The Use of SAS® Regression Procedures in Analyzing a Major Sporting Event: The Case of the 1994 Masters Tournament

William Orzech, U.S.A.F., Department of Defense, Los Angeles
Joseph Earley, Loyola Marymount University, Los Angeles

INTRODUCTION

Using a cross-section data set derived for the 1994 Masters professional golf tournament, the authors developed an econometric model formalizing the reasons for the various levels of performance. The Masters participant's score is used as the dependent variable. Regressors were selected which are believed to be integral to the performance of excellent golf. Due to the availability of data and theoretical analysis, the authors have selected the following variables as regressors: the number of fairways hit in regulation, driving distance, the number of greens hit in regulation, the number of times the player has previously played in the Masters and the number of putts. Using PROC CORR, PROC REG, PROC RSQUARE, PROC STEPWISE and PROC RSREG, the key factors responsible for excellent performance were determined. Graphical output using PROC PLOT and PROC CHART are presented. For the purpose of comparison, the Ladies Professional Golf Association National Championship for 1994 is also analyzed.

HISTORY OF THE MASTERS

Every spring one of the world's most important golf tournaments is contested, the Masters Tournament at Augusta National Golf Course in Augusta, Georgia. Numerous factors add to the uniqueness of the event, such as the fact that the Masters is the first of the major golf championships and it is held at the legendary Augusta National GC known not only for its challenge to the golfer, but also for its reflecting the classic beauty of a southern springtime. The Masters tournament draws its competitors from the very best golfers in the world, based on a criteria which emphasizes excellent past golfing performance.

ECONOMETRIC MODELS

The following OLS model was estimated in order to account for how well a golfer performed in the 1994 Masters:

$$\text{Score} = \beta_0 + \beta_1 \text{FWHIT} + \beta_2 \text{DD} + \beta_3 \text{GIR} + \beta_4 \text{PUTTS} + \beta_5 \text{STARTS} + \beta_6 \text{HI} + u_i$$

where:

FWHIT	Total number of fairways hit in regulation
DD	Average driving Distance per measured hole
GIR	Total number of greens hit in regulation
PUTTS	Total number of putts for the 72 hole tournament
STARTS	Number of times player has participated in Masters
HI	Handicap index - the U.S. Golf Association describes a "USGA Handicap Index" as a number that represents the ability of a player on a course of average difficulty. It is expressed in strokes and decimal fractions of a stroke.

Using the SAS PROC REG procedure, the following statistics were determined:

Dependent Variable: SCORE

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	6	939.12168	156.52028	9.544	0.0001
Error	32	524.77575	16.39924		
C Total	38	1463.89744			

Root MSE	4.04960	R-square	0.6415
Dep Mean	292.71795	Adj R-sq	0.5743
C.V.	1.38345		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	227.894493	22.73118832	10.026	0.0001
FWHIT	1	-0.191337	0.22535770	-0.849	0.4022
DD	1	0.055469	0.07041291	0.788	0.4366
GIR	1	-1.074572	0.19494692	-5.512	0.0001
PUTTS	1	0.842668	0.17884660	4.712	0.0001
STARTS	1	0.045490	0.10471250	0.434	0.6649
HI	1	0.764862	0.98880337	0.774	0.4449

UNIVARIATE PROCEDURE

Variable=SCORE

Moments

N	51	Sum Wgts	51
Mean	293.451	Sum	14966
Std Dev	6.552293	Variance	42.93255
Skewness	-0.20334	Kurtosis	-0.67596
USS	4393934	CSS	2146.627
CV	2.232841	Std Mean	0.917505
T:Mean=0	319.836	Prob>|T|	0.0001
Sgn Rank	663	Prob>|S|	0.0001
Num ^= 0	51		

Quantiles(Def=5)

100% Max	305	99%	305
75% Q3	299	95%	304
50% Med	293	90%	301
25% Q1	288	10%	285
0% Min	279	5%	282
		1%	279
Range	26		
Q3-Q1	11		
Mode	292		

Extremes

Lowest	Obs	Highest	Obs
279(	1)	301(	47)
281(	2)	304(	48)
282(	3)	304(	49)
283(	4)	305(	50)
285(	7)	305(	51)

Variable=FWHIT

Moments

N	51	Sum Wgts	51
Mean	46.2549	Sum	2359
Std Dev	3.491952	Variance	12.19373
Skewness	-0.78224	Kurtosis	1.001282
USS	109725	CSS	609.6863
CV	7.549365	Std Mean	0.488971
T:Mean=0	94.59641	Prob>|T|	0.0001
Sgn Rank	663	Prob>|S|	0.0001
Num ^= 0	51		

Quantiles(Def=5)

100% Max	53	99%	53
75% Q3	49	95%	51
50% Med	46	90%	50
25% Q1	45	10%	42
0% Min	35	5%	40
		1%	35
Range	18		
Q3-Q1	4		
Mode	46		

Extremes

Lowest	Obs	Highest	Obs
35(	26)	50(	38)
40(	50)	51(	8)
40(	29)	51(	18)
40(	24)	51(	41)
41(	49)	53(	10)

Variable=DD

Moments			
N	51	Sum Wgts	51
Mean	274.098	Sum	13979
Std Dev	11.97415	Variance	143.3802
Skewness	0.072786	Kurtosis	0.96311
USS	3838786	CSS	7169.01
CV	4.368564	Std Mean	1.676716
T:Mean=0	163.4732	Prob>|T|	0.0001
Sgn Rank	663	Prob>|S|	0.0001
Num ^= 0	51		

Quantiles(Def=5)			
100% Max	302.1	99%	302.1
75% Q3	280.5	95%	301.4
50% Med	274.5	90%	284.1
25% Q1	265.8	10%	261.8
0% Min	242.3	5%	256.1
		1%	242.3
Range	59.8		
Q3-Q1	14.7		
Mode	265		

Extremes			
Lowest	Obs	Highest	Obs
242.3(	26)	290.1(	22)
248.4(	23)	292(	29)
256.1(	38)	301.4(	48)
256.3(	28)	301.9(	40)
259.8(	50)	302.1(	45)

Variable=GIR

Moments			
N	51	Sum Wgts	51
Mean	41.88235	Sum	2136
Std Dev	4.283209	Variance	18.34588
Skewness	0.484739	Kurtosis	-0.18501
USS	90378	CSS	917.2941
CV	10.22676	Std Mean	0.599769
T:Mean=0	69.83077	Prob>|T|	0.0001
Sgn Rank	663	Prob>|S|	0.0001
Num ^= 0	51		

Quantiles(Def=5)			
100% Max	53	99%	53
75% Q3	44	95%	49
50% Med	42	90%	48
25% Q1	39	10%	36
0% Min	35	5%	36
		1%	35
Range	18		
Q3-Q1	5		
Mode	42		

Extremes			
Lowest	Obs	Highest	Obs
35(	20)	49(	6)
36(	48)	49(	8)
36(	38)	49(	10)
36(	26)	51(	4)
36(	23)	53(	7)

Variable=PUTTS

Moments			
N	51	Sum Wgts	51
Mean	118.3333	Sum	6035
Std Dev	4.077581	Variance	16.62667
Skewness	-0.07928	Kurtosis	-0.2372
USS	714973	CSS	831.3333
CV	3.445843	Std Mean	0.570976
T:Mean=0	207.2476	Prob>|T|	0.0001
Sgn Rank	663	Prob>|S|	0.0001
Num ^= 0	51		

Quantiles(Def=5)			
100% Max	128	99%	128
75% Q3	121	95%	125
50% Med	118	90%	123
25% Q1	116	10%	113
0% Min	110	5%	111
		1%	110
Range	18		
Q3-Q1	5		
Mode	117		

Extremes			
Lowest	Obs	Highest	Obs
110(	3)	123(	48)
110(	1)	125(	19)
111(	9)	125(	27)
112(	20)	125(	39)
112(	5)	128(	30)

Moments			
N	51	Sum Wgts	51
Mean	7.411765	Sum	378
Std Dev	7.141923	Variance	51.00706
Skewness	1.043249	Kurtosis	0.436285
USS	5352	CSS	2550.353
CV	96.35927	Std Mean	1.000069
T:Mean=0	7.411252	Prob>|T|	0.0001
Sgn Rank	517.5	Prob>|S|	0.0001
Num ^= 0	45		

Quantiles(Def=5)			
100% Max	29	99%	29
75% Q3	12	95%	21
50% Med	5	90%	18
25% Q1	1	10%	0
0% Min	0	5%	0
		1%	0
Range	29		
Q3-Q1	11		
Mode	1		

Extremes			
Lowest	Obs	Highest	Obs
0(	50)	20(	4)
0(	45)	20(	13)
0(	41)	21(	23)
0(	29)	22(	21)
0(	8)	29(	10)

Variable=HI

Moments			
N	39	Sum Wgts	39
Mean	4.682051	Sum	182.6
Std Dev	0.726517	Variance	0.527827
Skewness	-0.38892	Kurtosis	-0.15782
USS	875	CSS	20.05744
CV	15.51707	Std Mean	0.116336
T:Mean=0	40.24599	Prob>|T|	0.0001
Sgn Rank	390	Prob>|S|	0.0001
Num ^= 0	39		

Quantiles(Def=5)			
100% Max	6.1	99%	6.1
75% Q3	5.3	95%	5.9
50% Med	4.8	90%	5.5
25% Q1	4.2	10%	3.5
0% Min	2.9	5%	3.4
		1%	2.9
Range	3.2		
Q3-Q1	1.1		
Mode	4.2		

Extremes			
Lowest	Obs	Highest	Obs
2.9(	26)	5.5(	35)
3.4(	21)	5.5(	37)
3.4(	6)	5.5(	39)
3.5(	16)	5.9(	29)
3.8(	30)	6.1(	36)

Stepwise Procedure for Dependent Variable SCORE

Step 1 Variable GIR Entered R-square = 0.30960773 C(p) = 26.62866891

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	1	453.23396186	453.23396186	16.59	0.0002
Error	37	1010.66347404	27.31522503		
Total	38	1463.89743590			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	325.45638640	5.08055080	44310.69080755	1622.20	0.0001
GIR	-0.77786774	0.19089526	453.23396186	16.59	0.0002

Bounds on condition number: 1, 1

Step 2 Variable PUTTS Entered R-square = 0.61625670 C(p) = 1.07676002

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	2	905.06439158	452.53219579	29.15	0.0001
Error	36	558.83304432	15.52314012		
Total	38	1463.89743590			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	234.99840526	17.83903146	2693.80940708	173.54	0.0001
GIR	-1.11235045	0.15671312	782.08167077	50.38	0.0001
PUTTS	0.88277585	0.16362607	451.83042971	29.11	0.0001

Bounds on condition number: 1.185918, 4.743673

All variables in the model are significant at the 0.1500 level.
No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable SCORE

Step	Variable Entered	Variable Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	GIR		1	0.3096	0.3096	26.6287	16.5927	0.0002
2	PUTTS		2	0.3086	0.6183	1.0768	29.1069	0.0001

Simple Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
SCORE	51	293.45098	6.55229	279.00000	305.00000
FWHIT	51	46.25490	3.49195	35.00000	53.00000
DD	51	274.09804	11.97415	242.30000	302.10000
GIR	51	41.84235	4.28321	35.00000	53.00000
PUTTS	51	118.33333	4.07758	110.00000	128.00000
STARTS	51	7.41176	7.14192	0	29.00000
HI	39	4.68205	0.72652	2.90000	6.10000

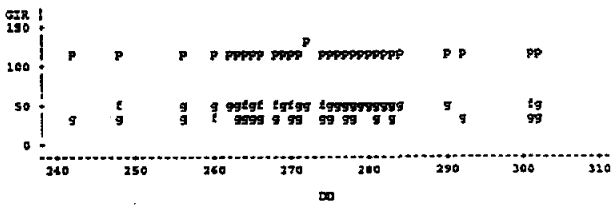
Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

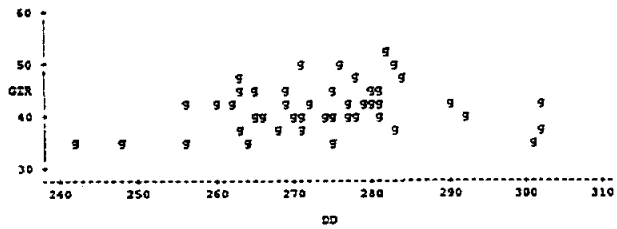
	SCORE	FWHIT	DD	GIR	PUTTS	STARTS	HI
SCORE	1.00000	-0.30058	0.08650	-0.57103	0.36256	-0.29296	0.15058
	0.0	0.0321	0.5461	0.0001	0.0089	0.0370	0.3602
	51	51	51	51	51	51	39
FWHIT	-0.30058	1.00000	0.03330	0.40588	0.11190	0.17775	0.17766
	0.0321	0.0	0.8165	0.0031	0.4343	0.2121	0.2792
	51	51	51	51	51	51	39

DD	0.08650 0.5461 51	0.03330 0.8165 51	1.00000 0.0 51	0.17973 0.2070 51	0.25173 0.0748 51	-0.19099 0.1796 51	0.20691 0.2083 39
GIR	-0.57103 0.0001 51	0.40588 0.0831 51	0.17973 0.2070 51	1.00000 0.0 51	0.27827 0.0480 51	0.30986 0.0271 51	-0.13806 0.4019 39
PUTTS	0.36256 0.0089 51	0.11198 0.4343 51	0.25173 0.0748 51	0.27827 0.0480 51	1.00000 0.0 51	-0.08722 0.5428 51	-0.08255 0.6173 39
STARTS	-0.29296 0.0370 51	0.17775 0.2121 51	-0.19099 0.1796 51	0.30956 0.0271 51	-0.08722 0.5428 51	1.00000 0.0 51	0.00900 0.9567 39
NI	0.15058 0.3692 39	0.17766 0.2792 39	0.20691 0.2083 39	-0.13806 0.4019 39	-0.08255 0.6173 39	0.00900 0.9567 39	1.00000 0.0 39

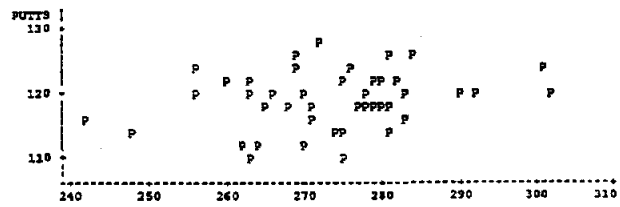
Plot of GIR*DD. Symbol used is 'g'.
Plot of PUTTS*DD. Symbol used is 'p'.
Plot of FWHIT*DD. Symbol used is 'f'.



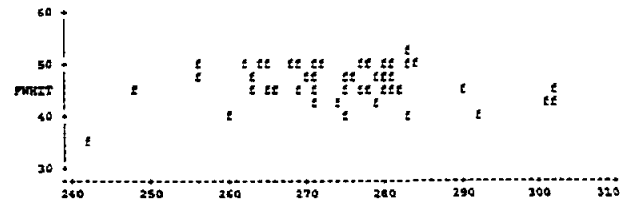
Plot of GIR*DD. Symbol used is 'g'.



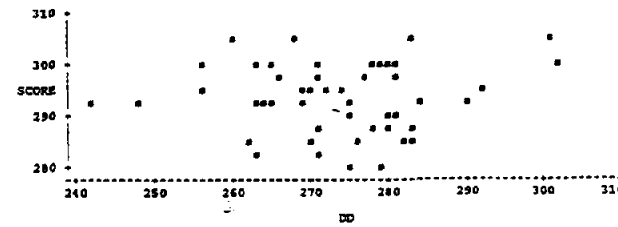
Plot of PUTTS*DD. Symbol used is 'p'.



Plot of FWHIT*DD. Symbol used is 'f'.



Plot of SCORE*DD. Symbol used is '*'.



Ladies Professional Golf Association 1994 Championship

Econometric Model

The following OLS model was estimated using PROC REG to explain how well a golfer performed in the 1994 LPGA:

$$Score = \beta_0 + \beta_1 FWHIT + \beta_2 DD + \beta_3 GIR + \beta_4 PUTTS + u_i$$

where:
- FWHIT — Total number of fairways hit in regulation
- DD — Average driving Distance per measured hole
- GIR — Total number of greens hit in regulation
- PUTTS — Total number of putts for the 72 hole tournament

Dependent Variable: SCORE

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	2049.54694	512.38674	48.333	0.0001
Error	75	795.25302	10.60337		
C Total	79	2844.80000			

Root MSE	3.25628	R-square	0.7205
Dep Mean	293.20000	Adj R-sq	0.7055
C.V.	1.11060		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	273.155000	14.85813758	18.384	0.0001
FWHIT	1	-0.077191	0.07722868	-1.000	0.3208
DD	1	-0.093123	0.04903516	-1.899	0.0614
GIR	1	-0.892479	0.08036037	-11.106	0.0001
PUTTS	1	0.669927	0.08157128	8.213	0.0001

