APPENDIX D
EXAMPLES IN SUDAAN

In SUDAAN, the NCVS sample design must be appropriately specified for all SUDAAN procedures using the *design* and *nest* statements. The *design=wr* option must be included within the *proc* statement of each analysis. The following *nest* statement must be included to provide the stratification and PSU variables.

```
nest yr_grp v2117 v2118 / psulev=3;
```

Weight statements are also required, but vary depending on the type of estimate (as shown in *Table 3-1*). Victimization rates also utilize the adjustment factor (*ADJINC_WT*), as outlined in the examples below. Before interpreting results produced by SUDAAN, the design summary should always be examined to ensure that it matches with the input dataset and the specified design.

Examples 1 and *2* demonstrate how to estimate the total number of victimizations. *Examples 3* and *4* demonstrate how to calculate the proportion of victimizations with given characteristics. *Examples 5* and *6* demonstrate the calculation of personal and property victimization rates for victimization characteristics included on the modified person and household files, while *Example* 7 demonstrates the calculation of victimization rates for victimization characteristics not on the modified files. Examples are included for both single- and pooled-year estimates. Finally, *Example 8* demonstrates how to identify low quality estimates and implement the rounding recommended as discussed in *Section 4* of the user's guide.

D.1   Victimization Totals

SUDAAN's *descript* procedure is used to estimate the total number of victimizations from the modified incident-level file. The specific crime type of interest is specified in the *var* statement, and the domain(s) of interest (i.e. subsets of the population based on characteristics of the victimization or victim) are specified on the *class* and *tables* statements. *Examples 1A-1C* demonstrate the calculation of victimization totals for single years, and *Examples 2A-2C* demonstrate pooled year estimates.

Example 1:   *Number of victimizations, single year*

Examples *1A* and *1B* below are estimates of personal crimes, while *Example 1C* is an estimate of property crimes. *Examples 1A* and *1C* provide overall crime estimates, while *Example 1B* provides estimated totals for person-level characteristics of interest.

Example 1A - Total number of violent victimizations, 2011

Year(s): 2011
Crime Type: violent victimizations (*vcrime*)
Domain(s): n/a
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**)

Code:
```
proc descript data=adjincident design=wr; *a;
  subpopn exclude_outus=0 and dummy=0 and year=2011; *b;
  nest yr_grp v2117 v2118 / psulev=3; *c;
  weight newwgt; *d;
  var vcrime; *e;
  print total setotal; *f;
run;
```

Code Comment(s):
```
a) Specifies input dataset and sample design
b) Limits analysis to records of interest
c) Specifies stratification and primary sampling units
d) Specifies analysis weight
e) Outcome of interest
f) Specify which estimates to print
```

Output:

```
                            S U D A A N
            Software for the Statistical Analysis of Correlated Data
        Copyright      Research Triangle Institute      August 2012
                            Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   7255    Weighted count : 23041441
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


-------------------------------------------------------------------
|               |           | SUDAAN Reserved Variable      |
| Variable      |           | One                           |
|               |           |-------------------------------|
|               |           | Total         | 1             |
-------------------------------------------------------------------
|               |           |               |               |
| VCRIME        | Total     |   5812522.79  |   5812522.79  |
|               | SE Total  |    357912.36  |    357912.36  |
-------------------------------------------------------------------
```

Example 1B - Total number of violent victimizations by sex, age category, and race/ethnicity, 2011

Year(s): 2011
Crime Type: violent victimizations (*vcrime*)
Domain(s): sex (*sex*), age category (*agecat*), and race/ethnicity (*race_eth*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**)

Code:
```
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var vcrime;
  class sex agecat race_eth / nofreqs; *a;
  tables sex agecat race_eth; *b;
  print total setotal;
run;
```

Code Comment(s):
   a) Domain variables (victim characteristics), *nofreqs* option excludes frequency tables from output for *class* variables
   b) Domain variables (victim characteristics)

Output:
```
                          S U D A A N
          Software for the Statistical Analysis of Correlated Data
        Copyright      Research Triangle Institute      August 2012
                        Release 11.0.0
DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118

 Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   7255    Weighted count : 23041441
Denominator degrees of freedom :    160
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, SEX.
```

Variable			SEX		
			Total	Male	Female
VCRIME	Total		5812522.79	3209724.89	2602797.90
	SE Total		357912.36	236960.55	252785.86

