

# **AIRBNB PRICE PREDICTION**



# BUSINESS PROBLEM

Hosts wanted to rent out their property on the Airbnb on nightly basis, but being new in the field, they are unaware of the market value of the property. As a result, they would set the price of the property as higher or least. When the price set is high in comparison to market value, he would not get any customers for renting out his property. While, on the contrary, when he will set the least price of the property, more customers would be attracted to rent out his property, but in return he will face the problem in making profits from the property.

# DATA SOURCE

1	id	log_price	property_type	room_type	amenities	accommodates	bathrooms	bed_type	cancellation	cleaning_fee	city	description	host_has_profile_pic	host_identity_verified	host_since	instant_bookable	latitude	longitude	name	neighbourhood	bedrooms	beds
2	6901257	5.01063529	Apartment	Entire home/apt	["Wireless Internet", "TV", "Cable TV"]	3	1	Real Bed	strict	TRUE	NYC	Beautiful, sunny apt with a view of the city	t	f	2012-03-26	f	40.6965236	-73.991617	Beautiful brooklyn Heights	Brooklyn Heights	1	1
3	6304928	5.12989872	Apartment	Entire home/apt	["Wireless Internet", "TV", "Cable TV"]	7	1	Real Bed	strict	TRUE	NYC	Enjoy traveling in the heart of the city	t	f	2017-06-19	t	40.7661154	-73.98904	Superb 3BR A Hell's Kitchen	Hell's Kitchen	3	3
4	7919400	4.97673374	Apartment	Entire home/apt	["TV", "Cable TV"]	5	1	Real Bed	moderate	TRUE	NYC	The Oasis in the heart of the city	t	t	2016-10-25	t	40.80811	-73.943756	The Garden of Eden	Harlem	1	3
5	13418779	6.62007321	House	Entire home/apt	["TV", "Cable TV"]	4	1	Real Bed	flexible	TRUE	SF	This light-filled home with a view of the city	t	t	2015-04-19	f	37.7720045	-122.43162	Beautiful Flat in the heart of the city	Lower Haight	2	2
6	3808709	4.74493213	Apartment	Entire home/apt	["TV", "Internet", "Cable TV"]	2	1	Real Bed	moderate	TRUE	DC	Cool, cozy, and modern	t	t	2015-03-01	t	38.9256269	-77.034596	Great studio in the heart of the city	Columbia Heights	0	1
7	12422935	4.44265126	Apartment	Private room	["TV", "Wireless Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	SF	Beautiful private room with a view of the city	t	t	2017-06-07	t	37.7531641	-122.42953	Comfort Suite in the heart of the city	Noe Valley	1	1
8	11825529	4.41884061	Apartment	Entire home/apt	["TV", "Internet", "Cable TV"]	3	1	Real Bed	moderate	TRUE	LA	Warm and cozy in the heart of the city	f	f	2017-03-03	t	33.9804544	-118.46282	Beach Town Studio and Par	Beach Town	1	1
9	13971273	4.78749174	Condominium	Entire home/apt	["TV", "Cable TV"]	2	1	Real Bed	moderate	TRUE	LA	Arguably the best in the city	t	t	2013-05-18	f	34.0467374	-118.26044	Near LA Live, Downtown	Downtown	1	1
10	180792	4.78749174	House	Private room	["TV", "Cable TV"]	2	1	Real Bed	moderate	TRUE	SF	Garden Studio in the heart of the city	f	f	2015-06-04	f	37.781128	-122.5011	Cozy Garden in the heart of the city	Richmond District	1	1
11	5385260	3.58351894	House	Private room	["Wireless Internet", "Cable TV"]	2	1	Real Bed	moderate	TRUE	LA	Quiet community in the heart of the city	f	f	2017-03-12	t	33.992563	-117.896	No.7 Queen Size Cozy Room	East Hollywood	1	1
12	5578513	4.60517019	Apartment	Private room	["Internet", "Wireless Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	NYC	This is a bright and modern apartment	t	t	2013-03-27	f	40.7238833	-73.98388	Large East Village	Alphabet City	1	1
13	17423675	5.01063529	House	Entire home/apt	["TV", "Cable TV"]	4	1.5	Real Bed	strict	TRUE	LA	A 1044 sq. ft. home in the heart of the city	t	t	2015-06-11	f	33.8758624	-118.40329	Sand Section in the heart of the city	Hermosa Beach	2	2
14	14066228	4.24849524	Apartment	Private room	["TV", "Internet", "Cable TV"]	2	1	Real Bed	flexible	TRUE	LA	Newly furnished in the heart of the city	f	f	2016-05-10	f	33.8132284	-118.38943	Beach Pad 1 in the heart of the city	Torrance	1	1
