Data Analysis 3 - Assignment 2
Airbnb Price Prediction - Montreal, Canada

Xibei Chen

07 February 2022

Introduction

The purpose of this project is to help a company set price to their new apartments that are not yet on the airbnb market. The company is operating small and mid-size apartments hosting 2-6 guests in Montreal. To help them set reasonable price, my approach is to analyze the data of listings in Montreal on 11th Dec 2021 (the most recent date) from Inside Airbnb (a website providing fairly cleaned web scrapped data on listings throughout all the major cities of the world) and then I will try to build several price prediction models based on the data with a variety of methods, such as OLS, Lasso, CART, Random Forest and GBM. Eventually with prediction performance, interpretability, external validity etc. taken into consideration, I will decide which prediction model to use for helping the company set price to their apartments.

Data Preparation

Data Cleaning

I selected only the listings that meet the following criteria.

- With valid Airbnb's unique identifier for the listing;
- The maximum capacity is lower than or equal to 6;
- Room type is entire apartment or flat;
- Property type is entire place (boat and RV excluded).