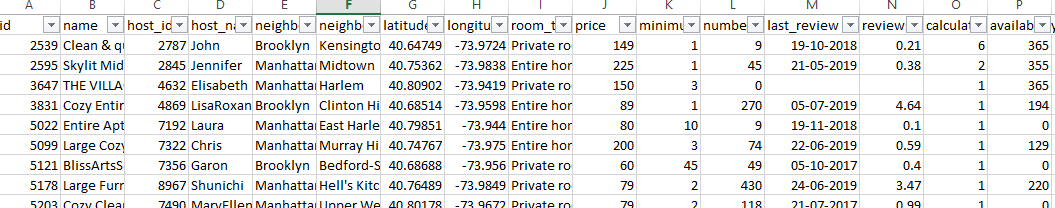
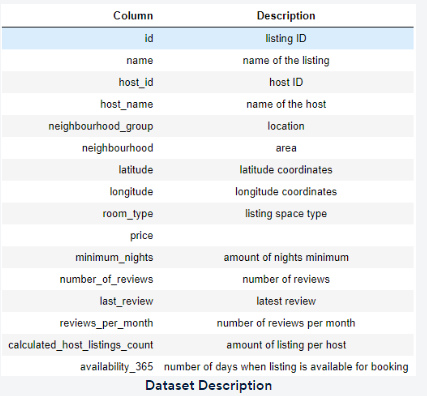
DATA METHODOLOGY

* Used Jupiter Python to visualize data from NYC Airbnb in order to obtain accurate insights.
* Data set “AB\_NYC\_2019” is provided in the form of excel with 48895 rows and 16 columns.



* Below are the dataset description of AB\_NYC\_2019 dataset

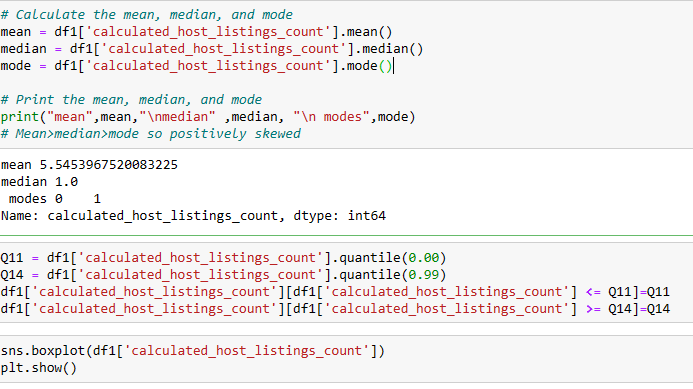


* Data Understanding and Preparation
* Imported necessary libraries
* **Load the dataset**
* **Data cleaning:**
  + Checking missing values in the given data set.
  + Handled null values only in “last\_review” by dropping rows.
  + Columns like id, name and host name have been imputed assuming name can change and id doesn’t add significance in further analysis as it can be become cumbersome for analyzing in this case study.
  + Outliers in numerical columns such as

1. minimum\_nights
2. number\_of\_reviews
3. reviews\_per\_month
4. calculates\_hos\_listings\_count
5. availability\_365 are handled

Except price.

* Skewness was checked for each column and appropriate measures were taken, such as dropping columns with high skewness.



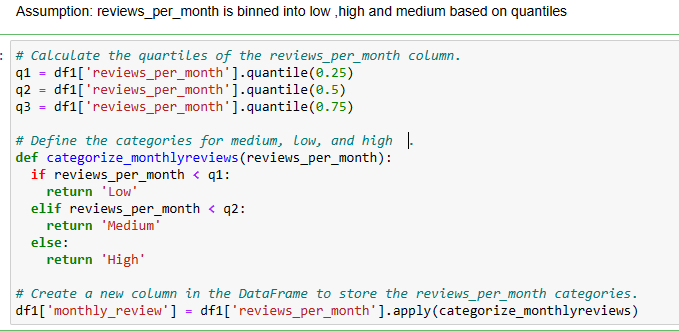
* Added new column Revenue.



* Added new column price\_category with low, medium and high category based on quantile(0.25,0.50,0.75) of price



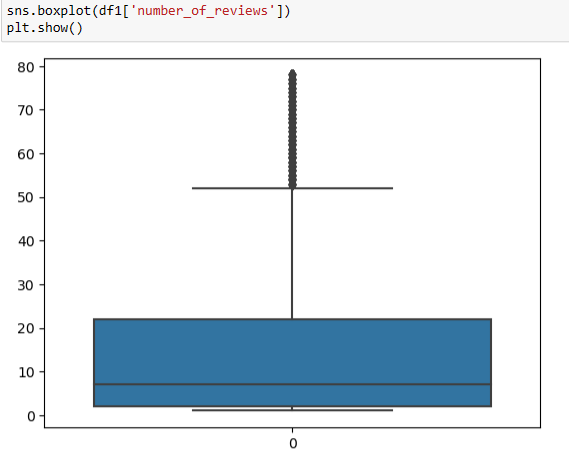
* Added new column monthly\_review with low, medium and high category based on quantile(0.25,0.50,0.75) of review\_per\_month





