Q6

(a) $H_0: \sigma_M = \sigma_F$, $H_1: \sigma_M \neq \sigma_F$

$n_F = 423$

$\hat{\sigma}_M^2 = \frac{SSE_M}{n_m - k} = \frac{97161.9174}{573} = 169.567$

$GQ = \frac{\hat{\sigma}_M^2}{\hat{\sigma}_F^2} = \frac{169.567}{(12.024)^2} = 1.1729$, $F(0.975, 573, 419) = 1.196$, $GQ = 1.1729 < 1.196 = F$

$\Rightarrow$ we don't reject H_0, 在 5% 信心水準下, 無證據顯示 $\sigma_M^2 \neq \sigma_F^2$

(b) $H_0: \sigma_S = \sigma_M$, $H_1: \sigma_S < \sigma_M$

$\hat{\sigma}_S^2 = \frac{SSE_S}{400-5} = \frac{56231.0382}{395} = 142.3571$, $\hat{\sigma}_M^2 = \frac{SSE_M}{600-5} = \frac{100703.0471}{595} = 169.2488$

$GQ = \frac{169.2488}{143.3571} = 1.1806$, $F(0.95, 595, 395) = 1.1647$, $GQ = 1.1806 > 1.1647 = F$

$\Rightarrow$ we reject H_0, 在 5% 信心水準下, 顯示 $\sigma_S < \sigma_M$

(c) $NR^2 = 59.03$, $\chi^2_{0.05,4} = 9.488$, $NR^2 > \chi^2_{0.05,4}$, reject H_0, support that heteroskedasticity is exist.

consistent with part (b)

(d) test statistic = 78.82 , $df = p + p^2 + C_2^p = 4 + 2 + C_2^4 = 12$ (dummy variable 平方項相同)

$\chi^2_{0.95,12} = 21.026$ we reject H_0.

