In [1]:

```
import matplotlib.pyplot as plt
from sklearn import datasets
from sklearn.cluster import KMeans
from sklearn import preprocessing
from sklearn.mixture import GaussianMixture
import pandas as pd
import numpy as np
```

In [2]:

```
iris = datasets.load_iris()
X = pd.DataFrame(iris.data)
X.columns = ['Sepal_Length','Sepal_Width','Petal_Length','Petal_Width']
y = pd.DataFrame(iris.target)
y.columns = ['Targets']
```

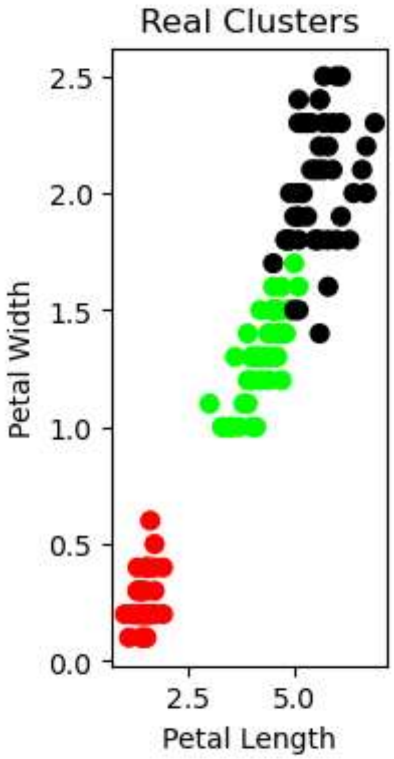
In [3]:

```
model = KMeans(n_clusters=3)
model.fit(X)
plt.figure(figsize=(6,4))
colormap = np.array(['red', 'lime', 'black'])
plt.subplot(1, 3, 1)
plt.scatter(X.Petal_Length, X.Petal_Width, c=colormap[y.Targets], s=40)
plt.title('Real Clusters')
plt.xlabel('Petal Length')
plt.ylabel('Petal Width')
```

```
C:\Users\bappo\AppData\Roaming\Python\Python39\site-packages\sklearn\cluster\_k
means.py:1429: UserWarning: KMeans is known to have a memory leak on Windows wi
th MKL, when there are less chunks than available threads. You can avoid it by
setting the environment variable OMP_NUM_THREADS=1.
  warnings.warn(
```

Out[3]:

```
Text(0, 0.5, 'Petal Width')
```

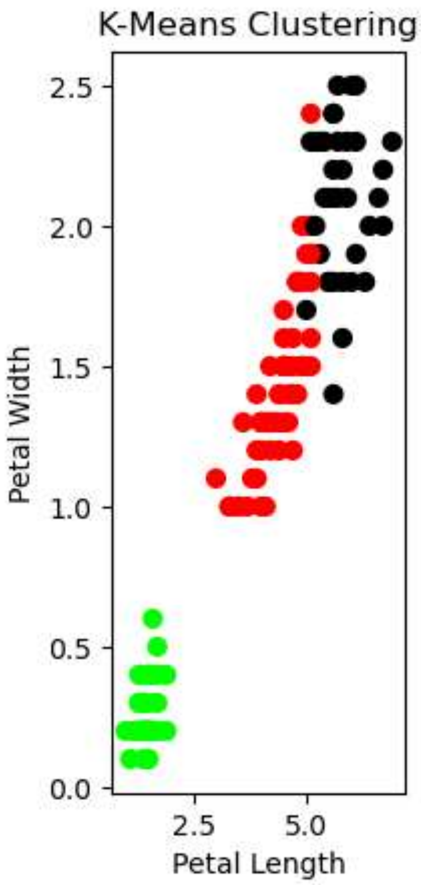


In [4]:

```
plt.subplot(1, 3, 2)
plt.scatter(X.Petal_Length, X.Petal_Width, c=colormap[model.labels_], s=40)
plt.title('K-Means Clustering')
plt.xlabel('Petal Length')
plt.ylabel('Petal Width')
```

Out[4]:

```
Text(0, 0.5, 'Petal Width')
```



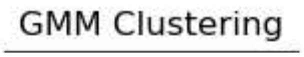
In [5]:

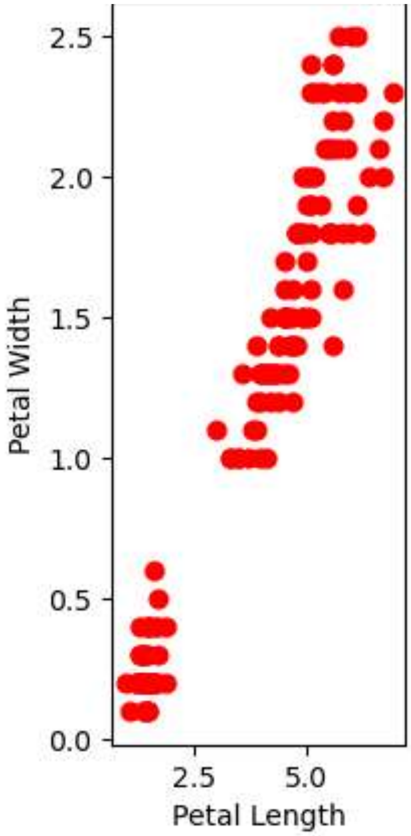
```
scaler = preprocessing.StandardScaler()
scaler.fit(X)
xsa = scaler.transform(X)
xs = pd.DataFrame(xsa, columns = X.columns)
gmm = GaussianMixture(n_components=40)
gmm.fit(xs)
plt.subplot(1, 3, 3)
plt.scatter(X.Petal_Length, X.Petal_Width, c=colormap[0], s=40)
plt.title('GMM Clustering')
plt.xlabel('Petal Length')
plt.ylabel('Petal Width')
```

```
C:\Users\bappo\AppData\Roaming\Python\Python39\site-packages\sklearn\cluster\_k
means.py:1429: UserWarning: KMeans is known to have a memory leak on Windows wi
th MKL, when there are less chunks than available threads. You can avoid it by
setting the environment variable OMP_NUM_THREADS=1.
  warnings.warn(
```

Out[5]:

```
Text(0, 0.5, 'Petal Width')
```





In [6]:

```
print('Observation: The GMM using EM algorithm based clustering matched the t
```

```
Observation: The GMM using EM algorithm based clustering matched the true label
s more closely than the Kmeans.
```