Airbnb Consumer Valuation and Pricing Predictions

Capstone 3 - Project Proposal

Steven Ho

December 2022

Introduction & Business Problem

Airbnb was started by two roommates in 2007 when they were trying to figure out a way to pay their rent. A conference was held in town that weekend in San Francisco and all hotels were sold out. The 2 roommates decided to turn their apartment into an airbed & breakfast. Fast forward many years, they turned one airbnb listing to over 4 million hosts all over the world. With the help from their investors, they were able to go public in 2020.

We have been hired by Airbnb to understand and study the consumer behavior of the New York City (NYC) market of the short term rental market. NYC is one the world's largest cities and the short term rental market has been a huge headache for the hospitality world in the Big Apple. Airbnb needs our help to better understand consumer behavior and how Airbnb host can create better pricing structures depending on the location, sizing of the Airbnb location, and overall experience.

The objective of this analysis is to predict the price and best time to stay at an Airbnb in the NYC area. Airbnb has become one of a kind service that is used and recognized by the entire world. Data analysis on millions of listings provided in a specific region of the world provided through Airbnb is a crucial factor for the company and spending factors to be analyzed and further enhancements and investments can be made to better understand the company and consumer behavior.

Datasets

This dataset will include a dataset found from Kaggle that will present characteristics for over 48,000 airbnb listings during the years of 2011 - 2019. Characteristics include: NYC neighborhood, room types, minimum nights of stay, reviews per month, and the estimated availability on a calendar year.

Envisioned Approach

1. With the datasets from Kaggle, we will start by cleaning the data
2. Visualize the data using an EDA analysis to create analysis and separation of different characteristics of the data
3. Model the data using machine learning applications to predict the price and best overall airbnb listings/areas to stay within the NYC area
4. Generate different aspects and training cases
5. Conclusion and analysis

Deliverables

1. Data presentation using Python displaying data visualization and analysis
2. Code for modeling and analysis using Python and Machine learning applications
3. A written final report displaying my analysis and research
4. A Presentation deck displaying my visuals and analysis