# **IMPORTING LIBRARIES**

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import seaborn as sns
# Import BaseSettings from pydantic_settings
%pip install ydata_profiling
from ydata_profiling import ProfileReport
```

```
Collecting ydata_profiling
  Using cached ydata_profiling-4.7.0-py2.py3-none-any.whl.metadata (20 kB)
Collecting scipy<1.12,>=1.4.1 (from ydata profiling)
 Using cached scipy-1.11.4-cp312-cp312-win_amd64.whl.metadata (60 kB)
Requirement already satisfied: pandas!=1.4.0,<3,>1.1 in c:\users\surya\appdata\local
\programs\python\python312\lib\site-packages (from ydata_profiling) (2.2.1)
Requirement already satisfied: matplotlib<3.9,>=3.2 in c:\users\surya\appdata\local
\programs\python\python312\lib\site-packages (from ydata_profiling) (3.8.3)
Requirement already satisfied: pydantic>=2 in c:\users\surya\appdata\local\programs
\python\python312\lib\site-packages (from ydata_profiling) (2.6.3)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in c:\users\surya\appdata\local\pr
ograms\python\python312\lib\site-packages (from ydata_profiling) (6.0.1)
Requirement already satisfied: jinja2<3.2,>=2.11.1 in c:\users\surya\appdata\local\p
rograms\python\python312\lib\site-packages (from ydata_profiling) (3.1.3)
Collecting visions<0.7.7,>=0.7.5 (from visions[type_image_path]<0.7.7,>=0.7.5->ydata
_profiling)
 Using cached visions-0.7.6-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: numpy<2,>=1.16.0 in c:\users\surya\appdata\local\prog
rams\python\python312\lib\site-packages (from ydata_profiling) (1.26.4)
Requirement already satisfied: htmlmin==0.1.12 in c:\users\surya\appdata\local\progr
ams\python\python312\lib\site-packages (from ydata_profiling) (0.1.12)
Requirement already satisfied: phik<0.13,>=0.11.1 in c:\users\surya\appdata\local\pr
ograms\python\python312\lib\site-packages (from ydata_profiling) (0.12.4)
Requirement already satisfied: requests<3,>=2.24.0 in c:\users\surya\appdata\local\p
rograms\python\python312\lib\site-packages (from ydata_profiling) (2.31.0)
Requirement already satisfied: tqdm<5,>=4.48.2 in c:\users\surya\appdata\local\progr
ams\python\python312\lib\site-packages (from ydata_profiling) (4.66.2)
Collecting seaborn<0.13,>=0.10.1 (from ydata_profiling)
 Using cached seaborn-0.12.2-py3-none-any.whl.metadata (5.4 kB)
Requirement already satisfied: multimethod<2,>=1.4 in c:\users\surya\appdata\local\p
rograms\python\python312\lib\site-packages (from ydata_profiling) (1.11.2)
Collecting statsmodels<1,>=0.13.2 (from ydata_profiling)
  Using cached statsmodels-0.14.1-cp312-cp312-win_amd64.whl.metadata (9.8 kB)
Collecting typeguard<5,>=4.1.2 (from ydata_profiling)
  Using cached typeguard-4.1.5-py3-none-any.whl.metadata (3.7 kB)
Requirement already satisfied: imagehash==4.3.1 in c:\users\surya\appdata\local\prog
rams\python\python312\lib\site-packages (from ydata_profiling) (4.3.1)
Collecting wordcloud>=1.9.1 (from ydata profiling)
  Using cached wordcloud-1.9.3-cp312-cp312-win_amd64.whl.metadata (3.5 kB)
Collecting dacite>=1.8 (from ydata_profiling)
 Using cached dacite-1.8.1-py3-none-any.whl.metadata (15 kB)
Collecting numba<1,>=0.56.0 (from ydata_profiling)
  Using cached numba-0.59.1-cp312-cp312-win_amd64.whl.metadata (2.8 kB)
Requirement already satisfied: PyWavelets in c:\users\surya\appdata\local\programs\p
ython\python312\lib\site-packages (from imagehash==4.3.1->ydata_profiling) (1.5.0)
Requirement already satisfied: pillow in c:\users\surya\appdata\local\programs\pytho
n\python312\lib\site-packages (from imagehash==4.3.1->ydata_profiling) (10.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\surya\appdata\local\progr
ams\python\python312\lib\site-packages (from jinja2<3.2,>=2.11.1->ydata_profiling)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\surya\appdata\local\prog
rams\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata_profiling)
(1.2.0)
Requirement already satisfied: cycler>=0.10 in c:\users\surya\appdata\local\programs
\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata_profiling) (0.
12.1)
```

Requirement already satisfied: fonttools>=4.22.0 in c:\users\surya\appdata\local\pro grams\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata\_profilin g) (4.49.0)

Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\surya\appdata\local\pro grams\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata\_profilin g) (1.4.5)

Requirement already satisfied: packaging>=20.0 in c:\users\surya\appdata\local\programs\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata\_profiling) (23.2)

Requirement already satisfied: pyparsing>=2.3.1 in c:\users\surya\appdata\local\prog rams\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata\_profiling) (3.1.1)

Requirement already satisfied: python-dateutil>=2.7 in c:\users\surya\appdata\local \programs\python\python312\lib\site-packages (from matplotlib<3.9,>=3.2->ydata\_profiling) (2.8.2)

Collecting llvmlite<0.43,>=0.42.0dev0 (from numba<1,>=0.56.0->ydata\_profiling)
Using cached llvmlite-0.42.0-cp312-cp312-win\_amd64.whl.metadata (4.9 kB)

Requirement already satisfied: pytz>=2020.1 in c:\users\surya\appdata\local\programs \python\python312\lib\site-packages (from pandas!=1.4.0,<3,>1.1->ydata\_profiling) (2 024.1)

Requirement already satisfied: tzdata>=2022.7 in c:\users\surya\appdata\local\progra ms\python\python312\lib\site-packages (from pandas!=1.4.0,<3,>1.1->ydata\_profiling) (2024.1)

Requirement already satisfied: joblib>=0.14.1 in c:\users\surya\appdata\local\progra ms\python\python312\lib\site-packages (from phik<0.13,>=0.11.1->ydata\_profiling) (1.1.1)

Requirement already satisfied: annotated-types>=0.4.0 in c:\users\surya\appdata\loca l\programs\python\python312\lib\site-packages (from pydantic>=2->ydata\_profiling) (0.6.0)

Requirement already satisfied: pydantic-core==2.16.3 in c:\users\surya\appdata\local \programs\python\python312\lib\site-packages (from pydantic>=2->ydata\_profiling) (2. 16.3)

Requirement already satisfied: typing-extensions>=4.6.1 in c:\users\surya\appdata\lo cal\programs\python\python312\lib\site-packages (from pydantic>=2->ydata\_profiling) (4.10.0)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\surya\appdata\lo cal\programs\python\python312\lib\site-packages (from requests<3,>=2.24.0->ydata\_pro filing) (3.3.2)

Requirement already satisfied: idna<4,>=2.5 in c:\users\surya\appdata\local\programs \python\python312\lib\site-packages (from requests<3,>=2.24.0->ydata\_profiling) (3.6)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\surya\appdata\local\pr ograms\python\python312\lib\site-packages (from requests<3,>=2.24.0->ydata\_profilin g) (2.2.1)

Requirement already satisfied: certifi>=2017.4.17 in c:\users\surya\appdata\local\pr ograms\python\python312\lib\site-packages (from requests<3,>=2.24.0->ydata\_profilin g) (2024.2.2)

Collecting patsy>=0.5.4 (from statsmodels<1,>=0.13.2->ydata\_profiling)

Using cached patsy-0.5.6-py2.py3-none-any.whl.metadata (3.5 kB)

Requirement already satisfied: colorama in c:\users\surya\appdata\local\programs\python\python312\lib\site-packages (from tqdm<5,>=4.48.2->ydata\_profiling) (0.4.6)

Requirement already satisfied: attrs>=19.3.0 in c:\users\surya\appdata\local\program s\python\python312\lib\site-packages (from visions<0.7.7,>=0.7.5->visions[type\_image\_path]<0.7.7,>=0.7.5->ydata\_profiling) (23.2.0)

Requirement already satisfied: networkx>=2.4 in c:\users\surya\appdata\local\program s\python\python312\lib\site-packages (from visions<0.7.7,>=0.7.5->visions[type\_image

