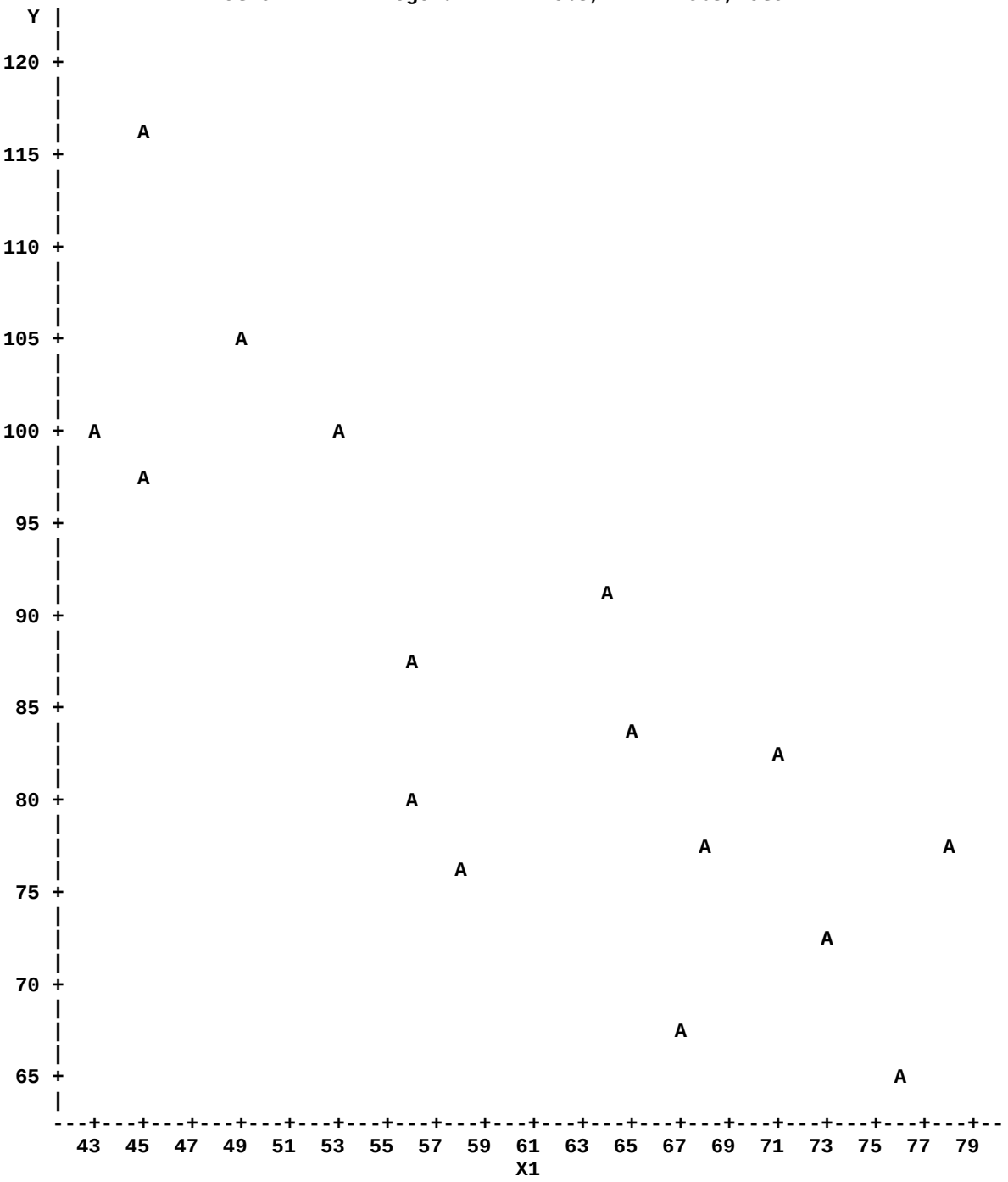
```
1          ************************************************************;
2          *** EXST7034 Homework Example 1                          ***;
3          *** Problem from Neter, Wasserman & Kuttner 1989, #11.16   ***;
4          ************************************************************;
5          OPTIONS LS=82 PS=61 NOCENTER NODATE NONUMBER;
6
7          DATA ONE; INFILE CARDS MISSOVER;
8               TITLE1 'EXST7034 - Homework Example NWK 11.16 (based on # 1.27) :
8        ! Muscle mass';
9            *     LABEL X1  = 'Age (years)';
10           *     LABEL X2  = '(X1-60)*Indicator of Age';
11           *     LABEL X3  = 'Indicator of Age > 60';
12           *     LABEL Y = 'Muscle mass';
13              INPUT Y X1;
14                X3 = 0; IF X1 GE 60 THEN X3 = 1;
15                X2 = (X1-60)*X3;
16                X1X3 = X1*X3;
17                LogY = LOG(Y);
18         CARDS;
NOTE: The data set WORK.ONE has 16 observations and 6 variables.
NOTE: DATA statement used:
      real time           0.06 seconds
      cpu time            0.06 seconds
18       !        RUN;
35         ;
36         PROC SORT DATA=ONE; BY X1; RUN;
NOTE: There were 16 observations read from the data set WORK.ONE.
NOTE: The data set WORK.ONE has 16 observations and 6 variables.
NOTE: PROCEDURE SORT used:
      real time           0.05 seconds
      cpu time            0.04 seconds
37         PROC PRINT DATA=ONE; RUN;
NOTE: There were 16 observations read from the data set WORK.ONE.
NOTE: The PROCEDURE PRINT printed page 1.
NOTE: PROCEDURE PRINT used:
      real time           0.04 seconds
      cpu time            0.04 seconds
```

EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass

Obs	Y	X1	X3	X2	X1X3	LogY
1	100	43	0	0	0	4.60517
2	116	45	0	0	0	4.75359
3	97	45	0	0	0	4.57471
4	105	49	0	0	0	4.65396
5	100	53	0	0	0	4.60517
6	87	56	0	0	0	4.46591
7	80	56	0	0	0	4.38203
8	76	58	0	0	0	4.33073
9	91	64	1	4	64	4.51086
10	84	65	1	5	65	4.43082
11	68	67	1	7	67	4.21951
12	78	68	1	8	68	4.35671
13	82	71	1	11	71	4.40672
14	73	73	1	13	73	4.29046
15	65	76	1	16	76	4.17439
16	77	78	1	18	78	4.34381

```
38         proc plot data=one; plot Y*X1; run;
NOTE: There were 16 observations read from the data set WORK.ONE.
NOTE: The PROCEDURE PLOT printed page 2.
NOTE: PROCEDURE PLOT used:
      real time           0.00 seconds
      cpu time            0.00 seconds
```

EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass




```
39         PROC REG; TITLE2 'SLR'; MODEL  Y = X1; RUN;
NOTE: 16 observations read.
NOTE: 16 observations used in computations.
NOTE: The PROCEDURE REG printed page 3.
NOTE: PROCEDURE REG used:
      real time           0.08 seconds
      cpu time            0.08 seconds
```

Simple linear regression for comparison.

```
EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
SLR

The REG Procedure
Model: MODEL1
Dependent Variable: Y

                          Analysis of Variance
                                 Sum of           Mean
Source                    DF       Squares         Square    F Value    Pr > F
Model                      1     2059.78145     2059.78145      29.59    <.0001
Error                     14      974.65605       69.61829
Corrected Total           15     3034.43750

Root MSE              8.34376    R-Square     0.6788
Dependent Mean       86.18750    Adj R-Sq     0.6559
Coeff Var             9.68094

                     Parameter Estimates
                    Parameter       Standard
Variable     DF      Estimate          Error    t Value    Pr > |t|
Intercept     1     148.05068       11.56292      12.80      <.0001
X1            1      -1.02359        0.18818      -5.44      <.0001
```

