Daniel Park

DATA 606

**Predicting Airbnb Prices in Washington D.C.**

My capstone project will be to predict Airbnb prices in Washington D.C. Airbnb is an internet marketplace where people can rent out their property for a fixed rate per day. The main issue for hosts is that Airbnb can’t really determine the best price to sell their property, so it is up to the host to determine it. There are many factors that can determine the price of a property. It mainly depends on what the majority is looking for when they’re looking for places to stay. Machine Learning can be used to find important features from thousands of Airbnb properties to show what people look for when they look for places to stay. The datasets that I plan on using is from a website that is made by Airbnb called insideairbnb.com. The website provides Airbnb data from multiple cities throughout the world. One of data file is called listings.csv, which is the most important one. This file provides the ID, accommodation, bathroom, daily rate, location, etc. Using machine learning I will be able to extract information based on feature importance in order to accomplish my goal. For my project, I plan on using two models: Hedonic Regression and Gradient Boosting. Hedonic Regression is great for data that is related to real estate (AKA, multiple properties), and gradient boosting is an overall popular model that people use due to its accuracy and speed. I plan on using these two models to accurately depict future prices for such locations based on feature importance and other factors like seasonality.