Sr. No.

Date:

30/sep/17

ECON

Assignment - 5

Piyush Kulkarni

pxk161130

Q.6.4

a) $H_0 : \beta_K = 0$

$H_1 : \beta_K \neq 0 \quad t_c = \pm 2$

Equation A

C → 3.96616 → significant

Edu → 1.254 → Not significant

Edu^2 → 1.8875 → Not Significant

Exper → 4.604 → Significant

Exper^2 → −5.38 → significant

Edu×Exper → −1.051 → Not significant

HRSWK → 8.357 → significant

Equation - B → calculating using R.

Significant	Not significant
C	edu
edu^2	
exper^2	
exper	
hrswk	

Signature

Date : Sr. No.

- Equation C → using R

significant	Not significant
C	edu
hrswk	edu^2

- Equation D

significant	Not significant
C	—
$exper^2$	
exper	
hswk	

- Eq - E

significant	Not significant
C	
edu	
exper	
$exper^2$	
hrswk	

b) Restriction → $\beta_6 = 0$
that is edu×exper has no effect on the model.

Signature

Date: ________ Sr. No. ________

$H_0: \beta_6 = 0$
$H_1: \beta_6 \neq 0$

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/N-K}$$

$$= \frac{2226674 - 2224166 / 1}{2224166 / 993}$$

$$= 1.119$$

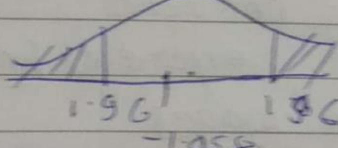
$F_c = 3.85084$

$F_c > F_{stat}$

We fail to reject the Null. the term eduxexper is not signifi cant

$t_c(0.975, 993) = 1.9623$
$t_{stat}\ \beta_6 = -1.058$

$H_0: \beta_6 = 0$
$H_1: \beta_6 \neq 0$



$t_c > |t_{stat}|$

We fail to reject the Null hypothesis

eduxexper is not significant

Signature ________

c) Restrictions imposed are

~~$\beta_4 \neq 0$~~ $\beta_4, \beta_5, \beta_6 = 0$

that is

$H_0: \beta_4 = 0, \beta_5 = 0, \beta_6 = 0$

$H_1:$ At least β_4 or β_5 or $\beta_6 \neq 0$

$$F_{val} = \frac{\frac{SSE_R - SSE_U}{J}}{SSE_U / N-K}$$

$$= \frac{233.8317 - 222.416}{222.416/993}$$

$$= 16.985$$

$$F_{(0.95, 3, 993)} = 2.61$$

$F_{val} > F_{critical}$

We ~~fail to~~ reject the null hypothesis

At least on of β_4 or β_5 or β_6 is non zero

$\therefore$ experience has an effect on wages

Date: ______ Sr. No. ______

d) Restriction → β_2 & $\beta_3 = 0$.

$H_0: \beta_2 = 0, \beta_3 = 0$

H_1: At least one of β_2 or β_3 is non zero

$$F_{val} = \frac{SSE_R - SSE_U / J}{SSE_U / N-K}$$

$= \cancel{280.5061/2}$

$$= \frac{280.5061 - 222.6674 / 2}{222.6674 / 994}$$

$= 129.1$

$F_{stat} = 3.004 = F(0.95, 2, 994)$

We reject the null

∴ We can say that edu has an effect on wages & is relevant.

e) Restriction → β_3 & $\beta_6 = 0$

$H_0: \beta_3 = 0, \beta_6 = 0$

H_1: At least one of β_3 or β_6 non zero

$$F_{val} = \frac{(223.6716 - 222.4166)/2}{222.4166 / 993}$$

$= 2.802$

$F_{stat} = 3.004$

Signature ______

We fail to reject the Null hypothesis
The test suggest adding the interaction term is not helping the model & ~~be~~ not relevant

f) Model E is most preffered of all. All the coeficients are significant and the term edu^2 & edu x expr is omitted which was irrelevant as shown by above calculations.

g) $$(AIC)_D = \ln\left(\frac{SSE}{N}\right) + \frac{2K}{N}$$

$$= \ln\left(\frac{280.5061}{1000}\right) + \frac{8}{1000}$$

$$= -1.263$$

$$(SC)_A = \ln\left(\frac{SSE}{N}\right) + \frac{K(\ln(N))}{N}$$

$$= -1.4548$$

eqn B favoured by AIC
eqn E favoured by SC

Date :

Sr. No.

$edu = exp$ & $edu^2 = exp^2$

6.5 $H_0: \beta_2 = \beta_4$ & $\beta_3 = \beta_5$

a) H_1: ~~$\beta_2 =$~~ At least one of $\beta_2 \neq 4$ & $\beta_3 \neq \beta_5$

b) ~~ln(wage)~~ ~~w~~ Though have same effect, we can't omit, one of the variable completly.

