

Airbnb Berlin

Data Source: The data being referred to is publicly accessible and sourced from Inside Airbnb <http://insideairbnb.com/get-the-data>

Data Collection: The dataset available for use was produced and released by Airbnb on Berlin, Germany March 19, 2023.

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

Data Profile:

- The Data has 19 columns and 12049 rows. After data cleaning 6059 rows and 18 columns.
- The missing Value have been dropped.
- There is no duplicates Values.

Index	Columns	Description	Time Variant/ Invariant	Data type
1	listing_id	Id number of the accommodation	Invariant	Qualitative
2	listing_name	Name of the accommodation	Invariant	Qualitative
3	host_id	Host Id number	Invariant	Qualitative
4	Host_name	Name of the host	Invariant	Qualitative
5	District	Surrounding area of a city	Invariant	Qualitative
6	neighbourhood	District or community within a city	Invariant	Qualitative
7	latitude	Latitude measures the distance north or south of the equator	Invariant	Quantitative

8	longitude	Longitude measures distance east or west of the prime meridian.	Invariant	Quantitative
9	room_type	Room description	Invariant	Qualitative
10	price	Price of the accommodation	Variant	Quantitative
11	minimum_nights	the minimum number of nights that a guest can book your Airbnb vacation rental.	Variant	Quantitative
12	number_of_reviews	Ratings for hosts	Variant	Quantitative
13	last_review	Last rating for hosts	Variant	Quantitative
14	reviews_per_month	Each month review rating for hosts	Variant	Quantitative
15	host_total_listings	Number of host listings	Variant	Quantitative
16	availability_365	Availability time in year	Variant	Quantitative
17	review_in_the_last_year	In last 12 months number of review	Variant	Quantitative
18	license	License of the accommodation	Invariant	Qualitative

Limitations and Ethics

Inside Airbnb Berlin is a dataset that provides information on Airbnb listings in Berlin. The dataset has several limitations and ethical considerations that must be taken into account when using it.

Limitations:

1. Bias: The dataset may be biased as it only includes information on Airbnb listings that have been publicly listed on the platform. Therefore, there

may be listings that are not included in the dataset, which could impact the overall analysis.

2. **Incomplete information:** The dataset may not include all the information needed to conduct a comprehensive analysis of the Airbnb market in Berlin. For instance, information such as the availability of the listings, the number of bookings, and the length of stay may not be available in the dataset.
3. **Data quality:** The quality of the data in the dataset may be variable. For instance, some listings may have incomplete or incorrect information, which could impact the accuracy of any analysis conducted using the data.
4. **Privacy:** The dataset includes personal information about Airbnb hosts, such as their names, addresses, and contact details. Care must be taken to protect this information and ensure that it is not misused or disclosed to unauthorized parties.

Ethical considerations:

1. **Respect for privacy:** As mentioned, the dataset includes personal information about Airbnb hosts. Therefore, any analysis conducted using the data must respect the privacy of these individuals and ensure that their personal information is not used in a way that could harm them.
2. **Transparency:** The use of the dataset must be transparent, and individuals must be informed about the use of their data. This includes providing information about how the data will be used, who will have access to it, and how it will be protected.
3. **Fairness:** The analysis conducted using the dataset must be fair and not discriminate against any individuals or groups. For instance, it must not be used to discriminate against hosts based on their race, gender, or other personal characteristics.
4. **Legal compliance:** The use of the dataset must comply with all applicable laws and regulations, including data protection laws, and any other legal requirements that may apply.

Questions:

1. How does the type of room (private room, entire apartment, shared room) impact the pricing of Airbnb listings in Berlin?
2. What is the average minimum stay required by Airbnb hosts in Berlin?
3. Who are the top hosts in Berlin, and what factors contribute to their success?