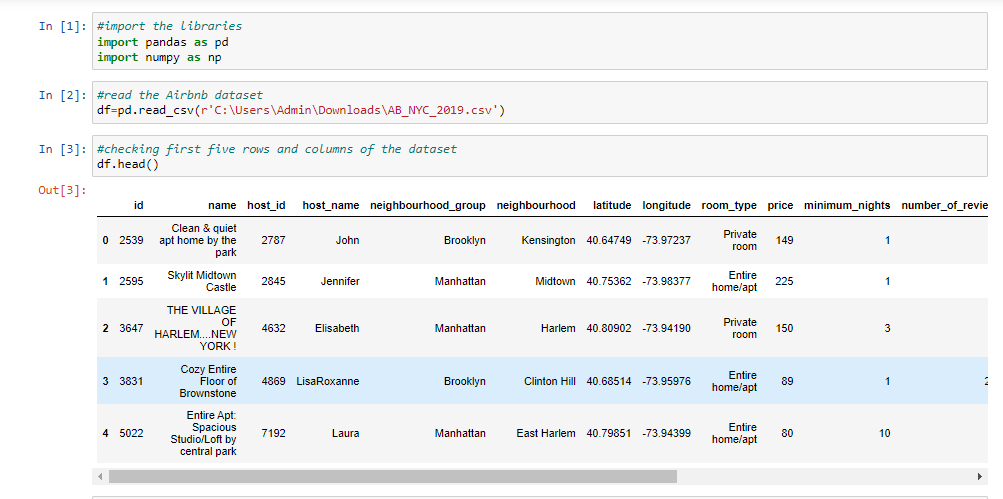
**Storytelling Case Study : AIRBNB ,NYC IIIT-B and Upgrad Vyankatesh kale , Laveena achotani and Kavitha choudhary**

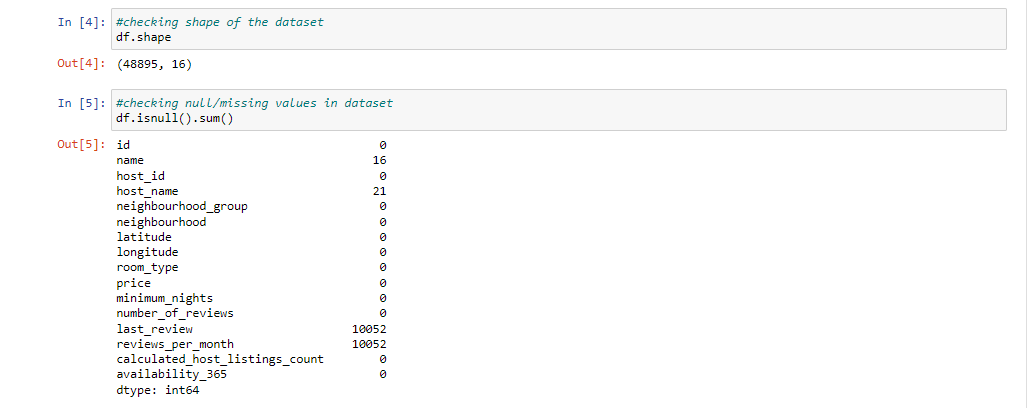
**Methodology Document PPT 1 -**

In the case study we have used Jupyter notebook to perform initial analysis of the data and Tableau for data analysis and visualization.

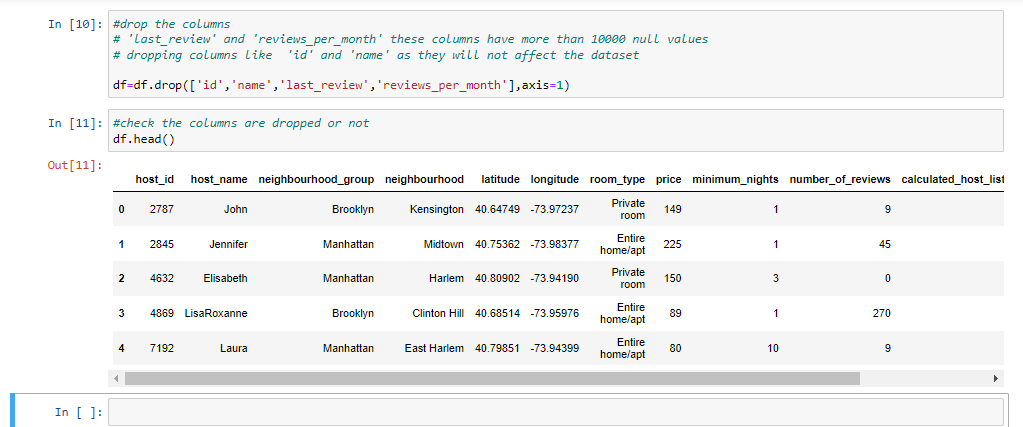
**Initial Analysis using Jupyter Notebook :** Data Set Used: AB\_NYC\_2019.csv

