

Exploring Airbnb prices in London: which factors influence price?

Analysing Airbnb data can help boost your listing's earnings



Laura Lewis

Follow

May 20, 2019 · 9 min read ★



Source: plumguide.com

Project aims and background

Airbnb is a home-sharing platform that allows home-owners and renters ('hosts') to put their properties ('listings') online, so that guests can pay to stay in them. Hosts set their own prices for their listings, and although Airbnb and other sites provide some general

guidance, there are currently no free and accurate services which help hosts price their properties using a wide range of data points.

It's really important to get pricing right on Airbnb, especially in big cities like London where there is a lot of competition. I recently conducted a project to use machine learning and deep learning to predict the price for Airbnb properties in London. I've explored the preparation and cleaning of Airbnb data in a previous post, and in a future post I'll explore the modelling. But this post is all about exploring the listing data to extract interesting and useful insights, in order to help hosts maximise their earnings.

Additional context: I previously worked for a year and a half at an Airbnb property management company, as head of the team responsible for pricing, revenue and analysis. Decisions made during the course of this project are therefore informed by domain expertise in this industry.



The dataset

The dataset used for this project comes from Insideairbnb.com, an anti-Airbnb lobby group that scrapes Airbnb listings, reviews and calendar data from multiple cities around the world. The dataset was scraped on 9 April 2019 and contains information on all London Airbnb listings that were live on the site on that date (about 80,000). A GeoJSON file of London borough boundaries was also downloaded from the same site.

The data is quite messy, and has some limitations. The major one is that it only includes the advertised price (sometimes called the 'sticker' price). The sticker price is the overall nightly price that is advertised to potential guests, rather than the actual average amount paid per night by previous guests. The advertised prices can be set to any

arbitrary amount by the host, and hosts that are less experienced with Airbnb will often set these to very low (e.g. £0) or very high (e.g. £10,000) amounts.

Nevertheless, this dataset can still be used as a proof of concept. A more accurate version could be built using data on the actual average nightly rates paid by guests, e.g. from sites like AirDNA that scrape and sell higher quality Airbnb data.

Exploring the data for insights

Location, location, location

Question: which areas of London have the most Airbnb properties, and which are the most expensive?

Answer: Westminster (west of the City) has the most Airbnb properties, followed by Tower Hamlets (east of the City). Inner London boroughs have significantly more listings than outer London boroughs.

However, the pattern with prices is slightly different. Kensington and Chelsea (to the west of Westminster) is the most expensive area. This is a famously expensive area to live, with some of the highest house prices in the world. Although inner London is generally more expensive than outer London, there are also some more expensive listings spread out to the west of the city along the Thames (which has some very beautiful areas).

There are very few listings in the very centre of London (as there is relatively little residential property here), but as a result they are very expensive.

How did I find this out? With a fun library called GeoPandas, which I discovered when working on this project. It's basically Pandas for geospatial data, and allows you to add area boundaries as vectors (or polygons or shapefiles) as a new column in your dataframe. The code to produce this is in my GitHub repo if you want to check it out.

Assessing the competition

Question: what are the most common property and room types?

Answer: about 70% of properties are apartments. The remainder are houses or more uncommon property types (e.g. ‘bed and breakfast’ or ‘yurt’).

About 55% of listings are entire homes (i.e. you are renting the entire property on your own). Most of the remainder are private rooms (i.e. you are renting a bedroom and possibly also a bathroom, but there will be other people in the property). Fewer than 1%

are shared rooms (i.e. you are sharing a room with either the property owner or other guests).



Changes over time

Question: how long have hosts been listing properties on Airbnb in London?