$\therefore \ln(wage) = \beta_1 + \beta_2(edu + exper) + \beta_3(edu^2 + exper^2) + \beta_6 \, hrswk + e$

c) ~~$F = (254.1726 - 222.4166)/2$~~

Using model B

$$F = \frac{(254.1726 - 222.6674)/2}{222.6674/994} = 70.3281$$

$F_{val} = 3.004$

$\therefore$ we reject the null

Edu & experience don't have same effect on wages.

Signature

```
model <-lm(log(wage)~ educ + I(educ^2) + exper + I(exper^2) + I(educ*exper) + hrswk, data = cps4c_small)
>
> summary(model)

Call:
lm(formula = log(wage) ~ educ + I(educ^2) + exper + I(exper^2) +
    I(educ * exper) + hrswk, data = cps4c_small)

Residuals:
     Min       1Q   Median       3Q      Max
-2.30371 -0.29260 -0.00782  0.31469  1.82924

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      1.055e+00  2.659e-01   3.969 7.74e-05 ***
educ             4.983e-02  3.969e-02   1.255   0.2096
I(educ^2)        3.193e-03  1.693e-03   1.886   0.0595 .
exper            3.727e-02  8.144e-03   4.577 5.32e-06 ***
I(exper^2)      -4.849e-04  9.013e-05  -5.380 9.29e-08 ***
I(educ * exper) -5.104e-04  4.824e-04  -1.058   0.2903
hrswk            1.145e-02  1.374e-03   8.336 2.53e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4733 on 993 degrees of freedom
Multiple R-squared:  0.3173,	Adjusted R-squared:  0.3132
F-statistic: 76.93 on 6 and 993 DF,  p-value: < 2.2e-16

>
> model <-lm(log(wage)~ educ + I(educ^2) + exper + I(exper^2) +  hrswk, data = cps4c_small)
> summary(model)
```

```
Call:
lm(formula = log(wage) ~ educ + I(educ^2) + exper + I(exper^2) +
    hrswk, data = cps4c_small)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3093 -0.2941 -0.0135  0.3201  1.8066

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.252e+00  1.901e-01   6.585 7.35e-11 ***
educ         2.894e-02  3.444e-02   0.840   0.4008
I(educ^2)    3.523e-03  1.664e-03   2.117   0.0345 *
exper        3.034e-02  4.834e-03   6.276 5.17e-10 ***
I(exper^2)  -4.564e-04  8.602e-05  -5.306 1.38e-07 ***
hrswk        1.156e-02  1.370e-03   8.434  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4733 on 994 degrees of freedom
Multiple R-squared:  0.3166,	Adjusted R-squared:  0.3131
F-statistic: 92.08 on 5 and 994 DF,  p-value: < 2.2e-16

>
> model <-lm(log(wage)~ educ + I(educ^2) +  hrswk, data = cps4c_small)
> summary(model)

Call:
lm(formula = log(wage) ~ educ + I(educ^2) + hrswk, data = cps4c_small)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-2.23636 -0.29448 -0.01522  0.30921  1.86710

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.573398   0.187816   8.377   <2e-16 ***
educ        0.036606   0.035043   1.045   0.2965
I(educ^2)   0.002930   0.001696   1.728   0.0844 .
hrswk       0.013453   0.001363   9.869   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4845 on 996 degrees of freedom
Multiple R-squared:  0.2823,	Adjusted R-squared:  0.2801
F-statistic: 130.6 on 3 and 996 DF,  p-value: < 2.2e-16

>
> model <-lm(log(wage)~  exper + I(exper^2) +  hrswk, data = cps4c_small)
> summary(model)

Call:
lm(formula = log(wage) ~ exper + I(exper^2) + hrswk, data = cps4c_small)

Residuals:
     Min       1Q   Median       3Q      Max
-2.06876 -0.33812 -0.02965  0.35144  1.66099

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.917e+00  8.046e-02  23.823  < 2e-16 ***
exper        2.790e-02  5.401e-03   5.166 2.89e-07 ***
I(exper^2)  -4.703e-04  9.604e-05  -4.897 1.13e-06 ***
hrswk        1.524e-02  1.508e-03  10.106  < 2e-16 ***
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5307 on 996 degrees of freedom
Multiple R-squared:  0.139,	Adjusted R-squared:  0.1364
F-statistic: 53.61 on 3 and 996 DF,  p-value: < 2.2e-16

>
> model <-lm(log(wage)~ educ + exper + I(exper^2)  + hrswk, data = cps4c_small)
> summary(model)

Call:
lm(formula = log(wage) ~ educ + exper + I(exper^2) + hrswk, data = cps4c_small)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2790 -0.2966 -0.0168  0.3218  1.8889

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.044e-01  9.603e-02   9.417  < 2e-16 ***
educ         1.006e-01  6.328e-03  15.901  < 2e-16 ***
exper        2.951e-02  4.826e-03   6.114 1.39e-09 ***
I(exper^2)  -4.401e-04  8.582e-05  -5.128 3.52e-07 ***
hrswk        1.188e-02  1.364e-03   8.709  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4741 on 995 degrees of freedom
Multiple R-squared:  0.3135,	Adjusted R-squared:  0.3107
F-statistic: 113.6 on 4 and 995 DF,  p-value: < 2.2e-16
```

```
>

> qt(.975,993)

[1] 1.962356

> ((254.176-222.6674)/2)/(222.6674/994)

[1] 70.3281

>

> qf(.95,2,994)

[1] 3.004779

> ((280.5061-222.6674)/2)/(222.6674/994)

[1] 129.0976

>

> log(222.4166/1000)+(7*log(1000)/1000)

[1] -1.454849

>
```

