Qualification

Nick Lauerman

Monday, November 24, 2014

Quallification of R

Explorations Using NIST datasets

32-bit R and 32-bit Windows System 7

Linear Regression using lm()

```
options(digits=15)
path="~/R/workspace/qual/raw data/Linear Regression/"
```

Using Norris.dat file. The file was cleaned up removing the header informatin and saving as Norris.txt.

The data file constist of 2 variables and 36 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-0.262323073774029	0.232818234301152
X	1.00211681802045	0.429796848199937E-03

Residual	.
Standard Deviation	0.884796396144373
R-Squared	0.999993745883712

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	4255954.13232369	4255954.13232369	5436385.54079785
Residual	34	26.6173985294224	0.782864662630069	

```
Norris <- read.table(file=paste0(path,"Norris.txt"), header=TRUE)
Norris.lm <- lm(y~x, data=Norris)
summary(Norris.lm)
```

```
##
## Call:
## lm(formula = y ~ x, data = Norris)
##
## Residuals:
##              Min               1Q           Median               3Q
## -2.3523781286600 -0.5326967162014 -0.0296292259639  0.6000277736811
##              Max
##  1.7897858528800
##
## Coefficients:
##                      Estimate       Std. Error    t value Pr(>|t|)
## (Intercept) -0.2623230737740  0.2328182343012   -1.12673  0.26775
## x            1.0021168180205  0.0004297968482 2331.60579  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.884796396144 on 34 degrees of freedom
## Multiple R-squared:  0.999993745884, Adjusted R-squared:  0.999993561939
## F-statistic:  5436385.5408 on 1 and 34 DF,  p-value: < 2.220446049e-16
```

```
anova(Norris.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df         Sum Sq        Mean Sq      F value     Pr(>F)
## x          1 4255954.132324 4255954.132324 5436385.5408 < 2.22e-16 ***
## Residuals 34      26.617399       0.782865
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Pontius.dat file. The file was cleaned up removing the header informatin and saving as Norris.txt.

The data file constist of 2 variables and 40 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression with a quadric term

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	0.673565789473684E-03	0.107938612033077E-03
X	0.732059160401003E-06	0.157817399981659E-09
x^2	-0.316081871345029E-14	0.486652849992036E-16

Residual	.
Standard Deviation	0.205177424076185E-03
R-Squared	0.999999900178537

Residual	.

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	2	15.6040343244198	7.80201716220991	185330865.995752
Residual	37	0.155761768796992E-05	0.420977753505385E-07	

```
Pontius <- read.table(file=paste0(path,"Pontius.txt"), header=TRUE)
Pontius.lm <- lm(y~x + I(x^2), data=Pontius)
summary(Pontius.lm)
```

```
##
## Call:
## lm(formula = y ~ x + I(x^2), data = Pontius)
##
## Residuals:
##                Min                 1Q             Median
## -4.46840225564e-04 -1.57827067669e-04  3.81729323309e-05
##                 3Q                Max
##  1.08788533835e-04  4.23453007519e-04
##
## Coefficients:
##                       Estimate         Std. Error    t value   Pr(>|t|)
## (Intercept)  6.73565789473e-04  1.07938612033e-04    6.24027 2.9705e-07
## x            7.32059160401e-07  1.57817399982e-10 4638.64669 < 2.22e-16
## I(x^2)      -3.16081871345e-15  4.86652849992e-17  -64.95017 < 2.22e-16
##
## (Intercept) ***
## x           ***
## I(x^2)      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000205177424076 on 37 degrees of freedom
## Multiple R-squared:  0.999999900179, Adjusted R-squared:  0.999999894783
## F-statistic: 185330865.996 on 2 and 37 DF,  p-value: < 2.220446049e-16
```

```
anova(Pontius.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df          Sum Sq         Mean Sq         F value     Pr(>F)
## x          1 15.603856733899 15.603856733899 370657513.46648 < 2.22e-16
## I(x^2)     1  0.000177590520  0.000177590520      4218.52506 < 2.22e-16
## Residuals 37  0.000001557618  0.000000042098
##
```

```
## x         ***
## I(x^2)    ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using NoInt1.dat file. The file was cleaned up removing the header informatin and saving as NoInt1.txt.

The data file constist of 2 variables and 11 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, with no intercept Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept		
X	2.07438016528926	0.165289256198347E-01

Residual	.
Standard Deviation	3.56753034006338
R-Squared	0.999365492298663

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	200457.727272727	200457.727272727	15750.2500000000
Residual	10	127.272727272727	12.7272727272727	

```
NoInt1 <- read.table(file=paste0(path,"NoInt1.txt"), header=TRUE)
NoInt1.lm <- lm(y~x + 0, data=NoInt1)
summary(NoInt1.lm)
```

```
##
## Call:
## lm(formula = y ~ x + 0, data = NoInt1)
##
## Residuals:
##             Min              1Q          Median              3Q
## -5.206611570248 -2.520661157025  0.165289256198  2.851239669421
##             Max
##  5.537190082645
##
## Coefficients:
##          Estimate      Std. Error t value   Pr(>|t|)
## x 2.0743801652893 0.0165289256198   125.5 < 2.22e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.56753034006 on 10 degrees of freedom
## Multiple R-squared:  0.999365492299, Adjusted R-squared:  0.999302041529
## F-statistic:      15750.25 on 1 and 10 DF,  p-value: < 2.220446049e-16
```

```
anova(NoInt1.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df          Sum Sq         Mean Sq  F value     Pr(>F)
## x          1 200457.72727273 200457.72727273 15750.25 < 2.22e-16 ***
## Residuals 10    127.27272727     12.72727273
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using NoInt2.dat file. The file was cleaned up removing the header informatin and saving as NoInt2.txt.

The data file constist of 2 variables and 3 observations.

The first variable is the response variable (y) and second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, with no intercept

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept		
X	0.727272727272727	0.420827318078432E-01

Residual	.
Standard Deviation	0.369274472937998
R-Squared	0.993348115299335

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	40.7272727272727	40.7272727272727	298.666666666667
Residual	2	0.272727272727273	0.136363636363636	

```
NoInt2 <- read.table(file=paste0(path,"NoInt2.txt"), header=TRUE)
NoInt2.lm <- lm(y~x + 0, data=NoInt2)
summary(NoInt2.lm)
```

```
##
```

```
## Call:
## lm(formula = y ~ x + 0, data = NoInt2)
##
## Residuals:
##                1                2                3
##  0.0909090909091  0.3636363636364 -0.3636363636364
##
## Coefficients:
##          Estimate      Std. Error  t value  Pr(>|t|)
## x 0.7272727272727 0.0420827318078 17.28198 0.0033315 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.369274472938 on 2 degrees of freedom
## Multiple R-squared:  0.993348115299, Adjusted R-squared:  0.990022172949
## F-statistic: 298.666666667 on 1 and 2 DF,  p-value: 0.00333149176904
```

```
anova(NoInt2.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df         Sum Sq        Mean Sq   F value    Pr(>F)
## x          1 40.72727272727 40.72727272727 298.66667 0.0033315 **
## Residuals  2  0.27272727273  0.13636363636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Filip.dat file. The file was cleaned up removing the header informatin and saving as Filip.txt.

The data file constist of 2 variables and 82 observations.

The first variable is the response variable (y) and second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 10th order poly

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-1467.48961422980	298.084530995537
X	-2772.17959193342	559.779865474950
X^2	-2316.37108160893	466.477572127796
X^3	-1127.97394098372	227.204274477751
X^4	-354.478233703349	71.6478660875927
X^5	-75.1242017393757	15.2897178747400
X^6	-10.8753180355343	2.23691159816033
X^7	-1.06221498588947	0.221624321934227
X^8	-0.670191154593408E-01	0.142363763154724E-01
X^9	-0.246781078275479E-02	0.535617408889821E-03
X^10	-0.402962525080404E-04	0.896632837373868E-05

Residual	.
Standard Deviation	0.334801051324544E-02
R-Squared	0.996727416185620

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	10	0.242391619837339	0.242391619837339E-01	2162.43954511489
Residual	71	0.795851382172941E-03	0.112091743968020E-04	

```
Filip <- read.table(file=paste0(path,"Filip.txt"), header=TRUE)

Filip.lm <- lm(y ~ poly(x,10,raw = TRUE), data=Filip)
summary(Filip.lm)
```

```
##
## Call:
## lm(formula = y ~ poly(x, 10, raw = TRUE), data = Filip)
##
## Residuals:
##                 Min                  1Q              Median
## -0.009908657615127 -0.002461025793474  0.000338470309395
##                  3Q                 Max
##  0.002074343933875  0.007165411220747
##
## Coefficients: (1 not defined because of singularities)
##                                    Estimate          Std. Error  t value
## (Intercept)                -1.74280442892e+02  8.75611625513e+01 -1.99039
## poly(x, 10, raw = TRUE)1   -3.26882205366e+02  1.48049618747e+02 -2.20792
## poly(x, 10, raw = TRUE)2   -2.66056537165e+02  1.09512084969e+02 -2.42947
## poly(x, 10, raw = TRUE)3   -1.23921613038e+02  4.65247149238e+01 -2.66357
## poly(x, 10, raw = TRUE)4   -3.63816705806e+01  1.25145201106e+01 -2.90716
## poly(x, 10, raw = TRUE)5   -6.97918831615e+00  2.21116610921e+00 -3.15634
## poly(x, 10, raw = TRUE)6   -8.74660171314e-01  2.56744909479e-01 -3.40673
## poly(x, 10, raw = TRUE)7   -6.90600968728e-02  1.89005629387e-02 -3.65386
## poly(x, 10, raw = TRUE)8   -3.11832187934e-03  8.00878988759e-04 -3.89362
## poly(x, 10, raw = TRUE)9   -6.13867079287e-05  1.48905169468e-05 -4.12254
## poly(x, 10, raw = TRUE)10                  NA                 NA       NA
##                              Pr(>|t|)
## (Intercept)                 0.05034548 .
## poly(x, 10, raw = TRUE)1    0.03043560 *
## poly(x, 10, raw = TRUE)2    0.01761673 *
## poly(x, 10, raw = TRUE)3    0.00953389 **
## poly(x, 10, raw = TRUE)4    0.00484493 **
## poly(x, 10, raw = TRUE)5    0.00233320 **
## poly(x, 10, raw = TRUE)6    0.00107912 **
## poly(x, 10, raw = TRUE)7    0.00048725 ***
## poly(x, 10, raw = TRUE)8    0.00021857 ***
```

```
## poly(x, 10, raw = TRUE)9   9.907e-05 ***
## poly(x, 10, raw = TRUE)10         NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00376801219438 on 72 degrees of freedom
## Multiple R-squared:  0.995796453084, Adjusted R-squared:  0.99527100972
## F-statistic: 1895.15468352 on 9 and 72 DF,  p-value: < 2.220446049e-16
```

```
anova(Filip.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##                          Df           Sum Sq           Mean Sq    F value
## poly(x, 10, raw = TRUE)  9 0.24216522127493 0.026907246808326 1895.15468
## Residuals               72 0.00102224994458 0.000014197915897
##                             Pr(>F)
## poly(x, 10, raw = TRUE) < 2.22e-16 ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Longley.dat file. The file was cleaned up removing the header informatin and saving as Longley.txt.

The data file constist of 7 variables and 16 observations.

The first variable is the response variable (y) and 2-7 variables are the predictor variables (x1, x2, x3, x4, x5, x6).

Procedure: Linear Least Squares Regression, 6 predictor terms

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-3482258.63459582	890420.383607373
X1	15.0618722713733	84.9149257747669
X2	-0.358191792925910E-01	0.334910077722432E-01
X3	-2.02022980381683	0.488399681651699
X4	-1.03322686717359	0.214274163161675
X5	-0.511041056535807E-01	0.226073200069370
X6	1829.15146461355	455.478499142212

Residual	.
Standard Deviation	304.854073561965
R-Squared	0.995479004577296

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	6	184172401.944494	30695400.3240823	330.285339234588
Residual	9	836424.055505915	92936.0061673238	

```
Longley <- read.table(file=paste0(path,"Longley.txt"), header=TRUE)

Longley.lm <- lm(y ~ .,
                 data=Longley)
summary(Longley.lm)
```

```
##
## Call:
## lm(formula = y ~ ., data = Longley)
##
## Residuals:
##             Min              1Q          Median              3Q
## -410.1146219309 -157.6747192954  -28.1619848188  101.5503832581
##             Max
##  455.3940945519
##
## Coefficients:
##                        Estimate         Std. Error  t value   Pr(>|t|)
## (Intercept) -3.48225863460e+06  8.90420383607e+05 -3.91080 0.00356040 **
## x1           1.50618722714e+01  8.49149257748e+01  0.17738 0.86314083
## x2          -3.58191792926e-02  3.34910077722e-02 -1.06952 0.31268106
## x3          -2.02022980382e+00  4.88399681652e-01 -4.13643 0.00253509 **
## x4          -1.03322686717e+00  2.14274163162e-01 -4.82199 0.00094437 ***
## x5          -5.11041056536e-02  2.26073200069e-01 -0.22605 0.82621180
## x6           1.82915146461e+03  4.55478499142e+02  4.01589 0.00303680 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 304.854073562 on 9 degrees of freedom
## Multiple R-squared:  0.995479004577, Adjusted R-squared:  0.992465007629
## F-statistic: 330.285339235 on 6 and 9 DF,  p-value: 4.98403052872e-10
```

```
anova(Longley.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df           Sum Sq          Mean Sq    F value     Pr(>F)
## x1         1 174397449.77913 174397449.77913 1876.53265 9.2954e-12 ***
## x2         1   4787181.04445   4787181.04445   51.51051 5.2109e-05 ***
## x3         1   2263971.10982   2263971.10982   24.36054 0.00080706 ***
## x4         1    876397.16186    876397.16186    9.43011 0.01333568 *
## x5         1    348589.39965    348589.39965    3.75085 0.08475523 .
## x6         1   1498813.44959   1498813.44959   16.12737 0.00303680 **
```

```
## Residuals  9    836424.05551     92936.00617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Wampler1.dat file. The file was cleaned up removing the header informatin and saving as Wampler1.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 0.000000000000000
X | 1.00000000000000 | 0.000000000000000 **X^2** | 1.00000000000000 | 0.000000000000000
X^3 | 1.00000000000000 | 0.000000000000000 **X^4** | 1.00000000000000 | 0.000000000000000 **X^5** | 1.00000000000000 | 0.000000000000000

Residual	.
Standard Deviation	0.000000000000000
R-Squared	1.000000000000000

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	Infinity
Residual	15	0.000000000000000	0.000000000000000	

```
Wampler1 <- read.table(file=paste0(path,"Wampler1.txt"), header=TRUE)

Wampler1.lm <- lm(y ~ poly(x,5,raw=T),
                  data = Wampler1)
summary(Wampler1.lm)
```

```
## Warning in summary.lm(Wampler1.lm): essentially perfect fit: summary may
## be unreliable

##
## Call:
## lm(formula = y ~ poly(x, 5, raw = T), data = Wampler1)
##
## Residuals:
##                 Min                  1Q              Median
## -1.11110167972e-10 -4.71266092371e-11 -1.21725378234e-11
##                  3Q                 Max
##  3.11073134364e-11  1.88884728249e-10
##
```

```
## Coefficients:
##                               Estimate        Std. Error           t value
## (Intercept)          9.99999999774e-01 7.45787925731e-11     13408637566.7
## poly(x, 5, raw = T)1 9.99999999964e-01 8.18978232604e-11     12210336736.1
## poly(x, 5, raw = T)2 1.00000000003e+00 2.70045022827e-11     37030862096.9
## poly(x, 5, raw = T)3 9.99999999996e-01 3.51615877944e-12    284401263629.8
## poly(x, 5, raw = T)4 1.00000000000e+00 1.95624101448e-13   5111844566186.9
## poly(x, 5, raw = T)5 1.00000000000e+00 3.89209212611e-15 256931225571572.9
##                         Pr(>|t|)
## (Intercept)          < 2.22e-16 ***
## poly(x, 5, raw = T)1 < 2.22e-16 ***
## poly(x, 5, raw = T)2 < 2.22e-16 ***
## poly(x, 5, raw = T)3 < 2.22e-16 ***
## poly(x, 5, raw = T)4 < 2.22e-16 ***
## poly(x, 5, raw = T)5 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.17797796384e-11 on 15 degrees of freedom
## Multiple R-squared:              1,  Adjusted R-squared:              1
## F-statistic: 5.6263479359e+32 on 5 and 15 DF,  p-value: < 2.220446049e-16
```

```
anova(Wampler1.lm)
```

```
## Warning in anova.lm(Wampler1.lm): ANOVA F-tests on an essentially perfect
## fit are unreliable

## Analysis of Variance Table
##
## Response: y
##                      Df         Sum Sq       Mean Sq            F value
## poly(x, 5, raw = T)  5 18814317208117 3762863441623 5.626347935901e+32
## Residuals           15              0             0
##                         Pr(>F)
## poly(x, 5, raw = T) < 2.22e-16 ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Wampler2.dat file. The file was cleaned up removing the header informatin and saving as Wampler2.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 0.000000000000000
X | 0.10000000000000 | 0.000000000000000 **X^2** | 0.100000000000000E-01 | 0.000000000000000
X^3 | 0.100000000000000E-02 | 0.000000000000000 **X^4** | 0.100000000000000E-03 | 0.000000000000000 **X^5** | 0.100000000000000E-04 | 0.000000000000000

Residual	.
Standard Deviation	0.000000000000000
R-Squared	1.000000000000000

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	6602.91858365167	1320.58371673033	Infinity
Residual	15	0.000000000000000	0.000000000000000	

```
Wampler2 <- read.table(file=paste0(path,"Wampler2.txt"),
                       header=TRUE)
Wampler2.lm <- lm(y ~ poly(x,5,raw=T),
                  data = Wampler2)
summary(Wampler2.lm)
```

```
## Warning in summary.lm(Wampler2.lm): essentially perfect fit: summary may
## be unreliable
```

```
##
## Call:
## lm(formula = y ~ poly(x, 5, raw = T), data = Wampler2)
##
## Residuals:
##                 Min                  1Q              Median
## -2.80817271630e-15 -6.30572214793e-16  6.38723617449e-17
##                  3Q                 Max
##  4.82605680278e-16  2.31461555360e-15
##
## Coefficients:
##                               Estimate        Std. Error         t value
## (Intercept)          1.00000000000e+00 1.16053772926e-15 861669530239097
## poly(x, 5, raw = T)1 1.00000000000e-01 1.27443084769e-15  78466399476704
## poly(x, 5, raw = T)2 1.00000000000e-02 4.20223265593e-16  23796873754436
## poly(x, 5, raw = T)3 1.00000000000e-03 5.47157547721e-17  18276271691134
## poly(x, 5, raw = T)4 1.00000000000e-04 3.04415159659e-18  32849875187509
## poly(x, 5, raw = T)5 1.00000000000e-05 6.05657399683e-20 165109846015880
##                         Pr(>|t|)
## (Intercept)          < 2.22e-16 ***
## poly(x, 5, raw = T)1 < 2.22e-16 ***
## poly(x, 5, raw = T)2 < 2.22e-16 ***
## poly(x, 5, raw = T)3 < 2.22e-16 ***
## poly(x, 5, raw = T)4 < 2.22e-16 ***
## poly(x, 5, raw = T)5 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.27259394376e-15 on 15 degrees of freedom
```

```
## Multiple R-squared:              1,  Adjusted R-squared:              1
## F-statistic: 8.15429152201e+32 on 5 and 15 DF,  p-value: < 2.220446049e-16
```

```
anova(Wampler2.lm)
```

```
## Warning in anova.lm(Wampler2.lm): ANOVA F-tests on an essentially perfect
## fit are unreliable
```

```
## Analysis of Variance Table
##
## Response: y
##                      Df         Sum Sq       Mean Sq           F value
## poly(x, 5, raw = T)  5 6602.918583652 1320.58371673 8.15429152201e+32
## Residuals           15    0.000000000    0.00000000
##                          Pr(>F)
## poly(x, 5, raw = T) < 2.22e-16 ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Wampler3.dat file. The file was cleaned up removing the header informatin and saving as Wampler3.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 2152.32624678170
X | 1.00000000000000 | 2363.55173469681
X^2 | 1.00000000000000 | 779.343524331583
X^3 | 1.00000000000000 |101.475507550350
X^4 | 1.00000000000000 | 5.64566512170752
X^5 | 1.00000000000000 | 0.112324854679312

Residual	.
Standard Deviation	2360.14502379268
R-Squared	0.999995559025820

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	675524.458240122
Residual	15	83554268.0000000	5570284.53333333	

```
Wampler3 <- read.table(file=paste0(path,"Wampler3.txt"),
                       header=TRUE)
Wampler3.lm <- lm(y ~ poly(x,5,raw=T),
                  data = Wampler3)
summary(Wampler3.lm)
```

```
##
## Call:
## lm(formula = y ~ poly(x, 5, raw = T), data = Wampler3)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -2048  -2048    759   2048   2523
##
## Coefficients:
##                             Estimate        Std. Error t value
## (Intercept)             0.999999999864 2152.326246781733 0.00046
## poly(x, 5, raw = T)1    0.999999999936 2363.551734696791 0.00042
## poly(x, 5, raw = T)2    1.000000000046  779.343524331578 0.00128
## poly(x, 5, raw = T)3    0.999999999992  101.475507550349 0.00985
## poly(x, 5, raw = T)4    1.000000000000    5.645665121707 0.17713
## poly(x, 5, raw = T)5    1.000000000000    0.112324854679 8.90275
##                          Pr(>|t|)
## (Intercept)               0.99964
## poly(x, 5, raw = T)1      0.99967
## poly(x, 5, raw = T)2      0.99899
## poly(x, 5, raw = T)3      0.99227
## poly(x, 5, raw = T)4      0.86178
## poly(x, 5, raw = T)5 2.2534e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2360.14502379 on 15 degrees of freedom
## Multiple R-squared:  0.999995559026, Adjusted R-squared:  0.999994078701
## F-statistic:  675524.45824 on 5 and 15 DF,  p-value: < 2.220446049e-16
```

```
anova(Wampler3.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##                       Df         Sum Sq       Mean Sq      F value
## poly(x, 5, raw = T)  5 18814317208117 3762863441623 675524.45824
## Residuals           15       83554268       5570285
##                          Pr(>F)
## poly(x, 5, raw = T) < 2.22e-16 ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Wampler4.dat file. The file was cleaned up removing the header informatin and saving as Wampler4.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**|
1.00000000000000 | 215232.624678170
X | 1.00000000000000 | 236355.173469681
X^2 | 1.00000000000000 | 77934.3524331583
X^3 | 1.00000000000000 | 10147.5507550350
X^4 | 1.00000000000000 | 564.566512170752
X^5 | 1.00000000000000 | 11.2324854679312

Residual	.
Standard Deviation	236014.502379268
R-Squared	0.957478440825662

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	67.5524458240122
Residual	15	835542680000.000	55702845333.3333	

```
Wampler4 <- read.table(file=paste0(path,"Wampler4.txt"),
                       header=TRUE)
Wampler4.lm <- lm(y ~ poly(x,5,raw=T),
                  data = Wampler4)
summary(Wampler4.lm)
```

```
##
## Call:
## lm(formula = y ~ poly(x, 5, raw = T), data = Wampler4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -204800 -204800   75900  204800  252300
##
## Coefficients:
##                               Estimate        Std. Error t value Pr(>|t|)
## (Intercept)          1.00000000428e+00 2.15232624678e+05 0.00000  1.00000
## poly(x, 5, raw = T)1 9.99999991568e-01 2.36355173470e+05 0.00000  1.00000
## poly(x, 5, raw = T)2 1.00000000300e+00 7.79343524332e+04 0.00001  0.99999
## poly(x, 5, raw = T)3 9.99999999609e-01 1.01475507550e+04 0.00010  0.99992
## poly(x, 5, raw = T)4 1.00000000002e+00 5.64566512171e+02 0.00177  0.99861
## poly(x, 5, raw = T)5 1.00000000000e+00 1.12324854679e+01 0.08903  0.93024
##
## Residual standard error: 236014.502379 on 15 degrees of freedom
```

```
## Multiple R-squared:  0.957478440826, Adjusted R-squared:  0.943304587768
## F-statistic:  67.552445824 on 5 and 15 DF,  p-value: 9.51904357155e-10
```

```
anova(Wampler4.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##                      Df         Sum Sq       Mean Sq  F value    Pr(>F)
## poly(x, 5, raw = T)   5 18814317208117 3762863441623 67.55245 9.519e-10 ***
## Residuals            15   835542680000   55702845333
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using Wampler5.dat file. The file was cleaned up removing the header informatin and saving as Wampler5.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 21523262.4678170
X | 1.00000000000000 | 23635517.3469681
X^2 | 1.00000000000000 | 7793435.24331583
X^3 | 1.00000000000000 | 1014755.07550350
X^4 | 1.00000000000000 | 56456.6512170752
X^5 | 1.00000000000000 | 1123.24854679312

Residual	.
Standard Deviation	23601450.2379268
R-Squared	0.224668921574940E-02

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	0.675524458240122E-02
Residual	15	0.835542680000000E+16	557028453333333	

```
Wampler5 <- read.table(file=paste0(path,"Wampler5.txt"),
                       header=TRUE)
Wampler5.lm <- lm(y ~ poly(x,5,raw=T),
                  data = Wampler5)
summary(Wampler5.lm)
```

```
##
## Call:
## lm(formula = y ~ poly(x, 5, raw = T), data = Wampler5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -20480000 -20480000   7590000  20480000  25230000
##
## Coefficients:
##                               Estimate        Std. Error t value Pr(>|t|)
## (Intercept)          1.00000042671e+00 2.15232624678e+07 0.00000  1.00000
## poly(x, 5, raw = T)1 9.99999184393e-01 2.36355173470e+07 0.00000  1.00000
## poly(x, 5, raw = T)2 1.00000028970e+00 7.79343524332e+06 0.00000  1.00000
## poly(x, 5, raw = T)3 9.99999962086e-01 1.01475507550e+06 0.00000  1.00000
## poly(x, 5, raw = T)4 1.00000000208e+00 5.64566512171e+04 0.00002  0.99999
## poly(x, 5, raw = T)5 9.99999999960e-01 1.12324854679e+03 0.00089  0.99930
##
## Residual standard error: 23601450.2379 on 15 degrees of freedom
## Multiple R-squared:  0.00224668921575,   Adjusted R-squared:  -0.330337747712
## F-statistic: 0.0067552445824 on 5 and 15 DF,  p-value: 0.99998618409
```

```
anova(Wampler5.lm)
```

```
## Analysis of Variance Table
##
## Response: y
##                      Df           Sum Sq         Mean Sq F value  Pr(>F)
## poly(x, 5, raw = T)   5   18814317208117   3762863441623 0.00676 0.99999
## Residuals            15 8355426800000000 557028453333333
```

Linear Regression using glm()

Using Norris.dat file. The file was cleaned up removing the header informatin and saving as Norris.txt. The data file constist of 2 variables and 36 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-0.262323073774029	0.232818234301152
X	1.00211681802045	0.429796848199937E-03

Residual	.
Standard Deviation	0.884796396144373
R-Squared	0.999993745883712

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	4255954.13232369	4255954.13232369	5436385.54079785
Residual	34	26.6173985294224	0.782864662630069	

```
Norris <- read.table(file=paste0(path,"Norris.txt"), header=TRUE)
Norris.glm <- glm(y~x, data=Norris)
summary(Norris.glm)
```

```
##
## Call:
## glm(formula = y ~ x, data = Norris)
##
## Deviance Residuals:
##              Min                1Q            Median                3Q
## -2.3523781286599  -0.5326967162014  -0.0296292259639   0.6000277736810
##              Max
##  1.7897858528801
##
## Coefficients:
##                      Estimate       Std. Error    t value Pr(>|t|)
## (Intercept) -0.2623230737740  0.2328182343012   -1.12673  0.26775
## x            1.0021168180205  0.0004297968482 2331.60579  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.782864662630083)
##
##     Null deviance: 4.255980749722e+06  on 35  degrees of freedom
## Residual deviance: 2.661739852942e+01  on 34  degrees of freedom
## AIC: 97.29323555918
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Norris.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##      Df       Deviance Resid. Df     Resid. Dev
## NULL                           35 4255980.749722
## x     1 4255954.132324        34      26.617399
```

Using Pontius.dat file. The file was cleaned up removing the header informatin and saving as Norris.txt.

The data file constist of 2 variables and 40 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression with a quadric term

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	0.673565789473684E-03	0.107938612033077E-03
X	0.732059160401003E-06	0.157817399981659E-09
x^2	-0.316081871345029E-14	0.486652849992036E-16

Residual	.
Standard Deviation	0.205177424076185E-03
R-Squared	0.999999900178537

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	2	15.6040343244198	7.80201716220991	185330865.995752
Residual	37	0.155761768796992E-05	0.420977753505385E-07	

```
Pontius <- read.table(file=paste0(path,"Pontius.txt"), header=TRUE)
Pontius.glm <- glm(y~x + I(x^2), data=Pontius)
summary(Pontius.glm)
```

```
##
## Call:
## glm(formula = y ~ x + I(x^2), data = Pontius)
##
## Deviance Residuals:
##                 Min                   1Q               Median
## -4.46840225564e-04  -1.57827067669e-04   3.81729323307e-05
##                  3Q                  Max
##  1.08788533835e-04   4.23453007519e-04
##
## Coefficients:
##                        Estimate         Std. Error    t value   Pr(>|t|)
## (Intercept)  6.73565789473e-04  1.07938612033e-04    6.24027 2.9705e-07
## x            7.32059160401e-07  1.57817399982e-10 4638.64669 < 2.22e-16
## I(x^2)      -3.16081871345e-15  4.86652849992e-17  -64.95017 < 2.22e-16
##
## (Intercept) ***
## x           ***
```

```
## I(x^2)      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 4.20977753505387e-08)
##
##     Null deviance: 1.560403588204e+01  on 39  degrees of freedom
## Residual deviance: 1.557617687970e-06  on 37  degrees of freedom
## AIC: -560.9342165898
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Pontius.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##        Df       Deviance Resid. Df      Resid. Dev
## NULL                            39 15.604035882038
## x       1 15.60385673390        38  0.000179148138
## I(x^2)  1  0.00017759052        37  0.000001557618
```

Using NoInt1.dat file. The file was cleaned up removing the header informatin and saving as NoInt1.txt.

The data file constist of 2 variables and 11 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, with no intercept Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept		
X	2.07438016528926	0.165289256198347E-01

Residual	.
Standard Deviation	3.56753034006338
R-Squared	0.999365492298663

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	200457.727272727	200457.727272727	15750.2500000000

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Residual	10	127.272727272727	12.7272727272727	

```
NoInt1 <- read.table(file=paste0(path,"NoInt1.txt"), header=TRUE)
NoInt1.glm <- glm(y~x + 0, data=NoInt1)
summary(NoInt1.glm)
```

```
##
## Call:
## glm(formula = y ~ x + 0, data = NoInt1)
##
## Deviance Residuals:
##              Min                1Q            Median                3Q
## -5.206611570248  -2.520661157025   0.165289256198   2.851239669421
##              Max
##  5.537190082645
##
## Coefficients:
##          Estimate      Std. Error t value   Pr(>|t|)
## x 2.0743801652893 0.0165289256198   125.5 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 12.7272727272727)
##
##     Null deviance: 200585.0000000000  on 11  degrees of freedom
## Residual deviance:    127.2727272727  on 10  degrees of freedom
## AIC: 62.14945440058
##
## Number of Fisher Scoring iterations: 2
```

```
anova(NoInt1.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##      Df       Deviance Resid. Df      Resid. Dev
## NULL                           11 200585.00000000
## x     1 200457.7272727         10    127.27272727
```

Using NoInt2.dat file. The file was cleaned up removing the header informatin and saving as NoInt2.txt. The data file constist of 2 variables and 3 observations.

The first variable is the response variable (y) and second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, with no intercept

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept		
X	0.727272727272727	0.420827318078432E-01

Residual	.
Standard Deviation	0.369274472937998
R-Squared	0.993348115299335

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	1	40.7272727272727	40.7272727272727	298.666666666667
Residual	2	0.272727272727273	0.136363636363636	

```
NoInt2 <- read.table(file=paste0(path,"NoInt2.txt"), header=TRUE)
NoInt2.glm <- glm(y~x + 0, data=NoInt2)
summary(NoInt2.glm)
```

```
##
## Call:
## glm(formula = y ~ x + 0, data = NoInt2)
##
## Deviance Residuals:
##                1                2                3
##  0.0909090909091  0.3636363636364 -0.3636363636364
##
## Coefficients:
##          Estimate      Std. Error  t value  Pr(>|t|)
## x 0.7272727272727 0.0420827318078 17.28198 0.0033315 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.136363636363636)
##
##     Null deviance: 41.0000000000000  on 3  degrees of freedom
## Residual deviance:  0.2727272727273  on 2  degrees of freedom
## AIC: 5.319945380833
##
## Number of Fisher Scoring iterations: 2
```

```
anova(NoInt2.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##      Df       Deviance Resid. Df     Resid. Dev
## NULL                            3 41.00000000000
## x     1 40.72727272727          2  0.27272727273
```

Using Filip.dat file. The file was cleaned up removing the header informatin and saving as Filip.txt.

The data file constist of 2 variables and 82 observations.

The first variable is the response variable (y) and second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 10th order poly

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-1467.48961422980	298.084530995537
X	-2772.17959193342	559.779865474950
X^2	-2316.37108160893	466.477572127796
X^3	-1127.97394098372	227.204274477751
X^4	-354.478233703349	71.6478660875927
X^5	-75.1242017393757	15.2897178747400
X^6	-10.8753180355343	2.23691159816033
X^7	-1.06221498588947	0.221624321934227
X^8	-0.670191154593408E-01	0.142363763154724E-01
X^9	-0.246781078275479E-02	0.535617408889821E-03
X^10	-0.402962525080404E-04	0.896632837373868E-05

Residual	.
Standard Deviation	0.334801051324544E-02
R-Squared	0.996727416185620

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	10	0.242391619837339	0.242391619837339E-01	2162.43954511489
Residual	71	0.795851382172941E-03	0.112091743968020E-04	

```
Filip <- read.table(file=paste0(path,"Filip.txt"), header=TRUE)

Filip.glm <- glm(y ~ poly(x,10,raw = TRUE), data=Filip)
summary(Filip.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 10, raw = TRUE), data = Filip)
##
## Deviance Residuals:
##                 Min                   1Q               Median
## -0.008804382818029  -0.002176048806714   0.000045016427861
##                  3Q                  Max
##  0.002028837646824   0.007096030979698
##
## Coefficients:
##                                     Estimate         Std. Error  t value
## (Intercept)                -1.46748962941e+03  2.98084531819e+02 -4.92307
## poly(x, 10, raw = TRUE)1   -2.77217962035e+03  5.59779867001e+02 -4.95227
## poly(x, 10, raw = TRUE)2   -2.31637110528e+03  4.66477573397e+02 -4.96566
## poly(x, 10, raw = TRUE)3   -1.12797395255e+03  2.27204275103e+02 -4.96458
## poly(x, 10, raw = TRUE)4   -3.54478237372e+02  7.16478662892e+01 -4.94751
## poly(x, 10, raw = TRUE)5   -7.51242025294e+01  1.52897179192e+01 -4.91338
## poly(x, 10, raw = TRUE)6   -1.08753181525e+01  2.23691160495e+00 -4.86176
## poly(x, 10, raw = TRUE)7   -1.06221499764e+00  2.21624322640e-01 -4.79286
## poly(x, 10, raw = TRUE)8   -6.70191162255e-02  1.42363763632e-02 -4.70760
## poly(x, 10, raw = TRUE)9   -2.46781081206e-03  5.35617410786e-04 -4.60741
## poly(x, 10, raw = TRUE)10 -4.02962530070e-05  8.96632840723e-06 -4.49418
##                              Pr(>|t|)
## (Intercept)                5.3468e-06 ***
## poly(x, 10, raw = TRUE)1   4.7835e-06 ***
## poly(x, 10, raw = TRUE)2   4.5449e-06 ***
## poly(x, 10, raw = TRUE)3   4.5637e-06 ***
## poly(x, 10, raw = TRUE)4   4.8712e-06 ***
## poly(x, 10, raw = TRUE)5   5.5476e-06 ***
## poly(x, 10, raw = TRUE)6   6.7487e-06 ***
## poly(x, 10, raw = TRUE)7   8.7537e-06 ***
## poly(x, 10, raw = TRUE)8   1.2051e-05 ***
## poly(x, 10, raw = TRUE)9   1.7486e-05 ***
## poly(x, 10, raw = TRUE)10 2.6515e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.12091743398651e-05)
##
##     Null deviance: 0.2431874712195122  on 81  degrees of freedom
## Residual deviance: 0.0007958513781304  on 71  degrees of freedom
```

```
## AIC: -689.8051030665
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Filip.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##                          Df        Deviance Resid. Df       Resid. Dev
## NULL                                            81 0.24318747121951
## poly(x, 10, raw = TRUE) 10 0.2423916198414        71 0.00079585137813
```

Using Longley.dat file. The file was cleaned up removing the header informatin and saving as Longley.txt.

The data file constist of 7 variables and 16 observations.

The first variable is the response variable (y) and 2-7 variables are the predictor variables (x1, x2, x3, x4, x5, x6).

Procedure: Linear Least Squares Regression, 6 predictor terms

Expected results:

Parameter	Estimate	Standard Deviation of Estimate
intercept	-3482258.63459582	890420.383607373
X1	15.0618722713733	84.9149257747669
X2	-0.358191792925910E-01	0.334910077722432E-01
X3	-2.02022980381683	0.488399681651699
X4	-1.03322686717359	0.214274163161675
X5	-0.511041056535807E-01	0.226073200069370
X6	1829.15146461355	455.478499142212

Residual	.
Standard Deviation	304.854073561965
R-Squared	0.995479004577296

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	6	184172401.944494	30695400.3240823	330.285339234588
Residual	9	836424.055505915	92936.0061673238	

```
Longley <- read.table(file=paste0(path,"Longley.txt"), header=TRUE)

Longley.glm <- glm(y ~ .,
                   data=Longley)
summary(Longley.glm)
```

```
##
## Call:
## glm(formula = y ~ ., data = Longley)
##
## Deviance Residuals:
##              Min                1Q            Median                3Q
## -410.1146219302  -157.6747192953   -28.1619848182   101.5503832585
##              Max
##  455.3940945528
##
## Coefficients:
##                         Estimate         Std. Error  t value   Pr(>|t|)
## (Intercept) -3.48225863460e+06  8.90420383608e+05 -3.91080 0.00356040 **
## x1           1.50618722714e+01  8.49149257748e+01  0.17738 0.86314083
## x2          -3.58191792926e-02  3.34910077723e-02 -1.06952 0.31268106
## x3          -2.02022980382e+00  4.88399681652e-01 -4.13643 0.00253509 **
## x4          -1.03322686717e+00  2.14274163162e-01 -4.82199 0.00094437 ***
## x5          -5.11041056536e-02  2.26073200069e-01 -0.22605 0.82621180
## x6           1.82915146461e+03  4.55478499142e+02  4.01589 0.00303680 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 92936.006167419)
##
##     Null deviance: 185008826.0000000  on 15  degrees of freedom
## Residual deviance:    836424.0555068  on  9  degrees of freedom
## AIC: 235.234869617
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Longley.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
```

```
##
##      Df        Deviance Resid. Df      Resid. Dev
## NULL                           15 185008826.00000
## x1    1 174397449.77913        14  10611376.22087
## x2    1   4787181.04445        13   5824195.17642
## x3    1   2263971.10982        12   3560224.06660
## x4    1    876397.16186        11   2683826.90474
## x5    1    348589.39965        10   2335237.50509
## x6    1   1498813.44959         9    836424.05551
```

Using Wampler1.dat file. The file was cleaned up removing the header informatin and saving as Wampler1.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 0.000000000000000
X | 1.00000000000000 | 0.000000000000000 **X^2** | 1.00000000000000 | 0.000000000000000
X^3 | 1.00000000000000 | 0.000000000000000 **X^4** | 1.00000000000000 | 0.000000000000000 **X^5** | 1.00000000000000 | 0.000000000000000

Residual	.
Standard Deviation	0.000000000000000
R-Squared	1.000000000000000

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	Infinity
Residual	15	0.000000000000000	0.000000000000000	

```
Wampler1 <- read.table(file=paste0(path,"Wampler1.txt"), header=TRUE)

Wampler1.glm <- glm(y ~ poly(x,5,raw=T),
                    data = Wampler1)
summary(Wampler1.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 5, raw = T), data = Wampler1)
##
## Deviance Residuals:
##                Min                  1Q              Median
## 0.00000000000e+00  0.00000000000e+00  5.82076609135e-11
```

```
##                3Q                Max
## 1.17324816529e-10  2.40090614057e-10
##
## Coefficients:
##                              Estimate        Std. Error           t value
## (Intercept)          9.99999999774e-01 1.21323524761e-10 8.24242455650e+09
## poly(x, 5, raw = T)1 9.99999999964e-01 1.33230000720e-10 7.50581696733e+09
## poly(x, 5, raw = T)2 1.00000000003e+00 4.39304698873e-11 2.27632438850e+10
## poly(x, 5, raw = T)3 9.99999999996e-01 5.72002793320e-12 1.74824321083e+11
## poly(x, 5, raw = T)4 1.00000000000e+00 3.18237996313e-13 3.14230233846e+12
## poly(x, 5, raw = T)5 1.00000000000e+00 6.33158997544e-15 1.57938212026e+14
##                        Pr(>|t|)
## (Intercept)          < 2.22e-16 ***
## poly(x, 5, raw = T)1 < 2.22e-16 ***
## poly(x, 5, raw = T)2 < 2.22e-16 ***
## poly(x, 5, raw = T)3 < 2.22e-16 ***
## poly(x, 5, raw = T)4 < 2.22e-16 ***
## poly(x, 5, raw = T)5 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.76991012867714e-20)
##
##     Null deviance: 1.881431720812e+13  on 20  degrees of freedom
## Residual deviance: 2.654865193016e-19  on 15  degrees of freedom
## AIC: -888.5667334507
##
## Number of Fisher Scoring iterations: 1
```

```
anova(Wampler1.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##                     Df       Deviance Resid. Df     Resid. Dev
## NULL                                        20 18814317208117
## poly(x, 5, raw = T)  5 18814317208117        15              0
```

Using Wampler2.dat file. The file was cleaned up removing the header informatin and saving as Wampler2.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 0.000000000000000
X | 0.100000000000000 | 0.000000000000000 **X^2** | 0.100000000000000E-01 | 0.000000000000000
X^3 | 0.100000000000000E-02 | 0.000000000000000 **X^4** | 0.100000000000000E-03 | 0.000000000000000 **X^5** | 0.100000000000000E-04 | 0.000000000000000

Residual	.
Standard Deviation	0.000000000000000
R-Squared	1.000000000000000

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	6602.91858365167	1320.58371673033	Infinity
Residual	15	0.000000000000000	0.000000000000000	

```
Wampler2 <- read.table(file=paste0(path,"Wampler2.txt"),
                       header=TRUE)
Wampler2.glm <- glm(y ~ poly(x,5,raw=T),
                    data = Wampler2)
summary(Wampler2.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 5, raw = T), data = Wampler2)
##
## Deviance Residuals:
##                Min                  1Q              Median
## 0.00000000000e+00  1.77635683940e-15  3.55271367880e-15
##                 3Q                 Max
## 3.99680288865e-15  6.55031584529e-15
##
## Coefficients:
##                              Estimate        Std. Error         t value
## (Intercept)          1.00000000000e+00 3.75176818855e-15 266540988073984
## poly(x, 5, raw = T)1 1.00000000000e-01 4.11996007737e-15  24272079855632
## poly(x, 5, raw = T)2 1.00000000000e-02 1.35849118920e-15   7361107734446
## poly(x, 5, raw = T)3 1.00000000000e-03 1.76884234773e-16   5653415078414
## poly(x, 5, raw = T)4 1.00000000000e-04 9.84108558751e-18  10161480571509
## poly(x, 5, raw = T)5 1.00000000000e-05 1.95795975261e-19  51073572818113
##                        Pr(>|t|)
## (Intercept)          < 2.22e-16 ***
## poly(x, 5, raw = T)1 < 2.22e-16 ***
## poly(x, 5, raw = T)2 < 2.22e-16 ***
## poly(x, 5, raw = T)3 < 2.22e-16 ***
## poly(x, 5, raw = T)4 < 2.22e-16 ***
## poly(x, 5, raw = T)5 < 2.22e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.69251750675387e-29)
##
##     Null deviance: 6.602918583652e+03  on 20  degrees of freedom
## Residual deviance: 2.538776260131e-28  on 15  degrees of freedom
## AIC: -1324.694261749
##
## Number of Fisher Scoring iterations: 1
```

```
anova(Wampler2.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##                      Df       Deviance Resid. Df     Resid. Dev
## NULL                                        20 6602.918583652
## poly(x, 5, raw = T)  5 6602.918583652        15    0.000000000
```

Using Wampler3.dat file. The file was cleaned up removing the header informatin and saving as Wampler3.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 2152.32624678170
X | 1.00000000000000 | 2363.55173469681
X^2 | 1.00000000000000 | 779.343524331583
X^3 | 1.00000000000000 |101.475507550350
X^4 | 1.00000000000000 | 5.64566512170752
X^5 | 1.00000000000000 | 0.112324854679312

