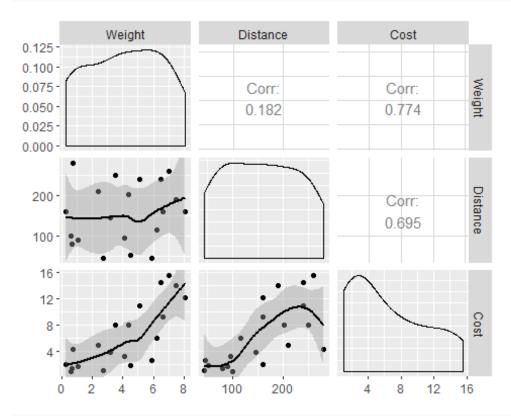
Routput 1 for Question 1

```
express=read.csv("C:\\Users\\thuntida.ngamkham\\OneDrive - University of Calgary\\d
ataset603\\EXPRESS.csv",header=TRUE)
#express=read.csv("~/OneDrive - University of Calgary/MyCoursesThierry/DATA603/data
/dataset603/EXPRESS.csv",header=TRUE)
summary(express)
##
        Weight
                       Distance
                                          Cost
## Min.
          :0.300
                   Min.
                          : 45.00
                                    Min.
                                            : 1.000
## 1st Ou.:2.075
                   1st Qu.: 93.75
                                     1st Ou.: 1.975
## Median :4.250
                   Median :160.00
                                     Median : 4.700
## Mean
          :4.058
                   Mean
                          :156.05
                                    Mean
                                          : 6.335
                                     3rd Qu.: 9.650
## 3rd Qu.:6.275
                    3rd Qu.:216.75
## Max.
         :8.100
                   Max. :280.00
                                    Max.
                                          :15.500
model<-lm(Cost~Distance+Weight,data=express)</pre>
summary(model)
##
## Call:
## lm(formula = Cost ~ Distance + Weight, data = express)
## Residuals:
     Min
             10 Median
                            30
                                 Max
## -2.239 -1.101 -0.129 1.283 2.313
##
## Coefficients:
                Estimate Std. Error t value Pr(>|t|)
##
## (Intercept) -4.672757
                          0.891147 -5.244 6.60e-05 ***
                           0.004602 8.026 3.49e-07 ***
## Distance
               0.036936
## Weight
               1.292414
                          0.137842 9.376 3.95e-08 ***
## ---
                  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Signif. codes:
## Residual standard error: 1.493 on 17 degrees of freedom
## Multiple R-squared: 0.9162, Adjusted R-squared: 0.9063
## F-statistic: 92.89 on 2 and 17 DF, p-value: 7.066e-10
reducedmodel<-lm(Cost~1, data=express)</pre>
coefficients(model)
## (Intercept)
                 Distance
                                Weight
## -4.67275662 0.03693551 1.29241393
anova(reducedmodel, model)
## Analysis of Variance Table
## Model 1: Cost ~ 1
```

```
## Model 2: Cost ~ Distance + Weight
              RSS Df Sum of Sq F
                                         Pr(>F)
    Res.Df
## 1
         19 452.09
        17 37.90 2
                        414.18 92.888 7.066e-10 ***
## 2
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
confint(model)
##
                    2.5 %
                               97.5 %
## (Intercept) -6.55291261 -2.79260064
## Distance
               0.02722672 0.04664431
## Weight
               1.00159293 1.58323493
confint(model,level=0.99)
                    0.5 %
                               99.5 %
## (Intercept) -7.25550628 -2.09000697
## Distance
               0.02359865 0.05027238
## Weight
               0.89291631 1.69191154
newdata1 = data.frame(Weight=150, Distance=6.5)
predict(model, newdata1, interval="predict")
         fit
                  lwr
                           upr
## 1 189.4294 146.5771 232.2817
newdata2 = data.frame(Weight=6.5, Distance=150)
predict(model, newdata2, interval="predict")
         fit
                  lwr
                            upr
## 1 9.268261 5.960167 12.57636
```

R output 2 for Question 2

```
library(GGally)
express=read.csv("C:\\Users\\thuntida.ngamkham\\OneDrive - University of Calgary\\d
ataset603\\EXPRESS.csv",header=TRUE)
#express=read.csv("~/OneDrive - University of Calgary/MyCoursesThierry/DATA603/data
/dataset603/EXPRESS.csv",header=TRUE)
#ggpairs(express)
ggpairs(express,lower = list(continuous = "smooth_loess", combo =
    "facethist", discrete = "facetbar", na = "na"))
```



```
additivemodel<-lm(Cost~Distance+Weight,data=express)</pre>
summary(additivemodel)
##
## Call:
## lm(formula = Cost ~ Distance + Weight, data = express)
##
## Residuals:
     Min
              1Q Median
                            3Q
                                  Max
## -2.239 -1.101 -0.129 1.283 2.313
##
## Coefficients:
               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.672757
                           0.891147 -5.244 6.60e-05 ***
## Distance 0.036936 0.004602 8.026 3.49e-07 ***
```

```
## Weight
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.493 on 17 degrees of freedom
## Multiple R-squared: 0.9162, Adjusted R-squared: 0.9063
## F-statistic: 92.89 on 2 and 17 DF, p-value: 7.066e-10
interacmodel<-lm(Cost~Distance+Weight+Distance*Weight,data=express)
summary(interacmodel)
##
## Call:
## lm(formula = Cost ~ Distance + Weight + Distance * Weight, data = express)
## Residuals:
      Min
               10 Median
##
                              3Q
                                     Max
## -0.7573 -0.4254 -0.1820 0.5079 1.4214
##
## Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                  -0.1405007 0.6481000 -0.217
                                                 0.8311
## Distance
                   0.0077208 0.0039057
                                        1.977
                                                 0.0656 .
## Weight
                   0.0190880 0.1582116
                                         0.121
                                                 0.9055
## Distance:Weight 0.0077957 0.0008977 8.684 1.88e-07 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.6439 on 16 degrees of freedom
## Multiple R-squared: 0.9853, Adjusted R-squared: 0.9826
## F-statistic: 358.2 on 3 and 16 DF, p-value: 7.125e-15
quadmodel<-lm(Cost~Distance+Weight+I(Weight^2)+Distance*Weight,data=express);</pre>
summary(quadmodel)
##
## Call:
## lm(formula = Cost ~ Distance + Weight + I(Weight^2) + Distance *
      Weight, data = express)
##
## Residuals:
##
      Min
               10 Median
                              3Q
                                     Max
## -0.7487 -0.2558 0.0532 0.2266 0.9142
## Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                   0.4746969 0.4584500
                                         1.035 0.316870
## Distance
                   0.0090777 0.0026535
                                         3.421 0.003791 **
## Weight
                  -0.5781705 0.1706879 -3.387 0.004062 **
## I(Weight^2)
                                        4.485 0.000436 ***
                   0.0867388 0.0193380
## Distance:Weight 0.0072587 0.0006176 11.753 5.74e-09 ***
```

```
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.4346 on 15 degrees of freedom
## Multiple R-squared: 0.9937, Adjusted R-squared: 0.9921
## F-statistic: 594.6 on 4 and 15 DF, p-value: 2.541e-16
quadmodel2<-lm(Cost~Distance+I(Distance^2)+Weight+I(Weight^2)+Distance*Weight,data=
express);
summary(quadmodel2)
##
## Call:
## lm(formula = Cost ~ Distance + I(Distance^2) + Weight + I(Weight^2) +
       Distance * Weight, data = express)
##
## Residuals:
                      Median
##
       Min
                 1Q
                                   3Q
                                           Max
## -0.86027 -0.19898 -0.00885 0.16531 0.94396
##
## Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                   8.270e-01 7.023e-01 1.178 0.258588
## Distance
                   4.021e-03 7.998e-03
                                          0.503 0.622999
## I(Distance^2)
                  1.507e-05 2.243e-05 0.672 0.512657
                  -6.091e-01 1.799e-01 -3.386 0.004436 **
## Weight
                   8.975e-02 2.021e-02 4.442 0.000558 ***
## I(Weight^2)
## Distance:Weight 7.327e-03 6.374e-04 11.495 1.62e-08 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.4428 on 14 degrees of freedom
## Multiple R-squared: 0.9939, Adjusted R-squared: 0.9918
## F-statistic: 458.4 on 5 and 14 DF, p-value: 5.371e-15
```

R output 3 for Question 3

```
library(olsrr)
condo=read.csv("C:\\Users\\thuntida.ngamkham\\\\OneDrive - University of Calgary\\d
ataset603\\CONDO.CSV", header = TRUE)
#condo=read.csv("~/OneDrive - University of Calgary/MyCoursesThierry/DATA603/data/d
ataset603/CONDO.CSV", header = TRUE)
model<-lm(PRICE100 ~FLOOR+ DIST+ factor(VIEW)+factor(END)+factor(FURNISH) ,data=con
do )
summary(model)
## Call:
## Im(formula = PRICE100 ~ FLOOR + DIST + factor(VIEW) + factor(END) +
## factor(FURNISH), data = condo)</pre>
```

```
## Residuals:
              1Q Median
##
      Min
                            30
                                    Max
## -70.878 -17.141 -1.102 14.253 66.830
## Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                  183.5701 5.2207 35.162 < 2e-16 ***
                              0.7482 -5.089 8.18e-07 ***
## FLOOR
                   -3.8076
## DIST
                            0.3750 4.644 6.14e-06 ***
                   1.7414
                            3.4565 11.667 < 2e-16 ***
                   40.3251
## factor(VIEW)1
                             3.6023 1.188 0.236267
## factor(END)1
                   4.2792
## factor(FURNISH)1 -32.7162
                             9.5814 -3.415 0.000771 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.39 on 203 degrees of freedom
## Multiple R-squared: 0.4943, Adjusted R-squared: 0.4819
## F-statistic: 39.69 on 5 and 203 DF, p-value: < 2.2e-16
goodmodel=ols_step_both_p(model, pent = 0.1, prem = 0.3, details=TRUE)
## Stepwise Selection Method
## -----
## Candidate Terms:
## 1. FLOOR
## 2. DIST
## 3. factor(VIEW)
## 4. factor(END)
## 5. factor(FURNISH)
##
## We are selecting variables based on p value...
## Stepwise Selection: Step 1
## - factor(VIEW) added
##
                      Model Summary
## -----
## R 0.579 RMSE
## R-Squared 0.335 Coef. Var
## Adj. R-Squared 0.332 MSE
## Pred R-Squared 0.322 MAE
                                             27.701
                                              13.762
                                              767.357
                                              22.134
## ------
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
                            ANOVA
## -----
##
               Sum of
                           DF Mean Square
                                                      Sig.
              Squares
                                              F
## ------
## Regression 80049.921 1
## Residual 158842.854 207
## Total 238892.775 208
                        1
207
                                 80049.921
                                           104.319
                                                      0.0000
                                  767.357
```

					rameter 					
m	odel	Beta	Std.	Error	Std.	Beta	t	Sig	lower	
(Interc	ept)	181.050		2.756			65.684	0.000	175.615	186.4
factor(VI	EW)1	39.163		3.834		0.579	10.214			46.7
Stepwise	Select	ion: Step 2								
- FLOOR a	dded									
			odel S	ummary						
 R				RMS			26.304			
R-Squared		0.4	103	Coe	f. Var		13.068			
Adi. R-Sa	uared	0.3	398	MSE			691.900			
Pred R-Sq	uared			Coe MSE MAE			20.616			
RMSE: Ro MSE: Mea	ot Mea n Squa	n Square Err	ror							
				ANOVA						
		Sum of								
		Squares						Sig.		
Regressio	 n	96361.421 .42531.355 .38892.775		2	48180	 .710	69.635	0.0000		
Residual	1	42531.355		206	691	.900				
Totol	~									
		.38892 . 775		208 						
				 Pa	 rameter	Estima				
m	 odel	Beta	Std.	Pa Error	rameter Std.	Estima Beta	ates t	 Sig		 upṛ
m	 odel 	Beta	Std.	Pa Error	rameter Std.	Estima Beta	ates	Sig		
m m 	 odel 	Beta	Std.	Pa Error	rameter Std.	Estima Beta	ates t 45.012	Sig 		206.9
 m (Interc factor(VI F	 odel ept) EW)1 LOOR	Beta 198.249 39.708 -3.895	Std.	Pa Error 4.404 3.643 0.802	rameter Std.	Estima Beta	ates t 45.012 10.900	Sig 0.000 0.000	189.566	206.9 46.8
m (Interc factor(VI F	odel ept) EW)1 LOOR	Beta 	Std.	Pa Error 4.404 3.643	rameter Std.	Estima Beta 	ates t 45.012 10.900	Sig 0.000 0.000	189.566 32.526	206.9 46.8
(Interc	odel ept) EW)1 LOOR	Beta 198.249 39.708 -3.895	Std.	Pa Error 4.404 3.643 0.802 	rameter Std.	Estima Beta 	45.012 10.900 -4.855	Sig 0.000 0.000	189.566 32.526	206.9 46.8
(Intercfactor(VIF	odel ept) EW)1 LOOR	Beta 198.249 39.708 -3.895	Std.	Pa Error 4.404 3.643 0.802 ummary	rameter Std.	Estima Beta 0.587 0.261	ates t 45.012 10.900	Sig 0.000 0.000	189.566 32.526	206.9 46.8
m (Interc factor(VI F	odel ept) EW)1 LOOR	Beta 198.249 39.708 -3.895	Std. odel S	Pa Error 4.404 3.643 0.802 ummary	rameter Std. E f. Var	Estima Beta 0.587 0.261	45.012 10.900 -4.855	Sig 0.000 0.000	189.566 32.526	206.9 46.8
	odel ept) EW)1 LOOR uared uared	Beta 198.249 39.708 -3.895 -0.6	Stdodel S	Pa Error 4.404 3.643 0.802 ummary RMS Coe MSE MAE	rameter Std. E f. Var	Estima Beta 0.587 0.261	45.012 10.900 -4.855 	Sig 0.000 0.000	189.566 32.526	206.9 46.8
(Intercfactor(VIF	odel cept) EW)1 LOOR constant uared uared uared oot Mea	Beta 198.249 39.708 -3.895 Mc 0.6 0.2 0.3	Std	Pa Error 4.404 3.643 0.802 ummary RMS Coe MSE MAE	rameter Std. E f. Var	Estima Beta 0.587 0.261	45.012 10.900 -4.855 	Sig 0.000 0.000	189.566 32.526	206.9 46.8
(Intercfactor(VIF	odel cept) EW)1 LOOR constant uared uared uared oot Mea	Beta 198.249 39.708 -3.895 Mo 0.6 0.2 0.3 In Square Error lute Error	Std	Pa Error 4.404 3.643 0.802 ummary RMS Coe MSE MAE	rameter Std. E f. Var	Estima Beta 0.587 0.261	45.012 10.900 -4.855 	Sig 0.000 0.000	189.566 32.526	206.9 46.8
(Intercfactor(VIF	odel cept) EW)1 LOOR constant uared uared uared oot Mea	Beta 198.249 39.708 -3.895 Mc 0.6 0.2 0.3 In Square Error	Std	Pa Error 4.404 3.643 0.802 ummary RMS Coe MSE MAE	rameter Std. Std.	Estima Beta 0.587 0.261	45.012 10.900 -4.855 	Sig 0.000 0.000 0.000	189.566 32.526	upp 206.9 46.8 -2.3

Total	142531.355 238892.775	206 208	48180.710 691.900	69.635	0.0000		
			arameter Estim				
model		Std. Error	Std. Beta		Sig	lower	սրլ
			0.587 -0.261	45.012 10.900	0.000 0.000	189.566 32.526	206.9 46.8
FLOOR	-3.895	0.802	-0.261	-4.855	0.000	-5.477	-2.
	ction: Step 3						
- DIST added							
	Мс	odel Summary	,				
R		 580 RM	ISE	25.034			
- 1	0.4	162 Co	ef. Var	12.437			
Adj. R-Square	d 0.4		Ε	626.685			
Pred R-Square	d 0.4	41 MA	λΕ 	19.725			
	Sum of Squares	DF	Mean Square	F	Sig.		
Regression Residual Total	Squares 110422.268 128470.507	3	36807.423 626.685	58.733	0.0000		
Regression Residual Total	Squares 110422.268 128470.507	3 205 208	36807.423	58.733	0.0000		
Regression Residual Total	Squares	3 205 208	36807.423 626.685 	58.733 ates	0.0000	lower	·
Regression Residual Total model	Squares 110422.268 128470.507 238892.775	3 205 208 	36807.423 626.685 Parameter Estim	58.733 ates	0.0000 Sig		
Regression Residual Total model (Intercept)	Squares	3 205 208 	36807.423 626.685 Parameter Estim	58.733 ates t35.672	0.0000 Sig	 173.838	194.
Regression Residual Total model (Intercept)	Squares	3 205 208 	36807.423 626.685 Parameter Estim	58.733 ates t35.672	0.0000 Sig	 173.838	194.
Regression Residual Total model (Intercept) factor(VIEW)1 FLOOR DIST	Squares 110422.268 128470.507 238892.775 Beta 184.008 38.119 -3.664 1.797	3 205 208 	36807.423 626.685 Parameter Estim Std. Beta 0.563 -0.246 0.244	35.672 10.944 -4.789 4.737	0.0000 Sig 0.000 0.000 0.000 0.000	173.838 31.252 -5.172 1.049	194.
Regression Residual Total model (Intercept) factor(VIEW)1 FLOOR DIST	Squares 110422.268 128470.507 238892.775 Beta 184.008 38.119 -3.664 1.797	3 205 208 	36807.423 626.685 Parameter Estim Std. Beta 0.563 -0.246 0.244	58.733 ates t35.672	0.0000 Sig 0.000 0.000 0.000 0.000	 173.838	194.
Regression Residual Total model (Intercept) factor(VIEW)1 FLOOR DIST	Squares 110422.268 128470.507 238892.775 Beta 184.008 38.119 -3.664 1.797	3 205 208 	36807.423 626.685 Parameter Estim Std. Beta 0.563 -0.246 0.244	35.672 10.944 -4.789 4.737	0.0000 Sig 0.000 0.000 0.000 0.000	173.838 31.252 -5.172 1.049	194.: 44.: 2.:
Regression Residual Total model(Intercept) factor(VIEW)1 FLOOR DIST R R R-Squared	Squares 110422.268 128470.507 238892.775 Beta 184.008 38.119 -3.664 1.797 0.6 0.4	3 205 208 	36807.423 626.685 Parameter Estim Std. Beta 0.563 0.246 0.244	58.733 ates t 35.672 10.944 -4.789 4.737	0.0000 Sig 0.000 0.000 0.000 0.000	173.838 31.252 -5.172 1.049	194.
Regression Residual Total model(Intercept) factor(VIEW)1 FLOOR DIST R R-Squared Adj. R-Square	Squares	3 205 208 	36807.423 626.685 Parameter Estim Std. Beta 0.563 0.244 0.244	58.733 ates t 35.672 10.944 -4.789 4.737 25.034 12.437 626.685	0.0000 Sig 0.000 0.000 0.000 0.000	173.838 31.252 -5.172 1.049	194.
Regression Residual Total model (Intercept) factor(VIEW)1 FLOOR DIST R R-Squared Adj. R-Square Pred R-Square	Squares	3 205 208 Std. Error 5.158 3.483 0.765 0.379 odel Summary 580 RM 62 C0 54 MS	36807.423 626.685 Parameter Estim Std. Beta 0.563 0.244 0.244	58.733 ates t 35.672 10.944 -4.789 4.737	0.0000 Sig 0.000 0.000 0.000 0.000	173.838 31.252 -5.172 1.049	194.

## ## ##			ANOVA					
## ##		Sum of Squares	DF	Mean Square	F	Sig.		
## ##	Regression		3 205	36807,423	58.733	0.0000		
## ##				rameter Estima	 etec			
##		Beta 						upper
##	(Intercept) factor(VIEW)1	184.008 38.119	5.158 3.483	0.563	35.672 10.944	0.000 0.000	173.838 31.252	44.987
##		-3.664						
##	DIST		0.379	0.244	4.737	0.000	1.049	2.545
	Stepwise Sele							
##	factor/FUDN	TCII) addad						
##	- factor(FURN	ish) added						
##			del Summary					
	D			·				
##	к R-Squared	0.7	01 RMS	ot of Van	24.419 12.131			
	Adj. R-Square	d 0.4	91 Coe 81 MSE 69 MAE	. vai	596.281			
##	Pred R-Square	d 0.4	69 MAE		19.058			
	PMSE: Poot M	ean Square Err	 on					
##			OI*					
##	MAE: Mean Ab							
##			ANOVA					
##								
##		Sum of			_	٠.		
##		Squares		Mean Square		Sig.		
	Regression		4	29312.883	49.16	0.0000		
	Residual	121641.245	204	596.281				
	Total	238892.775	208					
##								
ш			Parameter Es	timates				
##				Std. Beta	t	Sig	lower	upper
## ##	(Intercep	t) 185.193)1 40.335 OR -3.736	5.044	0.596	36.717	0.000	175.248 33.513	195.138
##	FLO	OR -3.736	0.747	-0.251	-5.004	0.000	-5.208	-2.264
##	DI	ST 1.679	0.372	0.228				
	factor(FURNISH)1 -32.450	9.588	-0.173			-51.355	
##								
##								
##		Мо	del Summary					
##								

: R							
			MSE	24.419			
R-Squared	0	.491 C	oef. Var	12.131			
: Adj. R-Square	ed 0	.481 M	SE	596.281			
Pred R-Square	ed 0	.469 M	AE	19.058			
·					-		
MSE: Mean So MAE: Mean Ab	Mean Square E quare Error osolute Error						
		ANOV	Α				
:							
;	Sum of						
:	Squares	DF	Mean Square	F	Sig.		
:							
Regression	117251 530	4	29312.883 596.281	49 16	0 0000		
Posidual	1216/1 2/6	204	E06 201	43.10	0.0000		
Kesiduai	238892.775	204	390.201				
Total							
		P	arameter Estima	tes			
			or Std. Beta	t	Sig	lower	uppe
(Interce	ot) 185.193	5.0	44 50 0.596 47 -0.251 72 0.228	36.717	0.000	175.248	195.13
factor(VIE	W)1 40.335	3.4	60 0.596	11.658	0.000	33.513	47.15
· FL(OOR -3.736	0.7	47 -0.251	-5.004	0.000	-5.208	-2.26
D.	IST 1.679	0.3	72 0.228	4.518	0.000	0.946	2.41
factor(FURNTS)	1)1 -32.456	9.5	88 -0.173	-3.384	0.001	-51.355	-13.54
	-, - 52, 750						
No more varia	ables to be a	dded/removed	•				
Final Model O							
: : Final Model (:	•	Model Summar					
Final Model (· 						
Final Model C		.701 R	 MSE	 24 . 419			
Final Model C		.701 R	 MSE				
Final Model C R R-Squared Adi. R-Square		.701 R .491 C	MSE oef. Var	24.419 12.131	-		
Final Model C R R-Squared Adi. R-Square		.701 R .491 C	MSE oef. Var	24.419 12.131 596.281			
Final Model C R R-Squared Adj. R-Square Pred R-Square	0 0 ed 0 ed 0	.701 R .491 C .481 M .469 M	MSE oef. Var	24.419 12.131			
R-Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Sc MAE: Mean Ab	0 0 ed 0 ed 0 	.701 R .491 C .481 M .469 M	MSE oef. Var	24.419 12.131 596.281			
R.Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ad	ed 0 Mean Square Equare Error	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281			
Final Model C R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Sc MAE: Mean Ab	ed 0 ed 0 ed 0 Mean Square E quare Error osolute Error	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281			
Final Model C R R-Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Sc MAE: Mean Ab	0 ed 0 ed 0 Mean Square E quare Error osolute Error	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058			
Final Model C R R-Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Sc MAE: Mean Ab	ed 0 ed 0 Mean Square E quare Error psolute Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058	Sig.		
Final Model C R R-Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Sc MAE: Mean Ab	Mean Square Equare Error ssolute Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058	Sig.		
R.Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Ab MAE: Mean Ab	Mean Square Equare Error Solute Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root N MSE: Mean Ab Regression Residual	Mean Square Equare Error Solute Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square Equare Error Solute Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square Equare Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square Equare Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square Equare Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square Equare Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058	Sig.		
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square Equare Error Sum of Squares	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058 	Sig. 0.0000	lower	ирре
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ad MAE: Mean Ad Regression Residual Total	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE Mean Square 29312.883 596.281	24.419 12.131 596.281 19.058 	Sig.	lower	uppe
R R-Squared Adj. R-Square Pred R-Square RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total	ded 0 ded 0 ded 0 ded 0 dean Square E quare Error ssolute Error Sum of Squares 117251.530 121641.245 238892.775	.701 R491 C481 M469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058 	Sig. 0.0000		uppe:
R R-Squared Adj. R-Squared R-Squared R-Squared R-Squared R-Squared R-Squared R-Squared RMSE: Rean Additional Residual Residual	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	.701 R .491 C .481 M .469 M	MSE oef. Var SE AE	24.419 12.131 596.281 19.058 	Sig. 0.0000	lower 175.248 33.513	uppe 195.13

```
##
           FLOOR
                  -3.736
                              0.747
                                        -0.251
                                               -5.004
                                                       0.000
                                                               -5.208
                                                                        -2.264
##
           DIST
                   1.679
                              0.372
                                        0.228
                                               4.518
                                                       0.000
                                                               0.946
                                                                        2.412
## factor(FURNISH)1
                  -32.450
                              9.588
                                        -0.173
                                               -3.384
                                                       0.001
                                                              -51.355
                                                                       -13.544
summary(goodmodel$model)
##
## Call:
## lm(formula = paste(response, "~", paste(preds, collapse = " + ")),
       data = 1)
##
## Residuals:
##
       Min
                1Q Median
                                 3Q
                                        Max
## -68.527 -17.773 -0.287
                            14.184 66.490
##
## Coefficients:
##
                     Estimate Std. Error t value Pr(>|t|)
                                  5.0438
                                                  < 2e-16 ***
## (Intercept)
                     185.1932
                                          36.717
                                          11.658 < 2e-16 ***
## factor(VIEW)1
                     40.3347
                                  3.4599
                                  0.7465 -5.004 1.21e-06 ***
## FLOOR
                      -3.7359
## DIST
                                  0.3717
                                          4.518 1.06e-05 ***
                       1.6793
## factor(FURNISH)1 -32.4497
                                  9.5885 -3.384 0.000856 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.42 on 204 degrees of freedom
## Multiple R-squared: 0.4908, Adjusted R-squared: 0.4808
## F-statistic: 49.16 on 4 and 204 DF, p-value: < 2.2e-16
```

R output 4 for Question 4

```
promotion=read.csv("C:\\Users\\thuntida.ngamkham\\Desktop\\dataset603\\promote.csv"
,header=TRUE)
dummymodel<-lm(promote~factor(rank),data=promotion)
summary(dummymodel)
##
## Call:
## lm(formula = promote ~ factor(rank), data = promotion)
##</pre>
```

```
## Residuals:
##
       Min
                1Q
                     Median
                                 3Q
                                        Max
## -20.875
           -6.714
                      0.000
                              7.2573 29.385
##
## Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
##
                   42.000
                               2.259 18.593 < 2e-16 ***
## (Intercept)
                                       2.640 0.013613 *
## factor(rank)2
                   10.571
                               4.005
                                       3.884 0.000602 ***
## factor(rank)3
                   14.875
                               3.830
## factor(rank)4
                   25.102
                               1.275
                                       6.471 < 2e-16 ***
## ---
                  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Signif. codes:
## Residual standard error: 4.795 on 36 degrees of freedom
## Multiple R-squared: 0.889 , Adjusted R-squared:
## F-statistic: 14.515 on 3 and 36 DF, p-value: 0.001319
```

R output 5 for Question 6

```
ski=read.csv("C:\\Users\\thuntida.ngamkham\\OneDrive - University of Calgary\\datas
et603\\ski.CSV", header = TRUE)
#ski=read.csv("~/OneDrive - University of Calgary/MyCoursesThierry/DATA603/data/dat
aset603/ski.CSV", header = TRUE)
model<-lm(skiers ~factor(weekend)+snow+temperature ,data=ski )</pre>
summary(model)
##
## lm(formula = skiers ~ factor(weekend) + snow + temperature, data = ski)
##
## Residuals:
##
        Min
                  1Q
                       Median
                                    3Q
                                             Max
## -152.431 -38.371
                        2,996
                                46.566 117.806
##
## Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)
                     561.037
                                  72.403
                                           7.749 1.93e-08 ***
                                           3.157 0.00379 **
## factor(weekend)1
                      92.582
                                  29.321
## snow
                       4.303
                                  1.823
                                           2.360 0.02546 *
                                  2.144 -3.191 0.00348 **
## temperature
                      -6.843
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 64.06 on 28 degrees of freedom
## Multiple R-squared: 0.8234, Adjusted R-squared: 0.8045
## F-statistic: 43.52 on 3 and 28 DF, p-value: 1.136e-10
intmodel<-lm(skiers ~factor(weekend)+snow+temperature+factor(weekend)*temperature,d
ata=ski )
summary(intmodel)
## Call:
## lm(formula = skiers ~ factor(weekend) + snow + temperature +
##
      factor(weekend) * temperature, data = ski)
## Residuals:
      Min
               10 Median
                               3Q
                                      Max
## -153.60 -39.89
                    11.32
                            36.98 111.18
##
## Coefficients:
##
                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)
                                552.949
                                            65.946 8.385 5.38e-09 ***
## factor(weekend)1
                                208.279
                                            51.697 4.029 0.00041 ***
## snow
                                  3.913
                                            1.665 2.350 0.02635 *
                                             1.975 -3.056 0.00500 **
## temperature
                                 -6.037
## factor(weekend)1:temperature -13.021
                                             4.984 -2.613 0.01450 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '* 0.05 '.' 0.1 ' ' 1
## Residual standard error: 58.28 on 27 degrees of freedom
## Multiple R-squared: 0.859, Adjusted R-squared: 0.8382
## F-statistic: 41.14 on 4 and 27 DF, p-value: 4.1e-11
```

R output 6 for Question 7

```
wateroil=read.csv("C:\\Users\\thuntida.ngamkham\\Desktop\\dataset603\\wateroil.csv"
,header=TRUE)
model1<-lm(Voltage~Volume+Salinity+Temperature+Delay+Surfactant+SpanTriton+SolidPar
t,data=wateroil)
summary(model1)
##
## Call:
## Im(formula = Voltage ~ Volume + Salinity + Temperature + Delay +
## Surfactant + SpanTriton + SolidPart, data = wateroil)</pre>
```

```
##
## Residuals:
##
       Min
                 1Q
                      Median
                                   3Q
                                           Max
## -0.68444 -0.23788 0.03217 0.13755 0.74783
##
## Coefficients:
               Estimate Std. Error t value Pr(>|t|)
                                   4.032 0.001975 **
## (Intercept) 0.998082
                          0.247542
                          0.005039 -4.451 0.000977 ***
## Volume
              -0.022429
## Salinity
               0.155711
                          0.074291
                                   2.096 0.060018 .
## Temperature -0.017187
                          0.011860 -1.449 0.175188
                          0.009619 -0.990 0.343279
## Delay
             -0.009527
## Surfactant 0.421421
                          0.100782 4.182 0.001533 **
## SpanTriton 0.417123
                          0.437702 0.953 0.361070
## SolidPart -0.155244
                          0.148582 -1.045 0.318516
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4635 on 11 degrees of freedom
## Multiple R-squared: 0.771, Adjusted R-squared: 0.6253
## F-statistic: 5.292 on 7 and 11 DF, p-value: 0.007428
model2<-lm(Voltage~Volume+Salinity+Surfactant,data=wateroil)</pre>
summary(model2)
##
## Call:
## lm(formula = Voltage ~ Volume + Salinity + Surfactant, data = wateroil)
## Residuals:
               10 Median
                               30
##
      Min
                                      Max
## -0.7529 -0.3051 0.1074 0.1803 1.1318
##
## Coefficients:
              Estimate Std. Error t value Pr(>|t|)
##
## (Intercept) 0.93257 0.24819
                                    3.757 0.001902 **
                          0.00490 -4.953 0.000173 ***
## Volume
             -0.02427
## Salinity
               0.14206
                          0.07573
                                    1.876 0.010283 *
## Surfactant
               0.38457
                          0.09801 3.924 0.001354 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.4796 on 15 degrees of freedom
## Multiple R-squared: 0.6656, Adjusted R-squared:
## F-statistic: 9.954 on 3 and 15 DF, p-value: 0.000735
model3<-lm(Voltage~Volume+Surfactant,data=wateroil)</pre>
summary(model3)
##
## Call:
```

```
## lm(formula = Voltage ~ Volume + Surfactant, data = wateroil)
##
## Residuals:
                      Median
##
       Min
                 1Q
                                   30
                                           Max
## -0.86384 -0.25726 0.01945 0.25781 1.35945
##
## Coefficients:
##
               Estimate Std. Error t value Pr(>|t|)
                                    3.838 0.001451 **
## (Intercept) 1.01041
                          0.26326
## Volume
               -0.02208
                          0.00512 -4.313 0.000536 ***
## Surfactant
               0.42836
                          0.10240 4.183 0.000703 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.516 on 16 degrees of freedom
## Multiple R-squared: 0.5872, Adjusted R-squared:
## F-statistic: 11.38 on 2 and 16 DF, p-value: 0.000843
model4<-lm(Voltage~Volume+Salinity+Surfactant+Volume*Salinity+Volume*Surfactant+Sal
inity*Surfactant,data=wateroil)
summary(model4)
##
## Call:
## lm(formula = Voltage ~ Volume + Salinity + Surfactant + Volume *
       Salinity + Volume * Surfactant + Salinity * Surfactant, data = wateroil)
##
##
## Residuals:
##
       Min
                 1Q
                      Median
                                   30
                                           Max
## -0.76973 -0.29242 0.09955 0.24755 0.81376
##
## Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                       0.940454
                                  0.271180
                                             3.468
                                                    0.00465 **
## Volume
                                  0.011149 -0.922
                                                    0.00149 **
                       -0.010274
## Salinity
                       0.030410
                                  0.283332
                                             0.107
                                                    0.00030 ***
## Surfactant
                       0.181183
                                  0.222972
                                             0.813
                                                    0.03228 *
## Volume:Salinity
                                  0.003654 -1.065
                                                    0.00805 **
                       -0.003889
## Volume:Surfactant
                       -0.001351
                                  0.004171 -0.324
                                                    0.00162
## Salinity:Surfactant 0.115549
                                  0.073071
                                             1.581
                                                    0.03979
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4778 on 12 degrees of freedom
## Multiple R-squared: 0.7345, Adjusted R-squared: 0.7018
## F-statistic: 5.534 on 6 and 12 DF, p-value: 0.005857
model5<-lm(Voltage~Volume+Surfactant+Volume*Surfactant,data=wateroil)</pre>
summary(model5)
```

```
##
## Call:
## lm(formula = Voltage ~ Volume + Surfactant + Volume * Surfactant,
      data = wateroil)
##
## Residuals:
       Min
                      Median
                 10
                                   30
                                          Max
## -0.91709 -0.17411 0.01823 0.20823
                                      1.31354
##
## Coefficients:
##
                     Estimate Std. Error t value Pr(>|t|)
                     0.944304
                              0.297053
                                          3.179 0.00623 **
## (Intercept)
## Volume
                    -0.018915
                                0.007964
                                         -2.375
                                                 0.03131 *
## Surfactant
                                                 0.00751 **
                     0.491709
                                0.159277
                                          3.087
