## 1. Problem Statement and Background

Manhattan is an epicenter of business, financial, and cultural matters. Easily one of the most well-known cities, it draws a lot of tourists from all over the world. In Manhattan, hotels are spread far and wide, accommodating anyone who is in need of a place to stay. These hotels used to be the only option for travelers until the rise of Airbnbs. The continuous increase in the popularity and use cases of the Internet has allowed for an online marketplace for arranging and offering lodging. While the cost of staying in an Airbnb is drastically lower than staying in a hotel, the number of choices out there becomes daunting. To resolve this, for someone who is looking for a place to stay in Manhattan to get a bang for their back, an in-depth analysis of the available Airbnbs will be undertaken to recommend several places.

## 2. Data

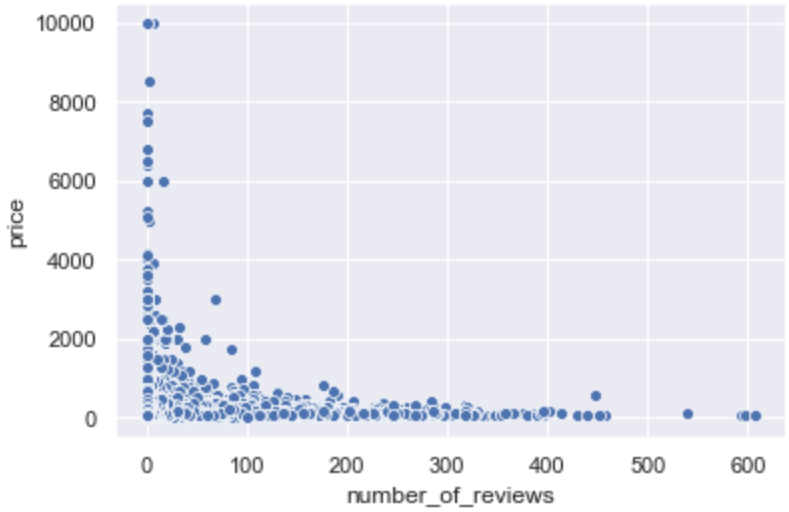
Leveraging location data, recommendations will be made as to which Airbnb rental will yield the most bang for buck. The data to be used will come from two main sources: Foursquare and the New York City Airbnb Open Data. Foursquare data will contain establishments that are within the vicinity of the Airbnb rental. The more establishments, the better as there will be more choices for entertainment and enjoyment. To augment this information, the NYC Airbnb Open Data contains details on the rentals such as price, reviews, location, amenities, and required length of stay. These restrictions may or may not affect a customer's choice but nonetheless it is important to analyze. By merging these two data sets and leveraging insights gathered, the customer can then make an informed decision on where to staying Manhattan.

## 3. Methodology

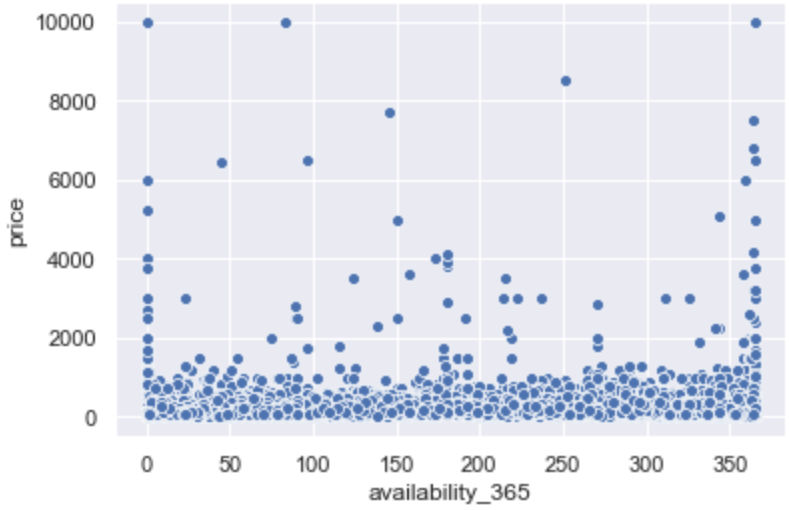
After collecting the data, it was processed and profiled. Using relevant features based on domain knowledge, it was then plugged into a clustering algorithm. The clustering algorithm used the K-Means method to group similar Airbnb rentals together.

## 4. Results and Discussion

First, data was collected from relevant resources. It was then trimmed down to only include data on Manhattan. A first pass at correlation analysis showed very weak relationships between the features. There was no linear relationship visible as well even in the visualizations.




There was no visible linear relationship between the number of reviews and price. This indicated that whenever a change in one variable occurs, there is no change in the other or it is considered inconsequential.




Similarly, availability was not related to price as well. The lack of distinct relationships between variables shows that prices of Manhattan Airbnbs are not easily influenced by restricting its

ability to take in customers. This is probably because a place to stay in Manhattan is sought after anyway.

There were five clusters that emerged after fitting a clustering algorithm to the data. Each having their own distinct characteristics. The cheaper choices but reviewed an acceptable number of times are in Cluster 4. These are also available most of the year and only require you to stay a night or two. This is advisable for those who are only in Manhattan for a short period of time. Unfortunately, your choices are limited to a shared room. Choosing to rent a private room will be more expensive. If the budget is slightly higher than average, Cluster 0 is your friend. This cluster offers a wide variety of choices ranging across different price budgets and room preferences; there are even entire apartments available at this price. These rentals get a decent number of reviews, but are rarely available. Required minimum nights of stay is still reasonable from one to four nights. Cluster 1 is fairly similar to Cluster 0 and can accommodate the same preferences except price. The prices of rentals in Cluster 1 can be higher with the same characteristics of rentals from Cluster 0. This is because there are more rentals for entire apartments in Cluster 1. Lastly, if luxurious is the preference, then Cluster 2 is the best choice. The prices of rentals in this cluster are really expensive, but you get premium locations: Tribeca, Battery Park City, Flatiron District, and SoHo. Really, the location is what you'll be paying for in these rentals.

## 5. Conclusion

Should you ever find yourself in need of accommodations in Manhattan, I have several recommendations depending on what you're looking for and of course, budget. The cheaper choices but reviewed an acceptable number of times are in Cluster 4. These are also available most of the year and only require you to stay a night or two. This is advisable for those who are only in Manhattan for a short period of time. Unfortunately, your choices are limited to a shared room. Choosing to rent a private room will be more expensive. If the budget is slightly higher than average, Cluster 0 is your friend. This cluster offers a wide variety of choices ranging across different price budgets and room preferences; there are even entire apartments available at this

price. These rentals get a decent number of reviews, but are rarely available. Required minimum nights of stay is still reasonable from one to four nights. Cluster 1 is fairly similar to Cluster 0 and can accommodate the same preferences except price. The prices of rentals in Cluster 1 can be higher with the same characteristics of rentals from Cluster 0. This is because there are more rentals for entire apartments in Cluster 1. Lastly, if luxurious is the preference, then Cluster 2 is the best choice. The prices of rentals in this cluster are really expensive, but you get premium locations: Tribeca, Battery Park City, Flatiron District, and SoHo. Really, the location is what you'll be paying for in these rentals.