Before using a segmented curve you might try alternative models that could be easier to interpret.

```
40         PROC REG; TITLE2 'Exponential decay'; MODEL  LogY = X1; RUN;
NOTE: 16 observations read.
NOTE: 16 observations used in computations.
NOTE: The PROCEDURE REG printed page 4.
NOTE: PROCEDURE REG used:
      real time           0.07 seconds
      cpu time            0.07 seconds


EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
Exponential decay

The REG Procedure
Model: MODEL1
Dependent Variable: LogY

                          Analysis of Variance
                                 Sum of           Mean
Source                    DF       Squares         Square    F Value    Pr > F
Model                      1       0.26879        0.26879      29.47    <.0001
Error                     14       0.12769        0.00912
Corrected Total           15       0.39647

Root MSE              0.09550    R-Square     0.6779
Dependent Mean        4.44403    Adj R-Sq     0.6549
Coeff Var             2.14897

                     Parameter Estimates
                    Parameter       Standard
Variable     DF      Estimate          Error    t Value    Pr > |t|
Intercept     1       5.15072        0.13235      38.92      <.0001
X1            1      -0.01169        0.00215      -5.43      <.0001
```

Basic Analysis of covariance with two intercepts and two slopes. Note that the parameter estimates are DIFFERENCES and that we do not have TYPE I tests.

```
41         PROC REG; TITLE2 'ANACOV with SAS REG procedure';
42            MODEL  Y = X1 X3 X1X3; RUN;
NOTE: 16 observations read.
NOTE: 16 observations used in computations.
NOTE: The PROCEDURE REG printed page 5.
NOTE: PROCEDURE REG used:
      real time           0.05 seconds
      cpu time            0.04 seconds


EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
ANACOV with SAS REG procedure

The REG Procedure
Model: MODEL1
Dependent Variable: Y
                            Analysis of Variance
                                   Sum of           Mean
Source                     DF        Squares         Square    F Value    Pr > F
Model                       3     2226.68272      742.22757      11.03    0.0009
Error                      12      807.75478       67.31290
Corrected Total            15     3034.43750

Root MSE              8.20444    R-Square     0.7338
Dependent Mean       86.18750    Adj R-Sq     0.6673
Coeff Var             9.51930

                       Parameter Estimates
                     Parameter       Standard
Variable     DF       Estimate          Error    t Value    Pr > |t|
Intercept     1      186.30233       26.86367       6.94      <.0001
X1            1       -1.80103        0.52754      -3.41      0.0051
X3            1      -42.63066       50.40227      -0.85      0.4142
X1X3          1        0.85553        0.80320       1.07      0.3078
```

We can use PROC GLM to get both TYPE I tests and TYPE II / TYPE III tests. Note that the parameter estimates are STILL DIFFERENCES (Is this what we want?) and that the classes statement has assigned a 1 to my 0 and a 0 to my 1.
Below the Analysis of Covariance is done as both a means models and effects models
Results for the effects model in GLM were as follows:

```
43         PROC GLM; CLASSES X3; TITLE2 'ANACOV with SAS REG procedure';
44            MODEL  Y = X1 X3 X1*X3 / solution; RUN;


NOTE: The PROCEDURE GLM printed pages 6-7.
NOTE: PROCEDURE GLM used:
      real time           0.12 seconds
      cpu time            0.09 seconds


EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
ANACOV with SAS REG procedure

The GLM Procedure

Class Level Information
Class         Levels    Values
X3                 2    0 1
Number of observations    16
```

```
Dependent Variable: Y
                                      Sum of
Source                     DF        Squares    Mean Square   F Value   Pr > F
Model                       3    2226.682718     742.227573     11.03   0.0009
Error                      12     807.754782      67.312899
Corrected Total            15    3034.437500

R-Square     Coeff Var      Root MSE        Y Mean
0.733804      9.519297      8.204444      86.18750

Source                     DF      Type I SS    Mean Square   F Value   Pr > F
X1                          1    2059.781452    2059.781452     30.60   0.0001
X3                          1      90.530865      90.530865      1.34   0.2687
X1*X3                       1      76.370401      76.370401      1.13   0.3078

Source                     DF    Type III SS    Mean Square   F Value   Pr > F
X1                          1    787.0932928    787.0932928     11.69   0.0051
X3                          1     48.1550950     48.1550950      0.72   0.4142
X1*X3                       1     76.3704012     76.3704012      1.13   0.3078

                                                 Standard
Parameter            Estimate                       Error    t Value    Pr > |t|
Intercept        143.6716621 B     42.64659609       3.37      0.0056
X1                -0.9455041 B      0.60566310      -1.56      0.1445
X3         0      42.6306635 B     50.40227224       0.85      0.4142
X3         1       0.0000000 B       .                .         .
X1*X3      0      -0.8555295 B      0.80319612      -1.07      0.3078
X1*X3      1       0.0000000 B       .                .         .
NOTE: The X'X matrix has been found to be singular, and a generalized inverse was
      used to solve the normal equations.  Terms whose estimates are followed by
      the letter 'B' are not uniquely estimable.
```

