In [5]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns

from sklearn.datasets import load_breast_cancer
data = load_breast_cancer()
data.keys()
```

Out[5]:
```
dict_keys(['data', 'target', 'frame', 'target_names', 'DESCR', 'fe
ature_names', 'filename', 'data_module'])
```

In [6]:
```
print(data['feature_names'])
```

```
['mean radius' 'mean texture' 'mean perimeter' 'mean area'
 'mean smoothness' 'mean compactness' 'mean concavity'
 'mean concave points' 'mean symmetry' 'mean fractal dimension'
 'radius error' 'texture error' 'perimeter error' 'area error'
 'smoothness error' 'compactness error' 'concavity error'
 'concave points error' 'symmetry error' 'fractal dimension error'
 'worst radius' 'worst texture' 'worst perimeter' 'worst area'
 'worst smoothness' 'worst compactness' 'worst concavity'
 'worst concave points' 'worst symmetry' 'worst fractal dimension'
]
```

In [8]:
```
df = pd.DataFrame(data['data'],columns = data['feature_names'])
print(df)
```

```
     mean radius  mean texture  mean perimeter  mean area  mean sm
oothness  \
0          17.99         10.38          122.80     1001.0
0.11840
1          20.57         17.77          132.90     1326.0
0.08474
2          19.69         21.25          130.00     1203.0
0.10960
3          11.42         20.38           77.58      386.1
0.14250
4          20.29         14.34          135.10     1297.0
0.10030
..           ...           ...             ...        ...
...
564        21.56         22.39          142.00     1479.0
0.11100
565        20.13         28.25          131.20     1261.0
0.09780
566        16.60         28.08          108.30      858.1
0.08455
567        20.60         29.33          140.10     1265.0
0.11780
568         7.76         24.54           47.92      181.0
```

```
0.05263

     mean compactness  mean concavity  mean concave points  mean s
ymmetry  \
0             0.27760         0.30010              0.14710
0.2419
1             0.07864         0.08690              0.07017
0.1812
2             0.15990         0.19740              0.12790
0.2069
3             0.28390         0.24140              0.10520
0.2597
4             0.13280         0.19800              0.10430
0.1809
..                ...             ...                  ...
...
564           0.11590         0.24390              0.13890
0.1726
565           0.10340         0.14400              0.09791
0.1752
566           0.10230         0.09251              0.05302
0.1590
567           0.27700         0.35140              0.15200
0.2397
568           0.04362         0.00000              0.00000
0.1587

     mean fractal dimension  ...  worst radius  worst texture  \
0                   0.07871  ...        25.380          17.33
1                   0.05667  ...        24.990          23.41
2                   0.05999  ...        23.570          25.53
3                   0.09744  ...        14.910          26.50
4                   0.05883  ...        22.540          16.67
..                      ...  ...           ...            ...
564                 0.05623  ...        25.450          26.40
565                 0.05533  ...        23.690          38.25
566                 0.05648  ...        18.980          34.12
567                 0.07016  ...        25.740          39.42
568                 0.05884  ...         9.456          30.37

     worst perimeter  worst area  worst smoothness  worst compactn
ess  \
0             184.60      2019.0           0.16220            0.66
560
1             158.80      1956.0           0.12380            0.18
660
2             152.50      1709.0           0.14440            0.42
450
3              98.87       567.7           0.20980            0.86
630
4             152.20      1575.0           0.13740            0.20
500
..               ...         ...               ...
```

```
...
564          166.10      2027.0          0.14100           0.21
130
565          155.00      1731.0          0.11660           0.19
220
566          126.70      1124.0          0.11390           0.30
940
567          184.60      1821.0          0.16500           0.86
810
568           59.16       268.6          0.08996           0.06
444

     worst concavity  worst concave points  worst symmetry  \
0             0.7119                0.2654          0.4601
1             0.2416                0.1860          0.2750
2             0.4504                0.2430          0.3613
3             0.6869                0.2575          0.6638
4             0.4000                0.1625          0.2364
..               ...                   ...             ...
564           0.4107                0.2216          0.2060
565           0.3215                0.1628          0.2572
566           0.3403                0.1418          0.2218
567           0.9387                0.2650          0.4087
568           0.0000                0.0000          0.2871

     worst fractal dimension
0                    0.11890
1                    0.08902
2                    0.08758
3                    0.17300
4                    0.07678
..                       ...
564                  0.07115
565                  0.06637
566                  0.07820
567                  0.12400
568                  0.07039

