**Exploratory Data Analysis on Airbnb Dataset**

|  |
| --- |
| **Name :** ANKIT RAI  **Email Id**: [Ankitskrai@gmail.com](mailto:Ankitskrai@gmail.com) |
| **PROJECT APPLICATION:**   * Data Wrangling – Handled Null Values, Created new columns of day month years * Plot Distribution of numeric features * Plot correlation matrix * Plot location, analysis by variation of price, reviews and availabilities * Analysis against Neighborhood Group * Analysis against occupancy * Analysis against reviews * Analysis against minimum nights to stay |
| **GitHub Repo & Drive Link:** |
| **Project Summary** |
| Since 2007, Airbnb has been an American company. It is an online marketplace that connects people who wish to rent out their homes with customers who are looking for lodging in certain locations.  We have our dataset from AirBnB in New York. New York is one of the most expensive cities in the United States. We would like to do an in-depth investigation of one of the world's most densely populated cities. Our dataset is rich in features, including location with coordinates,  prices, host name, room kinds, and availability throughout the season. We attempted to numerical characteristics, as well as performed univariate and bivariate analysis with numerous dependencies. Though we did not include outliers in our research, we did focus on neighbourhood groups or big areas of New York rather than local neighbourhoods. We were unable to undertake sentiment analysis or property quality research owing to a lack of data, but we can deduce that Manhattan and Brooklyn are among the most costly districts of New York, where people live a rich lifestyle and do not want sharing rooms, even if they are the cheapest.  Also, while location has a strong influence on property prices, it does not guarantee that a house in a desirable location will be occupied for the majority of the seasonextract information from these attributes such as the most costly places to reside in New York, whether location varies with occupancy rate, what style of room people like, and whether there is a peak season for renting. When we can track a rise in property values or occupancy rates, for example, we can track tourists or the locale.  We handled n/a values and built a new feature from the previous reviewed date. We also showed location-specific data and the distribution of our |