Selecting a Land from Manhattan Borough for a Luxury Housing Complex

Introduction

Buying a land which is suitable for building a housing complex is somewhat a tricky task. You cannot just go ahead and buy an empty land. Obviously, this a huge investment and the investor(s) expects a significant return. It’s true that factors like apartment size, number of bedrooms and bathrooms, number of parking slots a tenant gets, etc. play a key role in the marketability of an apartment. However, all of them are internal factors which can be controlled by the project team. Not only internal factors but also external factors such as infrastructure facilities (e.g. hospitals, schools, supermarkets, play grounds, restaurants, etc.) adds value to such projects. Therefore, it is advantageous for the project team to pick a land which is surrounded with at least some of these external facilities.

**Problem**

Assumption – All neighborhoods in Manhattan area have gotten at least one land which can be acquired for this project.

In this exercise, I will be looking at the infrastructure facilities available in each neighborhood (within a 500m radius). Once the facilities are identified, I will give points for each facility based on the importance of the facility (e.g. hospitals get the highest points). After calculating the total marks of each neighborhood, I will recommend the client to buy the land which scores the highest marks.

Data and Data Sources

New York City comprises of 5 boroughs and 306 neighborhoods. I have used the dataset given in one of the previous assignments which contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

The dataset has been downloaded from the following location.

<https://geo.nyu.edu/catalog/nyu_2451_34572>

However, since my task is based on ‘Manhattan’, I have filtered out all the neighbourhoods which are located inside Manhattan.

Methodology

As mentioned above, first, I have dropped the neighbourhoods located outside Manhattan borough. Then, using Foursquare API, I have queried the top 100 venues of each neighbourhood. Thanks to Foursquare service, I have managed to retrieve the venue category of each venue. My objective was to identify the most important venue categories which would enhance or add value to the project. As a result, the following set of venue categories has been identified as the most important ones.

* Doctor's Office, Medical Center
* School, High School, General College & University
* Pharmacy, Drugstore
* Supermarket, Grocery Store
* Clothing Store
* Bus Station, Bus Stop, Bus Line, Metro Station
* Department Store, Discount Store, Shopping Mall, Convenience Store, Electronics Store
* Bank
* Gas Station
* Gym, Gym / Fitness Center, Weight Loss Center

As you can see, I have also grouped the similar venue categories together.

Then, I have assigned a hypothetical score for each venue category based on the importance of the category. The categories which carries the same score have been assigned to a unique cluster (**Clustering**).

|  |  |  |
| --- | --- | --- |
| Cluster | Venue Category/Categories | Score |
| 1 | Doctor's Office, Medical Center | 10 |
| 2 | School, High School, General College & University | 9 |
| 3 | Pharmacy, Drugstore | 8 |
| 4 | Supermarket, Grocery Store | 7 |
| 5 | Clothing Store | 6 |
| 6 | Bus Station, Bus Stop, Bus Line, Metro Station | 5 |
| 7 | Department Store, Discount Store, Shopping Mall, Convenience Store, Electronics Store | 4 |
| 8 | Bank | 3 |
| 9 | Gas Station | 2 |
| 10 | Gym, Gym / Fitness Center, Weight Loss Center | 1 |

Then, I have calculated the aggregated score of each neighborhood.

For example, as per the results, there are 24 venues within the 500m radius in Marble Hills. Let’s assume that the following venues are included in those 24 venues.

|  |  |  |
| --- | --- | --- |
| Venue | Frequency | Score |
| Doctor's Office, Medical Center | 3 | 30 (3\*10) |
| Bank | 4 | 12 (4\*3) |
| Clothing Store | 2 | 12 (2\*6) |

As per this example, the aggregated score for Marble Hills is 54. Likewise, I have done the same for all the neighborhoods located in Manhattan. Please refer to the results below.

Results

|  |  |
| --- | --- |
| Neighborhood | Overall Score |
| Marble Hill | 21 |
| Chinatown | 16 |
| Washington Heights | 65 |
| Inwood | 27 |
| Hamilton Heights | 21 |
| Manhattanville | 15 |
| Central Harlem | 2 |
| East Harlem | 18 |
| Upper East Side | 27 |
| Yorkville | 18 |
| Lenox Hill | 13 |
| Roosevelt Island | 27 |
| Upper West Side | 20 |
| Lincoln Square | 38 |
| Clinton | 21 |
| Midtown | 46 |
| Murray Hill | 13 |
| Chelsea | 13 |
| Greenwich Village | 36 |
| East Village | 0 |
| Lower East Side | 14 |
| Tribeca | 18 |
| Little Italy | 26 |
| Soho | 65 |
| West Village | 0 |
| Manhattan Valley | 0 |
| Morningside Heights | 22 |
| Gramercy | 33 |
| Battery Park City | 46 |
| Financial District | 27 |
| Carnegie Hill | 40 |
| Noho | 15 |
| Civic Center | 23 |
| Midtown South | 35 |
| Sutton Place | 34 |
| Turtle Bay | 9 |
| Tudor City | 10 |
| Stuyvesant Town | 3 |
| Flatiron | 39 |
| Hudson Yards | 32 |

As per the results, Washington Heights and Soho have achieved the highest overall score of 65. Therefore, it can be decided that the investors can use one of the lands in Washington Heights or Soho to build the apartment based on the given information. However, please note that there are other external factors such as land value, quality of the neighborhood, etc. (which are not considered in this exercise) that need to be considered when making a decision since the investors are given with two options.

Discussion and Conclusion

During this exercise, I have found some interesting facts. If you look at the above table, you can see that there are three Neighborhoods which haven’t scored a single point. They are East Village, West Village and Manhattan Valley (highlighted in Red). This indicates that there is not even a hospital (Doctor's Office, Medical Center) located within a 500m radius in these three Neighbourhoods which is quite scary. Therefore, in my opinion, these three villages are not suitable for living at all. May be someone from USA can enlighten me on the reason behind this.

Then, I’ve filtered out the Neighbourhoods which have a doctor’s office or a medical centre and I’ve gotten only two results (please refer to the GITHUB code).

A doctor’s office is located in Financial district and a medical centre is located in Civic Center. Out of those 39 neighbourhoods, only two have a quick access to medical facilities.

Then, I’ve looked for the statistics of schools, high schools and universities. To my surprise, there are only 5 schools, 2 high schools and a ‘general college & university’ - Columbia University Club of New York. When it comes to drugstores and pharmacies, there are only 10 pharmacies and a drugstore.

**Conclusion**

Then, I’ve started analysing the two Neighbourhoods which got the highest score (please refer to the #53 and #54 of the code). As you can see, out of the 65 points scored by Soho, 54 points came from 9 clothing stores (more than 83%). Furthermore, it doesn’t have any of the venues mentioned below.

* Doctor's Office, Medical Centre
* School, High School, General College & University
* Pharmacy, Drugstore

**However, when it comes to Washington Heights, there is a pharmacy, grocery stores, supermarkets, banks, gyms, fitness centres and clothing stores (a wide variety of places than Soho including a pharmacy). Hence, I’d recommend the investor to buy the land in Washington Heights.**