MT 4420 (Brandt) Section Quiz 1: Written Portion Read these instructions!
• This portion of the quiz is open note and open book. You may refer to any of the course materials on our Canvas course page.

• You may not communicate with anyone else while completing this quiz. • Create a Word document and organize your work in report form as you do on the homework assignments. Please put your name and page numbers in the header of your Word document.
Always include your code, any relevant output (plots, summaries, etc.), and your conclusions. Save your report as a pdf and submit it to Canvas.

• You must submit your report by 11:00 pm on Thursday, Feb 15.

FILIPPO RIVA
MT4420

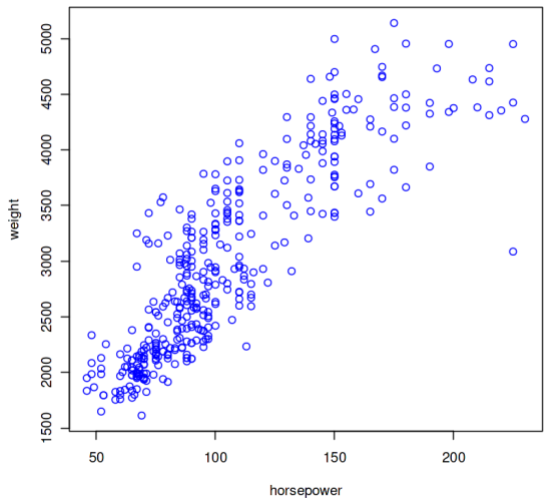
QUIZ 1: WRITTEN PORTION

1. (14 points) Working with the Auto data set.
Note: Recall that the Auto data set has some missing entries for horsepower. Be sure to use the na.strings="?" command when you read in the data set. For details, see ISL, page 49.

a. Plot horsepower (horiz. axis) versus weight (vert. axis). Explain why it makes sense for this plot to have a general upward trend.

CODE:

```
Auto = read.csv(file="Auto.csv",head=TRUE,sep=",",stringsAsFactors=FALSE, na.strings="?")
Auto = na.omit(Auto)
horsepower=Auto$horsepower
weight=Auto$weight
 plot(horsepower,weight,col="blue")
```



The trend makes sense because in general cars with a higher horsepower are also weighing more.

b. Use the lm command to find the least squares line for weight as a function of horsepower. Include commands, the lm summary output, and a plot of the data and the least squares line together. Please clean up the lm summary output so it looks like it does in R (or paste in a screenshot).

CODE:

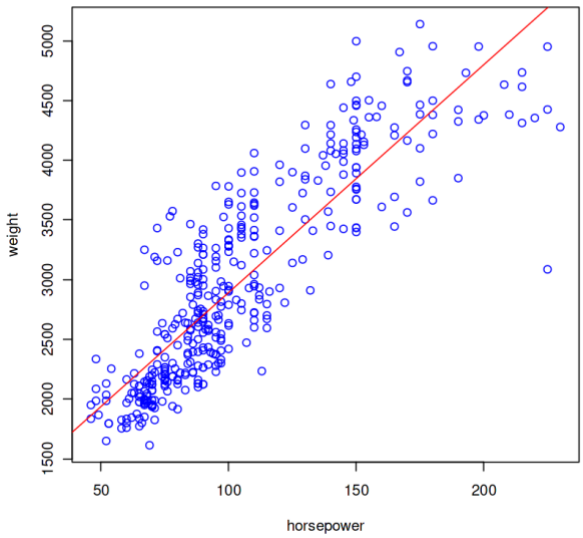
```
Auto = read.csv(file="Auto.csv",head=TRUE,sep=",",stringsAsFactors=FALSE, na.strings="?")
Auto = na.omit(Auto)
horsepower=Auto$horsepower
weight=Auto$weight
myline=lm(weight ~ horsepower)
summary(myline)
plot(horsepower,weight,col="blue")
abline(myline,col="red")
```

```
Call:
lm(formula = weight ~ horsepower)

Residuals:
    Min      1Q  Median      3Q     Max
-2191.1  -297.7   -80.1   330.8  1150.8

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 984.5003    62.5143   15.75   <2e-16 ***
horsepower   19.0782     0.5616   33.97   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 427.4 on 390 degrees of freedom
Multiple R-squared:  0.7474,    Adjusted R-squared:  0.7468
F-statistic:  1154 on 1 and 390 DF,  p-value: < 2.2e-16
```



Least square line: w = 19.0782 h + 984.5003

c. Explain what the p-values for slope and intercept tell you about your model.

The p-value for slope and intercept are really small numbers, telling us that there is a strong relationship between weight and horsepower because it is possible to reject the null hypothesis Ho stating that there is no relationship between horsepower and weight.

d. State you least squares model in the form W = mH+ b , where W and H are variables for weight and horsepower respectively, m is the slope, and b is the vertical intercept. Round slope and intercept to 1 place after the decimal.

W = 19.1 H + 984.5

e. Complete the following sentence. According to the model, a 75 increase in horsepower corresponds to a … (describe the weight change). Round the answer to 1 place after the decimal.

H1 = 0
According to the formula W1 = 19.1 (0) + 984.5 = 984.5

H2 = 75
According to the formula W2 = 19.1 (75) + 984.5 = 2417

An 75 increase in horsepower corresponds to a ΛW = 2417 – 984.5 = 1432.5 change in weight

f**. Use your model to answer the following. The easiest way may be to do pencil/paper algebra and calculations. Include your calculations so I can see how you got your answers. Round answers to 1 place after the decimal.**
o What is the predicted weight of a car that has 150 horsepower?

Car Horsepower = 150
W = 19.1 (150) + 984.5 = 3849.5

o What horsepower you would you expect for a car that weighs 3500 pounds?
Car weight = 3500

3500 = 19.1 H + 984.5

H= (3500 – 984.5) / 19.1 = 131.7

g**. Give 95% confidence intervals for the slope and intercept of your model. For full credit, use the approach outlined on page 66. For intercept, round answers to 1 place after the decimal. For slope, round answers to 2 places after the decimal.**

In order to calculate the 95% interval for $\beta 0$ and $\beta 1$ we need to use the following formula:

$[\hat{\beta}1 - 2 \cdot SE(\hat{\beta}1), \hat{\beta}1 + 2 \cdot SE(\hat{\beta}1)]$ horsepower
$[\hat{\beta}0 - 2 \cdot SE(\hat{\beta}0), \hat{\beta}0 + 2 \cdot SE(\hat{\beta}0)]$ Intercept

According to descriptive statistics obtained above:
$\hat{\beta}0 = 984.5003$
$SE(\hat{\beta}0) = 62.5143$
$\hat{\beta}1 = 19.0782$
$SE(\hat{\beta}1) = 0.5616$

By substituting these numbers in the formula above, the following intervals are obtained

$\hat{\beta}0$ (859.5, 1109.5)
$\hat{\beta}1$ (17.95, 20.20)