

```
In [10]: 1 import pandas as pd
2 from sklearn.preprocessing import StandardScaler
3 from sklearn.decomposition import PCA
4
5 # Step 1: Load the dataset
6 file_path = 'adult.csv' # Replace with the correct file path
7 data = pd.read_csv(file_path)
8
9 # Step 2: Select numeric columns
10 numeric_cols = ['age', 'education.num', 'capital.gain', 'capital.loss',
11 numeric_data = data[numeric_cols]
12
13 # Step 3: Standardize the data
14 scaler = StandardScaler()
15 standardized_data = scaler.fit_transform(numeric_data)
16
17 # Step 4: Apply PCA
18 pca = PCA(n_components=2) # Set the number of components
19 pca_data = pca.fit_transform(standardized_data)
20
21 # Step 5: Explained variance
22 explained_variance = pca.explained_variance_ratio_
23
24 # Step 6: Output results
25 print("Explained Variance Ratio (Top 2 Components):", explained_variance)
26 print("\nFirst 5 rows of PCA-transformed data:\n", pca_data[:5])
27
```

Explained Variance Ratio (Top 2 Components): [0.25934221 0.20604498]

First 5 rows of PCA-transformed data:

```
[[3.9314469  8.8879336 ]
 [2.77223222 8.78015287]
 [3.51594341 8.7432766 ]
 [1.57171734 7.79145044]
 [2.5455066  7.68936563]]
```

In [9]:

1 data

Out[9]:

	age	workclass	fnlwgt	education	education.num	marital.status	occupation	relation
0	90	?	77053	HS-grad	9	Widowed	?	Not-in-fa
1	82	Private	132870	HS-grad	9	Widowed	Exec-managerial	Not-in-fa
2	66	?	186061	Some-college	10	Widowed	?	Unma
3	54	Private	140359	7th-8th	4	Divorced	Machine-op-inspct	Unma
4	41	Private	264663	Some-college	10	Separated	Prof-specialty	Own-
...
32556	22	Private	310152	Some-college	10	Never-married	Protective-serv	Not-in-fa
32557	27	Private	257302	Assoc-acdm	12	Married-civ-spouse	Tech-support	
32558	40	Private	154374	HS-grad	9	Married-civ-spouse	Machine-op-inspct	Husl
32559	58	Private	151910	HS-grad	9	Widowed	Adm-clerical	Unma
32560	22	Private	201490	HS-grad	9	Never-married	Adm-clerical	Own-

32561 rows × 15 columns

In []:

1