

Regression Diagnostics with SAS

```
/****** cars2.sas *****/
options linesize=79 pagesize=35;
title 'Auto gas mileage: residual plots and so on';

proc format; /* Used to label values of the categorical variables */
  value carfmt
    1 = 'US'
    2 = 'Japanese'
    3 = 'Other' ;

data auto;
  infile 'cars.dat';
  input country mpg weight;
  zweight = (weight-3010.811)/783.8957;
           /* Got mean and sd from an earlier run */
  zweight2 = zweight**2;
/* Indicator dummy vars */
  if country = 1 then c1=1; else c1=0;
  if country = 3 then c2=1; else c2=0;
/* Interaction terms */
  wc1 = zweight*c1; wc2 = zweight*c2;
  w2c1 = zweight2*c1; w2c2 = zweight2*c2;
/* Variable labels */
  label country = 'Country of Origin'
        mpg = 'Miles per Gallon'
        zweight = 'Standardized Weight'
        zweight2 = 'Zweight Squared';
  format country carfmt.;

proc reg;
  model mpg = zweight c1 c2;
  output out=resdata predicted=prempg residual=rempg;

proc plot;
  plot rempg*(zweight zweight2) = country;

proc sort; by country;
proc univariate normal plot; var rempg;
  by country; /* side-by-side boxplots */

data peek;
  set resdata;
  if rempg > 10;
proc print;
```

```

proc reg data = auto; /* Don't want to use the data set peek */
  model mpg = zweight zweight2 c1 c2 / r influence;
/*
  i          prints (X'X)-inverse
  ssl       prints sequential sums of squares
  clm       prints confidence interval for E(Yh)
  cli       prints prediction interval for new observation
  r         prints residual analysis
  influence prints influence statistics
  partial   prints partial regression plots
*/
  output out=resdata2 predicted=prempg residual=rempg student=studres
  cookd=cook h=leverage press=press rstudent=sdelres dffits=dfit;

proc plot;
  plot sdelres*zweight;

proc means;
  class country;
  var rempg;

data peek2;
  set resdata2; if sdelres**2 > 4;
proc print;

data fix;
  set auto;
  if _n_ ne 71; /* Delete observation 71 */

proc reg;
  model mpg = zweight zweight2 c1 c2;
  output out=resdata3 predicted=prempg;
  USvsOt2: test c1=c2;
  countr2: test c1=c2=0;

```

The REG Procedure
 Model: MODEL1
 Dependent Variable: mpg Miles per Gallon

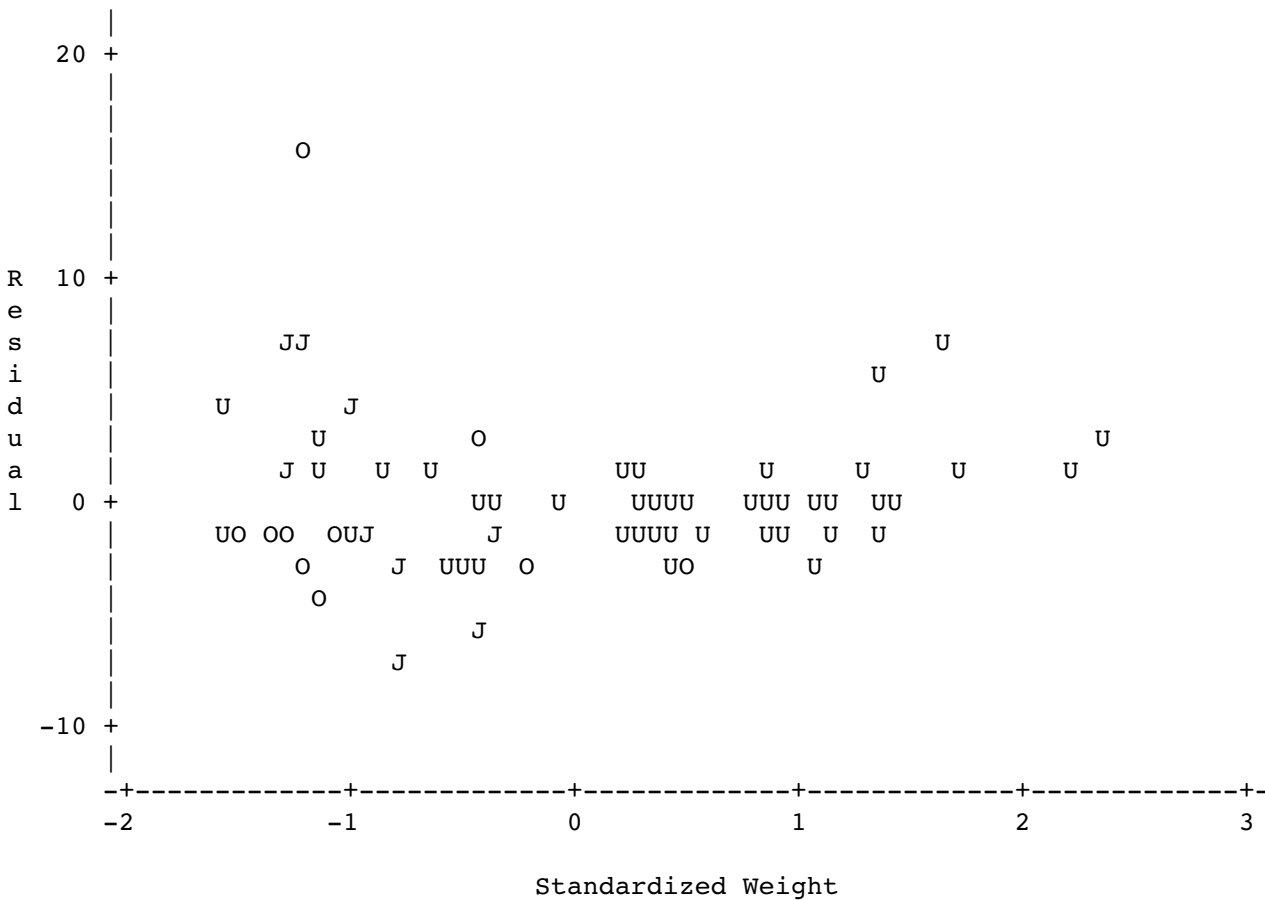
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1701.19566	567.06522	53.48	<.0001
Error	70	742.26379	10.60377		
Corrected Total	73	2443.45946			
Root MSE		3.25634	R-Square	0.6962	
Dependent Mean		21.29730	Adj R-Sq	0.6832	
Coeff Var		15.28993			