[569 rows x 30 columns]
```

In [10]:
```
print(df.head())
```

```
   mean radius  mean texture  mean perimeter  mean area  mean smoo
thness  \
0        17.99         10.38          122.80     1001.0          0
.11840
1        20.57         17.77          132.90     1326.0          0
.08474
2        19.69         21.25          130.00     1203.0          0
.10960
3        11.42         20.38           77.58      386.1          0
.14250
4        20.29         14.34          135.10     1297.0          0
.10030
```

```
.10030

   mean compactness  mean concavity  mean concave points  mean sym
metry  \
0           0.27760          0.3001              0.14710         0
.2419
1           0.07864          0.0869              0.07017         0
.1812
2           0.15990          0.1974              0.12790         0
.2069
3           0.28390          0.2414              0.10520         0
.2597
4           0.13280          0.1980              0.10430         0
.1809

   mean fractal dimension  ...  worst radius  worst texture  worst
perimeter  \
0                 0.07871  ...         25.38          17.33
184.60
1                 0.05667  ...         24.99          23.41
158.80
2                 0.05999  ...         23.57          25.53
152.50
3                 0.09744  ...         14.91          26.50
98.87
4                 0.05883  ...         22.54          16.67
152.20

   worst area  worst smoothness  worst compactness  worst concavit
y  \
0      2019.0            0.1622             0.6656           0.711
9
1      1956.0            0.1238             0.1866           0.241
6
2      1709.0            0.1444             0.4245           0.450
4
3       567.7            0.2098             0.8663           0.686
9
4      1575.0            0.1374             0.2050           0.400
0

   worst concave points  worst symmetry  worst fractal dimension
0                0.2654          0.4601                  0.11890
1                0.1860          0.2750                  0.08902
2                0.2430          0.3613                  0.08758
3                0.2575          0.6638                  0.17300
4                0.1625          0.2364                  0.07678

[5 rows x 30 columns]
```

In [15]:

```
from sklearn.preprocessing import StandardScaler

scalar = StandardScaler()
scalar.fit(df)
scaledData = scalar.transform(df)
print(df.head())
print(scaledData)
```

```
   mean radius  mean texture  mean perimeter  mean area  mean smoo
thness  \
0        17.99         10.38          122.80     1001.0          0
.11840
1        20.57         17.77          132.90     1326.0          0
.08474
2        19.69         21.25          130.00     1203.0          0
.10960
3        11.42         20.38           77.58      386.1          0
.14250
4        20.29         14.34          135.10     1297.0          0
.10030

   mean compactness  mean concavity  mean concave points  mean sym
metry  \
0           0.27760          0.3001              0.14710         0
.2419
1           0.07864          0.0869              0.07017         0
.1812
2           0.15990          0.1974              0.12790         0
.2069
3           0.28390          0.2414              0.10520         0
.2597
4           0.13280          0.1980              0.10430         0
.1809

   mean fractal dimension  ...  worst radius  worst texture  worst
perimeter  \
0                 0.07871  ...         25.38          17.33
184.60
1                 0.05667  ...         24.99          23.41
158.80
2                 0.05999  ...         23.57          25.53
152.50
3                 0.09744  ...         14.91          26.50
98.87
4                 0.05883  ...         22.54          16.67
152.20

   worst area  worst smoothness  worst compactness  worst concavit
y  \
0      2019.0            0.1622             0.6656           0.711
9
1      1956.0            0.1238             0.1866           0.241
6
```

```
2      1709.0           0.1444            0.4245           0.450
4
3       567.7           0.2098            0.8663           0.686
9
4      1575.0           0.1374            0.2050           0.400
0

   worst concave points  worst symmetry  worst fractal dimension
0                0.2654          0.4601                  0.11890
1                0.1860          0.2750                  0.08902
2                0.2430          0.3613                  0.08758
3                0.2575          0.6638                  0.17300
4                0.1625          0.2364                  0.07678

