Plotting neighborhoods based on venue counts and ratio of Gyms to total venues

K-Means Clustering ML Algorithm

Cluster neighborhoods based on Gym Categories popular among them

Analysis

Ranking

Based on the data, we can derive that the popular Gym types are Gym/Fitness Centers (a combined **32** places of interest), Yoga Studio (**7** Venues) and Cycle Studio (albeit with just **2** Venues!).

However, filtering the data even further, we find that there are only a few neighborhoods with Gyms as top 3 venues.

Civic Center, Clinton, **Flatiron,** Marble Hill, Roosevelt Island, **Sutton Place,** Yorkville

Flatiron and **Sutton Place** are highlighted because Gyms are their **topmost common places (Rank = 1).** These neighborhoods have a higher priority for Gyms than other category types. **Can we claim them to be Manhattan's most Health Conscious Neighborhoods? Probably.**

Analysis contd...

Visualization

With multiple visualization such as bar charts and heatmaps, we are able to find the neighborhoods with higher ratio of gyms and higher frequency of gym categories.

Histogram of Ratio of Gym vs. Total Venues per Neighborhood

Total Gym Venues by Neighborhood
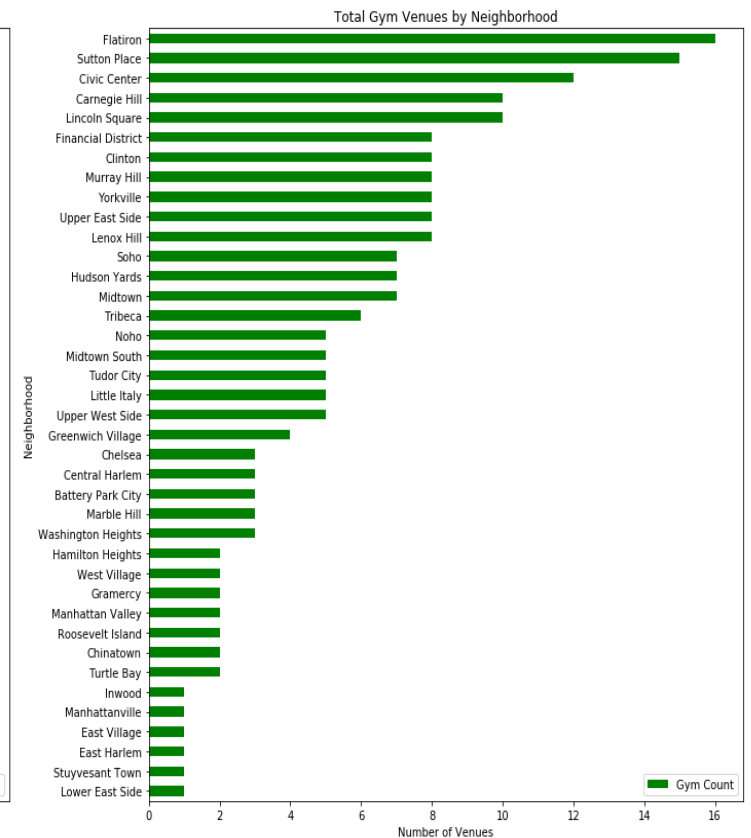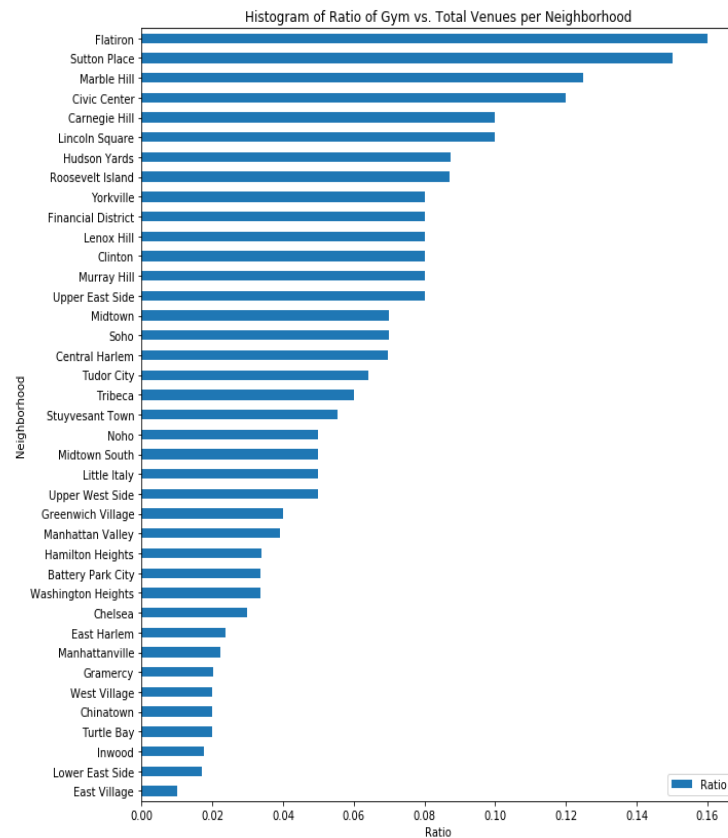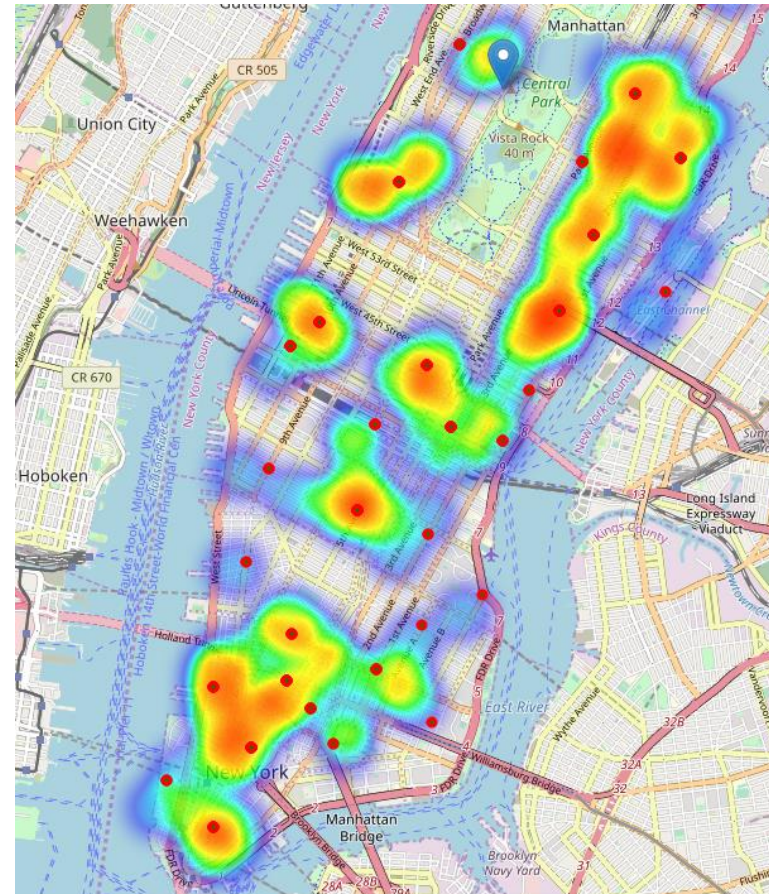
Analysis contd...

