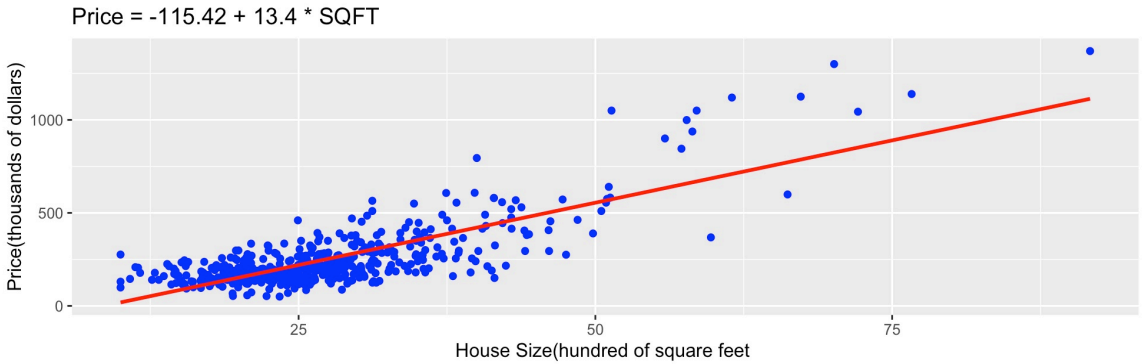
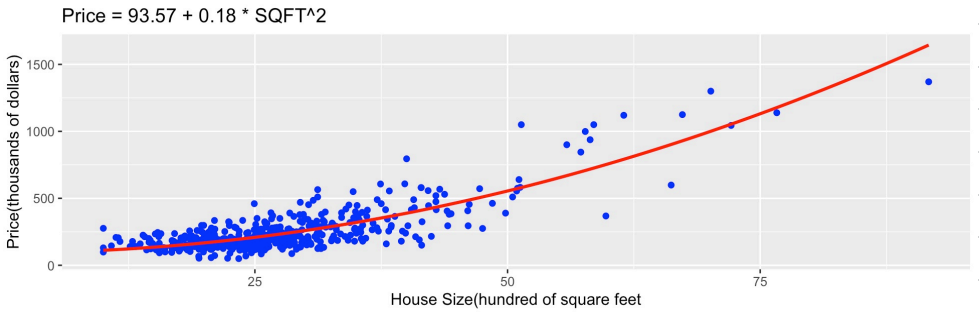
2.17

a、b.



b_2=13.4，代表當 house size 增加 100 平方呎，
以平均來說，價格會增加 13.4×1000=13,400元
b_1=-115.42，代表當house size為0時，以平均來說，
房屋價格為 -115.42×1000=-115,420元（無意義可不用解讀）

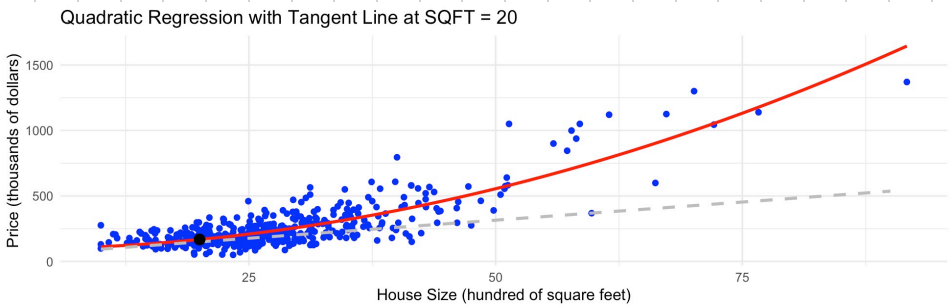
C.



$\frac{dPrice}{dSqft} = 0.18 \times 2 \times Sqft$

marginal effect $= 0.18 \times 2 \times 20 = 7.2$

d.



e.

