

```
* SAS Analysis Examples Replication for ASDA 3rd Edition, SAS v9.4 / SAS Stat 15.3 ;
* Berglund Fall 2024
* Chapter 5 ;
```

```
libname d "P:\ASDA3\Data Sets for Analysis Examples and Stata R Code" ;
```

```
ods listing ;
ods graphics off ;
options nodate nonumber ;
```

```
ods rtf style=minimal bodytitle ;
```

```
data c5_nhanes ;
  set d.nhanes1112;
  if age >= 18 and age ne . then age18p=1 ; else age18p=0 ;
  int wtme2yr = int(wtme2yr) ;
  female=0 ;
  if riagendr=2 then female = 1 ;
  if age > 45 then age45=1 ; else age45=0 ;
run ;
```

```
title "Example 5.1: Weighted histogram of cholesterol" ;
proc sgplot ;
  where age18p=1 ;
  histogram lbxtc / weight = wtme2yr ;
run ;
```

```
title "Example 5.2: Weighted boxplot of cholesterol by gender" ;
proc sgplot ;
  where age18p=1 ;
  vbox lbxtc / category=female weight=wtme2yr ;
run;
```

```
title "Example 5.3 : Population totals using NCSR data " ;
```

```
data c5_ncsr ;
  set d.ncsr ;
  * create variables needed for NCSR examples ;
  ncsrwtsh_pop = ncsrwtsh * (209128094 / 9282) ;
run ;
proc surveymeans sum stderr clsum ;
  strata sestrat ; cluster seclustr ; weight ncsrwtsh_pop ;
  var mde ;
run ;
proc surveymeans sum stderr clsum ;
  domain mar3cat ;
  strata sestrat ; cluster seclustr ; weight ncsrwtsh_pop ;
  var mde ;
run ;
```

```
data c5_hrs ;
  set d.hrs12 ;
  if nfinr=1 then finr=1 ; else if nfinr=0 then finr=0 ; else finr=. ;
  if gender=2 then female=1 ; else female=0 ;
  if nage >=70 then age70=1 ; else age70=0 ;
run ;
```

```
title "Figure 5.1 Scatter Plot of Weight and HH Total Income" ;
proc sgplot data=c5_hrs ;
  scatter y=nwgtth x=hllatota ;
run ;
```

```
title "Example 5.4 : Total HH Wealth using HRS 2012 data " ;
```

```
proc surveymeans all ;
  strata secu ; cluster stratum ; weight nwgtth ;
  var hllatota ;
  where finr=1 ;
run ;
```

```
title "Example 5.5: Estimating the Mean Value of Household Income using the 2012 HRS Data." ;
```

```
proc surveymeans all ;
  strata secu ; cluster stratum ; weight nwgtth ;
  var hllitot ;
  where finr=1 ;
run ;
```

```
title "Example 5.6: Estimating Mean Systolic Blood Pressure using the NHANES Data." ;
```

```
proc surveymeans data=c5_nhanes ;
  strata sdmvstra ; cluster sdmvpsu ; weight wtme2yr ;
  var bpxs1 ;
  domain age18p ;
```

```

run ;

title "Example 5.7: Estimating the Mean Value of Total Household Wealth using the HRS Data." ;
proc surveymeans data=c5_hrs ;
  strata stratum ; cluster secu ; weight nwgthh ;
  var hllatota ;
  where finr=1 ;
run ;

title "Example 5.8 Weighted HH Total Income" ;
proc surveymeans data=c5_hrs all ;
  weight nwgthh ;
  var hllatota ;
run ;

title "Example 5.9: Estimating Population Quantiles for Total Household Wealth Using the HRS Data." ;
proc surveymeans data=c5_hrs all ;
  strata stratum ; cluster secu ; weight nwgthh ;
  domain finr ;
  var hllatota ;
run ;

ods text="Example 5.10: Estimating the Lorenz Curve and Gini Coefficient for the 2012 HRS Population Distribution of
Total Household Wealth. Not available in SAS SURVEY procedures" ;

title "Figure 5.5 Weighted Scatter Plot with Regression Line for Total and High-Density Cholesterol Measures in the
2011-2012 NHANES" ;
proc sgplot data=c5_nhanes ;
  reg y=lbdhdd x=lbxtc / weight=wtmec2yr ;
  where age18p=1 ;
run ;

title "Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-
2012 NHANES." ;
proc means data=c5_nhanes n mean stddev min max vardef=wdf ;
  where age18p=1 ;
  var lbdhdd lbxtc ;
  weight wtmec2yr ;
run ;
proc corr data=c5_nhanes ;
  where age18p=1 ;
  var lbdhdd lbxtc ;
  weight wtmec2yr ;
run ;

data c5_nhanes_1 ;
  set c5_nhanes ;
  stdlbxtc=(lbxtc - 194.4355)/41.05184 ;
  stdlbdhdd = (lbdhdd - 52.83826) / 14.93157 ;
run ;

proc surveyreg data=c5_nhanes_1 ;
  strata sdmvstra ; cluster sdmvpsu ; weight wtmec2yr ;
  domain age18p ;
  model stdlbdhdd = stdlbxtc ;
run ;

title "Example 5.12: Estimating the Population Ratio of High Density to Total Cholesterol for U.S. Adults. " ;
proc surveymeans data=c5_nhanes ; * note need stat 15.1 for deff for mean but not ratio option ;
  strata sdmvstra ; cluster sdmvpsu ; weight wtmec2yr ;
  domain age18p ;
  ratio lbdhdd/lbxtc ;
run ;

title "Example 5.13: Estimating the Proportions of Males and Females Age >= 70 with Diabetes Using the HRS Data." ;
proc surveymeans data=c5_hrs ; * try using nocollapse for strata ;
  strata stratum ; cluster secu ; weight nwgtr ;
  domain age70 * gender ;
  var diabetes ;
run ;

title "Example 5.14: Estimating Mean Systolic Blood Pressure for Males and Females Age > 45 using the 2011-2012 NHANES
data." ;
proc surveymeans data=c5_nhanes mean stderr sumwgt deff ;
  strata sdmvstra ; cluster sdmvpsu ; weight wtmec2yr ;
  domain age45*riagendr ;
  var bpxsy1 ;
run ;

