

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
```

```
df = pd.read_csv(r'C:\ardhra\PCA\wine.data.csv')
df.head(10)
```

	Class	Alcohol	Malic acid	Ash	Alcalinity of ash	Magnesium \
0	1	14.23	1.71	2.43	15.6	127
1	1	13.20	1.78	2.14	11.2	100
2	1	13.16	2.36	2.67	18.6	101
3	1	14.37	1.95	2.50	16.8	113
4	1	13.24	2.59	2.87	21.0	118
5	1	14.20	1.76	2.45	15.2	112
6	1	14.39	1.87	2.45	14.6	96
7	1	14.06	2.15	2.61	17.6	121
8	1	14.83	1.64	2.17	14.0	97
9	1	13.86	1.35	2.27	16.0	98

	Total phenols	Flavanoids	Nonflavanoid phenols	Proanthocyanins \
0	2.80	3.06	0.28	2.29
1	2.65	2.76	0.26	1.28
2	2.80	3.24	0.30	2.81
3	3.85	3.49	0.24	2.18
4	2.80	2.69	0.39	1.82
5	3.27	3.39	0.34	1.97
6	2.50	2.52	0.30	1.98
7	2.60	2.51	0.31	1.25
8	2.80	2.98	0.29	1.98
9	2.98	3.15	0.22	1.85

	Color intensity	Hue	OD280/OD315 of diluted wines	Proline
0	5.64	1.04	3.92	1065
1	4.38	1.05	3.40	1050
2	5.68	1.03	3.17	1185
3	7.80	0.86	3.45	1480
4	4.32	1.04	2.93	735
5	6.75	1.05	2.85	1450
6	5.25	1.02	3.58	1290
7	5.05	1.06	3.58	1295
8	5.20	1.08	2.85	1045
9	7.22	1.01	3.55	1045

```
X = df.drop('Class',axis=1)
y = df['Class']
```

```

from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()

X = df.drop('Class',axis=1)
y = df['Class']

X = scaler.fit_transform(X)

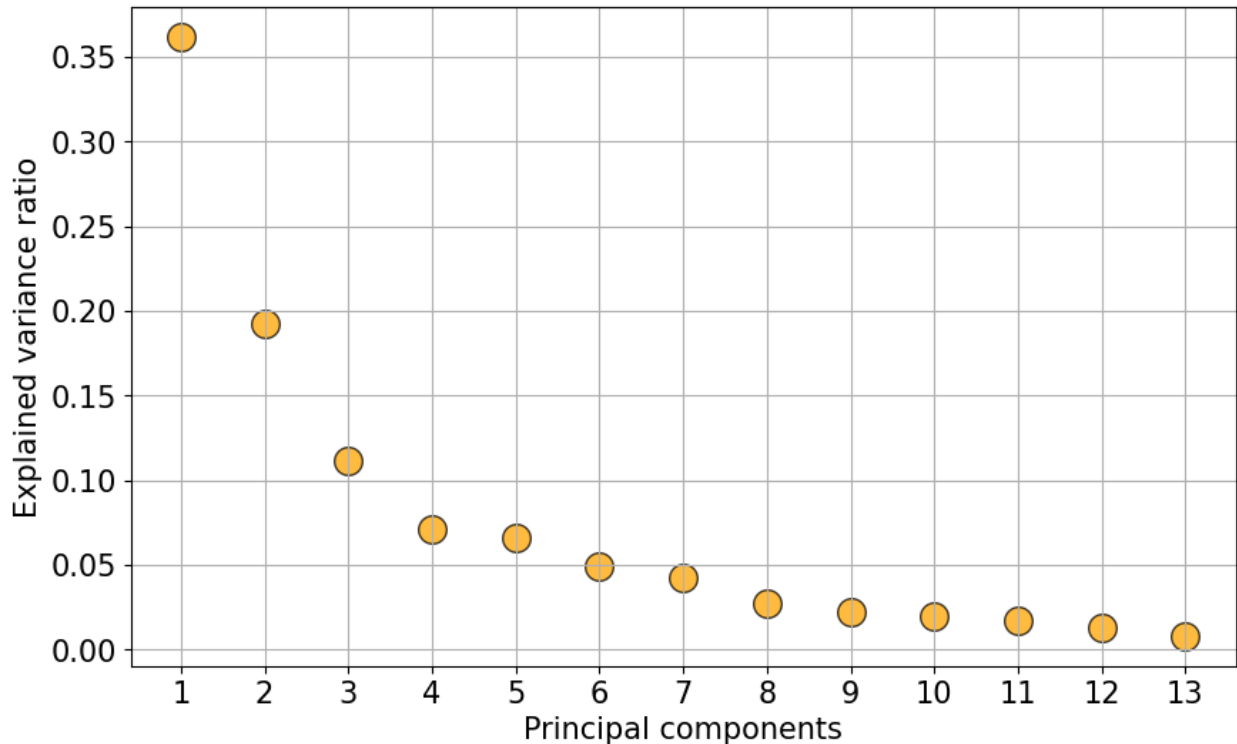
dfx = pd.DataFrame(data=X,columns=df.columns[1:])

from sklearn.decomposition import PCA
pca = PCA(n_components=None)
dfx_pca = pca.fit(dfx)

plt.figure(figsize=(10,6))
plt.scatter(x=[i+1 for i in
range(len(dfx_pca.explained_variance_ratio_))],
            y=dfx_pca.explained_variance_ratio_,
            s=200, alpha=0.75,c='orange',edgecolor='k')
plt.grid(True)
plt.title("Explained variance ratio of the \nfitted principal
component vector\n",fontsize=25)
plt.xlabel("Principal components",fontsize=15)
plt.xticks([i+1 for i in
range(len(dfx_pca.explained_variance_ratio_))],fontsize=15)
plt.yticks(fontsize=15)
plt.ylabel("Explained variance ratio",fontsize=15)
plt.show()

```

Explained variance ratio of the fitted principal component vector



```
dfx_trans = pca.transform(dfx)
dfx_trans = pd.DataFrame(data=dfx_trans)

plt.figure(figsize=(10,6))
plt.scatter(dfx_trans[0],dfx_trans[1],c=df['Class'],edgecolors='k',alpha=0.75,s=150)
plt.grid(True)
plt.title("Class separation using first two principal components\n",fontsize=20)
plt.xlabel("Principal component-1",fontsize=15)
plt.ylabel("Principal component-2",fontsize=15)
plt.show()
```

Class separation using first two principal components

