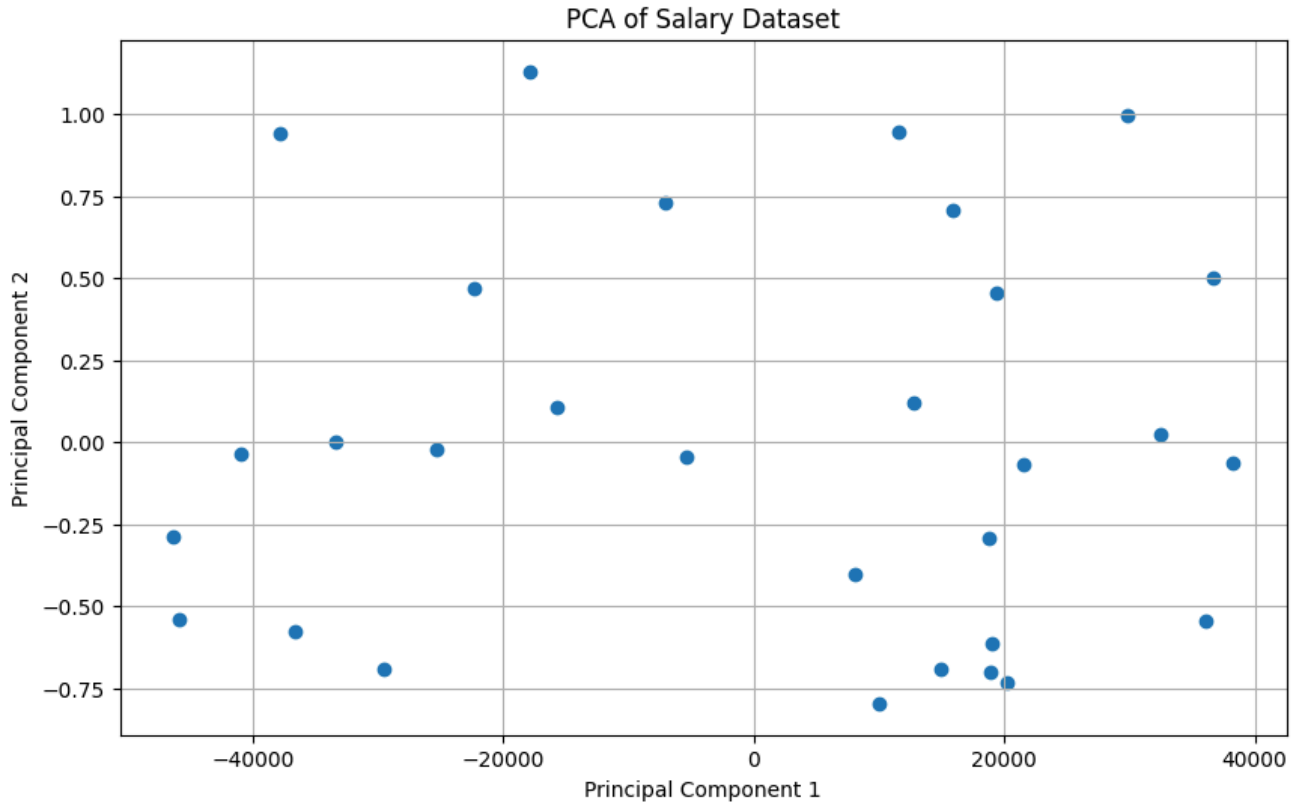
- Name: Prerna Sunil Jadhav
- Sap ID: 60004220127
- Batch: C22
- Branch: Computer Engineering
- Course: Machine Learning
- Experiment 7: PCA

```
import numpy as np

data = np.array([[4, 6],
                 [8, 2],
                 [13, 3],
                 [7, 15]])

mean = np.mean(data, axis=0)
centered_data = data - mean

print("Centered Data:")
print(centered_data)

covariance_matrix = np.cov(centered_data, rowvar=False)

print("\nCovariance Matrix:")
print(covariance_matrix)

eigen_values, eigen_vectors = np.linalg.eig(covariance_matrix)

print("\nEigenvalues:")
print(eigen_values)

print("\nEigenvectors:")
print(eigen_vectors)

new_values = np.dot(centered_data, eigen_vectors)

print("\nNew Values:")
print(new_values)
```

```
Centered Data:
[[-4.  -0.5]
 [ 0.  -4.5]
 [ 5.  -3.5]
 [-1.   8.5]]

Covariance Matrix:
[[14. -8.]
 [-8. 35.]]

Eigenvalues:
[11.29962122 37.70037878]
```

```
Eigenvectors:
[[-0.94747869  0.31981892]
 [-0.31981892 -0.94747869]]

New Values:
[[ 3.9498242  -0.80553633]
 [ 1.43918513  4.26365409]
 [-3.61802722  4.91526999]
 [-1.77098211 -8.37338775]]
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_csv("../content/salary_data.csv")

centered_data = df - df.mean()

covariance_matrix = np.cov(centered_data.values.T)

eigen_values, eigen_vectors = np.linalg.eig(covariance_matrix)

sorted_indices = np.argsort(eigen_values)[::-1]
eigen_values = eigen_values[sorted_indices]
eigen_vectors = eigen_vectors[:, sorted_indices]

new_values = np.dot(centered_data.values, eigen_vectors)

plt.figure(figsize=(10, 6))
plt.scatter(new_values[:, 0], new_values[:, 1])
plt.title('PCA of Salary Dataset')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.grid(True)
plt.show()

print("Centered Data:")
print(centered_data)
print("\nCovariance Matrix:")
print(covariance_matrix)
print("\nEigenvalues:")
print(eigen_values)
print("\nEigenvectors:")
print(eigen_vectors)
print("\nNew Values (Transformed Data):")
print(new_values)
```



