1(a) $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{SE(\hat{\beta}_2)} = \frac{0.0708503 - 0}{0.0052325} = 13.54043$

$t_{cri} = \pm 1.960$; $t_{cal} > t_{cri}$

Conclusion We can reject the null hypothesis at the level of significant of 95%.
In other word, we can sure that β_2 is significantly different from zero which mean education has an impact on logarithm of hourly wage.

1(b) $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$

H_a: Otherwise

$F_{cal} = \frac{ESS/df}{RSS/df} = \frac{168.697151/7}{276.282616/1,252} = 109.2094357$

$F_{cri(0.05;7,1252)} = 2.01$

; Since $F_{cal} > F_{cri}$, therefore we can reject the null hypothesis or we can say that all the parameters are not simultaneously equal to zero

---

1(c) $H_0: \beta_7 = \beta_8 = 0$

H_a: Otherwise

$F_{cal} = \frac{ESS_{new} - ESS_{old} / (\text{number of new regressors})}{RSS_{new}/(n-k_{new})} = \frac{168.697151 - 166.011417/2}{276.282816/1,252} = 6.0854205$

$F_{cri(2,1252)} = 3$

; Since $F_{cal} > F_{cri}$, therefore we can reject the null hypothesis or we can make sure that physical attractiveness has an impact on logarithm of hourly wage.

1(d) From table 1.2, we could expect the logarithm of hourly wage increase by 0.0070104 USD for women with above average looks

2a) Yes, normally households in municipal area tend to have higher expenditure because of higher cost of living and household that have more number of children aged under 15 may have higher expenditure because children tend to have higher expenses

---

2b) For β_1 $\quad H_0: \beta_1 = 0$ $\quad t_{cal}(\beta_1) = 43.83$

$H_a: \beta_1 \neq 0$

For β_2 $\quad H_0: \beta_2 = 0$ $\quad t_{cal}(\beta_2) = -15.8$ $\quad t_{cri\,(0.005,\,14905)} = \pm 2.576$

$H_a: \beta_2 \neq 0$

For β_3 $\quad H_0: \beta_3 = 0$ $\quad t_{cal}(\beta_3) = 6.82$

$H_a: \beta_3 \neq 0$

; Since $t_{cal} > t_{cri}$, we can reject the null hypothesis or we can say that β_1, β_2 and β_3 are significantly different from zero

---

2c) $E(\widehat{hhexp_i} \mid area_i = 1, child_i = 3) = 9{,}736 - 2{,}835(1) + 881(3)$

$= 9{,}544$

; Expected value of household expenditure not living in a municipal area with 3 children age under 15 is 9,544 per month

---

2d) For β_1 $\quad H_0: \beta_1 = 0$ $\quad t_{cal}(\beta_1) = 34.38$

$H_a: \beta_1 \neq 0$

For β_2 $\quad H_0: \beta_2 = 0$ $\quad t_{cal}(\beta_2) = -6.55$

$H_a: \beta_2 \neq 0$

For β_3 $\quad H_0: \beta_3 = 0$ $\quad t_{cal}(\beta_3) = 5.17$

$H_a: \beta_3 \neq 0$

For β_4 $\quad H_0: \beta_4 = 0$ $\quad t_{cal}(\beta_4) = -0.25$

$H_a: \beta_4 \neq 0$

$t_{cri\,(\frac{\alpha}{2} = 0.005,\, 14908-3)} = \pm 2.576$

β_1 : $34.38 > 2.576$ $\Rightarrow$ reject H_0

$t_{cal} > t_{cri}$

β_2 : $-6.55 < -2.576$ $\Rightarrow$ reject H_0

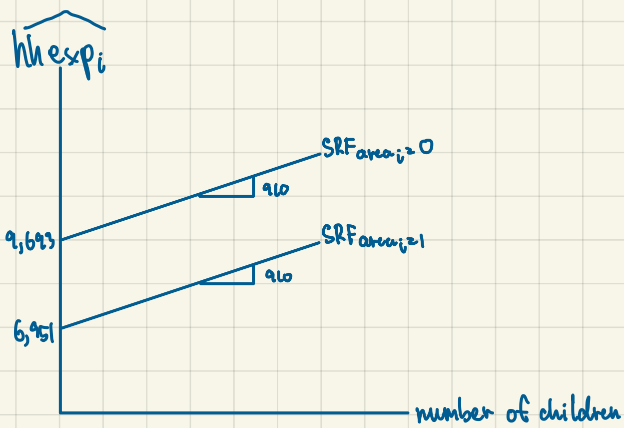
$t_{cal} < -t_{cri}$

β_3 : $5.17 > 2.576$ $\Rightarrow$ reject H_0

$t_{cal} > t_{cri}$

β_4 : $-0.25 > -2.576$ $\Rightarrow$ cannot reject H_0

$t_{cri} > t_{cal} > -t_{cri}$



3a) Following from the rule of thumb that VIF should not exceed 10 and TOL should be closer to 1 rather than 0, VIF of age_i & $agesq_i$ are both exceed 10 and TOL of age_i & $agesq_i$ is further from 1. Furthermore, when we look at the mean VIF, we could see that it's more than 10 which mean there might be some variable have problem. So we could suspect that age_i and $agesq_i$ might be linearly correlated.

3b) Yes, because the variable that VIF exceed 10 would lead to multicollinearity problem and we should choose to remove $agesq_i$ because this variable contain higher VIF.

3c) Yes, heteroscedasticity is present in this model because there is no statistical significant relationship between $\hat{u}_i^2$ and $weekhot_i$.

3d) H_0: homoscedasticity
H_a: otherwise

$$F_{cal} = \frac{R^2_{\hat{u}_i^2}/(k)}{(1-R^2_{\hat{u}_i^2})/(n-k-1)} = \frac{0.0184/5}{(1-0.0184)/(2032-5-1)} = 7.595436$$

$$F_{cri} = F_{0.05(5,2026)} = 4.36$$

; Since $F_{cal} > F_{cri}$, we can reject H_0 at significance level of 0.5%. In other word, we can sure that heteroscedasticity is present in our model