```
_path]<0.7.7,>=0.7.5->ydata_profiling) (3.2.1)
Requirement already satisfied: six in c:\users\surya\appdata\local\programs\python\p
ython312\lib\site-packages (from patsy>=0.5.4->statsmodels<1,>=0.13.2->ydata_profili
Using cached ydata_profiling-4.7.0-py2.py3-none-any.whl (357 kB)
Using cached dacite-1.8.1-py3-none-any.whl (14 kB)
Using cached numba-0.59.1-cp312-cp312-win_amd64.whl (2.7 MB)
Using cached scipy-1.11.4-cp312-cp312-win_amd64.whl (43.7 MB)
Using cached seaborn-0.12.2-py3-none-any.whl (293 kB)
Using cached statsmodels-0.14.1-cp312-cp312-win_amd64.whl (9.8 MB)
Using cached typeguard-4.1.5-py3-none-any.whl (34 kB)
Using cached visions-0.7.6-py3-none-any.whl (104 kB)
Using cached wordcloud-1.9.3-cp312-cp312-win_amd64.whl (301 kB)
Downloading llvmlite-0.42.0-cp312-cp312-win_amd64.whl (28.1 MB)
  ----- 0.0/28.1 MB ? eta -:--:-
  ----- 0.0/28.1 MB ? eta -:--:-
  ----- 0.1/28.1 MB 787.7 kB/s eta 0:00:36
  ----- 0.2/28.1 MB 1.5 MB/s eta 0:00:19
  ----- 0.3/28.1 MB 1.8 MB/s eta 0:00:16
  ----- 0.5/28.1 MB 2.2 MB/s eta 0:00:13
  ----- 0.5/28.1 MB 2.1 MB/s eta 0:00:14
  ----- 0.6/28.1 MB 2.0 MB/s eta 0:00:14
  - ----- 0.7/28.1 MB 2.1 MB/s eta 0:00:14
  - ----- 0.8/28.1 MB 2.0 MB/s eta 0:00:14
   ----- 0.9/28.1 MB 1.9 MB/s eta 0:00:15
  - ------ 1.0/28.1 MB 2.0 MB/s eta 0:00:14
   ------ 1.1/28.1 MB 2.0 MB/s eta 0:00:14
  - ------ 1.1/28.1 MB 1.9 MB/s eta 0:00:15
  - ----- 1.2/28.1 MB 1.8 MB/s eta 0:00:15
   ----- 1.2/28.1 MB 1.8 MB/s eta 0:00:16
   ----- 1.2/28.1 MB 1.7 MB/s eta 0:00:17
   ------ 1.3/28.1 MB 1.7 MB/s eta 0:00:17
  - ------ 1.3/28.1 MB 1.6 MB/s eta 0:00:17
  -- ----- 1.4/28.1 MB 1.6 MB/s eta 0:00:17
  -- ------ 1.6/28.1 MB 1.7 MB/s eta 0:00:16
  -- ----- 1.7/28.1 MB 1.7 MB/s eta 0:00:16
    ----- 1.7/28.1 MB 1.7 MB/s eta 0:00:16
  -- ----- 1.7/28.1 MB 1.6 MB/s eta 0:00:17
  -- ----- 1.8/28.1 MB 1.6 MB/s eta 0:00:17
  -- ----- 1.8/28.1 MB 1.6 MB/s eta 0:00:17
  -- ----- 1.8/28.1 MB 1.5 MB/s eta 0:00:18
   ------ 1.8/28.1 MB 1.5 MB/s eta 0:00:18
  -- ----- 1.9/28.1 MB 1.5 MB/s eta 0:00:18
  -- ----- 1.9/28.1 MB 1.4 MB/s eta 0:00:19
  -- ----- 2.0/28.1 MB 1.4 MB/s eta 0:00:19
  -- ----- 2.0/28.1 MB 1.4 MB/s eta 0:00:19
  -- ----- 2.0/28.1 MB 1.4 MB/s eta 0:00:19
  -- ----- 2.1/28.1 MB 1.4 MB/s eta 0:00:19
   - ----- 2.1/28.1 MB 1.4 MB/s eta 0:00:19
  --- 2.2/28.1 MB 1.4 MB/s eta 0:00:20
  --- 2.2/28.1 MB 1.4 MB/s eta 0:00:20
  --- 2.3/28.1 MB 1.3 MB/s eta 0:00:20
  --- ------ 2.3/28.1 MB 1.3 MB/s eta 0:00:20
    ----- 2.3/28.1 MB 1.3 MB/s eta 0:00:20
  --- ------ 2.4/28.1 MB 1.3 MB/s eta 0:00:21
  --- 2.4/28.1 MB 1.3 MB/s eta 0:00:21
```

 2.5/28.1 MB 1.2 MB/s eta 0:00:21
 2.5/28.1 MB 1.2 MB/s eta 0:00:22
 3.7/20.1 MD 02/.0 KD/S ELd 0.00:30

 3.8/28.1	MB	818.6	kB/s	eta	0:00:30
 3.8/28.1	MB	818.6	kB/s	eta	0:00:30
 4.5/28.1	MR	/8/.1	KB/S	eta	0:00:31
 4.5/28.1	MB	/83.8	KB/S	eta	0:00:31
 4.6/28.1	MR	/86.3	KB/S	eta	0:00:30
 5.0/28.1					
 5.7/28.1	MB	800.8	kB/s	eta	0:00:28

 5.7/28.1	MB	800.2	kB/s	eta	0:00:28
 5.7/28.1	MB	798.1	kB/s	eta	0:00:29
 5.8/28.1	MB	792.3	KB/S	eta	0:00:29
 5.8/28.1	MB	/89./	KB/S	eta	0:00:29
 5.9/28.1	MB	788.9	kB/s	eta	0:00:29
 5.9/28.1	MB	788.0	kB/s	eta	0:00:29
 6.0/28.1	MB	786.4	kB/s	eta	0:00:29
 6.0/28.1	MB	785.8	kB/s	eta	0:00:29
 6.0/28.1	MB	785.8	kB/s	eta	0:00:29
 6.1/28.1	MB	775.9	kB/s	eta	0:00:29
 6 1/28 1	MR	775 2	kB/s	eta	0.00.29
 6.2/28.1	MB	7/6.7	KB/S	eta	0:00:29
 6.3/28.1	MR	//1./	KB/S	ета	0:00:29
 6.3/28.1	MB	771.0	KB/S	eta	0:00:29
 6.3/28.1	MB	770.4	kB/s	eta	0:00:29
 6.4/28.1	MB	773.7	kB/s	eta	0:00:29
 6.5/28.1	MB	767.4	kB/s	eta	0:00:29
 6.5/28.1	MB	765.7	kB/s	eta	0:00:29
 6.6/28.1	MB	762.4	kB/s	eta	0:00:29
 6.7/28.1	MB	762.0	kB/s	eta	0:00:29
 6.7/28.1	MB	760.1	kB/s	eta	0:00:29
 6.7/28.1	MB	759.5	kB/s	eta	0:00:29
 6.7/28.1	MB	760.5	kB/s	eta	0:00:29
 -			-		
 7.5/28.1	MB	769.1	kB/s	eta	0:00:27

7.5/28.1 MB 768.4 kB/s eta 0:00:27	
7.6/28.1 MB 768.9 kB/s eta 0:00:27	
7.6/28.1 MB 769.6 kB/s eta 0:00:27	
7.6/28.1 MB 769.0 kB/s eta 0:00:27	
7.7/28.1 MB 770.6 kB/s eta 0:00:27	
7.7/28.1 MB 768.0 kB/s eta 0:00:27	
7.7/28.1 MB 768.5 kB/s eta 0:00:27	
7.8/28.1 MB 768.2 kB/s eta 0:00:27	
7.8/28.1 MB 766.2 kB/s eta 0:00:27	
7.8/28.1 MB 766.9 kB/s eta 0:00:27	
7.8/28.1 MB 765.4 kB/s eta 0:00:27	
7.9/28.1 MB 766.1 kB/s eta 0:00:27	
7.9/28.1 MB 765.5 kB/s eta 0:00:27	
8.0/28.1 MB 765.9 kB/s eta 0:00:27	
8.0/28.1 MB 765.4 kB/s eta 0:00:27	
8.0/28.1 MB 766.1 kB/s eta 0:00:27	
8.1/28.1 MB 765.4 kB/s eta 0:00:27	
8.1/28.1 MB 765.8 kB/s eta 0:00:27	
8.2/28.1 MB 764.7 kB/s eta 0:00:27	
8.2/28.1 MB 768.3 kB/s eta 0:00:26	
8.3/28.1 MB 771.6 kB/s eta 0:00:26	
8.4/28.1 MB 772.8 kB/s eta 0:00:26	
8.4/28.1 MB 773.4 kB/s eta 0:00:26	
8.4/28.1 MB 773.8 kB/s eta 0:00:26	
8.5/28.1 MB 773.3 kB/s eta 0:00:26	
8.5/28.1 MB 774.6 kB/s eta 0:00:26	
8.6/28.1 MB 778.6 kB/s eta 0:00:26	
8.7/28.1 MB 784.8 kB/s eta 0:00:25	
8.8/28.1 MB 787.9 kB/s eta 0:00:25	
8.8/28.1 MB 787.9 KB/S eta 0:00:25	
8.9/28.1 MB 791.9 kB/s eta 0:00:25	
8.9/28.1 MB 795.0 kB/s eta 0:00:25	
9.0/28.1 MB 800.1 kB/s eta 0:00:24	
9.1/28.1 MB 802.9 kB/s eta 0:00:24	
9.2/28.1 MB 807.7 kB/s eta 0:00:24	
9.3/28.1 MB 813.3 kB/s eta 0:00:24	
9.4/28.1 MB 818.1 kB/s eta 0:00:23	
9.5/28.1 MB 821.9 kB/s eta 0:00:23	
9.5/28.1 MB 823.0 kB/s eta 0:00:23	
9.6/28.1 MB 826.7 kB/s eta 0:00:23	
9.6/28.1 MB 827.8 kB/s eta 0:00:23	
9.7/28.1 MB 829.9 kB/s eta 0:00:23	
9.7/28.1 MB 829.2 kB/s eta 0:00:23	
9.7/28.1 MB 828.8 kB/s eta 0:00:23	
9.8/28.1 MB 828.5 kB/s eta 0:00:23	
9.8/28.1 MB 827.8 kB/s eta 0:00:23	
9.9/28.1 MB 828.4 kB/s eta 0:00:23	
9.9/28.1 MB 826.7 kB/s eta 0:00:23	
9.9/28.1 MB 826.0 kB/s eta 0:00:23	
9.9/28.1 MB 825.5 kB/s eta 0:00:23	
	,
10.0/28.1 MB 822.8 kB/s eta 0:00:23	
10.0/28.1 MB 822.1 kB/s eta 0:00:23	
10.0/28.1 MB 821.5 kB/s eta 0:00:23	
10.1/28.1 MB 822.7 kB/s eta 0:00:22	
10.1/28.1 MB 822.1 kB/s eta 0:00:22	
10.1/28.1 MB 821.5 kB/s eta 0:00:22	<u>,                                     </u>

 10.2/28.1	MB	817.5	kB/s	eta	0:00:22
10.2/28.1					
 -					
10.3/28.1					
10.3/28.1	MB	809.3	kB/s	eta	0:00:23
 10.3/28.1	MB	808.3	kB/s	eta	0:00:23
 10.3/28.1	MB	808.3	kB/s	eta	0:00:23
 10.3/28.1	MB	801.3	kB/s	eta	0:00:23
 10.4/28.1					
 10.4/28.1					
10.5/28.1					
 10.5/28.1			-		
 10.5/28.1					
 10.5/28.1	MB	780.4	kB/s	eta	0:00:23
 10.5/28.1	MB	777.6	kB/s	eta	0:00:23
 10.5/28.1	MB	774.8	kB/s	eta	0:00:23
 10.6/28.1	MB	773.9	kB/s	eta	0:00:23
 10.6/28.1	MB	771.2	kB/s	eta	0:00:23
 10.6/28.1					
10.6/28.1					
 ,					
 10.9/28.1	MB	752.5	kB/s	eta	0:00:23
 10.9/28.1	MB	750.8	kB/s	eta	0:00:23
 10.9/28.1	MB	747.4	kB/s	eta	0:00:24
 10.9/28.1	MB	745.7	kB/s	eta	0:00:24
 11.4/28.1	MB	733.2	kB/s	eta	0:00:23
 11.4/28.1	MB	733.2	kB/s	eta	0:00:23
 11.4/28.1	MB	732.3	kB/s	eta	0:00:23
 11.0/20.1	I'ID	/20.0			
		725 2	L.D.	- 1	0.00 00

 11.8/28.1	MB	719.4	kB/s	eta	0:00:23
 11.8/28.1	MB	715.5	kB/s	eta	0:00:23
 11.9/28.1	MB	713.9	kB/s	eta	0:00:23
 11.9/28.1	MB	714.0	kB/s	eta	0:00:23
 12.0/28.1	MB	714.0	kB/s	eta	0:00:23
 12.1/28.1	MB	716.3	kB/s	eta	0:00:23
 12.1/28.1	MB	717.9	kB/s	eta	0:00:23
 12.1/28.1	MB	716.3	kB/s	eta	0:00:23
 12.5/28.1	MB	717.1	kB/s	eta	0:00:22
 12.9/20.1	MR	721 5	VB/c	0+3	0.00.21
 13.1/20.1	MD	720 0	kD/s	ota	0.00.21
 13.1/20.1	MD	735.6	kD/s	ota	0.00.21
 -					
 -					
	סויו	//3./	KD/ S		
	MD	775 7	LD/c		0.00.13
 13.6/28.1					
 13.6/28.1 13.6/28.1	MB	775.7	kB/s	eta	0:00:19
 13.6/28.1 13.6/28.1 13.7/28.1	MB MB	775.7 774.8	kB/s kB/s	eta eta	0:00:19 0:00:19
 13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1	MB MB MB	775.7 774.8 773.9	kB/s kB/s kB/s	eta eta eta	0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1	MB MB MB MB	775.7 774.8 773.9 775.7	kB/s kB/s kB/s kB/s	eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1	MB MB MB MB	775.7 774.8 773.9 775.7 776.6	kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1	MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6	kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.8/28.1	MB MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6	kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.8/28.1	MB MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6 777.6 778.5	kB/s kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.8/28.1 13.8/28.1 13.8/28.1	MB MB MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6 777.6 778.5 779.4	kB/s kB/s kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.8/28.1 13.8/28.1 13.8/28.1 13.8/28.1	MB MB MB MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6 777.6 778.5 779.4 778.5	kB/s kB/s kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta eta eta eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19
13.6/28.1 13.6/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.7/28.1 13.8/28.1 13.8/28.1 13.8/28.1 13.9/28.1	MB MB MB MB MB MB MB MB MB	775.7 774.8 773.9 775.7 776.6 777.6 777.6 778.5 779.4 778.5 779.4	kB/s kB/s kB/s kB/s kB/s kB/s kB/s kB/s	eta	0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19 0:00:19

 13.9/28.1	MB	779.4	kB/s	eta	0:00:19
 14.0/28.1	MB	779.4	kB/s	eta	0:00:19
 14.0/28.1	MB	//2.1	KB/S	ета	0:00:19
 14.0/28.1	MB	772.1	kB/s	eta	0:00:19
 14.1/28.1	MB	772.9	kB/s	eta	0:00:19
 14.1/28.1	MB	772.1	kB/s	eta	0:00:19
 14.2/28.1	MB	771.1	kB/s	eta	0:00:19
 -			-		
 14.6/28.1	MB	773.9	kB/s	eta	0:00:18
 14.7/28.1	MB	775.7	kB/s	eta	0:00:18
 14.7/28.1	MB	775.7	kB/s	eta	0:00:18
 14.7/28.1	MB	777.5	kB/s	eta	0:00:18
 14.8/28.1	MB	778.5	kB/s	eta	0:00:18
 15.9/28.1	MB	785.9	kB/s	eta	0:00:16
 15.9/28.1	MB	784.1	kB/s	eta	0:00:16
 16.0/28.1	MB	784.1	kB/s	eta	0:00:16

 16.0/28.1	MB	785.9	kB/s	eta	0:00:16
 16.0/28.1	MB	786.9	kB/s	eta	0:00:16
 •					
 16.1/28.1					
 -			-		
			-		
 •			-		
 -					
 - · - •					
 16.6/28.1					
 -					
 17 0/20 1	MAD	700 -			

18	.8/28.1 .9/28.1 .9/28.1 .9/28.1 .9/28.1 .0/28.1 .0/28.1	MB MB MB MB MB MB	788.8 786.9 786.0 785.0 784.1 784.1	kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta	0:00:14 0:00:14 0:00:14
17	.9/28.1 .9/28.1 .9/28.1 .9/28.1 .0/28.1 .0/28.1	MB MB MB MB MB	786.9 786.0 785.0 784.1 784.1	kB/s kB/s kB/s kB/s	eta eta eta eta	0:00:14 0:00:14 0:00:13
17	.9/28.1 .9/28.1 .9/28.1 .9/28.1 .0/28.1 .0/28.1	MB MB MB MB MB	786.9 786.0 785.0 784.1 784.1	kB/s kB/s kB/s kB/s	eta eta eta eta	0:00:14 0:00:14 0:00:13
17	.9/28.1 .9/28.1 .9/28.1 .0/28.1 .0/28.1	MB MB MB MB	786.0 785.0 784.1 784.1	kB/s kB/s kB/s	eta eta eta	0:00:14 0:00:13
17	.9/28.1 .9/28.1 .0/28.1 .0/28.1	MB MB MB	785.0 784.1 784.1	kB/s kB/s	eta eta	0:00:13
17	.9/28.1 .0/28.1 .0/28.1 .0/28.1	MB MB MB	784.1 784.1	kB/s	eta	
18	.0/28.1 .0/28.1 .0/28.1	MB MB	784.1			0:00:13
18	.0/28.1 .0/28.1	MB		kB/s	$\circ$ + $\circ$	
	.0/28.1					
18. 18.						
18.		MB	780.3	kB/s	eta	0:00:13
	.0/28.1	MB	779.4	kB/s	eta	0:00:13
						0:00:13
18.	-			-		
18.						
18.						
18.						
18.						
18.	-			-		
18.						
18.	.0/28.1	MB	775.7	kB/s	eta	0:00:14
18.	.0/28.1	MB	775.7	kB/s	eta	0:00:14
18.	.0/28.1	MB	775.7	kB/s	eta	0:00:14
18.	.0/28.1	MB	775.7	kB/s	eta	0:00:14
18.						
18.						
18.						
						0:00:14
						0:00:14
						0:00:14
						0:00:14
18.						
18.						
18.	.0/28.1	MB	775.7	kB/s	eta	0:00:14
18.	.2/28.1	MB	736.5	kB/s	eta	0:00:14
18.	.2/28.1	MB	734.8	kB/s	eta	0:00:14
18.						
18.						
18.						
18.						
18.						
18.						
18.						
18.						
18.	.5/28.1	MB	726.6	kB/s	eta	0:00:14
18.	.5/28.1	MB	723.4	kB/s	eta	0:00:14
18.	.5/28.1	MB	721.8	kB/s	eta	0:00:14
18.	.5/28.1	MB	719.5	kB/s	eta	0:00:14
18.						
18.		MR		, -		
18. 18.	.5/28.1			kB/s	eta	0:00:14
	.5/28.1 .6/28.1	MB	714.8			
18	.5/28.1 .6/28.1 .6/28.1	MB MB	714.8 714.0	kB/s	eta	0:00:14
18	.5/28.1 .6/28.1 .6/28.1 .6/28.1	MB MB MB	714.8 714.0 713.2	kB/s kB/s	eta eta	0:00:14 0:00:14
18	.5/28.1 .6/28.1 .6/28.1 .6/28.1	MB MB MB	714.8 714.0 713.2 710.8	kB/s kB/s kB/s	eta eta eta	0:00:14 0:00:14 0:00:14
18	.5/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1	MB MB MB MB	714.8 714.0 713.2 710.8 710.8	kB/s kB/s kB/s kB/s	eta eta eta eta	0:00:14 0:00:14 0:00:14 0:00:14
	.5/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1	MB MB MB MB MB	714.8 714.0 713.2 710.8 710.8 710.8	kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta	0:00:14 0:00:14 0:00:14 0:00:14 0:00:14
18	.5/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1 .6/28.1	MB MB MB MB MB MB	714.8 714.0 713.2 710.8 710.8 710.8 710.8	kB/s kB/s kB/s kB/s kB/s kB/s	eta eta eta eta eta eta	0:00:14 0:00:14 0:00:14 0:00:14 0:00:14 0:00:14

 18.7/28.1	MB	689.9	kB/s	eta	0:00:14
 18.7/28.1	MB	689.2	kB/s	eta	0:00:14
 18.9/28.1	MB	681.3	kB/s	eta	0:00:14
 18.9/28.1	MB	679.9	kB/s	eta	0:00:14
 18.9/28.1	MB	679.1	kB/s	eta	0:00:14
 19.0/28.1	MB	677.7	kB/s	eta	0:00:14
 19.0/28.1	MB	676.3	kB/s	eta	0:00:14
 19.1/28.1	MB	666.7	kB/s	eta	0:00:14
 19.2/28.1	MB	665.4	kB/s	eta	0:00:14
 19.2/28.1	MB	662.7	kB/s	eta	0:00:14
 19.3/28.1	MB	650.8	kB/s	eta	0:00:14
 19.6/28.1	ΜВ	616.5	KB/S	eta	v:00:14

 19.6/28.1	MB	614.7	kB/s	eta	0:00:14
	MB	605.1	kB/s	eta	0:00:14
 19.7/28.1	MB	605.1	kB/s	eta	0:00:14
 19.7/28.1	MB	601.8	kB/s	eta	0:00:14
 19.7/28.1	MB	599.6	kB/s	eta	0:00:14
 19.8/28.1	MB	599.5	kB/s	eta	0:00:14
 19.8/28.1					
 -					
19.9/28.1					
 19.9/28.1					
 19.9/28.1			-		
 19.9/28.1					
 19.9/28.1	MB	589.3	kB/s	eta	0:00:14
 20.0/28.1	MB	588.8	kB/s	eta	0:00:14
 20.0/28.1	MB	588.8	kB/s	eta	0:00:14
 20.0/28.1	MB	588.2	kB/s	eta	0:00:14
 20.0/28.1	MB	587.7	kB/s	eta	0:00:14
 20.1/28.1	MB	586.1	kB/s	eta	0:00:14
 20.1/28.1	MB	585.6	kB/s	eta	0:00:14
 20.1/28.1	MB	585.6	kB/s	eta	0:00:14
 20.4/28.1					
 20.6/28.1	MB	572.8	kB/s	eta	0:00:14

 	20.7/28.1	MB	576.8	kB/s	eta	0:00:13
	20.7/28.1					
	20.7/28.1			-		
	20.9/28.1					
	20.9/28.1					
 	21.1/28.1	MB	563.9	kB/s	eta	0:00:13
 	21.1/28.1	MB	561.0	kB/s	eta	0:00:13
 	21.2/28.1	MB	560.1	kB/s	eta	0:00:13
 	21.2/28.1	MB	560.1	kB/s	eta	0:00:13
 	21.2/28.1	MB	558.1	kB/s	eta	0:00:13
	/20.1	יו י	J-72.7	110/3	CCa	0.00.10

 21.5/28.1	MB	540.2	kB/s	eta	0:00:13
 21.5/28.1	MB	540.6	kB/s	eta	0:00:13
 21.6/28.1	MB	539.3	kB/s	eta	0:00:13
 -			-		
 22.7/28.1	MB	514.3	kB/s	eta	0:00:11

 22.7/28.1	MB	512.7	kB/s	eta	0:00:11
 22.9/28.1	MB	507.1	kB/s	eta	0:00:11
 23.0/28.1	MB	506.7	kB/s	eta	0:00:11
 23.1/28.1	MB	507.1	kB/s	eta	0:00:10
 -			-		
 23.7/28.1	MB	509.1	kB/s	eta	0:00:09
 23.7/28.1	MB	508.3	kB/s	eta	0:00:09
 23.8/28.1	MB	507.9	kB/s	eta	0:00:09
 24.0/28.1	MB	511.9	KB/S	eta	0:00:08
 24.1/28.1	MB	513.1	kB/s	eta	0:00:08
 24.2/28.1	MB	513.1	kB/s	eta	0:00:08
 24.2/28.1	MB	513.1	kB/s	eta	0:00:08
 24.2/28.1	MB	513.1	kB/s	eta	0:00:08
 24.2/28.1	MB	515.5	kB/s	eta	0:00:08
 24.3/28.1	MB	517.1	kB/s	eta	0:00:08
 24.3/28.1	MB	516.7	kB/s	eta	0:00:08
 24.5/28.1	MB	513.5	kB/s	eta	0:00:08
 24.5/28.1	MB	512.7	kB/s	eta	0:00:08
 24.5/28.1	MB	512.3	kB/s	eta	0:00:08
24.5/28.1	MR	211.9	KD/3	CCU	0.00.00

 24.7/28.1	MB	506.3	kB/s	eta	0:00:07
 24.7/28.1	MB	506.3	kB/s	eta	0:00:07
 24.7/28.1	MB	504.4	kB/s	eta	0:00:07
 24.7/28.1	MB	503.6	kB/s	eta	0:00:07
 24.8/28.1	MB	500.2	kB/s	eta	0:00:07
 24.8/28.1	MB	498.6	kB/s	eta	0:00:07
 24.8/28.1	MB	497.9	kB/s	eta	0:00:07
 24.9/28.1	MB	497.1	kB/s	eta	0:00:07
 24.9/28.1	MB	496.0	kB/s	eta	0:00:07
 24.9/28.1	MB	496.3	kB/s	eta	0:00:07
 24.9/28.1	MB	496.0	kB/s	eta	0:00:07
 25.0/28.1	MB	495.6	kB/s	eta	0:00:07
 25.0/28.1	MB	495.6	kB/s	eta	0:00:07
 25.1/28.1	MB	495.2	kB/s	eta	0:00:07
 25.1/28.1	MB	494.9	kB/s	eta	0:00:07
 25.2/28.1	MB	494.5	kB/s	eta	0:00:06
 25.2/28.1	MB	492.3	kB/s	eta	0:00:06
 25.3/28.1	MB	490.4	kB/s	eta	0:00:06
 25.3/28.1	MB	489.3	kB/s	eta	0:00:06
 25.3/28.1	MB	489.3	kB/s	eta	0:00:06
 25.3/28.1	MB	487.9	kB/s	eta	0:00:06
 25.3/28.1	MB	487.1	kB/s	eta	0:00:06
 25.3/28.1	MB	486.4	kB/s	eta	0:00:06
 25.4/28.1	MB	484.2	kB/s	eta	0:00:06
 25.4/28.1	MB	483.9	kB/s	eta	0:00:06
 25.5/28.1	MB	481.7	kB/s	eta	0:00:06
 25.5/28.1	MB	481.7	kB/s	eta	0:00:06
 25.5/28.1	MB	480.0	kB/s	eta	0:00:06
 25.5/28.1	MB	478.9	kB/s	eta	0:00:06
 25.5/28.1	MB	478.9	kB/s	eta	0:00:06
 25.5/28.1	MB	477.2	kB/s	eta	0:00:06
 25.5/28.1	MB	477.2	kB/s	eta	0:00:06
 25.6/28.1	MB	472.4	kB/s	eta	0:00:06

 	25.6/28.1	MB	472.4	kB/s	eta	0:00:06
 	25.6/28.1	MB	472.4	kB/s	eta	0:00:06
	-			-		
 	26.2/28.1	MB	452.1	kB/s	eta	0:00:05
 	26.3/28.1	MB	452.5	kB/s	eta	0:00:05
 	26.3/28.1	MB	454.7	kB/s	eta	0:00:04
 	26.3/28.1	MB	455.6	kB/s	eta	0:00:04
 	26.//28.1	МΒ	452.8	KB/S	eta	v:00:04

 26.8/28.1	MB	452.1	kB/s	eta	0:00:04
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	451.5	kB/s	eta	0:00:03
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	439.1	kB/s	eta	0:00:04
 26.8/28.1	MB	432.7	kB/s	eta	0:00:04
 26.8/28.1	MB	432.4	kB/s	eta	0:00:03
 26.8/28.1	MB	432.4	kB/s	eta	0:00:03
 26.8/28.1					
 26.8/28.1			-		
 26.8/28.1					
 26.8/28.1					
 26.9/28.1					
 26.9/28.1					
 •			-		
 			•		
 26.9/28.1					
 27.0/28.1			-		
 27.1/28.1					
 27.1/28.1					
 27.2/28.1 27.2/28.1			-		
 27.2/28.1			•		
 27.2/28.1					
 27.3/28.1					
 27.3/28.1					
 27.3/28.1					
 27.4/28.1					
 27.4/28.1					
 27.4/28.1					
 27.5/28.1					
-, <del>-</del>	_		, -		

