

My IBM x *New story x New report x Technical Training Session x Data Analytics Session 41 x Untitled4.ipynb - Colaboratory x +

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=M1-i-3fLTXa

Gmail YouTube Maps top programs New Tab (727) YouTube

Untitled4.ipynb ☆

File Edit View Insert Runtime Tools Help Saving...

Comment Share

Files

- ..
- sample_data
- House Price India.csv

+ Code + Text

```
[1] import pandas as pd
```

```
df = pd.read_csv('/content/House Price India.csv')
df.head()
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0

5 rows x 23 columns

House Price India.csv x

1 to 10 of 14620 entries Filter

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810998	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show 10 per page

1 2 10 100 1000 1400 1450 1462

0s completed at 11:06 AM

Type here to search

archi... archi... Hou... Unti... Turb... Scre... Wha... AI A... ENG 11:06 04-10-2023

My IBM | *New story | *New report | Technical Training Session | Data Analytics Session 4 | Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=M1-i-3fLTXa

Google | Gmail | YouTube | Maps | top programs | New Tab | (727) YouTube

Untitled4.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

- sample_data
- House Price India.csv

+ Code + Text

```
[1] import pandas as pd
```

```
df = pd.read_csv('/content/House Price India.csv')
df.head()
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0

5 rows x 23 columns

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810998	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show 10 per page

1 2 10 100 1000 1400 1450 1462

0s completed at 11:06 AM

Type here to search

archi... archi... Hou... Unti... Turb... Scre... Wha... AI A... ENG 11:06 04-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9fYb?authuser=0#scrollTo=M1-i-3fLTXa

GmailYouTubeMapsTop programsNew Tab(727) YouTube

Untitled4.ipynb

FileEditViewInsertRuntimeToolsHelpAll changes saved

CommentShare

Files

sample_dataHouse

DownloadRename fileDelete fileCopy pathRefresh

```
[1] import pandas as pd

df = pd.read_csv('/content/House Price India.csv')
df.head()
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0

5 rows × 23 columns

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810998	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show 10 per page

12101001000140014501462

completed at 11:06 AM

Type here to search

archi...archi...Hou...Unti...Turb...Scre...Wha...AI A...ENG11:0704-10-2023

My IBM x New story x New report x Technical Training Session x Data Analytics Session 41 x Untitled4.ipynb - Colaboratory x

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=weGhI-CnPg2w

Gmail YouTube Maps top programs New Tab (727) YouTube

Untitled4.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- House Price India.csv

+ Code + Text

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14620 entries, 0 to 14619
Data columns (total 23 columns):
 #   Column              Non-Null Count  Dtype  
---  -
 0   id                  14620 non-null  int64  
 1   Date                14620 non-null  int64  
 2   number of bedrooms  14620 non-null  int64  
 3   number of bathrooms 14620 non-null  float64 
 4   living area         14620 non-null  int64  
 5   lot area            14620 non-null  int64  
 6   number of floors     14620 non-null  float64 
 7   waterfront present   14620 non-null  int64  
 8   number of views      14620 non-null  int64  
 9   condition of the house 14620 non-null  int64  
10  grade of the house   14620 non-null  int64  
11  Area of the house(excluding basement) 14620 non-null  int64  
12  Area of the basement 14620 non-null  int64  
13  Built Year           14620 non-null  int64  
14  Renovation Year       14620 non-null  int64  
15  Postal Code           14620 non-null  int64  
16  Latitude              14620 non-null  float64 
17  Longitude              14620 non-null  float64 
18  living_area_renov      14620 non-null  int64  
19  lot_area_renov         14620 non-null  int64  
20  Number of schools nearby 14620 non-null  int64  
21  Distance from the airport 14620 non-null  int64  
22  Price                 14620 non-null  int64  
dtypes: float64(4), int64(19)
memory usage: 2.6 MB
```

House Price India.csv

1 to 10 of 14620 entries Filter

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810898	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show 10 per page

1 2 10 100 1000 1400 1460 1462

11:22 04-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=VRsMhgtQRDgh

GmailYouTubeMaps

top programs

New Tab

(727) YouTube

Untitled4.ipynb

FileEditViewInsertRuntimeToolsHelp

CommentShare

Files

sample_data

House Price India.csv

```
plt.figure(figsize=(8,8))
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=140)
plt.title('sample pie Chart')
Text(0.5, 1.0, 'sample pie Chart')
```

sample pie Chart

Category	Percentage
Category C	38.2%
Category B	33.7%
Category A	16.9%
Category D	11.2%

House Price India.csv

1 to 10 of 14520 entries

id	Date	number of bedrooms	number of bathrooms	living area	lot area
6762810145	42491	5	2.5	3660	9050
6762810635	42491	4	2.5	2920	4000
6762810698	42491	5	2.75	2910	9480
6762812605	42491	4	2.5	3310	42880
6762812519	42491	3	2	2710	4500
6762813105	42491	3	2.5	2800	4750
6762813167	42491	5	3.25	3860	11995
6762813659	42491	3	1.75	2240	10578
6762813800	42491	3	2.5	2300	6650
6762814451	42491	4	2.25	2200	11250

Show 10 per page

0s completed at 11:31 AM

Type here to search

archi... archi... Hou... Unti... Turb... Scre... Wha... AI A...