```

```

title "Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level." ;
proc surveymeans data=c5_hrs ;
  strata stratum ; cluster secu ; weight nwgthh ;
  domain finr * edcat / diffmeans ;
  var hllatota ;
run ;
* alternate approach for diff in means ;
proc surveyreg data=c5_hrs ;
  strata stratum ; cluster secu ; weight nwgthh ;
  domain finr ;
  class edcat ;
  model hllatota=edcat / solution ;
  lsmeans edcat / diff ;
run ;

title "Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the HRS study. " ;

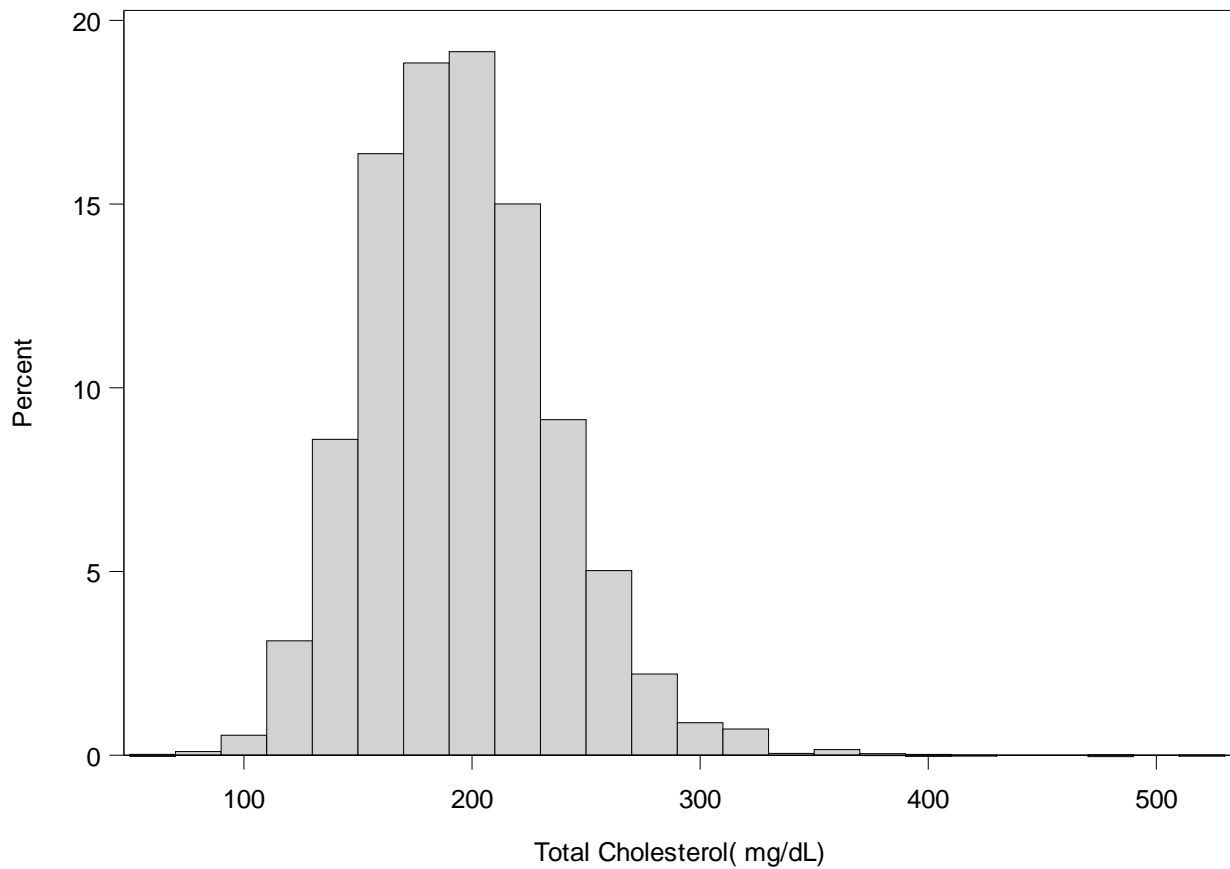
data hrs_2010_2012_c5 ;
  set d.hrs_2010_2012_both ;
  * prepare data for analysis ;
  hhweight = mwgthh ; if year=2012 then hhweight = nwgthh ;
  totwealth=h10atota ; if year=2012 then totwealth=hllatota ;
  finr2010 = 0 ; if (year = 2010 & mfinr = 1) then finr2010=1 ;
  finr2012 = 0 ; if (year = 2012 & nfinr = 1) then finr2012=1 ;
  finr2010_2012 = 0 ; if finr2010 = 1 | finr2012 = 1 then finr2010_2012=1 ;
run ;

proc surveyreg data=hrs_2010_2012_c5 ;
  strata stratum ; cluster secu ; weight hhweight ;
  domain finr2010_2012 ;
  class year ;
  model totwealth=year / solution ;
  lsmeans year / diff ;
run ;

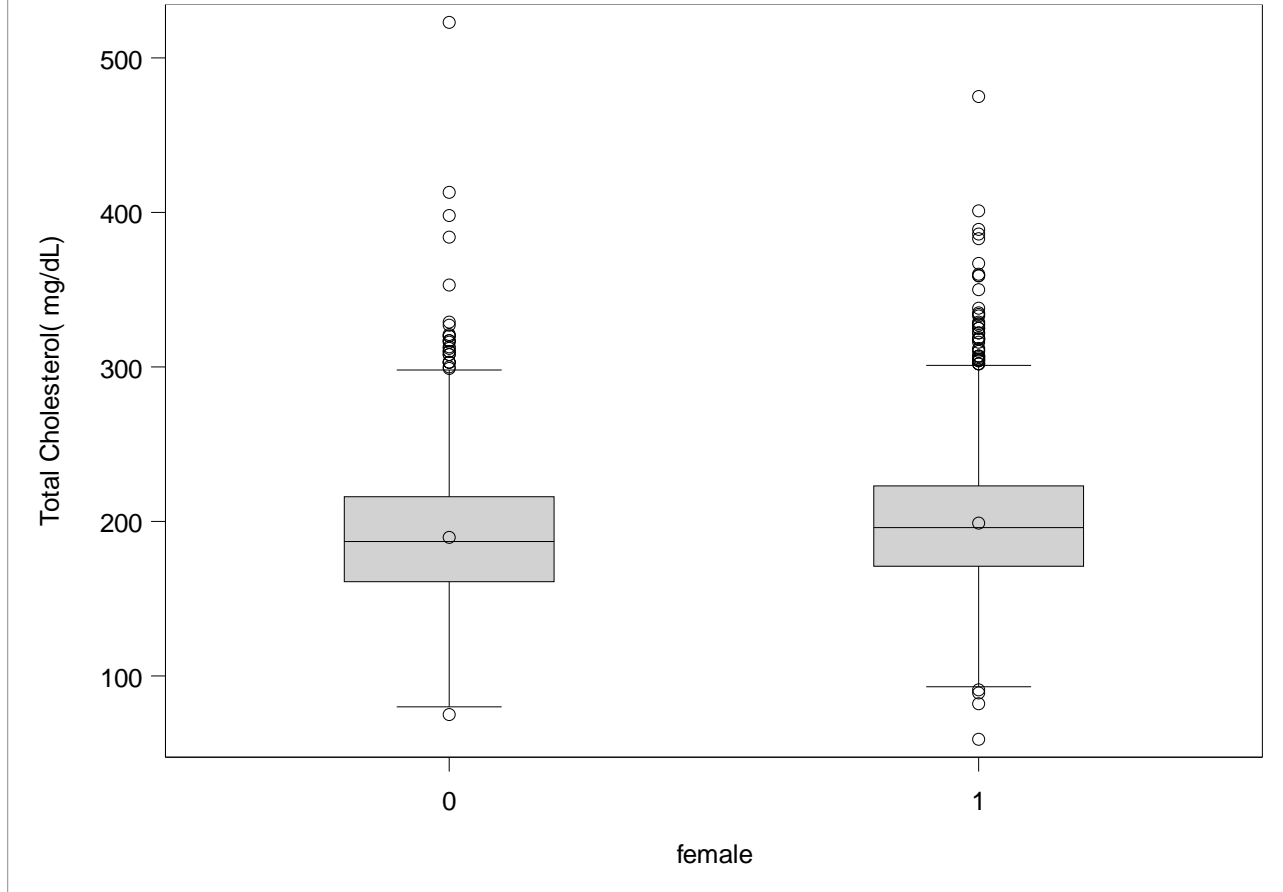
ods rtf close ;

