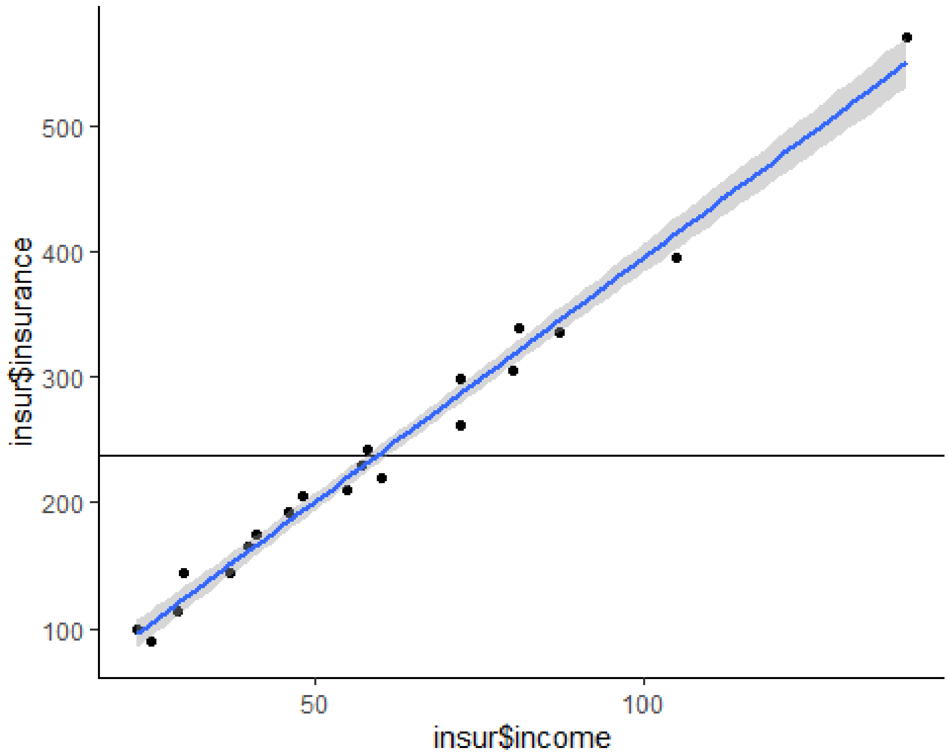
Assignment 3

```r
#if(!require("pacman")) install.packages("pacman")
pacman::p_load(tidyverse, reshape, gplots, RStata,haven)
theme_set(theme_classic())
options("RStata.StataPath")
options("RStata.StataVersion" = 13)

insur<- read_dta(file = "insur.dta")

head(insur)

## # A tibble: 6 x 2
##   insurance income
##       <dbl>  <dbl>
## 1        90     25
## 2       165     40
## 3       220     60
## 4       145     30
## 5       114     29
## 6       175     41

lm1<- lm(insur$insurance~ insur$income,data=insur)
summary(lm1)

##
## Call:
## lm(formula = insur$insurance ~ insur$income, data = insur)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.228 -10.766   2.456  11.295  21.739
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.8550     7.3835   0.928    0.365
## insur$income   3.8802     0.1121  34.606   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.36 on 18 degrees of freedom
## Multiple R-squared:  0.9852, Adjusted R-squared:  0.9844
## F-statistic:  1198 on 1 and 18 DF,  p-value: < 2.2e-16

mean(insur$insurance)

## [1] 236.95
```

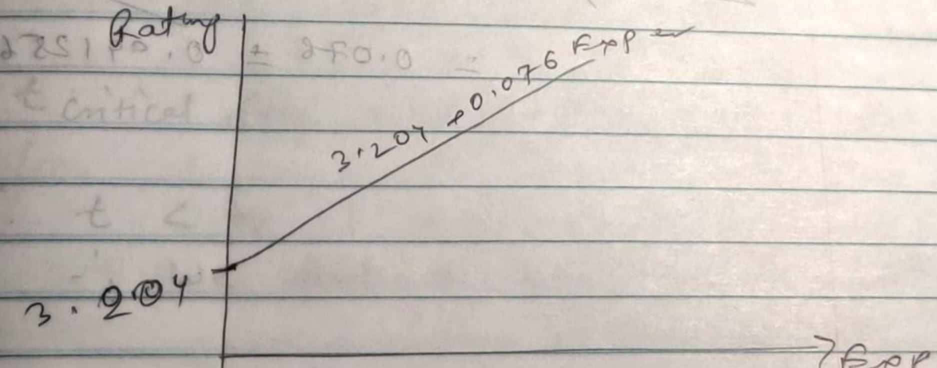
```
ggplot(insur,aes(x=insur$income,y=insur$insurance))+
  geom_point()+
  geom_hline(yintercept=236.95)+
  geom_smooth(method="lm")
```