The means model was run as follows:

```
45         PROC GLM DATA=ONE; CLASSES X3; TITLE2 'AnCova with SAS GLM and NOINT';
46            MODEL  Y = X3 X1*X3 / solution NOINT; RUN;
NOTE: Due to the presence of CLASS variables, an intercept is implicitly fitted.
      R-Square has been corrected for the mean.
NOTE: The PROCEDURE GLM printed pages 8-9.
NOTE: PROCEDURE GLM used:
      real time           0.08 seconds
      cpu time            0.08 seconds

The GLM Procedure
Dependent Variable: Y
                                      Sum of
Source                     DF        Squares    Mean Square   F Value   Pr > F
Model                       4    121079.2452     30269.8113    449.69   <.0001
Error                      12       807.7548        67.3129
Uncorrected Total          16    121887.0000

R-Square     Coeff Var      Root MSE        Y Mean
0.733804      9.519297      8.204444      86.18750

Source                     DF      Type I SS    Mean Square   F Value   Pr > F
X3                          2    120130.6250     60065.3125    892.33   <.0001
X1*X3                       2       948.6202       474.3101      7.05   0.0095

Source                     DF    Type III SS    Mean Square   F Value   Pr > F
X3                          2    4001.419587    2000.709793     29.72   <.0001
X1*X3                       2     948.620218     474.310109      7.05   0.0095

                                                 Standard
Parameter            Estimate                       Error    t Value    Pr > |t|
X3         0      186.3023256     26.86367229       6.94      <.0001
X3         1      143.6716621     42.64659609       3.37      0.0056
X1*X3      0       -1.8010336      0.52753789      -3.41      0.0051
X1*X3      1       -0.9455041      0.60566310      -1.56      0.1445
```

These models can also be run in PROC MIXED, but there are no sums of squares unless requested.

Analysis of Covariance – Effects model in PROC MIXED

```
47         proc mixed DATA=ONE; classes X3; title2 'AnCova using PROC MIXED';
48           model Y = X1 X3 X1X3 / htype=1 3 DDFM=Satterthwaite solution;
49         run;
NOTE: The PROCEDURE MIXED printed pages 10-11.
NOTE: PROCEDURE MIXED used:
      real time           0.05 seconds
      cpu time            0.05 seconds

EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
AnCova using PROC MIXED

The Mixed Procedure
                Model Information
Data Set                      WORK.ONE
Dependent Variable            Y
Covariance Structure          Diagonal
Estimation Method             REML
Residual Variance Method      Profile
Fixed Effects SE Method       Model-Based
Degrees of Freedom Method     Residual

Class Level Information
Class    Levels    Values
X3            2    0 1

            Dimensions
Covariance Parameters              1
Columns in X                       6
Columns in Z                       0
Subjects                           1
Max Obs Per Subject               16
Observations Used                 16
Observations Not Used              0
Total Observations                16

Covariance Parameter Estimates
Cov Parm     Estimate
Residual      67.3129

           Fit Statistics
-2 Res Log Likelihood            99.4
AIC (smaller is better)         101.4
AICC (smaller is better)        101.8
BIC (smaller is better)         101.9

Solution for Fixed Effects     Standard
Effect       X3    Estimate       Error      DF    t Value    Pr > |t|
Intercept            143.67     42.6466      12       3.37      0.0056
X1                  -0.9455      0.6057      12      -1.56      0.1445
X3           0      42.6307     50.4023      12       0.85      0.4142
X3           1            0           .       .          .           .
X1*X3        0      -0.8555      0.8032      12      -1.07      0.3078
X1*X3        1            0           .       .          .           .

Type 1 Tests of Fixed Effects
Effect     Num DF  Den DF    F Value    Pr > F
X1              1      12      30.60    0.0001
X3              1      12       1.34    0.2687
X1*X3           1      12       1.13    0.3078

        Type 3 Tests of Fixed Effects
Effect     Num DF  Den DF    F Value    Pr > F
X1              1      12      11.69    0.0051
X3              1      12       0.72    0.4142
X1*X3           1      12       1.13    0.3078
```

Analysis of Covariance – Means model in PROC MIXED

```
50         proc mixed DATA=ONE; classes X3; title2 'AnCova using PROC MIXED and NOINT';
51           model Y = X3 X1*X3 / htype=1 3 DDFM=Satterthwaite solution NOINT;
52         run;
NOTE: The PROCEDURE MIXED printed pages 12-13.
NOTE: PROCEDURE MIXED used:
      real time           0.05 seconds
      cpu time            0.05 seconds


EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
AnCova using PROC MIXED and NOINT

The Mixed Procedure
                  Model Information
Data Set                      WORK.ONE
Dependent Variable            Y
Covariance Structure          Diagonal
Estimation Method             REML
Residual Variance Method      Profile
Fixed Effects SE Method       Model-Based
Degrees of Freedom Method     Residual

Class Level Information
Class    Levels    Values
X3            2    0 1

Dimensions
Covariance Parameters             1
Columns in X                      4
Columns in Z                      0
Subjects                          1
Max Obs Per Subject              16
Observations Used                16
Observations Not Used             0
Total Observations               16

Covariance Parameter Estimates
Cov Parm     Estimate
Residual      67.3129

Fit Statistics
-2 Res Log Likelihood            99.4
AIC (smaller is better)         101.4
AICC (smaller is better)        101.8
BIC (smaller is better)         101.9

Solution for Fixed Effects
                            Standard
Effect    X3    Estimate       Error      DF    t Value    Pr > |t|
X3        0       186.30     26.8637      12       6.94      <.0001
X3        1       143.67     42.6466      12       3.37      0.0056
X1*X3     0      -1.8010      0.5275      12      -3.41      0.0051
X1*X3     1      -0.9455      0.6057      12      -1.56      0.1445

Type 1 Tests of Fixed Effects
Effect  Num DF  Den DF    F Value    Pr > F
X3           2      12     892.33    <.0001
X1*X3        2      12       7.05    0.0095

Type 3 Tests of Fixed Effects
Effect  Num DF  Den DF    F Value    Pr > F
X3           2      12      29.72    <.0001
X1*X3        2      12       7.05    0.0095
```