(e) Narrower: EXPER, METRO, FEMALE , Wider: EDUC , The result are inconsistent

(f) It is compatible. Because White test is robust

it could use in both homo / heteroskedasticity

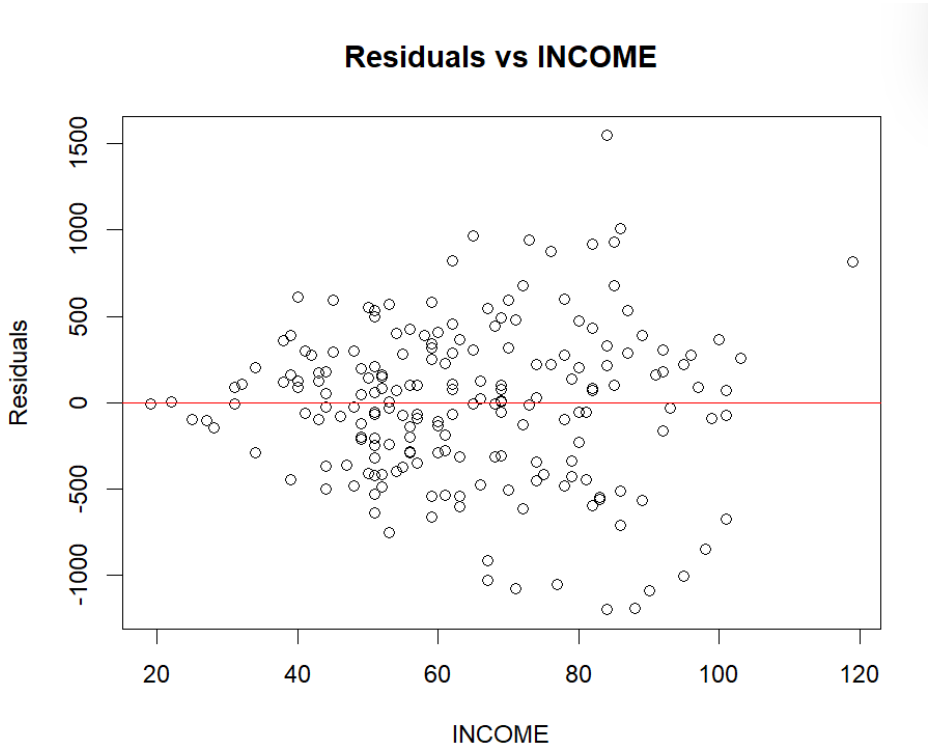
Q16

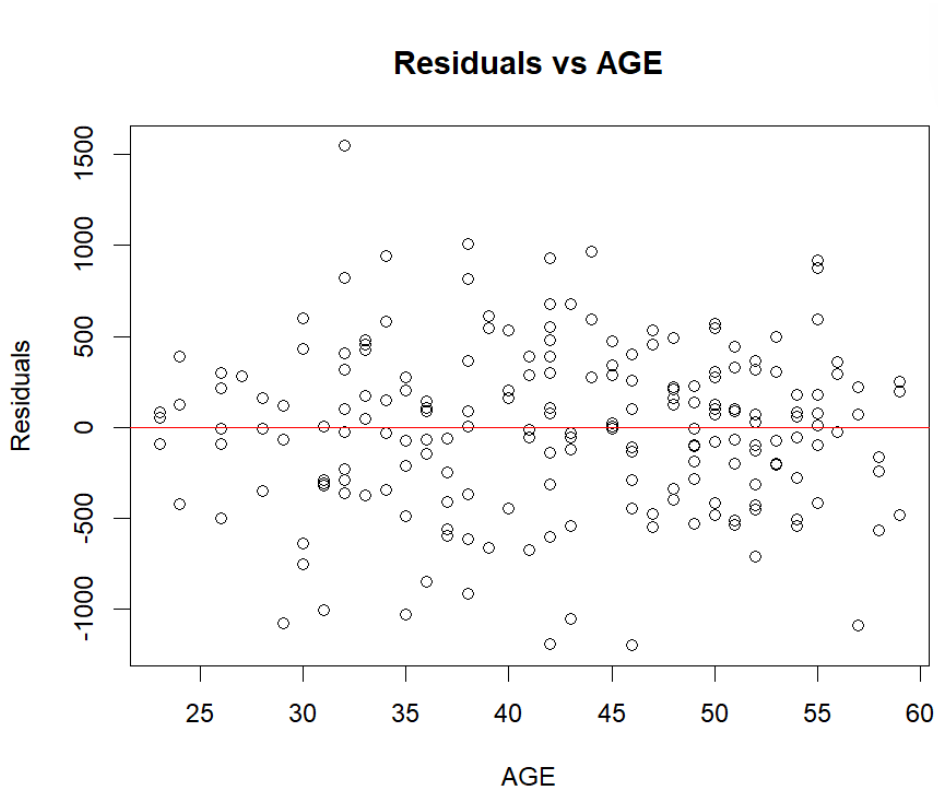
(a)

```
Call:
lm(formula = miles ~ income + age + kids, data = vacation)

Residuals:
     Min       1Q   Median       3Q      Max
-1198.14  -295.31    17.98   287.54  1549.41

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -391.548    169.775  -2.306   0.0221 *
income        14.201      1.800   7.889 2.10e-13 ***
age           15.741      3.757   4.189 4.23e-05 ***
kids         -81.826     27.130  -3.016   0.0029 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 452.3 on 196 degrees of freedom
Multiple R-squared:  0.3406,    Adjusted R-squared:  0.3305
F-statistic: 33.75 on 3 and 196 DF,  p-value: < 2.2e-16
```

```
> confint(ols_model, "kids", level = 0.95)
          2.5 %    97.5 %
kids -135.3298 -28.32302
```

(b)





Income 的殘差有隨 income 增加而變大的趨勢，AGE 則無明顯趨勢

(c)

H0: 殘差變異數相等，H1: 高收入的殘差變異數大於低收入組

```
        Goldfeld-Quandt test

data:  miles ~ income + age + kids
GQ = 3.1041, df1 = 86, df2 = 86, p-value = 1.64e-07
alternative hypothesis: variance increases from segment 1 to 2
```

p-value < 0.05, we reject H0

(d)

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept) -391.5480   142.6548 -2.7447 0.0066190 **
income        14.2013     1.9389  7.3246 6.083e-12 ***
age           15.7409     3.9657  3.9692 0.0001011 ***
kids         -81.8264    29.1544 -2.8067 0.0055112 **
```

Kids 的係數與 part (a)相同，標準誤增加。更加保守

```
                  2.5 %     97.5 %
(Intercept) -672.883378 -110.21263
income        10.377633   18.02503
age            7.919934   23.56191
kids        -139.322973  -24.32986
```

信賴區間相比(a)，變得更寬

(e)

conventional GLS (使用權重=1/income^2，讓低收入者有較大權重)

```
Call:
lm(formula = miles ~ income + age + kids, data = vacation, weights = weights)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-15.1907  -4.9555   0.2488   4.3832  18.5462