2.2 $\widehat{Rating} = 3.204 + 0.076\ Exper$

$Rating = \beta_1 + \beta_2 Exper + e$

(1) Rating increases by 0.076 if we increase experience by 1 year.



(2) 95% CI for $\beta_2 = b_2 \pm t_{(0.975, 22)}\ b_2$

$df = 24 - 2 = 22$

$= 0.076 \pm 2.074 \times 0.044$

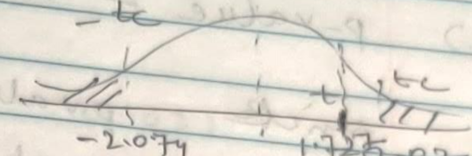
$= 0.076 \pm 0.091256$

$= 0.167256, -0.015256$

(c) $H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$

Two tail test



$t = \dfrac{b_2}{se(b_2)} = \dfrac{0.076}{0.044}$

$= 1.7272$

t critical from previous problem = 2.074

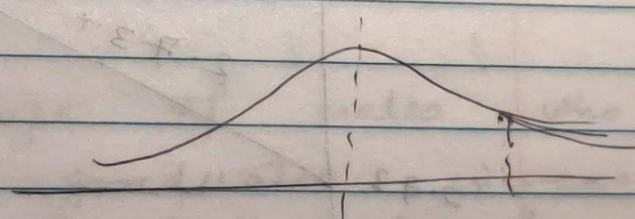
$\therefore$ $t < t_c$

$\therefore$ We fail to reject null hypothesis

d) $H_0 : \beta_2 = 0$

$H_1 : \beta_2 > 0$ One tail test

$t_{0.95, 22} = 1.717$



$b_2 = 0.076$ se = 0.044

$t_{stat} = \dfrac{0.076}{0.044} = 1.727$

tstat > t critical.

$\therefore$ We reject the null hypothesis

(e) p-value = 0.0982 $\alpha = 0.05$

since pvalue $> \alpha$

we cannot reject the null hypothes

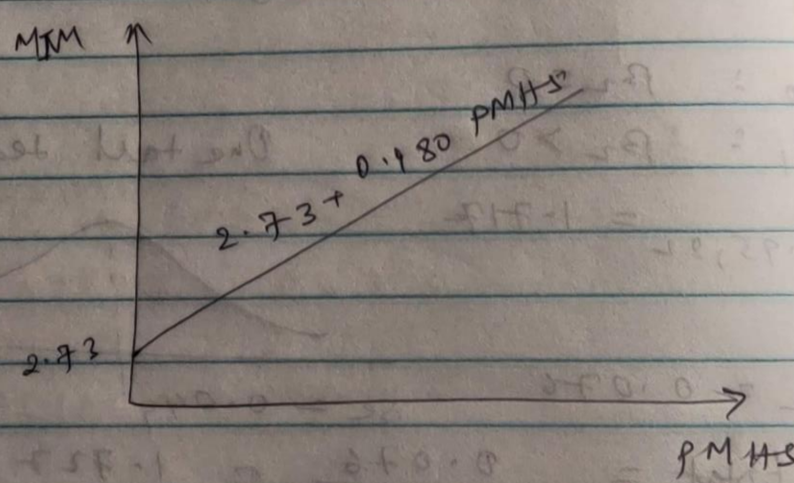
(3.4) $\widehat{MIM} = a + 0.180\,PMHS$

se (2.174) (b)

(t) 1.257 5.754

(a) $a = t_a \times se$

$= 1.257 \times 2.174 = 2.73$

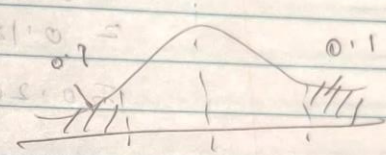


(b) $t_b = \frac{b}{se}$

$\hat{se}_b = \frac{0.180}{5.754} = 0.031$

(c) df = 51 − 2 = 49.

t = 1.257



$t_{(0.2,\ 49)} \sim 1.257$

from t table

p-value (1.257) ~ 0.1

∵ we have two tail distribution.

p-value = 2×0.1 = 0.2

(d) b = 0.180

If the percentage of males who are high school graduates increase by 1, the mean income of the males increase (0.180×1000) dollars.

(e) $CI = b_k \pm t_c \, se(\hat{b}_k)$

$= 0.180 \pm t_{(0.05, 49)} \times 0.031$

$= 0.180 \pm (2.678 \times 0.031)$

$= 0.180 \pm 0.083018$

$= 0.263018, 0.09698$

(f) $H_0 : b = 0.2$

$H_1 : b \neq 0.2$

$t\text{statistics} = \dfrac{0.2 - 0.18}{0.031} = \dfrac{0.02}{0.031}$

$= 0.645$

$t_{critical} = t_{c(0.975, 49)}$

$= 2.009$

$\therefore t_{stat} < t_{critical}$

We will not reject the null hypothesis.

This means that if we increase the number of male who are high school graduates by 1%, the mean income of the male will increase by (0.2×1000) dollars, 200

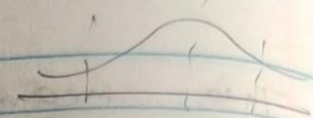
3.5 (a) from the R-code.

$$y = 6.855 + x * 3.8802$$

(b) (1) If the ~~income~~ income increases by thousand dollars then the life insurance amount increases by 3880.2 dollars.

(b)(ii) $\hat{se}(b_2) =$

$$\frac{CI - b_k}{t_c} = \hat{se}(b_k)$$



(c) $H_0 : b_2 = 5$ $\qquad \alpha = 0.05$

$H_1 : b_2 \neq 5$

$df = 20-2=$

$$t_{stat} = \frac{5-3.88}{0.1121} = 9.99$$

$$t_c = t(0.975, 18) = 2.101$$

$t_{stat} > t_c$

$\therefore$ We will reject the null

(d) $H_0 : b_2 = 1$

$H_1 : b_2 \neq 1$ two tail test

$$t_{stat} = \frac{3.88 - 1}{0.1121} = \frac{2.88}{0.1121}$$

$$= 25.6913$$

$$t_c = t(0.975, 18) = 2.101$$

$t_{stat} > t_c$

$\therefore$ We will reject the null.

(e) $\widehat{insurance} = 6.855 + 3.8802 * \widehat{income}$

→ The intercept is 6.855. That means even if the income is zero, the insurance amount will be 6.855.

→ The probability of income estimate for 2.8802 is ($2e^{-16}$) which is almost equal to zero. for any value to be significant we want p-value < 0.05, therefore this result is significant.

3.7 a) $H_0 : \beta_i = 1$

$H_1 : \beta_i \neq 1$

Disney : $b_2 = 0.894$

$\hat{se} = 0.123$

$$tstat = \frac{1 - 0.894}{0.123} = 0.862$$

$tc = t_{(0.975, 130)} = 1.978$

$tstat < tc$

∴ ~~we cannot rej~~

$\therefore$ we fail to reject the null

GE

$$t_{stat} = \frac{0.901 - 1}{0.0982} = -1.008$$

$-t_c \le t_{stat} < t_c$

$\therefore$ We fail to reject the null

IBM :

$$t_{stat} = \frac{1.187 - 1}{0.126} = 1.484$$

$-t_c < t_{stat} < t_c$

$\therefore$ we fail to reject the null

GM

$$t_{stat} = \frac{1.267 - 1}{0.201} = 1.328$$

$-t_c < t_{stat} < t_c$, We fail to reject the null

Exon

$$t_{stat} = \frac{0.413 - 1}{0.0895} = -6.558$$

$t_{stat} > t_c$ we reject the null

MS:

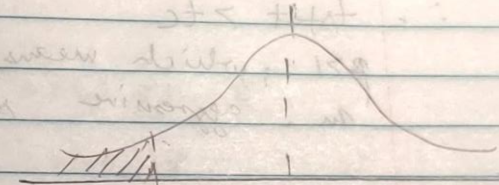
$$t_{stat} = \frac{1.318 - 1}{0.16}$$

$t_{stat} > t_c$

$\therefore$ we reject the null.

3.7) (b) $H_0 : \beta \geq 1$ Exxon

$H_1 : \beta < 1$



$$t_{stat} = \frac{0.412 - 1}{0.0875}$$

$$= -6.558$$

$t_c = -1.657$

$t_{stat} < t_c$ we reject the null

$\therefore \beta < 1$, which means that there is less variation than the market variation.

(c) $H_0 : \beta \leq 1$
$H_1 : \beta > 1$

$$tstat = \frac{1.318 - 1}{0.16} = 1.98'75$$

$t_c = 1.657$

$\therefore$ $tstat > t_c$ we reject the null.
$\beta > 1$, which means that Microsoft is an aggressive stock

(d) $CI = b_k \pm t_c * se(\hat{b_k})$
$= 1.318 \pm 1.978 * 0.160$

$UCI = 1.634$
$LCI = 1.002$

$\beta \in [1.002, 1.634]$

We have 95% confidence that β values will lie between 1.002 and 1.634.

(c) $H_0: b_1 = 0$

$H_1: b_1 \neq 0$

Disney

$$t_{stat} = \frac{-0.00354 - 0}{0.00597} = -0.593$$

GE

$$t_{stat} = \frac{-0.003578 - 0}{0.0048} = -0.745$$

GM

$$t_{stat} = \frac{-0.01494 - 0}{0.00974} = -1.534$$

IBM

$$t_{stat} = \frac{0.002875 - 0}{0.006093} = 0.439.$$

MS

$$t_{stat} = \frac{0.00257 - 0}{0.00775}$$

$$= 0.3316$$

Exxon

$$t_{stat} = \frac{0.006\cancel{7}76 - 0}{0.00433}$$

$$= 1.565$$

$-t_c < t_{stat} < t_c$

$\therefore$ We cannot reject the null hypoth

8.9) a) $\widehat{vote} = 51.8908 + 0.6545\ \widehat{growth}$

$H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$ two tail

$$tstat = \frac{0.6545 - 0}{0.1611}$$

$$= 4.0626$$

$df = 33 - 2 = 31$

$tc = t_{(0.975, 31)} = 2.040$

$(4.0626 > 2.040)$

(tstat > tcritical). Hence we can reject the null

) CI for $\beta_2 = b_k \pm t_c\ \widehat{se}(b_k)$

$= 0.6545 \pm 2.040 * 0.1611$

$= [0.983, 0.326]$

∴ We have 95% confidence that β value will be in the range $[0.326, 0.983]$

(c) $\widehat{vote} = 53.3 - 0.4502\ \widehat{inflation}$

$H_0: \beta_2 = 0$

$H_1: \beta_2 < 0$ one tail

$$tstat = \frac{-0.4502 - 0}{0.5101}$$

$$= -0.8825$$

$$t_c = t_{(0.05, 22)} = -1.717$$

$tstat < t_c$

∴ We cannot reject the null

This means that we don't have enough evidence that the inflation has negative effect on vote.

(d) with increase in inflation rate by 1 unit, the voting will decrease by 0.4502.

$$CI = b_k \pm t_c\ \widehat{se}(b_k)$$

$$= -0.4502 \pm t_{(0.975, 22)} * 0.5101$$

$$= -1.508, 0.924$$

e) $H_0 : \beta_2 \geq 50$
$H_1 : \beta_2 < 50$

$$t_{stat} = \frac{53.299 - 50}{1.7231} = 1.915$$

$t_c = -1.717$

$t_{stat} < t_c$
$\therefore$ we fail to reject the null.

f) $E(\hat{vote}) = b_1 + 2b_2$
$= 53.4077 + 2 \times (-0.44431)$
$= 52.5191$

$var(\hat{b}_1 + 2\hat{b}_2) = var(\hat{b}_1) + 2^2 var(\hat{b}_2) +$
$2 \cdot 2 cov(\hat{b}_1, \hat{b}_2)$

$= 5.0625 + 4 \times (0.3593) +$
4×-1.0592

$= 2.26153$

95% CI =

$(b_1 + 2b_2) \pm (t_{0.975, 22})\, se(b_1 + 2b_2)$

$= 52.5191 \pm 2.074 \sqrt{2.2653}$

$= 52.5191 \pm 3.1246$

$= (49.40, 55.64)$