

SAS Log

```
274 libname in v8 'c:\data\mysas';
NOTE: Libname IN refers to the same physical library as MYLIB.
NOTE: Libref IN was successfully assigned as follows:
      Engine:          V8
      Physical Name:  C:\Data\mysas
275
```

```
NOTE: PROCEDURE REG used:
      real time          5:31.40
      cpu time           0.06 seconds
```

```
276 data temp; set in.gss98;
277
278 if sex=2 then sex=0;
279 if age=99 then age=.;
280 if educ=98 or educ=99 then educ=.;
281 age2=age**2; /*creating age-squared var*/
282
283 rrincom=.;
284 if rincom98=1 then rrincom=500;
285 else if rincom98=2 then rrincom=2000;
286 else if rincom98=3 then rrincom=3500;
287 else if rincom98=4 then rrincom=4500;
288 else if rincom98=5 then rrincom=5500;
289 else if rincom98=6 then rrincom=6500;
290 else if rincom98=7 then rrincom=7500;
291 else if rincom98=8 then rrincom=9000;
292 else if rincom98=9 then rrincom=11250;
293 else if rincom98=10 then rrincom=13750;
294 else if rincom98=11 then rrincom=16250;
295 else if rincom98=12 then rrincom=18750;
296 else if rincom98=13 then rrincom=21250;
297 else if rincom98=14 then rrincom=23750;
298 else if rincom98=15 then rrincom=27500;
299 else if rincom98=16 then rrincom=32500;
300 else if rincom98=17 then rrincom=37500;
301 else if rincom98=18 then rrincom=45000;
302 else if rincom98=19 then rrincom=55000;
303 else if rincom98=20 then rrincom=67500;
304 else if rincom98=21 then rrincom=82500;
305 else if rincom98=22 then rrincom=99500;
306 else if rincom98=23 then rrincom=110000;
307 else rrincom=.;
308
309 lnrrincom=log(rrincom);
310
```

```
NOTE: Missing values were generated as a result of performing an operation on missing values.
      Each place is given by: (Number of times) at (Line):(Column).
```

```
      4 at 281:9      983 at 309:10
NOTE: There were 2832 observations read from the data set IN.GSS98.
NOTE: The data set WORK.TEMP has 2832 observations and 863 variables.
NOTE: DATA statement used:
      real time          3.82 seconds
      cpu time           0.48 seconds
```

```
311 proc freq;
312 table age educ sex rincom98 rrincom;
313 run;
```

```
NOTE: There were 2832 observations read from the data set WORK.TEMP.
NOTE: PROCEDURE FREQ used:
      real time          0.06 seconds
      cpu time           0.06 seconds
```

```
314
315 proc corr data=temp nomiss;
316 var age sex educ rrincom lnrrincom;
317 run;
```

```
NOTE: PROCEDURE CORR used:
      real time          0.07 seconds
      cpu time           0.03 seconds
```

```
318
319 proc reg data=temp simple;
320 model rrincom=age sex educ;
321 model lnrrincom=age sex educ;
322 run;
```

```
NOTE: 2832 observations read.
NOTE: 989 observations have missing values.
NOTE: 1843 observations used in computations.
```

SAS Listing

February 26, 2003 39

The SAS System

20:07 Wednesday,

The FREQ Procedure

Age of respondent

AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
18	6	0.21	6	0.21
19	36	1.27	42	1.49
20	36	1.27	78	2.76
21	48	1.70	126	4.46
22	37	1.31	163	5.76
23	42	1.49	205	7.25
24	37	1.31	242	8.56
25	50	1.77	292	10.33
26	61	2.16	353	12.48
27	50	1.77	403	14.25
28	77	2.72	480	16.97
29	57	2.02	537	18.99
30	53	1.87	590	20.86
31	56	1.98	646	22.84
32	66	2.33	712	25.18
33	74	2.62	786	27.79
34	83	2.93	869	30.73
35	79	2.79	948	33.52
36	71	2.51	1019	36.03
37	71	2.51	1090	38.54
38	63	2.23	1153	40.77
39	71	2.51	1224	43.28
40	70	2.48	1294	45.76
41	63	2.23	1357	47.98
42	63	2.23	1420	50.21
43	61	2.16	1481	52.37
44	67	2.37	1548	54.74
45	54	1.91	1602	56.65
46	52	1.84	1654	58.49
47	49	1.73	1703	60.22
48	51	1.80	1754	62.02
49	50	1.77	1804	63.79
50	42	1.49	1846	65.28
51	72	2.55	1918	67.82
52	40	1.41	1958	69.24
53	34	1.20	1992	70.44
54	37	1.31	2029	71.75
55	45	1.59	2074	73.34
56	42	1.49	2116	74.82

