1.

Scatterplot Matrix
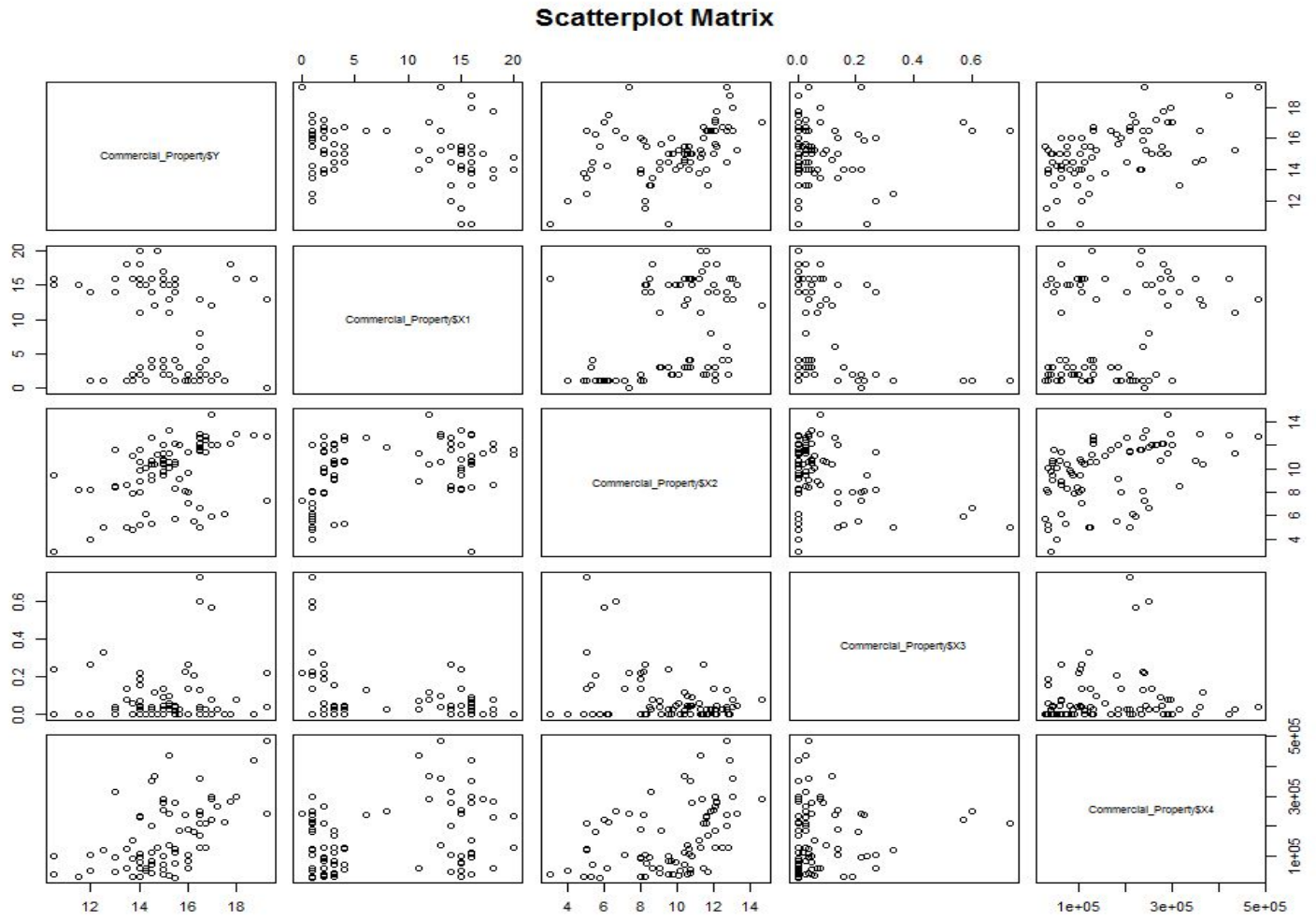
Y and X1 are clearly non linear.
Y and X2 are linear.
Y and X3 are clearly non linear.
Y and X4 are linear.

2.Yhat= 1.220 - 1.420X1+2.820X2+6.193X3+7.924X4

```
> fit<-lm(Commercial_Property$Y~Commercial_Property$X1+Commercial_Property$X2+Comm
ercial_Property$X3+Commercial_Property$X4,data=Commercial_Property)
> summary(fit)

Call:
lm(formula = Commercial_Property$Y ~ Commercial_Property$X1 +
    Commercial_Property$X2 + Commercial_Property$X3 + Commercial_Property$X4,
    data = Commercial_Property)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1872 -0.5911 -0.0910  0.5579  2.9441

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)             1.220e+01  5.780e-01  21.110  < 2e-16 ***
Commercial_Property$X1 -1.420e-01  2.134e-02  -6.655 3.89e-09 ***
Commercial_Property$X2  2.820e-01  6.317e-02   4.464 2.75e-05 ***
Commercial_Property$X3  6.193e-01  1.087e+00   0.570    0.57
Commercial_Property$X4  7.924e-06  1.385e-06   5.722 1.98e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.137 on 76 degrees of freedom
Multiple R-squared:  0.5847,     Adjusted R-squared:  0.5629
F-statistic: 26.76 on 4 and 76 DF,  p-value: 7.272e-14
```
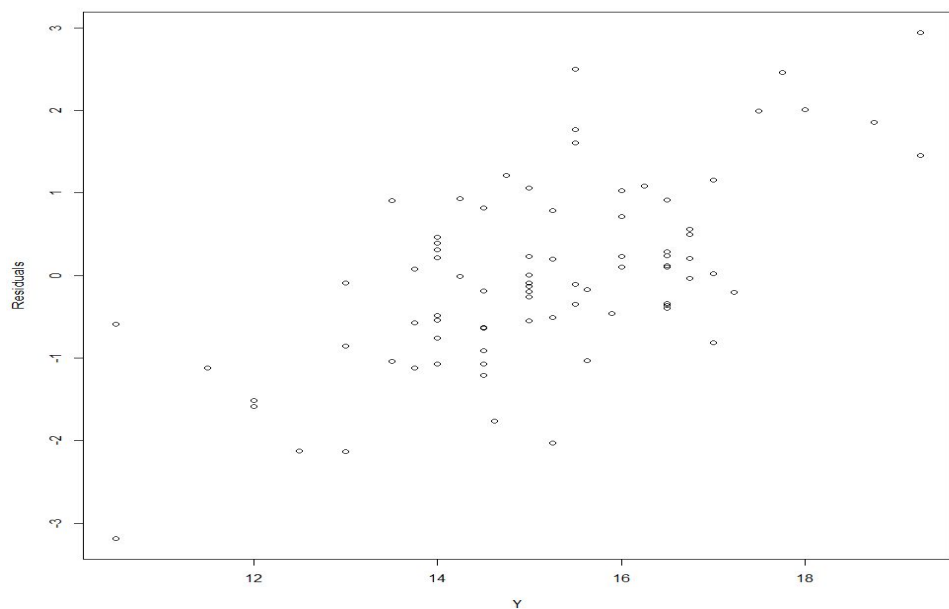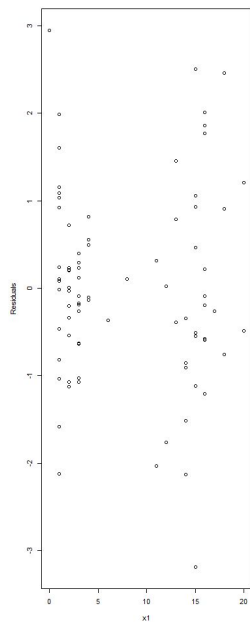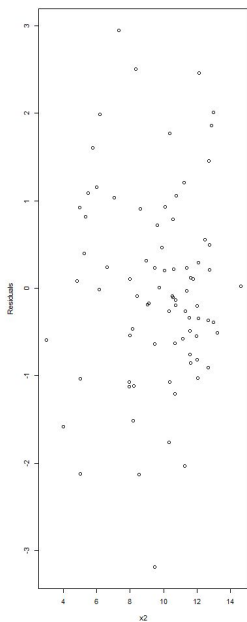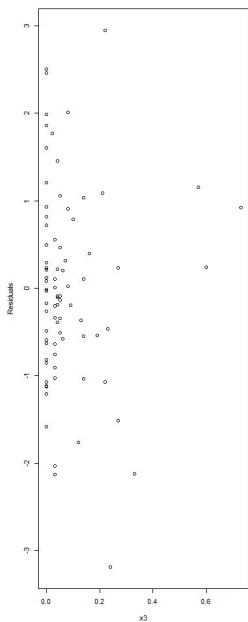
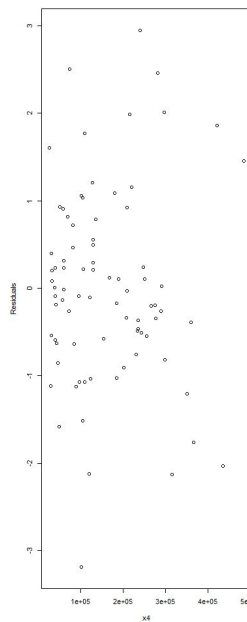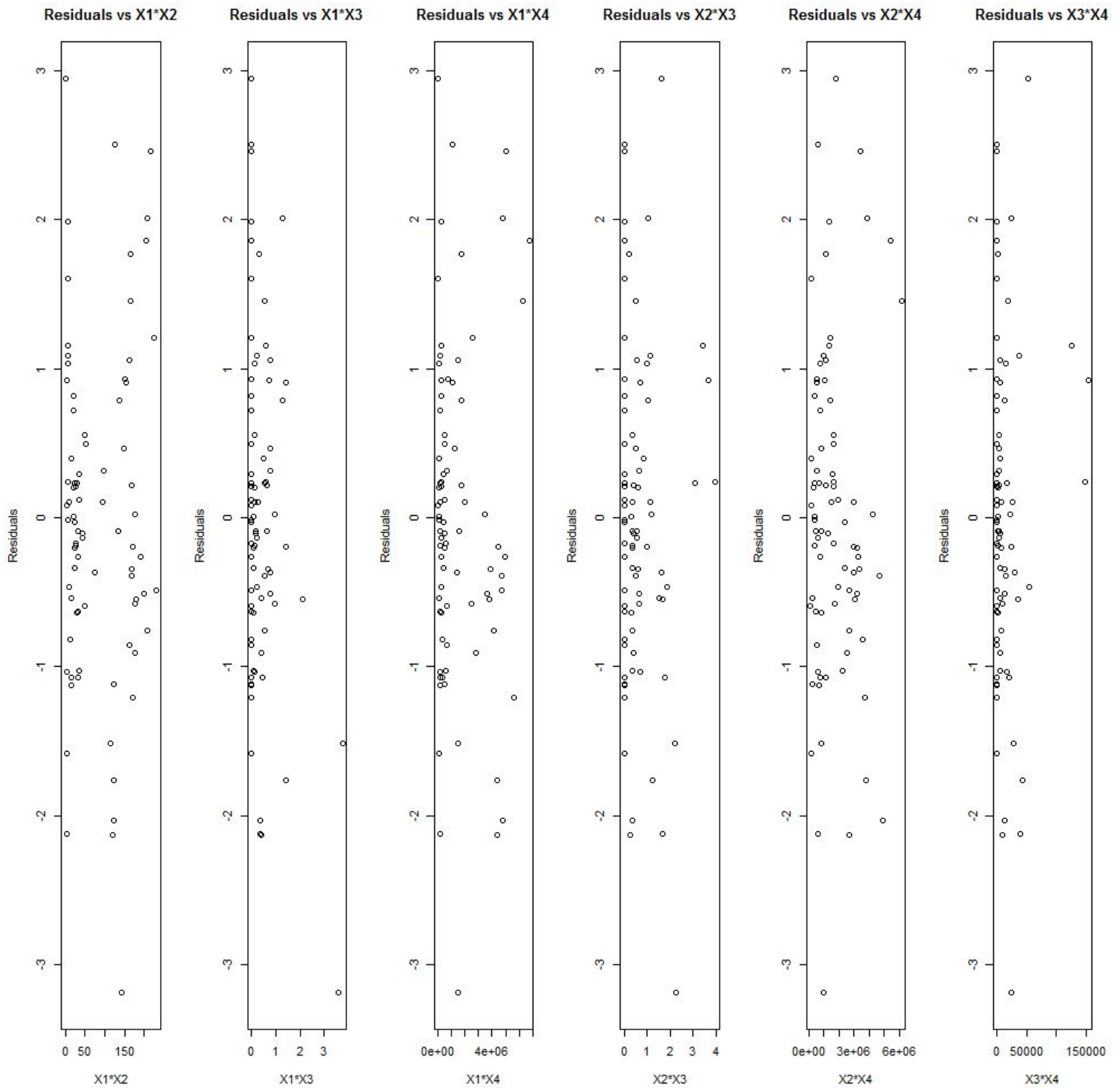**Residuals vs Y**

3.

The residuals look like i.i.d. normally distributed, i.e., the pattern is unsystematic random around zero

4. Yes, it is significant

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

Response: Commercial_Property$Y
                      Df Sum Sq Mean Sq F value    Pr(>F)
Commercial_Property$X1  1 14.819  14.819 11.4649  0.001125 **
Commercial_Property$X2  1 72.802  72.802 56.3262 9.699e-11 ***
Commercial_Property$X3  1  8.381   8.381  6.4846  0.012904 *
Commercial_Property$X4  1 42.325  42.325 32.7464 1.976e-07 ***
Residuals              76 98.231   1.293
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5. Rsq: 58.47%
Adj Rsq:56.29%

```
> pred.commerc                        > pred.coomerp
        fit      lwr      upr                  fit      lwr      upr
1  14.53567 13.97132 15.10003      1  14.53567 12.202099 16.86925
2  13.51381 12.89591 14.13170      2  13.51381 11.166710 15.86090
3  11.09105 10.09034 12.09176      3  11.09105  8.615475 13.56663
```
6.

7. Yhat=1.237 - 1.442X1 + 2.672X2 + 8.178X4

```
Call:
lm(formula = Commercial_Property$Y ~ Commercial_Property$X1 +
    Commercial_Property$X2 + Commercial_Property$X4, data = Commercial_Proper
ty)

Residuals:
   Min     1Q  Median     3Q    Max
-3.0620 -0.6437 -0.1013  0.5672  2.9583

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              1.237e+01  4.928e-01  25.100  < 2e-16 ***
Commercial_Property$X1  -1.442e-01  2.092e-02  -6.891 1.33e-09 ***
Commercial_Property$X2   2.672e-01  5.729e-02   4.663 1.29e-05 ***
Commercial_Property$X4   8.178e-06  1.305e-06   6.265 1.97e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.132 on 77 degrees of freedom
Multiple R-squared: 0.583,    Adjusted R-squared: 0.5667
F-statistic: 35.88 on 3 and 77 DF, p-value: 1.295e-14
```

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

Response: Commercial_Property$Y
                      Df Sum Sq Mean Sq F value    Pr(>F)
Commercial_Property$X1  1 14.819  14.819 11.4649  0.001125 **
Commercial_Property$X2  1 72.802  72.802 56.3262 9.699e-11 ***
Commercial_Property$X3  1  8.381   8.381  6.4846  0.012904 *
Commercial_Property$X4  1 42.325  42.325 32.7464 1.976e-07 ***
Residuals              76 98.231   1.293
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> anova(fit1)
Analysis of Variance Table

Response: Commercial_Property$Y
                      Df Sum Sq Mean Sq F value    Pr(>F)
Commercial_Property$X1  1 14.819  14.819 11.566  0.001067 **
Commercial_Property$X2  1 72.802  72.802 56.825 7.841e-11 ***
Commercial_Property$X4  1 50.287  50.287 39.251 1.973e-08 ***
Residuals              77 98.650   1.281
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Removing X3 makes the model better