```
27.6/28.1 MB 419.1 kB/s eta 0:00:02
                                     27.6/28.1 MB 418.9 kB/s eta 0:00:02
                                     27.6/28.1 MB 419.4 kB/s eta 0:00:02
  _____
  -----
                                     27.7/28.1 MB 419.7 kB/s eta 0:00:02
  _____
                                     27.7/28.1 MB 419.9 kB/s eta 0:00:02
                                     27.7/28.1 MB 420.0 kB/s eta 0:00:01
                                    27.8/28.1 MB 420.2 kB/s eta 0:00:01
                                     27.8/28.1 MB 420.8 kB/s eta 0:00:01
  _____
                                    27.9/28.1 MB 420.2 kB/s eta 0:00:01
  -----
                                    27.9/28.1 MB 420.8 kB/s eta 0:00:01
  _____
                                    27.9/28.1 MB 420.8 kB/s eta 0:00:01
     ------ 27.9/28.1 MB 420.5 kB/s eta 0:00:01
                                    28.0/28.1 MB 420.8 kB/s eta 0:00:01
  ----- 28.0/28.1 MB 422.4 kB/s eta 0:00:01
  ----- 28.1/28.1 MB 423.8 kB/s eta 0:00:01
  ----- 28.1/28.1 MB 423.8 kB/s eta 0:00:01
  ----- 28.1/28.1 MB 424.0 kB/s eta 0:00:00
Using cached patsy-0.5.6-py2.py3-none-any.whl (233 kB)
Installing collected packages: typeguard, scipy, patsy, llvmlite, dacite, numba, wor
dcloud, visions, statsmodels, seaborn, ydata_profiling
 Attempting uninstall: scipy
   Found existing installation: scipy 1.11.4
   Uninstalling scipy-1.11.4:
     Successfully uninstalled scipy-1.11.4
 Attempting uninstall: visions
   Found existing installation: visions 0.7.6
   Uninstalling visions-0.7.6:
     Successfully uninstalled visions-0.7.6
 Attempting uninstall: seaborn
   Found existing installation: seaborn 0.12.2
   Uninstalling seaborn-0.12.2:
     Successfully uninstalled seaborn-0.12.2
Successfully installed dacite-1.8.1 llvmlite-0.42.0 numba-0.59.1 patsy-0.5.6 scipy-
1.11.4 seaborn-0.12.2 statsmodels-0.14.1 typeguard-4.1.5 visions-0.7.6 wordcloud-1.
9.3 ydata_profiling-4.7.0
Note: you may need to restart the kernel to use updated packages.
ERROR: pip's dependency resolver does not currently take into account all the packag
es that are installed. This behaviour is the source of the following dependency conf
licts.
pandas-profiling 3.2.0 requires visions[type_image_path]==0.7.4, but you have vision
s 0.7.6 which is incompatible.
scikit-learn 1.4.1.post1 requires joblib>=1.2.0, but you have joblib 1.1.1 which is
```

### **IMPORTING DATASETS**

```
In []: # Reading The Data By Pandas
    data=pd.read_csv("D:\MAIN DRIVE\VSCODE\Regression-Project\data.csv")
    # Creating 'X' Matrix For Independent Features In The Dataset
    X=data.iloc[:,:-1].values
    # Creating 'Y' Matrix For Independent Features In The Dataset
    Y=data.iloc[:,-1].values
```

incompatible.