Date :

Sr. No.

Q 6.15

a) ~~SPrice = 111.54.54~~

$$\widehat{SPrice} = \underset{6555.11}{11154.29} + \underset{273.15}{10680\, Livarea}$$

se

$$- \underset{80.50}{11.33\, Age} - \underset{1970.01}{15552.44\, beds}$$

$$- \underset{2903.82}{7091.3\, baths}$$

All variables significant at $\alpha = 5$ except age.

b) diff. betn price age = 2 & 10.

$$(Sprice)_{Age=2} - (Sprice)_{Age=10}$$

$$= -11.33 \times 2 - (-11.33 \times 10)$$

$$= 90.61$$

Keeping All the things const, price diff betn 2 yr old house & 10 year hous is \$90.61.

~~Age~~

Age 95% conf interval

	-169.2429	146.575
Age 2	-338.4858	293.15
Age 10	-1692.429	1465.75

Signature

Date : ________ Sr. No. ____

95% conf interval ~~for esti~~ for difference

$= 1353.94 \ \& \ -1172.6$

$\Rightarrow -1172.6 \ \& \ 1353.94$

c) $(\text{SPrice})_{\text{Lar}=22} - (\text{SPrice})_{\text{Larea}=20}$

$= 10680 \times 22 - 10680 \times 20$

$= \cancel{106} \ \$21360$

Keeping all things const, increase in price $\rightarrow$ \$21,360

Price increase

$= 22\beta_2 - 20\beta_2$

$= 2\beta$

~~$H_0: \cancel{2}\beta = \cancel{20,000}\ 10000$~~

~~$\beta_2 \neq 10000$~~

~~$H_1: \beta_2 \neq 10,000$~~

t

$H_0: \beta_2 < 10,000$

$H_1: \beta_2 \geq 10,000$

$t_c = 1.645$

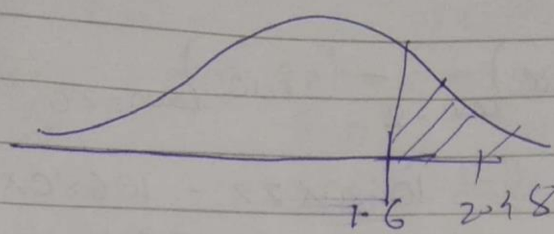
Signature ________

Date :

Sr. No.

$$t_{stat} = \frac{10680 - 10000}{273.15}$$

$$= 2.48$$



We reject the null hypothis
∴ increase in price would be
graterthan or equal ~~to~~ \$20,000

d) Addition of 1 bed & 200 sqft area

β_2 livArea + β_4 bed

$\cancel{20,000\ \beta_2 + \beta_4}$

$\cancel{22,000\ \beta_2 +}$

Change in liv Area = 20,000 - 18,000

change in bed = 1

∴ change in price = $2\beta_2 + \beta_4$

Signature

Date: Sr. No.

$= 2 \times 10680 - 15552.44$

$= \cancel{4}$

$= \$5807.56$

Keeping all things const, ~~Liv area~~ price will increase by $5807.56

95% conf interval Liv bed

$= 2\beta_2 + \beta_4 \pm t_c \times se(2\beta_2 + \beta_4)$

$se(2\beta_2 + b_4) = \sqrt{2^2 \times (273.15)^2 + (1970.01)^2 + 2 \times 2 \times (-170680.218)}$