Answer: the oldest London listing that is currently live on Airbnb was first listed on the site in August 2008. From 2011 onwards, the number of listings started increasing considerably. However, growth in the number of new hosts (of those currently listing on

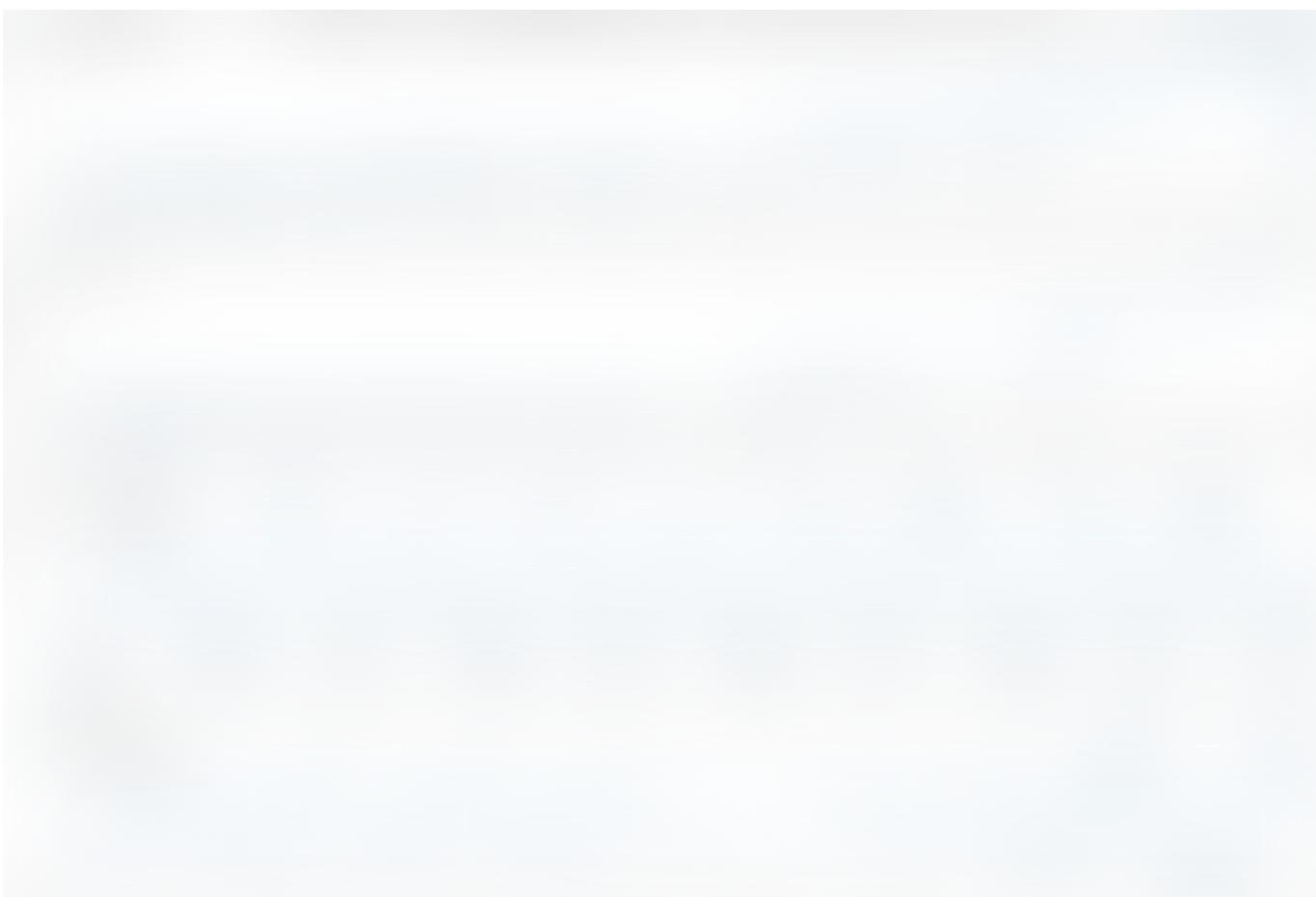
the site) has been decreasing since 2015, when the UK government introduced a law making it illegal to let short-term residential properties for more than 90 nights a year.

A high level of seasonality is evident, with notable peaks in the summer when people put properties online to take advantage of the increased number of tourists in the summer holidays.



