- 4.4 – **[10 pts]**
  1. The results of `summary()` is shown in **Table B.19**.

Table B.19: Summary of simple linear regression results of mpg on horsepower.

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 46.92659    1.92184   24.42  < 2e-16
hp          -0.21762    0.01829  -11.90 1.03e-14
---
Residual standard error: 3.096 on 40 degrees of freedom
Multiple R-squared: 0.7796,        Adjusted R-squared: 0.7741
F-statistic: 141.5 on 1 and 40 DF,  p-value: 1.027e-14
```

  - The "Intercept" "Estimate" is the estimated mean mpg for a car with no horsepower.
  - The "hp" "Estimate" (i.e., the slope) says that for every increase in one horsepower the mpg will decrease by 0.21762, on average.
  - The "Intercept" p-value shows that the mean mpg for a car with no horsepower is different than zero (a nonsensical significance test).
  - The "hp" p-value shows that the slope is significantly different from zero indicating that there is a significant relationship between the horsepower and mpg of a car.
  - The "residual standard error" (i.e., square root of $MS_{residual}$) is a measure of the standard deviation **about the line**.
  - The "residual degrees of freedom" is two less then the number of observations.
  - The "multipe r-squared" shows the proportion of the total variability in mpg (ignoring horsepower) that is explained away by knowing a horespower value.
  - The F and corresponding p-value show that the full model including the slope is significantly "better" than the simple model with no slope. Thus, a slope "is needed" and it can be concluded that there is a significant relationship between a car's horespower and its gas mileage.

  2. The results of `confint()` is shown in **Table B.20**.

Table B.20: Coefficient confidence intervals for simple linear regression results of mpg on horsepower.

```
                 2.5 %    97.5 %
(Intercept) 43.0424051 50.810780
hp          -0.2545932 -0.180651
```

  - The "intercept" CI says that the mean mpg for a car with no horsepower is between 43.04 and 50.81.
  - The "hp" CI says that the mpg will decrease between 0.18 and 0.25 for a one unit increase in horespower.

  3. The results of `anova()` is shown in **Table B.21**.

Table B.21: ANOVA table for simple linear regression results of mpg on horsepower.

```
          Df  Sum Sq Mean Sq F value    Pr(>F)
hp         1 1356.83 1356.83  141.53 1.027e-14
Residuals 40  383.48    9.59
Total     41 1740.30
```

  - The "hp" df (i.e., $df_{regression}$) is one less than the number of parameters estimate (two – intercept and slope).

  - The "Residuals" df is two less than the number of observations.
  - The "Total" df is one less than the number of observations.
  - The "Total" MS is the total variance in mpg (ignoring horsepower) or the total variability around the mean mpg (simple model).
  - The "Residual" MS is the total variance in mpg after considering horsepower or the total variabiltiy around the line (full model).
  - The "hp" MS (i.e., $MS_{regression}$) is the variance in mpg that can be explained by knowing the value of horespower.
  - The F test statistic is the ratio of variability in mpg explained by knowing the value of horespower to the variability left unexplained even after knowing the value of the horsespower.
4. There is a significant relationship as indicated by the very small slope and F-test p-values (p<0.00005).

R Commands

```
> car <- read.table("CarMPG.txt", head = TRUE)
> attach(car)
> lm1 <- lm(mpg ~ hp)
> summary(lm1)
> confint(lm1)
> anova(lm1)
```