$= 1869.\cancel{3}93$

~~= 2 × 10680 =~~

$= 5807.56 \pm 1.96 \times 1869.9$

$= \cancel{17}\ 4$

$= \cancel{9471.408 \ \& \ 2143.712}$

~~95% conf ⇒ $2143.712 & 9471.408~~

95% conf int ⇒ $2142.49 & $9472.6

Signature

Q.6.16 a) sprice = 7955.736 + 2994.652 livarea + 169.09 livarea^2 − 830.37 Age + 14.23 Age^2 − 11921.923 bed − 4971.063 baths

se: 8744.29, β_2: 772.29, 16.125, β_3; 197.79 (β_4), 3.35 (β_5), 1927.05, 2797.366

b) $H_0: \beta_3 = 0, \ \beta_5 = 0$

H_1: At least one of $\beta_3 \neq 0$ or $\beta_5 \neq 2$

$$F = \frac{SSE_R - SSE_U / J}{SSE_U / N-K}$$

$$= \frac{2.111 \times 10^{12} - 1.94435 \times 10^{12} / 2}{\frac{1.994 \times 10^{12}}{1493}}$$

$$= \frac{0.121/2 = 0.0605}{1.33555 \times 10^{3}}$$

$$= \frac{10^{12}(2.11 - 1.94)}{1.94 \times 10^{12}} \times \frac{1493}{2}$$

$$= 65.41$$

Date: ______ Sr. No. ______

$F_{c}val = (0.95, 2, 1493)$
$= 3.001$

As Fval > Fcritical
We reject the null
∴ At least one of Age^2 or $Livarea^2$ improves model

~~b~~

c→b Price diff ~~for~~ age 2 & 10

$= 2\beta_4 + 2^2\beta_5 - (10\beta_4 + \beta_5 10^2)$

$= -8\beta_4 - 96\beta_5$

$\cancel{= -8 \times 830.37 - 96 \times 14.23}$
$\cancel{-6642.96 - 1366.08}$

$=$

$= -8 \times -830.37 - 96 \times 14.23$

$= 6642.96 - 1366.08$

$= 5276.88$

Keeping all things const, ~~a~~ diff betn house 2 & 10 year old is \$5276.88.

95% conf $= -8\beta_4 - 96\beta_5 \pm t_c \times se(-8b_3 - 96b_4)$

$se(-8B_3 - 96B_4) = \sqrt{+64 \times 39116.36 + 96^2 \times 11.261 - 8 \times 96 \times 2 \times 610.74}$

$= 1291.93$

Signature ______

Date: Sr. No.

$95\% \text{ conf} = 5276.88 \pm 1.96 \times 1291.94$

$= \cancel{2757.54 \ \& \ 7796.163}$

$= \$ \ 2744.67 \ \& \ 7809.082$

c → c diff in price

$= (22\beta_2 + 22^2\beta_3) - (20\beta_2 + 20^2\beta_3)$

$= 2\beta_2 + 84\beta_3$

$= 2 \times 2994.65 + 84 \times 169.09$

$= \$20192.86$

Keeping all thing const, extension by 200 sq ft will increase \$ 20192.86

$H_0: 2\beta_2 + 84\beta_3 \leq 20{,}000$

$H_0: \beta_2 + 42\beta_3 \leq 10{,}000$

$H_1: \beta_2 + 42\beta_3 > 10{,}000$

$t_c(0.95, 1493) = 1.96$

Date: ______ Sr. No. ______

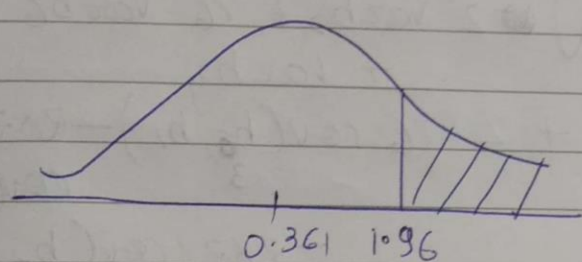
$$t = \frac{2994.65 + 42 \times 16909 - 10,000}{se(b_2 + 42 b_3)}$$

se $b_2 + 42 b_3$

$$= \sqrt{5.96 \times 10^5 + 42^2 \times 2.6 \times 10^2 + 2 \times 42 \times -1.171 \times 10^3}$$

$$= \sqrt{\quad}$$

$$= \frac{96.43}{266.45} = 0.361$$



We fail ~~to~~ to reject the null hypo.
∴ increase in price will not be grater than ~~to~~ 20,000.

d) Adding bedroom ~~of~~ of 200 sq ft.

