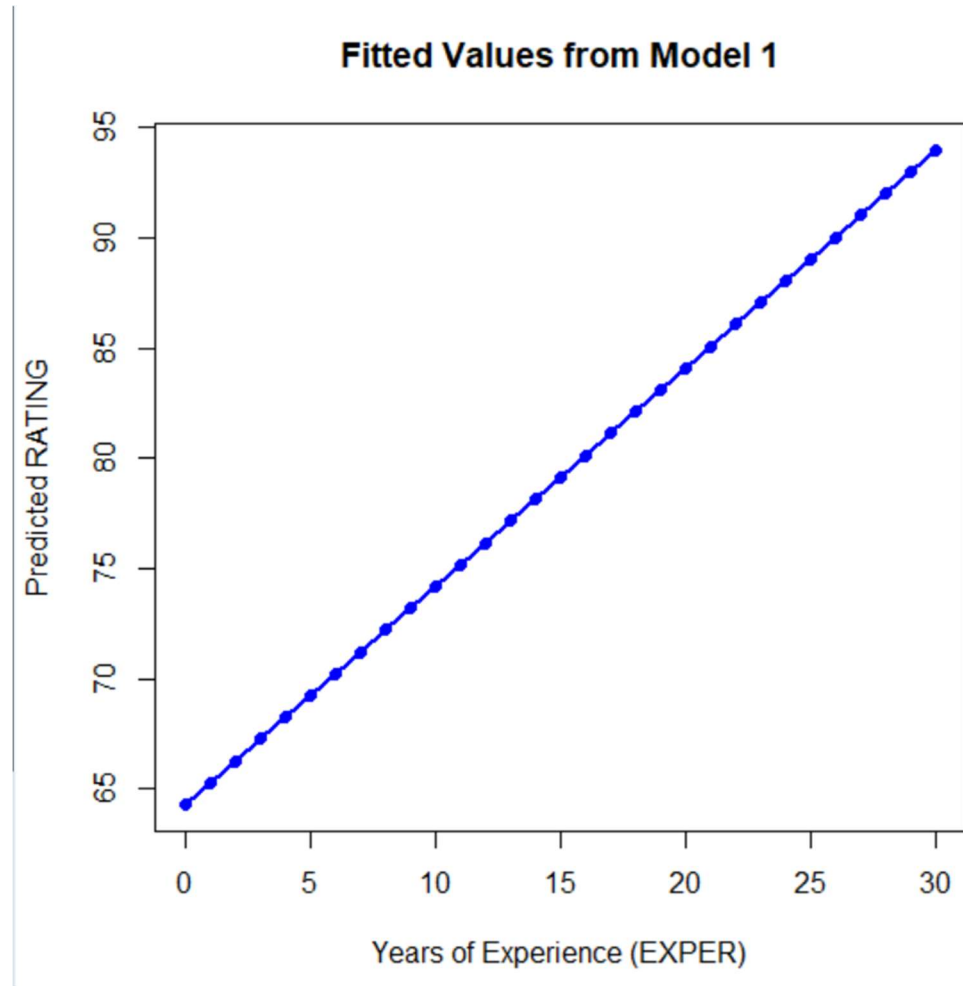


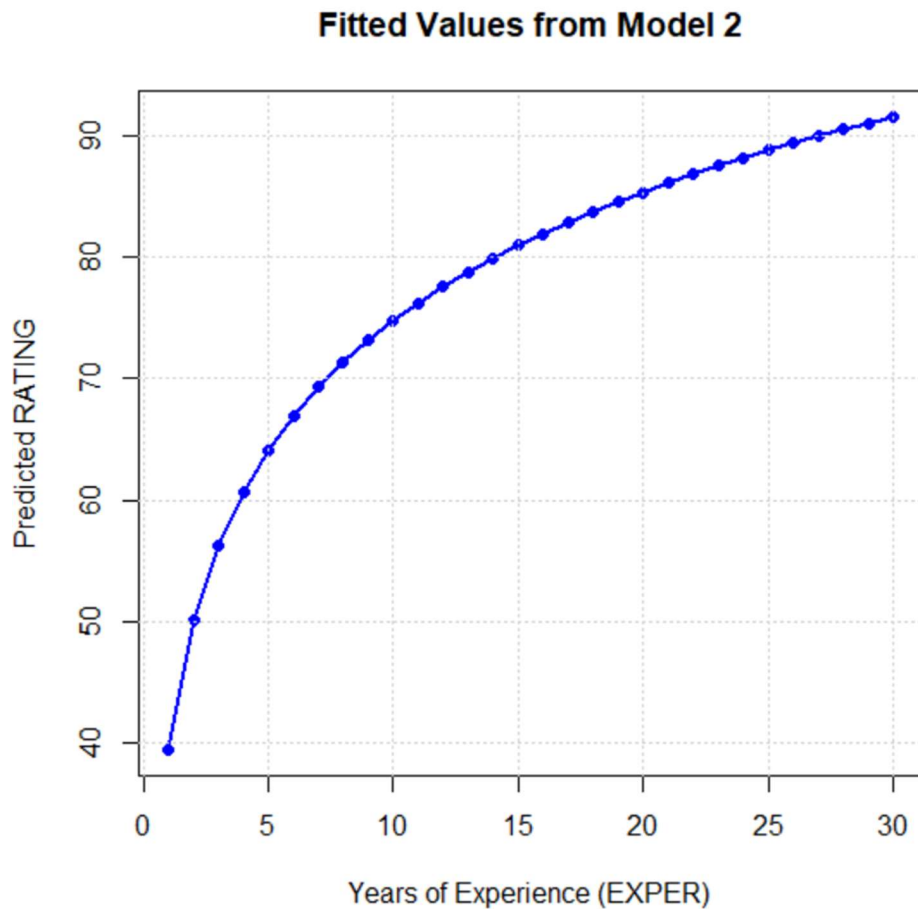
CH4

Q4

a.



b.



Model 2 變數中使用了對數轉換 $\ln(\text{EXPER})$ ，而 $\ln(0)$ 無意義，因此 Model 2 只適用於 $\text{EXPER} \geq 1$ 的數據。

c.

Model 1 中 EXPER 的係數是 0.990，因此，無論當前工作經驗是多少，額外一年經驗的影響都是 0.990 分。

對於擁有 10 年經驗的藝術家：邊際影響 = 0.990

對於擁有 20 年經驗的藝術家：邊際影響 = 0.990

d.

對於擁有 10 年經驗的藝術家 ($\text{EXPER}=10$):

$$15.312/10=1.5312$$

→ 額外 1 年的經驗預計使 **RATING** 增加 **1.531** 分。

對於擁有 20 年經驗的藝術家 (**EXPER=20**):

$15.312/20=0.7656$ 。額外 1 年的經驗預計使 **RATING** 增加 **0.766** 分

e.

模型二比較好，因為他的 R square = 0.6414 大於模型一 R square = 0.4858

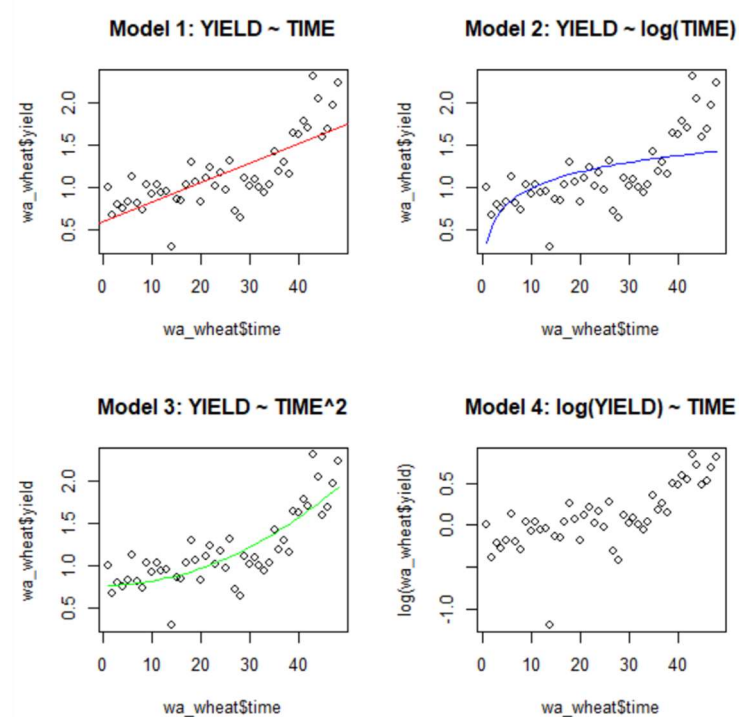
f.

線性模型 (**Model 1**) 假設影響固定，不符合現實，而對數模型 (**Model 2**) 能捕捉遞減效應，更貼近實際情況

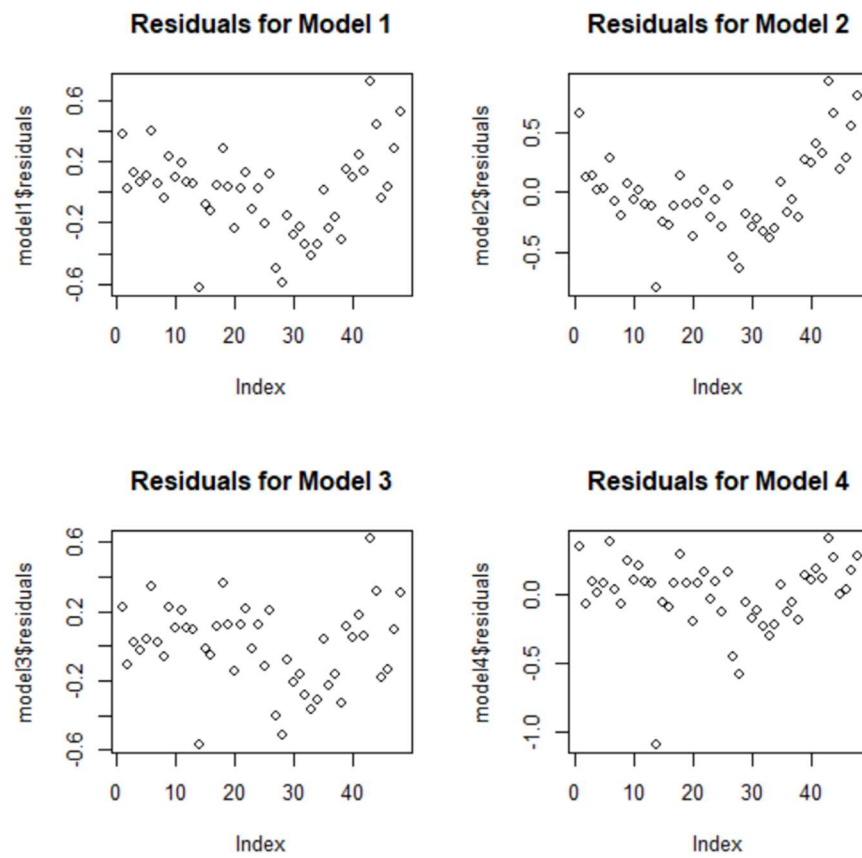
Q28

a.

i.



ii.



iii.

error test:

Model 1 : p 值 0.6792

Model 2 : p 值 0.1856

Model 3 : p 值 0.8266

Model 4 : p 值 7.205e-05

R square :

```

> summary(model1)$r.squared
[1] 0.5778369
> summary(model2)$r.squared
[1] 0.3385733
> summary(model3)$r.squared
[1] 0.6890101
> summary(model4)$r.squared
[1] 0.5073566

```

模型三最好因為他的 R square 解釋例最高。

b.

```

Call:
lm(formula = yield ~ time2, data = wa_wheat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.56899 -0.14970  0.03119  0.12176  0.62049

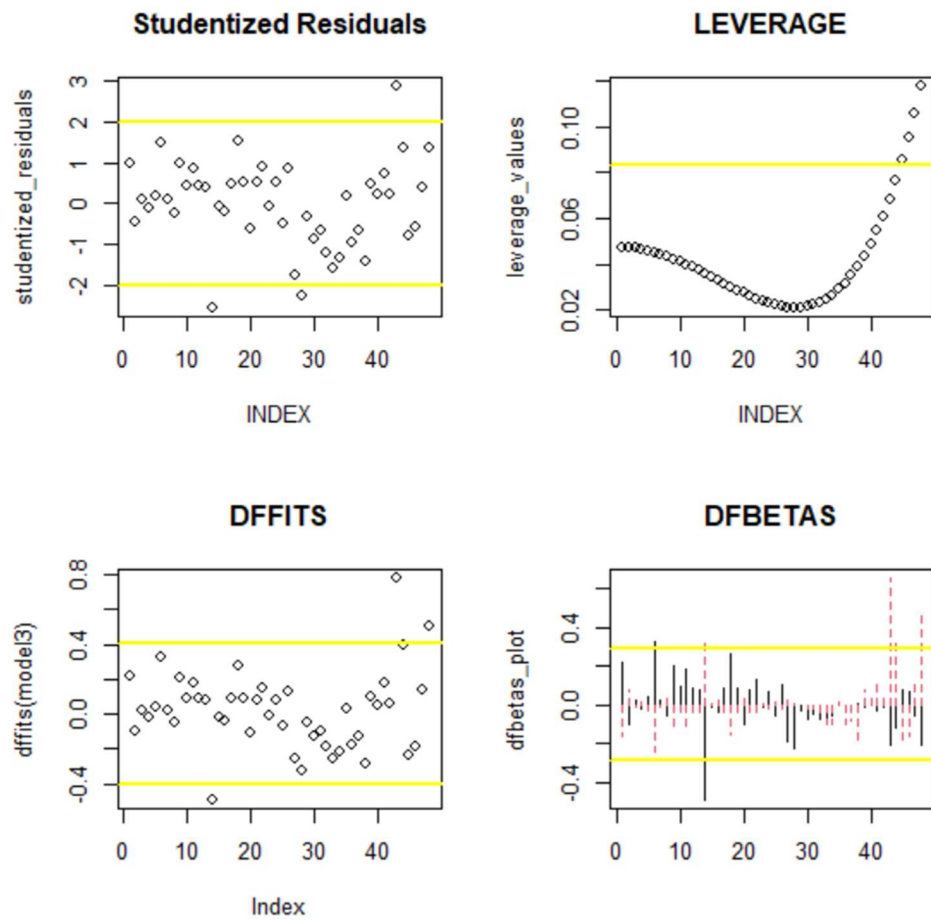
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.737e-01  5.222e-02   14.82  < 2e-16 ***
time2        4.986e-04  4.939e-05   10.10 3.01e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2396 on 46 degrees of freedom
Multiple R-squared:  0.689,    Adjusted R-squared:  0.6822
F-statistic: 101.9 on 1 and 46 DF,  p-value: 3.008e-13

```

係數正顯著，表示時間的平方對產量（Yield）有正向影響，其估計係數為 **0.0004986**，即隨著時間增加，小麥產量的成長速度加快（非線性增長）。

c.



異常觀察值: 14 28 43 45 46 47 48

d.

信賴區間 : $[1.372403, 2.389819]$ 。 YIELD = 2.2318 有落在信賴區間中。

Q29

a.

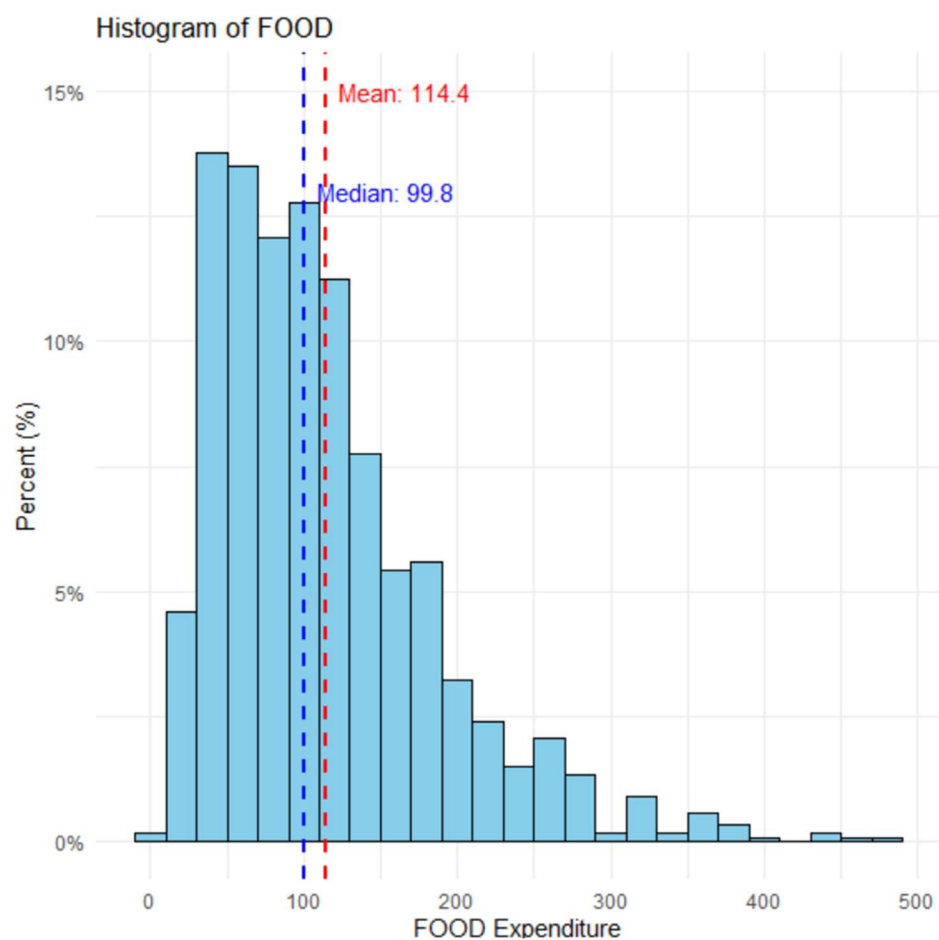
兩個變數的敘述性統計：

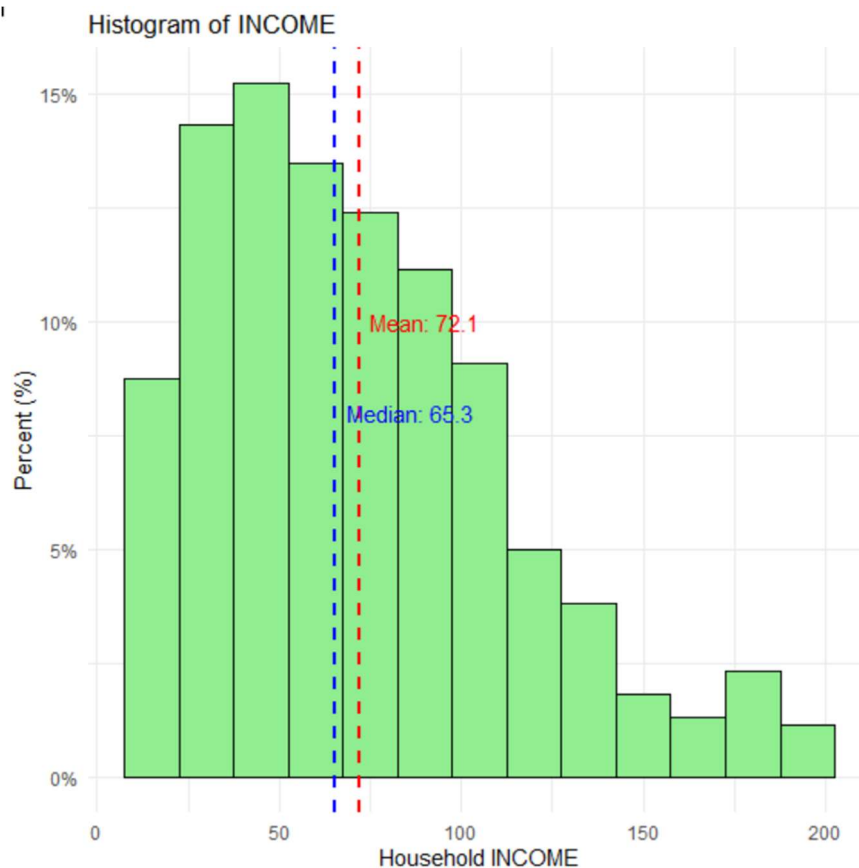
```
$summary_food
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  9.63  57.78   99.80  114.44  145.00  476.67

$summary_income
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 10.00  40.00   65.29   72.14  96.79  200.00

$sd_food
[1] 72.6575

$sd_income
[1] 41.65228
```





```
$jb_food  
  
Jarque Bera Test  
  
data:  cex5_small$food  
X-squared = 648.65, df = 2, p-value < 2.2e-16  
  
$jb_income  
  
Jarque Bera Test  
  
data:  cex5_small$income  
X-squared = 148.21, df = 2, p-value < 2.2e-16
```

說明:

直方圖形狀：FOOD 和 INCOME 的直方圖 不是對稱的，也不是鐘形分佈。

偏態性（右偏）：樣本均值大於中位數，表明 FOOD 和 INCOME 變數呈現右偏（**right-skewed**）。右偏分佈通常表示少數高值拉高了均值，例如少數高收入家庭影響了整體收入分佈。

Jarque-Bera 正態性檢定：FOOD 和 INCOME 的 Jarque-Bera 統計量均大於 5.99（5% 顯著水準的臨界值）。拒絕零假設（ H_0 ），即這兩個變數不服從正態分佈。

b.

回歸結果

```
Call:
lm(formula = food ~ income, data = cex5_small)

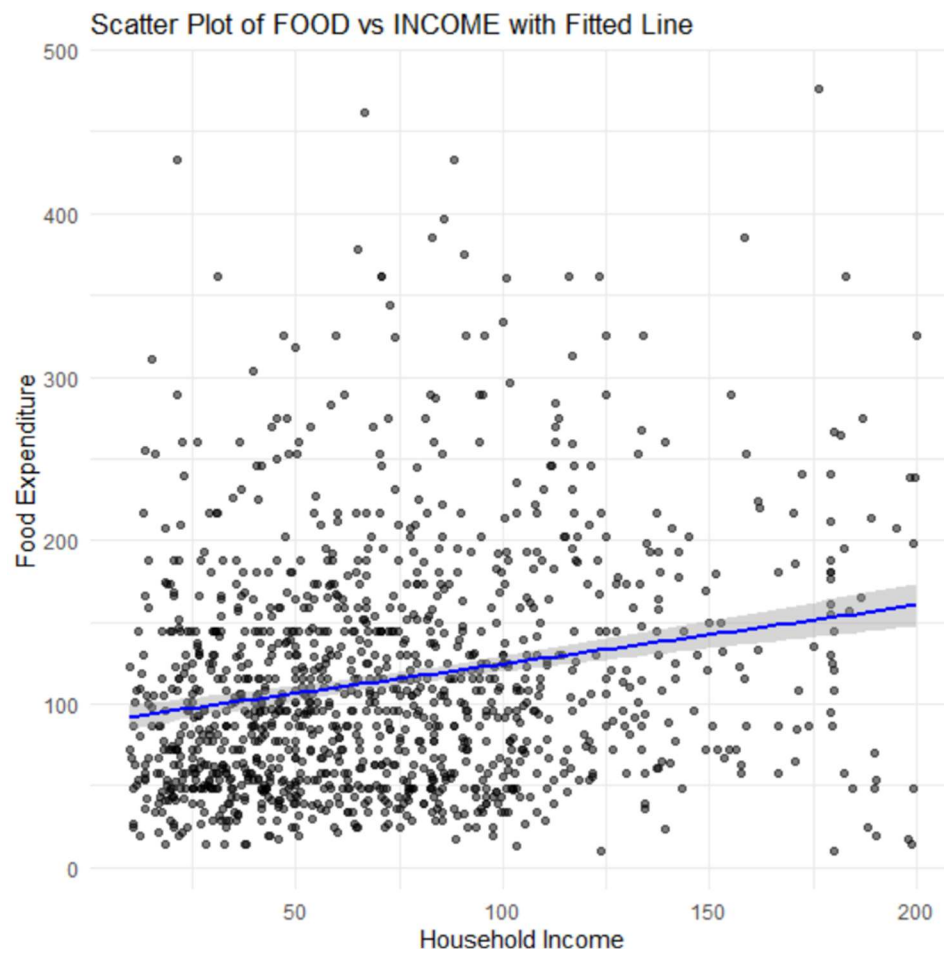
Residuals:
    Min       1Q   Median       3Q      Max
-145.37  -51.48  -13.52   35.50  349.81

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  88.56650     4.10819   21.559 < 2e-16 ***
income        0.35869     0.04932    7.272 6.36e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 71.13 on 1198 degrees of freedom
Multiple R-squared:  0.04228,    Adjusted R-squared:  0.04148
F-statistic: 52.89 on 1 and 1198 DF,  p-value: 6.357e-13
```

信賴區間

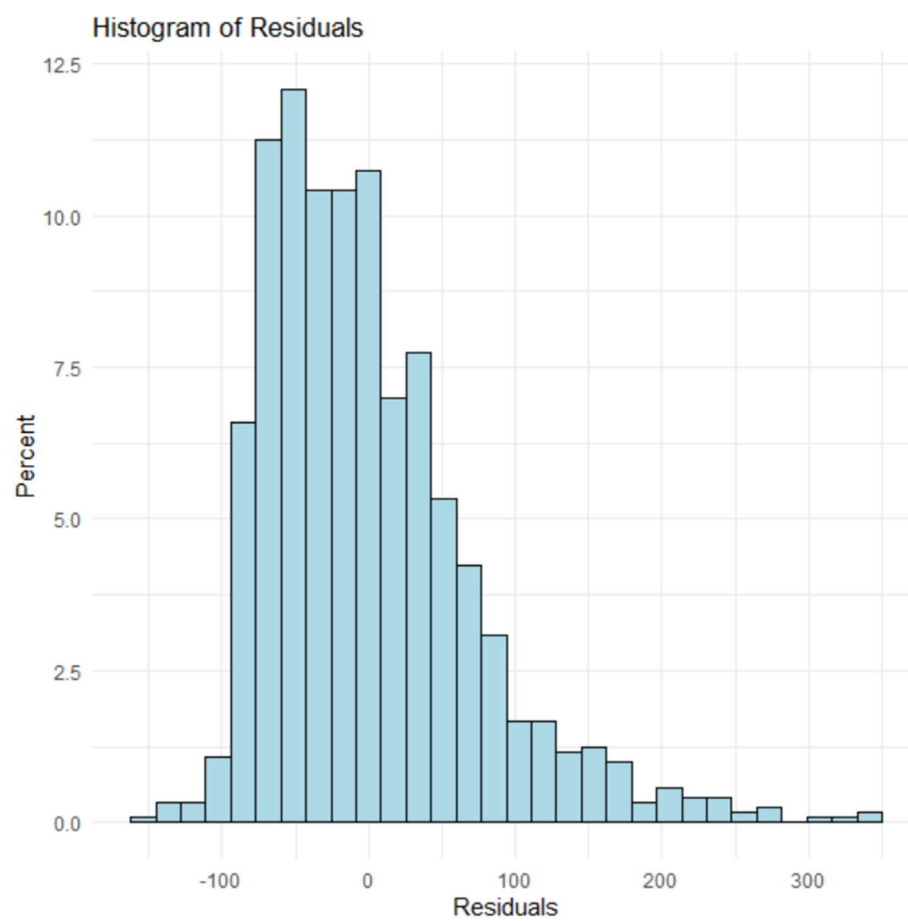
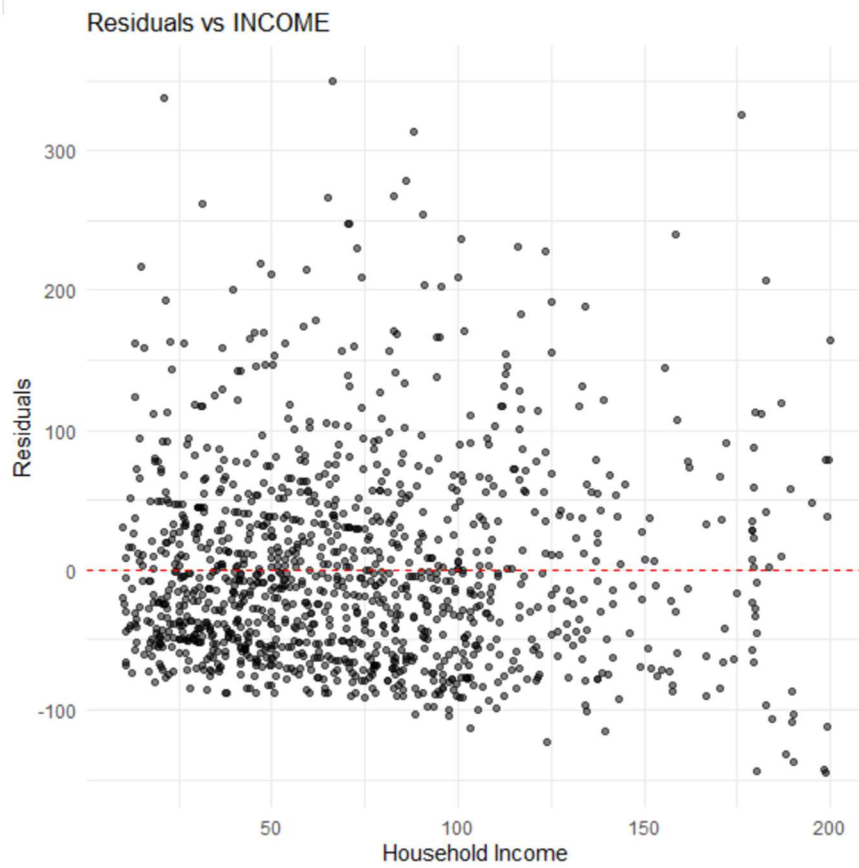
```
> confint(model, level = 0.95)
              2.5 %      97.5 %
(Intercept) 80.5064570 96.626543
income       0.2619215  0.455452
```



結果可得，由於信賴區間不包含 0，這表示 收入對食品支出有統計顯著影響。

然而，區間較寬，意味著對實際影響的估計存在一定的不確定性

c.



Jarque Bera Test

```
data: model$residuals  
X-squared = 624.19, df = 2, p-value < 2.2e-16
```

殘差不服從正態分佈，因為殘差的直方圖呈現右偏。在 OLS（最小二乘法）假設中，隨機誤差項服從正態分佈比 FOOD 和 INCOME 變數本身服從正態分佈更為重要。

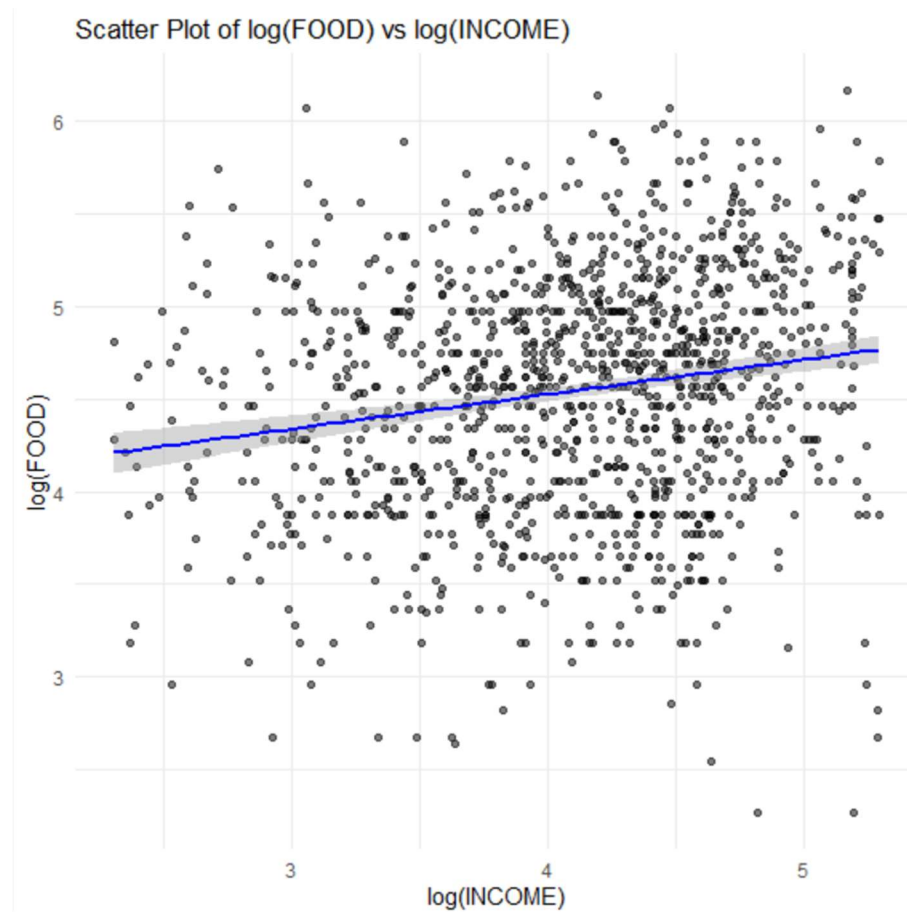
d.

```
INCOME Predicted_FOOD Elasticity Lower_CI Upper_CI  
1      19      95.38155 0.07145038 0.05217475 0.09072601  
2      65     111.88114 0.20838756 0.15216951 0.26460562  
3     160     145.95638 0.39319883 0.28712305 0.49927462  
> |
```

從表中可以看出：

- 所得彈性（elasticity）隨著收入（INCOME）增加而上升，從 0.07145（INCOME = 19）增至 0.39320（INCOME = 160）。
- 置信區間（95%）的上下界沒有重疊，顯示不同收入水準下的彈性估計統計上有顯著差異。
- 根據經濟學原理，食品的所得彈性通常較低（必需品需求對收入變化不敏感），但當收入增加時，消費者可能購買更多高價食品，導致彈性上升。

e.



```
> summary(log_model)$r.squared
[1] 0.03322915
> summary(model)$r.squared
[1] 0.0422812
```

線性回歸模型 R square 比較高，因此配適度比較好。

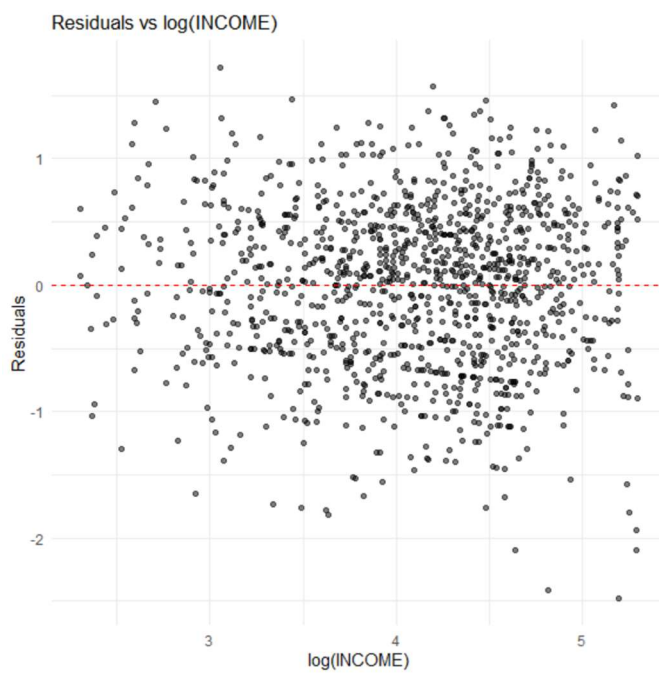
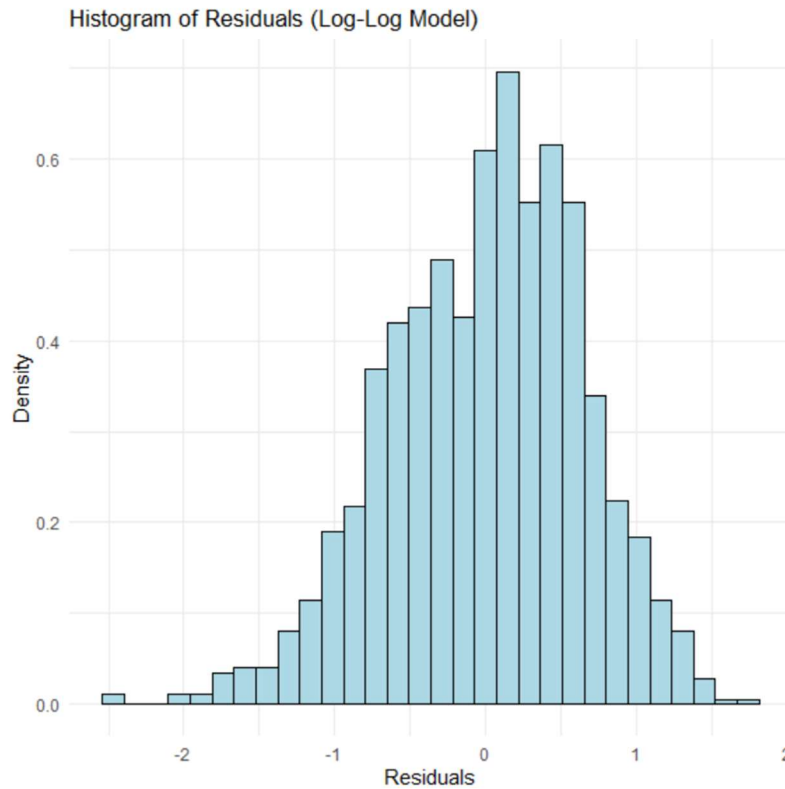
f.

彈性: 0.1863054

信賴區間

```
          2.5 %      97.5 %
log_income 0.1293432 0.2432675
```

g.



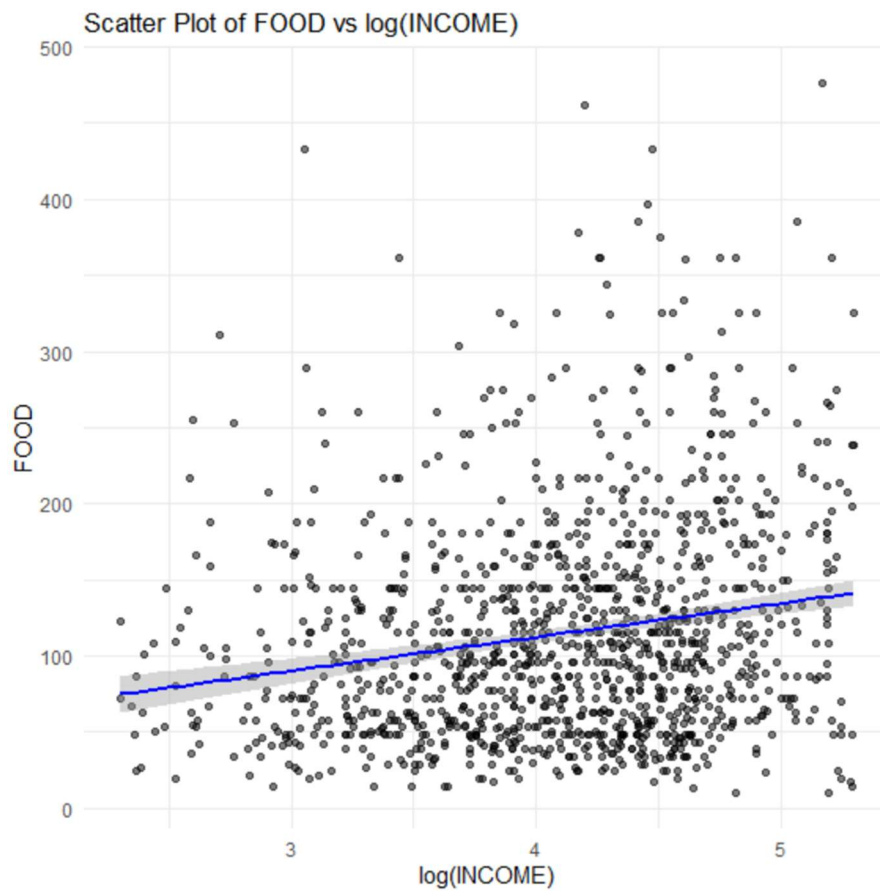
Jarque Bera Test

```
data: residuals_log  
X-squared = 25.85, df = 2, p-value = 2.436e-06
```

由於 $p\text{-value} < 0.05$ 顯著，拒絕常態假設，殘差非常態分佈

h.

R square = 0.0379



Linear-linear 配適度比較好。

i.

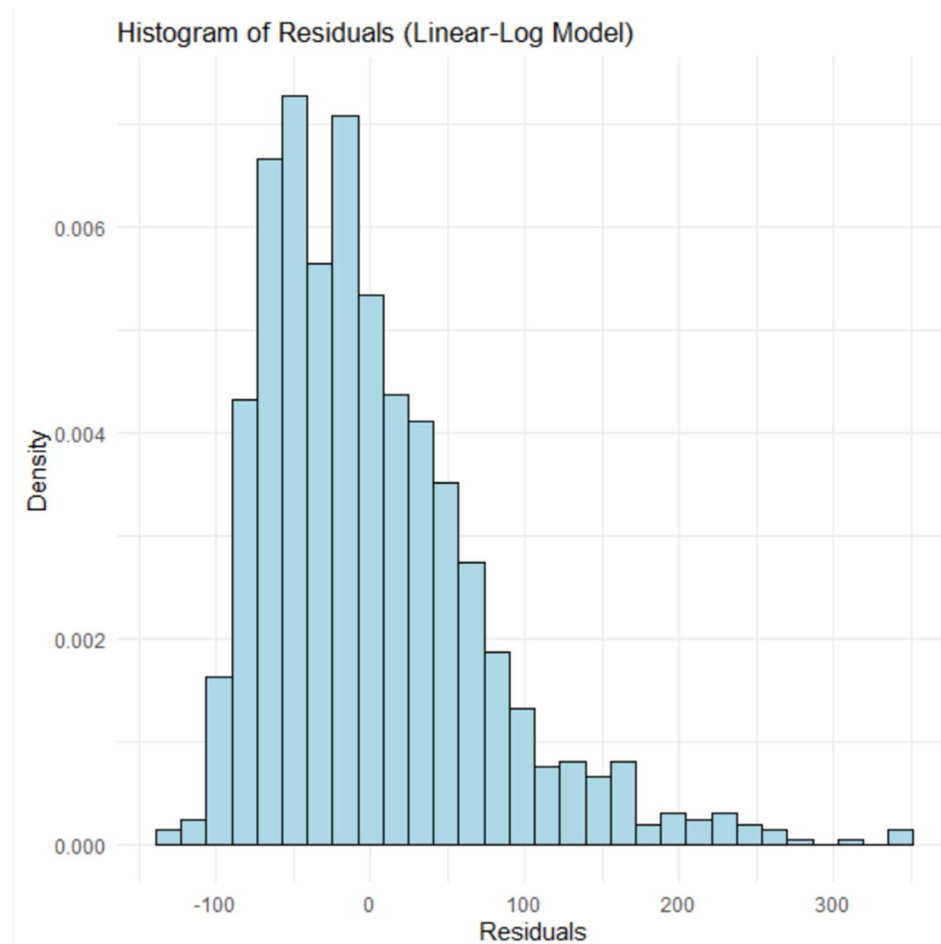
| | INCOME | Predicted_FOOD | Elasticity | Lower_CI | Upper_CI |
|---|--------|----------------|------------|-----------|-----------|
| 1 | 19 | 88.89788 | 0.2495828 | 0.1784009 | 0.3207648 |
| 2 | 65 | 116.18722 | 0.1909624 | 0.1364992 | 0.2454256 |
| 3 | 160 | 136.17332 | 0.1629349 | 0.1164652 | 0.2094046 |

當收入從 19 → 65 → 160 增加時，預測食品支出 也相應增加 (88.90 → 116.19 → 136.17)，這符合經濟學中「正常財」的概念，即收入增加，食品支出增加。

彈性 (Elasticity) 下降：收入較低時，食品支出的彈性較高 (0.2496)，隨著收入增加，彈性逐漸降低 (0.1909 → 0.1629)。這意味著食品支出對收入的敏感度

在降低。

j.



Jarque Bera Test

```
data: residuals_lin_log  
X-squared = 628.07, df = 2, p-value < 2.2e-16
```

結果顯示有點左偏，同時 p-value 小於 0.05，拒絕常態假設，殘差非常態。

k.

我會選擇 log-log 模型，因為與其他模型相比，它的殘差更接近常態分佈