CHAPTER 12
Multiple Regression and the General Linear Model

Multiple regression examines the relationship between the dependent variable and a group of quantitative independent variables. The equation for the multiple regression is

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon$$

The independent variables in the multiple regression may be:

a. powers of the other independent variables,
b. cross-products of the other independent variables, or
c. transformations of the other independent variables.

The only restriction is that no x may be a perfect linear function of the other independent variables.

The parameters (β_1, β_2, ..., β_k) are called **partial slopes**. The value β_i represents the expected change in y for a unit change in x_i, when all other x's are held constant.

Example: The Admissions Office of a local college would like to predict a student's college grade point ratio based on scores from standardized tests. Below is information on twenty students.

Obs	Grade Point Ratio	Creativity Score	Mechanical Ability Score	Abstract Thinking Score	Mathematical Ability Score
1	1.0	10	12	11	10
2	2.9	21	21	48	47
3	2.5	10	32	47	50
4	2.4	29	47	33	29
5	3.7	39	47	48	49
6	3.1	40	27	32	29
7	2.1	12	15	38	42
8	2.3	42	16	12	10
9	1.3	11	28	19	21
10	2.9	28	15	42	41
11	1.9	19	45	32	30
12	3.1	31	31	48	49
13	1.7	10	40	31	30
14	2.1	32	21	19	20
15	1.2	20	35	13	12
16	2.4	37	43	19	22
17	2.5	22	30	39	41
18	1.5	18	14	22	19
19	3.5	43	32	40	38
20	1.5	28	33	12	11

Grade Point Ratio Example

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 20

	GPR	CREATE	MECH	ABSTRACT	MATH
GPR	1.00000	0.66044	0.21658	0.76766	0.73849
	0.0	0.0015	0.3591	0.0001	0.0002
CREATE	0.66044	1.00000	0.18879	0.04956	0.00363
	0.0015	0.0	0.4254	0.8356	0.9879
MECH	0.21658	0.18879	1.00000	0.19425	0.18333
	0.3591	0.4254	0.0	0.4119	0.4391
ABSTRACT	0.76766	0.04956	0.19425	1.00000	0.98754
	0.0001	0.8356	0.4119	0.0	0.0001
MATH	0.73849	0.00363	0.18333	0.98754	1.00000
	0.0002	0.9879	0.4391	0.0001	0.0

Simple Regression Model Using Abstract Thinking Score to Predict GPR

Dependent Variable: GPR

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	6.64259	6.64259	25.828	0.0001
Error	18	4.62941	0.25719		
C Total	19	11.27200			

Root MSE	0.50714	R-square	0.5893
Dep Mean	2.28000	Adj R-sq	0.5665
C.V.	22.24291		

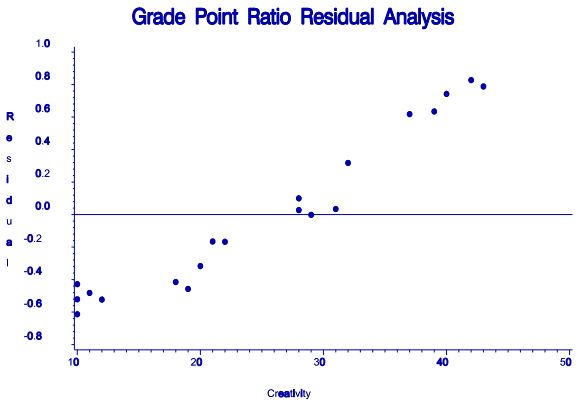
Parameter Estimates

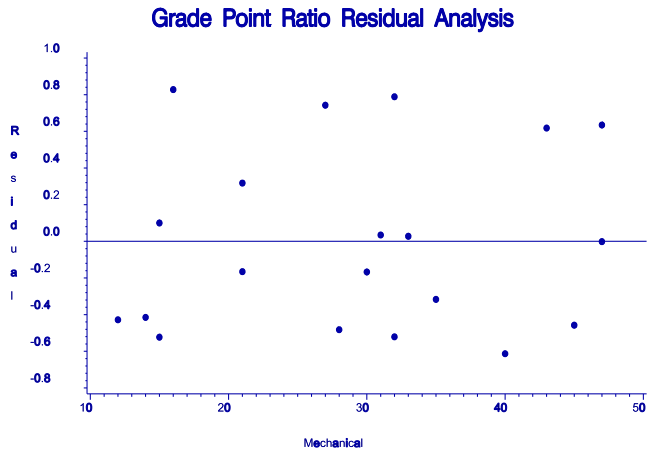
Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	0.941303	0.28678695	3.282	0.0041
ABSTRACT	1	0.044254	0.00870792	5.082	0.0001