Piecewise regression.

```
45         PROC REG; VAR X1 X3; TITLE2 'Piecewise regression with SAS REG procedure';
46            MODEL  Y = X1 X2 X3 / XPX I P CLM CLI; ID X1 X2;
47            TEST X2=0, X3=0; RUN;

NOTE: 16 observations read.
NOTE: 16 observations used in computations.
48            OUTPUT OUT=RESIDS PREDICTED=P1 RESIDUAL=E1;
49         OPTIONS PS=40;
NOTE: The data set WORK.RESIDS has 16 observations and 8 variables.
NOTE: The PROCEDURE REG printed pages 8-11.
NOTE: PROCEDURE REG used:
      real time           0.11 seconds
      cpu time            0.11 seconds

EXST7034 - Homework Example NWK 11.16 (based on # 1.27) : Muscle mass
Piecewise regression with SAS REG procedure

The REG Procedure
Model: MODEL1
                      Model Crossproducts X'X X'Y Y'Y
Variable      Intercept            X1            X2            X3             Y
Intercept            16           967            82             8          1379
X1                  967         60409          5944           562         81331
X2                   82          5944          1024            82          6161
X3                    8           562            82             8           618
Y                  1379         81331          6161           618        121887

                  X'X Inverse, Parameter Estimates, and SSE
Variable      Intercept            X1            X2            X3             Y
Intercept  10.720930233  -0.209302326  0.2093023256  1.8372093023  186.30232558
X1         -0.209302326  0.0041343669  -0.004134367   -0.03875969  -1.801033592
X2         0.2093023256  -0.004134367  0.0095839582  -0.017098621  0.8555295045
X3         1.8372093023   -0.03875969  -0.017098621  1.1859197769  8.7011068162
Y          186.30232558  -1.801033592  0.8555295045  8.7011068162  807.75478247


Analysis of Variance                Sum of           Mean
Source                    DF       Squares         Square    F Value    Pr > F
Model                      3    2226.68272      742.22757      11.03    0.0009
Error                     12     807.75478       67.31290
Corrected Total           15    3034.43750

Root MSE              8.20444    R-Square     0.7338
Dependent Mean       86.18750    Adj R-Sq     0.6673
Coeff Var             9.51930

Parameter Estimates
                     Parameter       Standard
Variable     DF       Estimate          Error    t Value    Pr > |t|
Intercept     1      186.30233       26.86367       6.94      <.0001
X1            1       -1.80103        0.52754      -3.41      0.0051
X2            1        0.85553        0.80320       1.07      0.3078
X3            1        8.70111        8.93463       0.97      0.3493
```

Output Statistics

Obs	X1	X2	Dep Var Y	Predicted Value	Std Error Mean Predict	95% CL Mean	
1	43	0	100.0000	108.8579	4.9593	98.0525	119.6632
2	45	0	116.0000	105.2558	4.1496	96.2145	114.2971
3	45	0	97.0000	105.2558	4.1496	96.2145	114.2971
4	49	0	105.0000	98.0517	3.0247	91.4614	104.6420
5	53	0	100.0000	90.8475	3.1597	83.9631	97.7320
6	56	0	87.0000	85.4444	4.0564	76.6063	94.2825
7	56	0	80.0000	85.4444	4.0564	76.6063	94.2825
8	58	0	76.0000	81.8424	4.8529	71.2688	92.4160
9	64	4	91.0000	83.1594	4.7690	72.7687	93.5502
10	65	5	84.0000	82.2139	4.3040	72.8362	91.5916
11	67	7	68.0000	80.3229	3.5055	72.6850	87.9608
12	68	8	78.0000	79.3774	3.2049	72.3946	86.3602
13	71	11	82.0000	76.5409	2.9361	70.1437	82.9380
14	73	13	73.0000	74.6499	3.3449	67.3620	81.9377
15	76	16	65.0000	71.8134	4.5324	61.9382	81.6885
16	78	18	77.0000	69.9223	5.5179	57.9000	81.9447

