

SINDUJA SEKAR

AirBNB accommodation in New York City

Dec 2019

Predicting and studying the factors influencing the choice of an airbnb accomodation

- The selection of a good bob location with proper research can avoid disappointments and hassles during the holidays.
- Choosing the right location in a busy city like Newyork should be of utmost importance and has to depend on features like proximity to places of visit, type of rooms, price and more.
- Reviews left by previous user about their first hand experience is also very valuable in making the decision .

Data Acquisition and Cleaning

- The data for this project is obtained from the kaggles data set 2019 and also the latitude and longitude for the city's centre is found using the four square API.
- The dataset is further combed and cleaned for null and duplicate records and unwanted columns are deleted .
- Values are encoded and datatypes are modified before applying the machine learning models to the dataset.
- The cleaned dataset contains 48852 records after cleaning the data set.

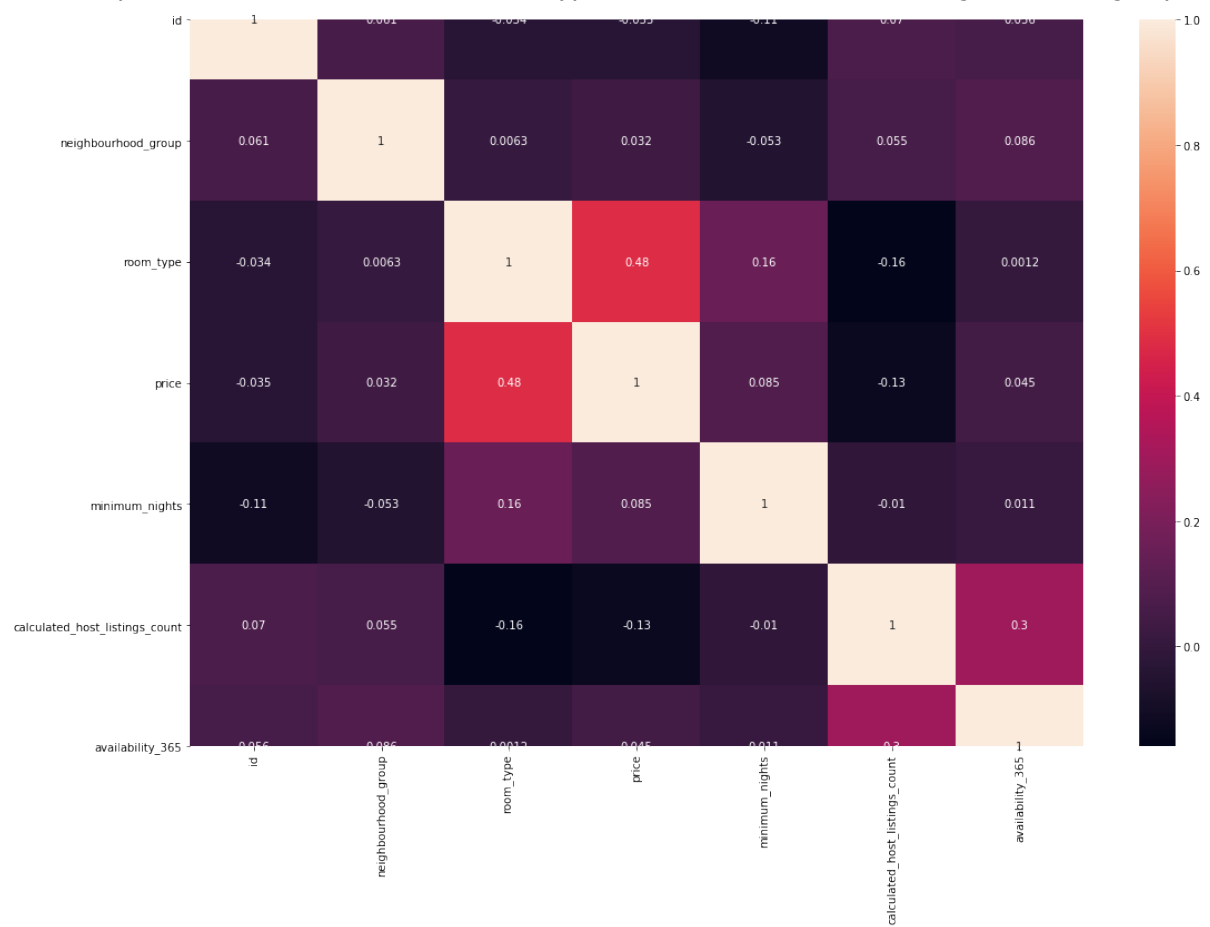
Feature selection

The following features are identified as the independent features to predict the dependent features that is the price for the stay at a bnb location.

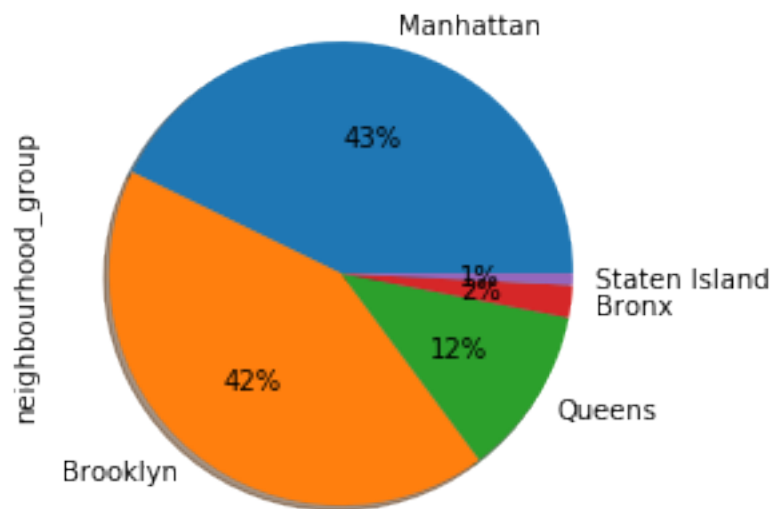
- Room type and availability
- Neighbourhood and neighbourhood group
- Days of stay
- Number of reviews in a month.

Identifying relationships between features

Correlation maps and scatter plots are used to visualise the effect of various independent features on the dependent features, the distribution of type of rooms across the the 5 neighbourhood groups



is identified using the below visualization.



Regression models and decision tree models are used to predict the price using the independent features as price is the most important features when it comes to choosing a bnb location.

Conclusion and future directions

While the process give a close to accurate prediction for the price based on the features selected. The prediction can be further improved by identifying dinning options around the bnb location. Also the latitude and longitude data can be further used to calculate the distance from the city centre to the given location.