This dataset has around 17,000 observations in it with 16 columns and it is a mix between categorical and numeric values.

**Application of machine learning algorithm**

* + Linear Regression
  + Decision Tree Regression
  + Random Forest Regression

**Seaborn**: [**Seaborn**](https://www.geeksforgeeks.org/introduction-to-seaborn-python/) is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and color palettes to make statistical plots more attractive.

**Correlation**:In statistics, correlation or dependence is any statistical relationship, whether causal or not, between two random variables or bivariate data

**Linear Regression:**

Advantage

* Easy to implement
* Non complex

Disadvantage:

* It assumes that there is a straight-line relationship between them. It assumes independence between attributes.
* It works on mean of the variable and this may not be enough for some datasets.

**Decision Tree Regression:**

Advantage:

* Easy to understand and interpret, perfect for visual representation
* Can work with numerical and categorical features.
* Feature selection happens automatically: unimportant features will not influence the result.

Disadvantage:

* They are unstable, meaning that a small change in the data can lead to a large change in the structure of the optimal decision tree
* Can some time lead to overfitting

**Random Forest Regression:**

Advantages:

* Often machine learning models are overfitted, random forest classifiers wouldn't get overfitted.
* Can handle large data easily
* It has an effective method for estimating missing data and maintains accuracy when large proportion of the data are missing.

Disadvantages:

* The main limitation of random forest is that a large number of trees can make the algorithm too slow and ineffective for real-time predictions.
* Random forest is a predictive modeling tool and not a descriptive tool.(For the present situation ther is aprediction analysis)

The source code contains all the statistical observations ,visualization and models.

**Recommendations:**

The Airbnb can add a feature of stars to rate the service so that a best review can be easily determined.

Reviews per month may be greater for a host than other but it could be possible that a host with review 0.5 may be providing good service than review0.6.This can be determined by stars(0-5) and will not flood the data(i.e data size want increase too much).