

NAME : Alok Kadtan

DIV : CS6

PRN : 202401080034

ROLL NO. : CS6-68

BATCH : CS63

topic : house price

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

Share

RAM Disk

[54741 rows x 21 columns]

[22] print("Properties by status:\n", df['Status'].value_counts())

Properties by status:

Status

Ready to Move 186916

Name: count, dtype: int64

[24] park_overlook = df[df['overlooking'].str.contains('Park', na=False)]
print("Properties overlooking park/garden:", park_overlook.shape[0])

Properties overlooking park/garden: 71738

[25] avg_price_ownership = df.groupby('Ownership')['Price (in rupees)'].mean()
print("Average price by ownership:\n", avg_price_ownership)

Average price by ownership:

Ownership

Co-operative Society 8314.178514

Freehold 8072.731045

Leasehold 7470.194874

Power Of Attorney 8604.022133

Name: Price (in rupees), dtype: float64

mands + Code + Text

4	TenX Habitat	Raymond Realty	2	NaN	1 Covered
---	--------------	----------------	---	-----	-----------

Ownership Super Area Dimensions Plot Area

0	NaN	NaN	NaN	NaN
1	Freehold	NaN	NaN	NaN
2	Freehold	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	Co-operative Society	NaN	NaN	NaN

[5 rows x 21 columns]

```
[4] import zipfile
```

```
# Unzip the uploaded file
with zipfile.ZipFile('house_prices.csv.zip', 'r') as zip_ref:
    zip_ref.extractall()
```

```
[6] import pandas as pd
```

```
# Load the data
df = pd.read_csv('/content/house_prices.csv') # Adjust path if needed
df.columns = df.columns.str.strip() # Clean spaces in column names
```

```
print("Total properties listed:", df.shape[0])
```

Total properties listed: 187531

Colab.research.google.com/drive/123...

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text

Share Gemini

RAM Disk

0s

[30] most_common_furnishing = df['Furnishing'].mode()[0]
print("Most common furnishing type:", most_common_furnishing)

Most common furnishing type: Semi-Furnished

0s

[31] df['Carpet Area'] = pd.to_numeric(df['Carpet Area'], errors='coerce')
df['Super Area'] = pd.to_numeric(df['Super Area'], errors='coerce')
wrong_area = df[df['Carpet Area'] > df['Super Area']]
print("Properties with carpet area greater than super area:\n", wrong_area)

Properties with carpet area greater than super area:
Empty DataFrame
Columns: [Index, Title, Description, Amount(in rupees), Price (in rupees), location, Carpet Area, Status, Floor, Transaction, Furnishing, facing, overlooking, Society, Bathroom, Balco
Index: []

[0 rows x 21 columns]

0s

df['Carpet Area'] = pd.to_numeric(df['Carpet Area'], errors='coerce')
df['Super Area'] = pd.to_numeric(df['Super Area'], errors='coerce')
wrong_area = df[df['Carpet Area'] > df['Super Area']]
print("Properties with carpet area greater than super area:\n", wrong_area)

Properties with carpet area greater than super area:
Empty DataFrame
Columns: [Index, Title, Description, Amount(in rupees), Price (in rupees), location, Carpet Area, Status, Floor, Transaction, Furnishing, facing, overlooking, Society, Bathroom, Balco
Index: []

FileEditViewInsertRuntimeToolsHelp

Commands+ Code+ Text

[1] from google.colab import files
uploaded = files.upload()

Choose Files

house_prices.csv.zip

- house_prices.csv.zip(application/x-zip-compressed) - 6926090 bytes, last modified: 4/29/2025 - 100% done

Saving house_prices.csv.zip to house_prices.csv.zip

import pandas as pd

df = pd.read_csv('house_prices.csv.zip') # make sure the file name matches exactly
print(df.head()) # Check first 5 rows

Index

title \

0 0 1 BHK Ready to Occupy Flat for sale in Srushti...

1 1 2 BHK Ready to Occupy Flat for sale in Dosti V...

2 2 2 BHK Ready to Occupy Flat for sale in Sunrise...

3 3 1 BHK Ready to Occupy Flat for sale Kasheli

4 4 2 BHK Ready to Occupy Flat for sale in TenX Ha...

Description

Amount(in rupees) \

0 Bhiwandi, Thane has an attractive 1 BHK Flat f...

1 One can find this stunning 2 BHK flat for sale...

2 Up for immediate sale is a 2 BHK apartment in ...

3 This beautiful 1 BHK Flat is available for sal...

4 This lovely 2 BHK Flat in Pokhran Road, Thane ...

Price (in rupees)

location

Carpet Area

Status

Floor \

0 6000.0 thane 500 sqft Ready to Move 10 out of 11

1 13799.0 thane 473 soft Ready to Move 3 out of 22



```
df['Balcony'] = pd.to_numeric(df['Balcony'], errors='coerce')
```

```
# First, convert Balcony to numeric
df['Balcony'] = pd.to_numeric(df['Balcony'], errors='coerce')
```

```
# Then filter
```

```
more_balcony = df[df['Balcony'] > 2]
```

```
print("Properties with more than 2 balconies:\n", more_balcony)
```

```
187522 Up for immediate sale is a 3 BHK apartment in ...      80 Lac
187523 This magnificent 5 BHK Flat is available for s...    1.19 Cr
187526 This magnificent 3 BHK Flat is available for s...     63 Lac
187530 This exquisite 3 BHK Flat is offered for sale ...    1.18 Cr
```