```
<>:2: SyntaxWarning: invalid escape sequence '\M'
<>:2: SyntaxWarning: invalid escape sequence '\M'
C:\Users\surya\AppData\Local\Temp\ipykernel_22704\2686243820.py:2: SyntaxWarning: in
valid escape sequence '\M'
   data=pd.read_csv("D:\MAIN DRIVE\VSCODE\Regression-Project\data.csv")
```

### GENRATING PROFILE USING PANDAS PROFILING

```
In [ ]: profile=ProfileReport(data)
profile.to_notebook_iframe()
```

Summarize dataset: 0% | 0/5 [00:00<?, ?it/s]

Generate report structure: 0% | 0/1 [00:00<?, ?it/s]

Render HTML: 0% | 0/1 [00:00<?, ?it/s]

# Overview

#### **Dataset statistics**

Number of variables	14
Number of observations	506
Missing cells	5
Missing cells (%)	0.1%
Duplicate rows	0
Duplicate rows (%)	0.0%
Duplicate 10W3 (70)	0.070
Total size in memory	55.5 KiB

### Variable types

Numeric	13
Categorical	1

#### **Alerts**

AGE is highly overall correlated with CRIM and 7 other fields (CRIM, DIS, INDUS, LSTAT, MEDV, NOX, TAX, ZN)

CRIM is highly overall correlated with AGE and 8 other fields (AGE, DIS, INDUS, LSTAT, MEDV, NOX, RAD, TAX,

<!-- Attribute Information:

- 1. CRIM per capita crime rate by town
- 2. ZN proportion of residential land zoned for lots over 25,000 sq.ft.
- 3. INDUS proportion of non-retail business acres per town

```
Charles River dummy variable (= 1 if tract bounds
4. CHAS
river; 0 otherwise)
5. NOX
             nitric oxides concentration (parts per 10 million)
6. RM
             average number of rooms per dwelling
             proportion of owner-occupied units built prior to 1940
7. AGE
8. DIS
            weighted distances to five Boston employment centres
9. RAD
             index of accessibility to radial highways
10. TAX
             full-value property-tax rate per $10,000
11. PTRATIO pupil-teacher ratio by town
12. B
             1000(Bk - 0.63)^2 where Bk is the proportion of blacks
by town
            % lower status of the population
13. LSTAT
14. MEDV
            Median value of owner-occupied homes in $1000's -->
```

| 2       0.02729       0.0       7.07       0       0.469       7.185       61.1       4.9671       2       242       17.8       392.83         3       0.03237       0.0       2.18       0       0.458       6.998       45.8       6.0622       3       222       18.7       394.63         4       0.06905       0.0       2.18       0       0.458       7.147       54.2       6.0622       3       222       18.7       396.90 <th>•</th> <th>CRIM</th> <th>ZN</th> <th>INDUS</th> <th>CHAS</th> <th>NOX</th> <th>RM</th> <th>AGE</th> <th>DIS</th> <th>RAD</th> <th>TAX</th> <th>PTRATIO</th> <th>В</th>  | •     | CRIM     | ZN    | INDUS | CHAS | NOX   | RM    | AGE  | DIS    | RAD | TAX | PTRATIO | В      |
|--|-------|----------|-------|-------|------|-------|-------|------|--------|-----|-----|---------|--------|
| 2       0.02729       0.0       7.07       0       0.469       7.185       61.1       4.9671       2       242       17.8       392.83         3       0.03237       0.0       2.18       0       0.458       6.998       45.8       6.0622       3       222       18.7       394.63         4       0.06905       0.0       2.18       0       0.458       7.147       54.2       6.0622       3       222       18.7       396.90 <th>0</th> <th>0.00632</th> <th>18.0</th> <th>2.31</th> <th>0</th> <th>0.538</th> <th>6.575</th> <th>65.2</th> <th>4.0900</th> <th>1</th> <th>296</th> <th>15.3</th> <th>396.90</th>  | 0     | 0.00632  | 18.0  | 2.31  | 0    | 0.538 | 6.575 | 65.2 | 4.0900 | 1   | 296 | 15.3    | 396.90 |
| 3       0.03237       0.0       2.18       0       0.458       6.998       45.8       6.0622       3       222       18.7       394.63         4       0.06905       0.0       2.18       0       0.458       7.147       54.2       6.0622       3       222       18.7       396.90 <t< th=""><th>1</th><th>0.02731</th><th>0.0</th><th>7.07</th><th>0</th><th>0.469</th><th>6.421</th><th>78.9</th><th>4.9671</th><th>2</th><th>242</th><th>17.8</th><th>396.90</th></t<>   | 1     | 0.02731  | 0.0   | 7.07  | 0    | 0.469 | 6.421 | 78.9 | 4.9671 | 2   | 242 | 17.8    | 396.90 |
| 4       0.06905       0.0       2.18       0       0.458       7.147       54.2       6.0622       3       222       18.7       396.90   | 2     | 0.02729  | 0.0   | 7.07  | 0    | 0.469 | 7.185 | 61.1 | 4.9671 | 2   | 242 | 17.8    | 392.83 |
| .  | 3     | 0.03237  | 0.0   | 2.18  | 0    | 0.458 | 6.998 | 45.8 | 6.0622 | 3   | 222 | 18.7    | 394.63 |
| <b>501</b> 0.06263       0.0       11.93       0       0.573       6.593       69.1       2.4786       1       273       21.0       391.99 <b>502</b> 0.04527       0.0       11.93       0       0.573       6.120       76.7       2.2875       1       273       21.0       396.90 <b>503</b> 0.06076       0.0       11.93       0       0.573       6.976       91.0       2.1675       1       273       21.0       396.90 <b>504</b> 0.10959       0.0       11.93       0       0.573       6.794       89.3       2.3889       1       273       21.0       393.45 <b>505</b> 0.04741       0.0       11.93       0       0.573       6.030       80.8       2.5050       1       273       21.0       396.90 | 4     | 0.06905  | 0.0   | 2.18  | 0    | 0.458 | 7.147 | 54.2 | 6.0622 | 3   | 222 | 18.7    | 396.90 |
| <b>502</b> 0.04527       0.0       11.93       0       0.573       6.120       76.7       2.2875       1       273       21.0       396.90 <b>503</b> 0.06076       0.0       11.93       0       0.573       6.976       91.0       2.1675       1       273       21.0       396.90 <b>504</b> 0.10959       0.0       11.93       0       0.573       6.794       89.3       2.3889       1       273       21.0       393.45 <b>505</b> 0.04741       0.0       11.93       0       0.573       6.030       80.8       2.5050       1       273       21.0       396.90  | •••   |          | •••   |       |      |       |       |      |        |     |     |         |        |
| <b>503</b> 0.06076       0.0       11.93       0       0.573       6.976       91.0       2.1675       1       273       21.0       396.90 <b>504</b> 0.10959       0.0       11.93       0       0.573       6.794       89.3       2.3889       1       273       21.0       393.45 <b>505</b> 0.04741       0.0       11.93       0       0.573       6.030       80.8       2.5050       1       273       21.0       396.90   | 501   | 0.06263  | 0.0   | 11.93 | 0    | 0.573 | 6.593 | 69.1 | 2.4786 | 1   | 273 | 21.0    | 391.99 |
| <b>504</b> 0.10959       0.0       11.93       0       0.573       6.794       89.3       2.3889       1       273       21.0       393.45 <b>505</b> 0.04741       0.0       11.93       0       0.573       6.030       80.8       2.5050       1       273       21.0       396.90  | 502   | 0.04527  | 0.0   | 11.93 | 0    | 0.573 | 6.120 | 76.7 | 2.2875 | 1   | 273 | 21.0    | 396.90 |
| <b>505</b> 0.04741 0.0 11.93 0 0.573 6.030 80.8 2.5050 1 273 21.0 396.90   | 503   | 0.06076  | 0.0   | 11.93 | 0    | 0.573 | 6.976 | 91.0 | 2.1675 | 1   | 273 | 21.0    | 396.90 |
|  | 504   | 0.10959  | 0.0   | 11.93 | 0    | 0.573 | 6.794 | 89.3 | 2.3889 | 1   | 273 | 21.0    | 393.45 |
| 506 rows × 14 columns  | 505   | 0.04741  | 0.0   | 11.93 | 0    | 0.573 | 6.030 | 8.08 | 2.5050 | 1   | 273 | 21.0    | 396.90 |
|  | 506 r | ows × 14 | colum | ns    |      |       |       |      |        |     |     |         |        |

file:///C:/Users/surya/OneDrive/Documents/RealEstateh.html

In [ ]: data.info()

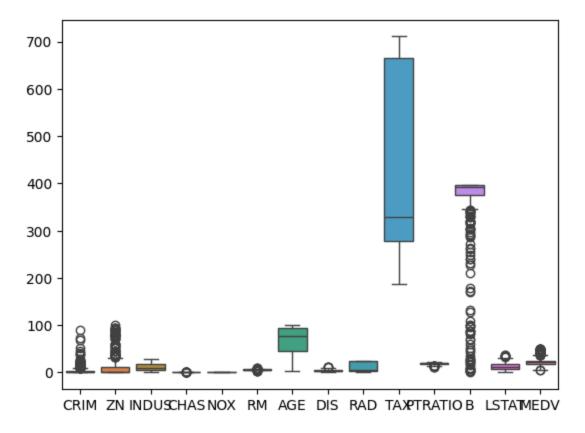
```
<class 'pandas.core.frame.DataFrame'>
       RangeIndex: 506 entries, 0 to 505
       Data columns (total 14 columns):
             Column
                      Non-Null Count Dtype
             -----
        0
             CRIM
                      506 non-null
                                        float64
        1
             ΖN
                       506 non-null
                                        float64
        2
             INDUS
                      506 non-null
                                        float64
        3
             CHAS
                       506 non-null
                                        int64
        4
             NOX
                       506 non-null
                                        float64
        5
             RM
                      501 non-null
                                        float64
        6
             AGE
                      506 non-null
                                        float64
        7
             DIS
                      506 non-null
                                        float64
        8
             RAD
                       506 non-null
                                        int64
        9
             TAX
                       506 non-null
                                        int64
             PTRATIO
        10
                      506 non-null
                                        float64
                      506 non-null
                                        float64
        11
             В
        12
             LSTAT
                      506 non-null
                                        float64
        13 MEDV
                                        float64
                      506 non-null
       dtypes: float64(11), int64(3)
       memory usage: 55.5 KB
In [ ]: # "CHAR" IS A UNBALANCED COLUMN <BIASED>
         data['CHAS'].value_counts()
Out[]: CHAS
              471
                35
         1
         Name: count, dtype: int64
        data.describe()
In [ ]:
Out[]:
                     CRIM
                                   ZN
                                            INDUS
                                                         CHAS
                                                                      NOX
                                                                                   RM
                                                                                              AGE
         count 506.000000
                            506.000000
                                        506.000000
                                                    506.000000
                                                                506.000000
                                                                            501.000000
                                                                                        506.000000
         mean
                  3.613524
                             11.363636
                                         11.136779
                                                      0.069170
                                                                  0.554695
                                                                              6.284341
                                                                                         68.574901
                  8.601545
           std
                             23.322453
                                          6.860353
                                                      0.253994
                                                                  0.115878
                                                                              0.705587
                                                                                         28.148861
           min
                  0.006320
                              0.000000
                                          0.460000
                                                      0.000000
                                                                  0.385000
                                                                              3.561000
                                                                                          2.900000
          25%
                  0.082045
                                          5.190000
                              0.000000
                                                      0.000000
                                                                  0.449000
                                                                              5.884000
                                                                                         45.025000
          50%
                  0.256510
                              0.000000
                                          9.690000
                                                      0.000000
                                                                  0.538000
                                                                              6.208000
                                                                                         77.500000
          75%
                  3.677083
                             12.500000
                                         18.100000
                                                      0.000000
                                                                  0.624000
                                                                              6.625000
                                                                                         94.075000
          max
                 88.976200
                            100.000000
                                         27.740000
                                                      1.000000
                                                                  0.871000
                                                                              8.780000
                                                                                        100.000000
```

# VISUALIZING THE DATASET

In [ ]: #For plotting histogram
print("HISTOGRAM")

```
%matplotlib inline
         data.hist(bins=50, figsize=(20, 15))
        HISTOGRAM
Out[ ]: array([[<Axes: title={'center': 'CRIM'}>, <Axes: title={'center': 'ZN'}>,
                   <Axes: title={'center': 'INDUS'}>,
                   <Axes: title={'center': 'CHAS'}>],
                  [<Axes: title={'center': 'NOX'}>, <Axes: title={'center': 'RM'}>,
                   <Axes: title={'center': 'AGE'}>, <Axes: title={'center': 'DIS'}>],
                  [<Axes: title={'center': 'RAD'}>, <Axes: title={'center': 'TAX'}>,
                  <Axes: title={'center': 'PTRATIO'}>,
                   <Axes: title={'center': 'B'}>],
                  [<Axes: title={'center': 'LSTAT'}>,
                   <Axes: title={'center': 'MEDV'}>, <Axes: >, <Axes: >]],
                dtype=object)
                                                                                           CHAS
       300
                                                                                 400
                                300
                                                         100
       250
                                250
                                                                                 300
       200
                                200
                                                         60
       150
                                                                                 200
                                150
                                                         40
       100
                                100
                                                                                 100
                  NOX
                                           RM
                                                         30
                                                         20
                                10
                                                         120
                                                                                 250
       100
                                100
                                                         100
                                                                                 200
        80
                                80
                                                                                 150
                                60
                                                         60
                                                                                 100
                                         400
                 LSTAT
        25
        20
                                25
                                15
In [ ]: print("COLORED - BOXPLOT")
         sns.boxplot(data=data)
        COLORED - BOXPLOT
```

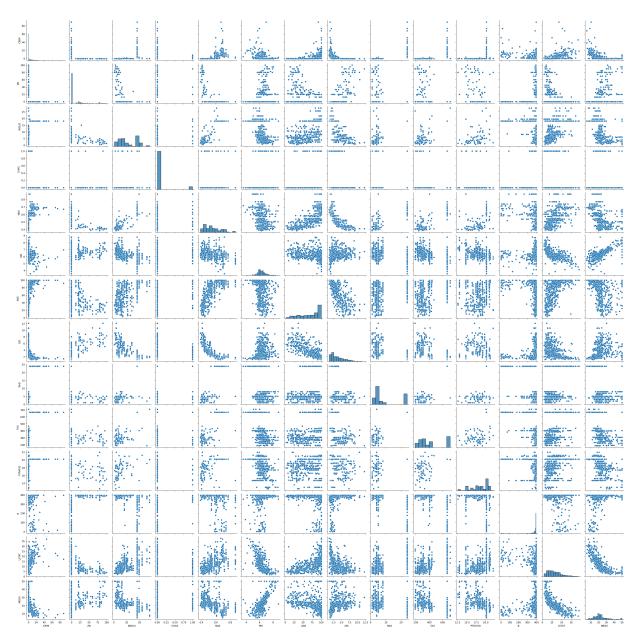
Out[]: <Axes: >



In [ ]: print("PAIR - PLOT OF ALL FEATURES WITH LABEL")
 sns.pairplot(data)

PAIR - PLOT OF ALL FEATURES WITH LABEL

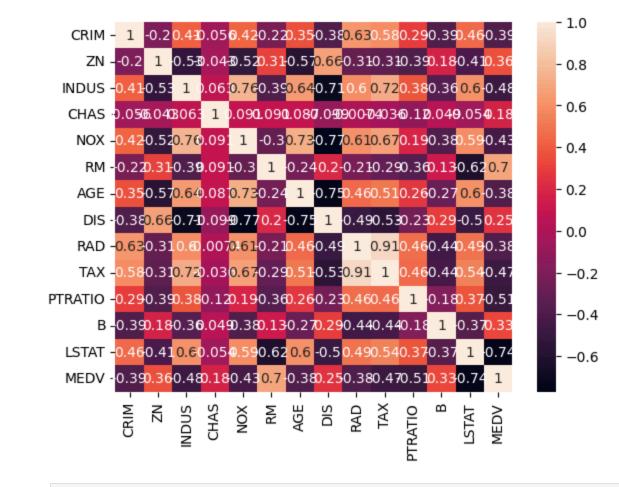
Out[]: <seaborn.axisgrid.PairGrid at 0x1c9800d6ff0>



```
In [ ]: print("HEATMAP")
sns.heatmap(data.corr(), annot=True)
```

HEATMAP

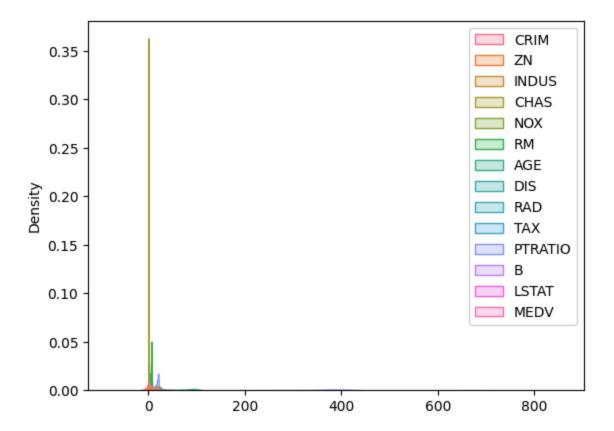
Out[]: <Axes: >



```
In [ ]: print("Density - Plot")
sns.kdeplot(data, fill=True)
```

Density - Plot

Out[]: <Axes: ylabel='Density'>



# **INDEPENDENT - FEATURES & DEPENDENT - LABEL**

```
In [ ]: print(F"Matrix For Independent Features In The Dataset -> \n{X}")

Matrix For Independent Features In The Dataset ->
[[6.3200e-03 1.8000e+01 2.3100e+00 ... 1.5300e+01 3.9690e+02 4.9800e+00]
[2.7310e-02 0.0000e+00 7.0700e+00 ... 1.7800e+01 3.9690e+02 9.1400e+00]
[2.7290e-02 0.0000e+00 7.0700e+00 ... 1.7800e+01 3.9283e+02 4.0300e+00]
...
[6.0760e-02 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9690e+02 5.6400e+00]
[1.0959e-01 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9345e+02 6.4800e+00]
[4.7410e-02 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9690e+02 7.8800e+00]]
In [ ]: print(F"Matrix For Dependent Features In The Dataset -> MEVD \n{Y.reshape(len(Y),1)}
```

Matrix For Dependent Features In The Dataset -> MEVD

[[24.]

[21.6]

[34.7]

[33.4]

[36.2]

[28.7]

[22.9]

[27.1]

[16.5]

[18.9]

[15.]

[18.9]

[21.7]

[20.4]

[18.2]

[19.9]

[23.1]

[17.5]

[20.2]

[18.2]

[13.6]

[19.6]

[15.2]

[14.5]

[15.6]

[13.9]

[16.6]

[14.8]

[18.4]

[21.]

[12.7]

[14.5]

[13.2]

[13.1]

[13.5]

[18.9]

[20.]

[21.]

[24.7]

[30.8]

[34.9]

[26.6]

[25.3]

[24.7]

[21.2]

[19.3]

[20.]

[16.6]

[14.4]

[19.4]

[19.7]

[20.5]

[25.]

[23.4] [18.9]

[35.4]

[24.7]

[31.6]

[23.3]

[19.6]

[18.7]

[16.]

[22.2]

[25.]

[33.]

[23.5]

[19.4]

[22.]

[17.4]

[20.9]

[24.2]

[21.7]

[22.8]

[23.4]

[24.1]

[21.4]

[20.]

[20.8]

[21.2]

[20.3]

[28.]

[23.9]

[24.8]

[22.9]

[23.9]

[26.6]

[22.5]

[22.2]

[23.6]

[28.7]

[22.6]

[22.]

[22.9]

[25.]

[20.6]

[28.4]

[21.4]

[38.7]

[43.8]

[33.2]

[27.5]

[26.5]

[18.6]

[19.3]

[20.1]

[19.5]

[19.5]

[20.4]

[19.8]

[19.4][21.7]

[22.8]

[18.8]

[18.7]

[18.5]

[18.3]

[21.2]

[19.2]

[20.4]

[19.3]

[22.]

[20.3]

[20.5]

[17.3]

[18.8]

[21.4]

[15.7]

[16.2]

[18.]

[14.3]

[19.2]

[19.6]

[23.]

[18.4]

[15.6]

[18.1]

[17.4]

[17.1]

[13.3]

[17.8]

[14.]

[14.4]

[13.4]

[15.6]

[11.8]

[13.8]

[15.6]

[14.6]

[17.8]

[15.4]

[21.5]

[19.6]

[15.3]

[19.4]

[17.]

[15.6]

[13.1]

[41.3] [24.3]

[23.3]

[27.]

[50.]

[50.]

[50.]

[22.7]

[25.]

[50.]

[23.8]

[23.8]

[22.3]

[17.4]

[19.1]

[23.1]

[23.6]

[22.6]

[29.4]

[23.2]

[24.6]

[29.9]

[37.2]

[39.8]

[36.2]

[37.9]

[32.5]

[26.4]

[29.6]

[50.]

[32.]

[29.8]

[34.9] [37.]

[30.5]

[36.4]

[31.1]

[29.1]

[50.]

[33.3]

[30.3]

[34.6]

[34.9]

[32.9]

[24.1]

[42.3]

[48.5]

[50.]

[22.6]

[24.4]

[22.5] [24.4]

[20.]

[21.7] [19.3]

[22.4]

[28.1]

[23.7]

[25.]

[23.3]

[28.7]

[21.5]

[23.]

[26.7]

[21.7]

[27.5]

[30.1]

[44.8]

[50.]

[37.6]

[31.6]

[46.7]

[31.5]

[24.3]

[31.7]

[41.7]

[48.3]

[29.]

[24.]

[25.1]

[31.5]

[23.7]

[23.3]

[22.]

[20.1]

[22.2]

[23.7]

[17.6]

[18.5]

[24.3]

[20.5]

[24.5]

[26.2]

[24.4]

[24.8]

[29.6]

[42.8]

[21.9]

[20.9]

[44.]

[50.] [36.]

[30.1]

[33.8]

[43.1]

[48.8]

[31.]

[36.5]

[22.8]

[30.7]

[50.]

[43.5]

[20.7]

[21.1]

[25.2]

[24.4]

[35.2]

[32.4]

[32.]

[33.2]

[33.1]

[29.1]

[35.1]

[45.4]

[35.4]

[46.]

[50.]

[32.2]

[22.]

[20.1]

[23.2]

[22.3]

[24.8]

[28.5]

[37.3]

[27.9]

[23.9]

[21.7]

[28.6]

[27.1]

[20.3]

[22.5]

[29.]

[24.8]

[22.]

[26.4]

[33.1]

[36.1]

[28.4]

[33.4]

[28.2]

[22.8]

[20.3]

[16.1]

[22.1]

[19.4]

[21.6]

[23.8]

[16.2]

[17.8]

[19.8]

[23.1]

[21.]

[23.8]

[23.1]

[20.4]

[18.5]

[25.]

[24.6]

[23.]

[22.2]

[19.3]

[22.6]

[19.8]

[17.1]

[19.4]

[22.2] [20.7]

[21.1]

[19.5]

[18.5]

[20.6]

[19.]

[18.7]

[32.7]

[16.5]

[23.9]

[31.2]

[17.5]

[17.2]

[23.1]

[24.5]

[26.6]

[22.9]

[24.1]

[18.6]

[30.1]

[18.2]

[20.6]

[17.8]

[21.7]

[22.7]

[22.6]

[25.]

[19.9]

[20.8]

[16.8]

[21.9]

[27.5]

[21.9]

[23.1]

[50.]

[50.]

[50.]

[50.]

[50.]

[13.8]

[13.8]

[15.]

[13.9]

[13.3]

[13.1]

[10.2]

[10.4] [10.9]

[11.3]

[12.3]

[8.8]

[ 7.2]

[10.5]

[ 7.4]

[10.2] [11.5]

[15.1]

[23.2]

[ 9.7]

[13.8]

[12.7]

[13.1]

[12.5]

[ 8.5]

[5.]

[ 6.3]

[ 5.6]

[ 7.2]

[12.1]

[ 8.3]

[ 8.5]

[5.]

[11.9]

[27.9]

[17.2]

[27.5]

[15.]

[17.2]

[17.9]

[16.3]

[ 7. ]

[ 7.2]

[ 7.5]

[10.4]

[8.8]

[ 8.4]

[16.7]

[14.2]

[20.8]

[13.4]

[11.7]

[ 8.3]

[10.2]

[10.9]

[11.]

[ 9.5]

[14.5]

[14.1]

[16.1]

[14.3]

[11.7]

[13.4]

[ 9.6]

[ 8.7]

[ 8.4]

[12.8]

[10.5]

[17.1]

[18.4]

[15.4]

[10.8] [11.8]

[14.9]

[12.6]

[14.1]

[13.]

[13.4]

[15.2]

[16.1]

[17.8]

[14.9]

[14.1]

[12.7]

[13.5]

[14.9]

[20.]

[16.4]

[17.7]

[19.5]

[20.2]

[21.4]

[19.9]

[19.]

[19.1]

[19.1]

[20.1] [19.9]

[19.6]

[23.2]

[29.8]

[13.8]

[13.3]

[16.7]

[12.]

[14.6]

[21.4]

[23.] [23.7]

[25.]

[21.8]

[20.6]

[21.2]

[19.1]

[20.6]

[15.2]

[ 7. ]

[ 8.1] [13.6]

[20.1]

[21.8]

[24.5]

[23.1]

[19.7]

[18.3]

[21.2]

[17.5]

[16.8]

[22.4]

[20.6]

[23.9] [22.] [11.9]]

### VALIDATING THE MISSING VALUES IN THE DATASET

```
In [ ]: from sklearn.impute import SimpleImputer
    imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
    X[:, :14] = imputer.fit_transform(X[:, :14])

