

# DATA SCIENCE WEEK III

## PROGRAM:

```
from sklearn.preprocessing import MinMaxScaler,StandardScaler
import pandas as pd
import matplotlib.pyplot as plt
tips = pd.read_csv('student_exam_scores.csv')
print(tips.head())
datasdt = tips[['previous_scores','exam_score']]
print(datasdt)
minmax = MinMaxScaler()
minmax = pd.DataFrame(minmax.fit_transform(datasdt),columns=['pre','scr'])
print(minmax)
dates = tips[['sleep_hours','attendance_percent']]
scalerstd = StandardScaler()
scalerstd = pd.DataFrame(scalerstd.fit_transform(dates),columns=['slp','atp'])
print(scalerstd)
plt.hist(minmax['scr'],bins=5,color='blue',edgecolor='black')
plt.xlabel('TOTAL PERCENTAGE')
plt.ylabel('TOTAL SCORE')
plt.title('STUDENT TOTAL MARKS')
plt.show()
```

# SCREENSHOTS:

The screenshot shows a Jupyter Notebook interface with several tabs at the top: 'exercises.ipynb', 'ds4q.ipynb', 'DATA Sales Data (2010-2014).csv', 'ds4q.ipynb', 'ds4q.ipynb', and 'ds4q.ipynb'. The main area displays a Python script named 'ds4q.py'.

```
1 #!/usr/bin/python
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 tips = pd.read_csv('student_exam_scores.csv')
5 print(tips.head())
6 datasdt = tips[['previous_scores','exam_score']]
7 print(datasdt)
8 minmax = MinMaxScaler()
9 minmax = pd.DataFrame(minmax.fit_transform(datasdt),columns=['pre','scr'])
10 print(minmax)
11 dates = tips[['sleep_hours','attendance_percent']]
12 scalerstd = StandardScaler()
13 scalerstd = pd.DataFrame(scalerstd.fit_transform(dates),columns=['slp','atp'])
14 print(scalerstd)
15 plt.hist(minmax['scr'],bins=5,color='blue',edgecolor='black')
16 plt.xlabel('TOTAL PERCENTAGE')
17 plt.ylabel('TOTAL SCORE')
18 plt.title('STUDENT TOTAL MARKS')
19 plt.show()
20
```

Below the code, there is a table output:

	previous_scores	exam_score
0	45	30.2
1	55	25.0
2	86	35.8
3	66	34.0
4	71	40.3
..	...	...
195	87	42.7
196	92	40.4
197	76	28.2
198	58	42.0

At the bottom right, it says 'Ln 19, Col 11 Spaces: 4'

```
  Exam3.py      ds4q.py      BMW Sales Data (2015-2017) (1).csv      ds4q.py      ds4q.py      dsEq.ipynb      dsEq.ipynb
  ds4q.py > ...
1  import pandas as pd
2  import matplotlib.pyplot as plt
3  tips = pd.read_csv('student_exam_scores.csv')
4  print(tips.head())
5  datasdt = tips[['previous_scores', 'exam_score']]
6  print(datasdt)
7  minmax = MinMaxScaler()
8  minmax = pd.DataFrame(minmax.fit_transform(datasdt), columns=['pre', 'scr'])
9  print(minmax)
10 dates = tips[['sleep_hours', 'attendance_percent']]
11 scalerstd = StandardScaler()
12 scalerstd = pd.DataFrame(scalerstd.fit_transform(dates), columns=['slp', 'atp'])
13 print(scalerstd)
14 plt.hist(minmax['scr'], bins=5, color='blue', edgecolor='black')
15 plt.xlabel('TOTAL PERCENTAGE')
16 plt.ylabel('TOTAL SCORE')
17 plt.title('STUDENT TOTAL MARKS')
18 plt.show()
19
20
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

	previous_scores	exam_score
0	45	30.2
1	55	25.0
2	86	35.8
3	66	34.0
4	71	40.3
..	...	...
195	87	42.7
196	92	40.4
197	76	28.2
198	58	42.0

Ln 19, Col 11 Spaces: 4

