

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [2]: # Here we are using inbuilt dataset of scikit learn
from sklearn.datasets import load_breast_cancer

# instantiating
cancer = load_breast_cancer()

# creating dataframe
df = pd.DataFrame(cancer['data'], columns = cancer['feature_names'])

# checking head of dataframe
df.head()
```

```
Out[2]:
```

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	dim
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0

5 rows × 30 columns

```
In [3]: # Importing standard scalar module
from sklearn.preprocessing import StandardScaler

scalar = StandardScaler()

# fitting
scalar.fit(df)
scaled_data = scalar.transform(df)

# Importing PCA
from sklearn.decomposition import PCA

# Let's say, components = 2
pca = PCA(n_components = 2)
pca.fit(scaled_data)
x_pca = pca.transform(scaled_data)

x_pca.shape
```

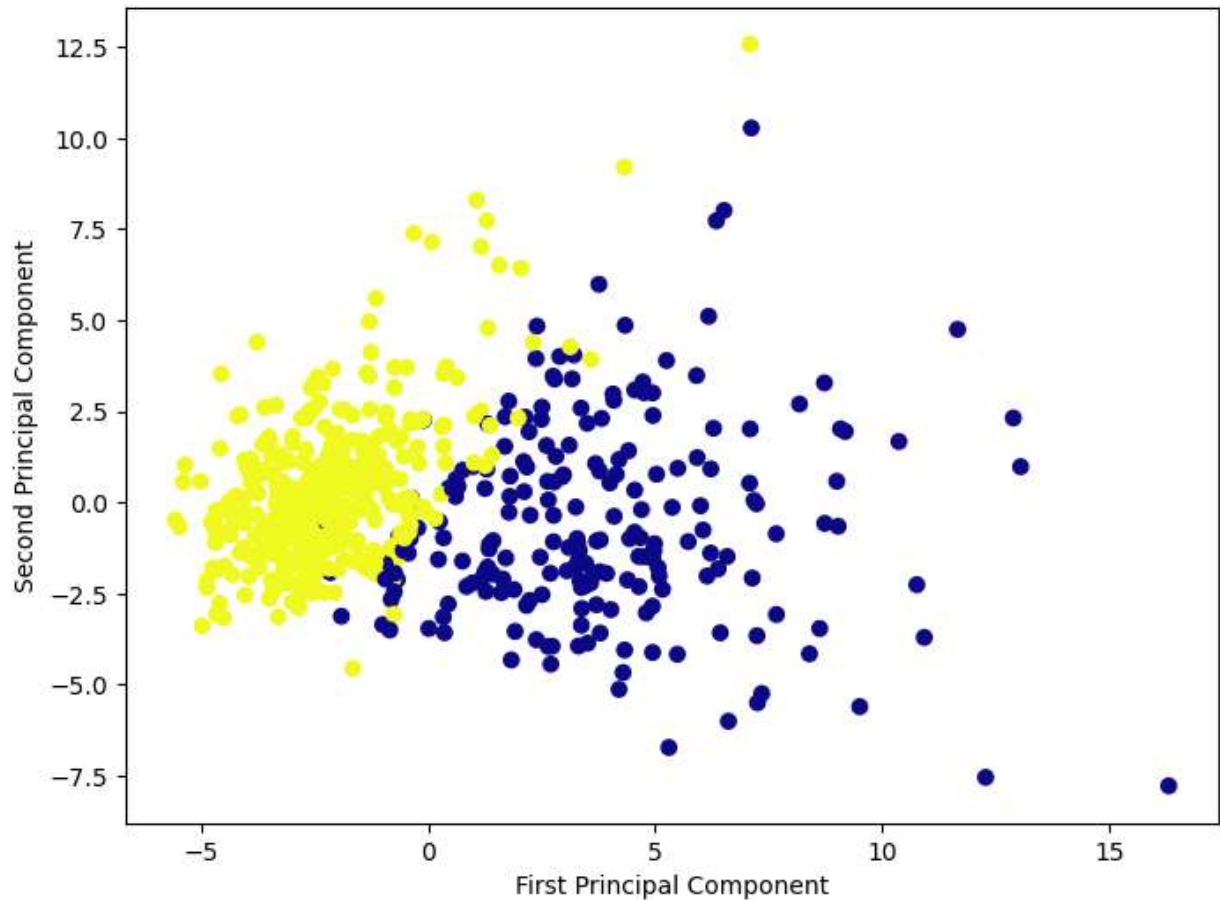
```
Out[3]: (569, 2)
```

```
In [4]: # giving a larger plot
plt.figure(figsize =(8, 6))

plt.scatter(x_pca[:, 0], x_pca[:, 1], c = cancer['target'], cmap ='plasma')

# Labeling x and y axes
plt.xlabel('First Principal Component')
plt.ylabel('Second Principal Component')
```

```
Out[4]: Text(0, 0.5, 'Second Principal Component')
```



```
In [5]: # components
pca.components_
```

```
Out[5]: array([[ 0.21890244,  0.10372458,  0.22753729,  0.22099499,  0.14258969,
  0.23928535,  0.25840048,  0.26085376,  0.13816696,  0.06436335,
  0.20597878,  0.01742803,  0.21132592,  0.20286964,  0.01453145,
  0.17039345,  0.15358979,  0.1834174 ,  0.04249842,  0.10256832,
  0.22799663,  0.10446933,  0.23663968,  0.22487053,  0.12795256,
  0.21009588,  0.22876753,  0.25088597,  0.12290456,  0.13178394],
 [-0.23385713, -0.05970609, -0.21518136, -0.23107671,  0.18611302,
  0.15189161,  0.06016536, -0.0347675 ,  0.19034877,  0.36657547,
 -0.10555215,  0.08997968, -0.08945723, -0.15229263,  0.20443045,
  0.2327159 ,  0.19720728,  0.13032156,  0.183848 ,  0.28009203,
 -0.21986638, -0.0454673 , -0.19987843, -0.21935186,  0.17230435,
  0.14359317,  0.09796411, -0.00825724,  0.14188335,  0.27533947]])
```

```
In [6]: df_comp = pd.DataFrame(pca.components_, columns = cancer['feature_names'])

plt.figure(figsize =(14, 6))
```

```
# plotting heatmap  
sns.heatmap(df_comp)
```

Out[6]: <AxesSubplot:>

