In [1]:

```
import sys
import scipy
import numpy
import matplotlib
import pandas
import sklearn
print('Python: {}'.format(sys.version))
print('scipy: {}'.format(scipy.__version__))
print('numpy: {}'.format(numpy.__version__))
print('matplotlib: {}'.format(matplotlib.__version__))
print('pandas: {}'.format(pandas.__version__))
print('sklearn: {}'.format(sklearn.__version__))
```

```
Python: 3.6.5 |Anaconda, Inc.| (default, Apr 26 2018, 08:42:37)
[GCC 4.2.1 Compatible Clang 4.0.1 (tags/RELEASE_401/final)]
scipy: 1.1.0
numpy: 1.14.3
matplotlib: 2.2.2
pandas: 0.23.0
sklearn: 0.19.1
```

In [2]:

```
import pandas
from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt
from sklearn import model_selection
from sklearn.neighbors import KNeighborsClassifier
names = ['sepal-length', 'sepal-width', 'petal-length', 'petal-width', 'class']
ds = pandas.read_csv('iris.data.txt', names=names)
```

In [3]:

```
array = ds.values
X = array[:,0:4]
Y = array[:,4]
print(X);
print(Y);
```

```
[[5.1 3.5 1.4 0.2]
 [4.9 3.0 1.4 0.2]
 [4.7 3.2 1.3 0.2]
 [4.6 3.1 1.5 0.2]
 [5.0 3.6 1.4 0.2]
 [5.4 3.9 1.7 0.4]
 [4.6 3.4 1.4 0.3]
 [5.0 3.4 1.5 0.2]
 [4.4 2.9 1.4 0.2]
 [4.9 3.1 1.5 0.1]
 [5.4 3.7 1.5 0.2]
 [4.8 3.4 1.6 0.2]
 [4.8 3.0 1.4 0.1]
 [4.3 3.0 1.1 0.1]
 [5.8 4.0 1.2 0.2]
 [5.7 4.4 1.5 0.4]
 [5.4 3.9 1.3 0.4]
 [5.1 3.5 1.4 0.3]
 [5.7 3.8 1.7 0.3]
 [5.1 3.8 1.5 0.3]
 [5.4 3.4 1.7 0.2]
 [5.1 3.7 1.5 0.4]
 [4.6 3.6 1.0 0.2]
 [5.1 3.3 1.7 0.5]
 [4.8 3.4 1.9 0.2]
 [5.0 3.0 1.6 0.2]
 [5.0 3.4 1.6 0.4]
 [5.2 3.5 1.5 0.2]
 [5.2 3.4 1.4 0.2]
 [4.7 3.2 1.6 0.2]
 [4.8 3.1 1.6 0.2]
 [5.4 3.4 1.5 0.4]
 [5.2 4.1 1.5 0.1]
 [5.5 4.2 1.4 0.2]
 [4.9 3.1 1.5 0.1]
 [5.0 3.2 1.2 0.2]
 [5.5 3.5 1.3 0.2]
 [4.9 3.1 1.5 0.1]
 [4.4 3.0 1.3 0.2]
 [5.1 3.4 1.5 0.2]
 [5.0 3.5 1.3 0.3]
 [4.5 2.3 1.3 0.3]
 [4.4 3.2 1.3 0.2]
 [5.0 3.5 1.6 0.6]
 [5.1 3.8 1.9 0.4]
 [4.8 3.0 1.4 0.3]
 [5.1 3.8 1.6 0.2]
 [4.6 3.2 1.4 0.2]