```

Example 5.1: Weighted histogram of cholesterol



Example 5.2: Weighted boxplot of cholesterol by gender



Example 5.3 : Population totals using NCSR data

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	42
Number of Clusters	84
Number of Observations	9282
Sum of Weights	209128097

Statistics						
Variable	Label	Std Error of Mean	Sum	Std Error of Sum	95% CL for Sum	
mde	Major Depressive Episode 1=Yes 0=No	0.004877	40092207	2567488	34910806.0	45273607.0

Example 5.3 : Population totals using NCSR data

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	42
Number of Clusters	84
Number of Observations	9282
Sum of Weights	209128097

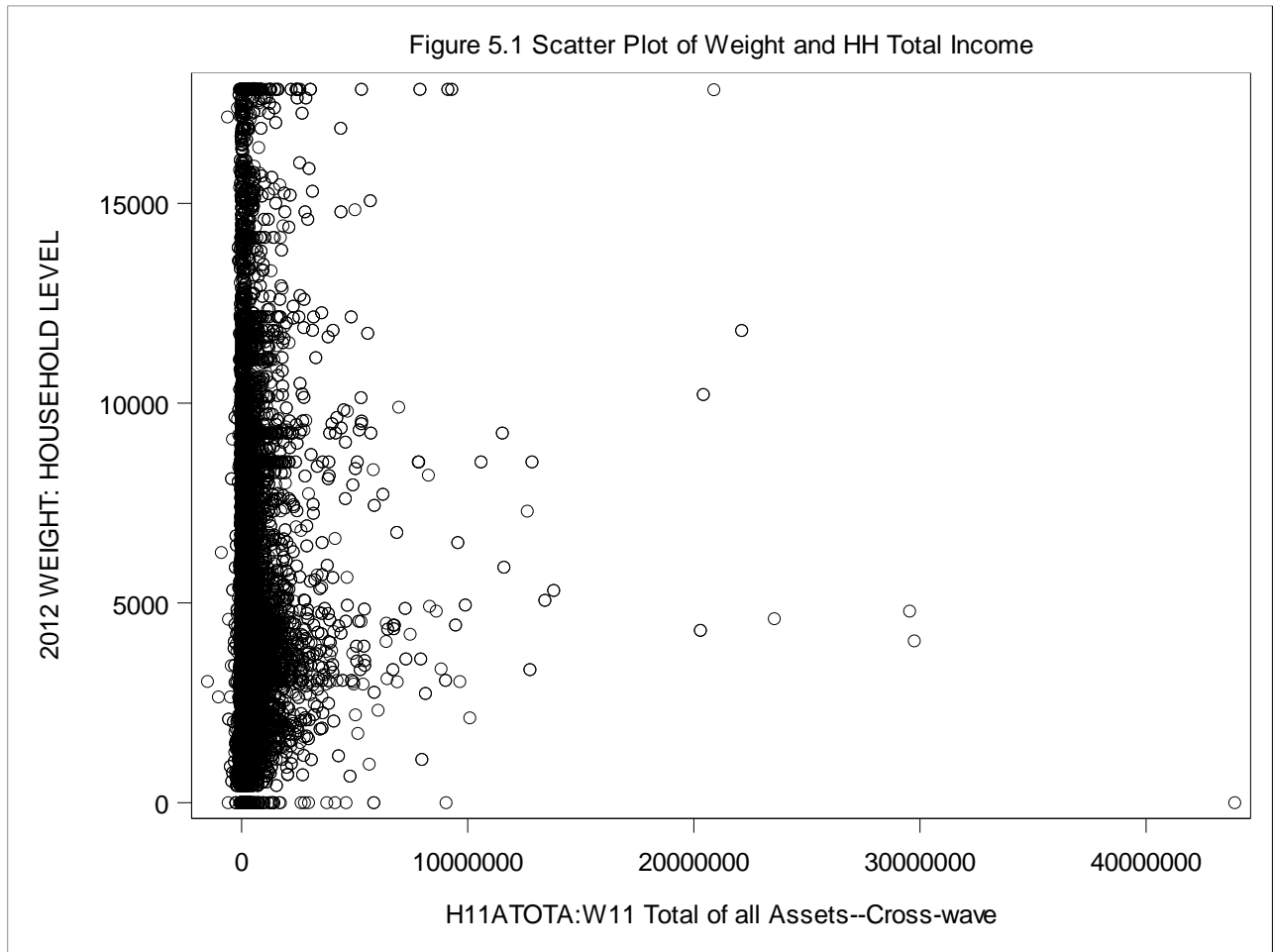
Statistics						
Variable	Label	Std Error of Mean	Sum	Std Error of Sum	95% CL for Sum	
mde	Major Depressive Episode 1=Yes 0=No	0.004877	40092207	2567488	34910806.0	45273607.0