The following two images show the decomposed time series for the number of (currently live) hosts joining the site, and the number of (currently live) listings getting their first review. This was done using the `seasonal_decompose` function from the statsmodels library. It decomposes a time series into the overall trend, seasonality (which is significant in the tourism industry) and the leftover unexplained residuals. The latter could be due to a wide variety of factors. For example, the big peak in the number of hosts joining Airbnb in 2015 could be a response to the government legislation, as existing hosts may have created new accounts in order to re-list their properties and get around the 90 day limit.





Decomposed time series

The more the merrier

Question: what is the distribution of the number of people accommodated in Airbnb listings in London, and how does this affect prices?

Answer: the most common property setup sleeps two people in one bed in one bedroom, with one bathroom. Unsurprisingly, properties that accommodate more people achieve

higher nightly rates, with diminishing returns coming after about 10 people.

As I discovered during the modelling part of this project, the biggest predictor of price is the number of people a property accommodates.

So, probably time to buy that sofa-bed then.



Guests heart Airbnb

Question: what is the distribution of reviews like?

Answer: people love their Airbnbs. For every review category (accuracy, cleanliness, check-in, communication, location, value and overall), the majority of listings that have had a review received a 10/10 rating for that category (or 95–100/100 overall). So, probably best to make sure you're in this category. Ratings of 8 or below are relatively rare. Guests seem to be most positive about communication, check-ins and accuracy.

About a quarter of listings have not yet been reviewed. These will mostly be properties which have not yet had a completed stay, many of which may be ‘inactive’ properties — i.e. those which, although listed on the site, do not have any dates available to book.



Amenities your guests will love (and those your property might hate)

Question: *which amenities are common, and which increase the price of an Airbnb listing?*

Answer: amenities can be split into four main groups:

Uncommon, but properties with it have a higher median price:

- Balcony
- Bed linen
- Coffee machine
- Basic cooking equipment
- Elevator
- Child friendly
- Outdoor space
- Long term stays allowed
- Private entrance
- Safe or security system

- Self check-in

Most properties have it, and properties with it have a higher median price:

- TV
- Washer, dryer and/or dishwasher

Most properties have it, and there is no major difference in price between properties with and without it:

- Internet

Uncommon, and properties with it have a lower median price:

- Breakfast (probably because these are bed & breakfasts and therefore single rooms rather than entire homes)
- Parking (presumably because these are less likely to be central properties)
- Greeted by host (surprisingly!)
- Pets allowed

The takeaway from this is that there are some things that hosts can do to try and improve their listing's price. However, there is not always a simple causative relationship — e.g. having a parking space might improve your price, but listings with parking spaces are cheaper on average because they are less likely to be in central London, and location is a more important factor than parking space.

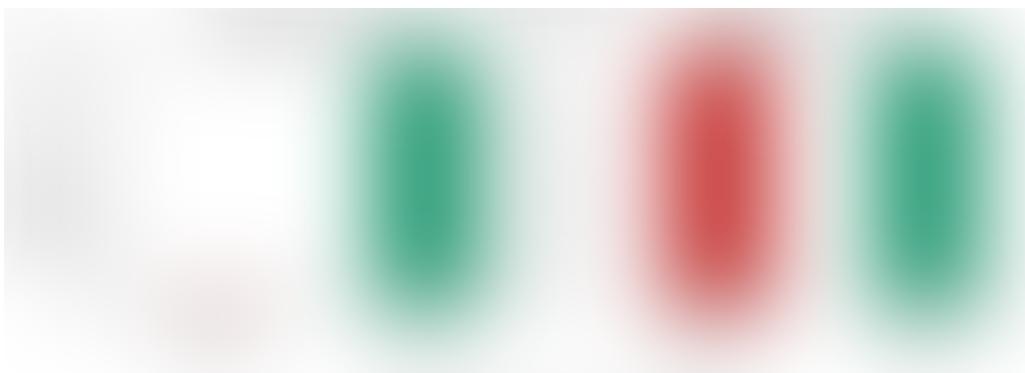
Here are some top tips:

- **Make sure you have the necessities that your competitors also have**, e.g. internet, a TV and white goods (a washer, dryer and/or dishwasher).
- If you can, **include some extras to make your property stand out and achieve a higher price**. For example, you could buy a coffee machine, allow self check-in and long-term stays, and make your property child-friendly.