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -424.996    121.444  -3.500 0.000577 ***
income        13.947      1.481   9.420  < 2e-16 ***
age           16.717      3.025   5.527 1.03e-07 ***
kids         -76.806     21.848  -3.515 0.000545 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.765 on 196 degrees of freedom
Multiple R-squared:  0.4573,    Adjusted R-squared:  0.449
F-statistic: 55.06 on 3 and 196 DF,  p-value: < 2.2e-16
```

```
                  2.5 %     97.5 %
(Intercept) -664.50116 -185.49119
income        11.02744   16.86718
age           10.75260   22.68240
kids        -119.89450  -33.71808
```

robust GLS

```
t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) -424.9962    95.8035 -4.4361 1.526e-05 ***
income        13.9473     1.3470 10.3545 < 2.2e-16 ***
age           16.7175     2.7974  5.9761 1.061e-08 ***
kids         -76.8063    22.6186 -3.3957 0.0008286 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
                  2.5 %     97.5 %
(Intercept) -613.93428 -236.05807
income        11.29086   16.60376
age           11.20062   22.23438
kids        -121.41339  -32.19919
```

信賴區間變得更窄，更為精準

Q18

(a)

H_0: $\sigma^2\_M = \sigma^2\_F$（男性與女性的誤差變異相等）

H_1: $\sigma^2\_M \neq \sigma^2\_F$（不相等）

```
df :  5415 & 4366
F-test statistic :  1.05076
5% thres :  0.9452566  and  1.058097
```

test statistic = 1.05076

we don't reject H0, no enough evidence support $\sigma^2\_M \neq \sigma^2\_F$

(b)

H0: homoscedasticity, H1: heteroskedasticity

```
> qchisq(0.99, df = 3)
[1] 11.34487
>
> LM1
[1] 23.55681
```

Use METRO, FEMALE, BLACK

NR^2= 23.55681 > 11.34487, we reject H0

METRO, FEMALE, BLACK 存在異質變異性

```
> qchisq(0.99, df = 9)
[1] 21.66599
>
> LM2
[1] 109.4243
```

Use all explanatory variables

NR^2= 109.4243 > 21.66599, we reject H0

使用所有解釋變數時，存在更顯著的異質變異性

(c)

```
        studentized Breusch-Pagan test

data:  model_ols
BP = 194.44, df = 44, p-value < 2.2e-16
```

H0: homoscedasticity, H1: heteroskedasticity

test statistic = 194.44, p-value < 0.05, reject H0

模型中的異質變異性，和解釋變數本身、平方項、交乘項有關

(d)

```
t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  1.2014e+00  3.2777e-02  36.6527 < 2.2e-16 ***
educ         1.0123e-01  1.9048e-03  53.1431 < 2.2e-16 ***
exper        2.9622e-02  1.3142e-03  22.5391 < 2.2e-16 ***
I(exper^2)  -4.4578e-04  2.7583e-05 -16.1615 < 2.2e-16 ***
female      -1.6550e-01  9.4834e-03 -17.4517 < 2.2e-16 ***
black       -1.1153e-01  1.6085e-02  -6.9333 4.371e-12 ***
metro        1.1902e-01  1.1576e-02  10.2814 < 2.2e-16 ***
south       -4.5755e-02  1.3895e-02  -3.2931 0.0009946 ***
midwest     -6.3943e-02  1.3717e-02  -4.6615 3.180e-06 ***
west        -6.5891e-03  1.4549e-02  -0.4529 0.6506470
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> confint(main_model)
                  2.5 %        97.5 %
(Intercept)  1.1384302204  1.2643338265
educ         0.0977830603  0.1046761665
exper        0.0270727569  0.0321706349
I(exper^2)  -0.0004974407 -0.0003941203
female      -0.1841810529 -0.1468229075
black       -0.1447358548 -0.0783146449
metro        0.0948966363  0.1431441846
south       -0.0723384657 -0.0191724010
midwest     -0.0915893895 -0.0362971859
west        -0.0348207138  0.0216425095
```

```
> confint_robust
                        [,1]          [,2]
(Intercept)  1.1371382559  1.2656257910
educ         0.0974961106  0.1049631162
exper        0.0270457914  0.0321976003
I(exper^2)  -0.0004998427 -0.0003917182
female      -0.1840894784 -0.1469144820
black       -0.1430527905 -0.0799977092
metro        0.0963309747  0.1417098462
south       -0.0729887326 -0.0185221340
midwest     -0.0908291013 -0.0370574741
west        -0.0351059530  0.0219277486
```

	變化方向
(Intercept)	變寬
educ	變寬
exper	變寬
I(exper^2)	變寬
female	變窄
black	變窄
metro	變窄
south	變寬
midwest	變窄
west	變寬

CI 變寬，表示原本可能低估不確定性，Robust 下更保守。

CI 變窄，表示原本可能高估不確定，Robust 下更精確。

不存在不一致性，Robust OLS 是一致性估計

(e)

FGLS 結果