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, AGECAT.
```

Variable			AGECAT			
			Total	12 to 14	15 to 17	18 to 20
VCRIME		Total	5812522.79	488382.86	428879.62	843753.11
		SE Total	357912.36	77367.96	64692.50	220103.55

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, AGECAT.

Variable			AGECAT			
			21 to 24	25 to 34	35 to 49	50 to 64
VCRIME		Total	617719.58	1114834.02	1365450.41	776857.48
		SE Total	69925.52	118262.08	141010.68	90922.04

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, AGECAT.

Variable			AGECAT
			65 or older
VCRIME		Total	176645.71
		SE Total	46321.04

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

Variable			RACE_ETH			
			Total	Non-Hispanic White	Non-Hispanic Black	Hispanic
VCRIME		Total	5812522.79	3719729.13	812173.25	895434.04
		SE Total	357912.36	242039.22	181536.14	130311.38

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

Variable			RACE_ETH	
			Non-Hispanic Other	Non-Hispanic More than One Race
VCRIME		Total	197081.81	188104.56
		SE Total	46736.13	43090.36

Example 1C - Total number of property victimizations, 2011

Year(s): 2011
Crime Type: property victimizations (*pcrime*)
Domain(s): n/a
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**)

Code:
```
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var pcrime;
  print total setotal;
run;
```

Output:

```
                       S U D A A N
           Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                       Release 11.0.0
DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118

 Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   7255    Weighted count : 23041441
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.
```

Variable		SUDAAN Reserved Variable One	
		Total	1
PCRIME	Total	17063147.83	17063147.83
	SE Total	543578.64	543578.64

Example 2:   Number of crimes, aggregated years

     Examples 2A and *2B* are estimates of personal crimes, while *Example 2C* is an estimate of a property crime. *Examples 2A* and *2C* provide overall crime estimates and *Example 2B* provides estimated totals for a victimization-level characteristic of interest. The code provided calculates the estimated number of victimization across the pooled years. To calculate the average number of victimizations per year, estimated totals and standard errors must be divided by the number of pooled years, as shown in the following examples.

Example 2A - Total and average number of aggravated assaults, 2009–2011

Year(s): 2009-2011
Crime Type: aggravated assaults (*aast*)
Domain(s): n/a
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))

Code:
```
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and (2009 <= year) and
          (year <=2011); *a;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var aast;
  print total setotal;
run;
```

Code Comment(s):
   a) *Subpopn* statement includes the range of years to be analyzed

Output:
```
                              S U D A A N
              Software for the Statistical Analysis of Correlated Data
           Copyright      Research Triangle Institute      August 2012
                          Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :  20701    Weighted count : 65552798
Denominator degrees of freedom :    160
```

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SUDAAN Reserved Variable One.

-------------------------------------------------------------------------
|                    |              | SUDAAN Reserved Variable       |
|  Variable          |              | One                            |
|                    |              |--------------------------------|
|                    |              | Total        | 1               |
-------------------------------------------------------------------------
|                    |              |              |                 |
|  AAST              | Total        |  2940415.60  |  2940415.60     |
|                    | SE Total     |   160814.15  |   160814.15     |
-------------------------------------------------------------------------
```

NOTE: The estimate above represents the total number of aggravated assaults from 2009-2011. To obtain the average number of aggravated assaults, both the estimate and the standard error must be divided by the number of pooled years (3), as follows:

$$avg\ number\ of\ aggravated\ assaults = \frac{total\ number\ of\ aggravated\ asaults}{number\ of\ pooled\ years}$$

$$= \frac{2940415.60}{3}$$

$$= 980138.53$$

$$SE(avg\ number\ of\ aggravated\ assaults) = \frac{se(total\ number\ of\ aggravated\ asaults)}{number\ of\ pooled\ years}$$

$$= \frac{160814.15}{3}$$

$$= 53604.72$$

Example 2B - Total and average number of aggravated assaults involving firearm, 2009–2011

Year(s): 2009-2011
Crime Type: aggravated assaults (*aast*)
Domain(s): weapon category (*weapcat*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))

Code:
```
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and (2009 <= year) and
        (year <=2011);
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var aast;
  class weapcat / nofreqs;
  tables weapcat;
  print total setotal;
run;
```

Output:
```
                            S U D A A N
          Software for the Statistical Analysis of Correlated Data
          Copyright     Research Triangle Institute     August 2012
                          Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :  20701    Weighted count : 65552798
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, WEAPCAT.
```

Variable			WEAPCAT			
			Total	No Weapon	Firearm	Knife or sharp object
AAST		Total	2940415.60	318541.93	778706.29	642022.34
		SE Total	160814.15	53739.51	81552.16	65520.62

