

Project Proposal

Predicting Air B&B Prices for Entire Homes

Scott Kroeger

Domain

Airbnb is a website that lists rooms, apartments and homes that users can rent for short term stays. Started in 2007, Brian Chesky and Joe Gebbia hosted the first guest on an air mattress in an effort to make rent money. Now the site has now has more than seven million listing world wide and accommodations that range from tree houses and tiny houses to castles. The site provide several data points for each listing. A successful model could be used by potential listing agents as a guild to the value of their listing prior to making it available on Airbnb.

An MVP would scrape data from Airbnb listings, and use that data to perform a linear regression that estimate the price of a list based on its features. The MVP will only consist of numeric data points, and all listings will be entire homes for the week of December 22, 2019 to December 28, 2019.

Data			
Variable	Type	Description	Used for Model
Number of guests	Int	The number of guests that the list can house	Y
Number of bedrooms	Int	The number of bathrooms in the listing	Y
Number of beds	Int	The number of beds	Y
Number of bathrooms	Int	The number of bathrooms	Y
Cleaning Fee	Float	A fee to clean the dwelling after the user departs	Y
Service Fee	Float	An Airbnb imposed fee that is income to Airbnb	Y
Number of Amenities	Int	The total number of amenities the listing has such as WIFI, washing machine, iron	Y
Latitude	Float	North/South coordinate of the listing	Y

Data			
Longitude	Float	The East/West coordinate of the listing	Y
Number of words in the house rules	int	The count of words in the house rules	Y
			Y Total: 10

Knowns and unknown

- The model will only capture numerical data and the individual attractiveness of the listing text and images will not be taken into account
- The Airbnb website does not lend itself to web scraping and not all variables have been proven to be available