11:31 04-10-2023

id	Date	number of bedrooms	number of bathroom	living area	lot area	number of floors
6762870146	42491 5	2.5	3600	9000	2	
6762810555	42491 4	2.5	2900	4000	1.5	
6762810068	42491 5	2.75	2910	9450	1.5	
6762812006	42491 4	2.5	3910	4200	2	
6762812519	42491 3	2	2700	4500	1.5	
6762813105	42491 3	2.5	2600	4750	1	
6762813157	42491 5	0.25	3600	11995	2	
6762813589	42491 3	1.75	2240	10578	2	
6762813600	42491 3	2.5	2390	6650	1	
6762814461	42491 4	2.25	2200	12550	1.5	

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TFFh3DO7ryVK8bopENFXKdF9fYb?authuser=0#scrollTo=VRsMhgtQRDgh

GmailYouTubeMapsTop programsNew Tab(727) YouTube

100%Reset

CommentShare

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

sample_dataHouse Price India.csv

CodeText

```
plt.figure(figsize=(8,8))
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=140)
plt.title('sample pie chart')
Text(0.5, 1.0, 'sample pie chart')
```

sample pie Chart

Category D

11.2%

Category C

38.2%

Category A

16.9%

33.7%

House Price India.csv

1 to 10 of 14620 entriesFilter

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810898	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show10per page

12101001000140014601462

completed at 11:31 AM

Type here to search

archi...archi...Hou...Unti...Turb...Scre...Wha...AI A...ENG11:3204-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9fYb?authuser=0#scrollTo=0R253h7LSA6Q

GmailYouTubeMapsTop programsNew Tab(727) YouTube

Colaboratory

FileEditViewInsertRuntimeToolsHelpAll changes saved

CommentShare

RAMDisk

Files

sample_dataHouse Price India.csv

[9]

CategoryB

```
plt.axis('equal')
plt.show()
```

1s completed at 11:33 AM

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	4/24/91	5	2.5
6762810635	4/24/91	4	2.5
6762810898	4/24/91	5	2.75
6762812605	4/24/91	4	2.5
6762812919	4/24/91	3	2
6762813105	4/24/91	3	2.5
6762813157	4/24/91	5	3.25
6762813599	4/24/91	3	1.75
6762813600	4/24/91	3	2.5
6762814481	4/24/91	4	2.25

Show 10 per page

12101001000140014601462

Type here to search

archi...archi...Hou...Unti...Turb...Scre...Wha...AI A...ENG11:3304-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bopENFXKdF9Yb?authuser=0#scrollTo=XUOG_dHG5dP

GmailYouTubeMaps

top programs

New Tab

(727) YouTube

Untitled4.ipynb

FileEditViewInsertRuntimeToolsHelp

All changes saved

CommentShare

RAMDisk

Files

sample_data

House Price India.csv

+ Code

+ Text

```
plt.figure(figsize=(8,6)) # Set the figure size (optional)
plt.scatter(x, y, c='blue', marker='o', label='Data Points') # Scatter plot
plt.xlabel('Variable1') # X-axis label
plt.ylabel('Variable2') # Y-axis label
plt.title('Scatter Plot of Variable1 vs. Variable2') # Title (optional)
plt.grid(True) # Display grid (optional)
plt.legend() # Display legend (optional)

# Show the plot
plt.show()
```

Scatter Plot of Variable1 vs. Variable2

House Price India.csv

1 to 10 of 14520 entries

id	Date	number of bedrooms	number of bathrooms	living area	lot area	r
6762810145	42491	5	2.5	3660	9050	2
6762810635	42491	4	2.5	2920	4000	1
6762810698	42491	5	2.75	2910	9480	1
6762812605	42491	4	2.5	3310	42880	2
6762812519	42491	3	2	2710	4500	1
6762813105	42491	3	2.5	2800	4750	1
6762813167	42491	5	3.25	3860	11965	2
6762813699	42491	3	1.75	2240	10578	2
6762813800	42491	3	2.5	2390	6650	1
6762814451	42491	4	2.25	2200	11250	1

Show 10 per page

0s completed at 11:35 AM

archi... archi... Hou... Unti... Turb... Scre... Wha... AI A...