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, WEAPCAT.
```

			WEAPCAT	

Variable			Other type weapon	Type weapon unknown	Do Not Know if off had weapon
AAST		Total	939927.30	228754.77	32462.99
		SE Total	96628.22	52287.67	11505.77

NOTE: The estimate above represents the total number of aggravated assaults involving a firearm from 2009-2011. To obtain the average number of aggravated assaults involving a firearm, both the estimate and the standard error must be divided by the number of pooled years (3), as follows:

$$avg\ number\ of\ aggravated\ assaults\ w\ firearm = \frac{total\ number\ of\ aggravated\ asaults\ w\ firearm}{number\ of\ pooled\ years}$$

$$= \frac{778706.29}{3}$$

$$= 259568.76$$

$$SE(avg\ number\ of\ aggravated\ assaults\ w\ firearm) = \frac{se(total\ number\ of\ aggravated\ asaults\ w\ firearm)}{number\ of\ pooled\ years}$$

$$= \frac{81552.16}{3}$$

$$= 27184.05$$

Example 2C - Total and average number of household burglaries, 2009-2011

Year(s): 2009-2011
Crime Type: household burglary (*hburg*)
Domain(s): n/a
Weight: *newwgt*
Subpopulation: Within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**);
2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))

Code:

```
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and (2009 <= year) and
          (year <=2011);
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var hburg;
  print total setotal;
run;
```

Output:

```
                         S U D A A N
          Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      August 2012
                         Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :  20701    Weighted count : 65552798
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SUDAAN Reserved Variable One.
```

		SUDAAN Reserved Variable One	
Variable		Total	1
HBURG	Total	10201317.31	10201317.31
	SE Total	336711.45	336711.45

NOTE: The estimate above represents the total number of household burglaries from 2009-2011. To obtain the average number of household burglaries, both the estimate and the standard error must be divided by the number of pooled years (3), as follows:

$$avg\ number\ of\ household\ burglaries = \frac{total\ number\ of\ household\ burglaries}{number\ of\ pooled\ years}$$

$$= \frac{10201317.31}{3}$$

$$= 3400439.10$$

$$SE(avg\ number\ of\ household\ burglaries) = \frac{se(total\ number\ of\ household\ burglaries)}{number\ of\ pooled\ years}$$

$$= \frac{336711.45}{3}$$

$$= 112237.15$$

D.2   Victimization Proportions

SUDAAN's *crosstab* procedure is used to estimate the percent distribution of victimizations in a specific domain across covariates from the modified incident-level file. If the analysis is limited to a specific domain (i.e. a subset of the population based on characteristics of the victimization or victim) then it is specified on the *subpopn* statement. Covariates of interest are specified on the *class* and *tables* statements.

Example 3: Percent distribution of victims across characteristics, single year

Example 3A is an estimate of personal crimes, while *Example 3B* is an estimate of property crimes. In *Example 3A*, both the analysis domain and covariates are specified based on person- and incident-level characteristics. In *Example 3B*, the analysis domain is specified based on an incident-level characteristic, while the covariates are household-level characteristics.

Example 3A - Distribution of female violent crime victims by age category, race/ethnicity, and victim-offender relationship, 2011

Year(s): 2011
Domain(s): female (*sex*=2); violent victimizations (*vcrime*=1)
Covariate(s): age category (*agecat*); race/ethnicity (*race_eth*); victim-offender relationship (*direl*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**); female (*sex*=2); violent victimizations (*vcrime*=1)

Code:
```
proc crosstab data=adjincident design=wr; *a;
  subpopn exclude_outus=0 and dummy=0 and year=2011 and sex=2
        and vcrime=1; *b;
  nest yr_grp v2117 v2118 / psulev=3; *c;
  weight newwgt; *d;
  class agecat race_eth direl / nofreqs; *e;
  tables  agecat race_eth direl; *f;
  print rowper serow; *g;
run;
```

Code Comment(s):
```
  a) Specifies input dataset and sample design
  b) Limits analysis to records of interest
  c) Specifies stratification and primary sampling units
  d) Specifies analysis weight
  e) Covariates of interest
  f) Covariates of interest
  g) Specify which estimates to print (row percents and standard errors)
```

Output:

```
                          S U D A A N
           Software for the Statistical Analysis of Correlated Data
      Copyright     Research Triangle Institute      August 2012
                        Release 11.0.0
```

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118

Number of observations read   : 37853   Weighted count :115824878
Observations in subpopulation :   642   Weighted count : 2602798
Denominator degrees of freedom :   160
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: AGECAT.

		AGECAT				
		Total	12 to 14	15 to 17	18 to 20	21 to 24
	Row Percent	100.00	7.64	6.26	15.50	9.31
	SE Row Percent	0.00	1.86	1.33	5.37	1.64

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: AGECAT.

		AGECAT			
		25 to 34	35 to 49	50 to 64	65 or older
	Row Percent	19.30	24.58	12.91	4.48
	SE Row Percent	2.36	3.05	1.99	1.72

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: RACE_ETH.

		RACE_ETH				
		Total	Non-Hispanic White	Non-Hispanic Black	Hispanic	Non-Hispanic Other

```
|                       | Row Percent    |   100.00 |   59.87 |   19.43 |   11.67 |    4.42 |
|                       | SE Row Percent |     0.00 |    4.46 |    5.18 |    1.98 |    1.51 |
-----------------------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: RACE_ETH.

```
-------------------------------------------------
|                       |                | RACE_ETH |
|                       |                |----------|
|                       |                | Non-     |
|                       |                | Hispanic |
|                       |                | More     |
|                       |                | than One |
|                       |                | Race     |
-------------------------------------------------
|                       |                |          |
|                       | Row Percent    |     4.61 |
|                       | SE Row Percent |     1.29 |
-------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: DIREL.

```
---------------------------------------------------------------------------------------------
|                       |                | DIREL    |          |          |          |          |
|                       |                |-------------------------------------------------------|
|                       |                | Total    | intimat- | other    | well     | strange- |
|                       |                |          | es       | relativ- | known/c- | rs       |
|                       |                |          |          | es       | asual    |          |
|                       |                |          |          |          | acquain- |          |
|                       |                |          |          |          | tances   |          |
---------------------------------------------------------------------------------------------
|                       |                |          |          |          |          |          |
|                       | Row Percent    |   100.00 |    23.23 |     9.98 |    35.28 |    25.34 |
|                       | SE Row Percent |     0.00 |     3.50 |     1.64 |     4.56 |     3.16 |
---------------------------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND SEX = 2 AND VCRIME = 1
by: DIREL.

```
-------------------------------------------------------------
|                       |                | DIREL    |          |
|                       |                |---------------------|
|                       |                | do not   | do not   |
|                       |                | know     | know     |
|                       |                | relatio- | number   |
|                       |                | nship    | of       |
|                       |                |          | offende- |
|                       |                |          | rs       |
-------------------------------------------------------------
|                       |                |          |          |
|                       | Row Percent    |     2.71 |     3.46 |
|                       | SE Row Percent |     0.61 |     0.94 |
-------------------------------------------------------------
```

Example 3B - Distribution of property crime victims by household income, MSA status, and region, 2011

Year(s): 2011
Domain(s): property victimizations (*pcrime*=1)
Covariate(s): household income (*hincome*); MSA status (*msa*); region (*region*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**); property victimizations (*pcrime*=1)

Code:
```
proc crosstab data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011
          and pcrime=1;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  class hincome msa region / nofreqs;
  tables hincome msa region;
  print rowper serow;
run;
```
Output:
```
                        S U D A A N
            Software for the Statistical Analysis of Correlated Data
            Copyright     Research Triangle Institute     August 2012
                          Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118

Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   5857    Weighted count : 17063148
Denominator degrees of freedom :    160


Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND PCRIME = 1
by: HINCOME.
```

			HINCOME				
			Total	Less than $7,500	$7,500 to $14,999	$15,000 to $24,999	$25,000 to $34,999
		Row Percent	100.00	6.45	8.68	10.25	8.89
		SE Row Percent	0.00	0.68	0.66	0.66	0.60

Variance Estimation Method: Taylor Series (WR)

For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND PCRIME = 1
by: HINCOME.

| | | HINCOME | | | |
		$35,000 to $49,999	$50,000 to $74,999	$75,000 or more	Unknown
	Row Percent	11.72	11.50	16.47	26.06
	SE Row Percent	0.63	0.70	0.73	1.13

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND PCRIME = 1
by: HOUSEHOLD LOCALE.

| | | HOUSEHOLD LOCALE | | | |
		Total	Urban	Suburban	Rural
	Row Percent	100.00	40.60	46.62	12.79
	SE Row Percent	0.00	1.51	1.62	1.63

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND PCRIME = 1
by: Region.

| | | Region | | | | |
		Total	Northeast	Midwest	South	West
	Row Percent	100.00	13.40	21.41	34.81	30.38
	SE Row Percent	0.00	1.08	1.45	1.52	1.40

Example 4: Percent distribution of victims across characteristics, aggregated years

    Example 4A is an estimate of personal crimes, while *Example 4B* is an estimate of property crimes. Both examples specify multiple covariates based on incident-level characteristics.

