Heart disease Capstone Milestone Report:

Problem Statement: This report hopes to identify certain characteristics that make one more vulnerable or susceptible to heart disease. This analysis can be useful to a wide variety of audiences, most obviously doctors who can identify traits/characteristics in their current patients to help prevent heart disease in the future.

The dataset was taken from Kaggle. There were 300 records all with features and labels. I used a train test split in order to set aside test records to use to determine the accuracy of the model trained on the training set. Because the data set was taken from Kaggle, it was mostly preprocessed, with no missing values, transformations or replacements needed.

Data visualizations and correlation tables were created in order to do some preliminary analysis of the features and which features were the strongest indicators of heart disease. Visualizations were created using the matplotlib package in python. I will include them in other reports.