In [6]:
```
import pandas as pd
from sklearn.cluster import KMeans
from sklearn.mixture import GaussianMixture
import matplotlib.pyplot as plt
```

In [8]:
```
dataset=pd.read_csv("Clustering_gmm.csv")
```

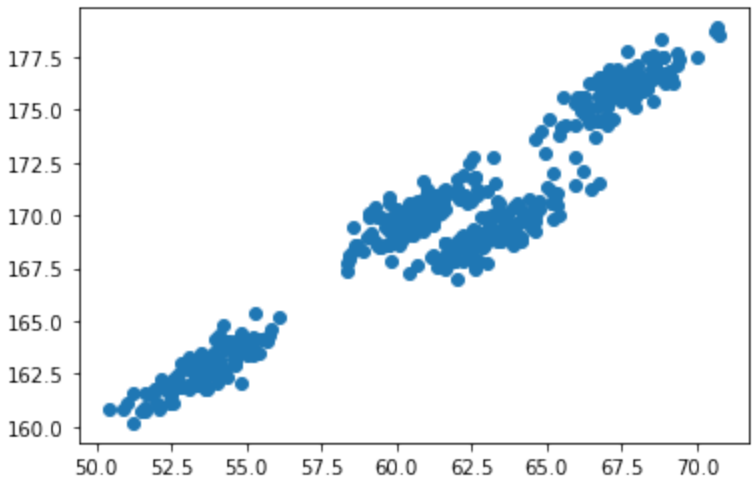
In [10]:
```
dataset.head(10)
```

Out[10]:

	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
5	61.577160	168.091751
6	63.341866	170.642516
7	61.041643	170.096682
8	62.633623	171.862972
9	53.407860	162.756843

In [11]:
```
plt.figure()
plt.scatter(dataset['Weight'],dataset['Height'])
```

Out[11]:
```
<matplotlib.collections.PathCollection at 0x236daa9b400>
```



KMeans

In [13]:
```
kmean=KMeans(n_clusters=4)
```

In [14]:
```
kmean.fit(dataset)
```

Out[14]:
```
KMeans(n_clusters=4)
```

In [16]:
```
dataset1=pd.DataFrame(dataset)#Duplicating the same dataset
```

In [19]:
```
predict=kmean.predict(dataset)
```

In [46]:
```
dataset1['Predictions']=predict
dataset1
#adding column to the dataset and adding prediction column which shows the id's of
```

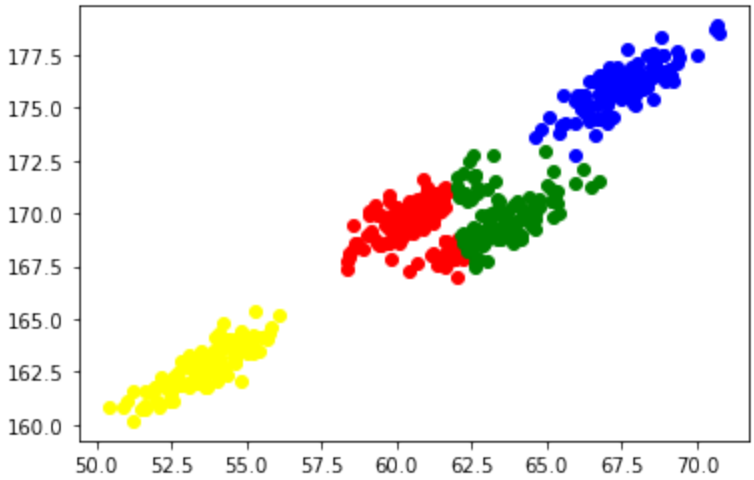
Out[46]:

	Weight	Height	Predictions
0	67.062924	176.086355	2
1	68.804094	178.388669	2
2	60.930863	170.284496	0
3	59.733843	168.691992	0
4	65.431230	173.763679	2
...	...	...	...
495	59.976983	169.679741	0
496	66.423814	174.625574	2
497	53.604698	161.919208	1
498	50.433644	160.794875	1
499	60.224392	169.689709	0

500 rows × 3 columns

In [26]:
```
color=['red','yellow','blue','green']
```

In [27]:
```
for i in range(0,4):
    data=dataset1[dataset1['Predictions']==i]
    plt.scatter(data['Weight'],data['Height'],c =color[i])
plt.show()
```



Guasssian Model

In [28]:
```
dataframe=pd.read_csv("Clustering_gmm.csv")
```

In [29]:
```
dataframe.head(10)
```

Out[29]:

	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
5	61.577160	168.091751
6	63.341866	170.642516
7	61.041643	170.096682
8	62.633623	171.862972
9	53.407860	162.756843

In [32]:
```
gmm=GaussianMixture(n_components=4)
```

In [33]:
```
gmm.fit(dataframe)
```

Out[33]:
```
GaussianMixture(n_components=4)
```

In [34]:
```
dataframe1=pd.DataFrame(dataframe)
```

In [35]:
```
predict1=gmm.predict(dataframe)
```

In [45]:
```
dataframe1['Predictions']=predict1
dataframe1
```

Out[45]:

	Weight	Height	Predictions
0	67.062924	176.086355	1
1	68.804094	178.388669	1
2	60.930863	170.284496	0
3	59.733843	168.691992	0
4	65.431230	173.763679	1
...	...	...	...
495	59.976983	169.679741	0
496	66.423814	174.625574	1
497	53.604698	161.919208	2
498	50.433644	160.794875	2
499	60.224392	169.689709	0

500 rows × 3 columns

In [41]:
```
for i in range(0,4):
    data1=dataframe1[dataframe1['Predictions']==i]
    plt.scatter(data1['Weight'],data1['Height'],c =color[i])
plt.show()
```