Gaussian Mixture Models

Name - Amith Korada

Course - AI & ML
(Batch - 4)

Duration - 12 Months

Problem Statement - Using Apriori algorithm try to find the rules that describe the relation between each of the products that were brought by the customers a

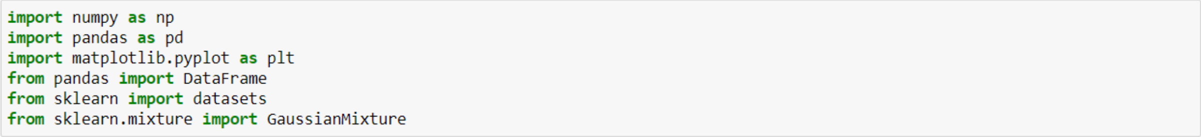
Prerequisites -

What things you need to install the software and how to install them:

Python 3.6 This setup requires that your machine has the latest version of python. The following URL https://www.python.org/downloads/ can be referred to as download python.

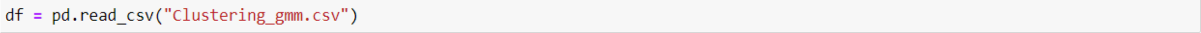
The second and easier option is to download anaconda and use its anaconda prompt to run the commands. To install anaconda check this URL https://www.anaconda.com/download/ You will also need to download and install the below 3 packages after you install either python or anaconda from the steps above Sklearn (scikit-learn) numpy scipy if you have chosen to install python 3.6 then run the below commands in command prompt/terminal to install these packages pip install -U sci-kit-learn pip install NumPy pip install scipy if you have chosen to install anaconda then run the below commands in anaconda prompt to install these packages conda install -c sci-kit-learn conda install -c anaconda numpy conda install -c anaconda scipy.

1. Importing necessary libraries-

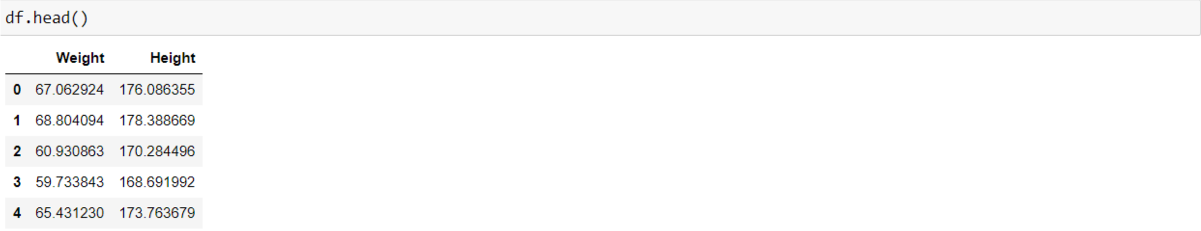


```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from pandas import DataFrame
from sklearn import datasets
from sklearn.mixture import GaussianMixture
```

2. Loading the dataset-

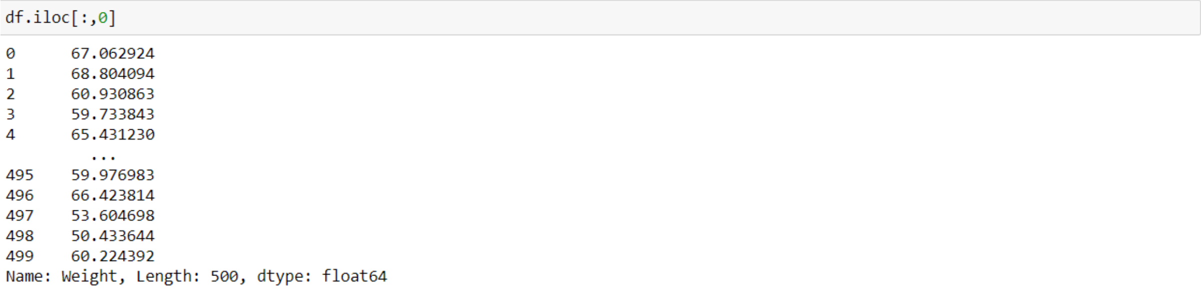


```
df = pd.read_csv("Clustering_gmm.csv")
```



```
df.head()
```

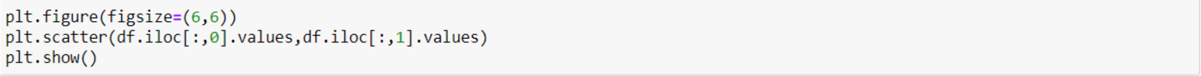
	Weight	Height
0	67.062924	176.086355
1	68.804094	178.388669
2	60.930863	170.284496
3	59.733843	168.691992
4	65.431230	173.763679