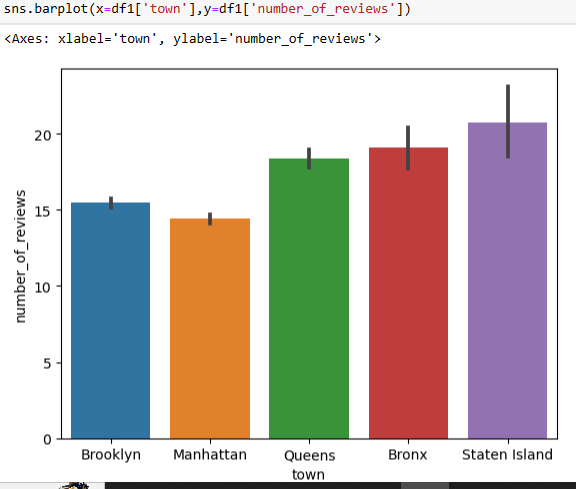
* Renaming neighborhood to area and neighbourhood\_group i.e. borough of New York city to town for easy analysis
* Assumption: reviews\_per\_month is binned into low ,high and medium based on quantiles
* Exploratory Data analysis
  + Univariate analysis

Example:



* + Bivariate analysis

Example:



* + Multivariate analysis done

Example:

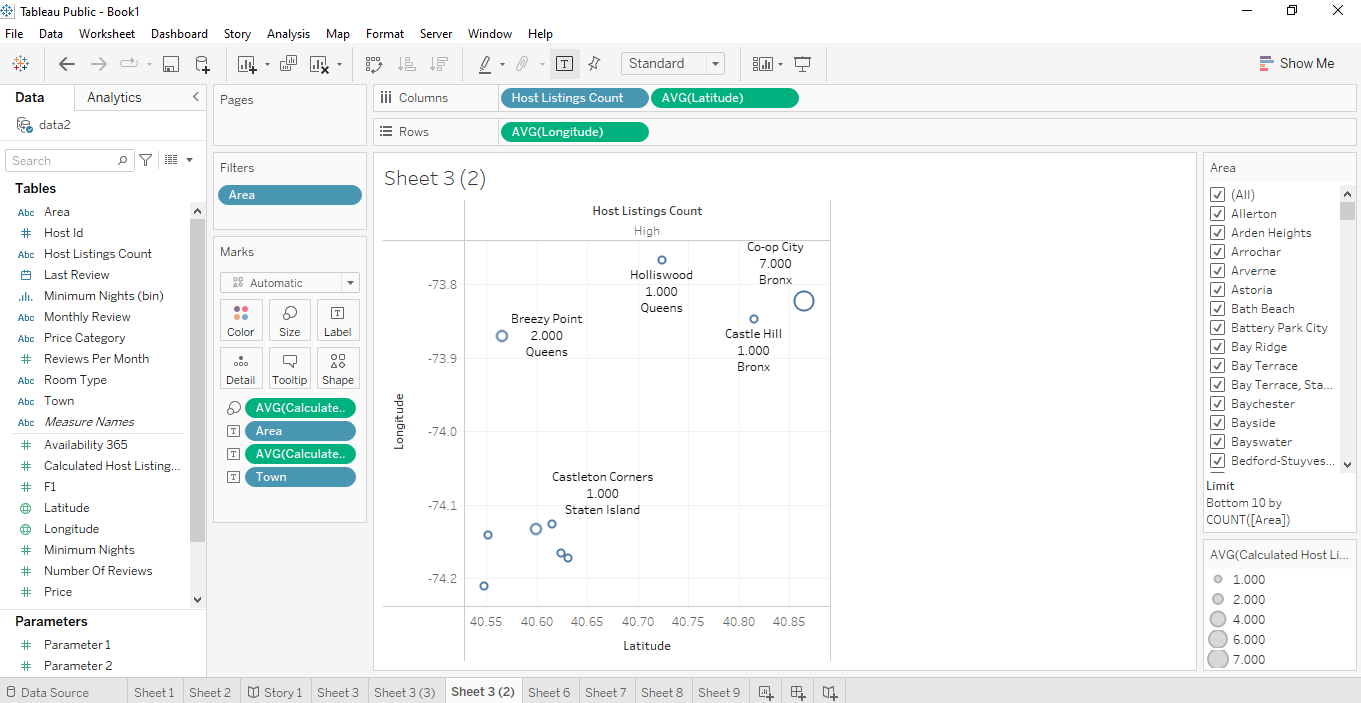
* 



* Cleaned data is saved under data2.csv



* Did further analysis in tableau desktop version with data2.csv cleaned data
* Imported data2 text file in tableau
* Following is one of the examples of analysis done is tableau



* Pyramid principle is used in preparing PPT’s