```
Call:
lm(formula = log(wage) ~ educ + exper + I(exper^2) + female +
    black + metro + south + midwest + west, data = cps5, weights = weights)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-4.7199 -0.6168 -0.0112  0.6182  6.1542

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.196e+00  3.184e-02  37.571  < 2e-16 ***
educ         1.015e-01  1.761e-03  57.604  < 2e-16 ***
exper        2.986e-02  1.299e-03  22.988  < 2e-16 ***
I(exper^2)  -4.510e-04  2.657e-05 -16.971  < 2e-16 ***
female      -1.658e-01  9.505e-03 -17.446  < 2e-16 ***
black       -1.112e-01  1.697e-02  -6.553 5.91e-11 ***
metro        1.184e-01  1.186e-02   9.979  < 2e-16 ***
south       -4.527e-02  1.354e-02  -3.343 0.000833 ***
midwest     -6.355e-02  1.405e-02  -4.524 6.13e-06 ***
west        -6.060e-03  1.439e-02  -0.421 0.673671
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.948 on 9789 degrees of freedom
Multiple R-squared:  0.3193,    Adjusted R-squared:  0.3187
F-statistic: 510.2 on 9 and 9789 DF,  p-value: < 2.2e-16
```

SE、CI 比較

	OLS_Robust_SE	FGLS_SE	SE_Change	OLS_Robust_Width	FGLS_Width	CI_Change
(Intercept)	3.277743e-02	3.184437e-02	smaller	0.1284875351	0.124843079	narrower
educ	1.904848e-03	1.761461e-03	smaller	0.0074670055	0.006905656	narrower
exper	1.314237e-03	1.298873e-03	smaller	0.0051518089	0.005092118	narrower
I(exper^2)	2.758278e-05	2.657195e-05	smaller	0.0001081245	0.000104173	narrower
female	9.483417e-03	9.505454e-03	larger	0.0371749964	0.037265303	wider
black	1.608548e-02	1.696582e-02	larger	0.0630550813	0.066513034	wider
metro	1.157624e-02	1.186360e-02	larger	0.0453788716	0.046510222	wider
south	1.389454e-02	1.354227e-02	smaller	0.0544665986	0.053091297	narrower
midwest	1.371725e-02	1.404549e-02	larger	0.0537716271	0.055064111	wider
west	1.454941e-02	1.438967e-02	smaller	0.0570337016	0.056413445	narrower

係數的估計值變化不大。

(f)

```
t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  1.1964e+00  3.2561e-02  36.7443 < 2.2e-16 ***
educ         1.0147e-01  1.8986e-03  53.4432 < 2.2e-16 ***
exper        2.9858e-02  1.3097e-03  22.7982 < 2.2e-16 ***
I(exper^2)  -4.5096e-04  2.7508e-05 -16.3934 < 2.2e-16 ***
female      -1.6583e-01  9.4550e-03 -17.5388 < 2.2e-16 ***
black       -1.1118e-01  1.5986e-02  -6.9548 3.757e-12 ***
metro        1.1838e-01  1.1575e-02  10.2279 < 2.2e-16 ***
south       -4.5266e-02  1.3857e-02  -3.2667  0.001092 **
midwest     -6.3548e-02  1.3705e-02  -4.6369 3.582e-06 ***
west        -6.0599e-03  1.4522e-02  -0.4173  0.676472
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

CI 比較

	OLS.Robust	FGLS.Normal	FGLS.Robust	Compare_FGLS_Norm	Compare_FGLS_Rob
(Intercept)	0.1285	0.1163	0.1508	narrower	wider
educ	0.0075	0.0064	0.0094	narrower	wider
exper	0.0052	0.0047	0.0054	narrower	wider
I(exper^2)	0.0001	0.0001	0.0001	narrower	wider
female	0.0372	0.0368	0.0379	narrower	wider
black	0.0631	0.0628	0.0664	narrower	wider
metro	0.0454	0.0458	0.0452	wider	narrower
south	0.0545	0.0531	0.0569	narrower	wider
midwest	0.0538	0.0551	0.0554	wider	wider
west	0.0570	0.0568	0.0590	narrower	wider

(g)

選擇 FGLS Robust，因為 part (c)中得出數據存在異質變異性，一般 OLS 估計不可靠。

FGLS Robust 同時具備效率與穩健性，是最可靠的呈現方式