DATA SCIENCE PCA LAB 5

Roll no: 20K-0409

Screen Shots

DATAset#1

```
    return FinalSelectedVectors, matrix_rank

# path of  dataset
dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\Country-data.csv'

principal_components_housing, rank_housing = perform_pca(dataset_path)

# first 5 principal components and their rank
print("First 5 Principal Components for the housing dataset:")
print(principal_components_housing[:, :5])
print("Matrix Rank for the housing dataset:", rank_housing)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS AZURE

```
Shape of the dataset before PCA: (167, 10)
First 5 Principal Components for the housing dataset:
[[-5.09077639e-03+0.00000000e+00j -3.93196540e-03+0.00000000e+00j
  -2.79611783e-03-6.88621257e-03j -2.79611783e-03+6.88621257e-03j
  -5.86583601e-03+0.00000000e+00j]
 [ 2.33696329e-02+0.00000000e+00j -5.55072178e-04+0.00000000e+00j
  -1.83273883e-02-1.04200908e-02j -1.83273883e-02+1.04200908e-02j
   1.41289706e-02+0.00000000e+00j]
 [ 3.57057863e-02+0.00000000e+00j  1.79815974e-03+0.00000000e+00j
```

DATAset#2

```
# path of  dataset
dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\bike-sharing.csv\bike-sharing.csv'

principal_components_housing, rank_bike = perform_pca(dataset_path)

# first 5 principal components and their rank
print("First 5 Principal Components for the bike sharing dataset:")
print(principal_components_housing[:, :5])
print("Matrix Rank for the bike sharing dataset:", rank_bike)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS AZURE

```
PS C:\Users\Mukand\Desktop\data set for pca> & "C:/Program Files/Python310/python.exe" "c:/Users/Mukand/Desktop/
Shape of the dataset before PCA: (17379, 17)
First 5 Principal Components for the bike sharing dataset:
[[0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]
 [0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]
 [0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]
 ...
 [0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]
 [0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]
 [0.+0.j 0.+0.j 0.+0.j 0.+0.j 0.+0.j]]
Matrix Rank for the bike sharing dataset: 15
PS C:\Users\Mukand\Desktop\data set for pca>
```

```python
# dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\bike-sharing.csv\bike-sharing.csv'
# dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\bike-sharing.csv\bike-sharing.csv'
dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\autos.csv'
# dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\Country-data.csv'
# dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\customer.csv\customer.csv'
# dataset_path = r'C:\Users\Mukand\Desktop\data set for pca\abalone.csv'

principal_components_housing, rank_bike = perform_pca(dataset_path)

# first 5 principal components and their rank
print("First 5 Principal Components for the dataset:")
print(principal_components_housing[:, :5])
print("Matrix Rank for dataset:", rank_bike)
```

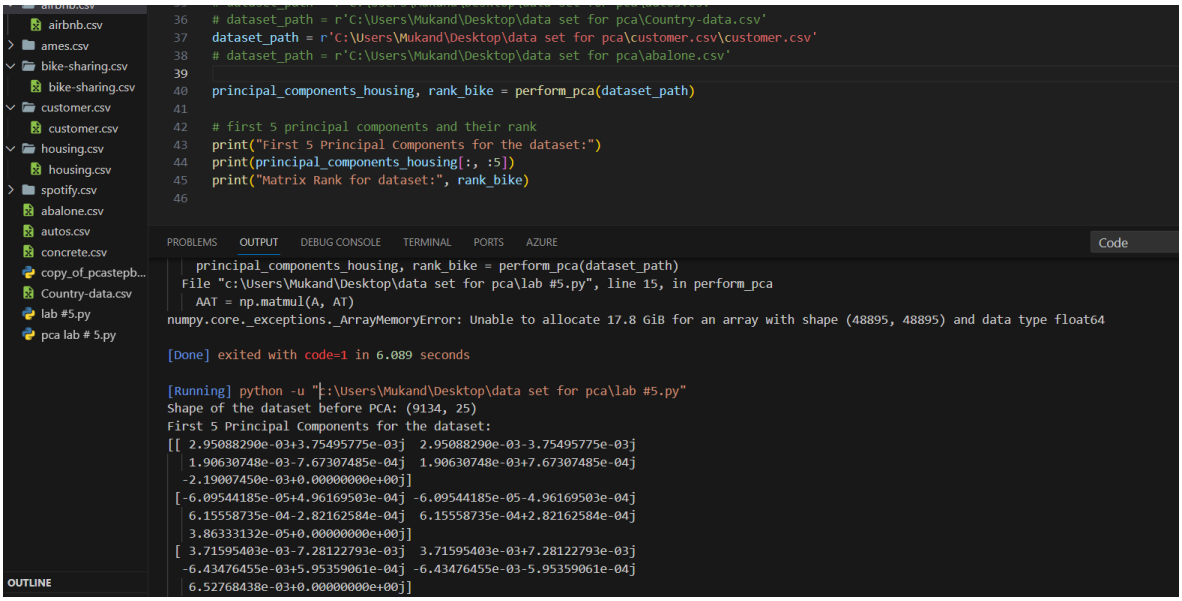
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS AZURE

```
[Done] exited with code=1 in 5.15 seconds

[Running] python -u "c:\Users\Mukand\Desktop\data set for pca\pca lab # 5.py"
Shape of the dataset before PCA: (167, 10)
First 5 Principal Components for the dataset:
[[-5.09077639e-03+0.00000000e+00j -3.93196540e-03+0.00000000e+00j
  -2.79611783e-03-6.88621257e-03j -2.79611783e-03+6.88621257e-03j
  -5.86583601e-03+0.00000000e+00j]
 [ 2.33696329e-02+0.00000000e+00j -5.55072178e-04+0.00000000e+00j
  -1.83273883e-02-1.04200908e-02j -1.83273883e-02+1.04200908e-02j
   1.41289706e-02+0.00000000e+00j]
 [ 3.57057863e-02+0.00000000e+00j  1.79815974e-03+0.00000000e+00j
   1.82111746e-02-4.78167975e-02j  1.82111746e-02+4.78167975e-02j
```

DataSet # 3

DATAset#4



DataSet #5

```
[Done] exited with code=1 in 1.011 seconds

[Running] python -u "c:\Users\Mukand\Desktop\7th sem 21st Aug -  2023\DS\lab 5\data set for pca\lab.py"
Shape of the dataset before PCA: (2930, 79)
First 5 Principal Components for the dataset:
[[ 0.0097941 +0.00000000e+00j -0.00383068-1.72529485e-04j
  -0.00383068+1.72529485e-04j  0.00172079+2.41449258e-03j
   0.00172079-2.41449258e-03j]
 [ 0.01704646+0.00000000e+00j -0.00020056+2.20056734e-05j
  -0.00020056-2.20056734e-05j -0.00034567+3.19054536e-04j
  -0.00034567-3.19054536e-04j]
 [ 0.01811726+0.00000000e+00j -0.01111475-2.05183456e-03j
  -0.01111475+2.05183456e-03j -0.01018182-6.61870288e-03j
  -0.01018182+6.61870288e-03j]
 ...
 [ 0.00598477+0.00000000e+00j -0.00706216+2.49251454e-03j
  -0.00706216-2.49251454e-03j  0.00313746+5.31285548e-03j
```

OUTLINE