8.Yes there is a curvature

X1 vs Y
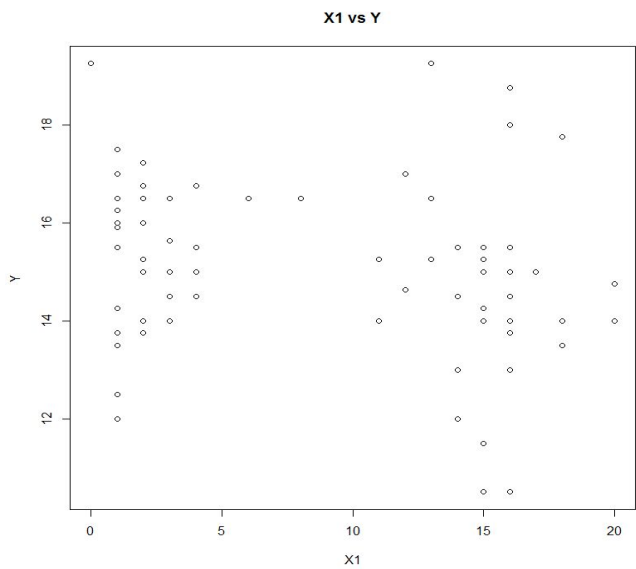
9.Yhat= 1.249 - 4.043X1 + 3.14X2 + 8.046X4 + 1.415X1^2

```
lm(formula = Commercial_Property$Y ~ Commercial_Property$X1 +
    Commercial_Property$X2 + Commercial_Property$X4 + I(Commercial_Property$X
1^2),
    data = Commercial_Property)

Residuals:
    Min      1Q   Median      3Q     Max
-2.89596 -0.62547 -0.08907  0.62793  2.68309

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                 1.249e+01  4.805e-01  26.000  < 2e-16 ***
Commercial_Property$X1     -4.043e-01  1.089e-01  -3.712  0.00039 ***
Commercial_Property$X2      3.140e-01  5.880e-02   5.340 9.33e-07 ***
Commercial_Property$X4      8.046e-06  1.267e-06   6.351 1.42e-08 ***
I(Commercial_Property$X1^2) 1.415e-02  5.821e-03   2.431  0.01743 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.097 on 76 degrees of freedom
Multiple R-squared:  0.6131,    Adjusted R-squared:  0.5927
F-statistic:  30.1 on 4 and 76 DF,  p-value: 5.203e-15
```
.
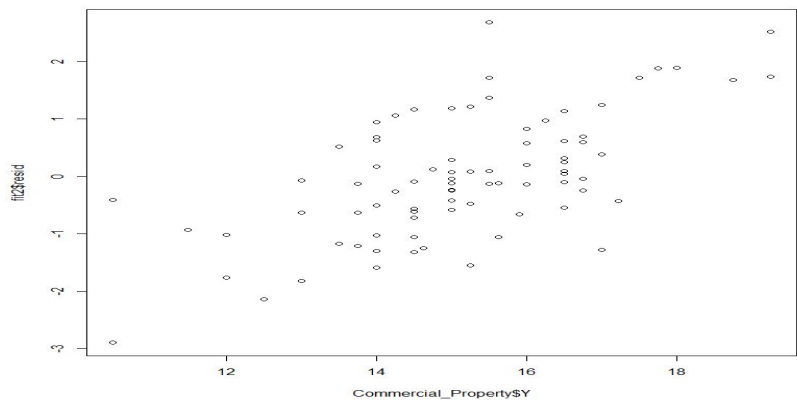
Yes, as the new addition of quadratic term makes the model better.

10.Clearly X1^2 is significant

```
> anova(fit1)
Analysis of Variance Table

Response: Commercial_Property$Y
                     Df Sum Sq Mean Sq F value    Pr(>F)
Commercial_Property$X1  1 14.819  14.819  11.566  0.001067 **
Commercial_Property$X2  1 72.802  72.802  56.825 7.841e-11 ***
Commercial_Property$X4  1 50.287  50.287  39.251 1.973e-08 ***
Residuals              77 98.650   1.281
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> anova(fit2)
Analysis of Variance Table

Response: Commercial_Property$Y
                           Df Sum Sq Mean Sq F value    Pr(>F)
Commercial_Property$X1      1 14.819  14.819 12.3036 0.0007627 ***
Commercial_Property$X2      1 72.802  72.802 60.4463 2.968e-11 ***
Commercial_Property$X4      1 50.287  50.287 41.7522 8.907e-09 ***
I(Commercial_Property$X1^2) 1  7.115   7.115  5.9078 0.0174321 *
Residuals                  76 91.535   1.204
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```