2nd_lab

February 6, 2018

In [15]:
```
#1st lab Task(AI)


import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import pyplot as plt
import matplotlib.patches as mp

%matplotlib inline
```

In [2]:
```
d = pd.read_csv("iris.csv")
df = pd.DataFrame(d)
```

In [3]:
```
# task 1
df.head(5)
```

Out[3]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

In [4]:
```
#task 1
df.tail(3)
```

Out[4]:

	sepal_length	sepal_width	petal_length	petal_width	species
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

In [5]:
```
# 2
df.sort_values('petal_length').head(2)
```

Out[5]:

	sepal_length	sepal_width	petal_length	petal_width	species
22	4.6	3.6	1.0	0.2	setosa
13	4.3	3.0	1.1	0.1	setosa

In [6]:
```
#task 3
df.groupby('species')['sepal_length','sepal_width','petal_length','petal_width'].mean()
```

Out[6]:

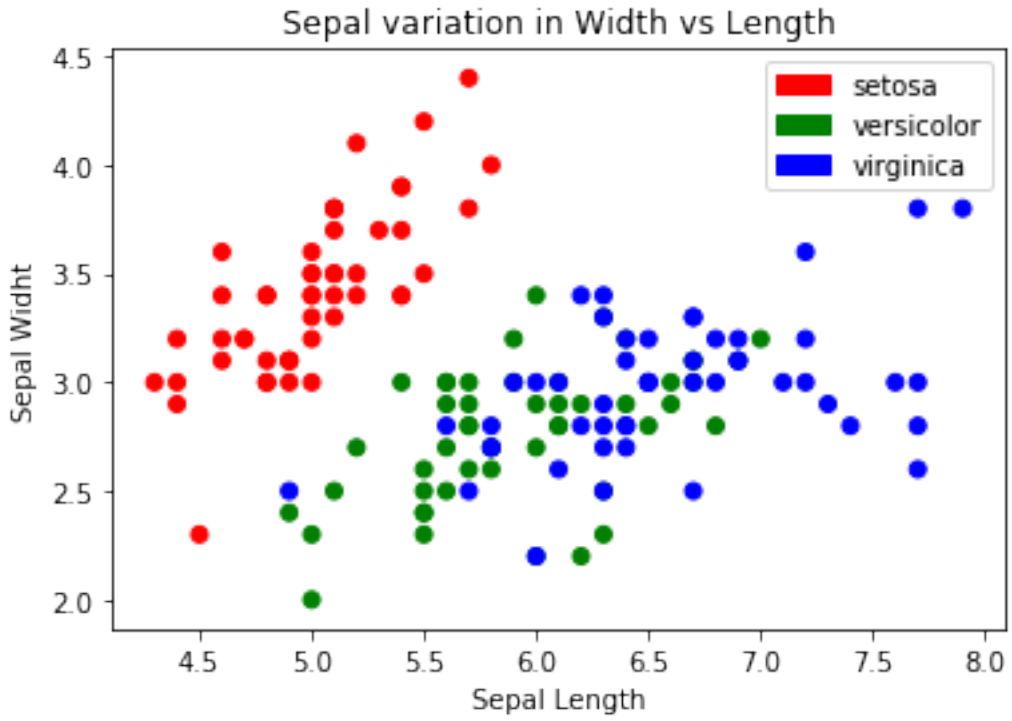
species	sepal_length	sepal_width	petal_length	petal_width
setosa	5.006	3.418	1.464	0.244
versicolor	5.936	2.770	4.260	1.326
virginica	6.588	2.974	5.552	2.026

In [18]:
```
#task 4
 # Assigning The Data
 sepal_length=df['sepal_length']
 sepal_width= df['sepal_width']
 species = df['species']
```

In [28]:
```
# create key, value pair dict to bind species to different colours
 pairs={'setosa' :'r','versicolor' :'g','virginica' :'b'}
 labels = [mp.Patch(color=cl, label=la) for la, cl in pairs.items()]

 plt.scatter(sepal_length, sepal_width, c=[pairs[i] for i in species], label=[pairs[i] f
 plt.ylabel('Sepal Widht')
 plt.xlabel('Sepal Length')
 plt.title('Sepal variation in Width vs Length')
 plt.legend(handles = labels)
```