Output Statistics

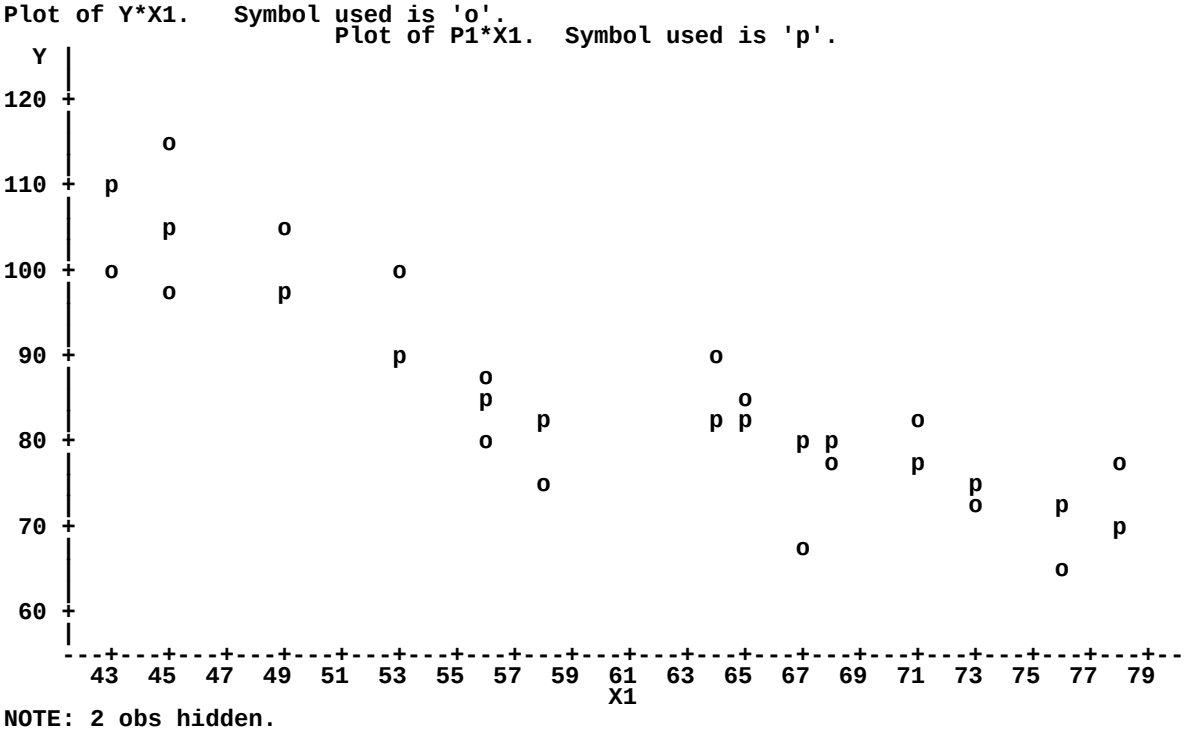
Obs	X1	X2	95% CL Predict		Residual
1	43	0	87.9700	129.7458	-8.8579
2	45	0	85.2235	125.2882	10.7442
3	45	0	85.2235	125.2882	-8.2558
4	49	0	78.9996	117.1038	6.9483
5	53	0	71.6917	110.0034	9.1525
6	56	0	65.5030	105.3859	1.5556
7	56	0	65.5030	105.3859	-5.4444
8	58	0	61.0734	102.6114	-5.8424
9	64	4	62.4829	103.8359	7.8406
10	65	5	62.0275	102.4003	1.7861
11	67	7	60.8836	99.7622	-12.3229
12	68	8	60.1860	98.5688	-1.3774
13	71	11	57.5548	95.5270	5.4591
14	73	13	55.3454	93.9543	-1.6499
15	76	16	51.3911	92.2356	-6.8134
16	78	18	48.3797	91.4650	7.0777

Sum of Residuals 0
Sum of Squared Residuals 807.75478
Predicted Residual SS (PRESS) 1525.85950

Test 1 Results for Dependent Variable Y

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	83.45063	1.24	0.3240
Denominator	12	67.31290		

```
50         PROC PLOT DATA=RESIDS;  PLOT Y*X1='o' P1*X1='p'/OVERLAY; RUN;
NOTE: There were 16 observations read from the data set WORK.RESIDS.
NOTE: The PROCEDURE PLOT printed page 12.
NOTE: PROCEDURE PLOT used:
      real time           0.01 seconds
      cpu time            0.01 seconds
51         PROC PLOT DATA=RESIDS;  PLOT E1*X1='x' / VREF=0; RUN;
52         OPTIONS PS=60;
NOTE: There were 16 observations read from the data set WORK.RESIDS.
NOTE: The PROCEDURE PLOT printed page 13.
NOTE: PROCEDURE PLOT used:
      real time           0.01 seconds
      cpu time            0.01 seconds
```



NOTE: 2 obs hidden.