

In []:

```
import pandas as pd
```

In []:

```
data = pd.read_csv("/content/Train.csv")
```

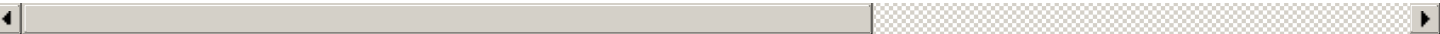
In []:

```
data
```

Out[]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY_TO_MOVE	RESALE	
0	Owner	0	0	2	BHK	1300.236407	1	1	Ksf
1	Dealer	0	0	2	BHK	1275.000000	1	1	
2	Owner	0	0	2	BHK	933.159722	1	1	
3	Owner	0	1	2	BHK	929.921143	1	1	
4	Dealer	1	0	2	BHK	999.009247	0	1	
...	
29446	Owner	0	0	3	BHK	2500.000000	1	1	
29447	Owner	0	0	2	BHK	769.230769	1	1	
29448	Dealer	0	0	2	BHK	1022.641509	1	1	
29449	Owner	0	0	2	BHK	927.079009	1	1	Shol
29450	Dealer	0	1	2	BHK	896.774194	1	1	

29451 rows × 12 columns



In []:

```
data.describe()
```

Out[]:

	UNDER_CONSTRUCTION	RERA	BHK_NO.	SQUARE_FT	READY_TO_MOVE	RESALE	LONGITUDE	
count	29451.000000	29451.000000	29451.000000	2.945100e+04	29451.000000	29451.000000	29451.000000	2
mean	0.179756	0.317918	2.392279	1.980217e+04	0.820244	0.929578	21.300255	
std	0.383991	0.465675	0.879091	1.901335e+06	0.383991	0.255861	6.205306	
min	0.000000	0.000000	1.000000	3.000000e+00	0.000000	0.000000	-37.713008	
25%	0.000000	0.000000	2.000000	9.000211e+02	1.000000	1.000000	18.452663	
50%	0.000000	0.000000	2.000000	1.175057e+03	1.000000	1.000000	20.750000	
75%	0.000000	1.000000	3.000000	1.550688e+03	1.000000	1.000000	26.900926	
max	1.000000	1.000000	20.000000	2.545455e+08	1.000000	1.000000	59.912884	



In []:

```
data.isna().sum
```

Out[]:

```
<bound method NDFrame._add_numeric_operations.<locals>.sum of
STRUCTION  RERA  BHK_NO.  BHK_OR_RK  SQUARE_FT  \
0          False          False  False  False  False  False
1          False          False  False  False  False  False
2          False          False  False  False  False  False
3          False          False  False  False  False  False
4          False          False  False  False  False  False
...          ...          ...    ...    ...    ...    ...
29446       False          False  False  False  False  False
29447       False          False  False  False  False  False
29448       False          False  False  False  False  False
29449       False          False  False  False  False  False
29450       False          False  False  False  False  False
```

```
READY_TO_MOVE  RESALE  ADDRESS  LONGITUDE  LATITUDE  \
0          False  False  False  False  False
1          False  False  False  False  False
2          False  False  False  False  False
3          False  False  False  False  False
4          False  False  False  False  False
...          ...    ...    ...    ...    ...
29446       False  False  False  False  False
29447       False  False  False  False  False
29448       False  False  False  False  False
29449       False  False  False  False  False
29450       False  False  False  False  False
```

```
TARGET(PRICE_IN_LACS)
0          False
1          False
2          False
3          False
4          False
...          ...
29446       False
29447       False
29448       False
29449       False
29450       False
```

```
[29451 rows x 12 columns]>
```

```
In [ ]:
```

```
data.isna().sum()
```

```
Out[ ]:
```

```
POSTED_BY          0
UNDER_CONSTRUCTION 0
RERA                0
BHK_NO.            0
BHK_OR_RK          0
SQUARE_FT          0
READY_TO_MOVE      0
RESALE              0
ADDRESS             0
LONGITUDE           0
LATITUDE            0
TARGET(PRICE_IN_LACS) 0
dtype: int64
```

```
In [ ]:
```

```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [ ]:
```

```
sns.stripplot(y=data["LONGITUDE"])
```

