In [ ]:

```python
# Import required libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix
from matplotlib.colors import ListedColormap
%matplotlib inline
# Load dataset
# Replace 'your_dataset.csv' with your actual dataset file
from google.colab import files
uploaded = files.upload()
```

Choose Files No file chosen

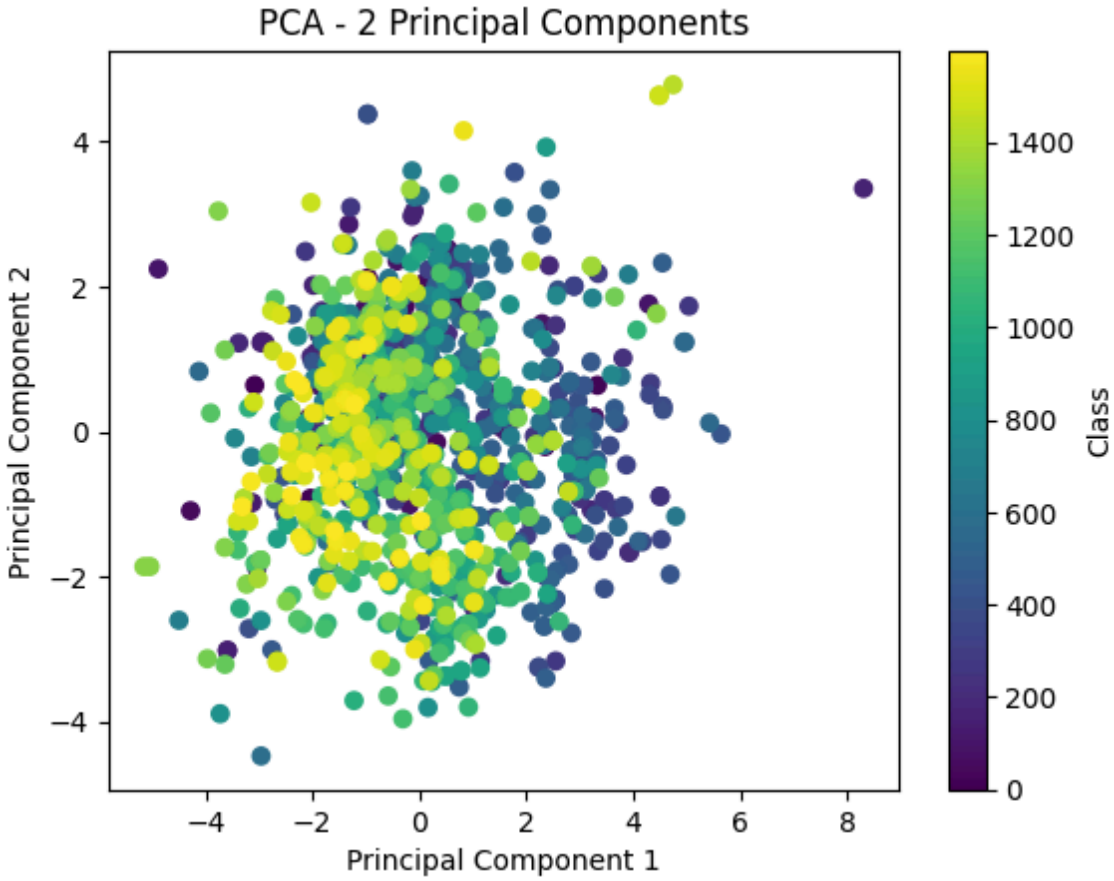
Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving WineQT.csv to WineQT (2).csv
```

In [4]:

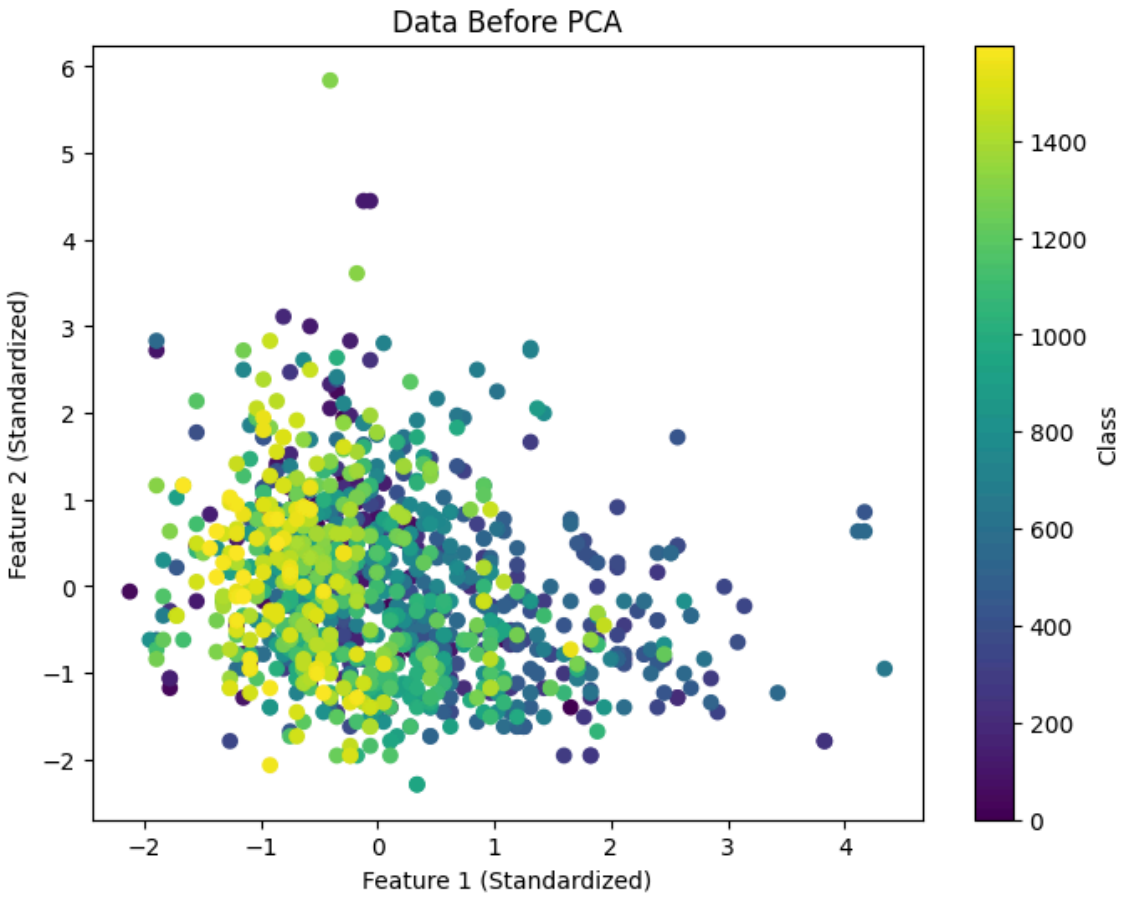
```python
# Import required libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler

# Load dataset
# Replace 'your_dataset.csv' with your actual dataset file
df = pd.read_csv('WineQT.csv')

# Assuming the features are in all columns except the last, which
contains the target variable
X = df.iloc[:, :-1].values
y = df.iloc[:, -1].values

# Standardize the data
sc = StandardScaler()
X_scaled = sc.fit_transform(X)

# Apply PCA to reduce to 2 principal components
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_scaled)

# Print the explained variance ratio
print("Explained Variance Ratio:", pca.explained_variance_ratio_)

# Plot the first two principal components
plt.scatter(X_pca[:, 0], X_pca[:, 1], c=y, cmap='viridis')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.title('PCA - 2 Principal Components')
plt.colorbar(label='Class')
plt.show()
```

```
Explained Variance Ratio: [0.26480487 0.18621688]
```


PCA - 2 Principal Components
Principal Component 2
Principal Component 1
Class
4
2
0
−2
−4
−4
−2
0
2
4
6
8
1400
1200
1000
800
600
400
200
0

In [6]:

```python
# Import required libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler

# Load dataset
# Replace 'WineQT.csv' with your actual dataset file
df = pd.read_csv('WineQT.csv')

# Assuming the features are in all columns except the last, which
contains the target variable
X = df.iloc[:, :-1].values
y = df.iloc[:, -1].values

# Standardize the data
sc = StandardScaler()
X_scaled = sc.fit_transform(X)

# Plot the data before applying PCA
plt.figure(figsize=(8, 6))
plt.scatter(X_scaled[:, 0], X_scaled[:, 1], c=y, cmap='viridis')
plt.xlabel('Feature 1 (Standardized)')
plt.ylabel('Feature 2 (Standardized)')
plt.title('Data Before PCA')
plt.colorbar(label='Class')
plt.show()

# Apply PCA to reduce to 2 principal components
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_scaled)

# Print the explained variance ratio
print("Explained Variance Ratio:", pca.explained_variance_ratio_)

# Plot the data after applying PCA
plt.figure(figsize=(8, 6))
plt.scatter(X_pca[:, 0], X_pca[:, 1], c=y, cmap='viridis')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.title('Data After PCA (2 Principal Components)')
plt.colorbar(label='Class')
plt.show()
```




```
Explained Variance Ratio: [0.26480487 0.18621688]
```