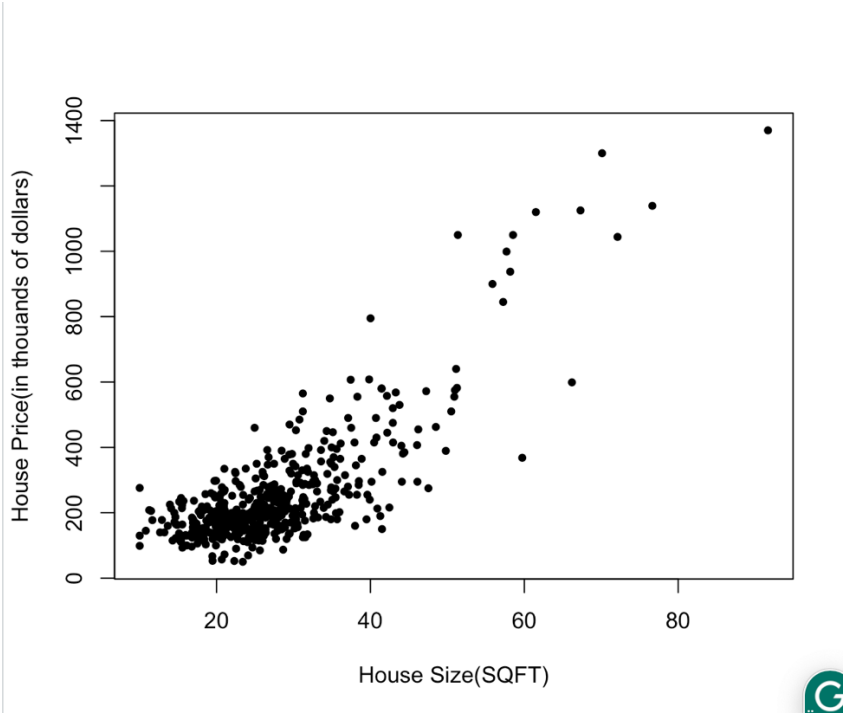
2.17

敘述統計

```
     price             sqft             age              pool          fireplace
 Min.   :  50.0   Min.   :10.00   Min.   : 1.00   Min.   :0.000   Min.   :0.000
 1st Qu.: 157.9   1st Qu.:20.56   1st Qu.: 7.00   1st Qu.:0.000   1st Qu.:0.000
 Median : 199.9   Median :25.55   Median : 9.00   Median :0.000   Median :1.000
 Mean   : 250.2   Mean   :27.28   Mean   : 8.19   Mean   :0.064   Mean   :0.546
 3rd Qu.: 279.0   3rd Qu.:31.20   3rd Qu.:10.00   3rd Qu.:0.000   3rd Qu.:1.000
 Max.   :1370.0   Max.   :91.67   Max.   :11.00   Max.   :1.000   Max.   :1.000
     close          twostory        occupied
 Min.   :0.000   Min.   :0.000   Min.   :0.00
 1st Qu.:0.000   1st Qu.:0.000   1st Qu.:0.00
 Median :0.000   Median :0.000   Median :1.00
 Mean   :0.378   Mean   :0.138   Mean   :0.56
 3rd Qu.:1.000   3rd Qu.:0.000   3rd Qu.:1.00
 Max.   :1.000   Max.   :1.000   Max.   :1.00
```

