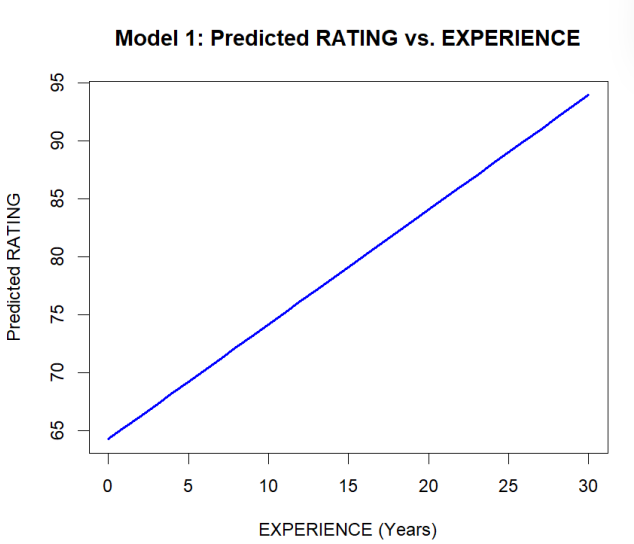
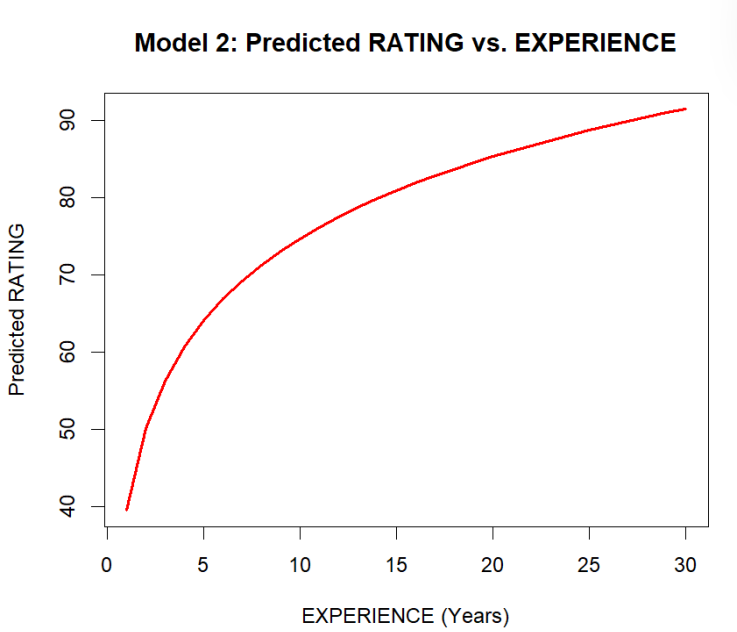
Q4

(a)



(b)



ln(EXPER)在 EXPER = 0 時是沒有定義，所以沒有包含經驗為 0 的 4 位藝術家，即為 Model 2 只用了 46 名藝術家的原因（N = 46）

(c)

將 Model 1 對 EXPER 微分，不管在經驗多少的情況下，得到的邊際效應皆為線性模型的斜率 0.990

(d)

將 Model 2 對 EXPER 微分，得出 15.312/EXPR。

當 EXPR =10，marginal effect = 15.312/10 = 1.5312

當 EXPR = 20，marginal effect = 15.312/20 = 0.7656

(e)

Model 2（對數模型）對數據的解釋力更強，因為它的 R^2 更高

(f)