15	2658946	5.29831737	Apartment	Entire home/apt	["TV", "Cable TV"]	6	1.5	Real Bed	strict	TRUE	DC	Amazing location in the heart of the city	t	t	2012-10-12	t	38.91963	-77.031189	Charming 2 b in the heart of the city	U Street Corridor	2	3
16	583490	4.95582706	Apartment	Entire home/apt	["Kitchen", "Heating", "Cable TV"]	2	1	Real Bed	strict	TRUE	LA	This apartment is in the heart of the city	t	t	2012-11-20	f	33.7785265	-118.14593	VINTAGE 1930s Mediteran	East Hollywood	1	1
17	6226658	4.09434456	Apartment	Private room	["Internet", "Wireless Internet", "Cable TV"]	2	1.5	Real Bed	moderate	TRUE	Chicago	Just west of the city in the heart of the city	t	t	2012-09-10	f	41.9082402	-87.695242	Sweet Home in the heart of the city	Humboldt Park	1	1
18	7499179	4.31748811	Loft	Private room	["TV", "Wireless Internet", "Cable TV"]	2	1	Real Bed	strict	FALSE	Chicago	Our place is in the heart of the city	f	f	2016-07-29	f	41.9103083	-87.675261	Private room in the heart of the city	Wicker Park	1	1
19	14648556	4.59511985	Townhouse	Private room	["Internet", "Wireless Internet", "Cable TV"]	2	2	Real Bed	strict	TRUE	Boston	This is a nice in the heart of the city	t	t	2014-07-27	f	42.339194	-71.049672	Comfy room in the heart of the city	South Boston	1	1
20	17589436	4.88280192	Apartment	Entire home/apt	["TV", "Cable TV"]	2	1	Real Bed	strict	TRUE	NYC	I am renting in the heart of the city	t	t	2015-08-17	f	40.7190863	-73.990285	Amazing LES in the heart of the city	Lower East Side	1	2
21	18224863	4.59511985	House	Entire home/apt	["TV", "Wireless Internet", "Cable TV"]	8	1	Real Bed	strict	TRUE	NYC	Huge 3 bedro in the heart of the city	f	f	2017-05-09	t	40.6437726	-73.950855	3 Bedroom A in the heart of the city	Flatbush	3	3
22	7264511	4.38202664	House	Private room	["Cable TV", "Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	LA	Lovely 1 BR in the heart of the city	t	t	2016-01-11	f	34.1657949	-118.44446	Chill out after in the heart of the city	Sherman Oaks	1	1
23	16679342	3.68887945	House	Private room	["Internet", "Air Conditioning", "Cable TV"]	2	1.5	Real Bed	strict	TRUE	NYC	Cozy room in the heart of the city	t	t	2014-08-27	f	40.6562778	-73.922645	THE LIBRARY in the heart of the city	East Flatbush	1	1
24	3563677	4.90527478	Apartment	Entire home/apt	["TV", "Cable TV"]	4	1	Real Bed	moderate	TRUE	LA	My place is in the heart of the city	t	t	2014-11-28	t	34.0470025	-118.26735	LOFT WITH V in the heart of the city	Downtown	1	1
25	17615783	4.00733319	House	Private room	["Internet", "Wireless Internet", "Cable TV"]	2	1	Real Bed	moderate	TRUE	LA	A cozy home in the heart of the city	t	t	2012-07-12	f	34.1753601	-118.4309	Private Rm. in the heart of the city	Valley Glen	1	1
26	2479317	6.95645444	Apartment	Entire home/apt	["TV", "Wireless Internet", "Cable TV"]	4	1	Real Bed	moderate	TRUE	DC	Our place is in the heart of the city	f	f	2015-05-19	f	38.9112753	-77.037272	Dupont 1 bed in the heart of the city	Dupont Circle	1	2
27	14122244	5.00394631	Apartment	Entire home/apt	["TV", "Cable TV"]	6	1	Real Bed	strict	TRUE	NYC	My place is in the heart of the city	t	t	2017-01-10	f	40.6920573	-73.745464	JFK LUXURIO in the heart of the city	Jamaica	3	3
28	19407360	4.55387689	Apartment	Entire home/apt	["TV", "Wireless Internet", "Cable TV"]	2	1	Real Bed	flexible	TRUE	LA	Retro, Melros in the heart of the city	t	t	2013-06-20	f	34.1006865	-118.3642	Retro West Hollywood Stud	West Hollywood	1	1
29	14490881	5.00394631	Apartment	Entire home/apt	["TV", "Cable TV"]	4	1	Real Bed	strict	TRUE	NYC	This beautiful in the heart of the city	f	f	2016-05-25	f	40.7301449	-73.851542	Luxurious 2 b in the heart of the city	Forest Hills	2	2