Example 4A - Percent of violent victimizations reported and not reported to police by type of crime, 2009-2011

Year(s): 2009-2011
Domain(s): violent victimizations (*vcrime*=1)
Covariate(s): report status (*notify*); type of crime (*newoff*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**)); violent victimizations (*vcrime*=1)

Code:
```
proc crosstab data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and (2009 <= year) and
          (year <=2011) and vcrime=1;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  class newoff notify / nofreqs;
  tables  newoff*notify; *a;
  print rowper serow;
run;
```

Code Comment(s):
   a) The goal is to obtain the distribution of victimizations by report status for each type of violent crime. Therefore, *newoff*notify* is specified on the *tables* statement and row percents and standard errors (*ROWPER* and *SEROW*) are requested. If the percent distribution of crimes among those reported or not reported to police were desired, column percents and standard errors (*COLPER* and *SECOL*)would be requested.

Output:

```
                           S U D A A N
            Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      August 2012
                           Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118




Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   3760    Weighted count : 16417743
Denominator degrees of freedom :    160
```

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 ) AND
VCRIME = 1
by: NEWOFF, NOTIFY.


---------------------------------------------------------------------------
|               |                | NOTIFY                                 |
| NEWOFF        |                |----------------------------------------|
|               |                | Total  | reported | not       | do not |
|               |                |        | to       | reported  | know   |
|               |                |        | police   | to        |        |
|               |                |        |          | police    |        |
---------------------------------------------------------------------------
|               |                |        |          |           |        |
| Total         | Row Percent    | 100.00 |   48.09  |   50.59   |  1.32  |
|               | SE Row Percent |   0.00 |    1.46  |    1.48   |  0.42  |
---------------------------------------------------------------------------
|               |                |        |          |           |        |
| Rape & Sexual | Row Percent    | 100.00 |   32.48  |   67.52   |  0.00  |
| Assault       | SE Row Percent |   0.00 |    6.41  |    6.41   |  0.00  |
---------------------------------------------------------------------------
|               |                |        |          |           |        |
| Robbery       | Row Percent    | 100.00 |   61.59  |   37.65   |  0.76  |
|               | SE Row Percent |   0.00 |    3.32  |    3.38   |  0.46  |
---------------------------------------------------------------------------
|               |                |        |          |           |        |
| Aggravated    | Row Percent    | 100.00 |   59.01  |   39.49   |  1.49  |
| Assault       | SE Row Percent |   0.00 |    2.55  |    2.67   |  0.93  |
---------------------------------------------------------------------------
|               |                |        |          |           |        |
| Simple Assault| Row Percent    | 100.00 |   44.12  |   54.42   |  1.46  |
|               | SE Row Percent |   0.00 |    1.79  |    1.79   |  0.45  |
---------------------------------------------------------------------------
```

Example 4B - Percent of property victimizations reported and not reported to police by type of crime, 2009-2011

Year(s): 2009-2011
Domain(s): property victimizations (*pcrime*=1)
Covariate(s): report status (*notify*); type of crime (*newoff*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**)); property victimizations (*pcrime*=1)

Code:
```
proc crosstab data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and (2009 <= year) and
          (year <=2011) and pcrime=1;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  class newoff notify / nofreqs;
  tables  newoff*notify;
  print rowper serow;
run;
```

Output:
```
                         S U D A A N
          Software for the Statistical Analysis of Correlated Data
        Copyright      Research Triangle Institute      August 2012
                       Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :  16813    Weighted count : 48697733
Denominator degrees of freedom :    160




Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND ( 2009 <= YEAR ) AND ( YEAR <= 2011 ) AND
PCRIME = 1
by: NEWOFF, NOTIFY.

--------------------------------------------------------------------------------
```

NEWOFF		NOTIFY			
		Total	reported to police	not reported to police	do not know
Total	Row Percent	100.00	38.19	61.13	0.68
	SE Row Percent	0.00	0.52	0.52	0.10
Household Burglary	Row Percent	100.00	54.55	44.65	0.80
	SE Row Percent	0.00	1.51	1.50	0.41
Motor Vehicle Theft	Row Percent	100.00	83.74	15.99	0.27
	SE Row Percent	0.00	1.40	1.39	0.19
Theft	Row Percent	100.00	31.17	68.17	0.67
	SE Row Percent	0.00	0.50	0.50	0.07

D.3   Victimization Rates

Victimization rates are calculated from the modified person-level file (for personal crimes) or the modified household-level file (for property crimes) using the *xmean* statement in SUDAAN's *varagen* procedure. The victimization rate is first given a unique user-specified name, followed by a colon. Then, the victimization count for the victimization type of interest is multiplied by the victimization adjustment factor (*ADJINC_WT*), and this product is multiplied by 1,000. To calculate victimization rates for person- or household-level characteristics, *class* and *tables* statements are included. However, if victimization rates are desired based on incident-level characteristics, a separate *xmean* statement is defined for each level of the incident characteristic of interest. Alternatively, SUDAAN's *descript* procedure can be used for the calculation of victimization rates, but because *descript* does not allow the user to specify a formula for the variable of interest, the analyst must first create a new variable equal to the product of the victimization count for the victimization type of interest and the victimization adjustment factor (*ADJINC_WT*), multiplied by 1,000. This new variable is used as the analysis variable in the *descript* procedure.

Exclusions based on the incident characteristics must be made when calculating victimization summaries, as outlined in ***Section 3.2***. For example, the modified person- and household-level files exclude victimizations occurring outside of the United States. Because there are no dummy records on the modified person and household files, no exclusions are needed to remove dummy records from the analysis.

The modified person and household files contain the victimization counts needed to calculate victimization rates for the most common victimization characteristics analyzed using NCVS data. *Examples 5A-5C* demonstrate the calculation of personal and property victimizations rates that can be calculated directly from the modified person and household level files provided for a single year and *Examples 6A-6B* demonstrate these calculations for pooled year estimates. *Example 5A* includes calculations using both the *vargen* and *descript* procedures, while remaining examples use only the *vargen* procedure for simplicity. If an analyst wants to calculate a victimization rate for an incident-level characteristic that is not included on the modified files, preprocessing steps are needed to calculate victimization summaries from the incident-level file and move these summaries to the person file (for personal crimes) or the household file (for property crimes). ***Section 3.2*** documents these steps. *Examples 7A-7B* demonstrate the calculation of

personal and property victimization rates for incident characteristics not included on the modified person and household files.

Example 5: Rate of crime, single year

Example 5A is an estimate of an overall personal victimization rate. *Example 5B* is an estimate of a property victimization rate with the inclusion of household-level domain characteristics. *Example 5C* is an estimate of a personal victimization rate, where the domain characteristics of interest are incident-level characteristics. All three estimates are computed with variables available on the modified person and household files.

Example 5A - Rate of simple assaults, 2011

Year(s): 2011
Crime Type(s): simple assault (*sast*)
Domain(s): n/a
Weight: *wgtpercy*
Subpopulation: 2011 (*year=2011*)
Calculated Directly from Adjusted Files?: yes

First Approach: vargen procedure
Code:
```
proc vargen data=adjper design=wr; *a;
 subpopn year=2011; *b;
 nest yr_grp v2117 v2118 / psulev=3; *c;
 weight wgtpercy; *d;
 xmean vrsast: adjinc_wt*sast*1000; *e;
 setenv colwidth=15 decwidth=5; *f;
 print estim seestim; *g;
run;
```

Code Comment(s):
```
   a) Specifies input dataset and sample design
   b) Limits analysis to records of interest
   c) Specifies stratification and primary sampling units
   d) Specifies analysis weight
   e) Specifies name and formula for victimization rate
   f) Specify format of output
   g) Specify which estimates to print (victimization rate and standard
      error)
```

Output:

```
                              S U D A A N
                Software for the Statistical Analysis of Correlated Data
                Copyright      Research Triangle Institute      August 2012
                              Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WGTPERCY
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read     : 708493    Weighted count :1270197175
Number of observations skipped : 106187
(WEIGHT variable nonpositive)
Observations in subpopulation  : 143122    Weighted count :257542238
Denominator degrees of freedom :    160
Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


-------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
-------------------------------------------------------

VRSAST
   Total                  15.36713         1.17685
   1                      15.36713         1.17685
-------------------------------------------------------
```

Second Approach: descript procedure
Code:

```
data anset;
      set adjper;
      vrsast=adjinc_wt*sast*1000; *a;
run;

proc descript data=anset design=wr; *b;
 subpopn year=2011; *c;
 nest yr_grp v2117 v2118 / psulev=3; *d;
 weight wgtpercy; *e;
 var vrsast; *f;
 setenv colwidth=15 decwidth=5; *g;
 print mean semean; *h;
run;
```

Code Comment(s):
   a) Specifies victimization recode
   b) Specifies input dataset and sample design
   c) Limits analysis to records of interest

d) Specifies stratification and primary sampling units
e) Specifies analysis weight
f) Specifies analysis variable
g) Specify format of output
h) Specify which estimates to print (victimization rate and standard error)

Output:

