In [2]: **import** numpy **as** np **import** pandas **as** pd

In [3]: data**=**pd**.**read\_csv("social\_network\_ads.csv") data**.**head()

Out[3]: **User ID Gender Age EstimatedSalary Purchased**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **0** | 15624510 | Male | 19 | | 19000 | 0 |
| **1** | | 15810944 | Male | 35 |  | 20000 | 0 | |
| **2** | | 15668575 | Female | 26 |  | 43000 | 0 | |
| **3** | | 15603246 | Female | 27 |  | 57000 | 0 | |
| **4** | | 15804002 | Male | 19 |  | 76000 | 0 | |

In [7]: print(data**.**describe()) print("-"**\***70) print(data**.**isnull()**.**sum())

User ID Age EstimatedSalary Purchased count 4.000000e+02 400.000000 400.000000 400.000000 mean 1.569154e+07 37.655000 69742.500000 0.357500 std 7.165832e+04 10.482877 34096.960282 0.479864 min 1.556669e+07 18.000000 15000.000000 0.000000 25% 1.562676e+07 29.750000 43000.000000 0.000000

50% 1.569434e+07 37.000000 70000.000000 0.000000 75% 1.575036e+07 46.000000 88000.000000 1.000000 max 1.581524e+07 60.000000 150000.000000 1.000000

----------------------------------------------------------------------

User ID 0 Gender 0

Age 0

EstimatedSalary 0 Purchased 0 dtype: int64

|  |  |
| --- | --- |
| x **=** | pd**.**get\_dummies(data**.**drop(columns**=**['Purchased']), drop\_first**=True**) |
| y **=** | data['Purchased'] |
|  |  |
| **from** sklearn.model\_selection **import** train\_test\_split  x\_train, x\_test, y\_train, y\_test **=** train\_test\_split(x, y, test\_size**=**0.25, random\_state**=**42) | |

In [14]:

In [15]:

In [16]: **from** sklearn.linear\_model **import** LogisticRegression lr **=** LogisticRegression(random\_state**=**0, solver**=**'lbfgs') lr**.**fit(x\_train, y\_train) pred **=** lr**.**predict(x\_test)

|  |
| --- |
| In [23]: |

print("Test Data:\n", x\_test[:10]) print("-"**\***40) print("Predicted Labels: ", pred[:10]) print("-"**\***40)

print("Actual Labels:\n", y\_test[:10])

Test Data:

User ID Age EstimatedSalary Gender\_Male

209 15697424 46 22000 False

280 15609669 59 88000 False

33 15776733 28 44000 False

210 15724536 48 96000 False 93 15699284 29 28000 False

84 15798659 30 62000 False

329 15639576 47 107000 False

94 15786993 29 83000 False 266 15721592 40 75000 True

126 15610801 42 65000 True

----------------------------------------

Predicted Labels: [0 0 0 0 0 0 1 0 0 0]

---------------------------------------Actual Labels:

209 0

280 1 33 0

210 1

93 0

84 0 329 1

94 0

266 0

126 0

Name: Purchased, dtype: int64

In [26]: **from** sklearn.metrics **import** confusion\_matrix, ConfusionMatrixDisplay, classification\_report, cm **=** confusion\_matrix(y\_test, pred, labels**=**lr**.**classes\_) print("Confusion Matrix:\n",cm)

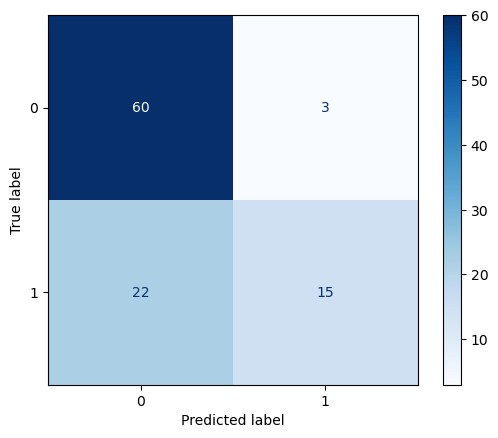
Confusion Matrix:

[[60 3]

[22 15]]

|  |
| --- |
| **import** matplotlib.pyplot **as** plt **import** seaborn **as** sns  display\_matrix **=** ConfusionMatrixDisplay(confusion\_matrix**=**cm, display\_labels**=**lr**.**classes\_) display\_matrix**.**plot(cmap**=**plt**.**cm**.**Blues) plt**.**show() |

In [29]:



|  |
| --- |
| accuracy **=** accuracy\_score(y\_test,pred) error\_rate **=** 1 **-** accuracy recall **=** recall\_score(y\_test, pred) precision **=** precision\_score(y\_test, pred)  print("Accuracy: ",accuracy,"\nAccuracy(%): ",accuracy**\***100) print("Error Rate: ",error\_rate,"\nError Rate(%): ",error\_rate**\***100) print("Recall: ",recall,"\nRecall(%): ",recall**\***100)  print("Precision: ",precision,"\nPrecision(%): ",precision**\***100) |

In [34]:

Accuracy: 0.75

Accuracy(%): 75.0

Error Rate: 0.25

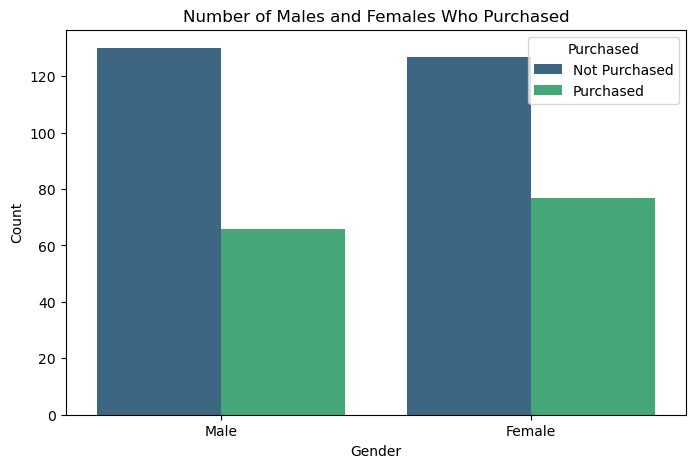
Error Rate(%): 25.0

Recall: 0.40540540540540543 Recall(%): 40.54054054054054

Precision: 0.8333333333333334

Precision(%): 83.33333333333334

In [40]: plt**.**figure(figsize**=**(8, 5)) sns**.**countplot(x**=**'Gender', hue**=**'Purchased', data**=**data, palette**=**'viridis') plt**.**xlabel('Gender') plt**.**ylabel('Count') plt**.**title('Number of Males and Females Who Purchased') plt**.**legend(title**=**'Purchased', labels**=**['Not Purchased', 'Purchased']) plt**.**show()



|  |
| --- |
|  |

In [ ]: