HW7
Osuke Sashida

Problem1

a.

The students' GPA who declared major is positively related to ACT score. But students' GPA who didn't declare major doesn't have relationship with ACT score.

b.

$$Grade = 3.226318 + (-1.649577) * Major - 0.002757 * ACT + 0.062245 * ACT * Major$$

$$Grade = \beta_0 + \beta_1 * Major + \beta_2 * ACT + \beta_3 * ACT * Major$$

c.

$$H_0: \beta_2 + \beta_3 = 0 \qquad H_\alpha: \beta_2 + \beta_3 \neq 0$$
$$t_{statistic} = 3.643$$
$$p_{value} = 0.000490$$

If you set significance level as 0.05, we can reject null hypothesis. It means the students who declared major tend to get more high grade when they get high score in ACT test.

d.

$$Grad_{20Dec} = 3.226318 + (-1.649577) * 1 - 0.002757 * 20 + 0.062245 * 20 * 1 = 2.766501$$

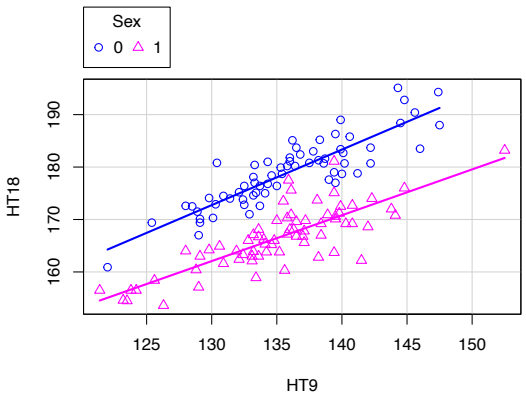
$$Grade_{20not} = 3.226318 - 0.002757 * 20 = 3.171178$$

$$Grade_{30Dec} = 3.226318 + (-1.649577) * 1 - 0.002757 * 30 + 0.062245 * 30 * 1 = 3.361381$$

$$Grade = 3.226318 - 0.002757 * 30 = 3.143608$$

Problem2

1.



$$HT18 = 35.07880 + 13.32748 * Sex - 1.05895 * HT9 + -0.18463 * HT9 * Sex$$

$$HT18 = \beta_0 + \beta_1 * Sex + \beta_2 * HT9 + \beta_3 * HT9 * Sex$$

These data are almost clearly divided into two groups by sex.
And these slope look like almost pararell.

2.

$$H_0: \beta_3 = 0, \quad H_\alpha: \beta_3 \neq 0$$
$$t_{statistics} = -1.722\,, \qquad p_{value} = 0.0875\,.$$

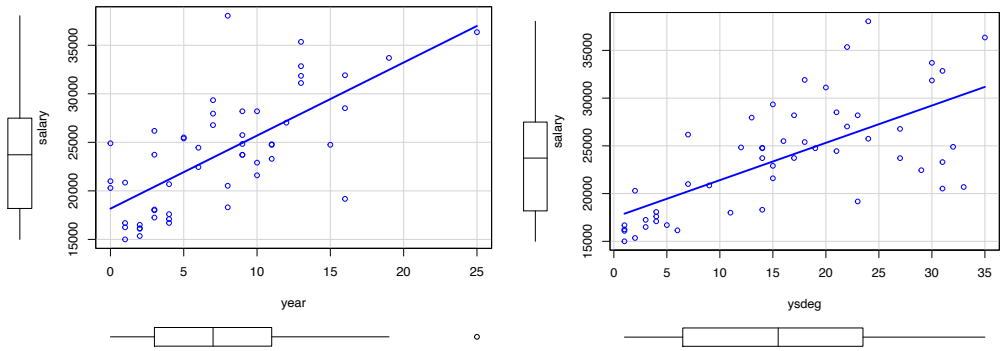
I can't reject null hypothesis. It means β_3 is equal to zero. So, these two lines slope are pararell.

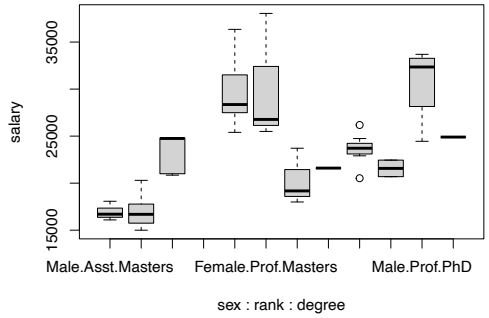
3.

$$confidence\ intervala(-15.4478287\ 42.10279381)$$

Problem3

1.





Year and ysdeg both have positive relationship with salary.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 15746.05     800.18  19.678  < 2e-16 ***
rankAssoc    5292.36    1145.40   4.621 3.22e-05 ***
rankProf    11118.76    1351.77   8.225 1.62e-10 ***
sexFemale    1166.37     925.57   1.260    0.214
degreePhD    1388.61    1018.75   1.363    0.180
ysdeg        -124.57      77.49  -1.608    0.115
year          476.31      94.91   5.018 8.65e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2398 on 45 degrees of freedom
Multiple R-squared:  0.855,     Adjusted R-squared:  0.8357
F-statistic: 44.24 on 6 and 45 DF,  p-value: < 2.2e-16
```

2.

$$H_0: \beta_1 \neq 0, \quad H_\alpha: \beta_1 = 0$$
$$t_{statistics} = 1.260, \qquad p_{value} = 0.214$$

We can't reject null hypothesis. So, the mean of the men and women is not same.

3.

$$confidence\ interval(-697.8183\ \ 3030.56452)$$

4.

Rank is necessary variable for prediction of salary. Because model's pvalue much fall down when we don't use rank.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 17183.57    1147.94  14.969  < 2e-16 ***
sexFemale   -1286.54    1313.09  -0.980 0.332209
degreePhD   -3299.35    1302.52  -2.533 0.014704 *
ysdeg         339.40      80.62   4.210 0.000114 ***
year          351.97     142.48   2.470 0.017185 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3744 on 47 degrees of freedom
Multiple R-squared:  0.6312,    Adjusted R-squared:  0.5998
F-statistic: 20.11 on 4 and 47 DF,  p-value: 1.048e-09
```