

STA 4320 – Homework 1

Prof. He Jiang

Name: _____

Student Number: _____

For the current assignment, Your report should be no longer than the following amount of pages:
5

You do not have to submit this instruction page.

Please compile your coding into a single PDF file.

Please submit your compiled (PDF file) report to the corresponding assignment on Gradescope.

Please label your solution. If there is only 1 question, please label as well.

1. The **Auto** dataset in the ISLR2¹ package contains Gas mileage, horsepower, and other information for $n = 392$ vehicles.

We would like to use simple linear regression to investigate the relationship between the response, Y , mpg, and the independent variable, X , horsepower.

Grading Method

The grading of this assignment will be based on: Completion.

Submissions that are empty or did not show sufficient efforts will receive 0. Submissions that showed significant efforts will receive full amount of points.

Tasks for this Assignment

- (a) (2 points) Load the **Auto** dataset, and print the first 6 rows.
- (b) (2 points) Conduct linear regression using the `lm` command, and print the summary of this regression.
- (c) (2 points) Create a scatterplot of X and Y . Overlay the regression line. Please label your plot.
- (d) (1 point) At the $\alpha = 0.05$ significance level, with the intercept already in the model, is there a negative relationship between Y and X ? Give support to your answer.
- (e) (1 point) At the $\alpha = 0.05$ significance level, is the entire model significant? Give support to your answer.
- (f) (2 points) What is the simple point prediction for mpg associated with a horsepower of 98? What are the associated 95% confidence and prediction intervals? Please round your answers to 4 decimal points.

¹This is the R package corresponding to our textbook.