Full-Time Airbnb:
What Can You Make With Your Home?

Jon Whelpley

December 2020

1 Introduction

Airbnb has been a popular alternative to hotels for any person interested in travelling. Moreover, it has given homeowners an opportunity to supplement their income by leasing out a spare bedroom or spending a few nights away from the home altogether. Naturally this lends itself to the question: What revenue can you realistically make if you were to lease a home as a full time airbnb?

2 Data

2.1 The Main Data Set

The data comes in the form of a table taken from Kaggle. Here is the Link

The table consists of 226030 records with 17 feature columns which are as follows:

Column Name	Description
'id'	unique listing id
'name'	name/description of listing
'host_id'	unique host Id
'host_name'	name of host
'neighbourhood_group'	group in which the neighbourhood lies
'neighbourhood'	name of the neighbourhood
'latitude'	latitude of listing
'longitude'	longitude of listing
'room_type'	room type
'price'	price of listing per night
'minimum_nights'	min no. of nights required to book.
'number_of_reviews'	total number of reviews on listing
'last_review'	date of last review
'reviews_per_month'	average reviews per month on listing
'calculated_host_listings_count'	total number of listings by host
'availability_365'	num of days in year listing is available
'city'	region of the listing

The goal here is to predict the reviews per month. According to airbnb users leave a review 70 percent of the time. This will give us an estimate for the number of consumers that actually rent the bnb. For example if we predict a home will have 10 reviews per month at 80 dollars per night, and a minimum of 2 nights per stay we will estimate the monthly revenue to be:

$$\frac{10 \text{ reviews per month}}{0.7 \text{ reviews per stay}} * \text{ 2 nights per stay} * 80 \text{ dollars per night}$$

$$= 2285.7 \text{ dollars per month}$$

Essentially, if we're able to predict reviews per month based on these features we can give a projection of revenue and even compare that to mortgage and maintenance costs.

2.2 Additional Location Data

While Location data for each airbnb is already given. Travellers don't just want to get an airbnb to have a place to stay. They travel to find things to do. So it will be very important to lookup places where there's interesting things to do. Due to the limitations of four-square's API it would be disadvantageous to check each airbnb and find what's nearby; however, if we cluster the neighborhoods we may find that searching the cluster centers may give enough to information to find what's close by. Minimally this will be used for the visualizations but it might be possible to get features that are just as valuable by getting locations that are nearby the centers of neighborhood cluster in a certain area.

2.3 Descriptive Data

Not only will extra location information be important but I suspect the text in the 'name' column will be important as it describes and 'sells' the airbnb. I will be looking for the most common words through the listings and see if those influence the number of reviews in any way. Alternatively I may even try to find the most common words to the top 10 percent of the most reviewed so showcase which descriptions are most successful. Each word will then become its own feature column that is '1' when the name column has the word and '0' when it doesn't.