Example 5.3 : Population totals using NCSR data

The SURVEYMEANS Procedure

Statistics for MAR3CAT Domains							
MAR3CAT	Variable	Label	Std Error of Mean	Sum	Std Error of Sum	95% CL for Sum	
1	mde	Major Depressive Episode 1=Yes 0=No	0.006415	20304191	1584109	17107329.8	23501051.2
2	mde	Major Depressive Episode 1=Yes 0=No	0.010533	10360671	702622	8942723.0	11778618.3
3	mde	Major Depressive Episode 1=Yes 0=No	0.010247	9427345	773138	7867090.6	10987600.2

Figure 5.1 Scatter Plot of Weight and HH Total Income



Example 5.4 : Total HH Wealth using HRS 2012 data

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	2
Number of Clusters	112
Number of Observations	14191
Number of Observations Used	13657
Number of Obs with Nonpositive Weights	534
Sum of Weights	58969863

Statistics											
Variable	Label	N	N Miss	Minimum	Maximum	Range	Clusters	Sum of Weights	DF	Mean	Std Error of Mean
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	13657	0	-1510000	29748000	31258000	112	58969863	110	428471	18367

Statistics									
Variable	Var of Mean	t Value	Pr > t	95% CL for Mean		Lower 95% One-Sided CL for Mean	Upper 95% One-Sided CL for Mean	Coeff of Variation	Sum
H11ATOTA	337347647	23.33	<.0001	392071.621	464869.890	398003	458938	0.042866	2.5266862E13

Statistics								
Variable	Std Error of Sum	Var of Sum	95% CL for Sum		Lower 95% One-Sided CL for Sum	Upper 95% One-Sided CL for Sum	Coeff of Variation for Sum	Design Effect
H11ATOTA	1.7154555E12	2.9427875E24	2.18672E13	2.86665E13	2.2421223E13	2.8112501E13	0.067893	3.714245

Quantiles						
Variable	Label	Percentile	Estimate	Std Error	95% Confidence Limits	
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	0	Min	-1510000	.	.
		10	D1	-24.278656	245.070299	-509.950 461.39
		20	D2	8943.267582	1530.727914	5909.724 11976.81
		25	Q1	21953	2733.354219	16536.594 27370.33
		30	D3	40001	3512.739696	33039.810 46962.66
		40	D4	83400	5237.092350	73021.390 93778.77
		50	Median	141907	9189.419106	123695.282 160117.83
		60	D6	223767	12216	199556.712 247976.49
		70	D7	345526	18776	308316.963 382734.56
		75	Q3	439965	20824	398697.591 481232.81
		80	D8	547396	24305	499229.032 595562.68
		90	D9	1018845	49816	920121.766 1117567.96
		100	Max	29748000	.	.

Example 5.5: Estimating the Mean Value of Household Income using the 2012 HRS Data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	2
Number of Clusters	112
Number of Observations	14191
Number of Observations Used	13657
Number of Obs with Nonpositive Weights	534
Sum of Weights	58969863

Statistics												
Variable	Label	N	N Miss	Minimum	Maximum	Range	Clusters	Sum of Weights	DF	Mean	Std Error of Mean	Var of Mean
H11ITOT	H11ITOT:W11 Incm: Total HHold / R+Sp only	13657	0	0	3666240	3666240	112	58969863	110	71382	2272.517174	5164334

Statistics									
Variable	t Value	Pr > t	95% CL for Mean		Lower 95% One-Sided CL for Mean	Upper 95% One-Sided CL for Mean	Coeff of Variation	Sum	Std Error of Sum
H11ITOT	31.41	<.0001	66878.8082	75885.9995	67613	75152	0.031836	4.2094106E12	256779341179

Statistics							
Variable	Var of Sum	95% CL for Sum		Lower 95% One-Sided CL for Sum	Upper 95% One-Sided CL for Sum	Coeff of Variation for Sum	Design Effect
H11ITOT	6.593563E22	3.70053E12	4.71829E12	3.7834588E12	4.6353624E12	0.061001	4.727330

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
H11ITOT	H11ITOT:W11 Incm: Total HHold / R+Sp only	0	Min	0	.	.	.
		10	D1	9635.570320	320.258945	9000.892	10270.248
		20	D2	14962	365.521885	14237.262	15686.019
		25	Q1	17899	446.424220	17014.660	18784.076
		30	D3	20979	618.466415	19752.897	22204.208
		40	D4	28951	777.173049	27411.055	30491.404
		50	Median	38604	1229.742145	36167.247	41041.367
		60	D6	51613	1500.495192	48639.173	54586.431
		70	D7	69644	1961.121046	65757.789	73530.752
		75	Q3	80991	2622.038737	75794.625	86187.156
		80	D8	98510	3082.560797	92401.027	104618.851
		90	D9	155669	6132.432234	143515.742	167821.824
		100	Max	3666240	.	.	.

Example 5.6: Estimating Mean Systolic Blood Pressure using the NHANES Data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	14
Number of Clusters	31
Number of Observations	9756
Number of Observations Used	9338
Number of Obs with Nonpositive Weights	418
Sum of Weights	306590681

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	6756	119.530464	0.546864	118.376682	120.684245

Example 5.6: Estimating Mean Systolic Blood Pressure using the NHANES Data.

The SURVEYMEANS Procedure

Statistics for age18p Domains							
age18p	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
0	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	1624	105.848426	0.285412	105.246260	106.450592
1	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	5132	122.029235	0.616339	120.728874	123.329597

Example 5.7: Estimating the Mean Value of Total Household Wealth using the HRS Data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	56
Number of Clusters	112
Number of Observations	14191
Number of Observations Used	13657
Number of Obs with Nonpositive Weights	534
Sum of Weights	58969863

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	13657	428471	17354	393706.970	463234.541

Example 5.8 Weighted HH Total Income

The SURVEYMEANS Procedure

Data Summary	
Number of Observations	20554
Number of Observations Used	19990
Number of Obs with Nonpositive Weights	564
Sum of Weights	89174512

Statistics											
Variable	Label	N	N Miss	Minimum	Maximum	Range	Sum of Weights	DF	Mean	Std Error of Mean	Var of Mean
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	19990	0	-1510000	29748000	31258000	89174512	19989	474129	11592	134381101

Statistics									
Variable	t Value	Pr > t	95% CL for Mean		Lower 95% One-Sided CL for Mean	Upper 95% One-Sided CL for Mean	Coeff of Variation	Sum	Std Error of Sum
H11ATOTA	40.90	<.0001	451406.978	496850.657	455060	493197	0.024450	4.2280206E13	1.0776493E12