#a



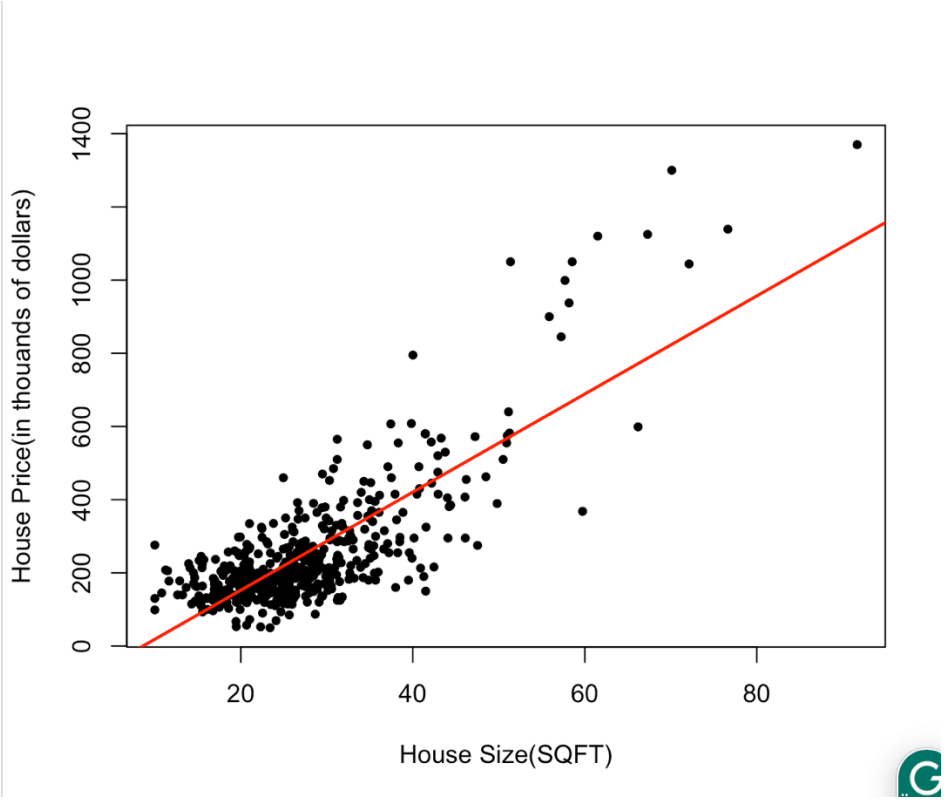
#b

b1=-115.4236

b2=13.4029

$$\widehat{PRICE} = -115.4236 + 13.4029 \times SQFT$$

當 SQFT=0，PRICE 為-115.4236，如果 SQFT 增加 1unit (100 平方英尺),PRICE 估計值增加 13.4029 unit ($1000)




```
Call:
lm(formula = price ~ sqft, data = collegetown)

Residuals:
    Min      1Q  Median      3Q     Max
-316.93  -58.90   -3.81   47.94  477.05

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -115.4236    13.0882  -8.819   <2e-16 ***
sqft          13.4029     0.4492  29.840   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 102.8 on 498 degrees of freedom
Multiple R-squared:  0.6413,    Adjusted R-squared:  0.6406
F-statistic: 890.4 on 1 and 498 DF,  p-value: < 2.2e-16
```

#c

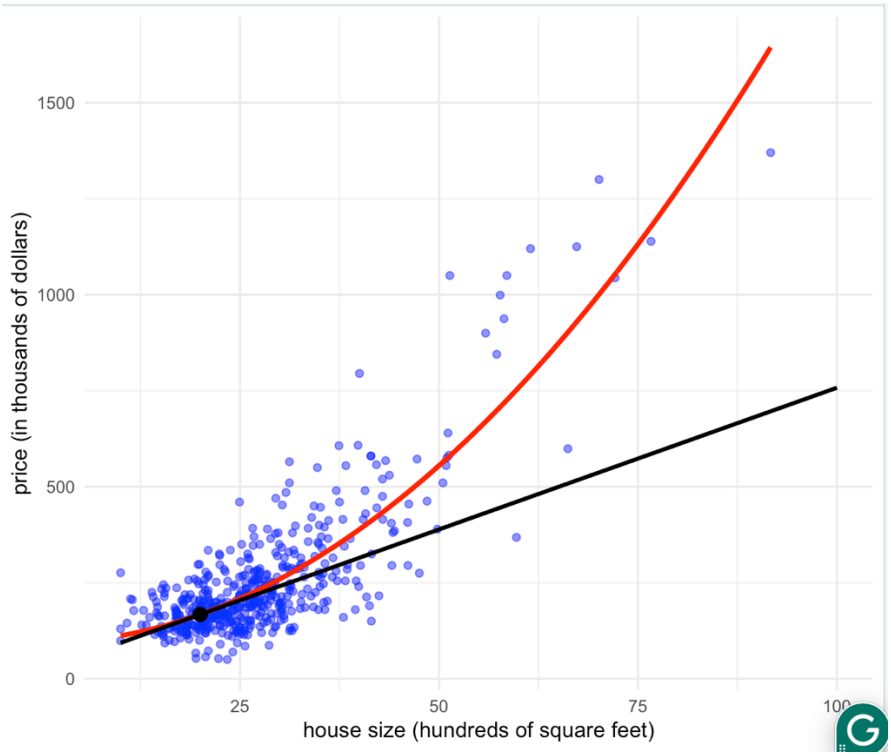
在房屋面積為 2000 平方英尺(sqft=20)的情況下，增加 100 平方英尺的邊際效果為

$\widehat{PRICE} = 93.56585 + 0.184519 \times SOFT^2$

$\frac{dPRICE}{dSOFT}$ = 2*a2*SOFT = 2*0.184519*20 = 7.38076

```
Call:
lm(formula = price ~ I(sqft^2), data = collegetown)

Residuals:
    Min      1Q  Median      3Q     Max
-383.67  -48.39   -7.50   38.75  469.70

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 93.565854   6.072226   15.41   <2e-16 ***
I(sqft^2)    0.184519   0.005256   35.11   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 92.08 on 498 degrees of freedom
Multiple R-squared:  0.7122,    Adjusted R-squared:  0.7117
F-statistic:  1233 on 1 and 498 DF,  p-value: < 2.2e-16
```

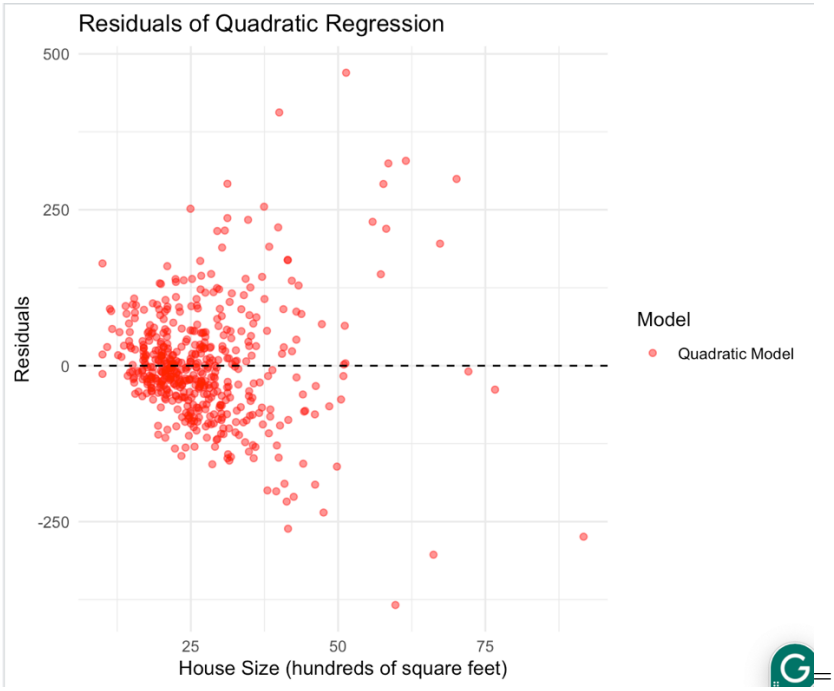
#d



#e

```
> elasticity <- ME * (sqft_2000 / price_2000)
> print(elasticity)
I(sqft^2)
0.8819511
```

#f



#g.

SSE(b)= 5262847

SSE(c)= 4222356

$$\widehat{PRICE} = 93.56585 + 0.184519 \times SOFT^2$$

SSE 較小，模型較好

```
print(c(SSE_Linear = sse_lm, SSE_Quadratic = sse_quad))
 SSE_Linear SSE_Quadratic
    5262847       4222356
```

2.25

#a

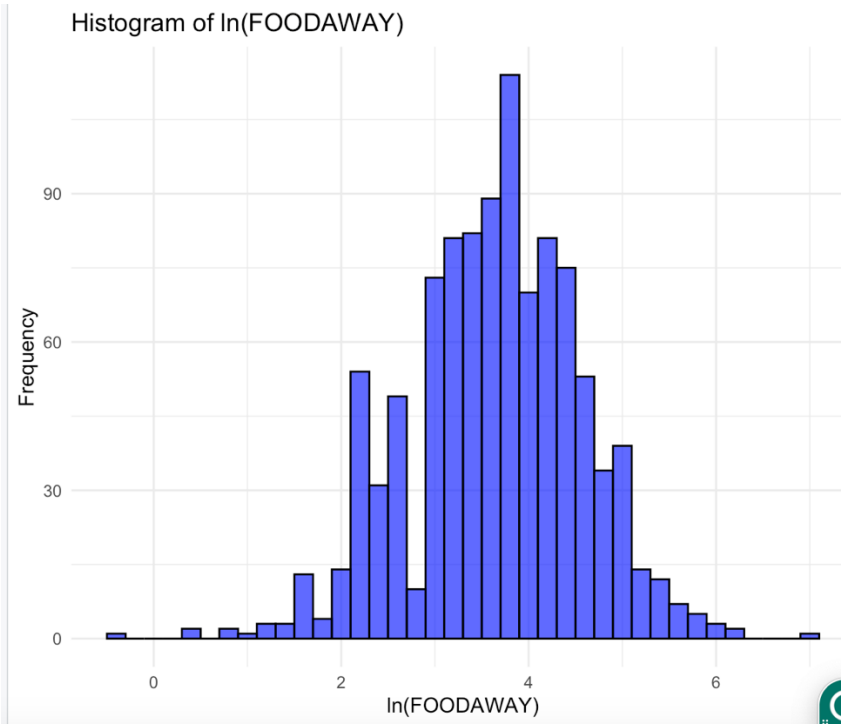
```
> summary(cex5_small$foodaway)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00   12.04   32.55   49.27   67.50 1179.00
```

```
> cat("25th percentile:     ", quantiles_25_75[1], "\n")
25th percentile:      12.04
> cat("75th percentile:     ", quantiles_25_75[2], "\n")
75th percentile:      67.5025
```

#b

```
> print(education_stats)
# A tibble: 3 × 3
  education_group   Mean_Foodaway Median_Foodaway
  <chr>                     <dbl>           <dbl>
1 Advanced Degree            73.2            48.2
2 College Degree             48.6            36.1
3 No College Degree          39.0            26.0
```

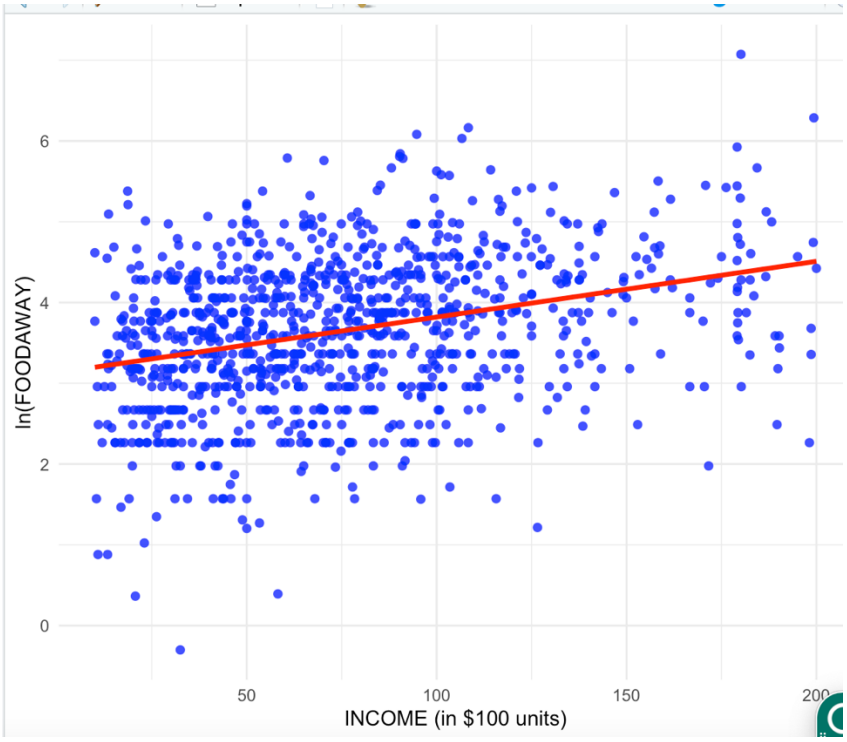
#c



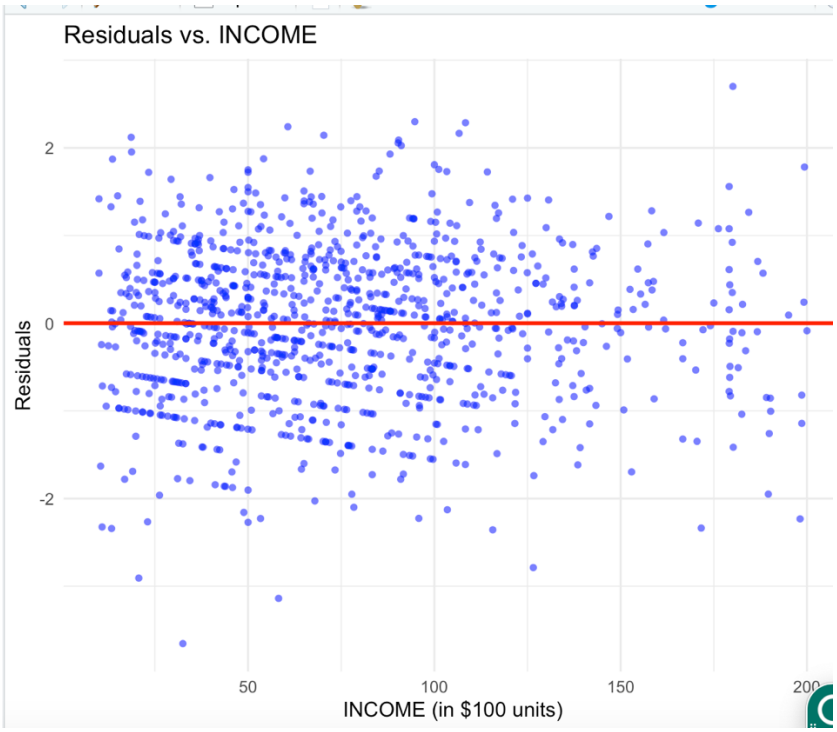
#d

```
Residuals:
    Min      1Q  Median      3Q     Max
-3.6547 -0.5777  0.0530  0.5937  2.7000

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.1293004  0.0565503   55.34   <2e-16 ***
income      0.0069017  0.0006546   10.54   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8761 on 1020 degrees of freedom
Multiple R-squared:  0.09826,   Adjusted R-squared:  0.09738
F-statistic: 111.1 on 1 and 1020 DF,  p-value: < 2.2e-16
```

#e

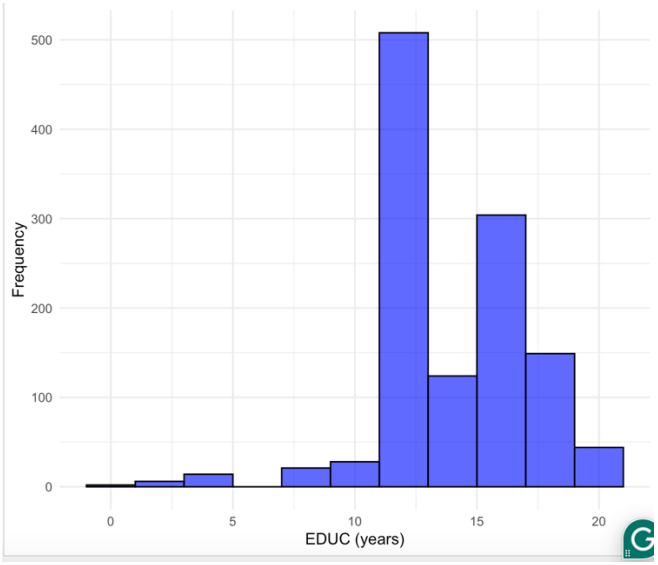


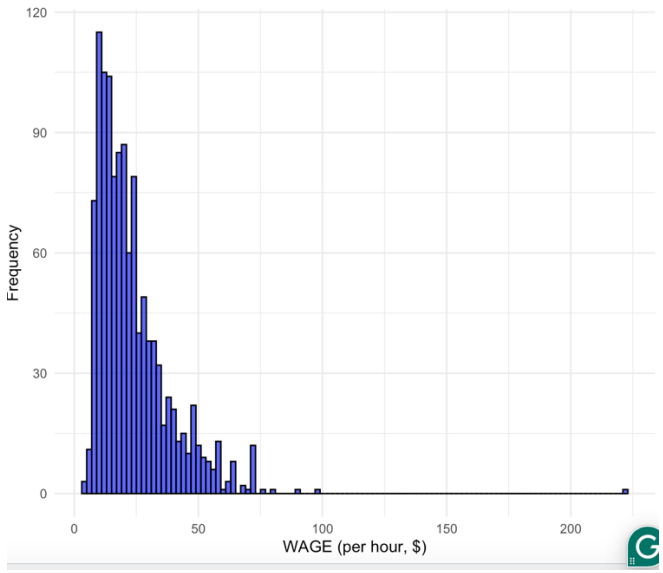
#f The least squares residuals=782.9716



#2.28

#a

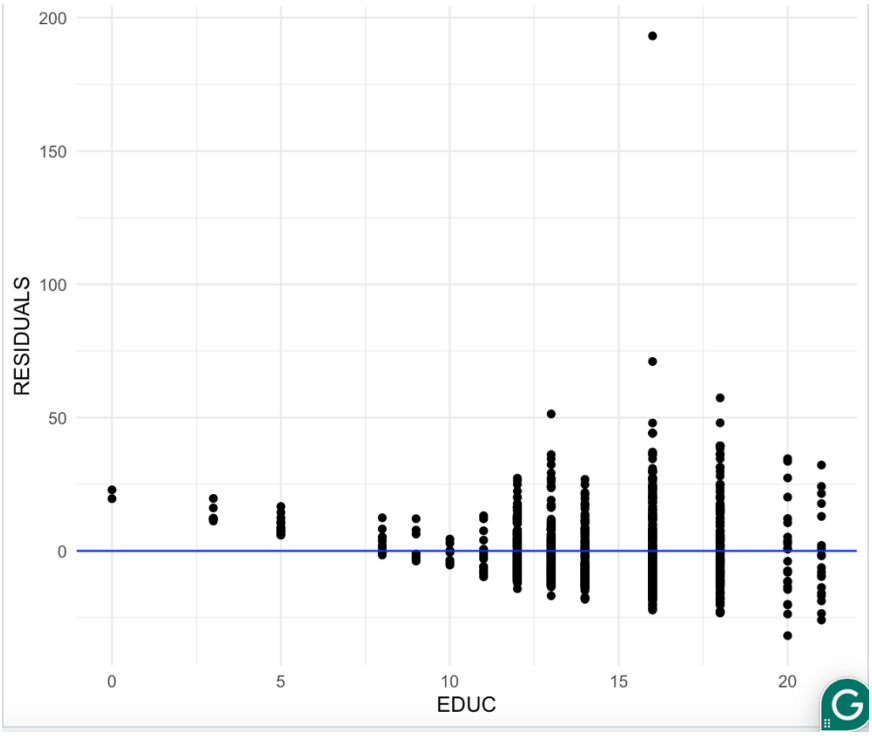




#b

```
Residuals:
    Min      1Q  Median      3Q     Max
-31.785  -8.381  -3.166   5.708 193.152

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -10.4000     1.9624    -5.3 1.38e-07 ***
educ          2.3968     0.1354    17.7  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.55 on 1198 degrees of freedom
Multiple R-squared:  0.2073,    Adjusted R-squared:  0.2067
F-statistic: 313.3 on 1 and 1198 DF,  p-value: < 2.2e-16
```

#c



#d

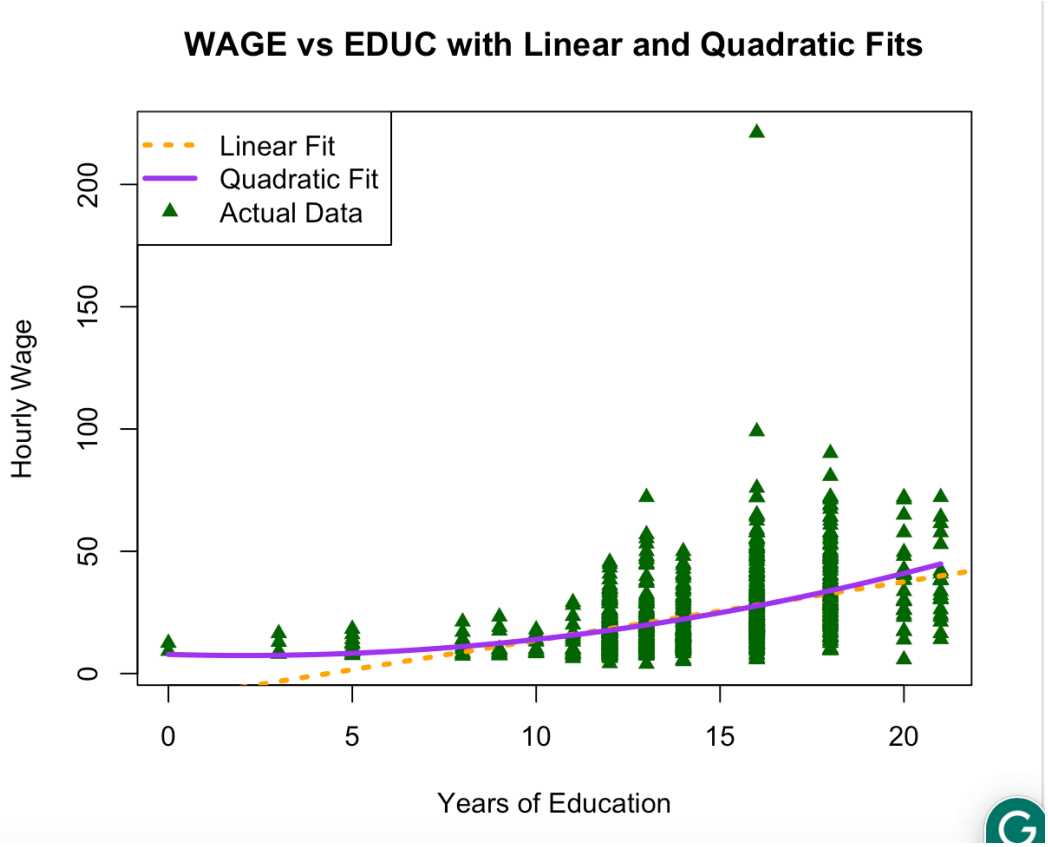
```
> summary(black_model)

Call:
lm(formula = wage ~ educ, data = cps5_small, subset = (black ==
    1))

Residuals:
    Min      1Q  Median      3Q     Max
-15.673  -6.719  -2.673   4.321  40.381

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -6.2541     5.5539  -1.126    0.263
educ          1.9233     0.3983   4.829 4.79e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.51 on 103 degrees of freedom
Multiple R-squared:  0.1846,    Adjusted R-squared:  0.1767
F-statistic: 23.32 on 1 and 103 DF,  p-value: 4.788e-06

> summary(white_model)

Call:
lm(formula = wage ~ educ, data = cps5_small, subset = (black ==
    0))

Residuals:
    Min      1Q  Median      3Q     Max
-32.131  -8.539  -3.119   5.960 192.890

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -10.475      2.081  -5.034  5.6e-07 ***
educ           2.418      0.143  16.902  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.79 on 1093 degrees of freedom
Multiple R-squared:  0.2072,    Adjusted R-squared:  0.2065
F-statistic: 285.7 on 1 and 1093 DF,  p-value: < 2.2e-16
```

```
>
> summary(male_model)

Call:
lm(formula = wage ~ educ, data = cps5_small, subset = (female ==
    0))

Residuals:
    Min      1Q  Median      3Q     Max
-27.643  -9.279  -2.957   5.663 191.329

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -8.2849     2.6738  -3.099  0.00203 **
educ          2.3785     0.1881  12.648  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.71 on 670 degrees of freedom
Multiple R-squared:  0.1927,    Adjusted R-squared:  0.1915
