index

Team Information

Team name: Team Team

Members: Grey Gergen, Tyler Davis, Jesus Rodriguez, Jack Macfadyen

Motivation

There are many considerations that must be made when finding temporary housing, either for vacations or business trips. Airbnb is a popular website that allows individuals or businesses to list potential places for users to choose. These users can choose where they wish to stay based on location, accommodations, how many beds they need, and price. We sought to create a model that predicts the price of an Airbnb listing based on various variables.

Airbnb data has been scraped by a team of contributors and gathered on a website here. Their motivation is to show transparency in how spaces are being rented to tourists in their communities.

When coming across this dataset, we found several missing values in the price of different Airbnb listings. Of the 36111 listings we found 14783 missing prices. We believe that we can use the latitude, longitude, beds, bathrooms, and neighborhood location among other variables to impute the missing pricing data for these listings.
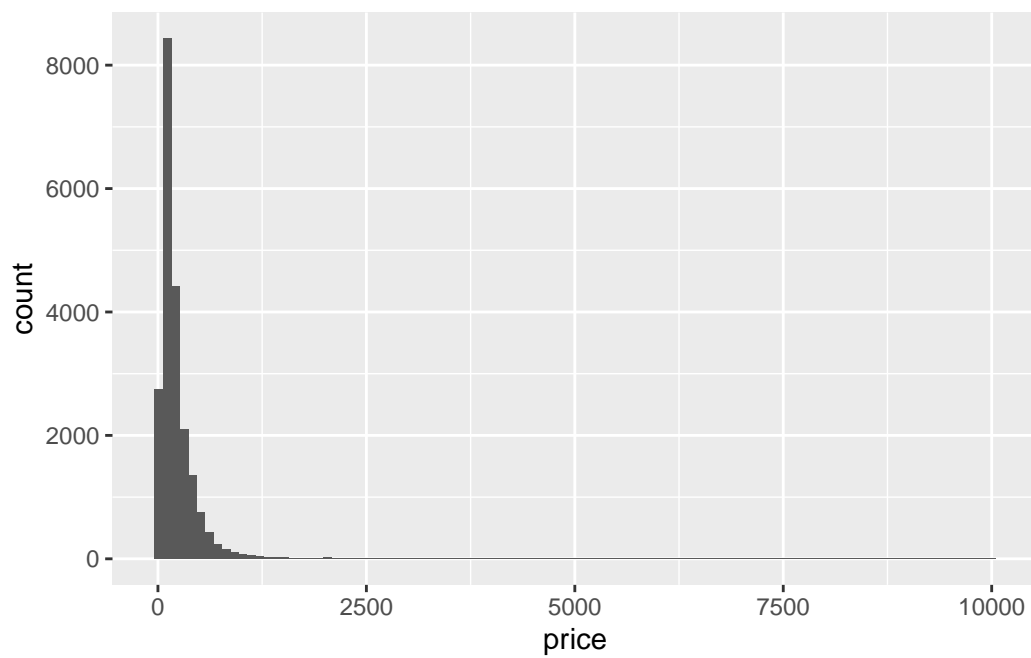
Data Documentation

Data Cleaning

Dataset Description

`Price`

Summary Statistics

	price
	Min. : 10.0
	1st Qu.: 88.0
	Median : 152.0
	Mean : 234.5
	3rd Qu.: 272.0
	Max. :10000.0
	NA's :14783

Distribution of Price

price	n
10000	3
9999	1
9895	1
9000	1
8758	2
8403	1
7500	1
7024	1
6112	2
6041	1

I am skipping ahead to write about the model construction and imputation methods

Model Construction

The next step in the analysis will be to build a model with price as the response variable. The predictors will be values that are suspected to be indicators of price such as beds, lat and long, accommodates, bathrooms, etc. This model will then be run and coefficients values and significance will be documented. This will all be done without imputation.

Matrix Plot to Assess Variable Correlation

Matrix plot needs to be described

Model Building

$$\log(\text{Price}) = \beta_0 + \beta_1 \cdot \text{bedrooms} + \beta_2 \cdot \text{beds}$$