30	851978	5.19295685	Apartment	Entire home/apt	["TV", "Internet", "Cable TV"]	3	1	Real Bed	flexible	TRUE	LA	This home is in the heart of the city	t	t	2015-09-26	f	34.0515841	-118.24292	Sweet Hollywood Home / 1 in the heart of the city	West Hollywood	1	2
31	4734170	4.65396035	Apartment	Private room	["TV", "Internet", "Cable TV"]	3	1	Real Bed	strict	TRUE	NYC	Nice, cozy roc in the heart of the city	f	f	2016-03-06	f	40.7503545	-73.980828	CENTER OF N in the heart of the city	Murray Hill	1	2
32	421056	4.31748811	Apartment	Private room	["TV", "Cable TV"]	2	1	Real Bed	strict	TRUE	NYC	The apartment in the heart of the city	t	t	2014-01-06	f	40.6612209	-73.948667	Modern Roo in the heart of the city	Lefferts Gardens	1	1
33	5520582	3.8286414	Apartment	Private room	["Wireless Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	LA	Our place is in the heart of the city	f	f	2017-03-18	f	34.0640766	-118.34591	Cosy private in the heart of the city	Mid-Wilshire	1	1
34	8908966	5.22035583	Apartment	Entire home/apt	["TV", "Cable TV"]	4	1	Real Bed	strict	TRUE	LA	Welcome to in the heart of the city	t	t	2011-10-07	f	33.997506	-118.47203	Light Bright V in the heart of the city	Venice	2	3
35	10994968	4.48863637	House	Private room	["Internet", "Wireless Internet", "Cable TV"]	1	1	Real Bed	moderate	TRUE	LA	LOVELY ROO in the heart of the city	t	t	2011-04-01	f	34.0863378	-118.36478	See Hollywoc in the heart of the city	West Hollywood	1	1
36	5559765	4.35670883	Apartment	Entire home/apt	["TV", "Internet", "Cable TV"]	4	1	Real Bed	flexible	TRUE	LA	Centrally loca in the heart of the city	f	f	2016-10-24	t	34.0863086	-118.27123	Hollywood SpAciouS 1 BeD	East Hollywood	1	2
37	15705862	4.38202664	Apartment	Private room	["TV", "Cable TV"]	2	1	Real Bed	moderate	TRUE	NYC	Guest are prc in the heart of the city	t	t	2012-08-07	t	40.6592186	-73.913513	Intimate Isol in the heart of the city	Brownsville	1	1
38	16904896	4.82831374	House	Private room	["TV", "Cable TV"]	2	1	Real Bed	strict	TRUE	SF	Beautiful, spa in the heart of the city	t	t	2013-12-26	t	37.7737278	-122.46313	Cozy Giants R in the heart of the city	Richmond District	1	1
39	4209927	3.91202301	Apartment	Private room	["TV", "Wireless Internet", "Cable TV"]	1	1	Real Bed	strict	TRUE	NYC	I'll come back in the heart of the city	f	f	2013-11-19	f	40.7151831	-73.9452	Queen-bed, h in the heart of the city	Williamsburg	1	1
40	12232154	4.74493213	Apartment	Entire home/apt	["Internet", "Wireless Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	NYC	Available July in the heart of the city	t	t	2011-01-06	f	40.731311	-73.986092	Amazing Pre in the heart of the city	East Village	1	1
41	4680055	4.68213123	Condominium	Private room	["TV", "Internet", "Cable TV"]	2	1	Real Bed	strict	TRUE	Boston	Tourists/Conf in the heart of the city	t	t	2013-06-16	t	42.3306283	-71.053148	Private Bedro in the heart of the city	South Boston	1	1
42	19813606	5.52146092	Condominium	Entire home/apt	["TV", "Wireless Internet", "Cable TV"]	5	2	Real Bed	strict	TRUE	Chicago	Modern two in the heart of the city	f	f	2016-07-17	t	41.861057	-87.619303	Modern cond in the heart of the city	South Loop/P	2	2
43	18191678	4.78749174	Apartment	Shared room	["TV", "Air conditioning", "Cable TV"]	4	1	Real Bed	moderate	TRUE	LA	This apartme in the heart of the city	f	f	2016-09-17	f	34.1078028	-118.32007	Luxury Apt Complex in the heart of the city	West Hollywood	1	1
44	12428036	4.31748811	Apartment	Private room	["Internet", "Wireless Internet", "Cable TV"]	8	1	Real Bed	moderate	TRUE	LA	This large roc in the heart of the city	t	t	2015-07-13	t	34.0554234	-118.2752	Spacious Roo in the heart of the city	Westlake	1	7
45	4014942	4.99721227	Apartment	Entire home/apt	["TV", "Internet", "Cable TV"]	4	1.5	Real Bed	moderate	TRUE	LA	Very unique in the heart of the city	t	t	2014-08-24	f	34.1025979	-118.35678	BEST HOLLYV in the heart of the city	Hollywood Hills	2	2