F-statistic:   160 on 1 and 670 DF,  p-value: < 2.2e-16

> summary(female_model)

Call:
lm(formula = wage ~ educ, data = cps5_small, subset = (female ==
    1))

Residuals:
    Min      1Q  Median      3Q     Max
-30.837  -6.971  -2.811   5.102  49.502

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -16.6028     2.7837  -5.964 4.51e-09 ***
educ          2.6595     0.1876  14.174  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.5 on 526 degrees of freedom
Multiple R-squared:  0.2764,    Adjusted R-squared:  0.275
F-statistic: 200.9 on 1 and 526 DF,  p-value: < 2.2e-16
```

#e

```
Call:
lm(formula = wage ~ I(educ^2), data = cps5_small)

Residuals:
    Min      1Q  Median      3Q     Max
-34.820  -8.117  -2.752   5.248 193.365

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.916477   1.091864   4.503 7.36e-06 ***
I(educ^2)   0.089134   0.004858  18.347  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.45 on 1198 degrees of freedom
Multiple R-squared:  0.2194,    Adjusted R-squared:  0.2187
F-statistic: 336.6 on 1 and 1198 DF,  p-value: < 2.2e-16

> alpha2 <- coef(model_quad)["I(educ^2)"]
>
>
> #在 12 年與 16 年教育
> ME_12 <- 2 * alpha2 * 12
> ME_16 <- 2 * alpha2 * 16
>
> cat("Marginal Effect at 12 years of education:", ME_12, "\n")
Marginal Effect at 12 years of education: 2.139216
> cat("Marginal Effect at 16 years of education:", ME_16, "\n")
Marginal Effect at 16 years of education: 2.852288
>
>
```

#f


WAGE vs EDUC with Linear and Quadratic Fits
Linear Fit
Quadratic Fit
Actual Data
Hourly Wage
0
50
100
150
200
0
5
10
15
20
Years of Education