Price
$$= (20\beta_2 + 20^2\beta_3 + \beta_6(bed+1)) - (18\beta_2 + 18^2\beta_3 + \beta_6\, bed)$$

$$= 2\beta_2 + 76\beta_3 + \beta_4$$

Signature ______

$2b_2 + 76b_3 + b_4$

price change $= 2 \times 2994.65 + 76 \times 1699.09$
$- 11921.92$

$= 6918.22$

$se = \sqrt{\mathrm{var}\, b_2 + \mathrm{var}\, b_3 + \mathrm{var}\, b_6 + 2\,\mathrm{cov}(b_3, b_6) - 2\,\mathrm{cov}(b_2 b_3) - \mathrm{cov}(b_2, b_6)}$

$= \sqrt{2^2 \mathrm{var}\, b_2 + 76^2 \mathrm{var}\, b_6 + \mathrm{var}\, b_4 + 2 \times 76\, \mathrm{cov}(b_3, b_4) - 2 \times 2 \times 76\, \mathrm{cov}(b_2, b_3) - 2 \times 2 \times \mathrm{cov}(b_3, b_4)}$

$= \sqrt{4 \times 5.96 \times 10^5 + 76^2 \times 2.60 \times 10^2 - 169.24}$

$\cancel{+ 2 \times 76 \times 5.89 \times 10^3} \cancel{+ 2 \times 2 \times 76 \times 42215.9}$

$\cancel{2 \times 2 \times} + \cancel{2 \times 2 \times 76 \times 1.71 \times 10^4}$
$\cancel{- 2 \times 2 \times}$

$+ 2 \times 76 \times 5897.07 + 2 \times 2 \times 76 \times 1.17 \times 10^4$
$+ 2 \times 2 \times 422159.08$

```
> model <- lm(sprice~ livarea + age + beds + baths, data =stckton4)
> summary(model)

Call:
lm(formula = sprice ~ livarea + age + beds + baths, data = stckton4)

Residuals:
    Min      1Q  Median      3Q     Max 
-209307  -17933   -2221   13873  616860 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  11154.29    6555.11   1.702   0.0890 .  
livarea      10680.00     273.15  39.100  < 2e-16 ***
age            -11.33      80.50  -0.141   0.8881    
beds        -15552.44    1970.01  -7.895 5.59e-15 ***
baths        -7019.30    2903.82  -2.417   0.0158 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37580 on 1495 degrees of freedom
Multiple R-squared:  0.648,	Adjusted R-squared:  0.647 
F-statistic:   688 on 4 and 1495 DF,  p-value: < 2.2e-16

> confint(model, 'stckton4.age', level =0.95)
             2.5 % 97.5 %
stckton4.age    NA     NA
> confint(model)
                 2.5 %     97.5 %
(Intercept)  -1703.9071  24012.484
livarea      10144.2059  11215.798
age           -169.2429    146.575
```

```
beds       -19416.7095 -11688.173

baths      -12715.2821  -1323.310

> qt(.975,1495)

[1] 1.961552

> cor(model)

Error in cor(model) : supply both 'x' and 'y' or a matrix-like 'x'

> vcov(model)

          (Intercept)    livarea       age        beds      baths

(Intercept)  42969508.8  390445.711 -287485.490 -7707291.996 -7951852.97

livarea       390445.7   74610.431   -2782.794  -170680.218  -477425.43

age          -287485.5   -2782.794    6480.578     9593.284    75436.21

beds        -7707292.0 -170680.218    9593.284  3880921.738 -1122463.21

baths       -7951853.0 -477425.434   75436.212 -1122463.209  8432146.17

> sqrt((4*74610.431)+3880921.738-(4*170680.218))

[1] 1869.931

> 5807.56 + (1.96*1869.931)

[1] 9472.625

> model <- lm(sprice~ livarea + I(livarea^2) + age + I(age^2) + beds + baths, data =stckton4)

> summary(model)


Call:

lm(formula = sprice ~ livarea + I(livarea^2) + age + I(age^2) +

    beds + baths, data = stckton4)


Residuals:

    Min      1Q  Median      3Q     Max

-233961  -16324   -2567   11364  618482


Coefficients:

             Estimate Std. Error t value Pr(>|t|)

(Intercept)   79755.736   8744.293   9.121  < 2e-16 ***

livarea       2994.652    772.295   3.878  0.00011 ***
```

```
I(livarea^2)    169.092     16.125  10.486  < 2e-16 ***
age           -830.379    197.779  -4.199 2.85e-05 ***
I(age^2)        14.233      3.356   4.241 2.36e-05 ***
beds        -11921.923   1927.050  -6.187 7.92e-10 ***
baths        -4971.063   2797.366  -1.777  0.07576 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 36080 on 1493 degrees of freedom
Multiple R-squared:  0.6759,      Adjusted R-squared:  0.6746
F-statistic:   519 on 6 and 1493 DF,  p-value: < 2.2e-16

> a<- lm(sprice~ livarea + age + beds + baths, data =stckton4)
> anova(a)
Analysis of Variance Table

Response: sprice
           Df     Sum Sq    Mean Sq   F value  Pr(>F)
livarea      1 3.7700e+12 3.7700e+12 2669.7728 < 2e-16 ***
age          1 4.7308e+09 4.7308e+09    3.3501 0.06740 .
beds         1 1.0286e+11 1.0286e+11   72.8435 < 2e-16 ***
baths        1 8.2513e+09 8.2513e+09    5.8432 0.01576 *
Residuals 1495 2.1111e+12 1.4121e+09
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(model)
Analysis of Variance Table

Response: sprice
             Df     Sum Sq    Mean Sq   F value    Pr(>F)
livarea        1 3.7700e+12 3.7700e+12 2896.1544 < 2.2e-16 ***
I(livarea^2)    1 1.9756e+11 1.9756e+11  151.7625 < 2.2e-16 ***
```

```
age           1 5.7820e+08 5.7820e+08    0.4442   0.50522
I(age^2)       1 2.3953e+10 2.3953e+10   18.4006 1.905e-05 ***
beds          1 5.7271e+10 5.7271e+10   43.9953 4.589e-11 ***
baths         1 4.1108e+09 4.1108e+09    3.1579   0.07576 .
Residuals    1493 1.9435e+12 1.3017e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> qf(.95,2,1493)
[1] 3.001751
> vcov(model)
            (Intercept)      livarea  I(livarea^2)          age     I(age^2)         beds
(Intercept)  76462668.198 -3.879328e+06  9.348354e+04 -609180.60806 5865.6686678 -5069570.6943
livarea      -3879327.806  5.964402e+05 -1.171059e+04    2426.37281  -32.0083074  -422159.0816
I(livarea^2)    93483.539 -1.171059e+04  2.600261e+02     -45.50172   -0.4937243     5897.0703
age          -609180.608  2.426373e+03 -4.550172e+01   39116.36592 -610.7487125    16384.9673
I(age^2)        5865.669 -3.200831e+01 -4.937243e-01    -610.74871   11.2617318     -169.2498
beds        -5069570.694 -4.221591e+05  5.897070e+03   16384.96731 -169.2498000  3713521.1288
baths       -6316381.550 -5.255475e+05  1.825909e+03   33088.47316  662.6230366 -1002664.0836
                  baths
(Intercept)  -6316381.550
livarea      -525547.476
I(livarea^2)     1825.909
age           33088.473
I(age^2)        662.623
beds        -1002664.084
baths        7825258.269
>
> sqrt((64*39116.36)+(96^2*11.261)-(8*96*2*610.74))
[1] 1291.949
> qt(.975,1493)
[1] 1.961554
>
```

```
> 5276.88+(1.96*1291.94)

[1] 7809.082

>

> sqrt((5.96*10^5)+(42^2*2.6*10^2)-(2*42*1.171*10^4))

[1] 266.4583
```

Date: ______

$cov(Liv, bed) = -422159.08$
$cov(Liv^2, bed) = 58970703$
$cov(Liv, Liv^2) = -1.171 \times 10^{4}$
c:

$= \sqrt{\cancel{4 \times 8.0}}$

$= 1802.5$

95% conf = $\cancel{6t}$
$= 69\cancel{8}.22 \pm (1.96 \times 18.02.5)$
18

$= 33853.24 \cancel{10451.12}$

Q. 6.22
a) $pizza = 161.46 - 2.97 Ag + 6.97\, Income - 0.12\, Age \times income$

$H_0: \beta_2 = 0, \ \beta_4 = 0$
H_1: At least one of $\beta_2 \neq 0$, $\beta_4 \neq 0$

$F = \dfrac{819286 - 580609/2}{580609/36} = 7.4$

$F_c = (0.95, 2, 36) = 3.2$

Signature ______

β_3 Income
β_4 Income Age

Date: ______ Sr. No. ______

As $F > F_{criti}$
We reject the null hypo
$\therefore$ Age has effect on pizza expenditaure.

b)

$$\cancel{\frac{\partial(Pizza)}{\partial(Age)} = -2.970}$$

$$\frac{\partial(Pizza)}{\partial(Income)}\Bigg| = 6.97 - 0.12 Age$$

$$Se_{(20)} = \sqrt{7.96 + 20^2 \times 0.0044 - \times 2 \times 20 \times 0.185}$$

Using R to do all the calculation

Age	Point estimate	95% conf
20	4.57	1.48 & 7.65
30	3.37	1.533 & 5.20
40	2.17	1.26 & 3.07
50	0.97	−0.40 & 2.34
55	0.37	−1.57 & 2.31

As age increases, propensity to spend on pizza decreases.

Signature ______

Date: ________

c) we ~~e~~ expect the sign ~~to~~ be negative but ~~gr~~ in model it is positive.

$$Pizza = 109.72 - 2.03\,Age + 14.09\,Inc - 0.42\,age \times incom + 0.004\,Age^2 \times Income$$

pval = 0.85

$\therefore$ the term is not significant

e)

	Pval
Age	0.569
Age×Income	0.264
AG^2×Income	0.401

None of them are significantly diff from zero

$H_0: \beta_2 = \beta_4 = \beta_5 = 0$

H_1: At least one of β_2, β_4 or $\beta_5 \neq 0$

$$Fval = \frac{819286 - 568869 / 3}{568869 / 35}$$

$= 5.135$

$F_{crit} = (0.95, 3, 35) = 2.87$

Signature ________

Date :

Sr. No.

$F_{val} > F_{crit}$

We reject the null hypothesis

$\therefore$ Age have an effect on pizza buying.

f) As we can see, there is a very high corelation betⁿ variable ~~of th~~ above 0.9, we can say that there is a multi-collinearity problem. in both model ~~c t 6~~. c and f.

```
> model <- lm(pizza ~  age + income + I(age*income), data = pizza4)
> summary(model)

Call:
lm(formula = pizza ~ age + income + I(age * income), data = pizza4)

Residuals:
    Min      1Q  Median      3Q     Max
-200.86  -83.82   20.70   85.04  254.23

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     161.46543  120.66341   1.338   0.1892
age              -2.97742    3.35210  -0.888   0.3803
income            6.97991    2.82277   2.473   0.0183 *
I(age * income)  -0.12324    0.06672  -1.847   0.0730 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 127 on 36 degrees of freedom
Multiple R-squared:  0.3873,	Adjusted R-squared:  0.3363
F-statistic: 7.586 on 3 and 36 DF,  p-value: 0.0004681

> anova(model)
Analysis of Variance Table

Response: pizza
                Df Sum Sq Mean Sq F value    Pr(>F)
age              1  44415   44415  2.7539 0.1057056
income           1 267600  267600 16.5923 0.0002432 ***
I(age * income)  1  55028   55028  3.4120 0.0729575 .
Residuals       36 580609   16128
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> model2 <-lm(pizza ~  income , data = pizza4)
> summary(model2)

Call:
lm(formula = pizza ~ income, data = pizza4)

Residuals:
    Min      1Q  Median      3Q     Max 
-260.17 -103.81  -49.86  122.59  337.12 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 128.9803    34.5913   3.729 0.000626 ***
income        1.1213     0.4595   2.440 0.019461 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 146.8 on 38 degrees of freedom
Multiple R-squared:  0.1355,	Adjusted R-squared:  0.1127 
F-statistic: 5.954 on 1 and 38 DF,  p-value: 0.01946

> anova(model2)
Analysis of Variance Table

Response: pizza
          Df Sum Sq Mean Sq F value  Pr(>F)  
income     1 128366  128366  5.9539 0.01946 *
Residuals 38 819286   21560                  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> vcov(model)
                (Intercept)         age       income I(age * income)
(Intercept)     14559.658424 -386.7975523 -269.1518796     6.552844006
age             -386.797552   11.2365799    6.4637905    -0.166083924
income          -269.151880    6.4637905    7.9680170    -0.185924335
I(age * income)    6.552844   -0.1660839   -0.1859243     0.004451389
> sqrt(7.96+ (20*20*0.0044)-(2*20*.185))
[1] 1.523155
> qt(.975,36)
[1] 2.028094
> (6.97-(0.12*20))
[1] 4.57
> (6.97-(0.12*30))
[1] 3.37
> (6.98-(0.12*30))
[1] 3.38
> point <- (6.97-(0.12*20))
> error <-sqrt(7.96+ (20^2*0.0044)-(2*20*.185))
> t <-qt(.975,36)
> point - error
[1] 3.046845
> point + error
[1] 6.093155
> point <- (6.97-(0.12*20))
> error <-sqrt(7.96+ (20^2*0.0044)-(2*20*.185))
> t <-qt(.975,36)
> point - (error*t)
[1] 1.480899
> point + (error*t)
[1] 7.659101
> point <- (6.97-(0.12*20))
> error <-sqrt(7.96+ (20^2*0.0044)-(2*20*.185))
```

```
> t <-qt(.975,36)

> point

[1] 4.57

> point - (error*t)

[1] 1.480899

> point + (error*t)

[1] 7.659101

> point <- (6.97-(0.12*30))

> error <-sqrt(7.96+ (30^2*0.0044)-(2*30*.185))

> t <-qt(.975,36)

> point

[1] 3.37

> point - (error*t)

[1] 1.533483

> point + (error*t)

[1] 5.206517

> point <- (6.97-(0.12*40))

> error <-sqrt(7.96+ (40^2*0.0044)-(2*40*0.185))

> t <-qt(.975,36)

> point

[1] 2.17

> point - (error*t)

[1] 1.263009

> point + (error*t)

[1] 3.076991

> point <- (6.97-(0.12*50))

> error <-sqrt(7.96+ (50^2*0.0044)-(2*50*0.185))

> t <-qt(.975,36)

> point

[1] 0.97

> point - (error*t)

[1] -0.4055203
```

```
> point + (error*t)
[1] 2.34552
> point <- (6.97-(0.12*55))
> error <-sqrt(7.96+ (55^2*0.0044)-(2*55*0.185))
> t <-qt(.975,36)
> point
[1] 0.37
> point - (error*t)
[1] -1.575279
> point + (error*t)
[1] 2.315279
> model3 <- lm(pizza ~  age + income + I(age*income) + I(age^2*income), data = pizza4)
> summary(model3)

Call:
lm(formula = pizza ~ age + income + I(age * income) + I(age^2 *
    income), data = pizza4)

Residuals:
     Min       1Q   Median       3Q      Max
-212.080  -79.979    7.395   81.429  260.074

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      109.720767 135.572473   0.809    0.424
age               -2.038273   3.541904  -0.575    0.569
income            14.096163   8.839862   1.595    0.120
I(age * income)   -0.470371   0.413908  -1.136    0.264
I(age^2 * income)  0.004205   0.004948   0.850    0.401

Residual standard error: 127.5 on 35 degrees of freedom
Multiple R-squared:  0.3997,	Adjusted R-squared:  0.3311
```

```
F-statistic: 5.826 on 4 and 35 DF,  p-value: 0.001057

> anova(model3)
Analysis of Variance Table

Response: pizza
                  Df Sum Sq Mean Sq F value    Pr(>F)
age                1  44415   44415  2.7327 0.1072594
income             1 267600  267600 16.4643 0.0002642 ***
I(age * income)    1  55028   55028  3.3856 0.0742600 .
I(age^2 * income)  1  11739   11739  0.7223 0.4011745
Residuals         35 568869   16253
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> model4 <- lm(pizza ~ income , data = pizza4)
> summary(model4)

Call:
lm(formula = pizza ~ income, data = pizza4)

Residuals:
    Min      1Q  Median      3Q     Max
-260.17 -103.81  -49.86  122.59  337.12

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 128.9803    34.5913   3.729 0.000626 ***
income        1.1213     0.4595   2.440 0.019461 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 146.8 on 38 degrees of freedom
```