當 SQFT=2000呎, Price $= 93.57 + 0.18 \times 20^2 = 165.57$

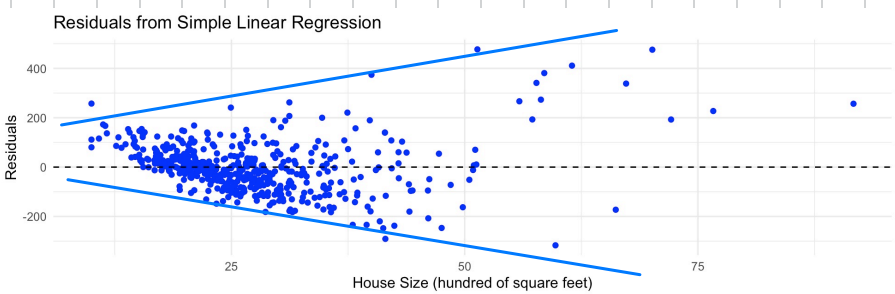
elasticity $E = \frac{dPrice/Price}{d(SQFT)/SQFT}$

$= \frac{dPrice}{dSQFT} \times \frac{SQFT}{price}$

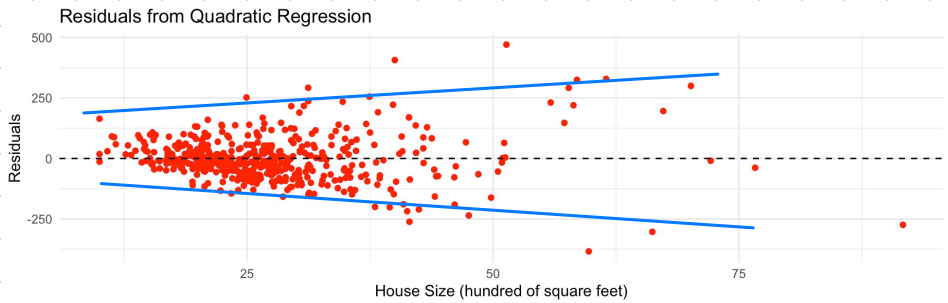
$= 0.18 \times 2 \times 20 \times \frac{20}{165.57} = 0.8697$ #

f.

Simple:



Quadratic:



根據兩張殘差圖可知，當 house size 愈大，殘差愈分散
違反同質性假設

g.

SSE for simple regression = 5262846.9471

SSE for quadratic regression: 4222356.3493

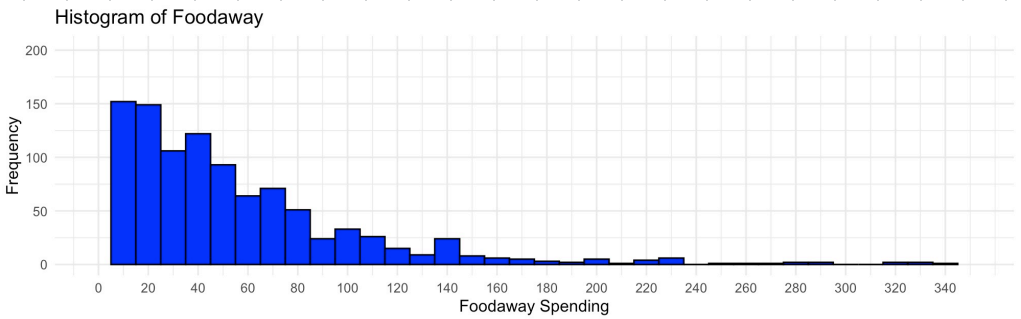
⇒ 二次項的模型具有較小 SSE

⇒ 較佳模型

⇒ 代表實際的 y 與預測 y 差距較小（殘差較小）

2.25

a.



mean: 49.27 median = 32.56

25th percentile (Q1): 12.04

75th percentile (Q3): 67.5

b.

mean_foodaway_advanced = 73.1549

median_foodaway_advanced: 48.15

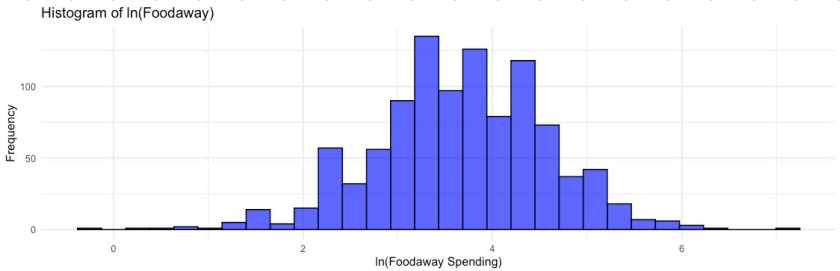
mean_foodaway_college: 48.5972

median_foodaway_college: 36.11

mean_foodaway_others: 39.0102

median_foodaway_others: 26.52

c.