data

+



# Data Acquisition and Cleansing

- When data was examined it contained some missing values.
- All the missing values were filled.
- Once column contained more than 10% missing values, so column was removed.
- Missing values were filled using the mode.

```
In [6]: 1 df_data.isnull().sum()
```

```
Out[6]: id                0
log_price                0
property_type            0
room_type                0
amenities                0
accommodates             0
bathrooms               200
bed_type                 0
cancellation_policy      0
cleaning_fee             0
city                    0
description              0
host_has_profile_pic     188
host_identity_verified   188
host_since               188
instant_bookable         0
latitude                 0
longitude                0
name                     0
neighbourhood            6872
bedrooms                 91
beds                     131
dtype: int64
```

# EDA and Insights

- After cleaning the data
- Independent features were plotted against dependent features using bar chart.
- Sweetviz library was used to see visualisation.

## Keypoints:

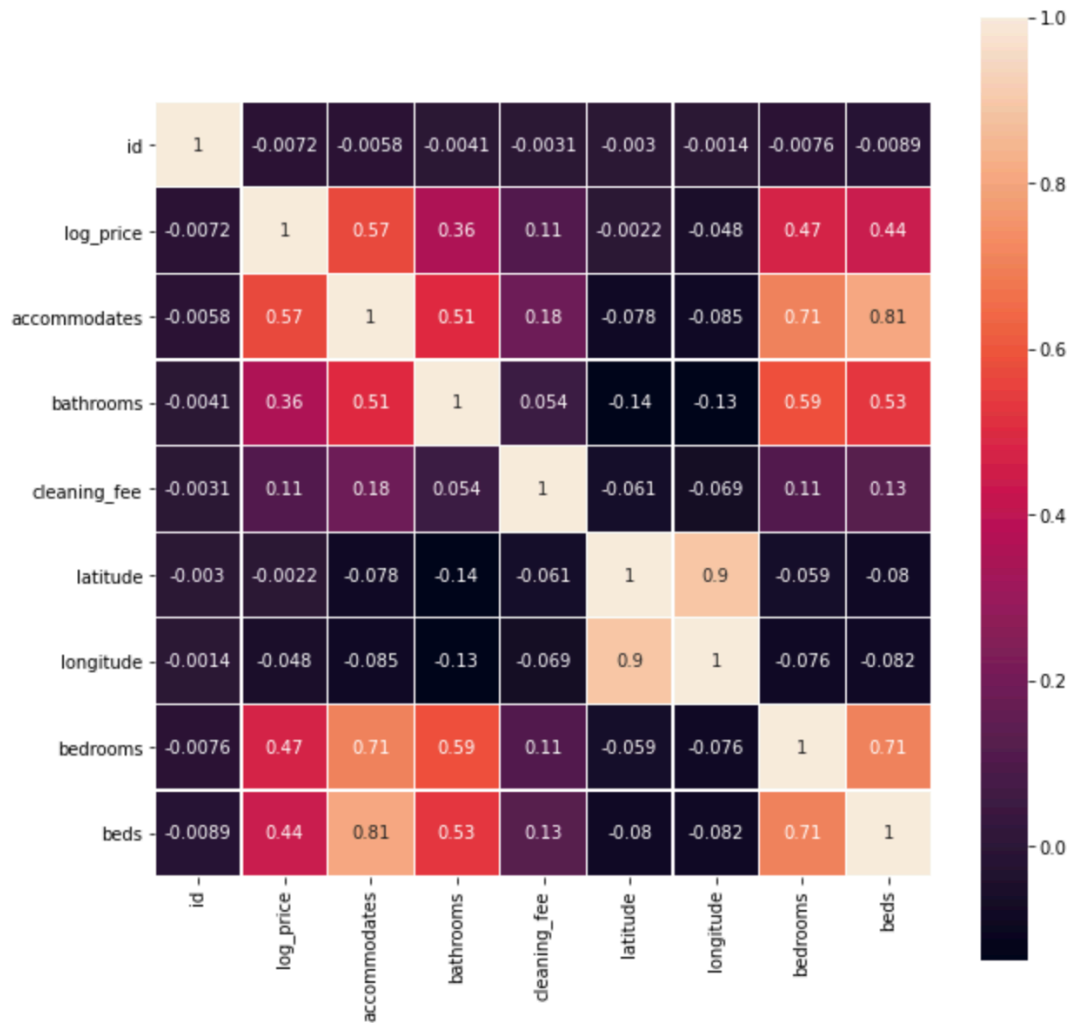
- 60% of property available on Airbnb in US is apartment.
- 56% of room type on Airbnb in US is entire home/apt.
- 40% of properties can make accomodation for 2 peoples.
- 80% of properties has 1 washroom.
- Almost all the property is furnished and has atleast bed.
- 44% of host charges for cancellation.
- 73% of host charges cleaning fees.
- 70% of records are from new york and los angeles whereas new york contain 40% of all roecords.
- 26% of host provide instant booking service.
- 67% properties contain one bedrooms.
- 60% properties has just one bed in the property.

## Important observations:

- Around 40% of the records are from new york city.