## Volume:Surfactant -0.001676
                                0.003173
                                         -0.528
                                                 0.00511 **
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5281 on 15 degrees of freedom
## Multiple R-squared: 0.5947, Adjusted R-squared: 0.7137
## F-statistic: 7.338 on 3 and 15 DF, p-value: 0.002968
model6<-lm(Voltage~Volume+I(Volume^2)+Salinity+Surfactant+Volume*Salinity+Volume*Su
rfactant+Salinity*Surfactant,data=wateroil)
summary(model6)
##
## Call:
## lm(formula = Voltage ~ Volume + I(Volume^2) + Salinity + Surfactant +
      Volume * Salinity + Volume * Surfactant + Salinity * Surfactant,
##
      data = wateroil)
##
## Residuals:
       Min
                 10
                      Median
##
                                   30
                                          Max
## -0.48500 -0.14500 0.01333 0.18000 0.58500
##
## Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
##
## (Intercept)
                       1.0666667 0.2002322
                                             5.327 0.000242 ***
                      -0.1052292   0.0288264   -3.650   0.003818 **
## Volume
## I(Volume^2)
                       0.0011693 0.0003407
                                             3.432 0.005603 **
## Salinity
                       0.5300000 0.2519533
                                             2.104 0.059234
## Surfactant
                       1.0883333 0.3099290 3.512 0.004870 **
## Volume:Salinity
                      ## Volume:Surfactant
                      -0.0120000 0.0043352 -2.768 0.018290 *
## Salinity:Surfactant 0.0366667 0.0578021 0.634 0.538822
## Signif. codes:
                  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.3168 on 11 degrees of freedom
```

```
## Multiple R-squared: 0.8718, Adjusted R-squared: 0.8202
## F-statistic: 10.69 on 7 and 11 DF, p-value: 0.0003923
model7<-lm(Voltage~Volume+I(Volume^2)+Surfactant+Volume*Surfactant,data=wateroil)</pre>
summary(model7)
##
## Call:
## lm(formula = Voltage ~ Volume + I(Volume^2) + Surfactant + Volume *
      Surfactant, data = wateroil)
##
##
## Residuals:
##
     Min
             10 Median
                          3Q
                                Max
##
   -1.03
         -0.07 -0.01
                        0.09
                               1.13
##
## Coefficients:
##
                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)
                     1.0666667 0.2801134
                                          3.808 0.00192 **
## Volume
                    -0.0751250 0.0298648 -2.515 0.02471 *
## I(Volume^2)
                     0.0007552 0.0003890 1.941 0.04262 .
## Surfactant
                                          3.076 0.00821 **
                     1.1800000 0.3835610
## Volume:Surfactant -0.0120000 0.0060646 -1.979 0.08787 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4852 on 14 degrees of freedom
## Multiple R-squared: 0.6807, Adjusted R-squared: 0.8100
## F-statistic: 7.462 on 4 and 14 DF, p-value: 0.001951
model8<-lm(Voltage~Volume+I(Volume^2)+Salinity+Surfactant+Volume*Salinity+Volume*Su</pre>
rfactant, data=wateroil)
summary(model8)
## Call:
## lm(formula = Voltage ~ Volume + I(Volume^2) + Salinity + Surfactant +
      Volume * Salinity + Volume * Surfactant, data = wateroil)
##
##
## Residuals:
                 10
                     Median
                                  30
                                          Max
## -0.54000 -0.09000 0.01333 0.12500
                                      0.64000
##
## Coefficients:
##
                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)
                     1.0666667 0.1951827
                                          5.465 0.000144 ***
## Volume
                    ## I(Volume^2)
                     0.0012552 0.0003047
                                         4.119 0.001423 **
                     0.6400000 0.1781766
## Salinity
                                          3.592 0.103700
                                          4.415 0.000843 ***
## Surfactant
                    1.1800000 0.2672650
## Volume:Salinity
                    ## Volume:Surfactant -0.0120000 0.0042258 -2.840 0.014906 *
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 0.01 '* 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3381 on 12 degrees of freedom
## Multiple R-squared: 0.8671, Adjusted R-squared: 0.8107
## F-statistic: 13.05 on 6 and 12 DF, p-value: 0.0001211
model9<-lm(Voltage~Volume+I(Volume^2)+Surfactant,data=wateroil)</pre>
summary(model9)
## Call:
## lm(formula = Voltage ~ Volume + I(Volume^2) + Surfactant, data = wateroil)
##
## Residuals:
       Min
                 1Q
                      Median
                                   3Q
                                           Max
## -0.79000 -0.27000 -0.02667 0.27000
                                       1.37000
## Coefficients:
                Estimate Std. Error t value Pr(>|t|)
##
## (Intercept) 1.067e+00 3.061e-01
                                      3.484 0.00333 **
## Volume
              -3.012e-02 2.115e-02 -1.424 0.01749 *
## I(Volume^2) 8.021e-05 2.043e-04
                                      0.393 0.0019 **
                                      3.470 0.00343 **
## Surfactant 4.600e-01 1.326e-01
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.5302 on 15 degrees of freedom
## Multiple R-squared: 0.5914, Adjusted R-squared: 0.8097
## F-statistic: 7.237 on 3 and 15 DF, p-value: 0.003149
```