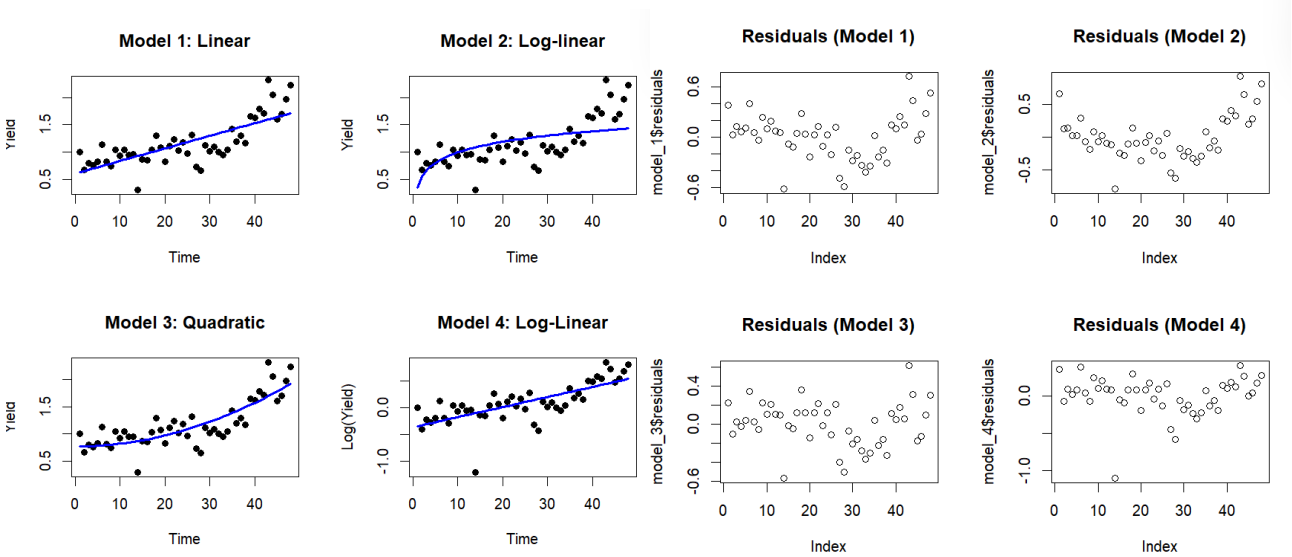
Model 1（線性模型）: 假設經驗的影響是恆定的，這樣的假設在長期可能不合理，因為一個藝術家從 20~30 年的進步，不可能跟 1~ 10 年的一樣大。

Model 2（對數模型）: 假設經驗的影響是遞減的（早期增長較快，後期趨於穩定）。符合一般的經驗學習曲線

所以 Model 2 更符合現實經濟現象，因為它允許經驗的邊際效應隨著年數增加而減少

Q28

(a)




```
> shapiro.test(model_3$residuals)

        Shapiro-Wilk normality test

data:  model_3$residuals
W = 0.98589, p-value = 0.8266
```

```
> names(r_squared) <- c("Linear", "Log-Time", "Quadratic", "Log-YIELD"
> print(r_squared)
   Linear  Log-Time Quadratic Log-YIELD
0.5778369 0.3385733 0.6890101 0.5073566
```

應該選擇 model3 Quadratic model，R^2 最高，殘差分配也最接近常態

(b)

```
Call:
lm(formula = northampton ~ I(time^2), 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 ***
I(time^2)   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
```

係數為正，代表 time^2 每上升一單位，yield 會增加 0.0004986 單位

(c)

```
異常觀察值 ID: 14
  - Studentized Residuals > 2: 誤差異常大，可能是異常值

異常觀察值 ID: 28
  - Studentized Residuals > 2: 誤差異常大，可能是異常值

異常觀察值 ID: 43
  - Studentized Residuals > 2: 誤差異常大，可能是異常值

異常觀察值 ID: 45
  - Leverage 遠大於平均值: 影響回歸結果

異常觀察值 ID: 46
  - Leverage 遠大於平均值: 影響回歸結果

異常觀察值 ID: 47
  - Leverage 遠大於平均值: 影響回歸結果

異常觀察值 ID: 48
  - Leverage 遠大於平均值: 影響回歸結果
```

```
> wa_wheat[outliers, ]
   northampton chapman mullewa greenough time
14      0.3024  0.4167  0.3965    0.4369   14
28      0.6539  0.5827  0.4252    0.9759   28
43      2.3161  2.0244  1.6880    1.8081   43
45      1.6040  1.4769  1.3871    1.5674   45
46      1.6980  1.4430  1.4558    1.6893   46
47      1.9691  1.7107  1.6571    1.7191   47
48      2.2318  1.8435  1.7992    2.2353   48
```

(d)

```
     fit    lwr    upr
1 1.8811 1.3724 2.3898

Actual northampton in 1997: 2.2318
```

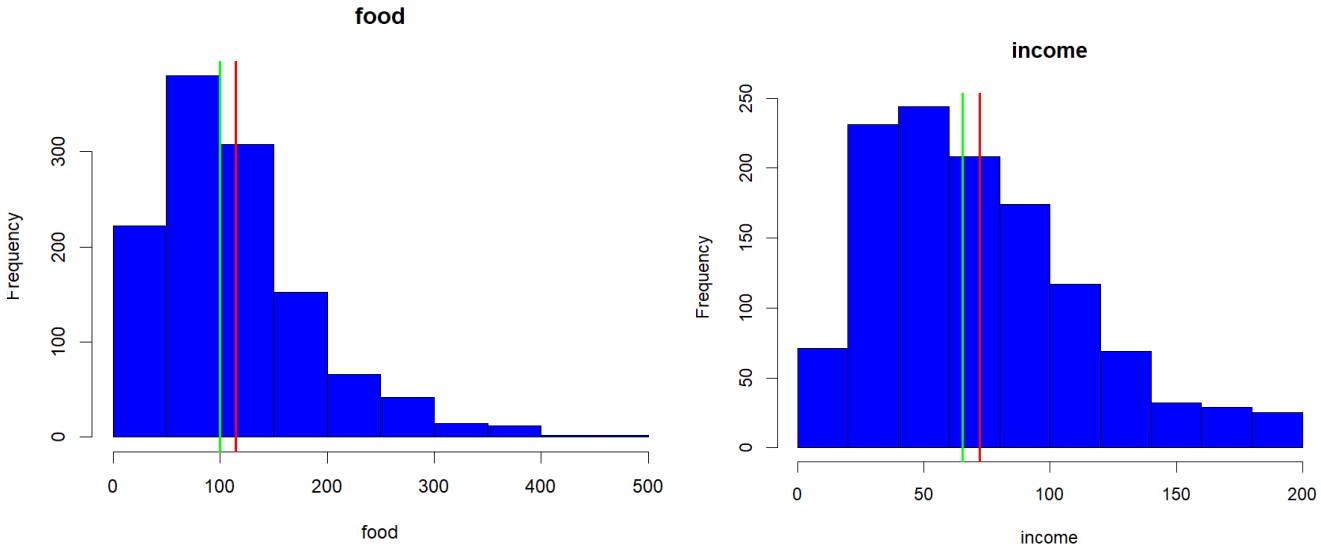
實際值有落在信賴區間內

Q29

(a)

```
> print(summary_food)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   9.63   57.78   99.80  114.44  145.00  476.67
> print(summary_income)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  10.00   40.00   65.29   72.14   96.79  200.00
```

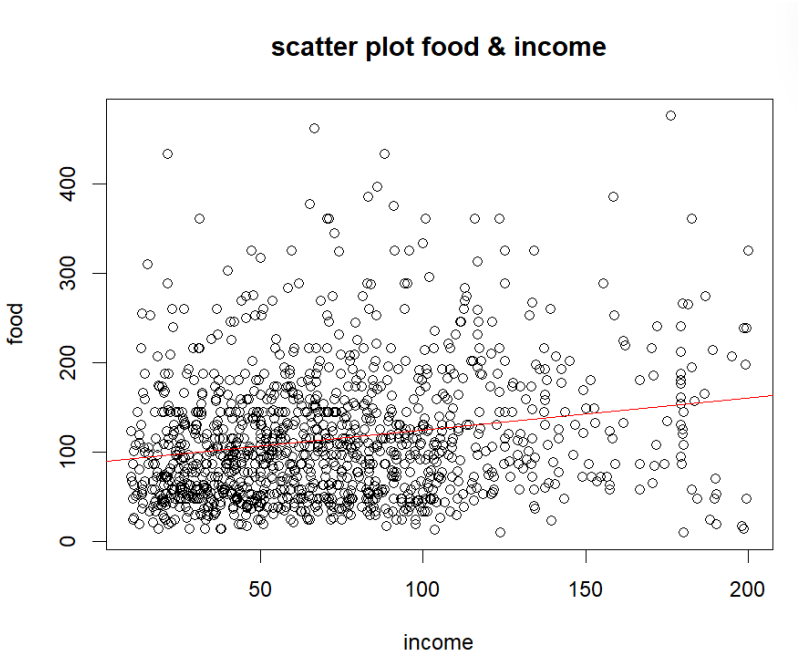
```
> print(sd_food)
[1] 72.6575
> print(sd_income)
[1] 41.65228
```




```
        Jarque-Bera Normality Test

data:  cex5_small$food
JB = 648.65, p-value < 2.2e-16
alternative hypothesis: greater


        Jarque-Bera Normality Test

data:  cex5_small$income
JB = 148.21, p-value < 2.2e-16
alternative hypothesis: greater
```

Food、Income 兩者皆不對稱，也無呈現鐘型分配，平均數大於中位數，Jarque–Bera test 的結果也顯示無法拒絕非常態分配，可能為正篇。

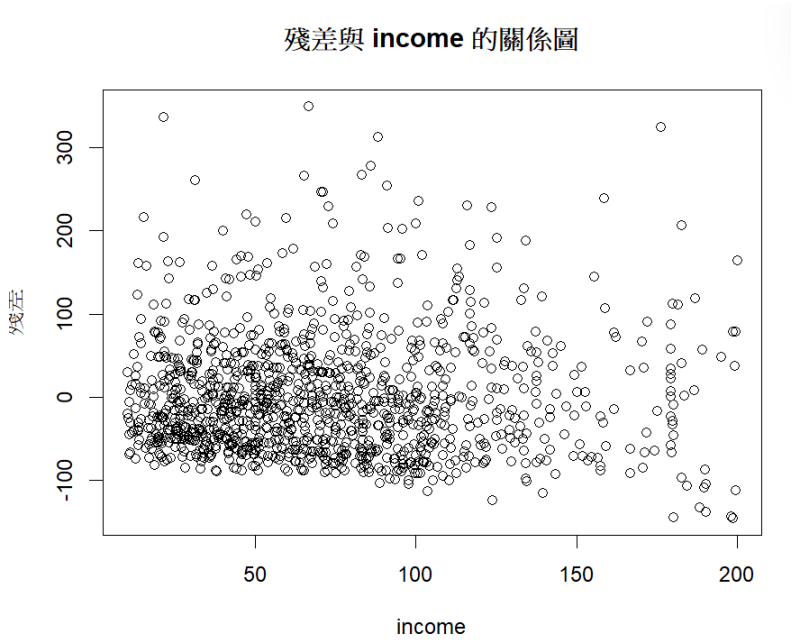
(b)

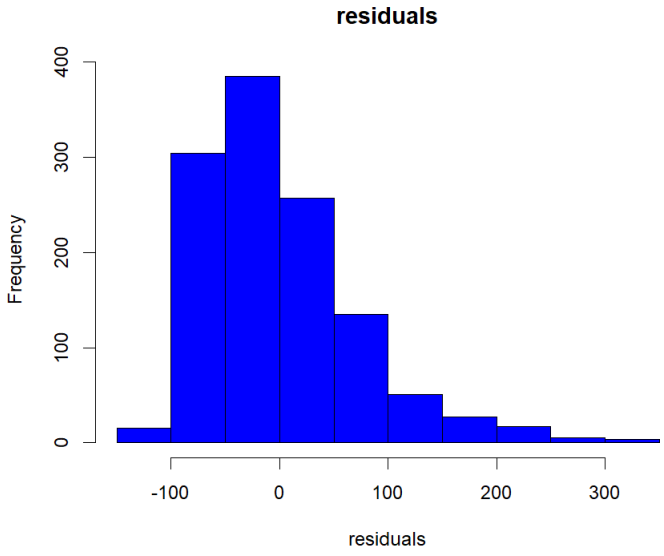



```
> confint(linear_model, level = 0.95)
                 2.5 %    97.5 %
(Intercept) 80.5064570 96.626543
income       0.2619215  0.455452

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
```

並不準確，因為 R^2 只有 0.04228

(c)






```
        Jarque-Bera Normality Test

data:  residuals_linear
JB = 624.19, p-value < 2.2e-16
alternative hypothesis: greater
```

無法拒絕非常態分配

(d)

```
> elasticity_estimates
          1           2           3
0.07145038 0.20838756 0.39319883
> elasticity_lower
          1           2           3
0.05217475 0.15216951 0.28712305
> elasticity_upper
          1           2           3
0.09072601 0.26460562 0.49927462
```

彈性隨收入增加單調遞增，且信賴區間無重疊，根據經濟學原理，通常認為食物的收入彈性會隨著收入的增加而減少，這裡可能是因為高收入群體的食物消費模式不同所導致不同的結果。

(e)

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

Residuals:
     Min       1Q   Median       3Q      Max
-2.48175 -0.45497  0.06151  0.46063  1.72315

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.77893    0.12035  31.400   <2e-16 ***
log(income)  0.18631    0.02903   6.417    2e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6418 on 1198 degrees of freedom
Multiple R-squared:  0.03323,   Adjusted R-squared:  0.03242
F-statistic: 41.18 on 1 and 1198 DF,  p-value: 1.999e-10
```


```
Generalized R² for log-log model:  0.0327
```

R^2 還是很小，並沒有比較好

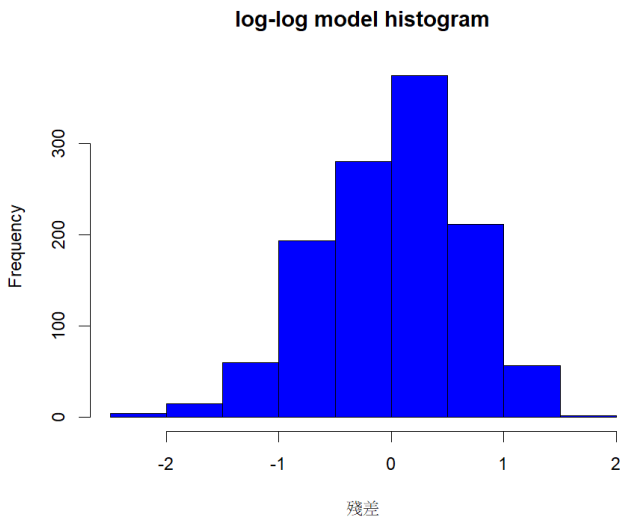
(f)

```
> gamma2 <- coef(log_log_model)[2]
> gamma2
ln_income
0.1863054
```

Log-log model 的彈性是固定的

```
> confint(log_log_model, level = 0.95)
                 2.5 %    97.5 %
(Intercept) 3.5428135 4.0150507
ln_income   0.1293432 0.2432675
```

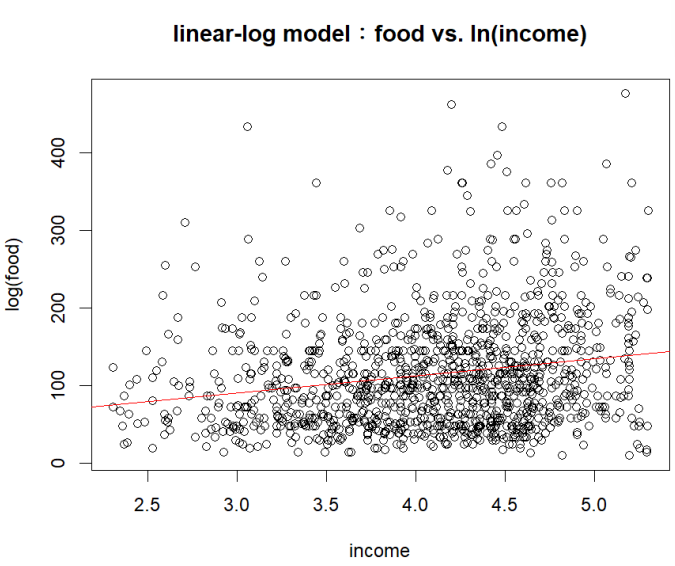
(g)




```
        Jarque-Bera Normality Test

data:  residuals_log_log
JB = 25.85, p-value = 2.436e-06
alternative hypothesis: greater
```

殘差分配依然屬於非常態分布

(h)



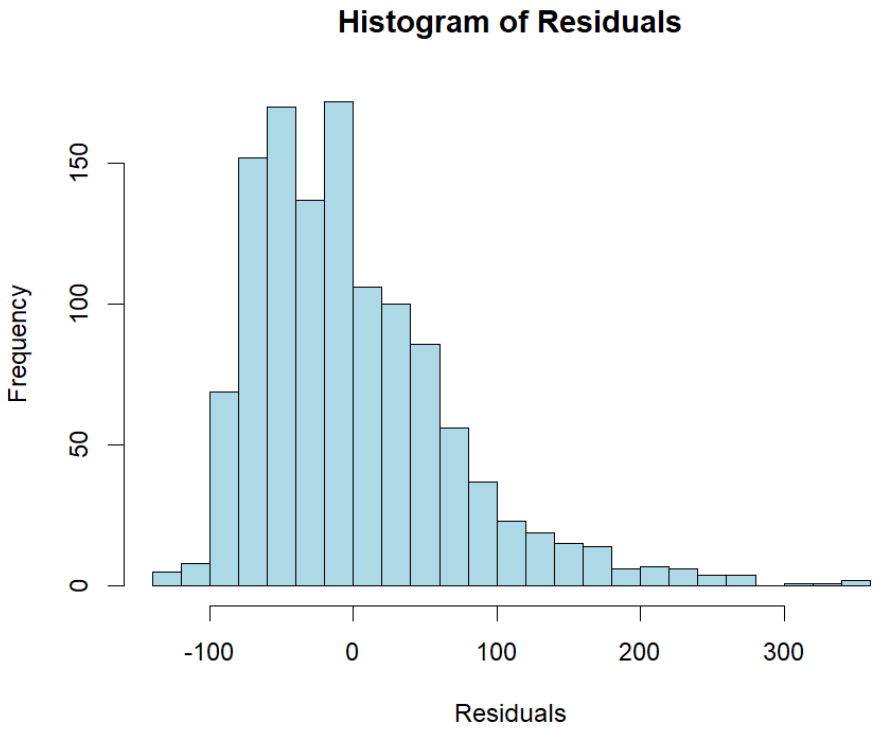

```
> summary(linear_log_model)$r.squared
[1] 0.03799984
```

跟前面的相比，R^2 仍然不太好

(i)

	income	Fitted_food	Elasticity	Lower_95CI	Upper_95CI
1	19	88.90	4.7421	3.3910	6.0932
2	65	116.19	12.4126	8.8760	15.9491
3	160	136.17	26.0696	18.6418	33.4974

(j)




```
        Jarque-Bera Normality Test

data:  residuals
JB = 628.07, p-value < 2.2e-16
alternative hypothesis: greater
```

仍然不支持常態分配

(k)

Log-log model 殘差較隨機分配