

# About Dataset

## New York City Airbnb Data Cleaning

Airbnb, Inc is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Based in San Francisco, California, the platform is accessible via website and mobile app. Airbnb does not own any of the listed properties; instead, it profits by receiving commission from each booking. The company was founded in 2008. Airbnb is a shortened version of its original name, AirBedandBreakfast.com.

### Context

Since 2008, guests and hosts have used Airbnb to travel in a more unique, personalized way. As part of the Airbnb Inside initiative, this dataset describes the listing activity of homestays in New York City

### Dataset

The following Airbnb activity is included in this New York dataset:

Listings, including full descriptions and average review score  
Reviews, including unique id for each reviewer and detailed comments  
Calendar, including listing id and the price and availability for that day

Column	Description	Data type
id	Unique id of each listing	numeric
name	Name of the airbnb listing	text
Host id	Unique id for the host	numeric
host_identity_verified	Whether the identify of the host is verified or not	Categorical
host name	Name of the host	text
neighbourhood group	District where the property is	Categorical
neighbourhood	Area or locality of the property	Categorical

lat	Latitude	numeric
long	Longitude	numeric
country	Country where the property is	Categorical
country code	ISO country code	Categorical
instant_bookable	If the property can be instantly booked or not	Categorical
cancellation_policy	Cancellation policy for the booking	Categorical
room type	Type of room	Categorical
Construction year	Year when the property was constructed	numeric
price	Price per night	numeric
service fee	Additional service fee	numeric
minimum nights	Minimum number of nights required for booking	numeric
number of reviews	Total number of reviews	numeric
last review	Last review date	date
reviews per month	Average number of reviews per month	numeric
review rate number	Rating score based on reviews	numeric
calculated host listings count	Total number of listing managed by the host	numeric
availability 365	Number of days the property is available for booking throughout the year	numeric
house_rules	Rules defined by the host for their guests	text
license	License number for legal compliance of the listing	text

**Steps -**

1. Explore the dataset
2. Data cleaning
3. Data manipulation
4. Describe
5. Visualization
6. Answer some questions for analysis
7. Use Quarto for rendering this project