Centered Data:

	YearsExperience	Salary
0	-4.213333	-36660.0
1	-4.013333	-29798.0
2	-3.813333	-38272.0
3	-3.313333	-32478.0
4	-3.113333	-36112.0
5	-2.413333	-19361.0
6	-2.313333	-15853.0
7	-2.113333	-21558.0
8	-2.113333	-11558.0
9	-1.613333	-18814.0
10	-1.413333	-12785.0
11	-1.313333	-20209.0
12	-1.313333	-19046.0
13	-1.213333	-18922.0
14	-0.813333	-14892.0
15	-0.413333	-8065.0
16	-0.213333	-9974.0
17	-0.013333	7085.0
18	0.586667	5360.0
19	0.686667	17937.0
20	1.486667	15735.0
21	1.786667	22270.0
22	2.586667	25299.0
23	2.886667	37809.0
24	3.386667	33428.0
25	3.686667	29579.0
26	4.186667	40966.0
27	4.286667	36632.0
28	4.986667	46388.0
29	5.186667	45869.0

Covariance Matrix:

```
[[8.05360920e+00 7.61063034e+04]
 [7.61063034e+04 7.51550960e+08]]
```

```
Eigenvalues:
[7.51550968e+08 3.46654177e-01]

Eigenvectors:
[[-1.01265659e-04 -9.99999995e-01]
 [-9.99999995e-01  1.01265659e-04]]

New Values (Transformed Data):
[[ 3.66600002e+04  5.00934262e-01]
 [ 2.97980003e+04  9.95819213e-01]
 [ 3.82720002e+04 -6.23059780e-02]
 [ 3.24780002e+04  2.44272513e-02]
 [ 3.61120001e+04 -5.43572152e-01]
 [ 1.93610001e+04  4.52728902e-01]
 [ 1.58530002e+04  7.07968833e-01]
 [ 2.15580001e+04 -6.97517489e-02]
 [ 1.15580002e+04  9.42904839e-01]
 [ 1.88140001e+04 -2.91878779e-01]
 [ 1.27850001e+04  1.18651879e-01]
 [ 2.02090000e+04 -7.33144371e-01]
 [ 1.90460000e+04 -6.15372410e-01]
 [ 1.89220000e+04 -7.02815468e-01]
 [ 1.48920000e+04 -6.94714861e-01]
 [ 8.06500000e+03 -4.03374207e-01]
 [ 9.97399997e+03 -7.96690348e-01]
 [-7.08499996e+03  7.30800526e-01]
 [-5.36000003e+03 -4.38827327e-02]
 [-1.79370000e+04  1.12973546e+00]
 [-1.57350001e+04  1.06748481e-01]
 [-2.22700001e+04  4.68519563e-01]
 [-2.52990001e+04 -2.47467526e-02]
 [-3.78090001e+04  9.42086640e-01]
 [-3.34280002e+04 -1.55820849e-03]
 [-2.95790002e+04 -6.91329727e-01]
 [-4.09660002e+04 -3.82176687e-02]
 [-3.66320002e+04 -5.77103033e-01]
 [-4.63880003e+04 -2.89155263e-01]
 [-4.58690003e+04 -5.41712139e-01]]
```