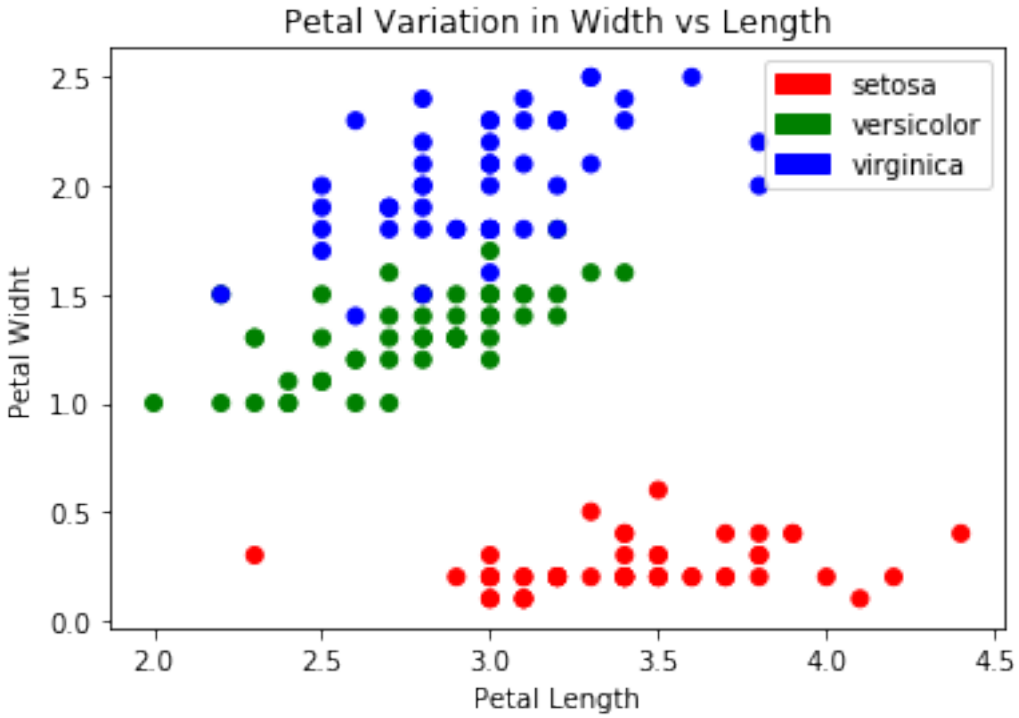
Out[28]: <matplotlib.legend.Legend at 0x21a795ec710>



In [23]:
```
#task 4
# Assigning The Data
petal_length = df['sepal_width']
petal_width = df['petal_width']
species = df['species']
```

In [27]:
```
# create key, value pair dict to bind species to different colours
pairs={'setosa' :'r','versicolor' :'g','virginica' :'b'}
labels = [mp.Patch(color=cl, label=la) for la, cl in pairs.items()]

plt.scatter(petal_length, petal_width, c=[pairs[i] for i in species], label=[pairs[i] f
plt.ylabel('Petal Widht')
plt.xlabel('Petal Length')
plt.title('Petal Variation in Width vs Length')
plt.legend(handles = labels)
plt.show()
```

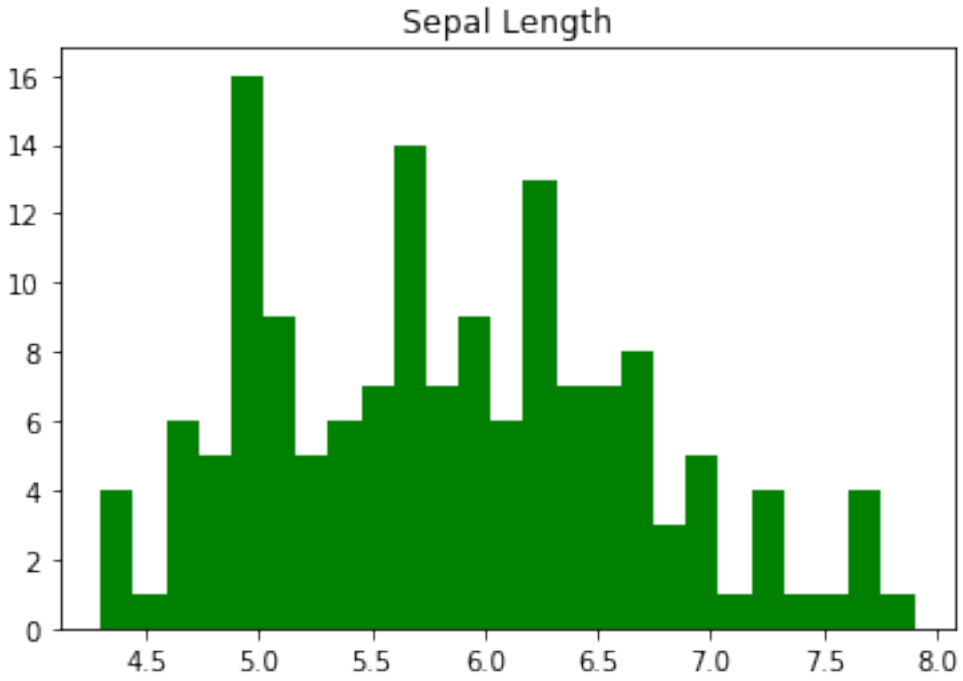


In [25]:
```
#task 5
df['Calyx Width'] = np.where(df['sepal_length'] <5, 0, 1)
df.head()
```

Out[25]:

	sepal_length	sepal_width	petal_length	petal_width	species	Calyx Width
0	5.1	3.5	1.4	0.2	setosa	1
1	4.9	3.0	1.4	0.2	setosa	0
2	4.7	3.2	1.3	0.2	setosa	0
3	4.6	3.1	1.5	0.2	setosa	0
4	5.0	3.6	1.4	0.2	setosa	1

In [29]:
```
sepal_length = df['sepal_length']
plt.hist(sepal_length, bins = 25, color = 'green')
plt.title("Sepal Length")
plt.show()
```



In [ ]: