## Problem Description

The customer of the battle of neighborhoods was not satisfied with the first results. He claimed that that the similarity in the quantity of venues is meaningless because he care about the quality of the venues. He said that he went to the theoretically similar neighborhoods and they truly have somewhat close similar quantities of venues around them, but one turned to be out more like a ghetto and the other is more modern and nicer place. Thus, the client rejected our previous work and asked for a new way to solve this problem.

## Needed Data

This project will require information about the neighborhood of Manhattan and the likes of the customers for each venue and the number of likes. This information is available freely from Venue's likes. The neighborhood will be clustered based on the likes that is given to each venue. This will emphasize the quality of the venue to be considered when clustering the neighborhoods. The screenshot below show the main data that are going to be used mainly the count of likes for each venue.

A screenshot of a social media post

Description automatically generated

The venue id will be fetched using Foursquare API for Venue explore as shown in the screen shot

A screenshot of a social media post

Description automatically generated

Methodology

In order to emphasize the quality of the venue some rules has to be made to distinguish the venues from each other. After fetching the surrounding venue for each neighborhood the number of likes for each venues are fetched.

The distribution of likes are examined using histogram as shown in the plot:

A picture containing screen, sitting, people, standing

Description automatically generated

The plot support our assumption that Foursquare explore is returning so many venues that have no likes. The plot shows that most of the venus have zero likes. To fix the complaint of our customer we apply the following rule:

1. If the venue has less than 25 or equal to likes then this venue is equivelant to one venue
2. If the venue has more than 25 like less than or equal to 50 then this venue is equivelant to two venue of same category
3. If the venue has more than 50 like less than or equalt to 100 then this venue is equivelant to three venue of same category
4. If the venue has more than 100 likes less than or equalt to 200 then this venue is equivelant to five venue of same category
5. If the venue has more than 200 likes less than or equalt to 300 then this venue is equivelant to 7 venue of same category
6. If the venue has more than 300 likes less than or equalt to 500 then this venue is equivelant to ten venue of same category

Results

After applying the rule above we got new results , hopefully, more convincing for our client. The clustering took into account the number of likes for each venues and emphasized the similarity based on quality in addition to the quantity.

A close up of a map

Description automatically generated

The screen shoot show that the clustering is different than the one that we have submitted first time. And further explanation will be analyzed in the next section .

Discussion

The results seems to reflect the quality of venue (evaluated by the number of likes) by dividing the neighborhoods into five clusters. The ordinary common ones are neighborhoods with variety of common venues (cafes, restaurants, pizzas, pub, bars …etc ). It is hard to conclude very specific character for these neighborhoods. But this time we gave the customer better list to choose the location as the top venues are sorted based on the quality also.

The number of clusters are increased to make more sense of the data.

The Battle of Nighborhoos V1 lack interpretation of the results and advice for customer on how to act on the basis of these results thus in V3 we add the following analysis :

* For ordinary customers who would like to have regular neighborhoods with sufficient varieties of restaurants and activities we recommend the following neighborhoods

A close up of a piece of paper

Description automatically generated

Comparing V3 with V1 in appendix we see that the recommended venues have been changed mostly in all the neighborhoods. For example Marble Hill most common quality venue become donut shop whereas it was sandwich place considering the quantity alone. However, in China Town the Chinese restaurant stay the most common venue in both cases which is reasonable outcome.

* For those who care the most about the atmosphere of the neighborhood with beautiful sceneries ,memorial sites and park there are not much options. Battery Park City is the best fit. The most quality venues has changed from gym, coffe shop , hotel and blaza to the venues the people appreciate the most in this neighborhood .

A screenshot of a social media post

Description automatically generated

* Our customer care the most about parks, books and coffee shops? Midtown and Morningsight Height are the best candidates for such customers

A screenshot of a cell phone

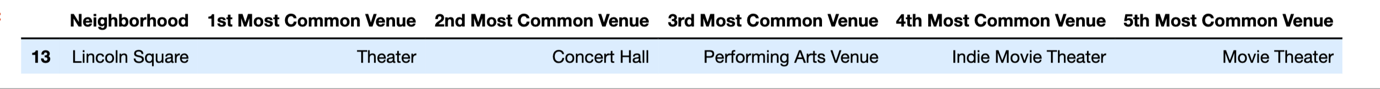
Description automatically generated

* The Mediterranean touches are visible strongly in particular neighborhoods where every neighborhood of these has Italy/Italian as its best venue

A screenshot of a cell phone

Description automatically generated

* The hippies, indies artists and designers have their fit neighborhoods as well. For those people we recommend Lincoln Square



Conclusion

An updated version of the battle of the neighborhoods was made to consider the quality of the venues in clustering the neighborhoods in addition to the quantity. The results show really promising and sound logic. The logic can be updated further based on our customers input. The rules that made the foundation for this analysis can be tailored in addition to the methodology.

Note: This study is conducted after failing to get the data for the other project idea to cluster neighborhood based on genders and ages preferences. The reason was that accessing this data from stats require to be the owner of the venue thus the idea was altered to this project.

Appendix A Results of V1

A close up of text on a white background

Description automatically generated