Import Library

```
from sklearn import datasets
import matplotlib.pyplot as plt
```

Importing the dataset

```
dataset = datasets.load_iris()
```

Data Segregation

```
X = dataset.data
y = dataset.target
names = dataset.target_names
print(X)
print(y)
print(names)
```

```
    [5.1 2.5 3.  1.1]
    [5.7 2.8 4.1 1.3]
    [6.3 3.3 6.  2.5]
    [5.8 2.7 5.1 1.9]
    [7.1 3.  5.9 2.1]
    [6.3 2.9 5.6 1.8]
    [6.5 3.  5.8 2.2]
    [7.6 3.  6.6 2.1]
    [4.9 2.5 4.5 1.7]
    [7.3 2.9 6.3 1.8]
    [6.7 2.5 5.8 1.8]
    [7.2 3.6 6.1 2.5]
    [6.5 3.2 5.1 2. ]
    [6.4 2.7 5.3 1.9]
    [6.8 3.  5.5 2.1]
    [5.7 2.5 5.  2. ]
    [5.8 2.8 5.1 2.4]
    [6.4 3.2 5.3 2.3]
    [6.5 3.  5.5 1.8]
    [7.7 3.8 6.7 2.2]
    [7.7 2.6 6.9 2.3]
    [6.  2.2 5.  1.5]
    [6.9 3.2 5.7 2.3]
    [5.6 2.8 4.9 2. ]
    [7.7 2.8 6.7 2. ]
    [6.3 2.7 4.9 1.8]
    [6.7 3.3 5.7 2.1]
    [7.2 3.2 6.  1.8]
    [6.2 2.8 4.8 1.8]
    [6.1 3.  4.9 1.8]
    [6.4 2.8 5.6 2.1]
    [7.2 3.  5.8 1.6]
    [7.4 2.8 6.1 1.9]
    [7.9 3.8 6.4 2. ]
    [6.4 2.8 5.6 2.2]
    [6.3 2.8 5.1 1.5]
    [6.1 2.6 5.6 1.4]
    [7.7 3.  6.1 2.3]
    [6.3 3.4 5.6 2.4]
    [6.4 3.1 5.5 1.8]
    [6.  3.  4.8 1.8]
    [6.9 3.1 5.4 2.1]
    [6.7 3.1 5.6 2.4]
    [6.9 3.1 5.1 2.3]
    [5.8 2.7 5.1 1.9]
    [6.8 3.2 5.9 2.3]
    [6.7 3.3 5.7 2.5]
    [6.7 3.  5.2 2.3]
    [6.3 2.5 5.  1.9]
    [6.5 3.  5.2 2. ]
    [6.2 3.4 5.4 2.3]
    [5.9 3.  5.1 1.8]]
    [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
     0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
     1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2
     2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
     2 2]
    ['setosa' 'versicolor' 'virginica']
```
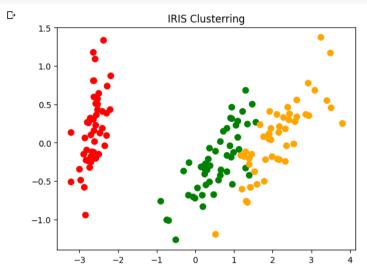
Finding the PCA clustering to the dataset with n=2

```
from sklearn.decomposition import PCA
model = PCA(n_components=2) # number of components to keep
y_means = model.fit(X).transform(X)
```

Variance Percentage

```
plt.figure()
colors = ['red', 'green', 'orange']

for color, i, target_name in zip(colors, [0, 1, 2], names):
  plt.scatter(y_means[y==i, 0], y_means[y==i, 1], color=color, lw=2, label=target_name)
plt.title('IRIS Clusterring')
plt.show()
```


IRIS Clusterring

✓ 0s   completed at 5:25 AM