**AI and Visualization Lab**

**Exp:1 Techniques for Data Pre-processing: Mean Removal, Scaling, Normalization**

**AIM**

To perform techniques for data pre-processing:Mean removal,scaling,normalization

**ALGORITHM:**

1. Import necessary libraries(pandas as pd, numpy as np ,from sklearn.preprocessing import StandardScaler)
2. Read the data from 'data.csv'.
3. Display the original data.

4.Mean Removal:

- Calculate the mean of the 'Values' column

- Subtract the mean from each value in the 'Values' column

5.Normalization:

- Calculate normalized values for the 'Values' column

- Add the normalized values as a new column

6.Standardization using StandardScaler:

- Initialize a StandardScaler

- Transform the 'Values' column and add it as a new column

7.Display Mean Removed Data

8.Display Normalized Data

9.Display Standardized data and scaler

**CODE:**

import pandas as pd

import numpy as np

from sklearn.preprocessing import StandardScaler

file\_path = 'data.csv'

data = pd.read\_csv(file\_path, header=None, names=['Values'])

print("Original Data:")

print(data)

mean\_values = np.mean(data['Values'])

data['Mean\_Removed\_Values'] = data['Values'] - mean\_values

normalized\_values = (data['Values'] - np.min(data['Values'])) / (np.max(data['Values']) - np.min(data['Values']))

data['Normalized\_Values'] = normalized\_values

scaler = ((data['Values'] - max(data['Values']))/max(data['Values']))

print("\nMean Removed Data:")

print(data[['Mean\_Removed\_Values']])

print("\nNormalized Data:")

print(data[['Normalized\_Values']])

print("\nStandardized Data:")

print(scaler)

**OUTPUT:**

Original Data:

Values

0 70

1 67

2 57

3 64

4 74

5 65

6 56

7 59

8 60

9 63

10 59

11 53

12 44

13 61

14 57

15 71

16 46

17 53

18 64

19 40

20 67

Mean Removed Data:

Mean\_Removed\_Values

0 10.47619

1 7.47619

2 -2.52381

3 4.47619

4 14.47619

5 5.47619

6 -3.52381

7 -0.52381

8 0.47619

9 3.47619

10 -0.52381

11 -6.52381

12 -15.52381

13 1.47619

14 -2.52381

15 11.47619

16 -13.52381

17 -6.52381

18 4.47619

19 -19.52381

20 7.47619

Normalized Data:

Normalized\_Values

0 0.882353

1 0.794118

2 0.500000

3 0.705882

4 1.000000

5 0.735294

6 0.470588

7 0.558824

8 0.588235

9 0.676471

10 0.558824

11 0.382353

12 0.117647

13 0.617647

14 0.500000

15 0.911765

16 0.176471

17 0.382353

18 0.705882

19 0.000000

20 0.794118

Standardized Data:

0 -0.054054

1 -0.094595

2 -0.229730

3 -0.135135

4 0.000000

5 -0.121622

6 -0.243243

7 -0.202703

8 -0.189189

9 -0.148649

10 -0.202703

11 -0.283784

12 -0.405405

13 -0.175676

14 -0.229730

15 -0.040541

16 -0.378378

17 -0.283784

18 -0.135135

19 -0.459459

20 -0.094595

Name: Values, dtype: float64

**RESULT:**

Thus techniques for data pre-processing-mean removal,scaling,normalization are executed successfully.