Statistics							
Variable	Var of Sum	95% CL for Sum		Lower 95% One-Sided CL for Sum	Upper 95% One-Sided CL for Sum	Coeff of Variation for Sum	Design Effect
H11ATOTA	1.161328E24	4.01679E13	4.43925E13	4.0507548E13	4.4052863E13	0.025488	2.053595

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	0	Min	-1510000	.	.	.
		10	D1	-9.711356	129.337208	-263.22	243.80
		20	D2	15194	1172.765638	12894.95	17492.38
		25	Q1	34927	1653.481123	31686.14	38168.06
		30	D3	55055	1819.281983	51489.47	58621.35
		40	D4	103320	2319.635335	98773.46	107866.81
		50	Median	171881	3865.040021	164304.73	179456.33
		60	D6	259799	4716.662361	250553.80	269043.90
		70	D7	399704	7457.264377	385087.45	414321.16
		75	Q3	495103	8542.816023	478358.00	511847.25
		80	D8	614866	11799	591739.57	637992.76
		90	D9	1129958	22658	1085547.61	1174369.26
		100	Max	29748000	.	.	.

Example 5.9: Estimating Population Quantiles for Total Household Wealth Using the HRS Data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	56
Number of Clusters	112
Number of Observations	20554
Number of Observations Used	19990
Number of Obs with Nonpositive Weights	564
Sum of Weights	89174512

Statistics											
Variable	Label	N	N Miss	Minimum	Maximum	Range	Clusters	Sum of Weights	DF	Mean	Std Error of Mean
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	19990	0	-1510000	29748000	31258000	112	89174512	56	474129	19681

Statistics									
Variable	Var of Mean	t Value	Pr > t	95% CL for Mean		Lower 95% One-Sided CL for Mean	Upper 95% One-Sided CL for Mean	Coeff of Variation	Sum
H11ATOTA	387331023	24.09	<.0001	434703.583	513554.052	441212	507045	0.041509	4.2280206E13

Statistics								
Variable	Std Error of Sum	Var of Sum	95% CL for Sum		Lower 95% One-Sided CL for Sum	Upper 95% One-Sided CL for Sum	Coeff of Variation for Sum	Design Effect
H11ATOTA	2.3237887E12	5.3999941E24	3.76251E13	4.69353E13	3.8393617E13	4.6166794E13	0.054962	5.919145

Quantiles							
Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	0	Min	-1510000	.	.	.
		10	D1	-9.711356	228.299260	-467.05	447.63
		20	D2	15194	1965.354774	11256.59	19130.74
		25	Q1	34927	2735.378006	29447.48	40406.72
		30	D3	55055	4223.690214	46594.34	63516.48
		40	D4	103320	5823.215388	91654.83	114985.44
		50	Median	171881	9463.115669	152923.63	190837.43
		60	D6	259799	12060	235640.13	283957.58
		70	D7	399704	18926	361790.23	437618.38
		75	Q3	495103	18861	457319.50	532885.75
		80	D8	614866	26257	562267.75	667464.58
		90	D9	1129958	47679	1034445.98	1225470.89
		100	Max	29748000	.	.	.

Example 5.9: Estimating Population Quantiles for Total Household Wealth Using the HRS Data.

The SURVEYMEANS Procedure

Statistics for finr Domains												
finr	Variable	Label	N	N Miss	Minimum	Maximum	Range	Clusters	Sum of Weights	DF	Mean	Std Error of Mean
1	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	13657	0	-1510000	29748000	31258000	112	58969863	56	428471	17354

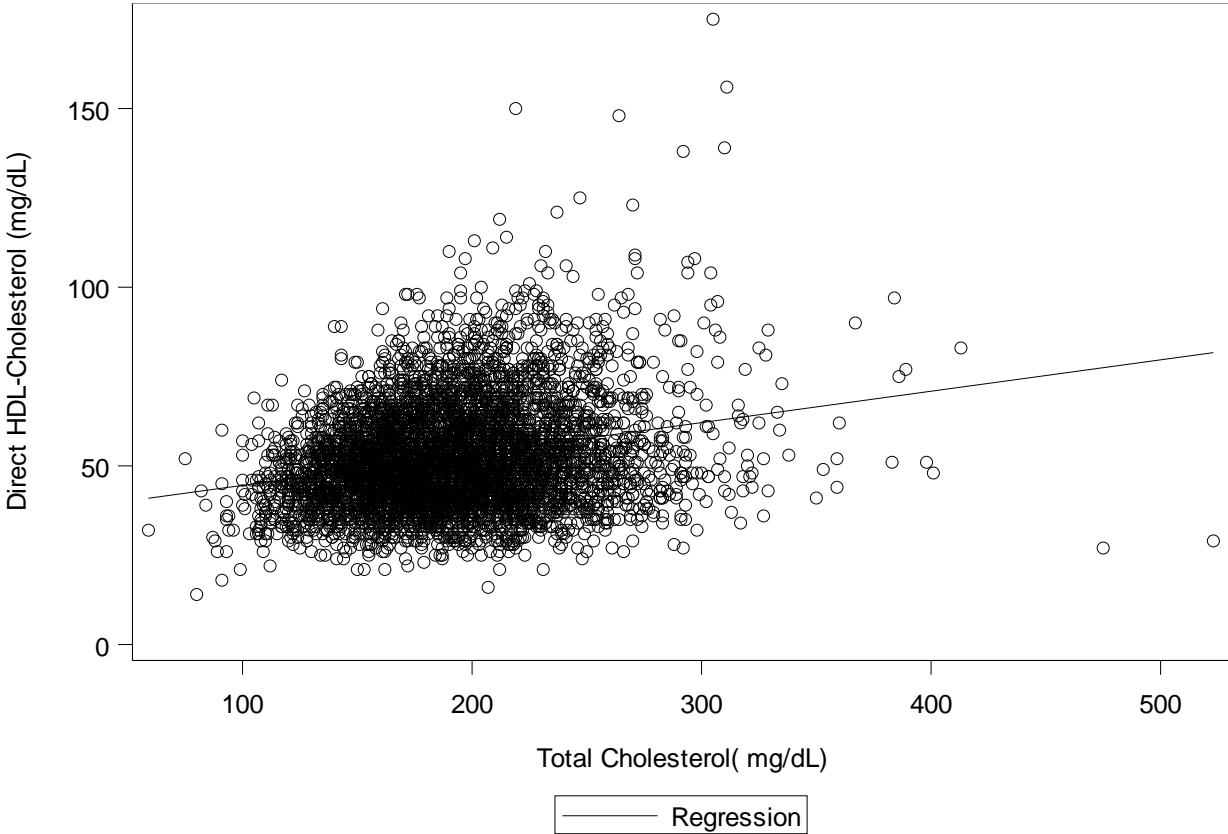
Statistics for finr Domains										
finr	Variable	Var of Mean	t Value	Pr > t	95% CL for Mean		Lower 95% One-Sided CL for Mean	Upper 95% One-Sided CL for Mean	Coeff of Variation	Sum
1	H11ATOTA	301153449	24.69	<.0001	393706.970	463234.541	399446	457495	0.040502	2.5266862E13

Statistics for finr Domains									
finr	Variable	Std Error of Sum	Var of Sum	95% CL for Sum		Lower 95% One-Sided CL for Sum	Upper 95% One-Sided CL for Sum	Coeff of Variation for Sum	Design Effect
1	H11ATOTA	1.3537104E12	1.8325319E24	2.25551E13	2.79787E13	2.3002751E13	2.7530973E13	0.053577	3.315742

Quantiles for finr Domains								
finr	Variable	Label	Percentile		Estimate	Std Error	95% Confidence Limits	
1	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	0	Min	-1510000	.	.	.
			10	D1	-24.278656	195.212051	-415.3	366.8
			20	D2	8943.267582	1284.307045	6370.5	11516.0
			25	Q1	21953	2214.513579	17517.3	26389.7
			30	D3	40001	3014.009018	33963.5	46039.0
			40	D4	83400	4533.978303	74317.4	92482.7
			50	Median	141907	7753.643211	126374.1	157439.0
			60	D6	223767	10417	202899.0	244634.2
			70	D7	345526	16408	312657.4	378394.1
			75	Q3	439965	18652	402600.2	477330.2
			80	D8	547396	21327	504673.4	590118.3
			90	D9	1018845	46214	926267.7	1111422.1
			100	Max	29748000	0	29748000.0	29748000.0

Example 5.10: Estimating the Lorenz Curve and Gini Coefficient for the 2012 HRS Population Distribution of Total Household Wealth. Not available in SAS SURVEY procedures

Figure 5.5 Weighted Scatter Plot with Regression Line for Total and High-Density Cholesterol Measures in the 2011-2012 NHANES



Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES.

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
lbdhdd	Direct HDL-Cholesterol (mg/dL)	5187	52.8382631	14.9301339	14.0000000	175.0000000
lbxtc	Total Cholesterol (mg/dL)	5187	194.4354654	41.0478781	59.0000000	523.0000000

Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES.

The CORR Procedure

2	Variables:	lbdhdd	lbxtc
	Weight Variable:	WTMEC2YR	

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
lbdhdd	5187	52.83826	3066	1.15527E10	14.00000	175.00000	Direct HDL-Cholesterol (mg/dL)
lbxtc	5187	194.43547	8428	4.25118E10	59.00000	523.00000	Total Cholesterol(mg/dL)

Pearson Correlation Coefficients, N = 5187 Prob > r under H0: Rho=0		
	lbdhdd	lbxtc
lbdhdd Direct HDL-Cholesterol (mg/dL)	1.00000	0.24144 <.0001
lbxtc Total Cholesterol(mg/dL)	0.24144 <.0001	1.00000

Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES.

The SURVEYREG Procedure

Regression Analysis for Dependent Variable stdlbdhdd

Data Summary	
Number of Observations	6988
Sum of Weights	260814685
Weighted Mean of stdlbdhdd	-0.0044356
Weighted Sum of stdlbdhdd	-1156879.7

Design Summary	
Number of Strata	14
Number of Clusters	31

Fit Statistics	
R-Square	0.05355
Root MSE	0.9396
Denominator DF	17

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	228.82	<.0001
Intercept	1	0.73	0.4040
stdlbxtc	1	228.82	<.0001

The denominator degrees of freedom for the F tests is 17.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	0.02742474	0.03204797	0.86	0.4040
stdlbxtc	0.22147079	0.01464098	15.13	<.0001

The degrees of freedom for the t tests is 17.

Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES.

The SURVEYREG Procedure

age18p=0

Domain Regression Analysis for Variable stldbdhdd

Domain Summary	
Number of Observations	6988
Number of Observations in Domain	1801
Number of Observations Not in Domain	5187
Sum of Weights in Domain	42172649
Weighted Mean of stldbdhdd	-0.02743
Weighted Sum of stldbdhdd	-1156924.9

Fit Statistics	
R-Square	0.05729
Root MSE	0.7419
Denominator DF	17

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	26.72	<.0001
Intercept	1	10.86	0.0043
stldbxtc	1	26.72	<.0001

The denominator degrees of freedom for the F tests is 17.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	0.21496136	0.06523494	3.30	0.0043
stldbxtc	0.27245142	0.05271004	5.17	<.0001

The degrees of freedom for the t tests is 17.

Example 5.11: Estimation of the Correlation of Adults' Total and High-Density Cholesterol Measures in the 2011-2012 NHANES.

The SURVEYREG Procedure

age18p=1

Domain Regression Analysis for Variable stdlbhdhd

Domain Summary	
Number of Observations	6988
Number of Observations in Domain	5187
Number of Observations Not in Domain	1801
Sum of Weights in Domain	218642036
Weighted Mean of stdlbhdhd	2.06593E-7
Weighted Sum of stdlbhdhd	45.16990

Fit Statistics	
R-Square	0.05829
Root MSE	0.9705
Denominator DF	17

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	288.54	<.0001
Intercept	1	0.00	1.0000
stdlbxtc	1	288.54	<.0001

The denominator degrees of freedom for the F tests is 17.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	0.00000041	0.03410036	0.00	1.0000
stdlbxtc	0.24144356	0.01421387	16.99	<.0001

The degrees of freedom for the t tests is 17.

Example 5.12: Estimating the Population Ratio of High Density to Total Cholesterol for U.S. Adults.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	14
Number of Clusters	31
Number of Observations	9756
Number of Observations Used	9338
Number of Obs with Nonpositive Weights	418
Sum of Weights	306590681

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
lbxtc	Total Cholesterol (mg/dL)	6988	188.529858	1.036759	186.342488	190.717227
lbdhdd	Direct HDL-Cholesterol (mg/dL)	6989	52.771964	0.470567	51.779153	53.764775

Example 5.12: Estimating the Population Ratio of High Density to Total Cholesterol for U.S. Adults.

The SURVEYMEANS Procedure

Statistics for age18p Domains							
age18p	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
0	lbxtc	Total Cholesterol (mg/dL)	1801	157.912524	1.148055	155.490340	160.334708
	lbdhdd	Direct HDL-Cholesterol (mg/dL)	1802	52.428317	0.443091	51.493476	53.363159
1	lbxtc	Total Cholesterol (mg/dL)	5187	194.435465	1.019552	192.284400	196.586531
	lbdhdd	Direct HDL-Cholesterol (mg/dL)	5187	52.838263	0.506560	51.769515	53.907011

Ratio Analysis						
Numerator	Denominator	N	Ratio	Std Error	95% CL for Ratio	
lbdhdd	lbxtc	6988	0.279913	0.002630	0.27436402	0.28546273

Example 5.12: Estimating the Population Ratio of High Density to Total Cholesterol for U.S. Adults.

The SURVEYMEANS Procedure

Ratios for age18p Domains							
age18p	Numerator	Denominator	N	Ratio	Std Error	95% CL for Ratio	
0	lbdhdd	lbxtc	1801	0.332011	0.002775	0.32615540	0.33786592
1	lbdhdd	lbxtc	5187	0.271752	0.002894	0.26564611	0.27785827

Example 5.13: Estimating the Proportions of Males and Females Age ≥ 70 with Diabetes Using the HRS Data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	56
Number of Clusters	112
Number of Observations	20554
Number of Observations Used	18851
Number of Obs with Nonpositive Weights	1703
Sum of Weights	90698760

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
diabetes	1=Yes Diabetes 0=No Diabetes	18838	0.215159	0.004414	0.20631766	0.22400109

Example 5.13: Estimating the Proportions of Males and Females Age ≥ 70 with Diabetes Using the HRS Data.

The SURVEYMEANS Procedure

Statistics for age70*gender Domains									
age70	gender	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean		
0	1	diabetes	1=Yes Diabetes 0=No Diabetes	4590	0.209950	0.006980	0.19596704	0.22393217	
	2	diabetes	1=Yes Diabetes 0=No Diabetes	6115	0.192096	0.006220	0.17963621	0.20455522	
1	1	diabetes	1=Yes Diabetes 0=No Diabetes	3494	0.273611	0.007468	0.25865017	0.28857234	
	2	diabetes	1=Yes Diabetes 0=No Diabetes	4639	0.226974	0.008555	0.20983745	0.24411115	

Example 5.14: Estimating Mean Systolic Blood Pressure for Males and Females Age > 45 using the 2011-2012 NHANES data.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	14
Number of Clusters	31
Number of Observations	9756
Number of Observations Used	9338
Number of Obs with Nonpositive Weights	418
Sum of Weights	306590681

Statistics					
Variable	Label	Sum of Weights	Mean	Std Error of Mean	Design Effect
BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	252389207	119.530464	0.546864	6.625878

Example 5.14: Estimating Mean Systolic Blood Pressure for Males and Females Age > 45 using the 2011-2012 NHANES data.

The SURVEYMEANS Procedure

Statistics for age45*riagendr Domains							
age45	riagendr	Variable	Label	Sum of Weights	Mean	Std Error of Mean	Design Effect
0	1	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	71978954	115.913072	0.455049	2.867775
	2	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	71046209	109.791492	0.491917	3.320528
1	1	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	51687896	128.300550	0.868705	2.893629
	2	BPXSY1	Systolic: Blood pres (1st rdg) mm Hg	57676147	128.181974	0.946016	3.162972

Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level.

The SURVEYMEANS Procedure

Data Summary	
Number of Strata	56
Number of Clusters	112
Number of Observations	20554
Number of Observations Used	19990
Number of Obs with Nonpositive Weights	564
Sum of Weights	89174512

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	19990	474129	19681	434703.583	513554.052

Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level.

The SURVEYMEANS Procedure

Statistics for finr*edcat Domains								
finr	edcat	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
1	1	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	2870	122089	10596	100863.103	143314.177
	2	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	4222	259027	9802.470020	239390.454	278663.868
	3	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	3265	336309	17202	301849.303	370767.939
	4	H11ATOTA	H11ATOTA:W11 Total of all Assets--Cross-wave	3232	834141	46478	741034.791	927247.195

Differences of H11ATOTA (H11ATOTA:W11 Total of all Assets--Cross-wave) Means for finr*edcat Domains								
finr	-finr	edcat	-edcat	Diff Estimate	Std Error	DF	t Value	Pr > t
1	1	1	2	-136939	12525	56	-10.93	<.0001
1	1	1	3	-214220	17182	56	-12.47	<.0001
1	1	1	4	-712052	48886	56	-14.57	<.0001
1	1	2	3	-77281	16245	56	-4.76	<.0001
1	1	2	4	-575114	47090	56	-12.21	<.0001
1	1	3	4	-497832	46277	56	-10.76	<.0001

Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level.

The SURVEYREG Procedure

Regression Analysis for Dependent Variable H11ATOTA

Data Summary	
Number of Observations	13589
Sum of Weights	58685309
Weighted Mean of H11ATOTA	429464.8
Weighted Sum of H11ATOTA	2.52033E13

Design Summary	
Number of Strata	56
Number of Clusters	112

Fit Statistics	
R-Square	0.05975
Root MSE	1082011
Denominator DF	56

Class Level Information			
CLASS Variable	Label	Levels	Values
edcat	Education 1=0-11 Yrs 2=12 Yrs 3=13-15 Yrs 4=16+ Yrs	4	1 2 3 4

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	3	99.77	<.0001
Intercept	1	753.81	<.0001
edcat	3	99.77	<.0001

The denominator degrees of freedom for the F tests is 56.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	834140.99	46482.9218	17.95	<.0001
edcat 1	-712052.35	48891.4538	-14.56	<.0001
edcat 2	-575113.83	47094.7482	-12.21	<.0001
edcat 3	-497832.37	46282.5555	-10.76	<.0001
edcat 4	0.00	0.0000	.	.

The degrees of freedom for the t tests is 56.

Matrix X'WX is singular and a generalized inverse was used to solve the normal equations. Estimates are not unique.

edcat Least Squares Means					
Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Estimate	Standard Error	DF	t Value	Pr > t
1	122089	10597	56	11.52	<.0001
2	259027	9803.55	56	26.42	<.0001
3	336309	17204	56	19.55	<.0001
4	834141	46483	56	17.95	<.0001

Differences of edcat Least Squares Means						
Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Estimate	Standard Error	DF	t Value	Pr > t
1	2	-136939	12527	56	-10.93	<.0001
1	3	-214220	17184	56	-12.47	<.0001
1	4	-712052	48891	56	-14.56	<.0001
2	3	-77281	16247	56	-4.76	<.0001
2	4	-575114	47095	56	-12.21	<.0001
3	4	-497832	46283	56	-10.76	<.0001

Example 5.15: Estimating Differences in Mean Total Household Wealth Between HRS Subpopulations Defined by Educational Attainment Level.

The SURVEYREG Procedure

finr=1

Domain Regression Analysis for Variable H11ATOTA

Domain Summary	
Number of Observations	13589
Number of Observations in Domain	13589
Number of Observations Not in Domain	0
Sum of Weights in Domain	58685309
Weighted Mean of H11ATOTA	429464.8
Weighted Sum of H11ATOTA	2.52033E13

Fit Statistics	
R-Square	0.05975
Root MSE	1082011
Denominator DF	56

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	3	99.77	<.0001
Intercept	1	753.81	<.0001
edcat	3	99.77	<.0001

The denominator degrees of freedom for the F tests is 56.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	834140.99	46482.9218	17.95	<.0001
edcat 1	-712052.35	48891.4538	-14.56	<.0001
edcat 2	-575113.83	47094.7482	-12.21	<.0001
edcat 3	-497832.37	46282.5555	-10.76	<.0001
edcat 4	0.00	0.0000	.	.

The degrees of freedom for the t tests is 56.

Matrix X'WX is singular and a generalized inverse was used to solve the normal equations. Estimates are not unique.

edcat Least Squares Means					
Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Estimate	Standard Error	DF	t Value	Pr > t
1	122089	10597	56	11.52	<.0001
2	259027	9803.55	56	26.42	<.0001

edcat Least Squares Means					
Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Estimate	Standard Error	DF	t Value	Pr > t
3	336309	17204	56	19.55	<.0001
4	834141	46483	56	17.95	<.0001

Differences of edcat Least Squares Means						
Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Education 1=0-11 Yrs 2=12 Yrs 3=13- 15 Yrs 4=16+ Yrs	Estimate	Standard Error	DF	t Value	Pr > t
1	2	-136939	12527	56	-10.93	<.0001
1	3	-214220	17184	56	-12.47	<.0001
1	4	-712052	48891	56	-14.56	<.0001
2	3	-77281	16247	56	-4.76	<.0001
2	4	-575114	47095	56	-12.21	<.0001
3	4	-497832	46283	56	-10.76	<.0001

Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the HRS study.

The SURVEYREG Procedure

Regression Analysis for Dependent Variable totwealth

Data Summary	
Number of Observations	37291
Sum of Weights	164299555
Weighted Mean of totwealth	479078.0
Weighted Sum of totwealth	7.87123E13

Design Summary	
Number of Strata	56
Number of Clusters	112

Fit Statistics	
R-Square	0.000013
Root MSE	1125104
Denominator DF	56

Class Level Information		
CLASS Variable	Levels	Values
year	2	2010 2012

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	0.71	0.4027
Intercept	1	718.42	<.0001
year	1	0.71	0.4027

The denominator degrees of freedom for the F tests is 56.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	483042.939	19048.5453	25.36	<.0001
year 2010	-7970.060	9451.6127	-0.84	0.4027
year 2012	0.000	0.0000	.	.

The degrees of freedom for the t tests is 56.

Matrix X'WX is singular and a generalized inverse was used to solve the normal equations. Estimates are not unique.

year Least Squares Means					
year	Estimate	Standard Error	DF	t Value	Pr > t
2010	475073	17908	56	26.53	<.0001

year Least Squares Means					
year	Estimate	Standard Error	DF	t Value	Pr > t
2012	483043	19049	56	25.36	<.0001

Differences of year Least Squares Means						
year	_year	Estimate	Standard Error	DF	t Value	Pr > t
2010	2012	-7970.06	9451.61	56	-0.84	0.4027

Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the HRS study.

The SURVEYREG Procedure

finr2010_2012=0

Domain Regression Analysis for Variable totwealth

Domain Summary	
Number of Observations	37291
Number of Observations in Domain	12109
Number of Observations Not in Domain	25182
Sum of Weights in Domain	57021932
Weighted Mean of totwealth	561362.6
Weighted Sum of totwealth	3.201E13

Fit Statistics	
R-Square	0.000045
Root MSE	1148621
Denominator DF	56

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	1.29	0.2616
Intercept	1	544.04	<.0001
year	1	1.29	0.2616

The denominator degrees of freedom for the F tests is 56.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	569090.113	25821.3345	22.04	<.0001
year 2010	-15421.085	13599.5262	-1.13	0.2616
year 2012	0.000	0.0000	.	.

The degrees of freedom for the t tests is 56.

Matrix X'WX is singular and a generalized inverse was used to solve the normal equations. Estimates are not unique.

year Least Squares Means					
year	Estimate	Standard Error	DF	t Value	Pr > t
2010	553669	24172	56	22.91	<.0001
2012	569090	25821	56	22.04	<.0001

Differences of year Least Squares Means						
year	_year	Estimate	Standard Error	DF	t Value	Pr > t
2010	2012	-15421	13600	56	-1.13	0.2616

Example 5.16: Estimating Differences in Mean Total Household Wealth from 2010 to 2012 using Data from the HRS study.

The SURVEYREG Procedure

finr2010_2012=1

Domain Regression Analysis for Variable totwealth

Domain Summary	
Number of Observations	37291
Number of Observations in Domain	25182
Number of Observations Not in Domain	12109
Sum of Weights in Domain	107277623
Weighted Mean of totwealth	435340.7
Weighted Sum of totwealth	4.67023E13

Fit Statistics	
R-Square	5.029E-6
Root MSE	1109915
Denominator DF	56

Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	0.39	0.5331
Intercept	1	736.78	<.0001
year	1	0.39	0.5331

The denominator degrees of freedom for the F tests is 56.

Estimated Regression Coefficients				
Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	437807.631	17016.5224	25.73	<.0001
year 2010	-4978.066	7936.9039	-0.63	0.5331
year 2012	0.000	0.0000	.	.

The degrees of freedom for the t tests is 56.

Matrix X'WX is singular and a generalized inverse was used to solve the normal equations. Estimates are not unique.

year Least Squares Means					
year	Estimate	Standard Error	DF	t Value	Pr > t
2010	432830	16011	56	27.03	<.0001
2012	437808	17017	56	25.73	<.0001

Differences of year Least Squares Means						
year	_year	Estimate	Standard Error	DF	t Value	Pr > t
2010	2012	-4978.07	7936.90	56	-0.63	0.5331