```
                        S U D A A N
        Software for the Statistical Analysis of Correlated Data
        Copyright      Research Triangle Institute      June 2013
                        Release 11.0.1



DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
     Sample Weight: WGTPERCY
     Stratification Variables(s): YR_GRP V2117
     Primary Sampling Unit: V2118


Number of observations read    :1902341    Weighted count :3155122787
Number of observations skipped : 281776
(WEIGHT variable nonpositive)
Observations in subpopulation  : 143122    Weighted count :257542238
Denominator degrees of freedom :    304


Variance Estimation Method: Taylor Series (WR)
For Subpopulation: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.
```

		SUDAAN Reserved Variable One
Variable		---
		Total	1

VRSAST	Mean	15.36713	15.36713
	SE Mean	1.17685	1.17685

Example 5B - Rate of household burglary by MSA status, household income, and family structure, 2011

Year(s): 2011
Crime Type(s): household burglary (*hburg*)
Domain(s): MSA status (*msa*); household income (*hincome*); family structure (*fam_structure2*)
Weight: *wgthhcy*
Subpopulation: 2011 (*year=2011*)
Calculated Directly from Adjusted Files?: yes

Code:
```
proc vargen data=adjhh design=wr;*a;
 subpopn year=2011;
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgthhcy;
 class msa hincome fam_structure2 / nofreqs; *b;
 tables msa hincome fam_structure2; *c;
 xmean vrburg: adjinc_wt*hburg*1000;
 setenv colwidth=15 decwidth=5;
 print estim seestim;
run;
```

Code Comment(s):
   a) Specifies input dataset and sample design
   b) Specifies domains of interest
   c) Specifies domains of interest

Output:
```
                        S U D A A N
          Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                        Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
     Sample Weight: WGTHHCY
     Stratification Variables(s): YR_GRP V2117
     Primary Sampling Unit: V2118


Number of observations read    : 398238    Weighted count :608895975
Number of observations skipped : 129435
(WEIGHT variable nonpositive)
Observations in subpopulation  :  79802    Weighted count :123038566
Denominator degrees of freedom :    160
```

```
Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, HOUSEHOLD LOCALE.


---------------------------------------------------------
Variable
   HOUSEHOLD LOCALE         Estimate      SE Estimate
---------------------------------------------------------
VRBURG
    Total                   29.37159          1.59444
    Urban                   33.48701          2.51543
    Suburban                25.44535          2.00861
    Rural                   33.01906          4.53445
---------------------------------------------------------


Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, HINCOME.


---------------------------------------------------------
Variable
   HINCOME                  Estimate      SE Estimate
---------------------------------------------------------

VRBURG
    Total                   29.37159          1.59444
    Less than $7,500        67.16497         14.00675
    $7,500 to $14,999       58.84510          9.94479
    $15,000 to
       $24,999              42.77625          5.55678
    $25,000 to
       $34,999              33.34014          5.83093
    $35,000 to
       $49,999              26.30117          2.68244
    $50,000 to
       $74,999              21.50634          3.00622
    $75,000 or more         12.28870          1.23588
    Unknown                 29.54958          2.70272
---------------------------------------------------------
```

```
Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, Family Structure.


---------------------------------------------------------
Variable
   Family Structure         Estimate      SE Estimate
---------------------------------------------------------
VRBURG
    Total                   29.37230          1.59447
    Two or more
      adults - W/O
```

```
     Children              24.20889           2.30429
Two or more
  adults - With
  Children              29.61002           2.38192
One Male Adult -
  W/O Children          30.99701           2.92235
One Male Adult -
  With Children         46.66971          14.31937
One Female Adult
  - W/O Children        30.10567           4.79787
One Female Adult
  - With Children       61.96275           8.31509
--------------------------------------------------------
```

Example 5C - Rate of violent crime by weapon involvement, injury, and victim-offender relationship, 2011

Year(s): 2011
Crime Type(s): violent victimizations (*rsa+rob+aast+sast*)
Domain(s): weapon involvement (*wpnuse*); injury (*inj*), victim-offender relationship (*rel*)
Weight: *wgtpercy*
Subpopulation: 2011 (*year=2011*)
Calculated Directly from Adjusted Files?: yes

Code:

```
data example5c_per;
 set adjper;

