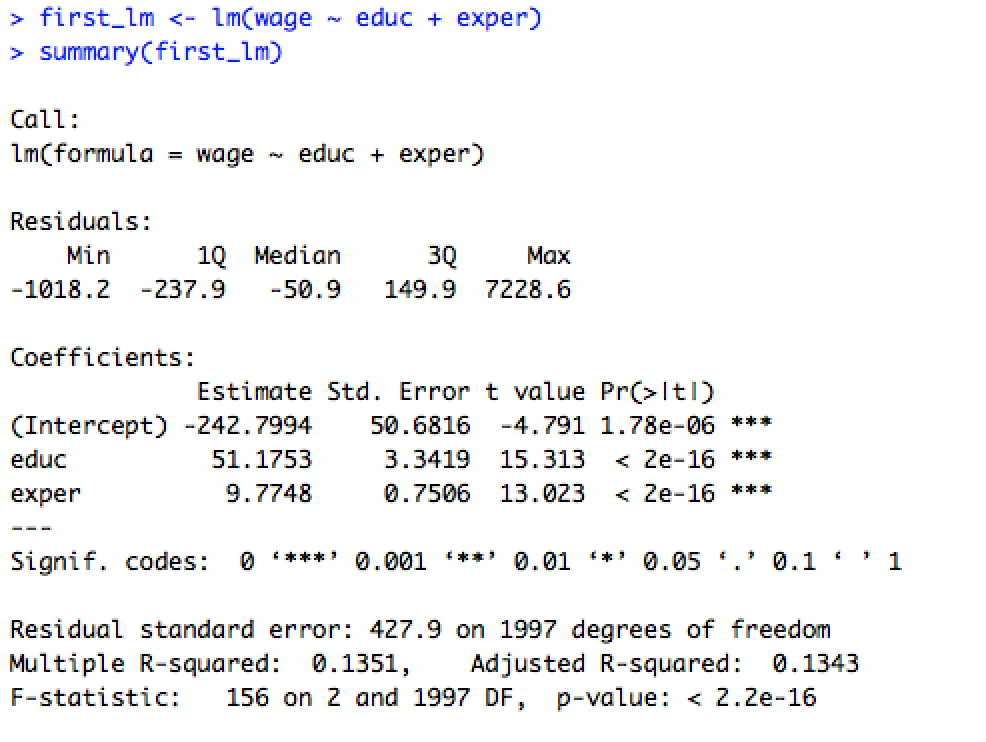
**STATS 500 Homework 2**

Changbai Liu

1. Based on Chapter 2, problems 1 and 2 (page 30)
2. **Fit a regression model with weekly wages as the response and years of education and experience as predictors. Present the output.**



1. **What percentage of variation in the response is explained by these predictors?**

R^2 = 0.1351

1. **Which observation has the largest (positive) residual? Give the case number.**

Lagest residual:

> max(first\_lm$residuals)

[1] 7228.612

case number:

> which.max(first\_lm$residuals)

1576

#1576 has the largest (positive) residual.

1. **Compute the mean and median of the residuals. Explain what the difference between the mean and the median indicates.**

> mean(first\_lm$residuals)

[1] -2.344291e-15

> median(first\_lm$residuals)

[1] -50.86827

The mean is bigger than median and the mean is much closer to the zero point, which indicates that this residual calculate function focus more on the mean.

But the median is about -50, which indicates that this set of residuals has some outliers.

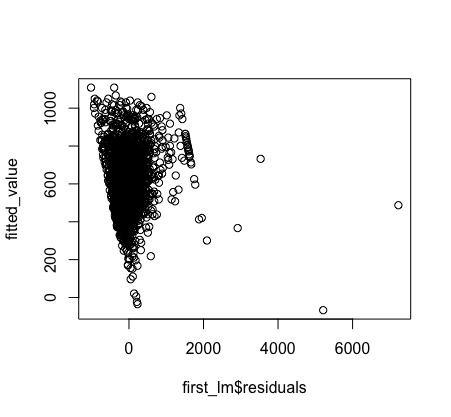
1. **Compute the correlation of the residuals with the fitted values. Plot residuals against fitted values.**

> fitted\_value <- wage - first\_lm$residuals

> cor(first\_lm$residuals,fitted\_value)

[1] 1.322904e-17

> plot(first\_lm$residuals,fitted\_value)



1. **For two people with the same education and one year dierence in experience, what would be the dierence in predicted weekly wages?**

Call:

lm(formula = wage ~ educ + exper)

Coefficients:

(Intercept) educ exper

-242.799 51.175 9.775

The difference would be 9.775 with one year difference of experience.

1. **Fit the same model but with log(weekly wages) as the response and interpret the regression coefficient for experience. Which model has a more natural interpretation?**

> second\_lm <- lm(log(wage) ~ educ + exper)

> second\_lm

Call:

lm(formula = log(wage) ~ educ + exper)

Coefficients:

(Intercept) educ exper

4.65032 0.09051 0.01808

The second model(lm(formula = log(wage) ~ educ + exper)) has a more natural interpretation.