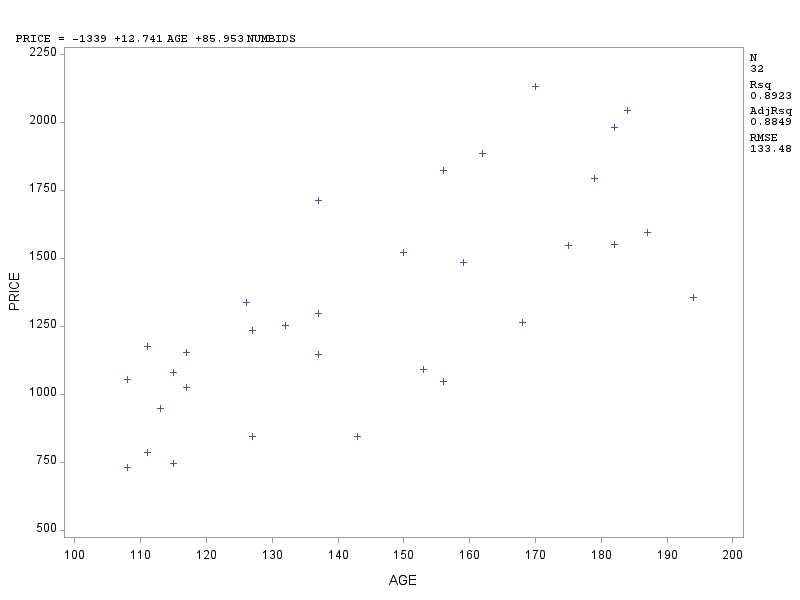
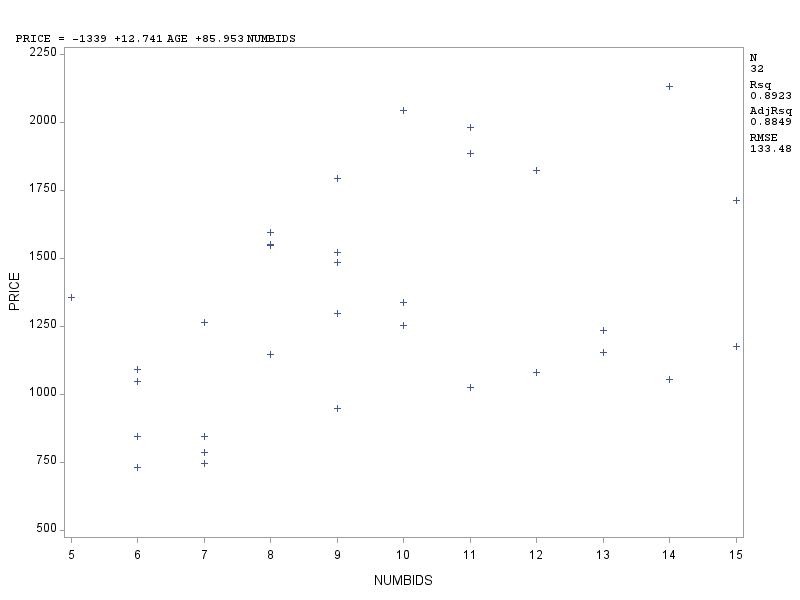
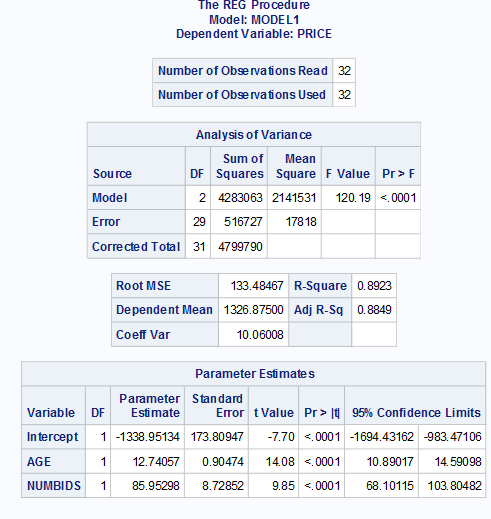
**Code for Chapter 4**

**Example 4.1 -4.4**







**data** gfclock;

input AGE NUMBIDS PRICE AGE\_BID;

datalines;

