```
In [1]:    #Name: Shreevatsa
           #Regno: 200970093
           #Date: 6th Dec 2021

           import pandas as pd
```

```
In [2]:    df = pd.read_csv('Dataset/Iris.csv')
```

```
In [3]:    df.head()
```

Out[3]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

```
In [4]:    df = df.drop(columns=['Id'])
```

```
In [5]:    from sklearn.preprocessing import LabelEncoder
           le = LabelEncoder()
```

```
In [6]:    df['Species'] = le.fit_transform(df['Species'])
           df.head()
```

Out[6]:

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

```
In [7]:    df['Species'].unique()
```
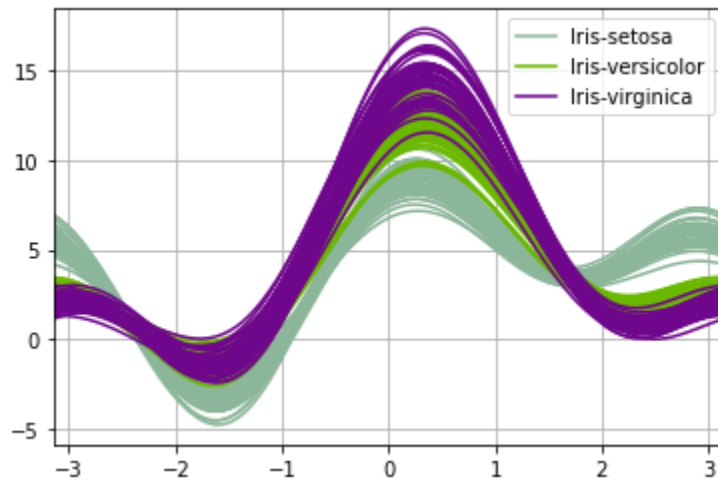
Out[7]:   array([0, 1, 2])

```
In [8]:    x = df.iloc[:,:-1]
           y = df.iloc[:,-1]
```

```
In [9]:    from sklearn.model_selection import train_test_split
           x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.30)
```

```
In [10]:   from pandas.plotting import andrews_curves
```

```
andrews_curves(pd.read_csv('Dataset/Iris.csv').drop(columns=['Id']),"Species")
```

Out[10]: `<AxesSubplot:>`

In [11]:
```python
#Unsupervised Learning
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=3,init = 'k-means++',   max_iter = 100, n_init = 10, random_sta
y_kmeans = kmeans.fit_predict(x)
```

In [12]:
```python
print(kmeans.cluster_centers_) #display cluster centers
```

```
[[5.9016129  2.7483871  4.39354839 1.43387097]
 [5.006      3.418      1.464      0.244     ]
 [6.85       3.07368421 5.74210526 2.07105263]]
```

In [13]:
```python
import matplotlib.pyplot as plt
```

In [14]:
```python
x = x.values
```

In [15]:
```python
#Visualising the clusters
plt.scatter(x[y_kmeans == 0, 0], x[y_kmeans == 0, 1], s = 100, c = 'red', label = 'Iris-se
plt.scatter(x[y_kmeans == 1, 0], x[y_kmeans == 1, 1], s = 100, c = 'blue', label = 'Iris-v
plt.scatter(x[y_kmeans == 2, 0], x[y_kmeans == 2, 1], s = 100, c = 'green', label = 'Iris-

#Plotting the centroids of the clusters
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:,1], s = 100, c = 'bla

plt.legend()
```

Out[15]: `<matplotlib.legend.Legend at 0x7f3dfb4968b0>`