Residual	.
Standard Deviation	2360.14502379268
R-Squared	0.999995559025820

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	675524.458240122

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Residual	15	83554268.0000000	5570284.53333333	

```
Wampler3 <- read.table(file=paste0(path,"Wampler3.txt"),
                       header=TRUE)
Wampler3.glm <- glm(y ~ poly(x,5,raw=T),
                    data = Wampler3)
summary(Wampler3.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 5, raw = T), data = Wampler3)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
##  -2048   -2048     759    2048    2523
##
## Coefficients:
##                             Estimate        Std. Error t value
## (Intercept)             0.999999999864 2152.326246781690 0.00046
## poly(x, 5, raw = T)1    0.999999999936 2363.551734696744 0.00042
## poly(x, 5, raw = T)2    1.000000000046  779.343524331562 0.00128
## poly(x, 5, raw = T)3    0.999999999992  101.475507550347 0.00985
## poly(x, 5, raw = T)4    1.000000000000    5.645665121707 0.17713
## poly(x, 5, raw = T)5    1.000000000000    0.112324854679 8.90275
##                         Pr(>|t|)
## (Intercept)              0.99964
## poly(x, 5, raw = T)1     0.99967
## poly(x, 5, raw = T)2     0.99899
## poly(x, 5, raw = T)3     0.99227
## poly(x, 5, raw = T)4     0.86178
## poly(x, 5, raw = T)5 2.2534e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 5570284.53333328)
##
##     Null deviance: 18814400762385  on 20  degrees of freedom
## Residual deviance:       83554268  on 15  degrees of freedom
## AIC: 392.721591995
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Wampler3.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
```

```
##
## Terms added sequentially (first to last)
##
##
##                      Df       Deviance Resid. Df     Resid. Dev
## NULL                                          20 18814400762385
## poly(x, 5, raw = T)   5 18814317208117        15       83554268
```

Using Wampler4.dat file. The file was cleaned up removing the header informatin and saving as Wampler4.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ————|————|———- **intercept**| 1.00000000000000 | 215232.624678170
X | 1.00000000000000 | 236355.173469681
X^2 | 1.00000000000000 | 77934.3524331583
X^3 | 1.00000000000000 | 10147.5507550350
X^4 | 1.00000000000000 | 564.566512170752
X^5 | 1.00000000000000 | 11.2324854679312

Residual	.
Standard Deviation	236014.502379268
R-Squared	0.957478440825662

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	67.5524458240122
Residual	15	835542680000.000	55702845333.3333	

```
Wampler4 <- read.table(file=paste0(path,"Wampler4.txt"),
                       header=TRUE)
Wampler4.glm <- glm(y ~ poly(x,5,raw=T),
                    data = Wampler4)
summary(Wampler4.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 5, raw = T), data = Wampler4)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -204800  -204800    75900   204800   252300
```

```
##
## Coefficients:
##                               Estimate        Std. Error t value Pr(>|t|)
## (Intercept)          1.00000000428e+00 2.15232624678e+05 0.00000  1.00000
## poly(x, 5, raw = T)1 9.99999991568e-01 2.36355173470e+05 0.00000  1.00000
## poly(x, 5, raw = T)2 1.00000000300e+00 7.79343524332e+04 0.00001  0.99999
## poly(x, 5, raw = T)3 9.99999999609e-01 1.01475507550e+04 0.00010  0.99992
## poly(x, 5, raw = T)4 1.00000000002e+00 5.64566512171e+02 0.00177  0.99861
## poly(x, 5, raw = T)5 1.00000000000e+00 1.12324854679e+01 0.08903  0.93024
##
## (Dispersion parameter for gaussian family taken to be 55702845333.3333)
##
##     Null deviance: 19649859888117  on 20  degrees of freedom
## Residual deviance:   835542680000  on 15  degrees of freedom
## AIC: 586.1387398065
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Wampler4.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
##
## Terms added sequentially (first to last)
##
##
##                      Df       Deviance Resid. Df     Resid. Dev
## NULL                                         20 19649859888117
## poly(x, 5, raw = T)  5 18814317208117        15   835542680000
```

Using Wampler5.dat file. The file was cleaned up removing the header informatin and saving as Wampler5.txt.

The data file constist of 2 variables and 21 observations.

The first variable is the response variable (y) and the second variable is the predictor variable (x).

Procedure: Linear Least Squares Regression, 5th order poly

Expected results:

Parameter|**Estimate**|**Standard Deviation of Estimate** ———|———|——- **intercept**| 1.00000000000000 | 21523262.4678170
X | 1.00000000000000 | 23635517.3469681
X^2 | 1.00000000000000 | 7793435.24331583
X^3 | 1.00000000000000 | 1014755.07550350
X^4 | 1.00000000000000 | 56456.6512170752
X^5 | 1.00000000000000 | 1123.24854679312

Residual	.
Standard Deviation	23601450.2379268

Residual	.
R-Squared	0.224668921574940E-02

Certified Analysis of Variance Table

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Statistic
Regression	5	18814317208116.7	3762863441623.33	0.675524458240122E-02
Residual	15	0.835542680000000E+16	557028453333333	

```
Wampler5 <- read.table(file=paste0(path,"Wampler5.txt"),
                       header=TRUE)
Wampler5.glm <- glm(y ~ poly(x,5,raw=T),
                    data = Wampler5)
summary(Wampler5.glm)
```

```
##
## Call:
## glm(formula = y ~ poly(x, 5, raw = T), data = Wampler5)
##
## Deviance Residuals:
##       Min         1Q     Median         3Q        Max
## -20480000  -20480000    7590000   20480000   25230000
##
## Coefficients:
##                               Estimate        Std. Error t value Pr(>|t|)
## (Intercept)          1.00000042671e+00 2.15232624678e+07 0.00000  1.00000
## poly(x, 5, raw = T)1 9.99999184393e-01 2.36355173470e+07 0.00000  1.00000
## poly(x, 5, raw = T)2 1.00000028970e+00 7.79343524332e+06 0.00000  1.00000
## poly(x, 5, raw = T)3 9.99999962086e-01 1.01475507550e+06 0.00000  1.00000
## poly(x, 5, raw = T)4 1.00000000208e+00 5.64566512171e+04 0.00002  0.99999
## poly(x, 5, raw = T)5 9.99999999960e-01 1.12324854679e+03 0.00089  0.99930
##
## (Dispersion parameter for gaussian family taken to be 557028453333333)
##
##     Null deviance: 8374241117208117  on 20  degrees of freedom
## Residual deviance: 8355426800000000  on 15  degrees of freedom
## AIC: 779.555887618
##
## Number of Fisher Scoring iterations: 2
```

```
anova(Wampler5.glm)
```

```
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: y
```

```
##
## Terms added sequentially (first to last)
##
##
##                      Df       Deviance Resid. Df       Resid. Dev
## NULL                                          20 8374241117208117
## poly(x, 5, raw = T)  5 18814317208117        15 8355426800000000
```