Data Assumptions

Data Resources

- Explore or Get the data
- Read Data Policies
- Make a Data Request
- Read Data Assumptions
- View the Data Dictionary
- Join the Data Community

Disclaimers

- This site is not associated with or endorsed by Airbnb or any of Airbnb's competitors.
- The data utilizes public information compiled from the Airbnb web-site including the availabiity calendar for 365 days in the future, and the reviews for each listing. Data is verified, cleansed, analyzed and aggregated.
- No "private" information is being used. Names, photographs, listings and review details are all publicly displayed on the Airbnb site.
- This site claims "fair use" of any information compiled in producing a non-commercial derivation to allow public analysis, discussion and community benefit.
- Accuracy of the information compiled from the Airbnb site is not the responsibility of Inside Airbnb. Due care has been taken with any processing and analysis.
- Location information for listings are anonymized by Airbnb.
    - In practice, this means the location for a listing on the map, or in the data will be from 0-450 feet (150 metres) of the actual address.
    - Listings in the same building are anonymized by Airbnb individually, and therefore may appear "scattered" in the area surrounding the actual address.

The Airbnb calendar for a listing does not differentiate between a booked night vs an unavailable night, therefore these bookings have been counted as "unavailable". This serves to understate the Availability metric because popular listings will be "booked" rather than being "blacked out" by a host.

- Some hosts might not keep their calendar updated, or have it highly available even though they live in the entire home/apartment.
    - To ensure you are only seeing listings that are both "highly available" and being booked frequently, use the "Only highly available" filter with the "Only recent and frequently booked" filter to only select listings that have only been rented recently (reviewed in the last 6 months), and are being rented regularly (number of nights per year greater than the threshold for that city).
- Neighbourhood names for each listing are compiled by comparing the listing's geographic coordinates with a city's definition of neighbourhoods. Airbnb neighbourhood names are not used because of their inaccuracies.
- All copyright and registered trademarks remain the property of their owners.

The Occupancy Model

One of the biggest issues with Airbnb is whether hosts are renting out residential properties permanently as hotels, as opposed to sharing the primary residence in which they live "occasionally".

An occupancy model can be used to estimate how often an Airbnb listing is being rented out, and also approximate a listing's income.

Inside Airbnb uses an occupancy model which we've christened the **"San Francisco Model"** in honor of the public policy and urban planners working for that fair city who created occupancy models to quantify the impact of Airbnb on housing.

The two models created by, respectively, Alex Marqusee for the San Francisco Planning Department; and the Budget and Legislative Analyst's Office, are detailed here:

- The Executive Summary of Amendments Relating to Short-Term Rentals refers to **Exhibit B: Marqusee Memo Airbnb and San Francisco: Descriptive Statistics and Academic Research** on Page 30. Note, that the Author and the Planning Department released an updated version of this Research in May 2015, but it is not yet available online.
- Analysis of the impact of short-term rentals on housing from the Budget and Legislative Analyst's Office.

    - Alex Marqusee uses a review rate of 72%, however this is attributed to an unreliable source: Airbnb's CEO and co-founder Brian Chesky.
    - The Budget and Legislative Analyst's Office (page 49) also use a value 72% for their review rate, and in addition, introduce a higher impact model using a review rate of 30.5% - based on comparing public data of reviews to the The New York Attorney General's report on Airbnb released in October 2014.
    - Inside Airbnb analysis found that using a review rate 30.5% is more fact based, however probably not conservative enough, given that the Budget and Legislative Analyst's Office did not take into account missing reviews because of deleted listings. A review rate of 72% is unverifiable - therefore 50% was chosen as it sits almost exactly between 72% and 30.5%.
- An **average length of stay**, where avaialble, is configured for each city, and this, multiplied by the **estimated bookings** for each listing over a period gives the **occupancy rate**
    - Where statements have been made about the average length of stay of Airbnb guests for a city, this was used.
    - For example, Airbnb reported 5.5 nights as the average length of stay for guests using Airbnb in San Francisco.
    - Where no public statements were made about average stays, a value of **3 nights per booking** was used.
    - If a listing has a **higher minimum nights** value than the average length of stay, the minimum nights value was used instead.
- The **occupancy rate** was **capped at 70%** - a relatively high, but reasonable number for a highly occupied "hotel".
    - This controls for situations where an Airbnb host might change their minimum nights during the high season, without the review data having a chance to catch up; or for a listing with a very high review rate.
    - It also ensures that the occupancy model remains conservative.
- **Number of nights** booked or available per year for the **high availability** and **frequently rented** metrics and filters were generally aligned with a city's short term rental laws designed to **protect residential housing**.

Inside Airbnb

Adding data to the debate

data and advocacy about Airbnb's impact on residential communities.

We work towards a vision where data and information empower communities to understand, decide and control the role of renting residential homes to tourists.

About Inside Airbnb

Donate

Collaborate

Partner

Inside Airbnb Network