Task 3 – Example Answer with Python Code

**from** sklearn.linear\_model **import** LogisticRegression

**from** sklearn **import** metrics

**import** numpy **as** np

**import** pandas **as** pd

*# Read in loan data from a CSV file*

df **=** pd**.**read\_csv('loan\_data\_created.csv')

*# Define the variable features*

features **=** ['credit\_lines\_outstanding', 'debt\_to\_income', 'payment\_to\_income', 'years\_employed', 'fico\_score']

*# Calculate the payment\_to\_income ratio*

df['payment\_to\_income'] **=** df['loan\_amt\_outstanding'] **/** df['income']

*# Calculate the debt\_to\_income ratio*

df['debt\_to\_income'] **=** df['total\_debt\_outstanding'] **/** df['income']

clf **=** LogisticRegression(random\_state**=**0, solver**=**'liblinear', tol**=**1e-5, max\_iter**=**10000)**.**fit(df[features], df['default'])

print(clf**.**coef\_, clf**.**intercept\_)

*# Use the following code to check yourself*

y\_pred **=** clf**.**predict(df[features])

fpr, tpr, thresholds **=** metrics**.**roc\_curve(df['default'], y\_pred)

print((1.0\*(abs(df['default']**-**y\_pred))**.**sum()) **/** len(df))

print(metrics**.**auc(fpr, tpr))