In [ ]: print(F"AFTER VALIDATING THE MISSING VALUES IN 'x' MATRIX : \n{X} ")

AFTER VALIDATING THE MISSING VALUES IN 'x' MATRIX :
    [[6.3200e-03 1.8000e+01 2.3100e+00 ... 1.5300e+01 3.9690e+02 4.9800e+00]
    [2.7310e-02 0.0000e+00 7.0700e+00 ... 1.7800e+01 3.9690e+02 9.1400e+00]
    [2.7290e-02 0.0000e+00 7.0700e+00 ... 1.7800e+01 3.9283e+02 4.0300e+00]
    ...
    [6.0760e-02 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9690e+02 5.6400e+00]
    [1.0959e-01 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9345e+02 6.4800e+00]
    [4.7410e-02 0.0000e+00 1.1930e+01 ... 2.1000e+01 3.9690e+02 7.8800e+00]]
```

### < NO ENCODNG IS NEEDED >

#### TRAIN AND TEST SPLIT

```
In []: from sklearn.model_selection import train_test_split
    X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.2, random_s

In []: print(F"The Independent Training Set :{X_train.shape} \n{X_train}") # Printing The

The Independent Training Set :(404, 13)
    [[3.5809e-01 0.0000e+00 6.2000e+00 ... 1.7400e+01 3.9170e+02 9.7100e+00]
    [1.5876e-01 0.0000e+00 1.0810e+01 ... 1.9200e+01 3.7694e+02 9.8800e+00]
    [1.1329e-01 3.0000e+01 4.9300e+00 ... 1.6600e+01 3.9125e+02 1.1380e+01]
    ...
    [1.5098e-01 0.0000e+00 1.0010e+01 ... 1.7800e+01 3.9451e+02 1.0300e+01]
    [2.2927e-01 0.0000e+00 6.9100e+00 ... 1.7900e+01 3.9274e+02 1.8800e+01]
    [1.3914e-01 0.0000e+00 4.0500e+00 ... 1.6600e+01 3.9690e+02 1.4690e+01]]

In []: print(F"The Dependent Training Set : {Y_train.shape} \n{Y_train}") # Printing The D
```

```
The Dependent Training Set : (404,)
       [26.7 21.7 22. 22.9 10.4 21.9 20.6 26.4 41.3 17.2 27.1 20.4 16.5 24.4
                  9.7 50. 30.5 12.3 19.4 21.2 20.3 18.8 33.4 18.5 19.6 33.2
        13.1 7.5 13.6 17.4 8.4 35.4 24. 13.4 26.2 7.2 13.1 24.5 37.2 25.
        24.1 16.6 32.9 36.2 11. 7.2 22.8 28.7 14.4 24.4 18.1 22.5 20.5 15.2
        17.4 13.6 8.7 18.2 35.4 31.7 33. 22.2 20.4 23.9 25. 12.7 29.1 12.
        17.7 27. 20.6 10.2 17.5 19.7 29.8 20.5 14.9 10.9 19.5 22.7 19.5 24.6
        25. 24.5 50. 14.3 11.8 31. 28.7 16.2 43.5 25. 22. 19.9 22.1 46.
        22.9 20.2 43.1 34.6 13.8 24.3 21.5 24.4 21.2 23.8 26.6 25.1 9.6 19.4
        19.4 9.5 14. 26.5 13.8 34.7 16.3 21.7 17.5 15.6 20.9 21.7 12.7 18.5
        23.7 19.3 12.7 21.6 23.2 29.6 21.2 23.8 17.1 22. 36.5 18.8 21.9 23.1
        20.2 17.4 37. 24.1 36.2 15.7 32.2 13.5 17.9 13.3 11.7 41.7 18.4 13.1
        25. 21.2 16. 34.9 25.2 24.8 21.5 23.4 18.9 10.8 21. 27.5 17.5 13.5
        28.7 14.8 19.1 28.6 13.1 19. 11.3 13.3 22.4 20.1 18.2 22.9 20.6 25.
        12.8 34.9 23.7 50. 29. 30.1 22. 15.6 23.3 30.1 14.3 22.8 50. 20.8
        6.3 34.9 32.4 19.9 20.3 17.8 23.1 20.4 23.2 7. 16.8 46.7 50. 22.9
        23.9 21.4 21.7 15.4 15.3 23.1 23.9 19.4 11.9 17.8 31.5 33.8 20.8 19.8
        22.4 5. 24.5 19.4 15.1 18.2 19.3 27.1 20.7 37.6 11.7 33.4 30.1 21.4
        45.4 20.1 20.8 26.4 10.4 21.8 32. 21.7 18.4 37.9 17.8 28. 28.2 36.
        18.9 15. 22.5 30.7 20. 19.1 23.3 26.6 21.1 19.7 20. 12.1 7.2 14.2
        17.3 27.5 22.2 10.9 19.2 32. 14.5 24.7 12.6 24. 24.1 50. 16.1 43.8
        26.6 36.1 21.8 29.9 50. 44. 20.6 19.6 28.4 19.1 22.3 20.9 28.4 14.4
        32.7 13.8 8.5 22.5 35.1 31.6 17.8 15.6 20.7 39.8 17.8 19.6 14.9 22.
        48.8 25. 48.5 23.9 20.3 15.2 10.5 19. 16.4 8.8 22. 24.8 50. 19.3
        22.7 37.3 31.6 8.3 23.1 50. 13.9 16.1 25.3 19.5 10.2 19.9 35.2 13.4
        24.7 11.5 23.4 16.7 15.4 18. 28.5 18.4 32.5 50. 50. 19.6 17.6 42.3
        14.5 13.2 16.2 29.6 16.7 13. 22.3 13.4 5. 19.5 14.6 22.
        33.2 20.6 14.1 14.1 21.1 30.3 23.7 21.4 18.3 8.5 22.8 22.8 22.2 13.9
        25. 18.5 7. 22.6 20.1 30.8 31.1 23.8 12.5 23.6 23.2 24.2 22.2 27.9
        22.2 33.1 19.3 18.9 22.6 50. 24.8 18.5 36.4 19.2 16.6 23.1]
In [ ]: print(F"The Independent Testing Set :{X_test.shape} \n{X_test}") # Printing The Ind
       The Independent Testing Set :(102, 13)
       [[6.7240e-02 0.0000e+00 3.2400e+00 ... 1.6900e+01 3.7521e+02 7.3400e+00]
        [9.2323e+00 0.0000e+00 1.8100e+01 ... 2.0200e+01 3.6615e+02 9.5300e+00]
        [1.1425e-01 0.0000e+00 1.3890e+01 ... 1.6400e+01 3.9374e+02 1.0500e+01]
        [1.4932e-01\ 2.5000e+01\ 5.1300e+00\ \dots\ 1.9700e+01\ 3.9511e+02\ 1.3150e+01]
        [1.4052e-01 0.0000e+00 1.0590e+01 ... 1.8600e+01 3.8581e+02 9.3800e+00]
        [1.2802e-01 0.0000e+00 8.5600e+00 ... 2.0900e+01 3.9524e+02 1.2270e+01]]
In [ ]: print(F"The Dependent Testing Set : {Y_test.shape} \n{Y_test}") # Printing The Inde
       The Dependent Testing Set : (102,)
                       8.3 21.2 19.9 20.6 18.7 16.1 18.6 8.8 17.2 14.9 10.5
       [22.6 50. 23.
        50. 29. 23. 33.3 29.4 21. 23.8 19.1 20.4 29.1 19.3 23.1 19.6 19.4
        38.7 18.7 14.6 20. 20.5 20.1 23.6 16.8 5.6 50. 14.5 13.3 23.9 20.
        19.8 13.8 16.5 21.6 20.3 17. 11.8 27.5 15.6 23.1 24.3 42.8 15.6 21.7
        17.1 17.2 15. 21.7 18.6 21. 33.1 31.5 20.1 29.8 15.2 15. 27.5 22.6
        20. 21.4 23.5 31.2 23.7 7.4 48.3 24.4 22.6 18.3 23.3 17.1 27.9 44.8
        50. 23. 21.4 10.2 23.3 23.2 18.9 13.4 21.9 24.8 11.9 24.3 13.8 24.7
        14.1 18.7 28.1 19.8]
```

### Resolved The Unbalanced Column 'CHAS'

```
In [ ]: # Convert it to a pandas Series
        X train_series = pd.DataFrame(X_train)
        # Now you can use value counts() on the Series
        counts = X_train_series[3].value_counts()
        print(counts)
       3
       0.0
              376
       1.0
               28
       Name: count, dtype: int64
In [ ]: # Convert it to a pandas Series
        X_test_series = pd.DataFrame(X_test)
        # Now you can use value_counts() on the Series
        counts = X_test_series[3].value_counts()
        print(counts)
       0.0
              95
       1.0
               7
       Name: count, dtype: int64
In [ ]: print(F"Ratio Of The CHAS IN X_train : {95//7}")
       Ratio Of The CHAS IN X_train : 13
In [ ]: print(F"Ratio Of The CHAS IN X_train : {376//28}")
       Ratio Of The CHAS IN X_train : 13
```

# The "CHAS" Is Been Resolved And It Has been Splited Evenly In Both 'X\_train' & 'X\_test' Respectively ABOVE

### **FEATURE SCALING**

```
In [ ]: from sklearn.preprocessing import StandardScaler
    FS=StandardScaler()
    X_FS= FS.fit_transform(X_train)
    Y_FS = FS.fit_transform(Y_train.reshape(len(Y_train),-1))
In [ ]: print(F"Featured Scaled Values of X_train :{X_FS.shape} \n{X_FS}")
```

```
Featured Scaled Values of X_train :(404, 13)
       [[-0.37257438 \ -0.49960763 \ -0.70492455 \ \dots \ -0.48463784 \ \ 0.3716906
         -0.41100022]
        [-0.39709866 -0.49960763 -0.04487755 ... 0.33649132 0.20501196
         -0.38768057]
                       0.77116771 -0.88675963 ... -0.84958414 0.36660893
        [-0.402693
         -0.18191902]
        . . .
        [-0.39805586 -0.49960763 -0.15941933 ... -0.30216469 0.40342278
         -0.33006734]
        [-0.38842357 \ -0.49960763 \ -0.60326872 \ \dots \ -0.25654641 \ \ 0.38343489
          0.8359148 ]
        [-0.39951258 - 0.49960763 - 1.01275558 \dots - 0.84958414 0.43041207
          0.27212814]]
In [ ]: print(F"Featured Scaled Values of Y_train :{Y_FS.shape} \n{Y_FS}")
```