57	33	1.17	2149	75.99
58	34	1.20	2183	77.19
59	26	0.92	2209	78.11

The SAS System

The FREQ Procedure

Age of respondent

AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
60	31	1.10	2240	79.21
61	26	0.92	2266	80.13
62	33	1.17	2299	81.29
63	22	0.78	2321	82.07
64	28	0.99	2349	83.06
65	27	0.95	2376	84.02
66	23	0.81	2399	84.83
67	29	1.03	2428	85.86
68	24	0.85	2452	86.70
69	24	0.85	2476	87.55
70	30	1.06	2506	88.61
71	24	0.85	2530	89.46
72	26	0.92	2556	90.38
73	41	1.45	2597	91.83
74	20	0.71	2617	92.54
75	27	0.95	2644	93.49
76	20	0.71	2664	94.20
77	21	0.74	2685	94.94
78	12	0.42	2697	95.37
79	18	0.64	2715	96.00
80	15	0.53	2730	96.53
81	12	0.42	2742	96.96
82	13	0.46	2755	97.42
83	14	0.50	2769	97.91
84	7	0.25	2776	98.16
85	13	0.46	2789	98.62
86	9	0.32	2798	98.94
87	12	0.42	2810	99.36
88	6	0.21	2816	99.58
89	12	0.42	2828	100.00

Frequency Missing = 4

The SAS System

The FREQ Procedure

Highest year of school completed

EDUC	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	0.07	2	0.07
2	5	0.18	7	0.25
3	10	0.35	17	0.60
4	9	0.32	26	0.92
5	8	0.28	34	1.21
6	23	0.82	57	2.02
7	21	0.74	78	2.77
8	82	2.91	160	5.67
9	75	2.66	235	8.33
10	113	4.01	348	12.34
11	138	4.89	486	17.23
12	851	30.18	1337	47.41
13	270	9.57	1607	56.99
14	350	12.41	1957	69.40
15	146	5.18	2103	74.57
16	412	14.61	2515	89.18
17	86	3.05	2601	92.23
18	109	3.87	2710	96.10
19	41	1.45	2751	97.55
20	69	2.45	2820	100.00

Frequency Missing = 12

Respondents sex

SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1600	56.50	1600	56.50
1	1232	43.50	2832	100.00

The SAS System

20:07 Wednesday,

February 26, 2003 42

The FREQ Procedure

Respondents income

RINCOM98	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	42	2.06	42	2.06
2	42	2.06	84	4.13
3	40	1.97	124	6.10
4	27	1.33	151	7.42
5	32	1.57	183	9.00
6	43	2.11	226	11.11
7	41	2.02	267	13.13
8	63	3.10	330	16.22

9	110	5.41	440	21.63
10	89	4.38	529	26.01
11	92	4.52	621	30.53
12	106	5.21	727	35.74
13	127	6.24	854	41.99
14	111	5.46	965	47.44
15	160	7.87	1125	55.31
16	169	8.31	1294	63.62
17	107	5.26	1401	68.88
18	178	8.75	1579	77.63
19	109	5.36	1688	82.99
20	64	3.15	1752	86.14
21	37	1.82	1789	87.95
22	20	0.98	1809	88.94
23	40	1.97	1849	90.90
24	118	5.80	1967	96.71
98	26	1.28	1993	97.98
99	41	2.02	2034	100.00

Frequency Missing = 798
The SAS System

20:07 Wednesday,

February 26, 2003 43

The FREQ Procedure

rrincom	Frequency	Percent	Cumulative Frequency	Cumulative Percent
500	42	2.27	42	2.27
2000	42	2.27	84	4.54
3500	40	2.16	124	6.71
4500	27	1.46	151	8.17
5500	32	1.73	183	9.90
6500	43	2.33	226	12.22
7500	41	2.22	267	14.44
9000	63	3.41	330	17.85
11250	110	5.95	440	23.80
13750	89	4.81	529	28.61
16250	92	4.98	621	33.59
18750	106	5.73	727	39.32
21250	127	6.87	854	46.19
23750	111	6.00	965	52.19
27500	160	8.65	1125	60.84
32500	169	9.14	1294	69.98
37500	107	5.79	1401	75.77
45000	178	9.63	1579	85.40
55000	109	5.90	1688	91.29
67500	64	3.46	1752	94.75
82500	37	2.00	1789	96.76
99500	20	1.08	1809	97.84
110000	40	2.16	1849	100.00

The CORR Procedure

5 Variables: AGE SEX EDUC rrincom lnrincom

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
AGE	1843	40.36517	12.44517	74393	18.00000	84.00000	Age of respondent
SEX	1843	0.47585	0.49955	877.00000	0	1.00000	Respondents sex
EDUC	1843	13.75366	2.65580	25348	0	20.00000	Highest year of school completed
rrincom	1843	29539	22946	54440250	500.00000	110000	
lnrincom	1843	9.92956	1.00814	18300	6.21461	11.60824	

Pearson Correlation Coefficients, N = 1843
Prob > |r| under H0: Rho=0

	AGE	SEX	EDUC	rrincom	lnrincom
AGE Age of respondent	1.00000	0.02294 0.3249	0.00716 0.7588	0.23102 <.0001	0.21328 <.0001
SEX Respondents sex	0.02294 0.3249	1.00000	-0.02822 0.2259	0.24953 <.0001	0.22296 <.0001
EDUC Highest year of school completed	0.00716 0.7588	-0.02822 0.2259	1.00000	0.36439 <.0001	0.31902 <.0001
rrincom	0.23102 <.0001	0.24953 <.0001	0.36439 <.0001	1.00000	0.81754 <.0001
lnrincom	0.21328 <.0001	0.22296 <.0001	0.31902 <.0001	0.81754 <.0001	1.00000

The REG Procedure

Descriptive Statistics

Variable	Sum	Mean	Uncorrected SS	Variance	Standard Deviation	Label
Intercept	1843.00000	1.00000	1843.00000	0	0	Intercept
AGE	74393	40.36517	3288179	154.88233	12.44517	Age of respondent
SEX	877.00000	0.47585	877.00000	0.24955	0.49955	Respondents sex

EDUC	25348	13.75366	361620	7.05329	2.65580	Highest year of school completed
rrincom	54440250	29539	2.577953E12	526518265	22946	
lnrincom	18300	9.92956	183585	1.01635	1.00814	

The SAS System

20:07 Wednesday,

February 26, 2003 46

The REG Procedure
 Model: MODEL1
 Dependent Variable: rrincom

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2.422922E11	80764073764	204.14	<.0001
Error	1839	7.275544E11	395625026		
Corrected Total	1842	9.698466E11			

Root MSE	19890	R-Square	0.2498
Dependent Mean	29539	Adj R-Sq	0.2486
Coeff Var	67.33597		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	-36559	2899.79948	-12.61	<.0001
AGE	Age of respondent	1	410.28859	37.24978	11.01	<.0001
SEX	Respondents sex	1	11707	928.33695	12.61	<.0001
EDUC	Highest year of school completed	1	3196.69181	174.57714	18.31	<.0001

The SAS System

20:07 Wednesday,

February 26, 2003 47

The REG Procedure
 Model: MODEL2
 Dependent Variable: lnrincom

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	370.54353	123.51451	151.27	<.0001
Error	1839	1501.57421	0.81652		
Corrected Total	1842	1872.11775			

Root MSE	0.90361	R-Square	0.1979
Dependent Mean	9.92956	Adj R-Sq	0.1966
Coeff Var	9.10024		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	7.34707	0.13174	55.77	<.0001
AGE	Age of respondent	1	0.01667	0.00169	9.85	<.0001
SEX	Respondents sex	1	0.45887	0.04217	10.88	<.0001
EDUC	Highest year of school completed	1	0.12298	0.00793	15.51	<.0001