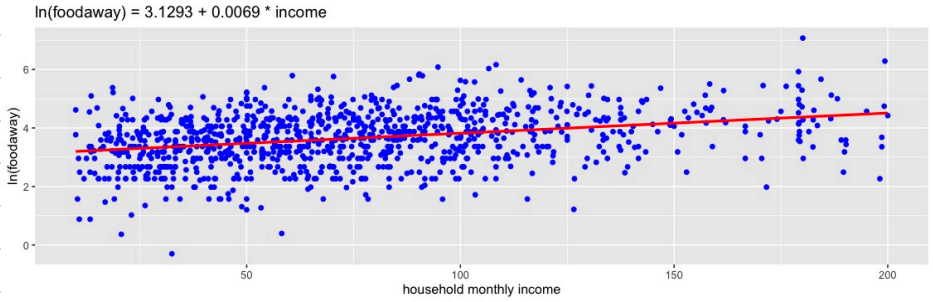


mean: 3.6508 median=3.6865

25%(Q1): 3.0759 75%(Q3): 4.2797

兩觀察值差異是 $\ln 0 = -\infty$，故在畫圖會排除

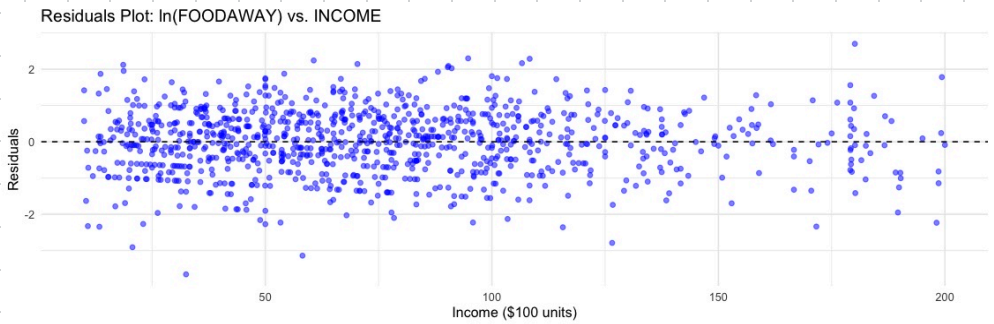
d. e.



$\ln(\text{foodaway}) = 3.1293 + 0.0069 \cdot \text{income}$

當 income 增加一單位，預期 foodaway spending 增加 0.69%

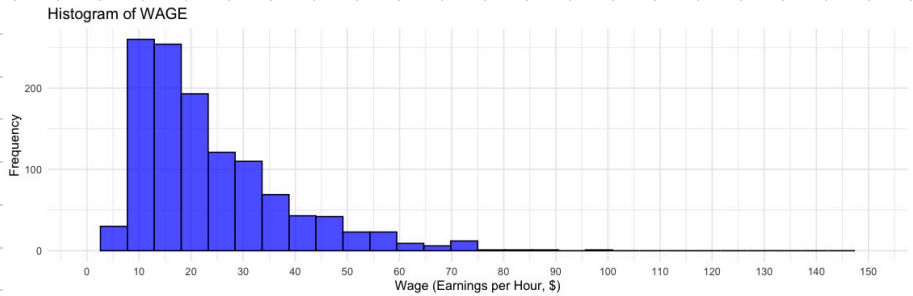
f



由殘差圖可知誤差項應滿足三大假設
常態獨立同質

2.2f

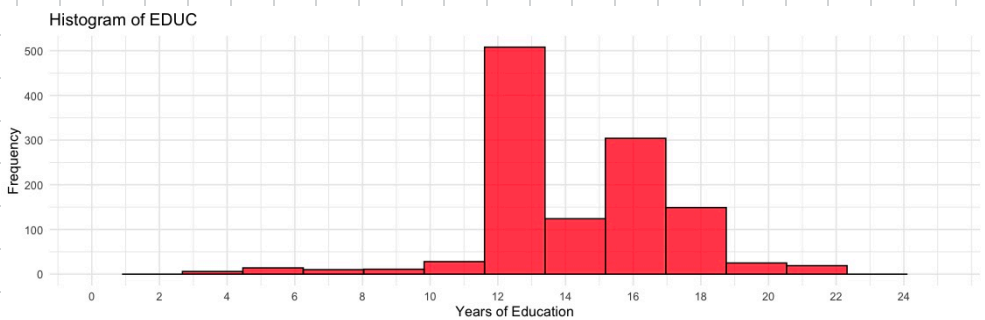
a.



