Methodology Airbnb

❖ Exploratory Data Analysis:

- Checked for the Null values/Missing values in the dataset. Missing values are present in the “name, host_name, last_reviews, reviews_per_month columns. We dropped the column last_review & impute missing values in reviews_per_month as 0.
- Checked the outliers in the dataset.
- There are 11 properties which have zero price. Values might not have been entered or might be an error. Binning of price ranges to be done to check the value to be imputed for the properties which have zero value. We will also have the null values present in the host_name and name columns as they are not required for now.
- The price range is heavily skewed with values mostly between 0 to 2000. Impute median values on the 11 properties which have zero price.
- Cleaned dataset then converted to csv for further tableau analysis. “data.to_csv(‘air_bnb.csv’)”.

❖ Data Analysis :

- Try to analyse the data using different columns based on their price, availability 365, minimum nights, and customer reviews.

❖ Conclusion after analyzing the data :

- We discovered that people like visiting New York's centre, from which they can enjoy the city's beauty.
- The number of shared room listings is limited, but their average price is low and availability is high.
- Number of reviews and reviews per month are more at less price than the higher price as there is less chance of people going for a high price room.
- Manhattan and Brooklyn Neighborhood groups are quite expensive.
- People are drawn to the host Blueground and stay for longer periods of time.
- Minimum number of nights to stay reduces with increase in price.
- Pay attention to popular areas like Manhattan and Brooklyn where people are interested.