UNIVARIATE PROCEDURE

Variable=SCORE

Moments

N	80	Sum Wgts	80
Mean	292.2	Sum	23456
Std Dev	6.000844	Variance	36.01013
Skewness	-0.10884	Kurtosis	-0.70632
USS	6880144	CSS	2844.8
CV	2.046673	Std Mean	0.670915
T:Mean=0	437.0153	Prob>|T|	0.0001
Sgn Rank	1620	Prob>|S|	0.0001
Num ^= 0	80		

Quantiles(Def=5)

100% Max	305	99%	305
75% Q3	298	95%	302
50% Med	293	90%	300.5
25% Q1	289	10%	285
0% Min	279	5%	283
		1%	279
Range	26		
Q3-Q1	9		
Mode	289		

Extremes

Lowest	Obs	Highest	Obs
279(	1)	301(	76)
282(	2)	303(	77)
283(	6)	304(	78)
283(	5)	305(	79)
283(	4)	305(	80)

Moments

N	80	Sum Wgts	80
Mean	38.875	Sum	3110
Std Dev	5.190972	Variance	26.9462
Skewness	-0.49488	Kurtosis	-0.42045
USS	123030	CSS	2131.75
CV	13.35299	Std Mean	0.580368
T:Mean=0	66.98331	Prob>|T|	0.0001
Sgn Rank	1620	Prob>|S|	0.0001
Num ^= 0	80		

Variable=FWHIT

Quantiles(Def=5)

100% Max	49	99%	49
75% Q3	43	95%	46
50% Med	40	90%	45
25% Q1	35.5	10%	32
0% Min	26	5%	29
		1%	26
Range	23		
Q3-Q1	7.5		
Mode	40		

Extremes

Lowest	Obs	Highest	Obs
26(	53)	46(	34)
27(	80)	46(	45)
29(	54)	46(	62)
29(	51)	47(	35)
29(	24)	49(	75)

Variable=DD

Moments

N	80	Sum Wgts	80
Mean	221.9462	Sum	17755.7
Std Dev	8.917753	Variance	79.52631
Skewness	1.349465	Kurtosis	5.156288
USS	3947094	CSS	6282.579
CV	4.017979	Std Mean	0.997035
T:Mean=0	222.6063	Prob>|T|	0.0001
Sgn Rank	1620	Prob>|S|	0.0001
Num ^= 0	80		

Quantiles(Def=5)

100% Max	263.5	99%	263.5
75% Q3	226.15	95%	237.3
50% Med	221.65	90%	231.15
25% Q1	217	10%	210.4
0% Min	204.5	5%	209
		1%	204.5
Range	59		
Q3-Q1	9.15		
Mode	219.8		

Extremes

Lowest	Obs	Highest	Obs
204.5(	40)	234.8(	16)
206.8(	75)	239.5(	13)
207(	61)	240.5(	27)
209(	74)	242.3(	38)
209(	54)	263.5(	1)

Variable=GIR

Moments			
N	80	Sum Wgts	80
Mean	42.2125	Sum	3377
Std Dev	6.026721	Variance	36.32136
Skewness	0.073603	Kurtosis	-0.02439
USS	145421	CSS	2869.387
CV	14.2771	Std Mean	0.673808
T:Mean=0	62.64768	Prob>|T|	0.0001
Sgn Rank	1620	Prob>|S|	0.0001
Num ^= 0	80		

Quantiles(Def=5)			
100% Max	55	99%	55
75% Q3	46	95%	55
50% Med	43	90%	49.5
25% Q1	37.5	10%	34.5
0% Min	27	5%	32
		1%	27
Range	28		
Q3-Q1	8.5		
Mode	43		

Extremes			
Lowest	Obs	Highest	Obs
27(	10)	55(	1)
31(	77)	55(	5)
32(	71)	55(	6)
32(	68)	55(	9)
32(	64)	55(	12)

Variable=PUTTS

Moments			
N	80	Sum Wgts	80
Mean	121.4875	Sum	9719
Std Dev	5.274716	Variance	27.82263
Skewness	-0.24459	Kurtosis	3.230721
USS	1182925	CSS	2197.988
CV	4.3417771	Std Mean	0.589721
T:Mean=0	206.0049	Prob>|T|	0.0001
Sgn Rank	1620	Prob>|S|	0.0001
Num ^= 0	80		

Quantiles(Def=5)			
100% Max	140	99%	140
75% Q3	125	95%	129
50% Med	121	90%	128
25% Q1	119	10%	115
0% Min	101	5%	113.5
		1%	101
Range	39		
Q3-Q1	6		
Mode	121		

Extremes			
Lowest	Obs	Highest	Obs
101(	56)	129(	12)
112(	33)	129(	13)
112(	10)	130(	66)
113(	2)	130(	76)
114(	44)	140(	67)

Stepwise Procedure for Dependent Variable SCORE

Step 1 Variable GIR Entered R-square = 0.43501602 C(p) = 75.58066569

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	1	1237.53357119	1237.53357119	60.06	0.0001
Error	78	1607.26642881	20.60597986		
Total	79	2844.80000000			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	320.92202604	3.61302199	162574.00807762	7889.65	0.0001
GIR	-0.65672552	0.08474265	1237.53357119	60.06	0.0001

Bounds on condition number: 1, 1

Step 2 Variable PUTTS Entered R-square = 0.70670740 C(p) = 4.68804837

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	2	2010.44122136	1005.22061099	92.77	0.0001
Error	77	834.35877802	10.83582829		
Total	79	2844.80000000			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	250.13211535	8.78177191	8790.96842962	811.29	0.0001
GIR	-0.96539414	0.07149887	1975.48204228	182.31	0.0001
PUTTS	0.64994411	0.08169231	772.90765080	71.33	0.0001

Bounds on condition number: 1.353709, 5.414837

Step 3 Variable DD Entered R-square = 0.71673018 C(p) = 3.99901729

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	3	2038.95402331	679.65134110	64.10	0.0001
Error	76	805.84597669	10.60323654		
Total	79	2844.80000000			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	265.84671014	12.93439004	4479.28941583	422.45	0.0001
DD	-0.07411795	0.04519836	28.51280132	2.69	0.1052
GIR	-0.81540169	0.07701797	1497.88258858	141.27	0.0001
PUTTS	0.67862855	0.08110486	742.35032066	70.01	0.0001

Bounds on condition number: 1.605221, 12.53774

All variables in the model are significant at the 0.1500 level.
No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable SCORE

Step	Variable Entered	Variable Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	GIR		1	0.4350	0.4350	75.5807	60.0570	0.0001
2	PUTTS		2	0.2717	0.7067	4.6880	71.3289	0.0001
3	DD		3	0.0100	0.7167	3.9990	2.6891	0.1052

Simple Statistics

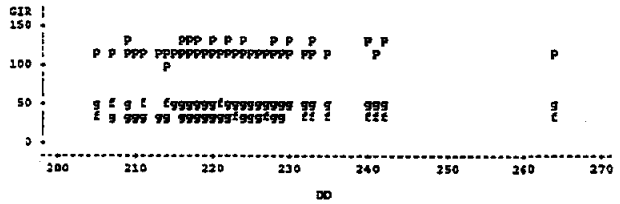
Variable	N	Mean	Std Dev	Minimum	Maximum
SCORE	80	293.20000	5.00084	279.0000	305.0000
FWHIT	80	38.87500	5.13097	26.0000	49.0000
DD	80	221.94625	8.91775	204.5000	243.5000
GIR	80	42.21250	6.02672	27.0000	55.0000
PUTTS	80	121.48750	5.27472	101.0000	140.0000

Correlation Analysis

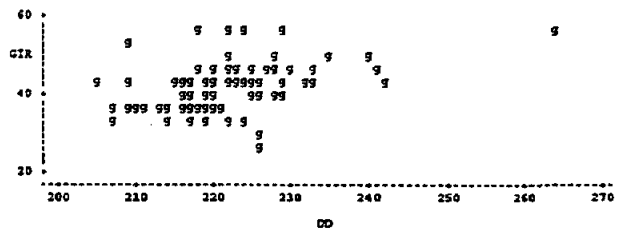
Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 80

	SCORE	FWHIT	DD	GIR	PUTTS
SCORE	1.00000 0.0	-0.12597 0.2655	-0.40165 0.0002	-0.65956 0.0001	0.11085 0.3276
FWHIT	-0.12597 0.2655	1.00000 0.0	-0.29806 0.0072	0.11294 0.3185	0.00133 0.9907
DD	-0.40165 0.0002	-0.29806 0.0072	1.00000 0.0	0.40967 0.0002	0.14270 0.2067
GIR	-0.65956 0.0001	0.11294 0.3185	0.40967 0.0002	1.00000 0.0	0.51116 0.0001
PUTTS	0.11085 0.3276	0.00133 0.9907	0.14270 0.2067	0.51116 0.0001	1.00000 0.0

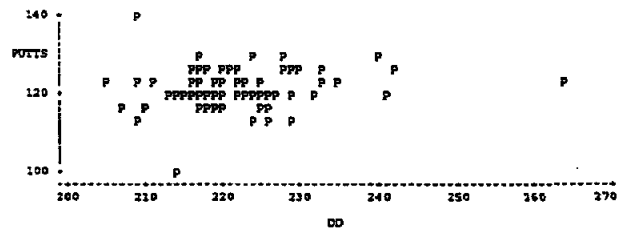
Plot of GIR*DD. Symbol used is 'g'.
Plot of PUTTS*DD. Symbol used is 'p'.
Plot of FWHIT*DD. Symbol used is 'f'.



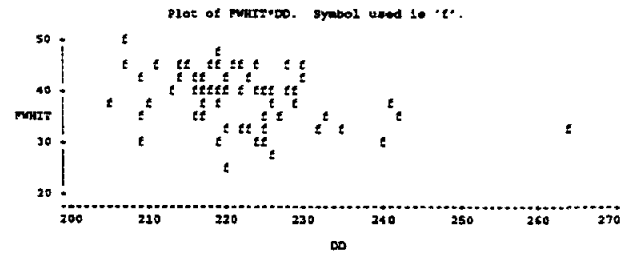
Plot of GIR*DD. Symbol used is 'g'.

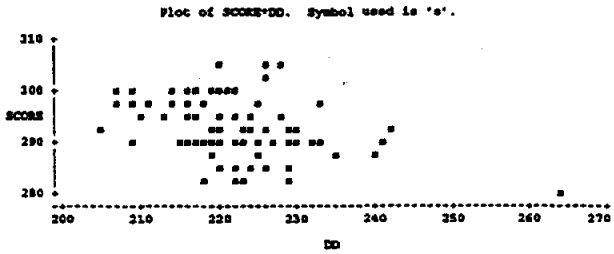


Plot of PUTTS*DD. Symbol used is 'p'.



Plot of FWHIT*DD. Symbol used is 'f'.





Analysis

The regression model was able to account for 64.15 % of the variation in score for the Masters players. The over-all model was statistically significant as judged by the F - value (p - value of 0.0001). The two statistically significant individual variables were GIR and PUTTS. This reflects the fact that the more accurate a player is in terms of getting the ball on the green in the appropriate amount of strokes the better the score. Putting, generally thought of as the most important part of golf, was directly related to score as was predicted. Apparently driving distance, the previous number of times a player was in the Masters and the player's handicap index were not statistically significant. This appears to give added proof to the golfer's addage: "You drive for show, and putt for dough".

For the Ladies Professional Golf Association 1994 Open Championship knowledge of fairways hit, driving distance, greens in regulation and the number of putts accounted for 72.05 % of the variation in the player's score. As with the Masters participants, GIR and PUTTS were statistically significant at the α = 0.05 level of significance. In addition, driving distance was very close to being significance with a p - value of 0.0614. A brief perusal of the driving distance of the leaders of the LPGA OPEN compared to the rest of the field indicated that for the ability to drive far was a distinct advantage in this tournament. The mean driving distance for the ladies was 221.95 with a standard deviation of 8.92 for a coefficient of variation of 4.02. For the men in the Masters, the mean driving distance was 274.10 with a standard deviation of 11.97 for a coefficient of variation of 4.37. It appears that there is a minimum average driving distance that must be attained for a competitor to be able to seriously contend for the championship. Those competitors whose driving distance places them in the lowest 10th percentile appear to have very little chance of victory in a major championship.

References

Golf World: International News Magazine of Golf, April 1, 15 and May 20, 1994.

Golf Digest: 1994 Annual Edition, pp. 164 - 165.

Gujarati, Damodar. Basic Econometrics. Second edition, McGraw-Hill, 1988.

SAS Institute Inc. SAS System for Elementary Statistical Analysis 1987.

SAS Institute Inc., SAS/ETS User's Guide. 1988 Version 6, First Edition, Cary, NC: SAS Institute Inc., 1988.

SAS Institute Inc., SAS/ETS Software: Applications Guide 1, Version 6, First Edition, 1991.

SAS Institute Inc., SAS/STAT User's Guide, Vols I and II, 1990 Version 6, Fourth Edition. Cary, N.C.: SAS Institute Inc.