Multiple R-squared: 0.1355, Adjusted R-squared: 0.1127

F-statistic: 5.954 on 1 and 38 DF, p-value: 0.01946

```
> anova(model4)
Analysis of Variance Table

Response: pizza
          Df Sum Sq Mean Sq F value  Pr(>F)
income     1 128366  128366  5.9539 0.01946 *
Residuals 38 819286   21560
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> qf(.95,3,35)

ageinc <- pizza$age*pizza$income
pizza1<-cbind(pizza,ageinc)

age2inc <- ageinc * pizza$age
pizza2<-cbind(pizza1,age2inc)

age3inc <- age2inc*pizza$age
pizza3<-cbind(pizza2,age3inc)

pizzacorr = pizza3[c( 'pizza', 'income', 'age', 'ageinc', 'age2inc', 'age3inc')]

cor(pizzacorr)

> pizzacorr1 = pizza3[c( 'pizza', 'income', 'age', 'ageinc', 'age2inc')]
> cor(pizzacorr1)
          pizza    income        age    ageinc   age2inc
pizza 1.0000000 0.3680448 -0.2164912 0.2669991 0.1924232
```

```
income   0.3680448 1.0000000  0.4684973 0.9812392 0.9436177
age     -0.2164912 0.4684973  1.0000000 0.5861949 0.6504045
ageinc   0.2669991 0.9812392  0.5861949 1.0000000 0.9892791
age2inc  0.1924232 0.9436177  0.6504045 0.9892791 1.0000000

> cor(pizzacorr)
            pizza    income        age    ageinc   age2inc   age3inc
pizza    1.0000000 0.3680448 -0.2164912 0.2669991 0.1924232 0.1334731
income   0.3680448 1.0000000  0.4684973 0.9812392 0.9436177 0.8974999
age     -0.2164912 0.4684973  1.0000000 0.5861949 0.6504045 0.6886959
ageinc   0.2669991 0.9812392  0.5861949 1.0000000 0.9892791 0.9635731
age2inc  0.1924232 0.9436177  0.6504045 0.9892791 1.0000000 0.9920828
age3inc  0.1334731 0.8974999  0.6886959 0.9635731 0.9920828 1.0000000
>
[1] 2.874187
```