 [5.3 3.7 1.5 0.2]
 [5.0 3.3 1.4 0.2]
 [7.0 3.2 4.7 1.4]
 [6.4 3.2 4.5 1.5]
 [6.9 3.1 4.9 1.5]
 [5.5 2.3 4.0 1.3]
 [6.5 2.8 4.6 1.5]
 [5.7 2.8 4.5 1.3]
 [6.3 3.3 4.7 1.6]
 [4.9 2.4 3.3 1.0]
 [6.6 2.9 4.6 1.3]
 [5.2 2.7 3.9 1.4]
 [5.0 2.0 3.5 1.0]
```

```
[5.9 3.0 4.2 1.5]
[6.0 2.2 4.0 1.0]
[6.1 2.9 4.7 1.4]
[5.6 2.9 3.6 1.3]
[6.7 3.1 4.4 1.4]
[5.6 3.0 4.5 1.5]
[5.8 2.7 4.1 1.0]
[6.2 2.2 4.5 1.5]
[5.6 2.5 3.9 1.1]
[5.9 3.2 4.8 1.8]
[6.1 2.8 4.0 1.3]
[6.3 2.5 4.9 1.5]
[6.1 2.8 4.7 1.2]
[6.4 2.9 4.3 1.3]
[6.6 3.0 4.4 1.4]
[6.8 2.8 4.8 1.4]
[6.7 3.0 5.0 1.7]
[6.0 2.9 4.5 1.5]
[5.7 2.6 3.5 1.0]
[5.5 2.4 3.8 1.1]
[5.5 2.4 3.7 1.0]
[5.8 2.7 3.9 1.2]
[6.0 2.7 5.1 1.6]
[5.4 3.0 4.5 1.5]
[6.0 3.4 4.5 1.6]
[6.7 3.1 4.7 1.5]
[6.3 2.3 4.4 1.3]
[5.6 3.0 4.1 1.3]
[5.5 2.5 4.0 1.3]
[5.5 2.6 4.4 1.2]
[6.1 3.0 4.6 1.4]
[5.8 2.6 4.0 1.2]
[5.0 2.3 3.3 1.0]
[5.6 2.7 4.2 1.3]
[5.7 3.0 4.2 1.2]
[5.7 2.9 4.2 1.3]
[6.2 2.9 4.3 1.3]
[5.1 2.5 3.0 1.1]
[5.7 2.8 4.1 1.3]
[6.3 3.3 6.0 2.5]
[5.8 2.7 5.1 1.9]
[7.1 3.0 5.9 2.1]
[6.3 2.9 5.6 1.8]
[6.5 3.0 5.8 2.2]
[7.6 3.0 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 2.0]
[6.4 2.7 5.3 1.9]
[6.8 3.0 5.5 2.1]
[5.7 2.5 5.0 2.0]
[5.8 2.8 5.1 2.4]
[6.4 3.2 5.3 2.3]
[6.5 3.0 5.5 1.8]
[7.7 3.8 6.7 2.2]
[7.7 2.6 6.9 2.3]
[6.0 2.2 5.0 1.5]
[6.9 3.2 5.7 2.3]
[5.6 2.8 4.9 2.0]
```

```
 [7.7 2.8 6.7 2.0]
 [6.3 2.7 4.9 1.8]
 [6.7 3.3 5.7 2.1]
 [7.2 3.2 6.0 1.8]
 [6.2 2.8 4.8 1.8]
 [6.1 3.0 4.9 1.8]
 [6.4 2.8 5.6 2.1]
 [7.2 3.0 5.8 1.6]
 [7.4 2.8 6.1 1.9]
 [7.9 3.8 6.4 2.0]
 [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 3.0 6.1 2.3]
 [6.3 3.4 5.6 2.4]
 [6.4 3.1 5.5 1.8]
 [6.0 3.0 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 3.0 5.2 2.3]
 [6.3 2.5 5.0 1.9]
 [6.5 3.0 5.2 2.0]
 [6.2 3.4 5.4 2.3]
 [5.9 3.0 5.1 1.8]]
['Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-setos
a'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
```