```
ds4q.py > ...
2  import pandas as pd
3  import matplotlib.pyplot as plt
4  tips = pd.read_csv('student_exam_scores.csv')
5  print(tips.head())
6  datasdt = tips[['previous_scores','exam_score']]
7  print(datasdt)
8  minmax = MinMaxScaler()
9  minmax = pd.DataFrame(minmax.fit_transform(datasdt),columns=['pre','scr'])
10 print(minmax)
11 dates = tips[['sleep_hours','attendance_percent']]
12 scalerstd = StandardScaler()
13 scalerstd = pd.DataFrame(scalerstd.fit_transform(dates),columns=['slp','atp'])
14 print(scalerstd)
15 plt.hist(minmax['scr'],bins=5,color='blue',edgecolor='black')
16 plt.xlabel('TOTAL PERCENTAGE')
17 plt.ylabel('TOTAL SCORE')
18 plt.title('STUDENT TOTAL MARKS')
19 plt.show()
20
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
[200 rows x 2 columns]
      pre      scr
0    0.090909  0.383041
1    0.272727  0.230994
2    0.836364  0.546784
3    0.472727  0.494152
4    0.563636  0.678363
..    ...
195  0.854545  0.748538
196  0.945455  0.681287
197  0.654545  0.324561
```

In 19. Col 11 Spaces:

Help ← → DS EXAM WORK

exam3.py ds4q.py ✎ BMW sales data (2010-2024) (1).csv ds1q.py ds2q.py ds2.2.py

ds4q.py > ...

```

2  import pandas as pd
3  import matplotlib.pyplot as plt
4  tips = pd.read_csv('student_exam_scores.csv')
5  print(tips.head())
6  datasdt = tips[['previous_scores','exam_score']]
7  print(datasdt)
8  minmax = MinMaxScaler()
9  minmax = pd.DataFrame(minmax.fit_transform(datasdt),columns=['pre','scr'])
10 print(minmax)
11 dates = tips[['sleep_hours','attendance_percent']]
12 scalerstd = StandardScaler()
13 scalerstd = pd.DataFrame(scalerstd.fit_transform(dates),columns=['slp','atp'])
14 print(scalerstd)
15 plt.hist(minmax['scr'],bins=5,color='blue',edgecolor='black')
16 plt.xlabel('TOTAL PERCENTAGE')
17 plt.ylabel('TOTAL SCORE')
18 plt.title('STUDENT TOTAL MARKS')
19 plt.show()
20

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

[200 rows x 2 columns]

	slp	atp
0	1.458426	-0.192061
1	1.324503	-0.994074
2	1.056656	-0.079498
3	-1.220042	1.426035
4	-0.148655	1.053170
..	...	...
195	-0.818272	1.348648
196	-0.349540	0.722515
197	0.186154	-0.775983

Ln 19, Col 11 Sp

occurrence.py tuple.py exam3.py ds4q.py ✎ BMW sales data (2010-2024) (1).csv ds1q.py ds2q.py ds2.2.py ds3q.py ds5q.py

ds4q.py > ...

```

2  import pandas as pd
3  import matplotlib.pyplot as plt
4  tips = pd.read_csv('student_exam_scores.csv')
5  print(tips.head())
6  datasdt = tips[['previous_scores','exam_score']]
7  print(datasdt)
8  minmax = MinMaxScaler()
9  minmax = pd.DataFrame(minmax.fit_transform(datasdt),columns=['pre','scr'])
10 print(minmax)
11 dates = tips[['sleep_hours','attendance_percent']]
12 scalerstd = StandardScaler()
13 scalerstd = pd.DataFrame(scalerstd.fit_transform(dates),columns=['slp','atp'])
14 print(scalerstd)
15 plt.hist(minmax['scr'],bins=5,color='aqua',edgecolor='black')
16 plt.xlabel('TOTAL PERCENTAGE')
17 plt.ylabel('TOTAL SCORE')
18 plt.title('STUDENT TOTAL MARKS')
19 plt.show()
20

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

	slp	atp
3	-1.238042	1.426035
4	-0.148655	1.053170
..	...	...
195	-0.818272	1.348648
196	-0.349540	0.722515
197	0.186154	-0.775983
198	0.454009	1.711664
199	-0.215617	1.587845

Figure 1