Type here to search

11:35 04-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=XUOG_dHGSIdP

GmailYouTubeMapsTop programsNew Tab(727) YouTube

100%Reset

CommentShare

FileEditViewInsertRuntimeToolsHelpAll changes saved

Files

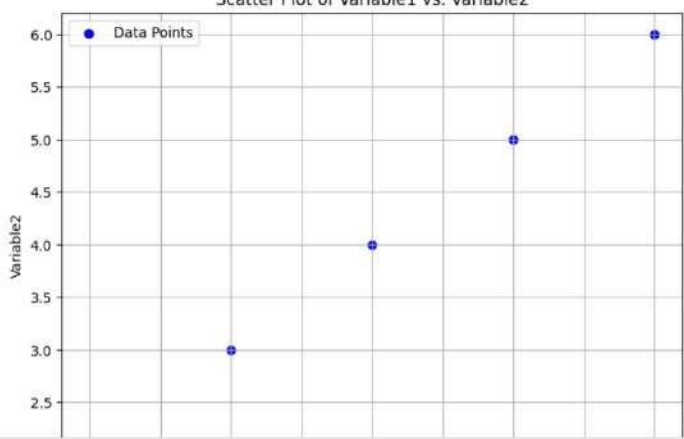
sample_dataHouse Price India.csv

+ Code+ Text

```
plt.ylabel('Variable2') # Y-axis label
plt.title('Scatter Plot of Variable1 vs. Variable2') #
plt.grid(True) # Display grid (optional)
plt.legend() # Display legend (optional)

# Show the plot
plt.show()
```

Scatter Plot of Variable1 vs. Variable2



House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	4/24/91	5	2.5
6762810635	4/24/91	4	2.5
6762810898	4/24/91	5	2.75
6762812605	4/24/91	4	2.5
6762812919	4/24/91	3	2
6762813105	4/24/91	3	2.5
6762813157	4/24/91	5	3.25
6762813599	4/24/91	3	1.75
6762813600	4/24/91	3	2.5
6762814481	4/24/91	4	2.25

Show10per page

12101001000140014601462

completed at 11:35AM

Type here to search

arch...arch...Hou...Unti...Turb...Scre...Wha...AI A...ENG

11:3604-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=XUOG_dHGSIdP

GmailYouTubeMapsTop programsNew Tab(727) YouTube

100%Reset

CommentShare

RAMDisk

Files

sample_dataHouse Price India.csv

Scatter Plot of Variable1 vs. Variable2

Data Points

Variable1	Variable2
1.0	2.0
2.0	3.0
3.0	4.0
4.0	5.0
5.0	6.0

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	4/24/91	5	2.5
6762810635	4/24/91	4	2.5
6762810898	4/24/91	5	2.75
6762812605	4/24/91	4	2.5
6762812919	4/24/91	3	2
6762813105	4/24/91	3	2.5
6762813157	4/24/91	5	3.25
6762813599	4/24/91	3	1.75
6762813600	4/24/91	3	2.5
6762814481	4/24/91	4	2.25

Show 10 per page

12101001000140014601462

completed at 11:35AM

Type here to search

arch...arch...Hou...Unti...Turb...Scre...Wha...AI A...ENG

11:3604-10-2023

Windows taskbar showing the search bar and various application icons (e.g., File Explorer, Microsoft Edge, Google Chrome, etc.). The system clock displays 11:49 on 04-10-2023.

The screenshot displays a Google Colab notebook interface. The notebook is titled 'Untitled4.ipynb' and contains a single code cell. The code performs a linear regression analysis on a dataset named 'House Price India.csv'. It uses numpy to create a design matrix X with a column of ones, and sklearn's LinearRegression class to fit the model. The output of the fit method is printed, showing the coefficients and the R-squared value.

```

X = np.column_stack((x1, x2, np.ones(len(x1)))) # Add 1 to x1 and x2
coefficients, residuals, _, _ = np.linalg.lstsq(X, y, rcond=None)

# Print the coefficients
print("coefficients:", coefficients)

# Calculate statistics like R-squared
y_predicted = np.dot(X, coefficients)
sse = np.sum((y - y_predicted) ** 2)
sst = np.sum((y - np.mean(y)) ** 2)
r_squared = 1 - (sse / sst)
print("R-squared:", r_squared)

```

The output of the code cell is:

```

Coefficients: [0.2666667 1.2666667 1.8      ]
R-squared: 0.867816091954023

```

On the right side of the notebook, a preview of the 'House Price India.csv' data is shown. It includes a table with columns: Id, Date, number of bedrooms, and number of bathrooms. The table shows 10 rows of data, with the first row being (6782810145, 42491, 5, 2.5).