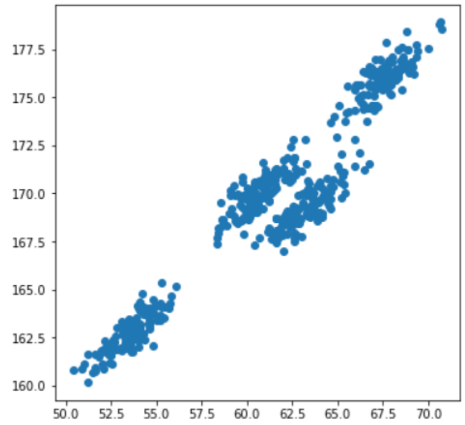
```
df.iloc[:,0]
```

```
0      67.062924
1      68.804094
2      60.930863
3      59.733843
4      65.431230
         ...
495    59.976983
496    66.423814
497    53.604698
498    50.433644
499    60.224392
Name: Weight, Length: 500, dtype: float64
```

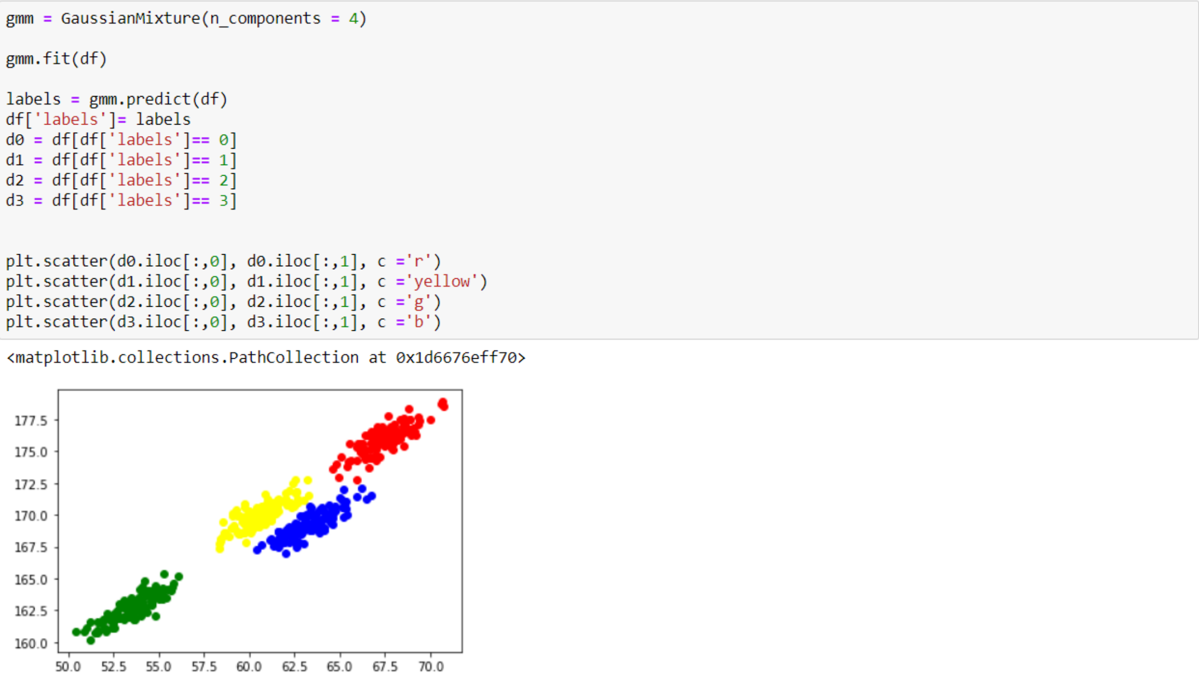
3. Plotting the data-



```
plt.figure(figsize=(6,6))
plt.scatter(df.iloc[:,0].values,df.iloc[:,1].values)
plt.show()
```



4. Implementing the model-



```
gmm = GaussianMixture(n_components = 4)

gmm.fit(df)

labels = gmm.predict(df)
df['labels']= labels
d0 = df[df['labels']== 0]
d1 = df[df['labels']== 1]
d2 = df[df['labels']== 2]
d3 = df[df['labels']== 3]


plt.scatter(d0.iloc[:,0], d0.iloc[:,1], c ='r')
plt.scatter(d1.iloc[:,0], d1.iloc[:,1], c ='yellow')
plt.scatter(d2.iloc[:,0], d2.iloc[:,1], c ='g')
plt.scatter(d3.iloc[:,0], d3.iloc[:,1], c ='b')
```

```
<matplotlib.collections.PathCollection at 0x1d6676eff70>
```