Data Wrangling

The data of listings in San Francisco, was taken from [www.insideairbnb.com](http://www.insideairbnb.com). First the data was viewed through pandas, and it was found that there were 96 columns of attributes. Most of the data types were strings and were therefore unable to be used by our machine learning algorithm. The columns containing these attributes were promptly removed by selecting only those columns with numbers for use with the data frame. Using exploratory data analysis we observed that there were no outliers so that step of data wrangling could be skipped. The remaining columns that contained NaN values which were replaced by zeros since NaN cannot be utilized in machine learning. The data was then normalized using min/max methods so it would be more centralized. After these steps were taken the data was able to be used for machine learning.