mean: 23.64 Median: 19.3 S.D: 15.2166 Min: 3.94

Max: 221.1 Q1: 13 Q3: 29.8

看起來像右尾分配，有極端高所得



mean: 14.2025 median: 14 S.D: 2.8908 Min: 0

Max: 21 Q1: 12 Q3: 16

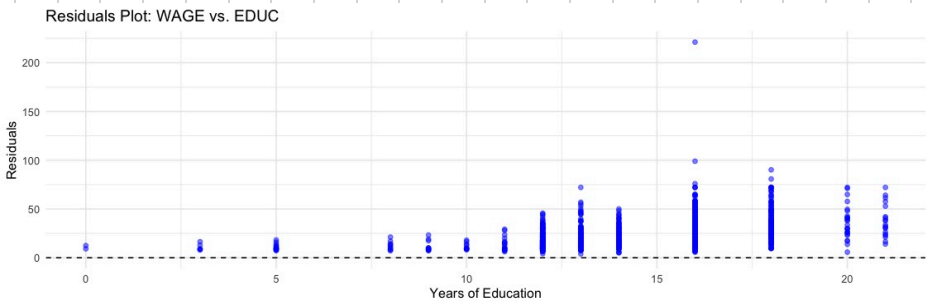
教育年限集中在12年(高中學歷)跟16年(大學學歷)

b.

Wage = -10.4 + 2.3968 edu

看起來合理，當 edu=12(高中學歷)，以平均來說可以賺到 -10.4 + 2.3968×12 = 18.3616

c.



由殘差圖可知，應該具有異質變異數，可能要加入 educ² 有極端值

d.

Male: wage = -8.285 + 2.378 educ

female: wage = -16.603 + 2.784 educ

black: wage = -6.254 + 5.554 educ

white: wage = -10.475 + 2.448 educ

e. (by R code)

quad: $wage = 4.9165 + 0.089 educ^2$

marginal effect at educ for linear regression

12 years: $2.3968 \times 12 = 28.761$

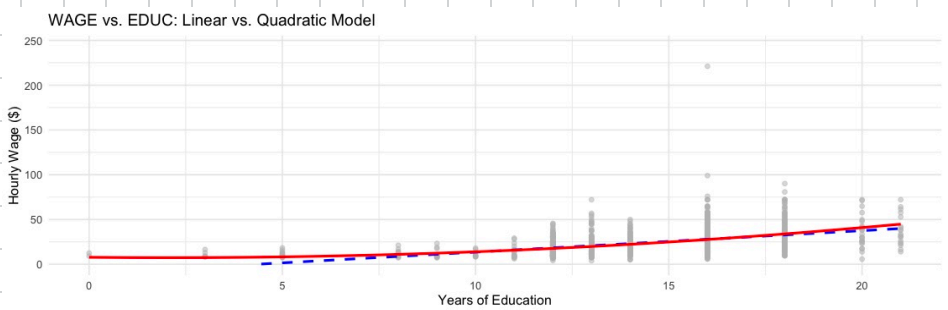
16 years: $2.3968 \times 16 = 38.348$

marginal effect at educ for quadratic regression

12 years: $0.089 \times 2 \times 12 = 2.139$

16 years: $0.089 \times 2 \times 16 = 2.852$

f.



by R code:

Adj R^2 for linear regression is 0.2067

Adj R^2 for quadratic regression is 0.2187

ANOVA:

F-value for linear regression is 313.34

F-value for quadratic regression is 336.63

$\therefore$ Adj-R^2 quad > Adj-R^2 liner + F-value quad > F-value linear

$\therefore$ quadratic regression is a better model *