2.25 Consumer expenditure data from 2013 are contained in the file *cex5_small*. [Note: *cex5* is a larger version with more observations and variables.] Data are on three-person households consisting of a husband and wife, plus one other member, with incomes between $1000 per month to $20,000 per month. *FOODAWAY* is past quarter's food away from home expenditure per month per person, in dollars, and *INCOME* is household monthly income during past year, in $100 units.

a. Construct a histogram of *FOODAWAY* and its summary statistics. What are the mean and median values? What are the 25th and 75th percentiles?

b. What are the mean and median values of *FOODAWAY* for households including a member with an advanced degree? With a college degree member? With no advanced or college degree member?

c. Construct a histogram of ln(*FOODAWAY*) and its summary statistics. Explain why *FOODAWAY* and ln(*FOODAWAY*) have different numbers of observations.

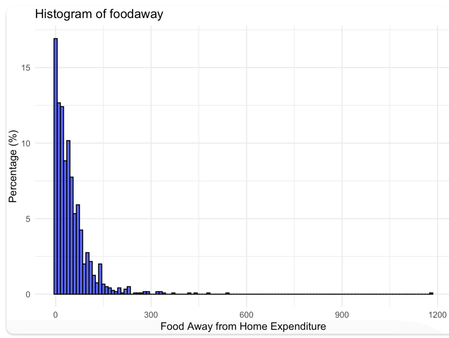
d. Estimate the linear regression $\ln(FOODAWAY) = \beta_1 + \beta_2 INCOME + e$. Interpret the estimated slope.

e. Plot ln(*FOODAWAY*) against *INCOME*, and include the fitted line from part (d).

f. Calculate the least squares residuals from the estimation in part (d). Plot them vs. *INCOME*. Do you find any unusual patterns, or do they seem completely random?

a.

```
$`foodaway 平均值`
[1] 49.27085

$`foodaway 中位數`
[1] 32.555

$`foodaway 第25百分位數`
   25%
12.04

$`foodaway 第75百分位數`
     75%
67.5025
```

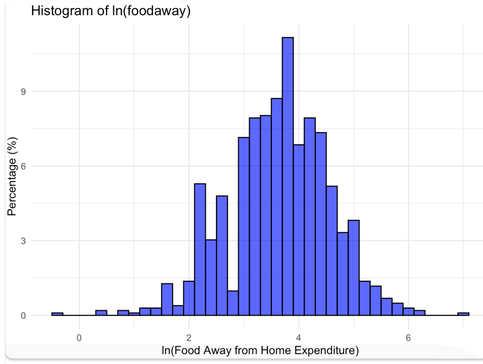


b.

	N	Mean	Median
ADVANCED = 1	257	73.15494	48.15
COLLEGE = 1	369	48.59718	36.11
NONE	574	39.01017	26.02

c. ln(foodaway) 無法對 0 或複數取 ln，導致 ln(foodaway) 的觀測數量少於原始 foodaway，共減少 178 筆

```
$`ln(foodaway) 平均值`
[1] 3.650804

$`ln(foodaway) 中位數`
[1] 3.686499

$`ln(foodaway) 第25百分位數`
     25%
3.075929

$`ln(foodaway) 第75百分位數`
     75%
4.279717

$原始觀測數量
[1] 1200

$過濾後觀測數量
[1] 1022

$減少的觀測數量
[1] 178
```

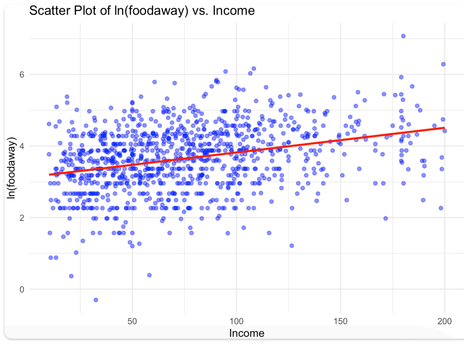


d. $\ln(\widehat{foodaway}) = 3.1293 + 0.0069\ income$

For every additional $100 in household income, the food away from home expenditure increases by approximately 0.69% per person.

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

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. 殘差隨機分佈，沒有明顯的非線性模式，顯示OLS線性模型適用