Id	Date	number of bedrooms	number of bathrooms
6782810145	42491	5	2.5
6782810835	42491	4	2.5
6782810888	42491	5	2.75
6782812605	42491	4	2.5
6782812919	42491	3	2
6782813105	42491	3	2.5
6782813157	42491	5	3.25
6782813599	42491	3	1.75
6782813800	42491	3	2.5
6782814461	42491	4	2.25

The bottom status bar indicates that the code was completed at 11:49 AM.

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=6Yrn7CqXLZl

GmailYouTubeMapsTop programsNew Tab(727) YouTube

Colaboratory

CommentShare

FileEditViewInsertRuntimeToolsHelpSaving...

Files

sample_dataHouse Price India.csv

+ Code+ Text

```
# Display the first few rows of the dataset
print(df.head())

# Get basic summary statistics for numeric columns
print(df.describe())

# Get information about the dataset, including data types and missing values
print(df.info())
```

variable1variable2

012

123

234

345

456

count5.0000005.000000

mean3.0000004.000000

std1.5811391.581139

min1.0000002.000000

25%2.0000003.000000

50%3.0000004.000000

75%4.0000005.000000

max5.0000006.000000

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 5 entries, 0 to 4

Data columns (total 2 columns):

columnNon-Null countdtype

0 variable15 non-nullint64

1 variable25 non-nullint64

dtypes: int64(2)

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810898	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show10per page

12101001000140014601462

0s completed at 11:55AM

Type here to search

11:5504-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13oU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=6Yrn7CqXLZl

GmailYouTubeMapsTop programsNew Tab(727) YouTube

Colaboratory

CommentShare

FileEditViewInsertRuntimeToolsHelpSaving...

Files

sample_dataHouse Price India.csv

+ Code+ Text

```
# Display the first few rows of the dataset
print(df.head())

# Get basic summary statistics for numeric columns
print(df.describe())

# Get information about the dataset, including data types and missing values
print(df.info())
```

variable1variable2

012

123

234

345

456

count5.0000005.000000

mean3.0000004.000000

std1.5811391.581139

min1.0000002.000000

25%2.0000003.000000

50%3.0000004.000000

75%4.0000005.000000

max5.0000006.000000

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 5 entries, 0 to 4

Data columns (total 2 columns):

columnNon-Null countdtype

0variable15 non-nullint64

1variable25 non-nullint64

dtypes: int64(2)

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810898	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814481	42491	4	2.25

Show10per page

12101001000140014601462

completed at 11:55AM

Type here to search

11:5504-10-2023

My IBM

New story

New report

Technical Training Session

Data Analytics Session 4

Untitled4.ipynb - Colaboratory

colab.research.google.com/drive/13cU07TPPh3DO7ryVK8bqpENFXKdF9Yb?authuser=0#scrollTo=IIQQGHAsXm9-

GmailYouTubeMapsTop programsNew Tab(727) YouTube

Colaboratory

CommentShare

RAMDisk

Files

sample_dataHouse Price India.csv

Display the first few rows of the dataset
print(df.head())

Get basic summary statistics for numeric columns
print(df.describe())

Get information about the dataset, including data types and missing values
print(df.info())

	variable1	variable2
0	1	2
1	2	3
2	3	4
3	4	5
4	5	6

	variable1	variable2
count	5.000000	5.000000
mean	3.000000	4.000000
std	1.581139	1.581139
min	1.000000	2.000000
25%	2.000000	3.000000
50%	3.000000	4.000000
75%	4.000000	5.000000
max	5.000000	6.000000

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 5 entries, 0 to 4  
Data columns (total 2 columns):  
#   column      Non-Null count  Dtype  
--  --  
0   variable1    5 non-null      int64  
1   variable2    5 non-null      int64
```

House Price India.csv

1 to 10 of 14620 entries

id	Date	number of bedrooms	number of bathrooms
6762810145	42491	5	2.5
6762810635	42491	4	2.5
6762810898	42491	5	2.75
6762812605	42491	4	2.5
6762812919	42491	3	2
6762813105	42491	3	2.5
6762813157	42491	5	3.25
6762813599	42491	3	1.75
6762813600	42491	3	2.5
6762814461	42491	4	2.25

Show 10 per page

12 2 10 100 1000 1400 1460 1462

0s completed at 12:00 PM

Type here to search

04-10-2023 12:00