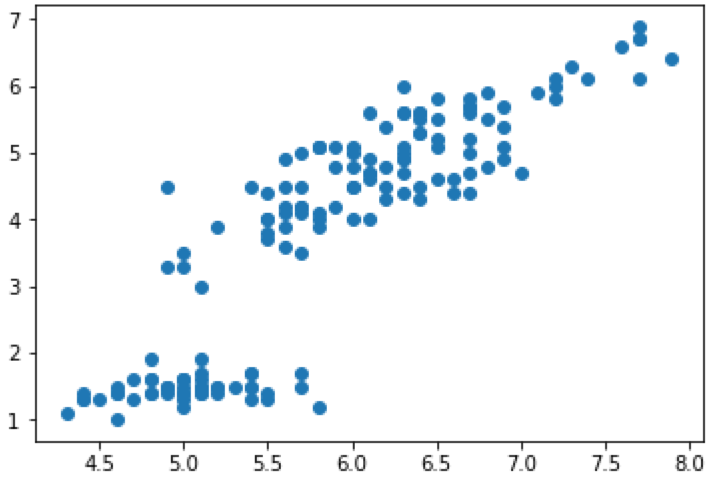
```
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versico
lor'
 'Iris-versicolor' 'Iris-versicolor' 'Iris-virginica' 'Iris-virginic
a'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'
 'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginic
a']
```

In [6]:

```
plt.scatter(X[:,0], X[:,2])
```

Out[6]:

```
<matplotlib.collections.PathCollection at 0x1a197c9390>
```

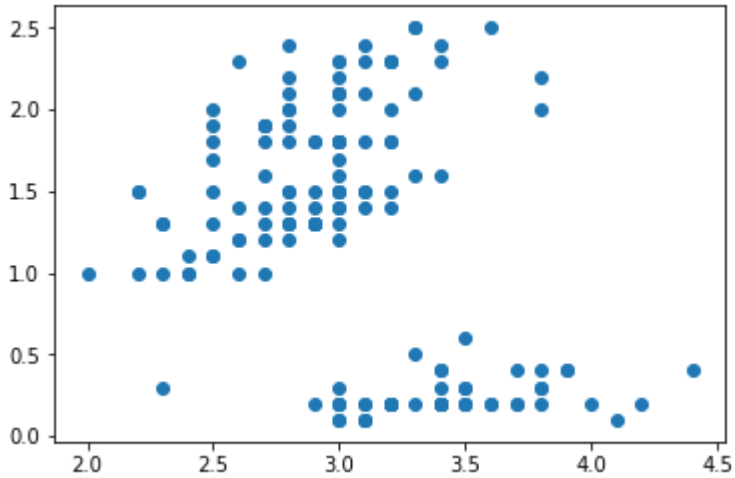


In [7]:

```
plt.scatter(X[:,1], X[:,3])
```

Out[7]:

```
<matplotlib.collections.PathCollection at 0x1a198daf98>
```



In [11]:

```
from sklearn.cluster import KMeans
model = KMeans(n_clusters=3)
model.fit(X)
```

Out[11]:

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=30
0,
    n_clusters=3, n_init=10, n_jobs=1, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

In [12]:

```
labels = model.predict(X)
print(labels)
```

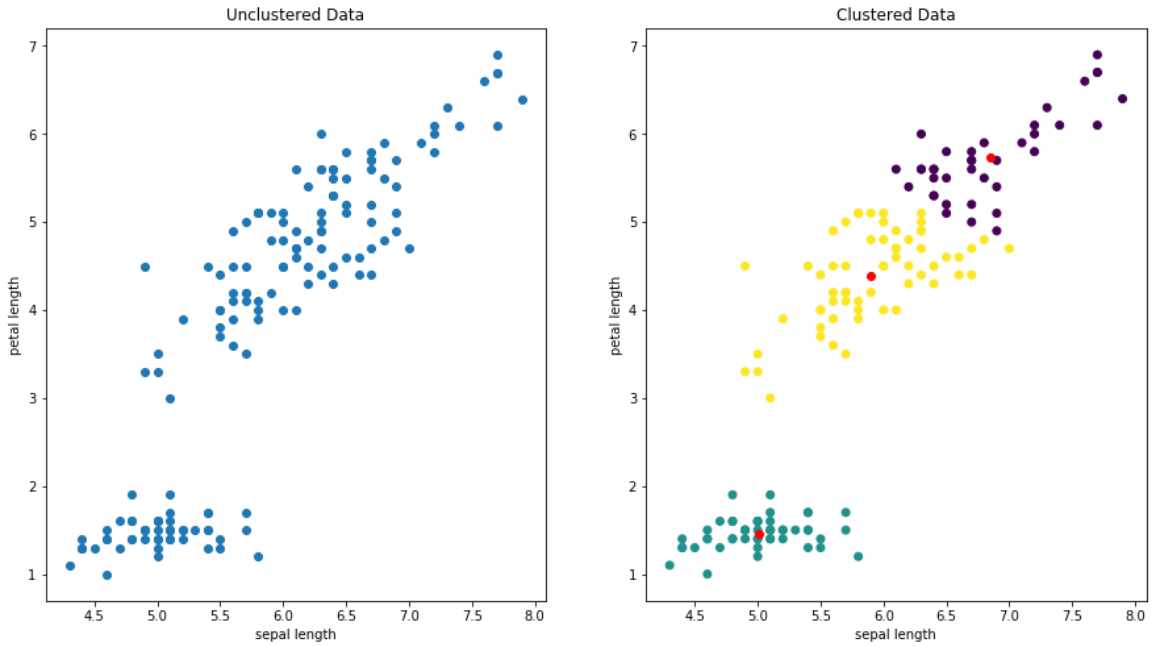
```
[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 2 2 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
 2 2 2 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 0 2 0 0 0 0 2 0
0 0 0
 0 0 2 2 0 0 0 0 2 0 2 0 2 0 0 2 2 0 0 0 0 0 2 0 0 0 0 2 0 0 0 2 0 0
0 2 0
 0 2]
```

In [14]:

```
fig = plt.figure(figsize=(15,8))
fig1 = fig.add_subplot(121)
fig2 = fig.add_subplot(122)
fig1.set_title('Unclustered Data')
fig2.set_title('Clustered Data')
fig1.scatter(X[:,0], X[:,2])
fig2.scatter(X[:,0], X[:,2], c=labels)
fig2.scatter(model.cluster_centers_[:,0] ,model.cluster_centers_[:,2], color='re
d')
fig1.set_xlabel('sepal length')
fig1.set_ylabel('petal length')
fig2.set_xlabel('sepal length')
fig2.set_ylabel('petal length')
```