Parameter Estimates

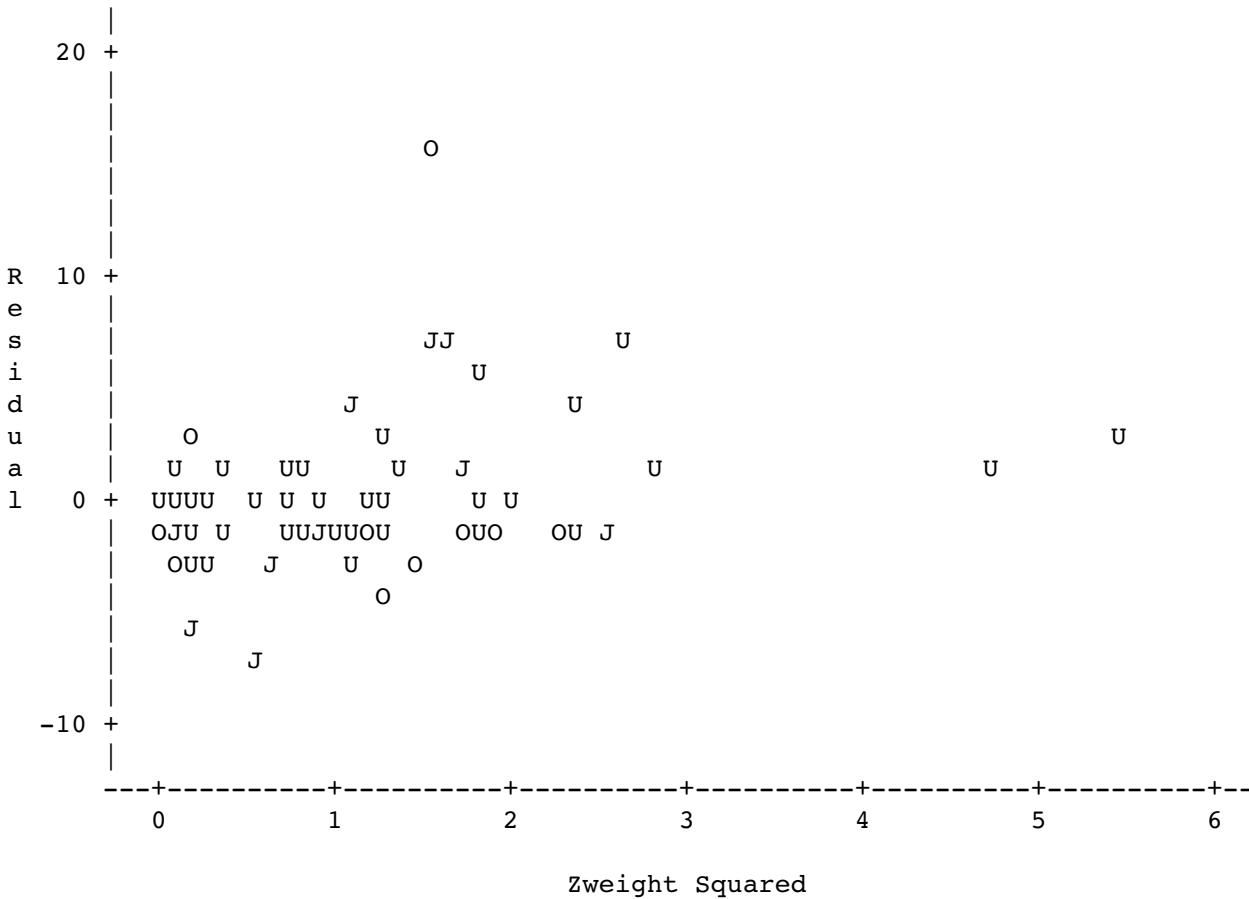
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	21.17135	1.08262	19.56	<.0001
zweight	Standardized Weight	1	-5.15019	0.46888	-10.98	<.0001
c1		1	0.58714	1.25193	0.47	0.6405
c2		1	-1.92825	1.39087	-1.39	0.1700

Plot of rempg*zweight. Symbol is value of country.



NOTE: 11 obs hidden.

Plot of rempg*zweight2. Symbol is value of country.



Now the proc univariate output. There's a lot, so only parts will be shown. The by statement produces separate output for each value of country.

----- Country of Origin=US -----

The UNIVARIATE Procedure
Variable: rempg (Residual)

Moments

N	52	Sum Weights	52
Mean	0	Sum Observations	0
Std Deviation	2.04986024	Variance	4.20192701
Skewness	1.52027841	Kurtosis	3.89229267
Uncorrected SS	214.298278	Corrected SS	214.298278
Coeff Variation	.	Std Error Mean	0.28426447

Tests for Normality

Test	--Statistic---		-----p Value-----	
Shapiro-Wilk	W	0.88851	Pr < W	0.0002
Kolmogorov-Smirnov	D	0.117612	Pr > D	0.0725
Cramer-von Mises	W-Sq	0.172739	Pr > W-Sq	0.0118
Anderson-Darling	A-Sq	1.240103	Pr > A-Sq	<0.0050

Quantiles (Definition 5)

Quantile	Estimate
100% Max	7.6457891
99%	7.6457891
95%	4.2864924
90%	1.9656917
75% Q3	0.8390394
50% Median	-0.0918601
25% Q1	-1.4242096
10%	-2.1947087
5%	-2.5889086
1%	-3.1290087
0% Min	-3.1290087

The UNIVARIATE Procedure
Variable: rempg (Residual)

Extreme Observations

-----Lowest-----		-----Highest-----	
Value	Obs	Value	Obs
-3.12901	25	2.25929	26
-2.94641	34	2.38889	20
-2.58891	29	4.28649	43
-2.53001	2	6.13469	35
-2.37641	28	7.64579	13

The country "Other" is interesting. Residuals fail the tests for normality decisively, and it has that seeming outlier.

----- Country of Origin=Other -----

The UNIVARIATE Procedure
Variable: rempg (Residual)

Tests for Normality

Test	--Statistic--		-----p Value-----	
Shapiro-Wilk	W	0.635912	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.387375	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.310103	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.67891	Pr > A-Sq	<0.0050

Extreme Observations

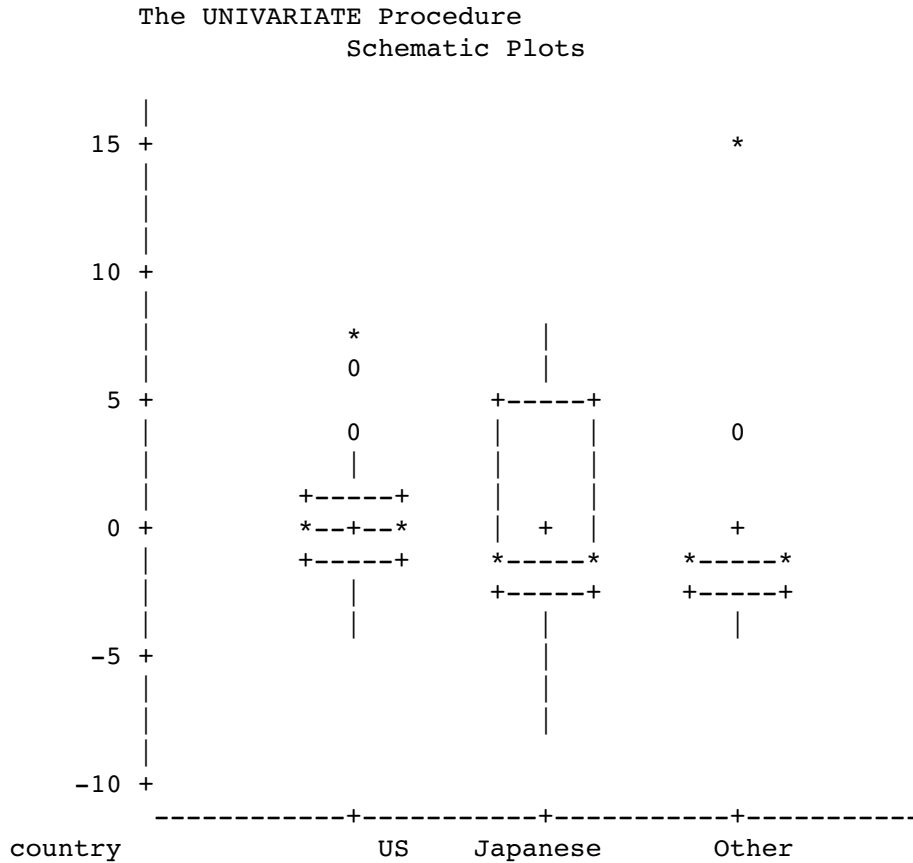
-----Lowest-----		-----Highest-----	
Value	Obs	Value	Obs
-4.03003	4	-1.197228	11
-3.43103	1	-1.001027	6
-2.55473	5	-0.949827	9
-2.42423	2	3.386372	3
-1.83293	10	15.378673	8

Watch out. It's not really observation 8. Remember, we sorted cases by country. But these commands let us locate it. (This output is a little out of order)

```
data peek;
  set resdata;
  if rempg > 10;
proc print;
```

z													
c				z				w			p		
o				w				e			r		
u				e				i			r		
n				i				g			e		
O				t				m			m		
b				r				p			p		
s				y				g			g		
1	Other	41	2040	-1.23844	1.53374	0	1	0	-1.23844	0	1.53374	25.6213	15.3787

One very good thing about proc univariate with a by option is that it produces side-by-side boxplots. Japan may have higher variance.



That was the end of proc univariate. Then came the preceding part, and now a regression with the quadratic term.

```
proc reg data = auto; /* Don't want to use the data set peek */
  model mpg = zweight zweight2 c1 c2 / r influence;
/*      i          prints (X'X)-inverse
  ssl      prints sequential sums of squares
  clm      prints confidence interval for E(Yh)
  cli      prints prediction interval for new observation
  r         prints residual analysis
  influence prints influence statistics
  partial  prints partial regression plots
*/
output out=resdata2 predicted=prempg residual=rempg student=studres
  cookd=cook h=leverage press=press rstudent=sdelres dffits=dffit;
```

We have seen the basic regression output before, so it is skipped. Here is the output caused by / r influence.

Output Statistics

Obs	Dep Var mpg	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual
1	22.0000	21.4584	0.5575	0.5416	3.044	0.178
2	17.0000	18.7021	0.5153	-1.7021	3.052	-0.558
3	22.0000	23.7174	0.5948	-1.7174	3.037	-0.565
4	17.0000	19.1466	1.0627	-2.1466	2.907	-0.739
5	23.0000	25.9467	0.9668	-2.9467	2.940	-1.002
6	25.0000	20.5768	1.0100	4.4232	2.925	1.512
7	20.0000	19.3031	0.5264	0.6969	3.050	0.229
8	15.0000	15.3621	0.6518	-0.3621	3.025	-0.120
9	18.0000	17.0113	0.4977	0.9887	3.055	0.324
10	26.0000	27.4071	0.7676	-1.4071	2.998	-0.469
11	20.0000	19.1192	0.5231	0.8808	3.050	0.289
12	16.0000	16.0940	0.5398	-0.0940	3.047	-0.0309
13	19.0000	18.4146	0.5099	0.5854	3.053	0.192
14	14.0000	14.6417	0.8982	-0.6417	2.962	-0.217
15	14.0000	16.0146	0.5475	-2.0146	3.046	-0.661
16	21.0000	14.7424	0.8513	6.2576	2.975	2.103
17	29.0000	28.5969	0.8632	0.4031	2.972	0.136
18	16.0000	16.9174	0.4990	-0.9174	3.054	-0.300
19	22.0000	19.7443	0.5340	2.2557	3.048	0.740
20	22.0000	19.4901	0.5297	2.5099	3.049	0.823
21	24.0000	22.8263	0.5768	1.1737	3.041	0.386
22	19.0000	18.2462	0.5069	0.7538	3.053	0.247
23	23.0000	24.8950	0.9505	-1.8950	2.945	-0.643

Output Statistics

Obs	-2 -1 0 1 2	Cook's D	RStudent	Hat Diag H	Cov Ratio	DFFITs
1		0.000	0.1766	0.0325	1.1093	0.0324
2	*	0.002	-0.5550	0.0277	1.0817	-0.0937
3	*	0.002	-0.5627	0.0369	1.0913	-0.1102
4	*	0.015	-0.7361	0.1179	1.1721	-0.2691
5	**	0.022	-1.0023	0.0976	1.1078	-0.3296
6	***	0.055	1.5265	0.1065	1.0173	0.5271
7		0.000	0.2269	0.0289	1.1036	0.0392
8		0.000	-0.1188	0.0444	1.1245	-0.0256
9		0.001	0.3216	0.0259	1.0959	0.0524
10		0.003	-0.4667	0.0615	1.1281	-0.1195
11		0.000	0.2868	0.0286	1.1007	0.0492
12		0.000	-0.0306	0.0304	1.1094	-0.0054
13		0.000	0.1904	0.0271	1.1028	0.0318
14		0.001	-0.2152	0.0842	1.1707	-0.0653
15	*	0.003	-0.6587	0.0313	1.0757	-0.1184

16		****	0.072	2.1581	0.0757	0.8356	0.6174
17			0.000	0.1347	0.0778	1.1649	0.0391
18			0.000	-0.2984	0.0260	1.0972	-0.0487
19		*	0.003	0.7375	0.0298	1.0654	0.1292
20		*	0.004	0.8212	0.0293	1.0548	0.1427
21			0.001	0.3836	0.0347	1.1025	0.0728
22			0.000	0.2452	0.0268	1.1005	0.0407
23	*		0.009	-0.6406	0.0943	1.1525	-0.2067

Output Statistics

Obs	-----DFBETAS-----				
	Intercept	zweight	zweight2	c1	c2
1	0.0026	-0.0077	-0.0165	0.0122	0.0005
2	-0.0246	-0.0177	0.0517	-0.0166	0.0010
3	0.0131	0.0599	0.0303	-0.0579	-0.0036
4	-0.0736	-0.0907	0.1116	0.0510	-0.1614
5	0.0500	0.0744	-0.0611	-0.0407	-0.2290
6	0.1119	0.1239	-0.1856	-0.0709	0.3363
7	0.0091	0.0035	-0.0226	0.0087	-0.0002
8	-0.0028	-0.0146	-0.0085	0.0012	0.0009
9	0.0149	0.0250	-0.0150	0.0026	-0.0015
10	0.0482	0.0990	-0.0277	-0.0758	-0.0060
11	0.0119	0.0059	-0.0281	0.0102	-0.0003
12	-0.0012	-0.0031	-0.0002	0.0001	0.0002
13	0.0087	0.0076	-0.0168	0.0050	-0.0004
14	0.0011	-0.0310	-0.0386	0.0034	0.0019
15	-0.0244	-0.0689	-0.0073	0.0022	0.0042
16	0.0009	0.3042	0.3454	-0.0326	-0.0185
17	-0.0180	-0.0334	0.0143	0.0247	0.0020
18	-0.0137	-0.0239	0.0127	-0.0021	0.0014
19	0.0266	0.0026	-0.0749	0.0327	-0.0001
20	0.0316	0.0086	-0.0826	0.0335	-0.0005
21	-0.0027	-0.0313	-0.0279	0.0343	0.0019
22	0.0114	0.0109	-0.0209	0.0058	-0.0006
23	-0.1869	-0.0239	0.0363	0.1705	0.1447

Now it starts again with observation 24 and displays the same material.

Output Statistics

Obs	Dep Var mpg	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual
24	35.0000	28.3350	0.9685	6.6650	2.939	2.267
25	24.0000	25.7391	0.9367	-1.7391	2.950	-0.590
26	21.0000	21.6394	1.0652	-0.6394	2.906	-0.220
27	30.0000	28.4959	0.8544	1.5041	2.975	0.506
28	18.0000	17.3509	0.4962	0.6491	3.055	0.212
29	16.0000	16.1342	0.5362	-0.1342	3.048	-0.0440

30	17.0000	16.6886	0.5045	0.3114	3.053	0.102
31	21.0000	25.3373	0.9501	-4.3373	2.945	-1.473
32	28.0000	31.9010	1.2162	-3.9010	2.846	-1.371
33	21.0000	23.6347	0.5929	-2.6347	3.038	-0.867
34	25.0000	26.1233	0.9339	-1.1233	2.951	-0.381
35	28.0000	31.1646	1.1484	-3.1646	2.874	-1.101
36	12.0000	13.8418	1.7103	-1.8418	2.579	-0.714
37	12.0000	13.9491	1.4859	-1.9491	2.715	-0.718
38	14.0000	16.2986	0.5235	-2.2986	3.050	-0.754
39	30.0000	28.7551	0.9856	1.2449	2.934	0.424
40	22.0000	24.2211	0.6082	-2.2211	3.034	-0.732
41	14.0000	15.4291	0.6371	-1.4291	3.029	-0.472
42	14.0000	15.2008	0.6919	-1.2008	3.017	-0.398
43	15.0000	16.7791	0.5020	-1.7791	3.054	-0.583
44	18.0000	18.5861	0.5131	-0.5861	3.052	-0.192
45	20.0000	22.2046	0.5673	-2.2046	3.042	-0.725
46	21.0000	15.4291	0.6371	5.5709	3.029	1.839

Output Statistics

Obs	-2 -1 0 1 2	Cook's D	RStudent	Hat Diag H	Cov Ratio	DFFITS
24	*****	0.112	2.3398	0.0979	0.8097	0.7709
25	*	0.007	-0.5868	0.0916	1.1547	-0.1863
26		0.001	-0.2185	0.1185	1.2160	-0.0801
27	*	0.004	0.5029	0.0762	1.1431	0.1445
28		0.000	0.2110	0.0257	1.1005	0.0343
29		0.000	-0.0437	0.0300	1.1089	-0.0077
30		0.000	0.1012	0.0266	1.1043	0.0167
31	**	0.045	-1.4854	0.0942	1.0124	-0.4791
32	**	0.069	-1.3797	0.1544	1.1081	-0.5896
33	*	0.006	-0.8658	0.0367	1.0571	-0.1690
34		0.003	-0.3783	0.0911	1.1711	-0.1198
35	**	0.039	-1.1029	0.1377	1.1417	-0.4407
36	*	0.045	-0.7115	0.3054	1.4923	-0.4718
37	*	0.031	-0.7154	0.2305	1.3466	-0.3916
38	*	0.003	-0.7512	0.0286	1.0626	-0.1289
39		0.004	0.4218	0.1014	1.1816	0.1417
40	*	0.004	-0.7295	0.0386	1.0762	-0.1462
41		0.002	-0.4692	0.0424	1.1053	-0.0987
42		0.002	-0.3956	0.0500	1.1194	-0.0907
43	*	0.002	-0.5798	0.0263	1.0779	-0.0953
44		0.000	-0.1907	0.0275	1.1032	-0.0321
45	*	0.004	-0.7221	0.0336	1.0714	-0.1346
46	***	0.030	1.8726	0.0424	0.8736	0.3939

Output Statistics

Obs	-----DFBETAS-----				
	Intercept	zweight	zweight2	c1	c2
24	0.4863	-0.1456	0.1786	-0.4922	-0.5156
25	-0.1594	-0.0076	0.0159	0.1480	0.1315
26	-0.0794	-0.0290	0.0321	0.0704	0.0513
27	-0.0657	-0.1232	0.0513	0.0912	0.0074
28	0.0100	0.0145	-0.0127	0.0026	-0.0009
29	-0.0017	-0.0044	-0.0001	0.0001	0.0003
30	0.0045	0.0087	-0.0032	0.0004	-0.0005
31	0.0510	0.0839	-0.0532	-0.0452	-0.3372
32	0.3175	0.4941	-0.3638	-0.3423	-0.0296
33	0.0187	0.0902	0.0483	-0.0880	-0.0055
34	-0.0993	-0.0007	0.0049	0.0930	0.0845
35	-0.1478	0.1628	-0.2337	0.1839	0.2431
36	0.0805	-0.1333	-0.3881	0.0115	0.0083
37	0.0565	-0.1257	-0.3092	0.0126	0.0078
38	-0.0307	-0.0728	0.0066	0.0001	0.0044
39	0.0832	-0.0314	0.0400	-0.0860	-0.0928
40	0.0240	0.0880	0.0303	-0.0808	-0.0053
41	-0.0119	-0.0566	-0.0300	0.0043	0.0034
42	-0.0074	-0.0503	-0.0358	0.0045	0.0031
43	-0.0261	-0.0487	0.0208	-0.0031	0.0029
44	-0.0086	-0.0067	0.0174	-0.0054	0.0004
45	-0.0025	0.0465	0.0604	-0.0578	-0.0028
46	0.0475	0.2259	0.1196	-0.0173	-0.0137

Now start with observaation 47

Output Statistics

Obs	Dep Var mpg	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual
47	19.0000	18.9983	0.5209	0.001711	3.051	0.00056
48	19.0000	18.9384	0.5197	0.0616	3.051	0.0202
49	18.0000	16.9174	0.4990	1.0826	3.054	0.354
50	19.0000	18.5861	0.5131	0.4139	3.052	0.136
51	24.0000	23.0652	0.5811	0.9348	3.040	0.308
52	16.0000	15.5321	0.6165	0.4679	3.033	0.154
53	14.0000	15.2441	1.1916	-1.2441	2.856	-0.436
54	28.0000	26.1743	0.6892	1.8257	3.017	0.605
55	34.0000	31.9010	1.2162	2.0990	2.846	0.738
56	25.0000	27.6999	0.7893	-2.6999	2.992	-0.902
57	26.0000	24.7372	0.6248	1.2628	3.031	0.417
58	18.0000	18.8196	0.5175	-0.8196	3.051	-0.269
59	18.0000	16.8709	0.4999	1.1291	3.054	0.370
60	18.0000	18.0266	0.5032	-0.0266	3.054	-0.0087

61	19.0000	19.5532	0.5308	-0.5532	3.049	-0.181
62	19.0000	19.6165	0.5318	-0.6165	3.049	-0.202
63	19.0000	18.3020	0.5079	0.6980	3.053	0.229
64	24.0000	23.3072	0.5858	0.6928	3.039	0.228
65	26.0000	28.5088	1.1124	-2.5088	2.888	-0.869
66	35.0000	28.0236	0.9580	6.9764	2.943	2.371
67	18.0000	24.5288	0.9592	-6.5288	2.942	-2.219
68	31.0000	26.5130	0.9335	4.4870	2.951	1.521
69	18.0000	22.2833	1.0382	-4.2833	2.916	-1.469

Output Statistics

Obs	-2	-1	0	1	2	Cook's D	RStudent	Hat H	Cov Ratio	DFFITs
47						0.000	0.000557	0.0283	1.1071	0.0001
48						0.000	0.0201	0.0282	1.1069	0.0034
49						0.001	0.3522	0.0260	1.0944	0.0575
50						0.000	0.1347	0.0275	1.1047	0.0226
51						0.001	0.3055	0.0353	1.1074	0.0584
52						0.000	0.1532	0.0397	1.1182	0.0311
53						0.007	-0.4330	0.1483	1.2457	-0.1806
54				*		0.004	0.6023	0.0496	1.1021	0.1376
55				*		0.020	0.7351	0.1544	1.2228	0.3142
56		*				0.011	-0.9010	0.0651	1.0843	-0.2377
57						0.001	0.4141	0.0408	1.1074	0.0854
58						0.000	-0.2668	0.0280	1.1009	-0.0452
59						0.001	0.3674	0.0261	1.0936	0.0601
60						0.000	-0.008644	0.0264	1.1049	-0.0014
61						0.000	-0.1801	0.0294	1.1057	-0.0314
62						0.000	-0.2008	0.0295	1.1052	-0.0350
63						0.000	0.2271	0.0269	1.1013	0.0378
64						0.000	0.2264	0.0358	1.1115	0.0436
65				*		0.022	-0.8671	0.1292	1.1692	-0.3340
66				****		0.119	2.4555	0.0958	0.7780	0.7993
67	****					0.105	-2.2858	0.0961	0.8219	-0.7452
68				***		0.046	1.5356	0.0910	0.9980	0.4858
69		**				0.055	-1.4818	0.1125	1.0340	-0.5276

Output Statistics

Obs	Intercept	zweight	zweight2	c1	c2
47	0.0000	0.0000	-0.0001	0.0000	-0.0000
48	0.0009	0.0005	-0.0019	0.0007	-0.0000
49	0.0162	0.0282	-0.0150	0.0025	-0.0017
50	0.0061	0.0047	-0.0123	0.0038	-0.0003
51	-0.0034	-0.0270	-0.0208	0.0284	0.0016
52	0.0043	0.0180	0.0081	-0.0013	-0.0011
53	-0.0639	-0.1038	0.0682	0.0560	-0.0936
54	-0.0448	-0.1061	0.0096	0.0854	0.0064
55	-0.1692	-0.2633	0.1938	0.1824	0.0158

56	0.0995	0.1990	-0.0635	-0.1508	-0.0120
57	-0.0179	-0.0560	-0.0115	0.0492	0.0034
58	-0.0116	-0.0077	0.0253	-0.0084	0.0004
59	0.0168	0.0299	-0.0148	0.0024	-0.0018
60	-0.0004	-0.0004	0.0007	-0.0002	0.0000
61	-0.0068	-0.0016	0.0182	-0.0075	0.0001
62	-0.0075	-0.0014	0.0203	-0.0085	0.0001
63	0.0105	0.0097	-0.0196	0.0056	-0.0006
64	-0.0035	-0.0215	-0.0143	0.0219	0.0013
65	0.1044	0.1311	-0.1553	-0.0735	-0.2056
66	0.5296	-0.1304	0.1545	-0.5286	-0.5418
67	-0.6862	-0.1093	0.1572	0.6225	0.5181
68	0.3885	-0.0144	0.0026	-0.3677	-0.3419
69	-0.5187	-0.1682	0.1965	0.4606	0.3448

Output Statistics

Obs	Dep Var mpg	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual
70	25.0000	27.4171	1.0357	-2.4171	2.916	-0.829
71	41.0000	26.2561	0.9779	14.7439	2.936	5.021
72	25.0000	26.7786	1.0008	-1.7786	2.929	-0.607
73	23.0000	25.0372	0.9444	-2.0372	2.947	-0.691
74	17.0000	16.7508	1.1475	0.2492	2.874	0.0867

Output Statistics

Obs	-2 -1 0 1 2	Cook's D	RStudent	Hat Diag H	Cov Ratio	DFFITS
70	*	0.017	-0.8269	0.1120	1.1523	-0.2937
71	*****	0.559	6.2575	0.0998	0.1230	2.0840
72	*	0.009	-0.6045	0.1046	1.1696	-0.2066
73	*	0.010	-0.6886	0.0931	1.1457	-0.2206
74		0.000	0.0861	0.1375	1.2465	0.0344

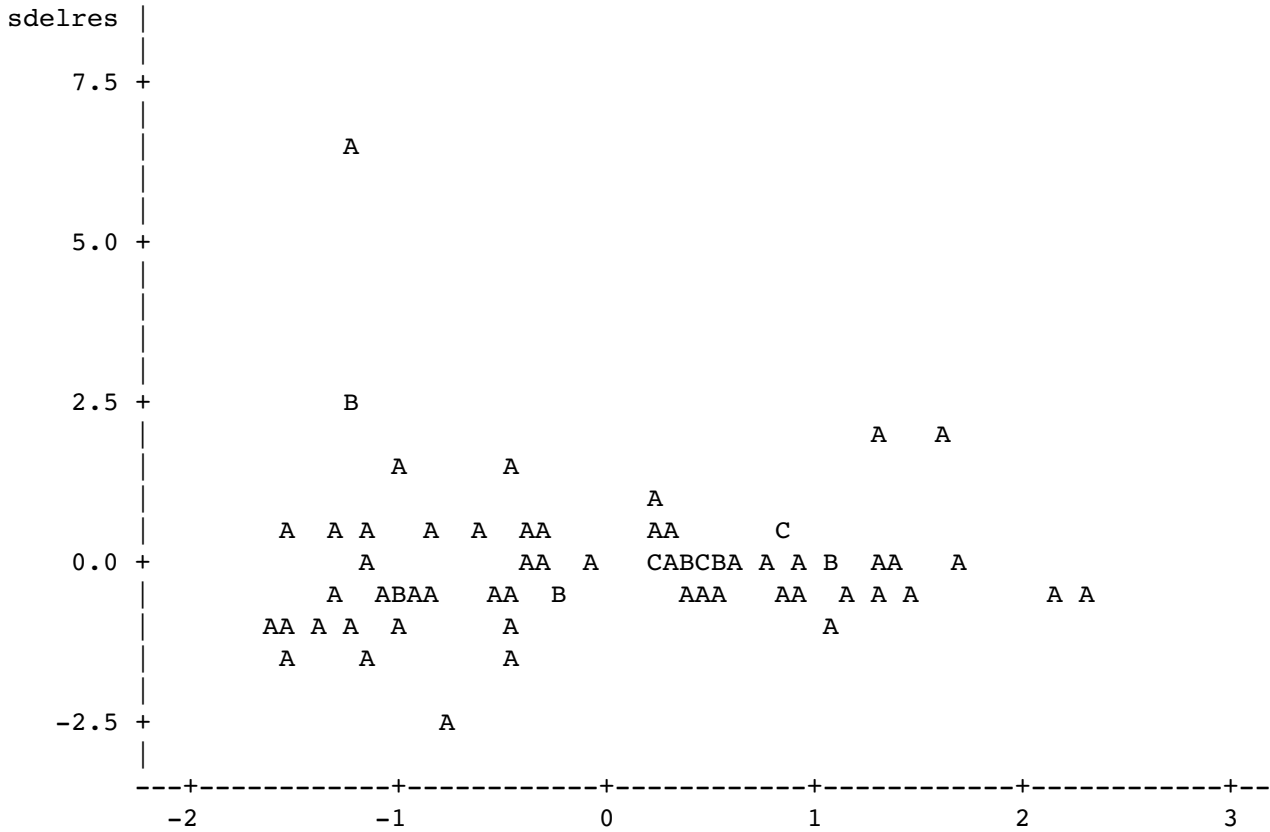
Output Statistics

Obs	Intercept	zweight	zweight2	c1	c2
70	0.0738	0.0974	-0.1044	-0.0542	-0.1926
71	-0.3624	-0.5213	0.4641	0.2862	1.4348
72	0.0434	0.0597	-0.0587	-0.0330	-0.1395
73	0.0185	0.0329	-0.0164	-0.0176	-0.1559
74	0.0116	0.0167	-0.0149	-0.0092	0.0187

Sum of Residuals 0
Sum of Squared Residuals 660.88273
Predicted Residual SS (PRESS) 805.50858

```
proc plot;
  plot sdelres*zweight;
```

Plot of sdelres*zweight. Legend: A = 1 obs, B = 2 obs, etc.



proc means output

Standardized Weight

Analysis Variable : rempg Residual

Country of Origin	N Obs	N	Mean	Std Dev	Minimum	Maximum
US	52	52	-2.25461E-15	1.8614966	-3.9009944	6.2575559
Japanese	11	11	-1.58257E-14	4.4050845	-6.5287582	6.9764381
Other	11	11	-3.22974E-15	5.3861988	-4.3372561	14.7439395

```

data peek2;
  set resdata2;  if sdelres**2 > 4;
proc print;

```

Obs	country	mpg	weight	zweight	zweight2	c1	c2	wc1
1	US	21	4290	1.63184	2.66289	1	0	1.63184
2	Japanese	35	2020	-1.26396	1.59759	0	0	0.00000
3	Japanese	35	2050	-1.22569	1.50231	0	0	0.00000
4	Japanese	18	2410	-0.76644	0.58743	0	0	0.00000
5	Other	41	2040	-1.23844	1.53374	0	1	0.00000

Obs	wc2	w2c1	w2c2	prempg	rempg	studres
1	0.00000	2.66289	0.00000	14.7424	6.2576	2.10306
2	0.00000	0.00000	0.00000	28.3350	6.6650	2.26746
3	0.00000	0.00000	0.00000	28.0236	6.9764	2.37065
4	0.00000	0.00000	0.00000	24.5288	-6.5288	-2.21884
5	-1.23844	0.00000	1.53374	26.2561	14.7439	5.02131

Obs	cook	leverage	press	sdelres	dfit
1	0.07240	0.075659	6.7698	2.15807	0.61742
2	0.11162	0.097922	7.3885	2.33983	0.77091
3	0.11911	0.095813	7.7157	2.45553	0.79934
4	0.10465	0.096069	-7.2226	-2.28575	-0.74517
5	0.55933	0.099844	16.3793	6.25751	2.08403

```

data fix;
  set auto;
  if _n_ ne 71; /* Delete observation 71 */

```

```

proc reg;
  model mpg = zweight zweight2 c1 c2;
  output out=resdata3 predicted=prempg;
  USvsOt2: test c1=c2;
  countr2: test c1=c2=0;

```

The REG Procedure
 Dependent Variable: mpg Miles per Gallon

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1630.55808	407.63952	66.10	<.0001
Error	68	419.38713	6.16746		
Corrected Total	72	2049.94521			

Root MSE	2.48344	R-Square	0.7954
Dependent Mean	21.02740	Adj R-Sq	0.7834
Coeff Var	11.81048		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	20.02737	0.92195	21.72	<.0001
zweight	Standardized Weight	1	-5.30160	0.37121	-14.28	<.0001
zweight2	Zweight Squared	1	0.92625	0.29314	3.16	0.0024
c1		1	0.90967	0.96996	0.94	0.3516
c2		1	-3.39304	1.08839	-3.12	0.0027

The REG Procedure
 Model: MODEL1

Test USVSOT2 Results for Dependent Variable mpg

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	124.41494	20.17	<.0001
Denominator	68	6.16746		

The REG Procedure
 Model: MODEL1

Test COUNTR2 Results for Dependent Variable mpg

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	63.90303	10.36	0.0001
Denominator	68	6.16746		