

## KMEANSCLUSTER

```
import pandas as pd

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

from sklearn.preprocessing import StandardScaler df= pd.read_csv

("studentdata.csv" )

print(df)x=df[['python_marks','english_marks']]

scaler=StandardScaler()

x_scaled=scaler.fit_transform(x)

km=KMeans(n_clusters=2,random_state=77)

km.fit(x_scaled)

df['cluster']=km.labels_

print(df)

c1=df[df['cluster']==0]

print('cluster1')

print(c1)

c2=df[df['cluster']==1]

print('cluster2')

print(c2)
```

## OUTPUT:

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
C:\microsoft visual studio\python workspace\week4>"c:/microsoft visual studio/python.exe" "c:/microsoft visual studio/python workspace/week4/meandata.py"

  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks
0           4      Sydney  24         95           99           87           82           75
1           5        Tokyo  24         99           95           89           86           82
2           6        Berlin 22         72           70           99           79           77
3           7        London 23         97           90           74           72           85
4           8        Tokyo  22         91           71           79           80           75
..         ...         ...         ...         ...         ...         ...         ...         ...
492        496    New York  23         71           92           98           78           88
493        497  Los Angeles 20         90          100           90           78           73
494        498        Paris 22         78           80           78           71           70
495        499  Los Angeles 21         96           93           78           78           82
496        500        London 22         85           87           93           70           86

[497 rows x 8 columns]
  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks  cluster
0           4      Sydney  24         95           99           87           82           75           0
1           5        Tokyo  24         99           95           89           86           82           0
2           6        Berlin 22         72           70           99           79           77           0
3           7        London 23         97           90           74           72           85           1
4           8        Tokyo  22         91           71           79           80           75           1
..         ...         ...         ...         ...         ...         ...         ...         ...
492        496    New York  23         71           92           98           78           88           0
493        497  Los Angeles 20         90          100           90           78           73           0
494        498        Paris 22         78           80           78           71           70           0
495        499  Los Angeles 21         96           93           78           78           82           1
496        500        London 22         85           87           93           70           86           0

[497 rows x 9 columns]
cluster1
  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks  cluster
0           4      Sydney  24         95           99           87           82           75           0
1           5        Tokyo  24         99           95           89           86           82           0
2           6        Berlin 22         72           70           99           79           77           0
6           10       Tokyo  18         77           87          100           98           93           0
7           11      Toronto 21         78           90           88           79           72           0
..         ...         ...         ...         ...         ...         ...         ...         ...
488        490        Tokyo 18         87          100           95           89           85           0
492        496    New York  23         71           92           98           78           88           0
493        497  Los Angeles 20         90          100           90           78           73           0
494        498        Paris 22         78           80           78           71           70           0
496        500        London 22         85           87           93           70           86           0

[255 rows x 9 columns]
cluster2
  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks  cluster
3           7        London 23         97           90           74           72           85           1
4           8        Tokyo  22         91           71           79           80           75           1
5           9      Toronto 20         93           88           75           93           72           1
8          12  Melbourne 24         80           95           87           90           94           1
10         14      Toronto 20         81           74           88           99           96           1
..         ...         ...         ...         ...         ...         ...         ...         ...
488        492        London 23         85           73           78           81           96           1
489        493  Los Angeles 24         94           74           80           92           95           1
490        494  Melbourne 25         96           72           93           86           93           1
491        495    New York 24         83           78           82           84           84           1
495        499  Los Angeles 21         96           93           78           78           82           1

[242 rows x 9 columns]
C:\microsoft visual studio\python workspace\week4>"c:/microsoft visual studio/python.exe" "c:/microsoft visual studio/python workspace/week4/meandata.py"
```

```
← → week4
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
cluster1
  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks  cluster
0           4      Sydney  24         95           99           87           82           75           0
1           5        Tokyo  24         99           95           89           86           82           0
2           6        Berlin 22         72           70           99           79           77           0
6           10       Tokyo  18         77           87          100           98           93           0
7           11      Toronto 21         78           90           88           79           72           0
..         ...         ...         ...         ...         ...         ...         ...         ...
488        490        Tokyo 18         87          100           95           89           85           0
492        496    New York  23         71           92           98           78           88           0
493        497  Los Angeles 20         90          100           90           78           73           0
494        498        Paris 22         78           80           78           71           70           0
496        500        London 22         85           87           93           70           86           0

[255 rows x 9 columns]
cluster2
  student_id  location  age  sql_marks  excel_marks  python_marks  power_bi_marks  english_marks  cluster
3           7        London 23         97           90           74           72           85           1
4           8        Tokyo  22         91           71           79           80           75           1
5           9      Toronto 20         93           88           75           93           72           1
8          12  Melbourne 24         80           95           87           90           94           1
10         14      Toronto 20         81           74           88           99           96           1
..         ...         ...         ...         ...         ...         ...         ...         ...
488        492        London 23         85           73           78           81           96           1
489        493  Los Angeles 24         94           74           80           92           95           1
490        494  Melbourne 25         96           72           93           86           93           1
491        495    New York 24         83           78           82           84           84           1
495        499  Los Angeles 21         96           93           78           78           82           1

[242 rows x 9 columns]
C:\microsoft visual studio\python workspace\week4>"c:/microsoft visual studio/python.exe" "c:/microsoft visual studio/python workspace/week4/meandata.py"
```