- Around 70% of the records property type is apartment.
- Around 95% of the records contains bed\_type as real bed.
- 99% of host has profile picture.
- 60% host identity is verified.

## Summary:

- Sanfrancisco is expensive i.e have high prices among other states.
- Different type of property has different price.
- Prices increase linearly with increase in number of accommodates.
- Entire home/apt is expensive among private room and shared room.
- Prices increase with increase in number of bedrooms.
- Host id verified or not does not effect the price.
- Host has profile picture also does not effect the price as we have just 226 cases of no profile picture.
- Instant booking feature does not effect the price as it is similar for both cases (i.e available or not).



# Feature Engineering

- Features on which the price was not dependent were eliminated.
- As two features which were highly correlated to one another, one feature was removed and one was kept.

# ML Algorithm

- The algorithm we will be using to create a model is Linear regression and Polynomial regression.
- Since we need to predict the real values.
- Few independent features has linear relationship with dependent variables.

# ML Model

- Three models were created using different algorithms:
  - Linear regression
  - Polynomial regression with degree 2
  - Polynomial regression with degree 3
- Model with polynomial regression with degree 3 has highest accuracy score of 60%. So, we will discard other two models.



# Conclusion

Model created can be used by person interested in registering their property with AirBnb to predict the price of their property. The UI can be designed and connected with the model in future so that the customer can smoothly use the model with the interactive UI.