A PYTHON PROGRAM TO IMPLEMENT DIMENSIONALITY REDUCTION USING PCA

Expt no. 10

Name: Jayasuriya.j

Roll no: 241801102

PROGRAM:

```
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
import seaborn as sns
import matplotlib.pyplot as plt

df =
pd.read_csv(r"C:\Users\Luqman\Downloads\IRIS
.csv")

first_col = df.columns[0] if
first_col.lower().startswith("unnamed") or
first_col.lower() in {"id", "#"}:
    df = df.iloc[:, 1:]

ncols = df.shape[1]

feature_names_4 = [
    'sepal length (cm)',
    'sepal width (cm)',
    'petal length (cm)',
```

```
    'petal width (cm)'
]

if ncols == 5:
    df.columns = feature_names_4 + ['species']
    X = df.iloc[:, :-1]
elif ncols >= 4:
    df = df.iloc[:, :4]
    df.columns = feature_names_4
    X = df
else:
    raise ValueError(f"Expected at least 4 columns,
but got {ncols}")

print("Original Data:")
display(X.head())

scaler = StandardScaler()
scaled_data =
pd.DataFrame(scaler.fit_transform(X))

print("Scaled Data:")
display(scaled_data.head())

plt.figure(figsize=(5,4))
sns.heatmap(scaled_data.corr(), annot=False,
cmap="coolwarm")
plt.show()

pca = PCA(n_components=3) data_pca
= pca.fit_transform(scaled_data)
```

```
data_pca = pd.DataFrame(data_pca,
columns=['PC1','PC2','PC3'])

print("PCA Data:")
display(data_pca.head())

plt.figure(figsize=(5,4))
sns.heatmap(data_pca.corr(), annot=False,
cmap="coolwarm")
plt.show()
```

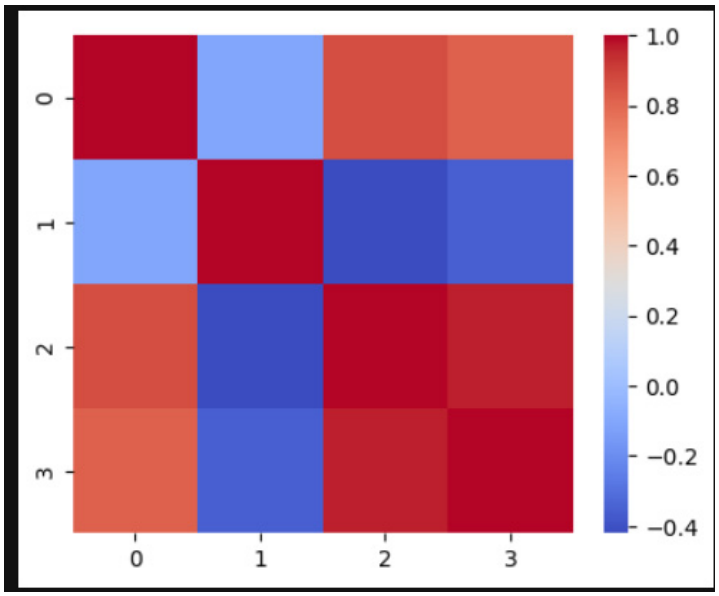
OUTPUT:

Original Data:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2

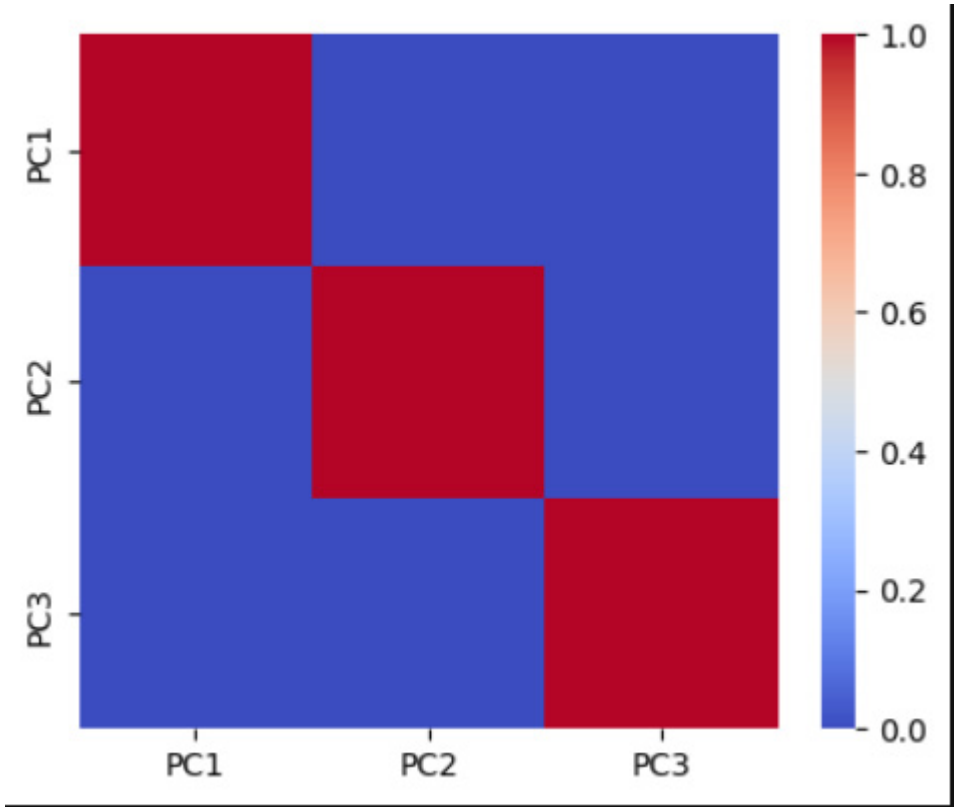
Scaled Data:

	0	1	2	3
0	-0.900681	1.032057	-1.341272	-1.312977
1	-1.143017	-0.124958	-1.341272	-1.312977
2	-1.385353	0.337848	-1.398138	-1.312977
3	-1.506521	0.106445	-1.284407	-1.312977
4	-1.021849	1.263460	-1.341272	-1.312977



PCA Data:

	PC1	PC2	PC3
0	-2.264542	0.505704	0.121943
1	-2.086426	-0.655405	0.227251
2	-2.367950	-0.318477	-0.051480
3	-2.304197	-0.575368	-0.098860
4	-2.388777	0.674767	-0.021428


PC1
PC2
PC3
PC1
PC2
PC3
1.0
0.8
0.6
0.4
0.2
0.0