Visualization

With multiple visualization such as bar charts and heatmaps, we are able to find the neighborhoods with higher ratio of gyms and higher frequency of gym categories.

Analysis contd...

Top 2 Neighborhoods with highest ration of Gyms to Total Venues are Flatiron and Sutton Place.

Interestingly, we would notice that although certain neighborhoods have a higher number of Gyms, the overall frequency per total venues is lower. Example, Civic Center and Carnegie Hill. Conversely, Marble Hill and Hudson Yards, scoring lower on overall Gym venues, scores higher on the Ratio scale. Especially, Marble Hill that is very low on the number of Gyms but in top 3 based on ratio.

Based on the heatmap, we can observe a few Neighborhoods have a higher density of Gyms. Prominent, as observed before, are Sutton Place, Flatiron, Civic Center, Carnegie Hill, Financial District.

Evidently, places such as Marble Hill, show lower concentration on the heat map due to lesser venues. However, ratio of Gyms vs. All venues is higher for Marble Hill.

Clustering

Using K-Means Clustering, all neighborhoods are classified into clusters. With K=10 (using elbow point method), clusters of neighborhoods can be mapped on a plot.
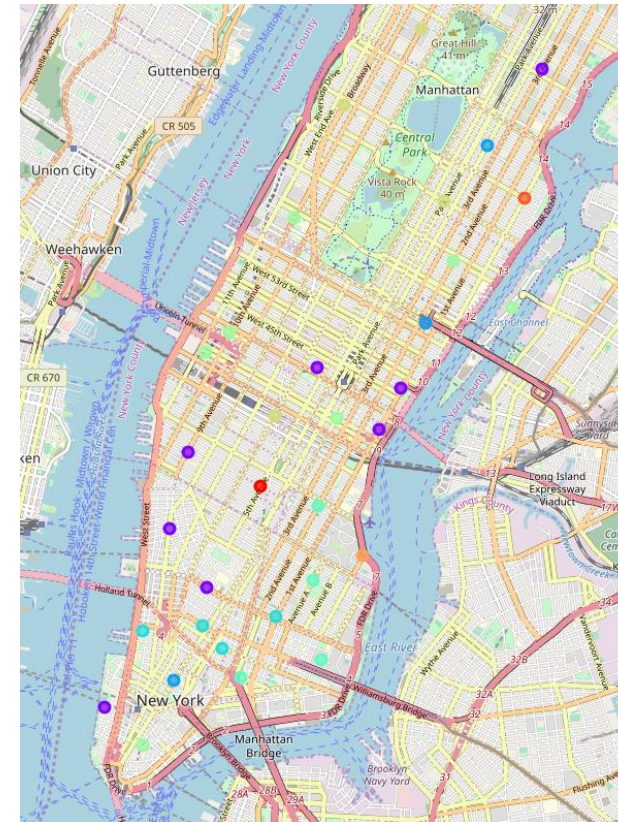
Clustering contd...

After analyzing all 10 clusters, it is observed that neighborhoods are clustered based on top 3 venue categories which in our case were Gym Categories. Clusters 1, 5 and 6 were largest clusters by number and had Gym/Fitness Center, Gym, Yoga Studio and Cycle Studio in top 3 categories. This is in line with the top 10 analysis performed earlier, where these very categories were among the most popular.

Out[36]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Cluster Labels	Most Common Venue 1	Most Common Venue 2	Most Common Venue 3	Most Common Venue 4	Most Common Venue 5	Most Common Venue 6	Most Common Venue 7	Most Common Venue 8	Most Common Venue 9	Most Common Venue 10
374	Central Harlem	40.815976	-73.943211	6	Gym / Fitness Center	Gym	Cycle Studio	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Climbing Gym
659	Lenox Hill	40.768113	-73.958860	6	Gym / Fitness Center	Gym	Cycle Studio	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Climbing Gym
759	Roosevelt Island	40.762160	-73.949168	6	Gym / Fitness Center	Gym	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Cycle Studio	Climbing Gym
882	Lincoln Square	40.773529	-73.985338	6	Gym / Fitness Center	Gym	Cycle Studio	Yoga Studio	Climbing Gym	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool
982	Clinton	40.759101	-73.996119	6	Gym / Fitness Center	Gym	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Cycle Studio	Climbing Gym
1182	Murray Hill	40.748303	-73.978332	6	Gym / Fitness Center	Gym	Martial Arts Dojo	Boxing Gym	Yoga Studio	Weight Loss Center	Pilates Studio	Gymnastics Gym	Gym Pool	Cycle Studio
2321	Financial District	40.707107	-74.010665	6	Gym	Gym / Fitness Center	Cycle Studio	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Climbing Gym
3217	Hudson Yards	40.756658	-74.000111	6	Gym / Fitness Center	Gym	Cycle Studio	Yoga Studio	Weight Loss Center	Pilates Studio	Martial Arts Dojo	Gymnastics Gym	Gym Pool	Climbing Gym

Conclusion

If you are part of the **Flatiron, Sutton Place** then most definitely you are part of a health-conscious neighborhood in Manhattan, NY. However, in addition to these two, there are other neighborhoods that qualify as health-conscious basis other criteria listed in the Discussion section above.

All in all, out of the 40 neighborhoods in Manhattan, about 10 have qualified as health-conscious in our study. Overall, **25% of all neighborhoods in Manhattan can be considered health conscious based on Gym/Fitness Center related data available as of today.**

This analysis is essential so as to help plan neighborhoods, keeping in mind, the health requirements of inhabitants. It is essential for people to get basic needs within the area but fitness related facilities are becoming essential day-by-day. Due to the health hazards, we face, everyday due to our lifestyles, staying fit has become a requirement, a need and an expectation. After all, a **health conscious neighborhood** has the potential to become a **healthy neighborhood.**

Assumptions

Factors such as day-time population of neighborhood along with other related demographics are excluded from the analysis. These are critical for a holistic analysis.

It is assumed that the facilities available are utilized by people who are living in the neighborhood. It is highly likely that Manhattan being a business district, people coming in to work, may be using the facilities as well.

By definition, healthy and health conscious are not always synonymous.

FourSquare API data is assumed to be complete. There may be other sources providing extensive and expansive data.