

```
In [2]: import numpy as np
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

```
In [3]: from sklearn.cluster import KMeans
```

```
In [4]: data=pd.read_csv('C:\\Users\\hi\\Downloads\\IRIS.csv')
```

```
In [5]: data
```

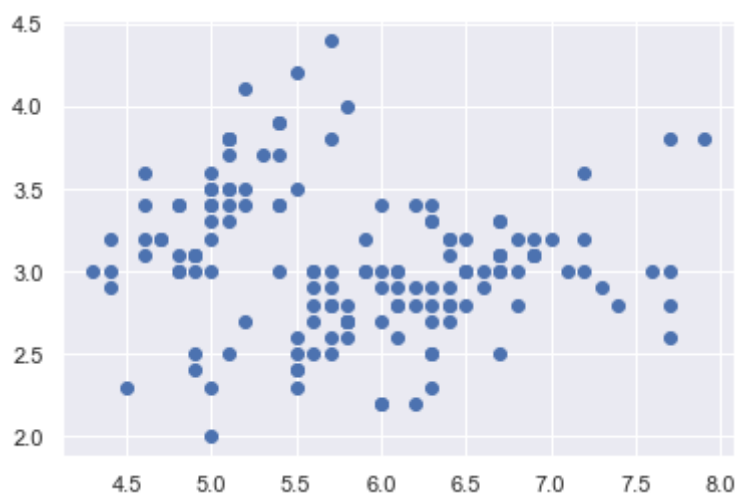
```
Out[5]:
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

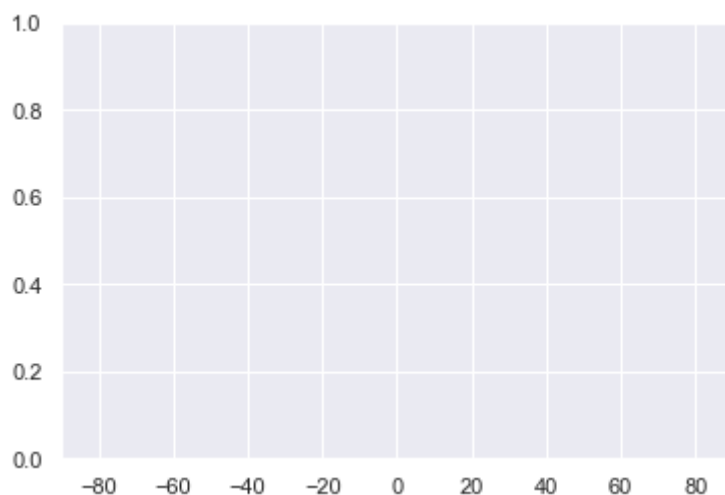
```
In [7]: plt.scatter(data['sepal_length'],data['sepal_width'])
```

```
Out[7]: <matplotlib.collections.PathCollection at 0x2079378c160>
```



```
In [11]: plt.xlim(-180,180)
plt.xlim(-90,90)
plt.show
```

```
Out[11]: <function matplotlib.pyplot.show(close=None, block=None)>
```



```
In [12]: x=data.iloc[:,1:3]
x
```

```
Out[12]:
```

	sepal_width	petal_length
--	-------------	--------------

	sepal_width	petal_length
0	3.5	1.4
1	3.0	1.4
2	3.2	1.3
3	3.1	1.5
4	3.6	1.4
...
145	3.0	5.2
146	2.5	5.0
147	3.0	5.2
148	3.4	5.4
149	3.0	5.1

150 rows × 2 columns

```
In [13]: kmeans=KMeans(3)
```

```
In [14]: kmeans.fit(x)
```

```
Out[14]: KMeans(n_clusters=3)
```

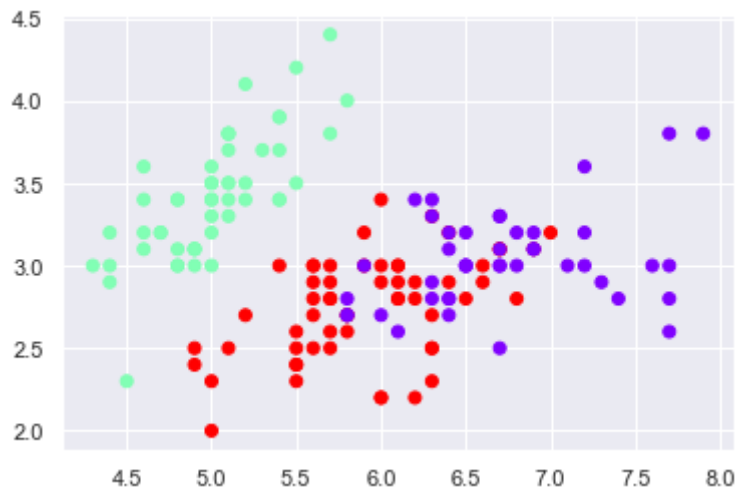
```
In [15]: identified_clusters=kmeans.fit_predict(x)
identified_clusters
```

```
Out[15]: array([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, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
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, 0, 0, 2, 0, 0, 0, 0, 0, 2, 0, 2, 0, 2, 0, 0, 2, 2, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0])
```

```
In [17]: data_with_clusters=data.copy()
data_with_clusters['clusters']=identified_clusters
```

```
In [18]: plt.scatter(data_with_clusters['sepal_length'],data_with_clusters['sepal_width'],c:
```

```
Out[18]: <matplotlib.collections.PathCollection at 0x20793989970>
```



```
In [ ]:
```