

```

1 *****;
2 *** Finish times in a 10 K race ***;
3 *** Data taken from various sites on the ***;
4 *** internet reporting race results ***;
5 *****;
6 options ps=256 ls=80 nocenter nodate nonumber;
7
8 data one; length hometown $ 23 sex $ 3;
9 infile "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\04a-MReg-Polynomial Marathons.DAT" missover;
10 TITLE1 'EXST7015: Marathon Footrace Example';
11 input Marathon $ Age Sex $ TIME HomeTown $ 35-57;
12 if age eq 99 then age = .;
13 *(apparently 99 represents missing for 5 PA race participants);
14 *-----1-----2-----3-----4-----5-----6;
15 cards;

```

NOTE: The infile "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\04a-MReg-Polynomial

Marathons.DAT" is:
 File Name=C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\04a-MReg-Polynomial
 Marathons.DAT,
 RECFM=V,LRECL=256

NOTE: 6150 records were read from the infile
 "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\04a-MReg-Polynomial
 Marathons.DAT".

The minimum record length was 29.
 The maximum record length was 57.

NOTE: The data set WORK.ONE has 6150 observations and 5 variables.

NOTE: DATA statement used:

real time 0.14 seconds
 cpu time 0.14 seconds

```

15 ! run;
16 ;
17
18 data two; set one; if marathon = 'VT052002';
19 /*
20 proc print data=two WIDTH=UNIFORM; TITLE2 'Raw data print';
21 var Marathon Age Sex TIME HomeTown; run;
22 */
23

```

NOTE: There were 6150 observations read from the data set WORK.ONE.

NOTE: The data set WORK.TWO has 1490 observations and 5 variables.

NOTE: DATA statement used:

real time 0.04 seconds
 cpu time 0.04 seconds

```

24 proc sort data=two; by sex; RUN;

```

NOTE: There were 1490 observations read from the data set WORK.TWO.

NOTE: The data set WORK.TWO has 1490 observations and 5 variables.

NOTE: PROCEDURE SORT used:

real time 0.07 seconds
 cpu time 0.07 seconds

```

25 proc means data=two; by sex; var age time; run;

```

NOTE: There were 1490 observations read from the data set WORK.TWO.

NOTE: The PROCEDURE MEANS printed page 1.

NOTE: PROCEDURE MEANS used:

real time 0.05 seconds
 cpu time 0.05 seconds

EXST7015: Marathon Footrace Example

The MEANS Procedure

sex=F					
Variable	N	Mean	Std Dev	Minimum	Maximum
Age	527	34.4686907	8.6562175	16.0000000	65.0000000
TIME	527	246.7373435	24.4075610	161.3300000	288.4700000

sex=M					
Variable	N	Mean	Std Dev	Minimum	Maximum
Age	963	39.9075805	9.3669019	18.0000000	69.0000000
TIME	963	230.7960332	29.3895704	146.4500000	288.3500000

```

27      options ls=132 ps=65;
27      !      proc plot data=two; by sex; plot time*age;
28      TITLE2 'Scatter plot'; run;
29      options ps=256 ls=80;
    
```

NOTE: There were 1490 observations read from the data set WORK.TWO.

NOTE: The PROCEDURE PLOT printed pages 2-3.

NOTE: PROCEDURE PLOT used:

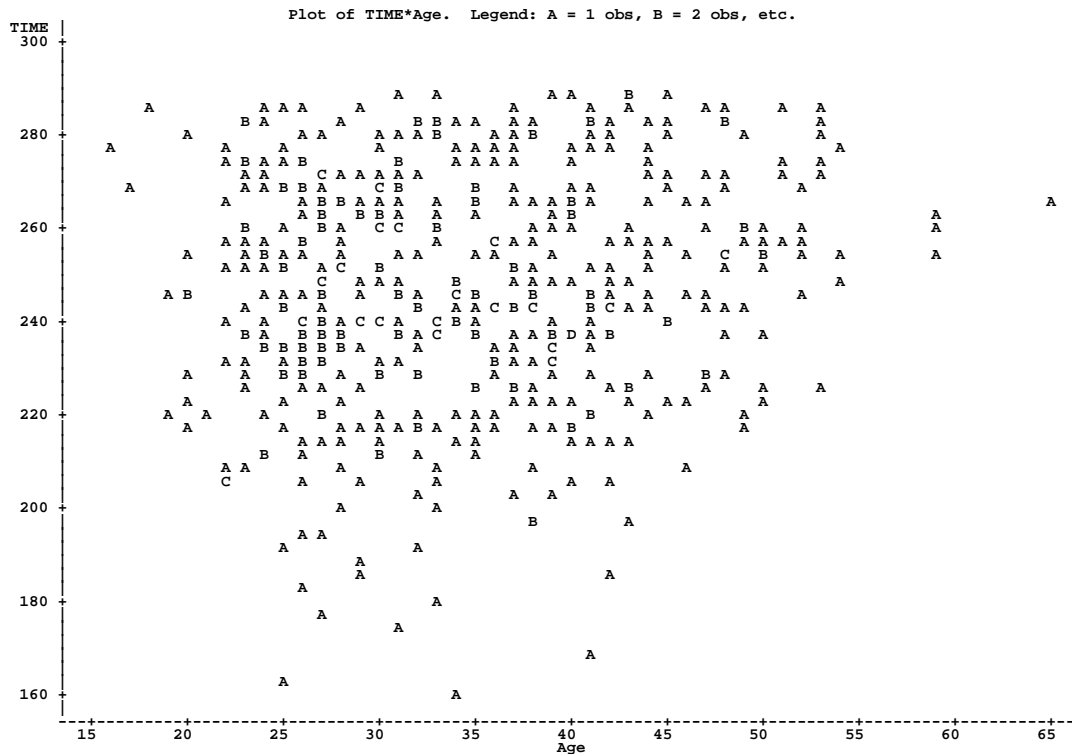
```

real time      0.01 seconds
cpu time       0.01 seconds
    
```