	Price (in rupees)	location	Carpet Area	Status	Floor
28	14344.0	thane	925 sqft	Ready to Move	8 out of 20
34	9643.0	thane	440 sqft	Ready to Move	2 out of 7
35	3768.0	thane	NaN	Ready to Move	5 out of 7
39	4070.0	thane	600 sqft	Ready to Move	2 out of 7
54	11905.0	thane	636 sqft	Ready to Move	10 out of 10
...
187521	5816.0	zirakpur	NaN	Ready to Move	2 out of 11
187522	4040.0	zirakpur	1200 sqft	Ready to Move	3 out of 3
187523	4547.0	zirakpur	1705 sqft	Ready to Move	5 out of 10
187526	3225.0	zirakpur	NaN	Ready to Move	2 out of 4
187530	6162.0	zirakpur	NaN	Ready to Move	5 out of 13

	Transaction	...	facing	overlooking
28	Resale	...	East	Garden/Park
34	Resale	...	East	Garden/Park

Commands + Code + Text

```
[12] ready_to_move = df[df['Status'] == 'Ready to move']  
     print("Ready to move properties:", ready_to_move.shape[0])
```

```
Ready to move properties: 0
```

```
[13] new_transaction = df[df['Transaction'] == 'New Property']  
     print("New transaction properties:", new_transaction.shape[0])
```

```
New transaction properties: 42565
```

```
[14] new_transaction = df[df['Transaction'] == 'New Property']  
     print("New transaction properties:", new_transaction.shape[0])
```

```
New transaction properties: 42565
```

```
print("Top 5 locations:\n", df['location'].value_counts().head(5))
```

```
Top 5 locations:  
location  
new-delhi    27599  
bangalore    24030  
kolkata      22380
```

```
[34] resale_properties = df[df['Transaction'] == 'Resale']  
     print("Number of resale properties:", resale_properties.shape[0])
```

➞ Number of resale properties: 144172

```
[35] df['Amount(in rupees)'] = pd.to_numeric(df['Amount(in rupees)'], errors='coerce')  
     ready_avg_amount = df[df['Status'] == 'Ready to move']['Amount(in rupees)'].mean()  
     print("Average amount of Ready to Move properties:", ready_avg_amount)
```

➞ Average amount of Ready to Move properties: nan

```
[36] freehold_count = df[df['Ownership'] == 'Freehold'].shape[0]  
     print("Number of Freehold ownership properties:", freehold_count)
```

➞ Number of Freehold ownership properties: 112229

▶ Start coding or generate with AI.

Commands | + Code + Text

↗ [54741 rows x 21 columns]

✓ [22] print("Properties by status:\n", df['Status'].value_counts())

↗ Properties by status:
Status
Ready to Move 186916
Name: count, dtype: int64

✓ [24] park_overlook = df[df['overlooking'].str.contains('Park', na=False)]
print("Properties overlooking park/garden:", park_overlook.shape[0])

↗ Properties overlooking park/garden: 71738

▶ avg_price_ownership = df.groupby('Ownership')['Price (in rupees)'].mean()
print("Average price by ownership:\n", avg_price_ownership)

↗ Average price by ownership:
Ownership
Co-operative Society 8314.178514
Freehold 8072.731045
Leasehold 7470.194874
Power Of Attorney 8604.022133
Name: Price (in rupees), dtype: float64

```
east_facing = df[df['facing'] == 'East']
print("East facing properties: \n", east_facing)
```

```
187514 this lovely 3 BHK flat in Gazipur, zirakpur is... 65 Lac
187522 up for immediate sale is a 3 BHK apartment in ... 80 Lac
187526 This magnificent 3 BHK flat is available for 5... 63 Lac
187528 Gazipur, Zirakpur has an appealing 3 BHK flat ... 76 Lac
```

	Price (in rupees)	Location	Carpet Area	Status	Floor
1	13799.0	thane	473 sqft	Ready to Move	3 out of 22
2	17500.0	thane	779 sqft	Ready to Move	10 out of 29
5	6618.0	thane	NaN	Ready to Move	2 out of 7
9	11150.0	thane	900 sqft	Ready to Move	3 out of 27
13	17526.0	thane	1820 sqft	Ready to Move	18 out of 27
...
187512	3810.0	zirakpur	950 sqft	Ready to Move	3 out of 3
187514	3939.0	zirakpur	1200 sqft	Ready to Move	1 out of 4
187522	4040.0	zirakpur	1200 sqft	Ready to Move	3 out of 3
187526	3225.0	zirakpur	NaN	Ready to Move	2 out of 4
187528	4343.0	zirakpur	1250 sqft	Ready to Move	1 out of 3

	Transaction	facing	overlooking
1	Resale	...	East
2	Resale	...	East
5	Resale	...	East
9	Resale	...	East
13	Resale	...	East
...
187512	Resale	...	East
187514	New Property	...	East

```
[8] print("Unique locations:", df['location'].nunique())
```

```
Unique locations: 81
```

```
[9] print("Average price:", df['Price (in rupees)'].mean())
```

```
Average price: 7583.771884897507
```

```
print("Property with max amount:\n", df.loc[df['Amount(in rupees)'].idxmax()])
```

```
Property with max amount:
```

Index	24
Title	1 BHK Ready to Occupy Flat for sale Mumbra
Description	This ready to move-in 1 BHK flat is available ...
Amount(in rupees)	Call for Price
Price (in rupees)	NaN
location	thane
Carpet Area	NaN
Status	Ready to Move
Floor	6 out of 7
Transaction	Resale
Furnishing	Unfurnished
facing	NaN
overlooking	NaN
Society	NaN
Bathroom	1
Balcony	NaN
Car Parking	NaN