AirBedandBreakfast Company

- Airbnb, 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.

- This dataset is used to predict price of homestays . Each row in the data provides relevant information about them.

Attribute Information :

id	integer	Airbnb's unique identifier for the listing
name	text	Name of the listing
host_id	integer	Airbnb's unique identifier for the host/user
host_name	text	Name of the host. Usually just the first name(s).
host_identity_verified	boolean	t=true; f=false
neighbourhood group		The neighbourhood group as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
neighbourhood	text	Neighbourhood name
latitude	numeric	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
longitude	numeric	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
room_type	text	All homes are grouped into the following three room types: Entire place Private room Shared room Entire place the entire space with other people. Shared rooms are popular among flexible travelers looking for new friends and budget-friendly stays.
price	currency	daily price in local currency
service fee	currency	Service fees in local currency
minimum_nights	integer	minimum number of night stay for the listing (calendar rules may be different)
number_of_reviews	integer	The number of reviews the listing has
license	text	The licence/permit/registration number
instant_bookable	boolean	[t=true; f=false]. Whether the guest can automatically book the listing without the host requiring to accept their booking request. An indicator of a commercial listing.
reviews_per_month	numeric	The number of reviews the listing has over the lifetime of the listing

country	text	Country name
country code	text	Country code
cancellation_policy	text	**cancellation_policy**: 'strict', 'moderate', 'flexible'
Construction year	numeric	Construction year of homes
last review	string	The date of last review of specific home
review rate number	numeric	Review rate
calculated host listings count	numeric	Counts of **calculated host listings**
availability 365	integer	avaliability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
house_rules	string	House rules to stay in it

Data Information :

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 102599 entries, 0 to 102598
Data columns (total 26 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   id                              102599 non-null  int64
 1   NAME                            102349 non-null  object
 2   host id                         102599 non-null  int64
 3   host_identity_verified          102310 non-null  object
 4   host name                       102193 non-null  object
 5   neighbourhood group             102570 non-null  object
 6   neighbourhood                   102583 non-null  object
 7   lat                             102591 non-null  float64
 8   long                            102591 non-null  float64
 9   country                         102067 non-null  object
 10  country code                    102468 non-null  object
 11  instant_bookable                102494 non-null  object
 12  cancellation_policy             102523 non-null  object
 13  room type                       102599 non-null  object
 14  Construction year               102385 non-null  float64
 15  price                           102352 non-null  object
 16  service fee                     102326 non-null  object
 17  minimum nights                  102190 non-null  float64
 18  number of reviews               102416 non-null  float64
 19  last review                     86706 non-null   object
 20  reviews per month               86720 non-null   float64
 21  review rate number              102273 non-null  float64
 22  calculated host listings count  102280 non-null  float64
 23  availability 365                102151 non-null  float64
 24  house_rules                     50468 non-null   object
 25  license                         2 non-null       object
dtypes: float64(9), int64(2), object(15)
memory usage: 20.4+ MB
```

Data Information (Cont.) :

- Data with (102599) record & (26) features (columns).
- Missing Values Percentage :

```
id                               0.000000
NAME                             0.243667
host id                          0.000000
host_identity_verified           0.281679
host name                        0.395715
neighbourhood group              0.028265
neighbourhood                    0.015595
lat                              0.007797
long                             0.007797
country                          0.518524
country code                     0.127682
instant_bookable                 0.102340
cancellation_policy              0.074075
room type                        0.000000
Construction year                0.208579
price                            0.240743
service fee                      0.266084
minimum nights                   0.398639
number of reviews                0.178364
last review                     15.490404
reviews per month               15.476759
review rate number               0.317742
calculated host listings count   0.310919
availability 365                 0.436651
house_rules                     50.810437
license                         99.998051
dtype: float64
```

Data overview :

	id	NAME	host id	host_identity_verified	host name	neighbourhood group	neighbourhood	lat	long	country	country code	instant_bookable	cancellation_policy	room type
0	1,001,254	Clean & quiet apt home by the park	80,014,485,718	unconfirmed	Madaline	Brooklyn	Kensington	40.6475	-73.9724	United States	US	☐	strict	Private room
1	1,002,102	Skylit Midtown Castle	52,335,172,823	verified	Jenna	Manhattan	Midtown	40.7536	-73.9838	United States	US	☐	moderate	Entire home/apt
2	1,002,403	THE VILLAGE OF HARLEM....NEW YORK !	78,829,239,556	None	Elise	Manhattan	Harlem	40.809	-73.9419	United States	US	☑	flexible	Private room
3	1,002,755	None	85,098,326,012	unconfirmed	Garry	Brooklyn	Clinton Hill	40.6851	-73.9598	United States	US	☑	moderate	Entire home/apt
4	1,003,689	Entire Apt: Spacious Studio/Loft by central park	92,037,596,077	verified	Lyndon	Manhattan	East Harlem	40.7985	-73.944	United States	US	☐	moderate	Entire home/apt

	Construction year	price	service fee	minimum nights	number of reviews	last review	reviews per month	review rate number	calculated host listings count	availability 365	house_rules	license
0	2,020	$966	$193	10	9	10/19/2021	0.21	4	6	286	Clean up and treat the home the way you'd like your home to be treated. No smoking	None
1	2,007	$142	$28	30	45	5/21/2022	0.38	4	2	228	Pet friendly but please confirm with me if the pet you are planning on bringing with y	None
2	2,005	$620	$124	3	0	None	None	5	1	352	I encourage you to use my kitchen, cooking and laundry facilities. There is no additio	None
3	2,005	$368	$74	30	270	7/5/2019	4.64	4	1	322	None	None
4	2,009	$204	$41	10	9	11/19/2018	0.1	3	1	289	Please no smoking in the house, porch or on the property (you can go to the nearby c	None