EXST7015: Marathon Footrace Example

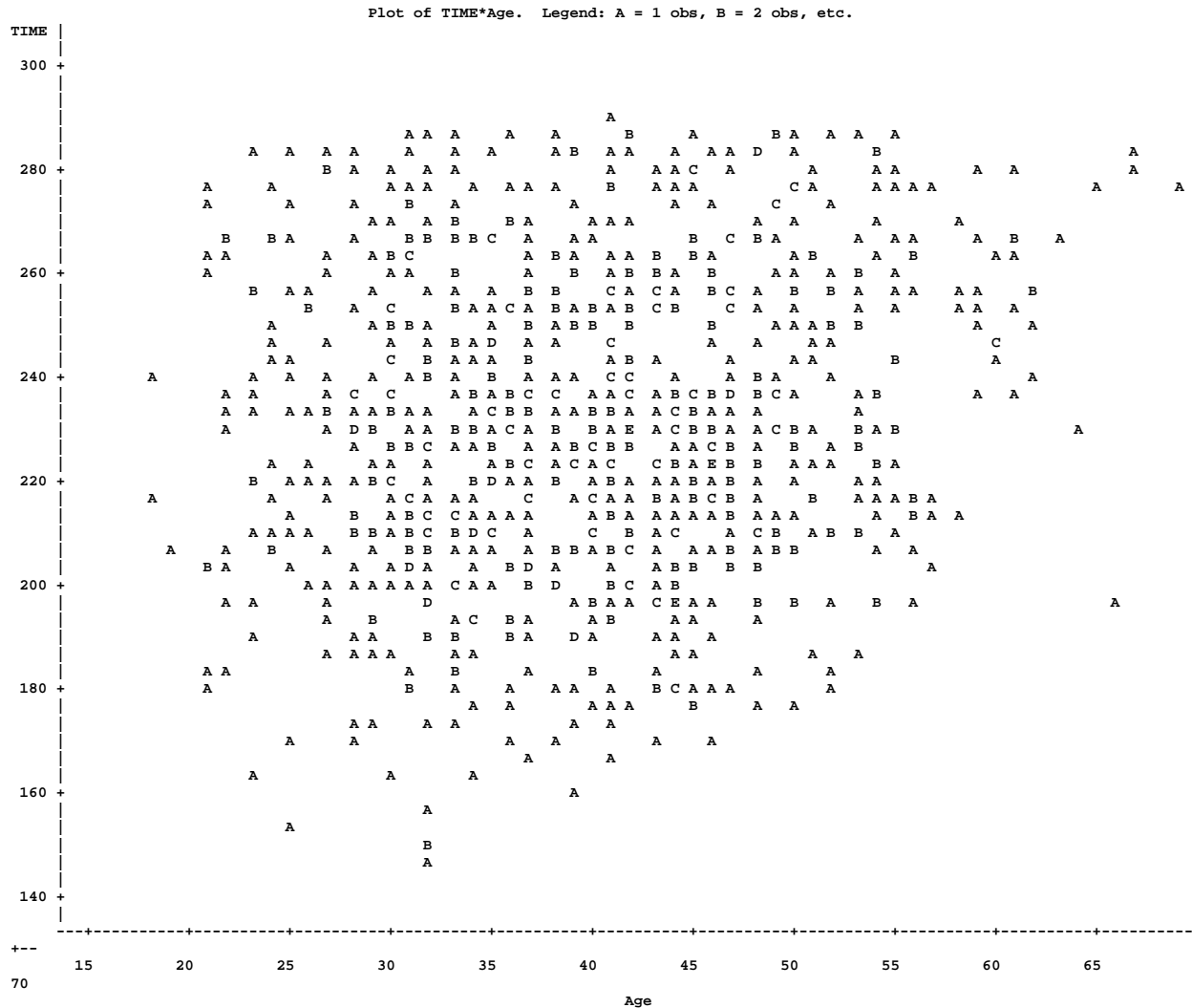
Scatter plot

sex=F



EXST7015: Marathon Footrace Example
Scatter plot

sex=M



```
31 proc mixed data=two CL; BY sex; TITLE2 'Quadratic model - separate by
sex';
32 model time= age age*age / htype=1 3 DDFM=Satterthwaite solution;
33 run;
```

NOTE: The PROCEDURE MIXED printed pages 4-5.

NOTE: PROCEDURE MIXED used:

```
real time          0.07 seconds
cpu time           0.07 seconds
```

EXST7015: Marathon Footrace Example
Quadratic model - separate by sex

sex=F

The Mixed Procedure

	Model Information
Data Set	WORK.TWO
Dependent Variable	TIME
Covariance Structure	Diagonal
Estimation Method	REML

Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual
Dimensions	
Covariance Parameters	1
Columns in X	3
Columns in Z	0
Subjects	1
Max Obs Per Subject	527
Observations Used	527
Observations Not Used	0
Total Observations	527

Covariance Parameter Estimates				
Cov Parm	Estimate	Alpha	Lower	Upper
Residual	581.96	0.05	517.46	659.40

Fit Statistics	
-2 Res Log Likelihood	4855.1
AIC (smaller is better)	4857.1
AICC (smaller is better)	4857.1
BIC (smaller is better)	4861.4

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	270.94	15.3541	524	17.65	<.0001
Age	-1.7668	0.8679	524	-2.04	0.0423
Age*Age	0.02906	0.01179	524	2.46	0.0140

Type 1 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Age	1	524	8.37	0.0040
Age*Age	1	524	6.08	0.0140

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Age	1	524	4.14	0.0423
Age*Age	1	524	6.08	0.0140

EXST7015: Marathon Footrace Example
 Quadratic model - separate by sex

sex=M

The Mixed Procedure

Model Information	
Data Set	WORK.TWO
Dependent Variable	TIME
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

Dimensions	
Covariance Parameters	1

Columns in X	3
Columns in Z	0
Subjects	1
Max Obs Per Subject	963
Observations Used	963
Observations Not Used	0
Total Observations	963

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
Residual	832.54	0.05	762.80	912.34

Fit Statistics

-2 Res Log Likelihood	9214.3
AIC (smaller is better)	9216.3
AICC (smaller is better)	9216.3
BIC (smaller is better)	9221.2

Solution for Fixed Effects

Effect	Estimate	Standard		DF	t Value	Pr > t
		Error				
Intercept	265.60	13.9782		960	19.00	<.0001
Age	-2.3003	0.6995		960	-3.29	0.0010
Age*Age	0.03392	0.008488		960	4.00	<.0001

Type 1 Tests of Fixed Effects

Effect	Num	Den	F Value	Pr > F
	DF	DF		
Age	1	960	22.10	<.0001
Age*Age	1	960	15.97	<.0001

Type 3 Tests of Fixed Effects

Effect	Num	Den	F Value	Pr > F
	DF	DF		
Age	1	960	10.81	0.0010
Age*Age	1	960	15.97	<.0001

```

35      proc mixed data=two CL; classes sex;
35          TITLE2 'Quadratic model - version for testing';
36          model time= age age*age sex sex*age sex*age*age
36              / htype=1 3 DDFM=Satterthwaite solution;
37      run;

```

NOTE: The PROCEDURE MIXED printed page 6.

NOTE: PROCEDURE MIXED used:

real time	0.06 seconds
cpu time	0.06 seconds

EXST7015: Marathon Footrace Example
Quadratic model - version for testing

The Mixed Procedure

Model Information

Data Set	WORK.TWO
Dependent Variable	TIME
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
sex	2	F M

Dimensions

Covariance Parameters	1
Columns in X	9
Columns in Z	0
Subjects	1
Max Obs Per Subject	1490
Observations Used	1490
Observations Not Used	0
Total Observations	1490

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
Residual	744.06	0.05	693.30	800.64

Fit Statistics

-2 Res Log Likelihood	14090.4
AIC (smaller is better)	14092.4
AICC (smaller is better)	14092.4
BIC (smaller is better)	14097.7

Solution for Fixed Effects

Effect	sex	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		265.60	13.2146	1484	20.10	<.0001
Age		-2.3003	0.6613	1484	-3.48	0.0005
Age*Age		0.03392	0.008024	1484	4.23	<.0001
sex	F	5.3368	21.8182	1484	0.24	0.8068
sex	M	0
Age*sex	F	0.5335	1.1834	1484	0.45	0.6522
Age*sex	M	0
Age*Age*sex	F	-0.00486	0.01556	1484	-0.31	0.7548
Age*Age*sex	M	0

Type 1 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Age	1	1484	5.65	0.0176
Age*Age	1	1484	27.68	<.0001
sex	1	1484	135.89	<.0001
Age*sex	1	1484	0.90	0.3423
Age*Age*sex	1	1484	0.10	0.7548

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Age	1	1484	11.81	0.0006
Age*Age	1	1484	16.38	<.0001
sex	1	1484	0.06	0.8068
Age*sex	1	1484	0.20	0.6522
Age*Age*sex	1	1484	0.10	0.7548

```

38      proc mixed data=two CL; classes sex; TITLE2 'Quadratic model - version for estimates';
39      model time= sex sex*age sex*age*age / htype=1 3 DDFM=Satterthwaite solution noint;
40      run;

```

NOTE: The PROCEDURE MIXED printed page 7.

NOTE: PROCEDURE MIXED used:

```

      real time          0.06 seconds
      cpu time           0.06 seconds

```

EXST7015: Marathon Footrace Example
 Quadratic model - version for estimates

The Mixed Procedure

Model Information

```

Data Set          WORK.TWO
Dependent Variable  TIME
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
sex      2      F M

```

Dimensions

```

Covariance Parameters      1
Columns in X                6
Columns in Z                0
Subjects                    1
Max Obs Per Subject        1490
Observations Used          1490
Observations Not Used      0
Total Observations         1490

```

Covariance Parameter Estimates

```

Cov Parm      Estimate      Alpha      Lower      Upper
Residual      744.06         0.05      693.30     800.64

```

Fit Statistics

```

-2 Res Log Likelihood      14090.4
AIC (smaller is better)    14092.4
AICC (smaller is better)   14092.4
BIC (smaller is better)    14097.7

```

Solution for Fixed Effects

Effect	sex	Estimate	Standard Error	DF	t Value	Pr > t
sex	F	270.94	17.3612	1484	15.61	<.0001
sex	M	265.60	13.2146	1484	20.10	<.0001
Age*sex	F	-1.7668	0.9814	1484	-1.80	0.0720
Age*sex	M	-2.3003	0.6613	1484	-3.48	0.0005
Age*Age*sex	F	0.02906	0.01333	1484	2.18	0.0294
Age*Age*sex	M	0.03392	0.008024	1484	4.23	<.0001

Type 1 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
sex	2	1484	56030.1	<.0001
Age*sex	2	1484	15.63	<.0001
Age*Age*sex	2	1484	11.31	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
sex	2	1484	323.76	<.0001
Age*sex	2	1484	7.67	0.0005
Age*Age*sex	2	1484	11.31	<.0001

```
41      proc mixed data=two CL; classes sex; TITLE2 'Quadratic model -
simplified';
42      model time= age age*age sex / htype=1 3 DDFM=Satterthwaite
solution;
43      run;
```

NOTE: The PROCEDURE MIXED printed page 8.

NOTE: PROCEDURE MIXED used:
 real time 0.06 seconds
 cpu time 0.06 seconds

EXST7015: Marathon Footrace Example
 Quadratic model - simplified

The Mixed Procedure

Model Information	
Data Set	WORK.TWO
Dependent Variable	TIME
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
sex	2	F M

Dimensions	
Covariance Parameters	1
Columns in X	5
Columns in Z	0
Subjects	1
Max Obs Per Subject	1490
Observations Used	1490
Observations Not Used	0
Total Observations	1490

Covariance Parameter Estimates				
Cov Parm	Estimate	Alpha	Lower	Upper
Residual	743.56	0.05	692.86	800.06

Fit Statistics	
-2 Res Log Likelihood	14083.3
AIC (smaller is better)	14085.3
AICC (smaller is better)	14085.3
BIC (smaller is better)	14090.6

Solution for Fixed Effects						
Effect	sex	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		258.20	10.0083	1486	25.80	<.0001
Age		-1.9676	0.5157	1486	-3.82	0.0001

Age*Age		0.03042	0.006469	1486	4.70	<.0001
sex	F	17.9370	1.5382	1486	11.66	<.0001
sex	M	0

Type 1 Tests of Fixed Effects

Effect	DF	DF	F Value	Pr > F
Age	1	1486	5.65	0.0175
Age*Age	1	1486	27.70	<.0001
sex	1	1486	135.98	<.0001

Type 3 Tests of Fixed Effects

Effect	DF	DF	F Value	Pr > F
Age	1	1486	14.56	0.0001
Age*Age	1	1486	22.12	<.0001
sex	1	1486	135.98	<.0001

```

45      GOPTIONS DEVICE=CGMflwa GSFMODE=REPLACE GSFNAME=OUT NOPROMPT noROTATE
46      ftext='TimesRoman' ftitle='TimesRoman' htext=1 httitle=1 cttitle=black
         ctext=black;
47
48      GOPTIONS GSFNAME=OUT1;
49      FILENAME OUT1 'C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\MarathonReg1VT.CGM';
50      PROC GPLOT DATA=TWO; BY SEX;
51          TITLE1 font='TimesRoman' H=1 'Polynomial Regression Example';
52          TITLE2 font='TimesRoman' H=1 'Marathon race';
53          PLOT TIME*AGE=1 TIME*AGE=2 TIME*AGE=3 / overlay HAXIS=AXIS1 VAXIS=AXIS2;
54          AXIS1 LABEL=(font='TimesRoman' H=1 'Age (years)') WIDTH=1 MINOR=(N=1)
55                  VALUE=(font='TimesRoman' H=1) color=black ORDER=10 TO 70 BY 10;
56          AXIS2 LABEL=(ANGLE=90 font='TimesRoman' H=1 'Time to run marathon (min)')
57                  WIDTH=1 VALUE=(font='TimesRoman' H=1) MINOR=(N=5) color=black
58                  ORDER=125 TO 325 BY 25;
59          SYMBOL1 color=red V=None I=RQclm95 L=1 MODE=INCLUDE;
60          SYMBOL2 color=green V=None I=RQcli95 L=1 MODE=INCLUDE;
61          SYMBOL3 color=blue V=dot I=None L=1 MODE=INCLUDE; RUN;
NOTE: Regression equation : TIME = 270.9395 - 1.766769*Age + 0.029057*Age^2.
NOTE: Regression equation : TIME = 270.9395 - 1.766769*Age + 0.029057*Age^2.
NOTE: Foreground color BLACK same as background. Part of your graph may not be
      visible.
NOTE: 404 RECORDS WRITTEN TO
      C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\MarathonReg1VT.CGM
NOTE: Regression equation : TIME = 265.6026 - 2.300266*Age + 0.033918*Age^2.
NOTE: Regression equation : TIME = 265.6026 - 2.300266*Age + 0.033918*Age^2.
NOTE: Foreground color BLACK same as background. Part of your graph may not be
      visible.
NOTE: 713 RECORDS WRITTEN TO
      C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\MarathonReg1VT.CGM
62      **** V = "dot" would place a dot for each point;
63      **** I = for regression: R requests fitted regression line, L, Q or C
63      ! requests Linear,
64          Quadratic or cubic, CLM or CLI requests corresponding
64      ! confidence interval and
65          95 specifies alpha level for CI (any value from 50 to 99);
66      **** I = for categories" requests STD (std dev) 1 (1 width, 2 or 3) M
66      ! (of mean=std err)
67          J (join means of bars) t (add top & bottom hash) p (use pooled
67      ! variance);
68      **** Other options for categories: omit M=std dev, use B to get bar
68      ! for min/max;
69      RUN:
    
```