Out[14]:

```
Text(0,0.5,'petal length')
```

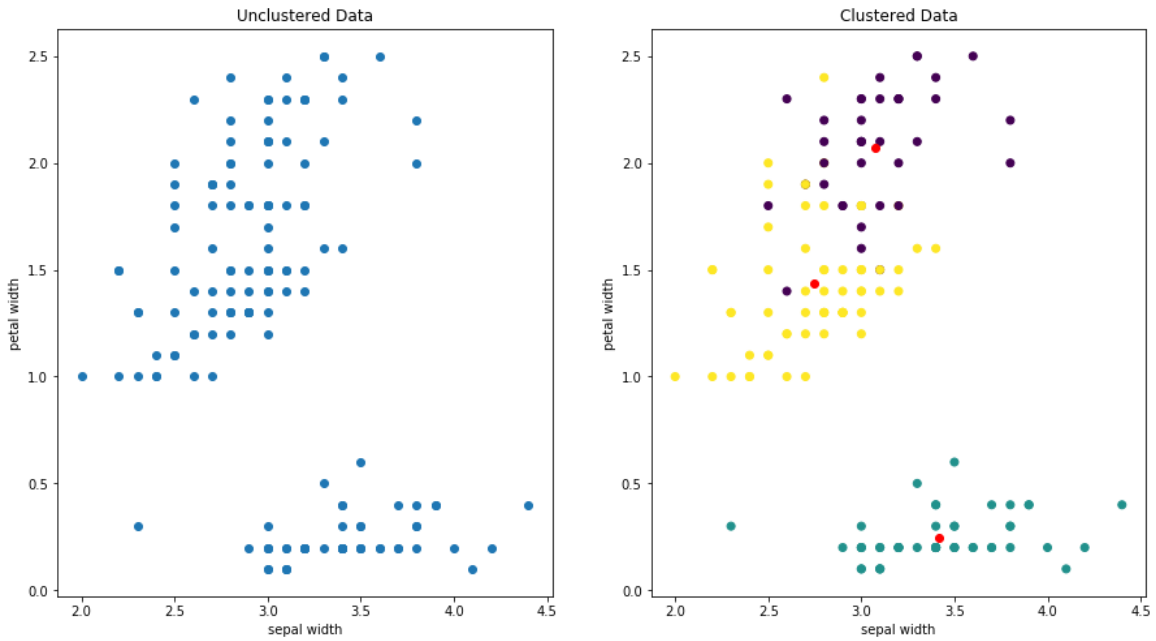


In [20]:

```
fig = plt.figure(figsize=(15,8))
fig1 = fig.add_subplot(121)
fig2 = fig.add_subplot(122)
fig1.set_title('Unclustered Data')
fig2.set_title('Clustered Data')
fig1.scatter(X[:,1], X[:,3])
fig2.scatter(X[:,1], X[:,3], c=labels)
fig2.scatter(model.cluster_centers_[:,1] ,model.cluster_centers_[:,3], color='re
d')
fig1.set_xlabel('sepal width')
fig1.set_ylabel('petal width')
fig2.set_xlabel('sepal width')
fig2.set_ylabel('petal width')
```

Out[20]:

```
Text(0,0.5,'petal width')
```



In [24]:

```
predicted_label = model.predict([[5, 0.5, 1.8, 2.6]])
print(predicted_label)
```

```
[2]
```

In [19]:

```
import pandas as pd
df = pd.DataFrame({'labels': labels, 'species': Y})
print(df)
```

	labels	species
0	1	Iris-setosa
1	1	Iris-setosa
2	1	Iris-setosa
3	1	Iris-setosa
4	1	Iris-setosa
5	1	Iris-setosa
6	1	Iris-setosa
7	1	Iris-setosa
8	1	Iris-setosa
9	1	Iris-setosa
10	1	Iris-setosa
11	1	Iris-setosa
12	1	Iris-setosa
13	1	Iris-setosa
14	1	Iris-setosa
15	1	Iris-setosa
16	1	Iris-setosa
17	1	Iris-setosa
18	1	Iris-setosa
19	1	Iris-setosa
20	1	Iris-setosa
21	1	Iris-setosa
22	1	Iris-setosa
23	1	Iris-setosa
24	1	Iris-setosa
25	1	Iris-setosa
26	1	Iris-setosa
27	1	Iris-setosa
28	1	Iris-setosa
29	1	Iris-setosa
..	...	...
120	0	Iris-virginica
121	2	Iris-virginica
122	0	Iris-virginica
123	2	Iris-virginica
124	0	Iris-virginica
125	0	Iris-virginica
126	2	Iris-virginica
127	2	Iris-virginica
128	0	Iris-virginica
129	0	Iris-virginica
130	0	Iris-virginica
131	0	Iris-virginica
132	0	Iris-virginica
133	2	Iris-virginica
134	0	Iris-virginica
135	0	Iris-virginica
136	0	Iris-virginica
137	0	Iris-virginica
138	2	Iris-virginica
139	0	Iris-virginica
140	0	Iris-virginica
141	0	Iris-virginica
142	2	Iris-virginica
143	0	Iris-virginica
144	0	Iris-virginica
145	0	Iris-virginica
146	2	Iris-virginica
147	0	Iris-virginica
148	0	Iris-virginica

```
149       2  Iris-virginica

[150 rows x 2 columns]
```

In [22]:

```
ct = pd.crosstab(df['labels'], df['species'])
print(ct)
```

```
species  Iris-setosa  Iris-versicolor  Iris-virginica
labels
0                  0                2              36
1                 50                0               0
2                  0               48              14
```

In [23]:

```
print(model.inertia_)
```

```
78.94084142614601
```