Python (Data Visualization)

1. Data Import

We can import the Iris dataset from the Python package **scikit-learn**.

Detailed information about **scikit-learn** can be found at scikit-learn.org.

```
from sklearn import datasets
iris = datasets.load_iris()
```

What does the Iris dataset look like?

```
iris.feature_names
```

Result: ['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)', 'petal width (cm)']

To display the names of the target classes:

```
iris.target_names
```

Result: array(['setosa', 'versicolor', 'virginica'], dtype='<U10')

To display the attribute values of the records:

```
iris.data
```

Result: array([5.1, 3.5, 1.4, 0.2], [4.9, 3. , 1.4, 0.2], [4.7, 3.2, 1.3, 0.2], [4.6, 3.1, 1.5, 0.2], [5. , 3.6, 1.4, 0.2] ……)

To display the target outputs of the records:

```
iris.target
```

Result: array ([0, 0, 0,…,1, 1, 1 ,…,2, 2, 2,…])

The classes 'setosa', 'versicolor' and 'virginica' are denoted by 0, 1, and 2, respectively.

2. Data Visualization

In this section, we use the package **matplotlib** to visualize data.

Detailed information about **matplotlib** can be found at matplotlib.org.

Package setup for visualization:

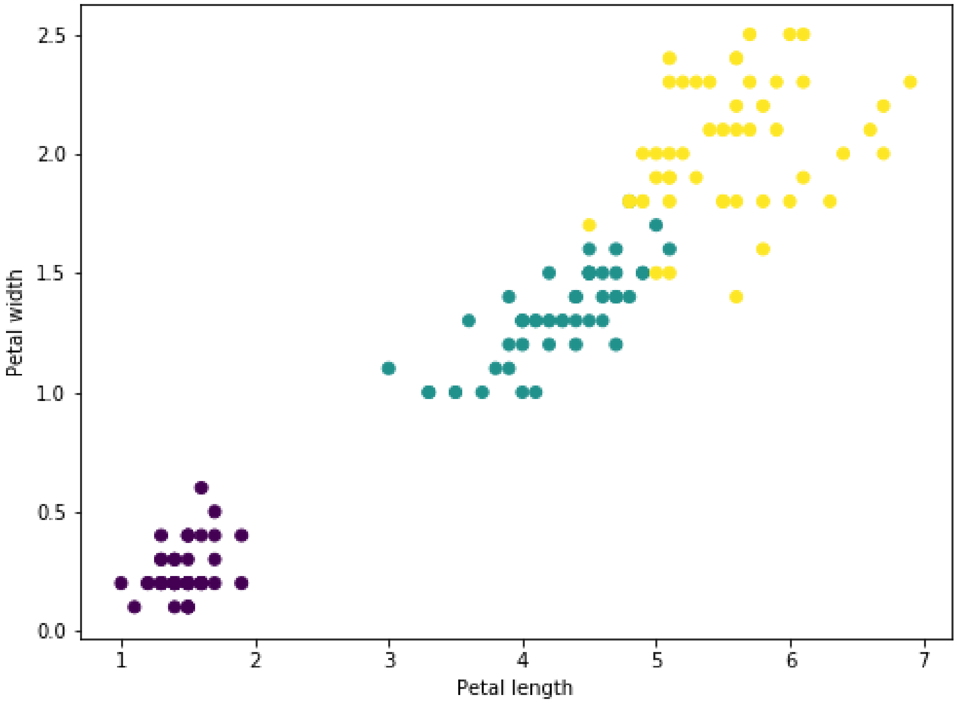
```
import matplotlib.pyplot as plt
```

We use a subset of attributes in the Iris dataset for visualization. First, we select the attributes "Petal length" and "Petal width" as follows.

```
X = iris.data[:, 2:4]
t = iris.target
```

We can now generate a scatter plot using the attribute values in X, and use the target outputs to distinguish the instances.

```
plt.scatter(X[:, 0], X[:, 1], c=t)
plt.xlabel('Petal length')
plt.ylabel('Petal width')
plt.show()
```

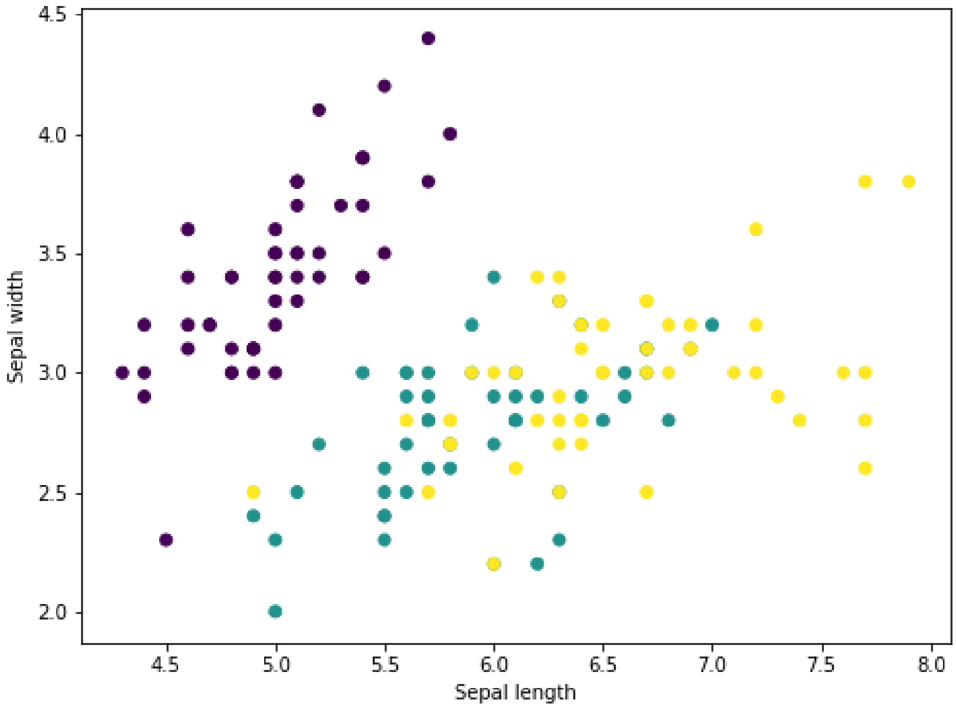


You can generate the scatter plot for other pairs of attributes. For example, the attribute pair (Sepal length, Sepal width) can be specified as follows:

```
X = iris.data[:, :2]
```

Accordingly, labels for the two axes should also be changed:

```
plt.scatter(X[:, 0], X[:, 1], c=t)
plt.xlabel('Sepal length')
plt.ylabel('Sepal width')
plt.show()
```


4.5
4.0
3.5
3.0
2.5
2.0
Sepal width
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
Sepal length