127 13 1235 1651

115 12 1080 1380

127 7 845 889

150 9 1522 1350

156 6 1047 936

182 11 1979 2002

156 12 1822 1872

132 10 1253 1320

137 9 1297 1233

113 9 946 1017

137 15 1713 2055

117 11 1024 1287

137 8 1147 1096

153 6 1092 918

117 13 1152 1521

126 10 1336 1260

170 14 2131 2380

182 8 1550 1456

162 11 1884 1782

184 10 2041 1840

143 6 845 858

159 9 1483 1431

108 14 1055 1512

175 8 1545 1400

108 6 729 648

179 9 1792 1611

111 15 1175 1665

187 8 1593 1496

111 7 785 777

115 7 744 805

194 5 1356 970

168 7 1262 1176

;

**run**;

/\* or import data by the following code and change the address where you store the sas data set

data gfclock;

set " ";

run;\*/

**proc print** data=gfclock;

**run**;

**proc reg** data=gfclock;

model PRICE = AGE NUMBIDS/CLB CLI CLM ALPHA=**0.05**; /\* construct regression model, regress y on x’s by format: y=x1 x2 (you should input the name of y or x variables corresponding to your imported data on SAS).\*/

plot PRICE\*AGE PRICE\*NUMBIDS/PRED;

**RUN**;

/\*Important informaiton can be found in this website: http://www.ats.ucla.edu/stat/sas/library/SASReg\_mf.htm\*/

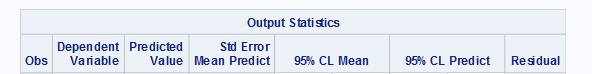
/\* CLB gives us the confidence interval for our parameters.\*/

/\* CLM gives us the confidence interval for the expected values of y \*/

/\* CLI gives us the prediction interval for the values of individual y \*/

/\* ALPHA=0.05 set up the corresponding confidence level 95% you can change it to ALPHA=0.1 to get the 90% intervals. \*/

**Example 4.5 ( Estimations and Predictions)**





**data** gfclock;

input AGE NUMBIDS PRICE AGE\_BID;

datalines;

127 13 1235 1651

115 12 1080 1380

127 7 845 889

150 9 1522 1350

156 6 1047 936

182 11 1979 2002

156 12 1822 1872

132 10 1253 1320

137 9 1297 1233

113 9 946 1017

137 15 1713 2055

117 11 1024 1287

137 8 1147 1096

153 6 1092 918

117 13 1152 1521

126 10 1336 1260

170 14 2131 2380

182 8 1550 1456

162 11 1884 1782

184 10 2041 1840

143 6 845 858

159 9 1483 1431

108 14 1055 1512

175 8 1545 1400

108 6 729 648

179 9 1792 1611

111 15 1175 1665

187 8 1593 1496

111 7 785 777

115 7 744 805

194 5 1356 970

168 7 1262 1176

150 10 . .

;

**run**;

**proc print** data=gfclock;

**run**;

**proc reg** data=gfclock;

model PRICE = AGE NUMBIDS/CLB CLI CLM ALPHA=**0.05**;

plot PRICE\*AGE PRICE\*NUMBIDS/PRED;

**RUN**;

/\* construct confidence interval and prediction interval for age=150, numbids=10. \*/

**data** a;

AGE=**150**;

NUMBIDS=**10**;

output;

**data** aa;

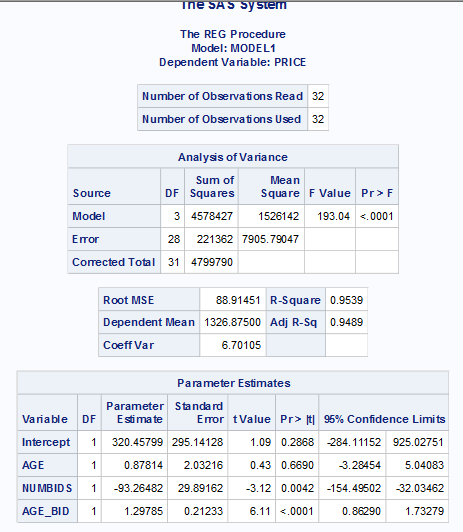
set gfclock a;/\*combine gfclock with new observation\*/

**proc** **reg** data=aa;

model PRICE = AGE NUMBIDS/clm cli;

**run**;

**Example 4.6 ( Interaction Model)**



**data** gfclock;

input AGE NUMBIDS PRICE AGE\_BID;

datalines;

127 13 1235 1651

115 12 1080 1380

127 7 845 889

150 9 1522 1350

156 6 1047 936

182 11 1979 2002

156 12 1822 1872

132 10 1253 1320

137 9 1297 1233

113 9 946 1017

137 15 1713 2055

117 11 1024 1287

137 8 1147 1096

153 6 1092 918

117 13 1152 1521

126 10 1336 1260

170 14 2131 2380

182 8 1550 1456

162 11 1884 1782

184 10 2041 1840

143 6 845 858

159 9 1483 1431

108 14 1055 1512

175 8 1545 1400

108 6 729 648

179 9 1792 1611

111 15 1175 1665

187 8 1593 1496

111 7 785 777

115 7 744 805

194 5 1356 970

168 7 1262 1176

;

**run**;

**proc print** data=gfclock;

**run**;

**proc reg** data=gfclock;

model PRICE = AGE NUMBIDS AGE\_BID/CLB CLI CLM ALPHA=**0.05**;

plot PRICE\*AGE PRICE\*NUMBIDS/PRED;

**RUN**;

/\*when iteration term is not contained in the dataset, we can create one by using glm procedure with the following code:

data a;

proc glm data=a;

model PRICE=AGE NUMBIDS AGE\*NUMBIDS;

run;\*/

**Example 4.7 ( Quadratic Model)**

**data** aerobic;

input SUBJECT IGG MAXOXY;

datalines;

1 881 34.6

2 1290 45.0

3 2147 62.3

4 1909 58.9

5 1282 42.5

6 1530 44.3

7 2067 67.9

8 1982 58.5

9 1019 35.6

10 1651 49.6

11 752 33.0

12 1687 52.0

13 1782 61.4

14 1529 50.2

15 969 34.1

16 1660 52.5

17 2121 69.9

18 1382 38.8

19 1714 50.6

20 1959 69.4

21 1158 37.4

22 965 35.1

23 1456 43.0

24 1273 44.1

25 1418 49.8

26 1743 54.4

27 1997 68.5

28 2177 69.5

29 1965 63.0

30 1264 43.2

;

**run**;

**proc print** data=aerobic;

**run**;

**proc gplot** data=aerobic;

plot IGG\*MAXOXY;

**run**;

**data** aerobic1; /\*Creating a new dataset containing the square of MAXOXY, "SQM".\*/

set aerobic;

SQM = MAXOXY\*\***2**;

**run**;

**proc print** data=aerobic1 (obs=**3**);

**run**;

**proc** **reg** data=aerobic1;

model IGG=MAXOXY SQM/CLB CLI CLM ALPHA=**0.01 p r**;

plot IGG\*MAXOXY IGG\*SQM/PRED;

**RUN**;

/\* or you can directly use the following glm procedure without creating another dataset:

ods graphics on;

proc glm data=aerobic;

model IGG=MAXOXY MAXOXY\*MAXOXY;

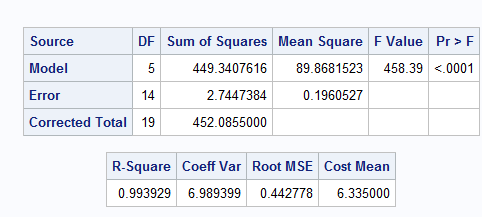
run;

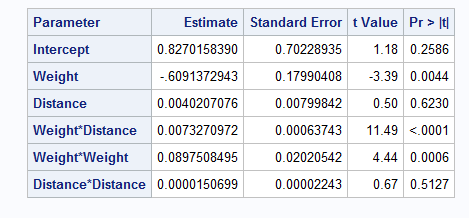
ods graphics off;

\*/

**Example 4.8 and 4.11**

**Completed model**

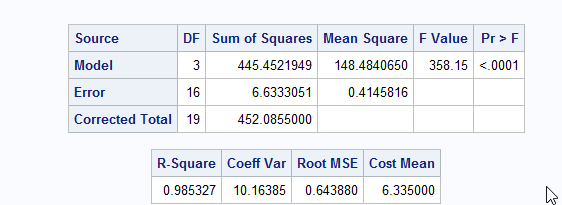








**Reduced model**



**data** express;

set C:\Users\yusha\Desktop\fall 2017\Data sets\SAS\SAS\Exercises&Examples\EXPRESS.sas7bdat';

**run**;

**proc** **glm** data=express;/\*complete model\*/

model cost=weight distance weight\*distance weight\*weight distance\*distance;

**run**;

**data** a;

weight=**5**;

distance=**100**;

output;

**data** aa;

set express a;

**proc** **glm** data=aa;

model cost=weight distance weight\*distance weight\*weight distance\*distance/p cli;

**run**;

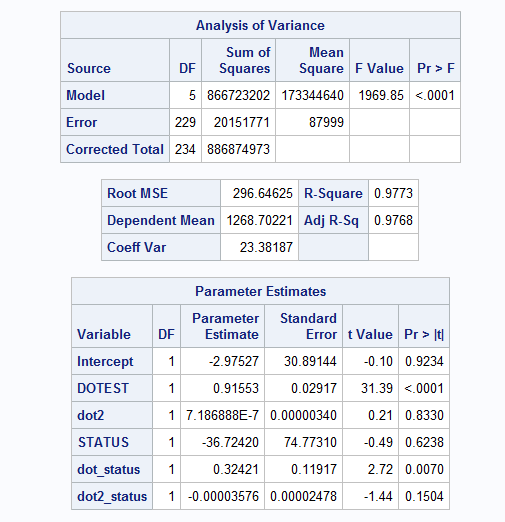
**proc** **glm** data=express;/\*reduced model\*/

model cost=weight distance weight\*distance;

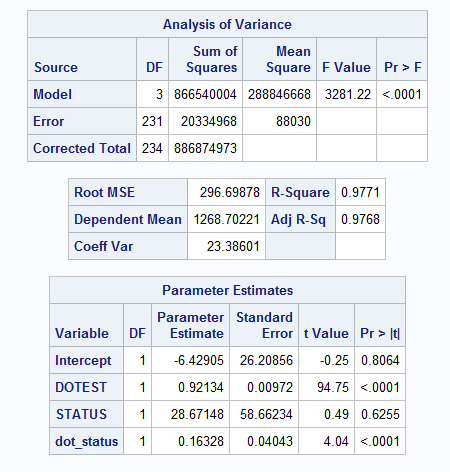
**run**;

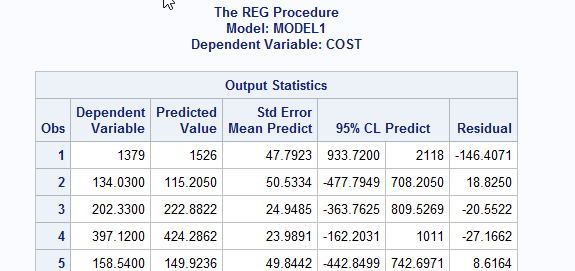
**4.14 A complete example**

**For full model:**

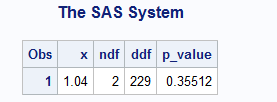


**For reduced model**





**For partial F test(getting p-value)**



**data** a;

set 'C:\Users\yusha\Desktop\fall 2017\Data sets\SAS\SAS\Exercises&Examples\FLAG.sas7bdat';

**run**;

**proc** **print** data=a;

**run**;

**data** b;

set a;

dot2=dotest\*dotest;

dot\_status=dotest\*status;

dot2\_status=dotest\*dotest\*status;

**run**;

**proc** **reg** data=b;

model cost=dotest dot2 status dot\_status dot2\_status;/\*complete model\*/

**run**;

**proc** **reg** data=b;

model cost=dotest status dot\_status/cli;/\*reduced model\*/

**run**;

/\*find p-value of partial F test\*/

**data** one;

input x ndf ddf;

p\_value=**1**-PROBF(x, ndf, ddf);/\*input cdf F function \*/

cards;

1.04 2 229

**run**;

**proc** **print** data=one;/\*print out p-value of partial F-test with F-statistic=1.04,df of numerator=2, df of denominator=229\*/

**run**;