DSC 423
Assignment 4

Based on Modules 8 and 9

Your submission must include your name and student ID. Your submission must include the honor statement: "I have completed this work independently. The solutions given are entirely my own work." You submission must be submitted as a PDF.

1) Regularization (35 points) Last week you created a model using the PISA dataset. Build a model again, this time…
    a. Use Ridge regression and present your model along with appropriate outputs.
        i. Show the ridge trace plot and discuss how this techniques handles multicolinearity.
        ii. Evaluate the residual plots. Present the appropriate plots, describe them and draw appropriate conclusions.
    b. Use LASSO regression and present your model along with appropriate outputs.
        i. LASSO is a form of feature selection. Discuss how it reduced the feature space.
    c. Are the two models from a and b the same? Explain.
2) REMISSION (15 points)
    a. Download "remission" and create a logistic model to predict remission.
    b. Perform logistic regression.
        i. Submit your model.
    c. Notice that you are using the *glm* function.
        i. Explain how this differs from lm().
    d. Provide an analysis.
        i. Evaluate the model?
        ii. Evaluate the independent variables?