DATA METHODOLOGY OF AIRBNB NYC DATASET

By

Derangula Laxmi Triveni

Najmabi Mainuddin Shaikh

Garima Gupta

Step 1: Story Telling

➢ Checked the data AB_NYC_2019, noted the important fields.
➢ Made a mind map of the essential slides for the presentation.

Step 2: Data Cleaning

➢ Using python Jupyter notebook, found the unique values, evaluated the numerical
   And categorical variables.
➢ Checked for null values, missing values and duplicate values. And replaced the Null
   Values appropriately.
➢ Found that Price is highly skewed and has many outliers based on below calculation

```
#Skewness,Kurtosis
print("Skewness:%f" % airbnb_nyc['price'].skew())
print("Kurtosis:%f" % airbnb_nyc['price'].kurt())
```

Skewness : 19.118939
Kurtosis : 585.672879
Box plot of Price with many outliers
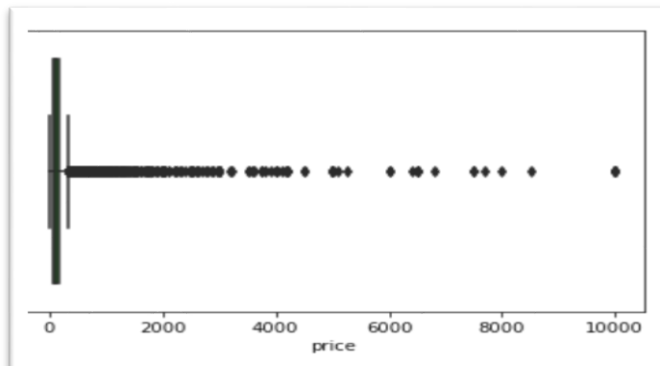
➢ Used Tableau to visualize data from the NYC Airbnb dataset in order to obtain accurate
   insights.

> Exploratory data analysis was used to identify customer preferences based on various pa rameters such as are preferences, property prices and listing preferences.

Step 3: Data analysis

> On the basis of overall analysis Manhattan has the highest listing for the most popular cate gory of entire home/apartment.

> Staten Island is on lower side for listing as well as detraction on it and based on that we nee d to focus more here.

Step 4: Presentation

> Made a presentation adhering to the best practices.
> Added recommendations to the respective departments.

THANK YOU