In [22]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.cluster import KMeans
from sklearn.mixture import GaussianMixture
```

In [23]:
```
raw_data = pd.read_csv("./EmployeeAttrition.csv", index_col=0)
```

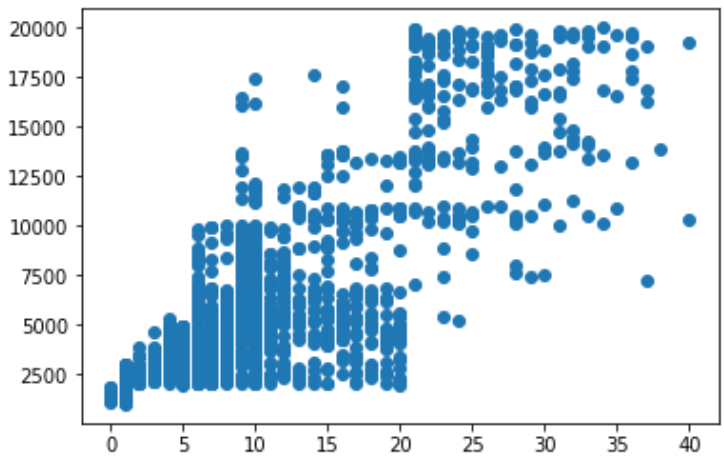
In [24]:
```
raw_data.head()
```

Out[24]:

	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EducationField	Emplc
0	41	Yes	Travel_Rarely	Sales	1	2	Life Sciences	
1	49	No	Travel_Frequently	Research & Development	8	1	Life Sciences	
2	37	Yes	Travel_Rarely	Research & Development	2	2	Other	
3	33	No	Travel_Frequently	Research & Development	3	4	Life Sciences	
4	27	No	Travel_Rarely	Research & Development	2	1	Medical	

5 rows × 27 columns

In [25]:
```
plt.scatter(raw_data["TotalWorkingYears"], raw_data["MonthlyIncome"])
plt.show()
```

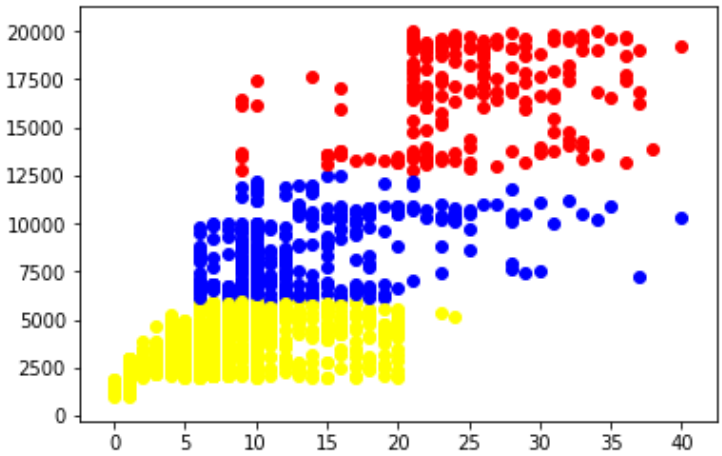


In [26]:
```
two_mode_data = np.array(raw_data[["TotalWorkingYears", "MonthlyIncome"]]
```

In [27]:
```
kmeans_model = KMeans(n_clusters=3, random_state=100).fit(two_mode_data)
```

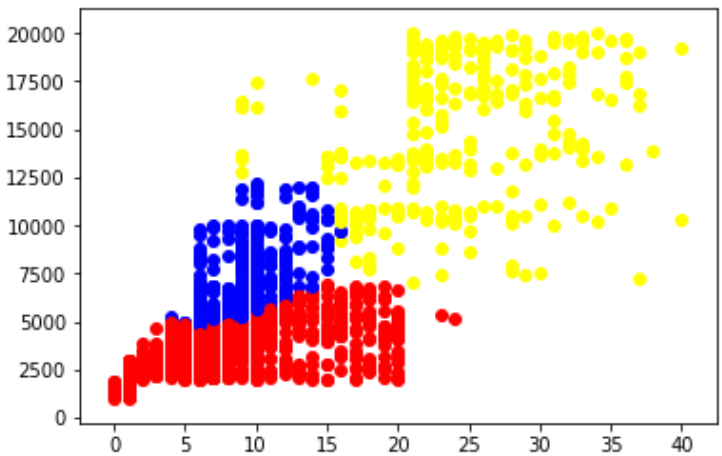
In [28]:
```
color_list = ["blue", "red", "yellow", "green", "black"]
```

In [29]:
```
for value, label in zip(two_mode_data, kmeans_model.labels_):
    x = value[0]
    y = value[1]
    color = color_list[label]
    plt.scatter(x, y, color=color)
plt.show()
```



In [30]:
```
gmm_model = GaussianMixture(n_components=3, random_state=0).fit(two_mode_
gmm_labels = gmm_model.predict(two_mode_data)
```

In [31]:
```
for value, label in zip(two_mode_data, gmm_labels):
    x = value[0]
    y = value[1]
    color = color_list[label]
    plt.scatter(x, y, color=color)
plt.show()
```



In [ ]:
```

```