[5 rows x 30 columns]
[[ 1.09706398 -2.07333501  1.26993369 ...  2.29607613  2.75062224
   1.93701461]
 [ 1.82982061 -0.35363241  1.68595471 ...  1.0870843  -0.24388967
   0.28118999]
 [ 1.57988811  0.45618695  1.56650313 ...  1.95500035  1.152255
   0.20139121]
 ...
 [ 0.70228425  2.0455738   0.67267578 ...  0.41406869 -1.10454895
  -0.31840916]
 [ 1.83834103  2.33645719  1.98252415 ...  2.28998549  1.91908301
   2.21963528]
 [-1.80840125  1.22179204 -1.81438851 ... -1.74506282 -0.04813821
  -0.75120669]]
```

In [16]:
```
print(scaledData.shape)
```

```
(569, 30)
```

In [19]:
```
from sklearn.decomposition import PCA

pca = PCA(n_components=2)
pca.fit(scaledData)
```

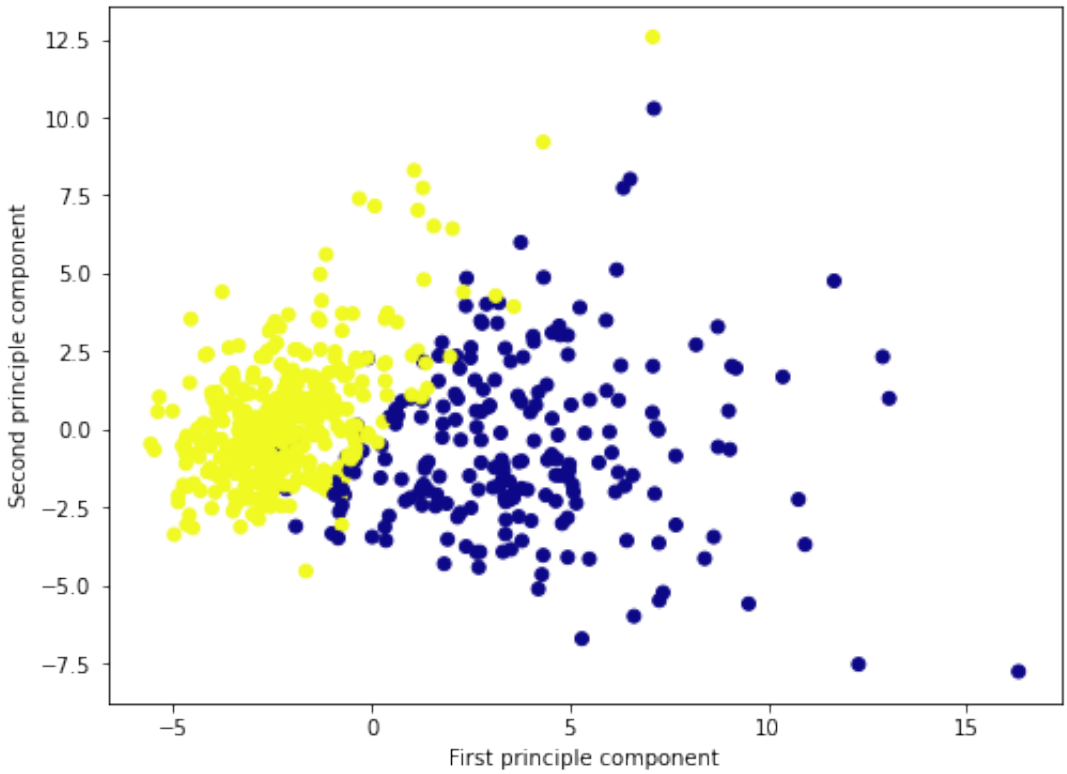
Out[19]:
```
PCA(n_components=2)
```

In [21]:
```
x_pca = pca.transform(scaledData)
print(x_pca)
```

```
[[ 9.19283683  1.94858307]
 [ 2.3878018  -3.76817174]
 [ 5.73389628 -1.0751738 ]
 ...
 [ 1.25617928 -1.90229671]
 [10.37479406  1.67201011]
 [-5.4752433  -0.67063679]]
```

In [22]:
```
plt.figure(figsize=(8,6))
plt.scatter(x_pca[:,0],x_pca[:,1],c = data['target'], cmap='plasma'
plt.xlabel('First principle component')
plt.ylabel('Second principle component')
```

Out[22]: Text(0, 0.5, 'Second principle component')




In [ ]: