**Module 12 Report**

**Overview of the Analysis**

This set of data was on the data of lending activity based on each person. The purpose of the data is to analyze the different types of borrowers and the lending patterns associated with them. Some of the variables we were trying to predict were the loan status and the number of loans that were predicted through .value\_counts, as well as the rest of the data such as loan size, interest rate, borrower income, etc. The stages of machine learning during the exercise include, importing and reshaping the data so we can use it in our model, splitting the training and testing datasets for the model, fitting a logistic regression model onto our model, predicting, and then classifying the predictions it has given us. We also resampled our data to test our model’s capabilities, also fitting this resampled data through a logistic regression model.

**Results**

* Machine Learning Model 1:
  + Accuracy: .99 (99%)
  + Precision:
    - “Healthy Loans” : 1.00 (100%)
    - “High Risk Loans” : .85 (85%)
  + Recall Scores:
    - “Healthy Loans” : .99 (99%)
    - “High Risk Loans” : .91 (91%)
* Machine Learning Model 2:
  + Accuracy: .99 (99%)
  + Precision:
    - “Healthy Loans” : 1.00 (100%)
    - “High Risk Loans” : .84 (84%)
  + Recall Scores:
    - “Healthy Loans” : .99 (99%)
    - “High Risk Loans” : .99 (99%)

**Summary**

Both models seem to perform quite well at 99 accuracy rates; Model 2 does a little better on the recall, however. This is said because the recall score in Model 2 tested at 99 percent while the recall in Model 1 tested at 91 percent for the high-risk loans. Even though it is a higher rate, Model 1 is more practical because it is based off of previous loan activities, and not resampled data used primarily for testing.