```
Featured Scaled Values of Y_train :(404, 1)
[[ 4.43044345e-01]
[-9.88238901e-02]
[-6.63117960e-02]
 [ 3.12244864e-02]
 [-1.32344610e+00]
 [-7.71491607e-02]
 [-2.18034902e-01]
 [ 4.10532251e-01]
 [ 2.02529959e+00]
 [-5.86505302e-01]
 [ 4.86393804e-01]
 [-2.39709631e-01]
[-6.62366855e-01]
 [ 1.93784957e-01]
 [-1.54019340e+00]
 [ 4.20618512e-02]
 [-1.39930766e+00]
 [ 2.96815032e+00]
 [ 8.54864204e-01]
 [-1.11753617e+00]
 [-3.48083278e-01]
 [-1.53010714e-01]
 [-2.50546996e-01]
 [-4.13107467e-01]
 [ 1.16914778e+00]
 [-4.45619561e-01]
 [-3.26408549e-01]
 [ 1.14747305e+00]
 [-1.03083726e+00]
 [-1.63772968e+00]
 [-9.76650432e-01]
 [-5.64830573e-01]
 [-1.54019340e+00]
 [ 1.38589508e+00]
 [ 1.50435498e-01]
 [-9.98325161e-01]
 [ 3.88857522e-01]
 [-1.67024177e+00]
 [-1.03083726e+00]
 [ 2.04622322e-01]
 [ 1.58096764e+00]
 [ 2.58809145e-01]
 [ 1.61272863e-01]
 [-6.51529490e-01]
 [ 1.11496096e+00]
 [ 1.47259399e+00]
 [-1.25842191e+00]
 [-1.67024177e+00]
 [ 2.03871217e-02]
 [ 6.59791640e-01]
 [-8.89951514e-01]
 [ 1.93784957e-01]
 [-4.88969020e-01]
 [-1.21249724e-02]
```

[-2.28872267e-01]

- [-8.03252596e-01]
- [-5.64830573e-01]
- [-9.76650432e-01]
- [-1.50768130e+00]
- [-4.78131655e-01]
- [ 1.38589508e+00]
- [ 9.84912581e-01]
- [ 1.12579832e+00]
- [-4.46370665e-02]
- [-2.39709631e-01]
- [ 1.39598134e-01]
- [ 2.58809145e-01]
- [-1.07418671e+00]
- [ 7.03141099e-01]
- [-1.15004827e+00]
- [-5.32318479e-01]
- [ 4.75556440e-01]
- [-2.18034902e-01]
- [-1.34512083e+00]
- [-5.53993208e-01]
- [-3.15571184e-01]
- [ 7.79002651e-01]
- [-2.28872267e-01]
- [-8.35764690e-01]
- [-1.26925928e+00]
- [-3.37245914e-01]
- [ 9.54975702e-03]
- [-3.37245914e-01]
- [ 2.15459687e-01]
- [ 2.58809145e-01]
- [ 2.04622322e-01]
- [ 2.96815032e+00]
- [-9.00788879e-01]
- [-1.17172300e+00]
- [ 9.09051028e-01]
- [ 6.59791640e-01]
- [-6.94878949e-01]
- [ 2.26372162e+00]
- [ 2.58809145e-01]
- [-6.63117960e-02]
- [-2.93896455e-01]
- [-5.54744312e-02]
- [ 2.53465573e+00]
- [ 3.12244864e-02]
- [-2.61384361e-01]
- [ 2.22037216e+00]
- [ 1.29919616e+00]
- [-9.54975702e-01]
- [ 1.82947592e-01]
- [-1.20498620e-01]
- [ 1.93784957e-01]
- [-1.53010714e-01]
- [ 1.28760769e-01] [ 4.32206981e-01]
- [ 2.69646510e-01]
- [-1.41014502e+00]

- [-3.48083278e-01]
- [-3.48083278e-01]
- [-1.42098238e+00]
- [-9.33300973e-01]
- [ 4.21369616e-01]
- [-9.54975702e-01]
- [ 1.31003352e+00]
- [-6.84041584e-01]
- [-9.88238901e-02]
- [-5.53993208e-01]
- [-7.59903137e-01]
- [-1.85522808e-01]
- [-9.88238901e-02]
- [-1.07418671e+00]
- [-4.45619561e-01]
- [ 1.17923404e-01]
- [-3.58920643e-01]
- [-1.07418671e+00]
- [-1.09661255e-01]
- [ 6.37365806e-02]
- [ 7.57327922e-01]
- [-1.53010714e-01]
- [ 1.28760769e-01]
- [-5.97342667e-01]
- [-6.63117960e-02]
- [ 4 =0=40400 00]
- [ 1.50510609e+00]
- [-4.13107467e-01]
- [-7.71491607e-02]
- [ 5.28992159e-02]
- [-2.61384361e-01]
- [-5.64830573e-01]
- [-3.048363736-61
- [ 1.55929291e+00] [ 1.61272863e-01]
- [ 1.47259399e+00]
- [-7.49065773e-01]
- [ 1.03909940e+00]
- [-9.87487796e-01]
- [-5.10643749e-01]
- [-1.00916253e+00]
- [-1.18256036e+00]
- [ 2.06864905e+00]
- [-4.56456926e-01]
- [-1.03083726e+00]
- [ 2.58809145e-01]
- [-1.53010714e-01]
- [-7.16553679e-01]
- [ 1.33170825e+00]
- [ 2.80483875e-01]
- [ 2.37134416e-01]
- [-1.20498620e-01]
- [ 8.54113100e-02]
- [-4.02270102e-01]
- [-1.28009664e+00]
- [-1.74685443e-01]
- [ 5.29743263e-01]
- [-5.53993208e-01]

- [-9.87487796e-01]
- [ 6.59791640e-01]
- [-8.46602055e-01]
- [-3.80595373e-01]
- [ 6.48954275e-01]
- [-1.03083726e+00]
- [-3.91432737e-01]
- [-1.22590982e+00]
- [-1.00916253e+00]
- [-2.29623371e-02]
- [-2.72221725e-01]
- [-4.78131655e-01]
- [ 3.12244864e-02]
- [-2.18034902e-01]
- [ 2.58809145e-01]
- [-1.06334935e+00]
- [ 1.33170825e+00]
- [ 1.17923404e-01]
- [ 2.96815032e+00]
- [ 6.92303734e-01]
- [ 8.11514746e-01]
- [-6.63117960e-02]
- [-7.59903137e-01]
- [ 7.45739453e-02]
- [ 8.11514746e-01]
- [-9.00788879e-01]
- [ 2.03871217e-02]
- [ 2.96815032e+00]
- [-1.96360172e-01]
- [-1.76777806e+00]
- [ 1.33170825e+00]
- [ 1.06077413e+00]
- [-2.93896455e-01]
- [-2.50546996e-01]
- [-5.21481114e-01]
- [ 5.28992159e-02] [-2.39709631e-01]
- [ 6.37365806e-02]
- [-1.69191650e+00]
- [-6.29854761e-01]
- [ 2.61051729e+00]
- [ 2.96815032e+00]
- [ 3.12244864e-02]
- [ 1.39598134e-01]
- [-1.31335984e-01] [-9.88238901e-02]
- [-7.81577867e-01]
- [-7.92415232e-01]
- [ 5.28992159e-02]
- [ 1.39598134e-01]
- [-3.48083278e-01]
- [-1.16088563e+00]
- [-5.21481114e-01] [ 9.63237852e-01]
- [ 1.21249724e+00]
- [-1.96360172e-01]

- [-3.04733820e-01]
- [-2.29623371e-02]
- [-1.90866380e+00]
- [ 2.04622322e-01]
- [-3.48083278e-01]
- [-8.14089961e-01]
- [-4.78131655e-01]
- [-3.58920643e-01]
- [ 4.86393804e-01]
- [-2.07197537e-01]
- [ 1.62431710e+00]
- [-1.18256036e+00]
- [ 1.16914778e+00]
- [ 8.11514746e-01]
- [-1.31335984e-01]
- [ 2.46963155e+00]
- [-2.72221725e-01]
- [-1.96360172e-01]
- [ 4.10532251e-01]
- [-1.32344610e+00]
- [-8.79865254e-02]
- [ 1.01742468e+00]
- [-9.88238901e-02]
- [-4.56456926e-01]
- [ 1.65682919e+00]
- [-5.21481114e-01]
- [ 5.83930087e-01]
- [ 6.05604816e-01]
- [ 1.45091926e+00]
- [-4.02270102e-01]
- [-8.24927326e-01]
- [-1.21249724e-02]
- [ 8.76538934e-01]
- [-2.83059090e-01] [-3.80595373e-01]
- [ 7.45739453e-02]
- [ 4.32206981e-01]
- [-1.63848078e-01]
- [-3.15571184e-01]
- [-2.83059090e-01]
- [-1.13921090e+00]
- [-1.67024177e+00]
- [-9.11626243e-01]
- [-5.75667937e-01]
- [ 5.29743263e-01]
- [-4.46370665e-02]
- [-1.26925928e+00]
- [-3.69758008e-01]
- [ 1.01742468e+00] [-8.79114149e-01]
- [ 2.26297051e-01]
- [-1.08502408e+00] [ 1.50435498e-01]
- [ 1.61272863e-01]
- [ 2.96815032e+00]
- [-7.05716314e-01]

- [ 2.29623371e+00]
- [ 4.32206981e-01]
- [ 1.46175663e+00]
- [-8.79865254e-02]
- [ 7.89840016e-01]
- [ 2.96815032e+00]
- [ 2.31790844e+00]
- [-2.18034902e-01]
- [-3.26408549e-01]
- [ 6.27279546e-01]
- [-3.80595373e-01]
- [-3.37997018e-02]
- [-1.85522808e-01]
- [ 6.27279546e-01] [-8.89951514e-01]
- [ 1.09328623e+00]
- [-9.54975702e-01]
- [-1.52935603e+00]
- [-1.21249724e-02]
- [ 1.35338298e+00]
- [ 9.74075216e-01]
- [-5.21481114e-01]
- [-7.59903137e-01]
- [-2.07197537e-01]
- [ 1.86273912e+00]
- [-5.21481114e-01]
- [-3.26408549e-01]
- [-8.35764690e-01]
- [-6.63117960e-02]
- [ 2.83810195e+00]
- [ 2.58809145e-01]
- [ 2.80558985e+00]
- [ 1.39598134e-01]
- [-2.50546996e-01]
- [-8.03252596e-01] [-1.31260874e+00]
- [-3.91432737e-01]
- [-6.73204220e-01]
- [-1.49684394e+00]
- [-6.63117960e-02]
- [ 2.37134416e-01]
- [ 2.96815032e+00]
- [-3.58920643e-01]
- [ 9.54975702e-03]
- [ 1.59180500e+00]
- [ 9.74075216e-01]
- [-1.55103076e+00]
- [ 5.28992159e-02]
- [ 2.96815032e+00]
- [-9.44138338e-01]
- [-7.05716314e-01]
- [ 2.91321239e-01] [-3.37245914e-01]
- [-1.34512083e+00]
- [-2.93896455e-01]
- [ 1.36422035e+00]

- [-9.98325161e-01]
- [ 2.26297051e-01]
- [-1.20423509e+00]
- [ 8.54113100e-02]
- [-6.40692126e-01]
- [-7.81577867e-01]
- [-4.99806384e-01]
- [ 6.38116910e-01]
- [-4.56456926e-01]
- [ 1.07161150e+00]
- [ 2.96815032e+00]
- [ 2.96815032e+00]
- [-3.26408549e-01]
- [-5.43155843e-01]
- [ 2.13367324e+00]
- [-8.79114149e-01]
- [-1.01999989e+00]
- [-6.94878949e-01]
- [ 7.57327922e-01]
- [-6.40692126e-01]
- [-1.04167462e+00]
- [-3.37997018e-02]
- [-9.98325161e-01]
- [-1.90866380e+00]
- [-3.37245914e-01]
- [-8.68276785e-01]
- [-6.63117960e-02]
- [-1.57270549e+00]
- [ 2.15459687e-01]
- [ 1.14747305e+00]
- [-2.18034902e-01]
- [-9.22463608e-01]
- [-9.22463608e-01]
- [-1.63848078e-01]
- [ 8.33189475e-01]
- [ 1.17923404e-01]
- [-1.31335984e-01]
- [-4.67294290e-01]
- [-1.52935603e+00]
- [ 2.03871217e-02]
- [ 2.03871217e-02]
- [-4.46370665e-02]
- [-9.44138338e-01]
- [ 2.58809145e-01]
- [-4.45619561e-01]
- [-1.69191650e+00]
- [-1.28760769e-03]
- [-2.72221725e-01]
- [ 8.87376299e-01]
- [ 9.19888393e-01]
- [ 1.28760769e-01]
- [-1.09586144e+00]
- [ 1.07086039e-01] [ 6.37365806e-02]
- [ 1.72110228e-01]
- [-4.46370665e-02]

```
[ 5.73092722e-01]
[-4.46370665e-02]
[ 1.13663569e+00]
[-3.58920643e-01]
[-4.02270102e-01]
[-1.28760769e-03]
[ 2.96815032e+00]
[ 2.37134416e-01]
[-4.45619561e-01]
[ 1.49426872e+00]
[ -3.69758008e-01]
[ -6.51529490e-01]
[ 5.28992159e-02]
```

# MODEL SELECTION < SELECTIVE BEST , < AS PER GIVEN DATA >

LINEAR REGRESSION MODEL

# Predicting the Test set results < LINEAR REGRESSION >

```
In [ ]: Y_pred = LR.predict(X_test)
    np.set_printoptions(precision=2)
    print(np.concatenate((Y_pred.reshape(len(Y_pred),1), Y_test.reshape(len(Y_test),1))
```

[[24.87 22.6] [23.73 50.] [29.38 23. [12.14 8.3] [21.44 21.2 ] [19.3 19.9] [20.51 20.6] [21.38 18.7] [18.91 16.1] [19.91 18.6] [ 5.14 8.8 ] [16.37 17.2] [17.08 14.9] [ 5.61 10.5 ] [39.98 50. ] [32.5 29. [22.47 23. [36.85 33.3 ] [30.85 29.4] [23.16 21. [24.78 23.8] [24.68 19.1] [20.57 20.4] [30.35 29.1] [22.44 19.3] [10.3 23.1] [17.65 19.6] [18.29 19.4] [35.49 38.7] [20.94 18.7] [18.3 14.6] [17.81 20. ] [20.01 20.5] [24.08 20.1] [29.09 23.6] [19.3 16.8] [11.15 5.6] [24.59 50.] [17.58 14.5] [15.48 13.3] [26.22 23.9] [20.87 20. ] [22.31 19.8] [15.63 13.8 ] [23. 16.5] [25.18 21.6] [20.14 20.3] [22.91 17. ] [10.05 11.8] [24.26 27.5] [20.95 15.6] [17.33 23.1 ] [24.53 24.3 ] [29.94 42.8] [13.43 15.6] [21.74 21.7 ]

```
[20.81 17.1]
[15.52 17.2]
[14.01 15. ]
[22.19 21.7]
[17.78 18.6]
[21.59 21. ]
[32.9 33.1]
[31.1 31.5]
[17.74 20.1 ]
[32.75 29.8]
[18.71 15.2]
[19.46 15.]
[19.
      27.5 ]
[22.9 22.6]
[22.96 20. ]
[24.02 21.4]
[30.75 23.5]
[28.83 31.2]
[25.9 23.7]
[ 5.24 7.4 ]
[36.68 48.3]
[23.77 24.4 ]
[27.25 22.6]
[19.29 18.3]
[28.63 23.3 ]
[19.16 17.1]
[18.99 27.9]
[37.78 44.8 ]
[39.16 50.]
[23.72 23. ]
[25.36 21.4 ]
[15.89 10.2]
[26.11 23.3 ]
[16.69 23.2]
[15.85 18.9]
[13.09 13.4]
[24.76 21.9]
[31.27 24.8 ]
[22.17 11.9]
[20.29 24.3 ]
[ 0.6 13.8 ]
[25.48 24.7 ]
[15.57 14.1 ]
[17.98 18.7 ]
[25.3 28.1]
[22.35 19.8 ]]
```

# Evaluating the Model Performance < LINEAR REGRESSION >

```
In [ ]: from sklearn.metrics import r2_score
# Assuming RMS is the R^2 score
RMS = r2_score(Y_test, Y_pred)
```

```
# Format the RMS value with two decimal points
formatted RMS = "{:.2f}".format(RMS)
# Convert formatted_RMS to float for comparison
formatted_RMS_float = float(formatted_RMS)
if formatted_RMS_float==1:
    print(f"The R^2 Score {formatted_RMS} %
                                               'Over Fitted Model'")
elif 0.8 < formatted RMS float < 0.9:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'Perfect Model'")
elif 0.7 < formatted_RMS_float < 0.8:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'accurate Model'")
elif 0.6 < formatted_RMS_float <0.7:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'Modrate Model'")
elif 0.5< formatted RMS float < 0.6:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'Under Fitted Model'")
    print(f"The R^2 Score {formatted_RMS} %
                                                'Weak Model'")
```

The R^2 Score 0.59 % 'Under Fitted Model'

### SELECTED MODEL < DECISION REGRESSION >

# Predicting the Test set results < DECISION REGRESSION >

```
In [ ]: Y_pred = DT.predict(X_test)
    np.set_printoptions(precision=2)
    print(np.concatenate((Y_pred.reshape(len(Y_pred),1), Y_test.reshape(len(Y_test),1))
```

[[23.7 22.6] [22.4 50.] [19.2 23.] [11.7 8.3] [21.7 21.2] [20.3 19.9] [21.2 20.6] [19.6 18.7] [19.9 16.1] [16.1 18.6] [ 7.2 8.8] [17.9 17.2] [14.2 14.9] [ 8.8 10.5] [48.5 50.] [34.6 29.] [21.4 23.] [34.9 33.3] [24.1 29.4] [20.3 21.] [24.7 23.8] [21.5 19.1] [20.1 20.4] [24.7 29.1] [21.8 19.3] [16.1 23.1] [17.5 19.6] [16.4 19.4] [39.8 38.7] [20.4 18.7] [12.5 14.6] [19.1 20.] [17.6 20.5] [22.2 20.1] [23.8 23.6] [17.7 16.8] [ 8.5 5.6] [20.3 50.] [11.7 14.5] [17.9 13.3] [23.4 23.9] [21.1 20.] [20.9 19.8] [12.5 13.8] [22. 16.5] [22. 21.6] [22.4 20.3] [16.4 17.] [13.4 11.8] [26.7 27.5] [19. 15.6] [15.6 23.1] [20.3 24.3] [35.2 42.8] [15.6 15.6]

[18.9 21.7]

```
[19.4 17.1]
[17.5 17.2]
[10.9 15.]
[22.5 21.7]
[20.6 18.6]
[23.1 21.]
[29.6 33.1]
[22.8 31.5]
[20.6 20.1]
[24.1 29.8]
[17.4 15.2]
[27.1 15.]
[16.3 27.5]
[23.1 22.6]
[20.6 20. ]
[26.4 21.4]
[24.1 23.5]
[30.5 31.2]
[25. 23.7]
[ 7.2 7.4]
[41.7 48.3]
[23.1 24.4]
[23.9 22.6]
[18.5 18.3]
[28.6 23.3]
[14.5 17.1]
[20.8 27.9]
[41.7 44.8]
[41.7 50.]
[25. 23.]
[21.2 21.4]
[13.8 10.2]
[24.4 23.3]
[12.7 23.2]
[16.6 18.9]
[11.8 13.4]
[20.6 21.9]
[30.1 24.8]
[21.1 11.9]
[20.9 24.3]
[11.9 13.8]
[22.9 24.7]
[12.7 14.1]
[17.6 18.7]
[25. 28.1]
[21.2 19.8]]
```

# Evaluating the Model Performance < DECISION REGRESSION >

```
In [ ]: from sklearn.metrics import r2_score

# Assuming RMS is the R^2 score
RMS = r2_score(Y_test, Y_pred)
```

```
# Format the RMS value with two decimal points
formatted RMS = "{:.2f}".format(RMS)
# Convert formatted_RMS to float for comparison
formatted_RMS_float = float(formatted_RMS)
if formatted_RMS_float==1.0:
    print(f"The R^2 Score {formatted_RMS} %
                                              'Over Fitted Model'")
elif 0.8 < formatted RMS float < 0.9:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'Perfect Model'")
elif 0.7 < formatted_RMS_float < 0.8:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'accurate Model'")
elif 0.6 < formatted_RMS_float <0.7:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                              'Modrate Model'")
elif 0.0 < formatted RMS float < 0.5:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                              'Under Fitted Model'")
elif formatted_RMS_float < 0.0:</pre>
                                              'BAD Model'")
    print(f"The R^2 Score {formatted_RMS} %
    print(f"The R^2 Score {formatted_RMS} %
                                               'Weak Model'")
```

The R^2 Score 0.63 % 'Modrate Model'

### SELECTED MODEL < RANDOM FOREST REGRESSION >

## Predicting the Test set results < RANDOM FOREST REGRESSION >

```
In [ ]: Y_predRFR = RFR.predict(X_test)
    np.set_printoptions(precision=2)
    print(np.concatenate((Y_predRFR.reshape(len(Y_predRFR),1), Y_test.reshape(len(Y_test)))
```

[[24.06 22.6] [27.99 50.] [22.08 23. [11.01 8.3] [20.77 21.2 ] [20.74 19.9] [21.21 20.6] [20.14 18.7] [20.5 16.1] [18.75 18.6] [ 8.24 8.8 ] [15.32 17.2] [15.02 14.9 ] [ 8.43 10.5 ] [47.33 50. [34. 29. [21.07 23. [34.69 33.3 ] [25.59 29.4] [21.19 21. [23.72 23.8] [21.96 19.1] [19.74 20.4] [24.5 29.1] [20.39 19.3] [17.77 23.1] [18.65 19.6] [16.05 19.4] [43.98 38.7] [19.2 18.7] [14.87 14.6] [17.57 20. ] [20.1 20.5] [21.47 20.1] [22.97 23.6] [17.88 16.8 ] [ 8.64 5.6 ] [27.94 50.] [14.52 14.5] [15.51 13.3] [22.76 23.9] [20.88 20. ] [22.42 19.8] [15.16 13.8 ] [23.69 16.5] [22.42 21.6] [21.15 20.3] [16.56 17. ] [14.54 11.8] [25.17 27.5] [16.32 15.6] [19.78 23.1 ] [21.96 24.3 ] [39.87 42.8] [15.04 15.6] [21.21 21.7]

```
[19.68 17.1]
[19.12 17.2]
[21.72 15. ]
[20.08 21.7]
[21.59 18.6]
[21.69 21. ]
[33.29 33.1]
[28.01 31.5]
[18.36 20.1]
[26.92 29.8]
[15.99 15.2]
[21.17 15.]
[17.28 27.5]
[22.02 22.6]
[20.15 20. ]
[22.89 21.4]
[24.37 23.5]
[30.66 31.2 ]
[29.56 23.7]
[ 8.63 7.4 ]
[43.01 48.3]
[22.41 24.4 ]
[22.8 22.6]
[20.27 18.3]
[26.57 23.3]
[18.22 17.1]
[22.46 27.9]
[42.5 44.8]
[41.7 50.]
[24.37 23.]
[22.7 21.4]
[14.92 10.2 ]
[26.44 23.3]
[16.07 23.2]
[19.2 18.9]
[12.04 13.4 ]
[22.51 21.9 ]
[29.91 24.8 ]
[21.13 11.9]
[22.02 24.3 ]
[11.61 13.8]
[23.42 24.7 ]
[14.88 14.1 ]
[19.22 18.7]
[24.05 28.1]
[20.06 19.8 ]]
```

# Evaluating the Model Performance < RANDOM FOREST REGRESSION >

```
In [ ]: from sklearn.metrics import r2_score

# Assuming RMS is the R^2 score
RMS = r2_score(Y_test, Y_predRFR)
```

```
# Format the RMS value with two decimal points
formatted RMS = "{:.2f}".format(RMS)
# Convert formatted_RMS to float for comparison
formatted_RMS_float = float(formatted_RMS)
if formatted_RMS_float==1.0:
    print(f"The R^2 Score {formatted_RMS} % 'Over Fitted Model'")
elif 0.9 < formatted RMS float < 1:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'Perfect Model'")
elif 0.8 < formatted_RMS_float < 0.9:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                               'accurate Model'")
elif 0.7 < formatted_RMS_float <0.8:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                              'Modrate Model'")
elif 0.5 < formatted RMS float < 0.6:</pre>
    print(f"The R^2 Score {formatted_RMS} % 'Under Fitted Model'")
elif formatted_RMS_float < 0.0:</pre>
    print(f"The R^2 Score {formatted_RMS} %
                                              'BAD Model'")
    print(f"The R^2 Score {formatted_RMS} %
                                               'Weak Model'")
```

The R^2 Score 0.77 % 'Modrate Model'

### **Grid Search CV**

```
In [ ]: from sklearn.model_selection import GridSearchCV
        from sklearn.ensemble import RandomForestRegressor
        # Define the parameter grid
        parameters = {
            'n_estimators': [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000],
            'criterion': ["squared_error", "absolute_error", "friedman_mse", "poisson"],
            'max_depth': [20]
        # Initialize a RandomForestRegressor (you can choose any model)
        model = RandomForestRegressor()
        # Initialize GridSearchCV with the model, parameter grid, and any other necessary a
        GCV = GridSearchCV(estimator=model, param_grid=parameters, cv=5) # cv is the numbe
        GCV.fit(X_train,Y_train)
Out[]:
                  GridSearchCV (1) (?)
         ▶ estimator: RandomForestRegressor
               RandomForestRegressor
In [ ]: Best_prarm=GCV.best_params_
        print(Best_prarm)
       {'criterion': 'squared_error', 'max_depth': 20, 'n_estimators': 100}
```

[[24.06 22.6] [27.99 50.] [22.08 23. [11.01 8.3] [20.77 21.2 ] [20.74 19.9] [21.21 20.6] [20.14 18.7] [20.5 16.1] [18.75 18.6] [ 8.24 8.8 ] [15.32 17.2] [15.02 14.9 ] [ 8.43 10.5 ] [47.33 50. ] [34. 29. [21.07 23. ] [34.69 33.3] [25.59 29.4] [21.19 21. [23.72 23.8] [21.96 19.1] [19.74 20.4] [24.5 29.1] [20.39 19.3] [17.77 23.1] [18.65 19.6] [16.05 19.4] [43.98 38.7] [19.2 18.7] [14.87 14.6] [17.57 20. ] [20.1 20.5] [21.47 20.1] [22.97 23.6] [17.88 16.8 ] [ 8.64 5.6 ] [27.94 50.] [14.52 14.5] [15.51 13.3] [22.76 23.9] [20.88 20. ] [22.42 19.8] [15.16 13.8 ] [23.69 16.5] [22.42 21.6] [21.15 20.3] [16.56 17. ] [14.54 11.8] [25.17 27.5] [16.32 15.6] [19.78 23.1 ] [21.96 24.3 ] [39.87 42.8] [15.04 15.6] [21.21 21.7]

```
[19.68 17.1]
[19.12 17.2]
[21.72 15. ]
[20.08 21.7]
[21.59 18.6]
[21.69 21. ]
[33.29 33.1 ]
[28.01 31.5]
[18.36 20.1]
[26.92 29.8]
[15.99 15.2]
[21.17 15.]
[17.28 27.5]
[22.02 22.6 ]
[20.15 20. ]
[22.89 21.4]
[24.37 23.5]
[30.66 31.2 ]
[29.56 23.7]
[ 8.63 7.4 ]
[43.01 48.3]
[22.41 24.4 ]
[22.8 22.6]
[20.27 18.3]
[26.57 23.3 ]
[18.22 17.1 ]
[22.46 27.9]
[42.5 44.8]
[41.7 50.]
[24.37 23.
[22.7 21.4]
[14.92 10.2 ]
[26.44 23.3 ]
[16.07 23.2]
[19.2 18.9]
[12.04 13.4 ]
[22.51 21.9]
[29.91 24.8]
[21.13 11.9]
[22.02 24.3 ]
[11.61 13.8]
[23.42 24.7 ]
[14.88 14.1 ]
[19.22 18.7]
[24.05 28.1]
[20.06 19.8 ]]
```

### Score

```
In [ ]: G=GCV.best_score_
print(G)
```

0.8626805578047279

### SAVING THE Grid Search CV

```
In []: from joblib import dump, load
dump(GCV, 'REAL.joblib')
Out[]: ['REAL.joblib']
```

#### **USING THE MODEL**

[29.55]

C:\Users\surya\AppData\Local\Temp\ipykernel\_22704\297953511.py:8: DeprecationWarnin g: Conversion of an array with ndim > 0 to a scalar is deprecated, and will error in future. Ensure you extract a single element from your array before performing this o peration. (Deprecated NumPy 1.25.)

PRICE=(int(K)\*1000)

### [29.55] By GCV

### [30.12] RFR

```
In [ ]: print(F"THE PRICE OF THE HOUSE IS {PRICE} $ IN BOSTON CITY")
```

THE PRICE OF THE HOUSE IS 29000 \$ IN BOSTON CITY

BY RFR <24,39,720 RUPEES> / < 30000 > / / BY GCV <24,07,951 RUPEES>/<29000 >