**Number of Rows:** 48895

**Number of Columns:** 16



* The dataset contains 48895 rows and 16 columns
* Now we have to find whether there are any null values in the dataset



* We removed the columns like Id, Name, Last Review and review per month which was not giving much information

**Approach :**

• Checked the Duplicate rows in our dataset and no duplicate data were found.

• Checked the Null Values in our dataset. Columns like name, host-name, last review and review-per-month have null values.

• We’ve dropped the column name as missing values are less and dropping it won’t have significant impact on analysis.

• Identified and review outliers using the tableau.

**Data Analysis and Visualizations using Tableau:**

We have used tableau to visualize the data for the assignment. Below are the detailed steps used for each visualization.

1. **Top 10 Host:**
   * We identified the top 10 Host Ids, Host Name with count of Host Ids using the tree map.



1. **Preferred Room type with respect to Neighbourhood group:** 
   * We created a pie chart for understanding the percentage of room type preferred w r t neighborhood group
   * We added Room Type to the colors Marks card to highlight the different Room Type in different colors and count of Host Id to the size
2. **For Variance of price with Neighbourhood Groups:** 
   * We used a box and whisker’s plot with Neighbourhood Groups in Columns and Price in Rows.
   * We changed the Price from a Sum Measure to the median measure
3. **Average price of Neighbourhood groups:**
   * We created a bubble chart with Neighbourhood Groups in Columns and Price column in Rows.
   * We added the Neighbourhood Groups to the colors Marks card to highlight the different neighbourhood Groups in different colors. Also Put Avg price in Label.
4. **Customer Booking w r t minimum nights:** 
   * We created the bin for Minimum nights as shown below.
   * The bins were used to display the distribution of minimum nights based on the number of ids booked for each neighbourhood group.



1. **Popular Neighborhoods:** 
   * We took neighbourhood in rows and sum of reviews in column and took neighbourhood groups in color.
   * We used filter to show Top 20 neighbours as per the sum of reviews.
2. **Neighbourhood vs Availability:** 
   * We created a dual axis chart using bar chart for availability 365 and line chart for price for top 10 neighbourhood group sorted by price.

**Methodology Document PPT 2:**

1. **Room type with respect to Neighborhood group:**
   * We created a pie chart for understanding the percentage of room type preferred w r t neighborhood group
   * We added Room Type to the colors Marks card to highlight the different Room Type in different colors and count of Host Id to the size
2. **Customer Booking with respect to minimum nights**:
   * We created the bin for Minimum nights as shown below.
   * The bins were used to display the distribution of minimum nights based on the number of ids booked for each neighborhood group.



1. **Neighborhood vs Availability:**
   * We created a dual axis chart using bar chart for availability 365 and line chart for price for top 10 neighborhood group sorted by price.
2. **Price range preferred by Customers:**
   * We have taken pricing preference based on volume of bookings done in a price range and no of Ids to create a bar chart. We have created bin for Price column with interval of $20.
3. **Understanding Price variation w.r.t Room Type & Neighbourhood:**
   * We created Highlights Table chat by taking Room Type in rows & Neighbourhood Group in column. We took the average price in color Marks card to highlight the different Room Type in different colors.
4. **Price variation w r t Geography:**
   * We used Geo location chart to plot neighbourhood, neighbourhood Group in map to show case the variation of prices across.
5. **Popular Neighborhoods:**
   * We took neighbourhood in rows and sum of reviews in column and took neighbourhood groups in color.
   * We used filter to show Top 20 neighbours as per the sum of reviews.
6. **Tools used:**
   * Data cleaning and preparation: Jupyter notebook – Python
   * Visualization and analysis: Tableau
   * Data Storytelling: Microsoft PPT