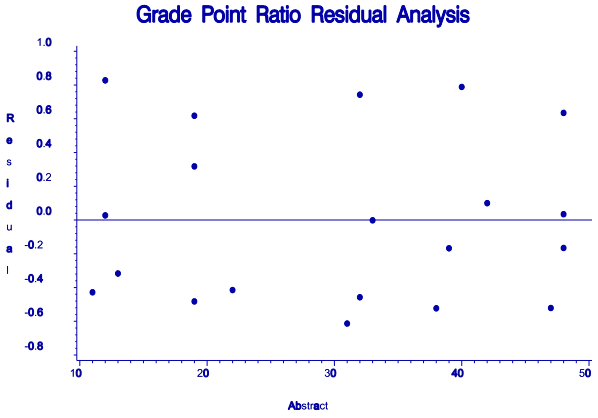
Obs	Dep Var GPR	Predict Value	Std Err Predict	Lower95% Predict	Upper95% Predict	Residual
1	1.0000	1.4281	0.202	0.2809	2.5753	-0.4281
2	2.9000	3.0655	0.192	1.9265	4.2046	-0.1655
3	2.5000	3.0213	0.185	1.8873	4.1552	-0.5213
4	2.4000	2.4017	0.116	1.3088	3.4946	-0.00170
5	3.7000	3.0655	0.192	1.9265	4.2046	0.6345
6	3.1000	2.3574	0.114	1.2652	3.4497	0.7426
7	2.1000	2.6230	0.132	1.5220	3.7239	-0.5230
8	2.3000	1.4724	0.195	0.3307	2.6140	0.8276
9	1.3000	1.7821	0.150	0.6711	2.8931	-0.4821
10	2.9000	2.8000	0.153	1.6873	3.9127	0.1000
11	1.9000	2.3574	0.114	1.2652	3.4497	-0.4574
12	3.1000	3.0655	0.192	1.9265	4.2046	0.0345
13	1.7000	2.3132	0.114	1.2213	3.4050	-0.6132
14	2.1000	1.7821	0.150	0.6711	2.8931	0.3179
15	1.2000	1.5166	0.188	0.3801	2.6531	-0.3166
16	2.4000	1.7821	0.150	0.6711	2.8931	0.6179
17	2.5000	2.6672	0.137	1.5638	3.7707	-0.1672
18	1.5000	1.9149	0.134	0.8127	3.0171	-0.4149
19	3.5000	2.7115	0.142	1.6052	3.8177	0.7885
20	1.5000	1.4724	0.195	0.3307	2.6140	0.0276

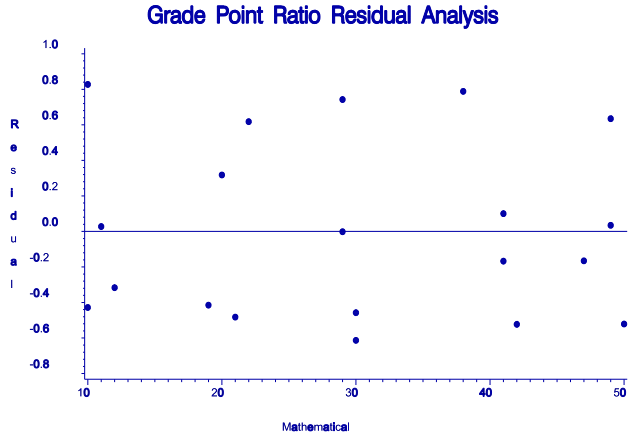
Sum of Residuals	0
Sum of Squared Residuals	4.6294
Predicted Resid SS (Press)	5.6970

Residual Analysis for GPR=0.941303+0.044254*Abstract









Multiple Regression Model Using Abstract Thinking Score and Creativity Score to Predict GPR

Dependent Variable: GPR

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	11.01981	5.50990	371.415	0.0001
Error	17	0.25219	0.01483		
C Total	19	11.27200			

Root MSE	0.12180	R-square	0.9776
Dep Mean	2.28000	Adj R-sq	0.9750
C.V.	5.34204		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-0.062321	0.09032044	-0.690	0.4995
ABSTRACT	1	0.042472	0.00209394	20.283	0.0001
CREATE	1	0.042133	0.00245284	17.177	0.0001

Obs	Dep Var GPR	Predict Value	Std Err Predict	Lower95% Predict	Upper95% Predict	Residual
1	1.0000	0.8262	0.060	0.5398	1.1126	0.1738
2	2.9000	2.8611	0.048	2.5853	3.1370	0.0389
3	2.5000	2.3552	0.059	2.0697	2.6407	0.1448
4	2.4000	2.5611	0.029	2.2968	2.8254	-0.1611
5	3.7000	3.6195	0.056	3.3365	3.9026	0.0805
6	3.1000	2.9821	0.046	2.7077	3.2565	0.1179
7	2.1000	2.0572	0.046	1.7827	2.3317	0.0428
8	2.3000	2.2169	0.064	1.9268	2.5071	0.0831
9	1.3000	1.2081	0.049	0.9310	1.4852	0.0919
10	2.9000	2.9012	0.037	2.6326	3.1699	-0.00123
11	1.9000	2.0973	0.031	1.8319	2.3627	-0.1973
12	3.1000	3.2825	0.048	3.0065	3.5585	-0.1825
13	1.7000	1.6756	0.046	1.4009	1.9504	0.0244
14	2.1000	2.0929	0.040	1.8223	2.3636	0.00709
15	1.2000	1.3325	0.046	1.0575	1.6075	-0.1325
16	2.4000	2.3036	0.047	2.0281	2.5791	0.0964
17	2.5000	2.5210	0.034	2.2543	2.7878	-0.0210
18	1.5000	1.6305	0.036	1.3624	1.8986	-0.1305
19	3.5000	3.4483	0.055	3.1665	3.7300	0.0517
20	1.5000	1.6271	0.048	1.3511	1.9031	-0.1271

Sum of Residuals	0
Sum of Squared Residuals	0.2522
Predicted Resid SS (Press)	0.3532