- **Avoid pets** as the additional wear and tear to your property might harm its financial performance (not to mention putting off guests who might prefer a non-smoking property).









Not worth the extra effort

Question: is it worth being a superhost? And is it worth being verified? (two questions that come up a lot)

Answer to both: not really.

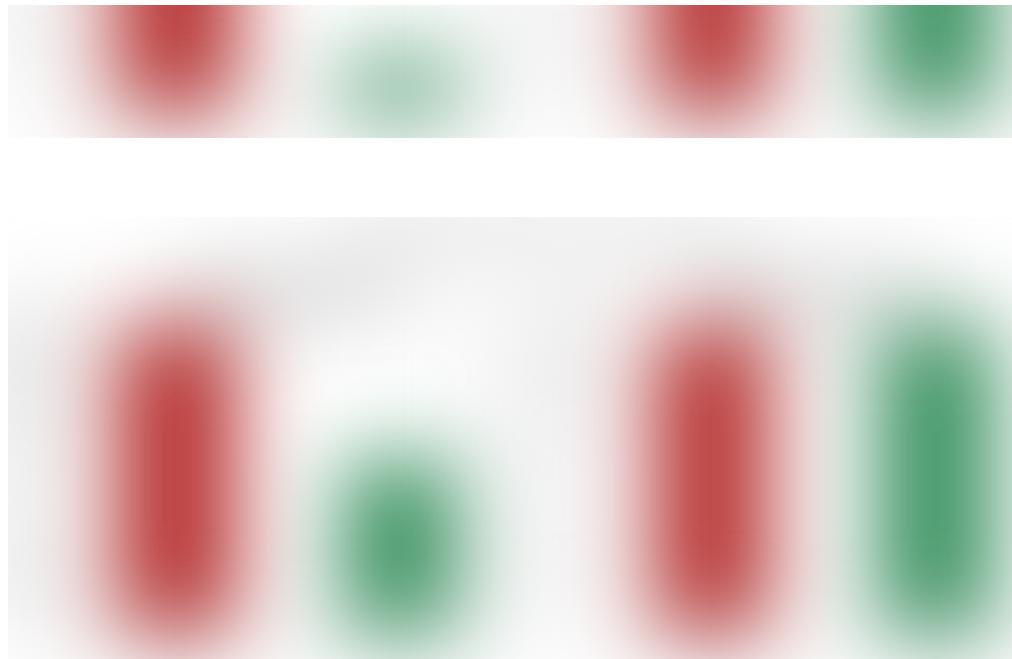
About 15% of hosts are superhosts. This requires meeting four criteria:

1. Host at least 10 stays per year
2. Respond to guests quickly and maintain a 90% response rate or higher
3. Have at least 80% 5-star reviews
4. Don't cancel confirmed reservations

However, it does not seem to improve the median price per night of their Airbnb listings, so it is arguable whether or not it is worth it (from a purely financial perspective).

About 35% of hosts are verified, e.g. by providing ID and verifying your phone number and email address. Interestingly, however, despite the fact that one might expect the additional trust commanded by verified hosts to result in higher nightly prices, there is no discernable difference between the median nightly prices of listings with verified and unverified hosts.

Finances aside, however, these are probably still good things to do for other (i.e. nice human) reasons.



Summary

The exploratory analysis above highlights some interesting trends and patterns, as well as some factors that can increase an Airbnb listing's price. In another post I'll go into the details of how I used this data to build a model to predict Airbnb prices. I hope you've found at least some of this useful (if you're an Airbnb host) and/or interesting (if you're a data junkie like me)!

• • •

If you found this post interesting or helpful, please let me know via the medium of claps and/or comments, and you can follow me in order to be notified about future posts. Thanks for reading!

Sign up for The Daily Pick

By Towards Data Science

Hands-on real-world examples, research, tutorials, and cutting-edge techniques delivered Monday to Thursday. Make learning your daily ritual. [Take a look](#)



Get this newsletter

Emails will be sent to
taikhoanxaichung2020@gmail.com.
[Not you?](#)

Airbnb Data Science Python Data Data Visualization

About Help Legal

Get the Medium app

