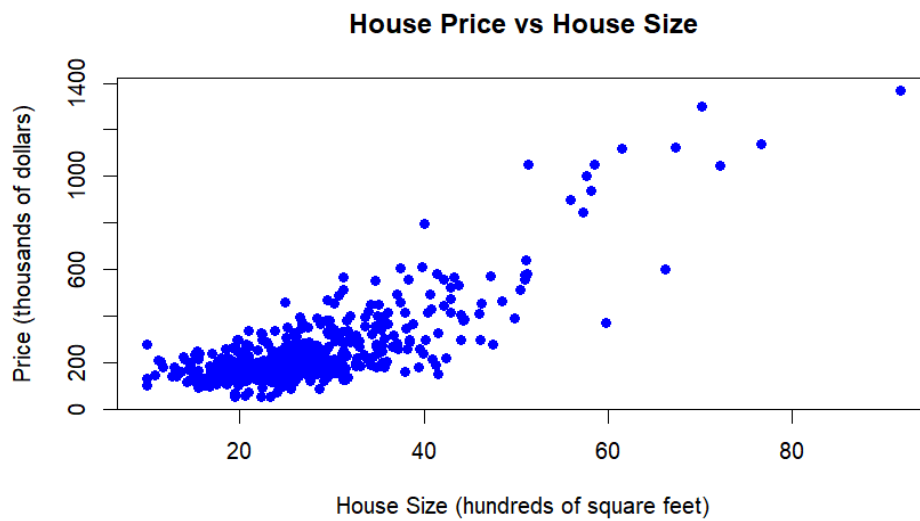


Q17

(a)



(b)右圖紅線為OLS迴歸線

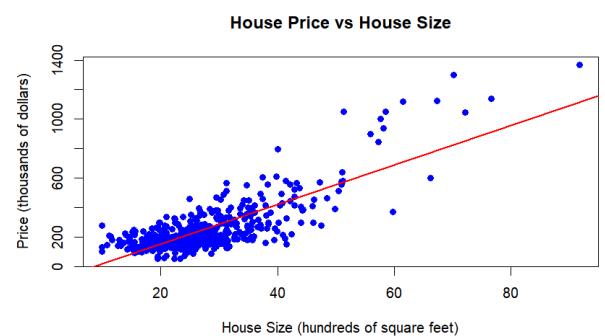
```

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

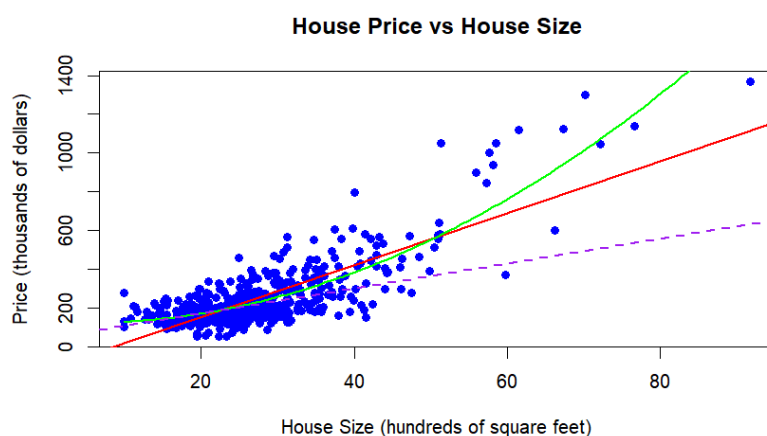
```



(c)

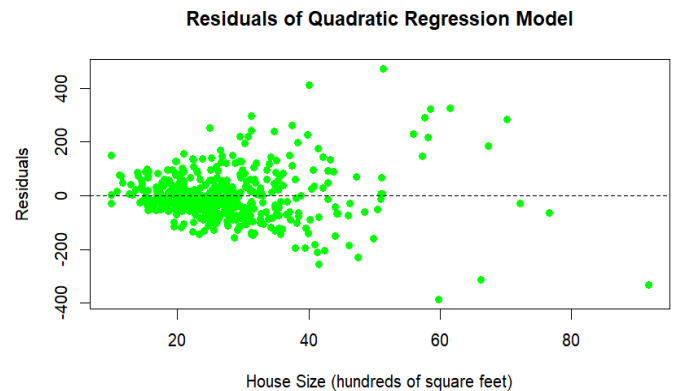
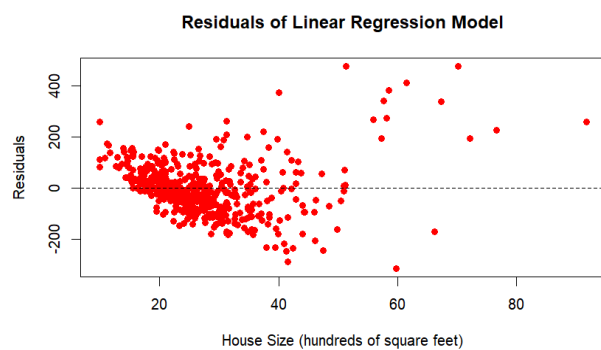
邊際效應為: 6.448092

(d)圖中綠線為二次迴歸線, 紫色為2000平方英尺之切線



(e) Quadratic Regression 在2000平方英尺之彈性為 3.782625

(f)左圖為OLS回歸之殘差, 右圖為二次回歸之殘差



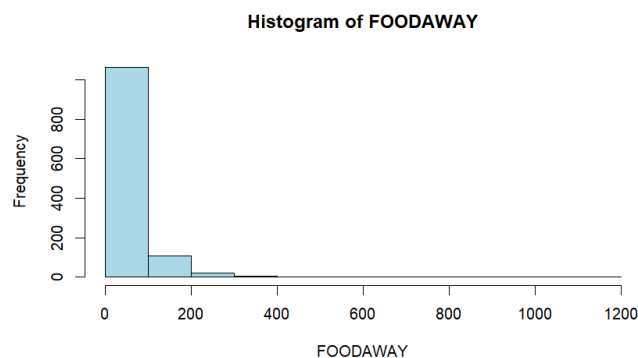
兩者殘差分布稍微有些不同, 可能違反假設SR3:Conditional Homoskedasticity, 即在給定x的情況下, 殘差之變異數應該為定值(constant)。

(g)利用殘差總和(Sum of squared residuals, SSE)選出較適合的模型

```
> sse_lm
[1] 5262847
> sse_quad
[1] 4207791
>
> # 比較 SSE, SSE 較小的模型較適合
> better_model <- ifelse(sse_lm < sse_quad, "Linear Model", "Quadratic Model")
> better_model
[1] "Quadratic Model"
```

Q25.

(a)繪製資料直方圖以及統計量和25,75百分位

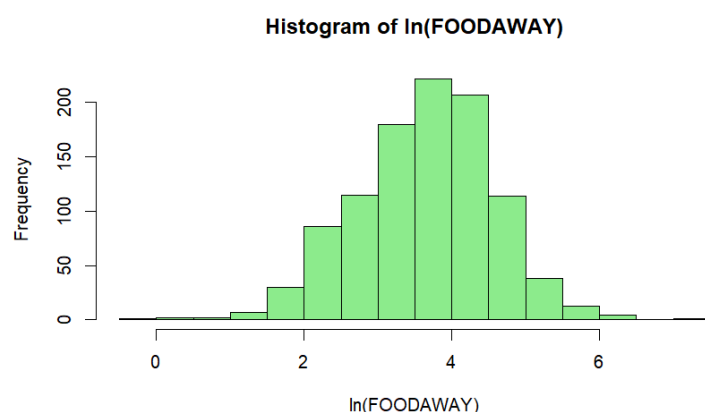


```
> summary(data$foodaway)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00   12.04   32.55   49.27   67.50  1179.00
> quantile(data$foodaway, probs = c(0.25, 0.75))
 25%    75%
12.0400 67.5025
```

(b)根據學歷進行分群, 計算中位數與平均

```
> cat("Mean FOODAWAY for households with an advanced degree:", mean_advanced, "\n")
Mean FOODAWAY for households with an advanced degree: 73.15494
> cat("Median FOODAWAY for households with an advanced degree:", median_advanced, "\n")
Median FOODAWAY for households with an advanced degree: 48.15
> cat("Mean FOODAWAY for households with a college degree:", mean_college, "\n")
Mean FOODAWAY for households with a college degree: 48.59718
> cat("Median FOODAWAY for households with a college degree:", median_college, "\n")
Median FOODAWAY for households with a college degree: 36.11
> cat("Mean FOODAWAY for households with no degree:", mean_no_degree, "\n")
Mean FOODAWAY for households with no degree: 39.01017
> cat("Median FOODAWAY for households with no degree:", median_no_degree, "\n")
Median FOODAWAY for households with no degree: 26.02
```

(c)計算取log後的統計量及直方圖, 應取log後會出現-Inf, 無法納入回歸, 應從樣本剷除。



```
> summary(data$ln_foodaway)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -Inf   2.488   3.483   -Inf   4.212   7.072
```

(d)

Residuals:

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

Coefficients:

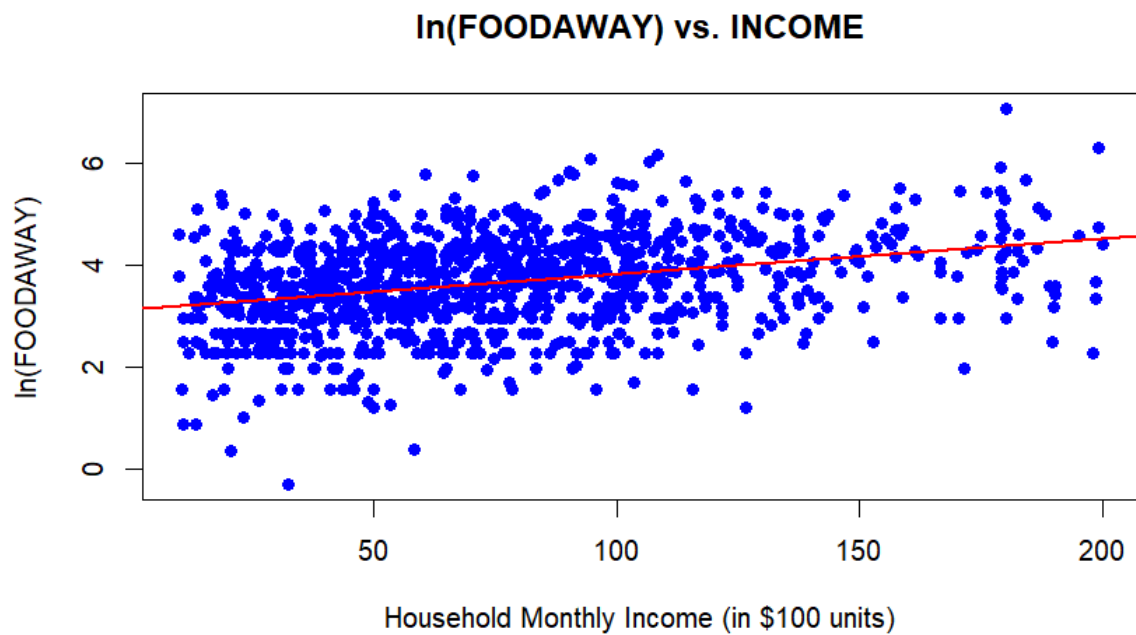
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	3.1293004	0.0565503	55.34	<2e-16 ***
income	0.0069017	0.0006546	10.54	<2e-16 ***

---

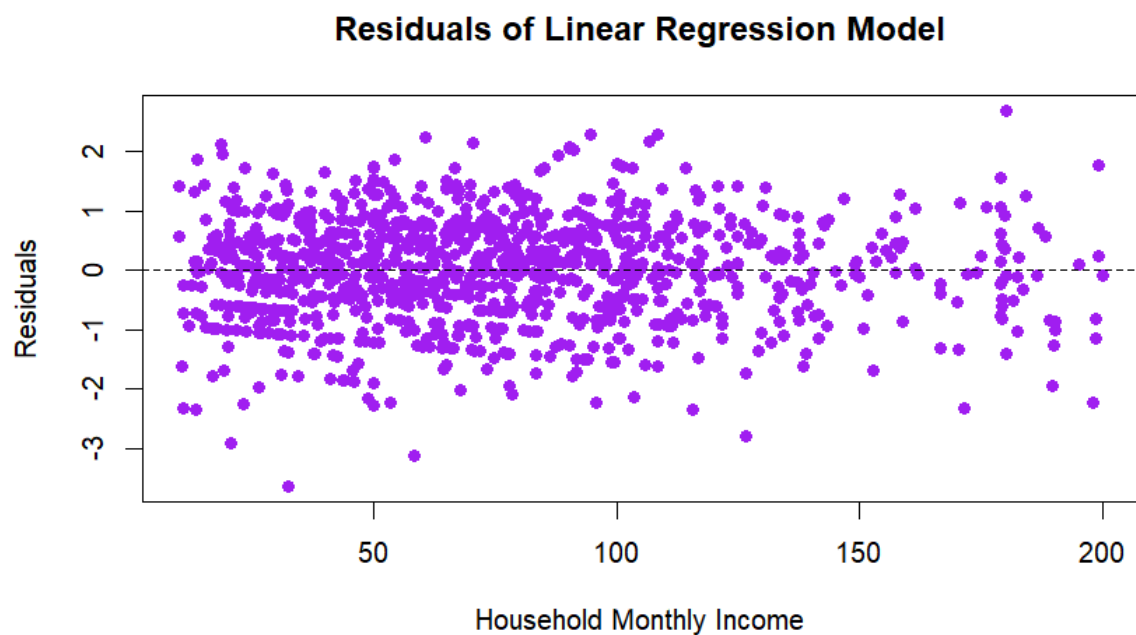
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

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)紅色直線為迴歸線



(f)

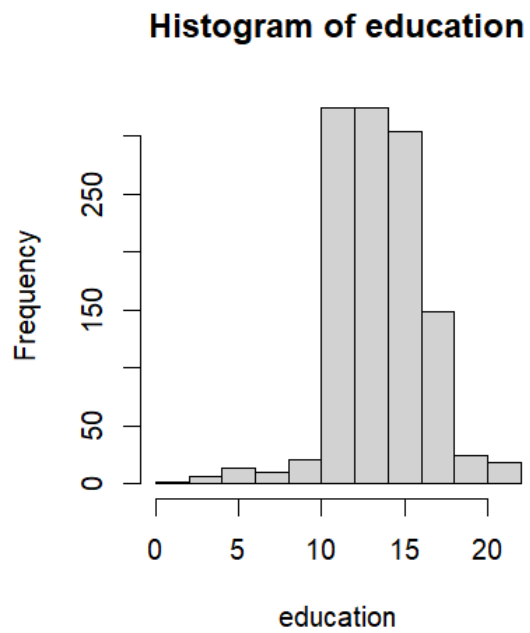
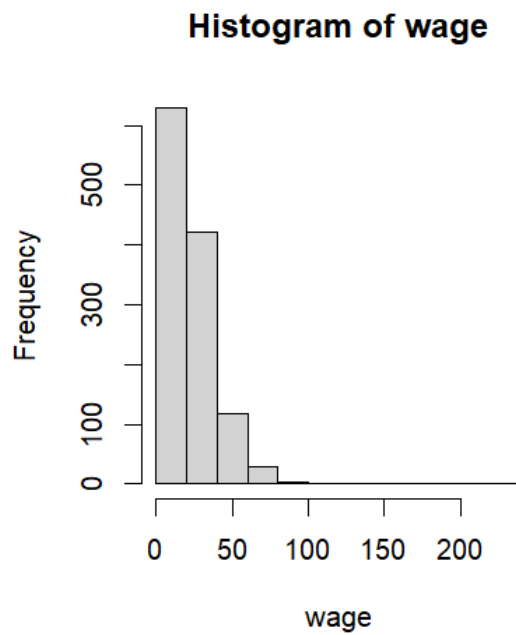


Q28.

(a)計算教育程度與薪資這兩個變數的統計量，教育程度為整數離散的變數，薪資全距則較大

```
> summary(data$wage)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.94  13.00   19.30   23.64  29.80   221.10

> summary(data$educ)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0   12.0   14.0   14.2   16.0   21.0
```



(b)計算回歸結果

Residuals:

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

Coefficients:

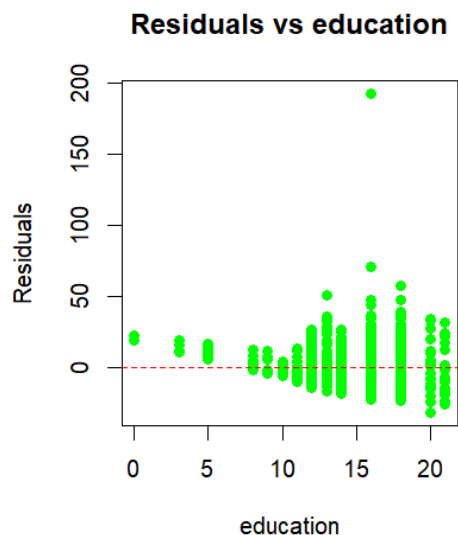
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-10.4000	1.9624	-5.3	1.38e-07 ***
educ	2.3968	0.1354	17.7	< 2e-16 ***

---

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)教育程度在16年的人，薪資的震盪幅度最大，表示變異數可能也是最大的一群。



(d)由上而下分別為男性、女性、黑人、白人之回歸結果

Call:

```
lm(formula = wage ~ educ, data = data[data$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

Call:

```
lm(formula = wage ~ educ, data = data[data$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

```
Call:
lm(formula = wage ~ educ, data = data[data$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
```

```
Call:
lm(formula = wage ~ educ, data = data[data$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
```

#### (e) Quadratic Regression

```
Residuals:
    Min       1Q   Median       3Q      Max
-35.219  -8.047  -2.708   5.307 193.439

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.82200     4.62512   1.691  0.0911 .
educ        -0.42951     0.66438  -0.646  0.5181
I(educ^2)    0.10434     0.02402   4.344 1.52e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.45 on 1197 degrees of freedom
Multiple R-squared:  0.2196,    Adjusted R-squared:  0.2183
F-statistic: 168.4 on 2 and 1197 DF,  p-value: < 2.2e-16
```

與線性回歸比較marginal effect, 可以看到教育程度越高, 邊際效應越大, 但在線性回歸為一常數。

```
> cat("Marginal effect at 12 years of education for quadratic:", marginal_effect_12_q, "\n")
Marginal effect at 12 years of education for quadratic: 2.50421
> cat("Marginal effect at 16 years of education for quadratic:", marginal_effect_16_q, "\n")
Marginal effect at 16 years of education for quadratic: 3.338946
> cat("Marginal effect for linear:", marginal_effect_l, "\n")
Marginal effect for linear: 2.396761
```

(f) Quadratic Model更適合用於預測這些資料。

