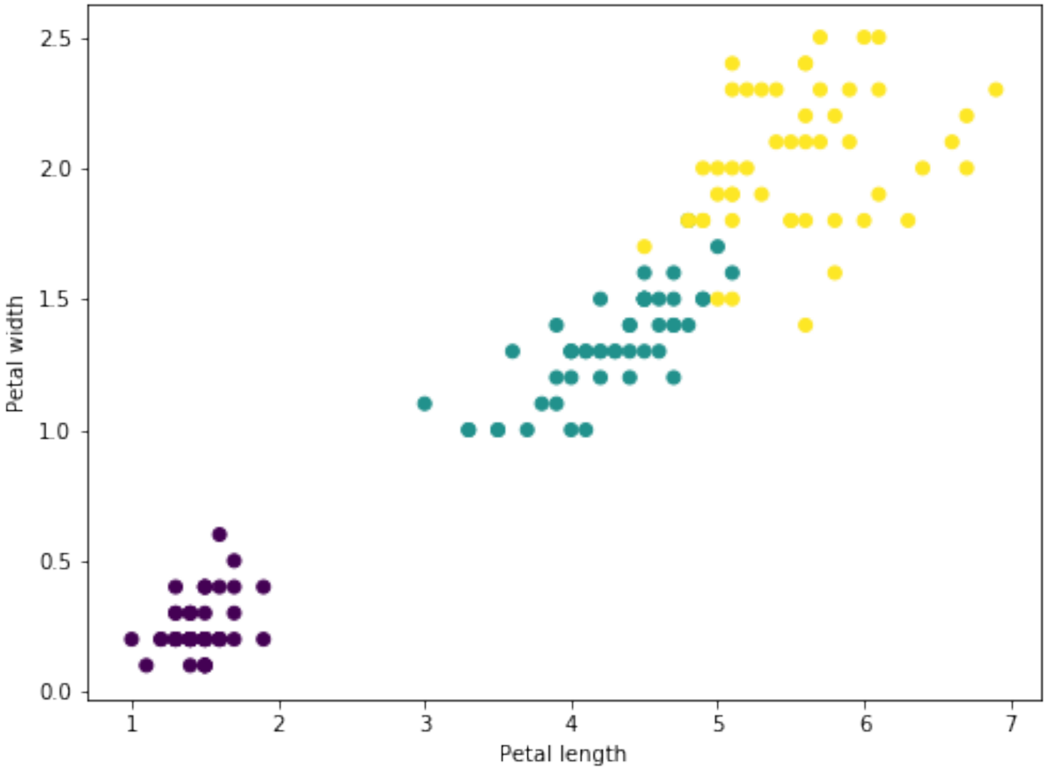
Playing arund with Iris

We will use Iris in class to practice some attribute transformations and computing similarities.

In [13]:

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

# import some data to play with
iris = datasets.load_iris()

X = iris.data[:, 2:4]  # we only take petal length and petal width.
Y = iris.target

plt.figure(2, figsize=(8, 6))
plt.clf()

# Plot the training points
plt.scatter(X[:, 0], X[:, 1], c=Y)
plt.xlabel('Petal length')
plt.ylabel('Petal width')

plt.show()
```



In [14]:

```
import numpy as np
A = iris.data
a = A[0,:]
b = A[-1,:]
print(a,b)
```

```
[5.1 3.5 1.4 0.2] [5.9 3.  5.1 1.8]
```

In [15]:

```
c = np.log(a)
```

In [16]:

```
d = np.abs(c)
print(d)
```

```
[1.62924054 1.25276297 0.33647224 1.60943791]
```

In [17]:

```
for i in range(A.shape[1]):
    print(np.min(A[:,i]), np.max(A[:,i]))
```

```
4.3 7.9
2.0 4.4
1.0 6.9
0.1 2.5
```

In [18]:

```
c = A[:,0]
c_mean = np.mean(c)
c_std = np.std(c)
d = (c-c_mean)/c_std
print(c_mean, c_std)
print(np.min(d), np.max(d), np.mean(d), np.std(d))
```

```
5.843333333333334 0.8253012917851409
-1.870024133847019 2.4920192021244283 -4.736951571734001e-16 1.0
```

In [ ]:

```

```