 *aggregate violent crimes into a single measure for each
  outcome of interest;
 viol_weap1=(rsa_wpnuse1+rob_wpnuse1+aast_wpnuse1+sast_wpnuse1);
 viol_weap2=(rsa_wpnuse2+rob_wpnuse2+aast_wpnuse2+sast_wpnuse2);
 viol_weap3=(rsa_wpnuse3+rob_wpnuse3+aast_wpnuse3+sast_wpnuse3);

 viol_inj1=(rsa_inj1+rob_inj1+aast_inj1+sast_inj1);
 viol_inj2=(rsa_inj2+rob_inj2+aast_inj2+sast_inj2);
 viol_inj3=(rsa_inj3+rob_inj3+aast_inj3+sast_inj3);
 viol_inj4=(rsa_inj4+rob_inj4+aast_inj4+sast_inj4);

 viol_rel1=(rsa_rel1+rob_rel1+aast_rel1+sast_rel1);
 viol_rel2=(rsa_rel2+rob_rel2+aast_rel2+sast_rel2);
 viol_rel3=(rsa_rel3+rob_rel3+aast_rel3+sast_rel3);
 viol_rel4=(rsa_rel4+rob_rel4+aast_rel4+sast_rel4);
 viol_rel5=(rsa_rel5+rob_rel5+aast_rel5+sast_rel5);
 viol_rel6=(rsa_rel6+rob_rel6+aast_rel6+sast_rel6);
run;
```

```
proc vargen data=example5c_per design=wr;
 subpopn year=2011;
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgtpercy;
 xmean vrweap1: adjinc_wt*viol_weap1*1000; *a;
 xmean vrweap2: adjinc_wt*viol_weap2*1000; *b;
 xmean vrweap3: adjinc_wt*viol_weap3*1000; *c;

 xmean vrinj1: adjinc_wt*viol_inj1*1000; *d;
 xmean vrinj2: adjinc_wt*viol_inj2*1000; *e;
 xmean vrinj3: adjinc_wt*viol_inj3*1000; *f;
 xmean vrinj4: adjinc_wt*viol_inj4*1000; *g;

 xmean vrrel1: adjinc_wt*viol_rel1*1000; *h;
 xmean vrrel2: adjinc_wt*viol_rel2*1000; *i;
 xmean vrrel3: adjinc_wt*viol_rel3*1000; *j;
 xmean vrrel4: adjinc_wt*viol_rel4*1000; *k;
 xmean vrrel5: adjinc_wt*viol_rel5*1000; *l;
 xmean vrrel6: adjinc_wt*viol_rel6*1000; *m;

 setenv colwidth=15 decwidth=5;
 print estim seestim;
run;
```

Code Comment(s):
   a) Yes, offender had a weapon
   b) No, offender did not have a weapon
   c) Do not know if offender had a weapon
   d) No injury
   e) Serious injury
   f) Minor injury
   g) Rape without other injuries
   h) Intimates
   i) Other relatives
   j) well known/casual acquaintances
   k) Strangers
   l) Do not know relationship
   m) Do not know number of offenders

Output:

```
                          S U D A A N
          Software for the Statistical Analysis of Correlated Data
          Copyright     Research Triangle Institute     August 2012
                          Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
      Sample Weight: WGTPERCY
      Stratification Variables(s): YR_GRP V2117
      Primary Sampling Unit: V2118



Number of observations read    : 708493    Weighted count :1270197175
Number of observations skipped : 106187
(WEIGHT variable nonpositive)
Observations in subpopulation  : 143122    Weighted count :257542238
Denominator degrees of freedom :    160



Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


-------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
-------------------------------------------------------
VRWEAP1
   Total                   4.63777         0.34818
   1                       4.63777         0.34818
-------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


-------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
-------------------------------------------------------
VRWEAP2
   Total                  16.45241         1.17620
   1                      16.45241         1.17620
-------------------------------------------------------




Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRWEAP3
   Total                   1.47902         0.30060
```

```
   1                       1.47902        0.30060
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRINJ1
   Total                  16.93653        1.22551
   1                      16.93653        1.22551
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRINJ2
   Total                   1.12991        0.19134
   1                       1.12991        0.19134
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRINJ3
   Total                   4.08535        0.38482
   1                       4.08535        0.38482
--------------------------------------------------------


Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRINJ4
   Total                   0.41741        0.09706
   1                       0.41741        0.09706
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One          Estimate      SE Estimate
--------------------------------------------------------
```

```
VRREL1
   Total                    3.30343           0.49965
   1                        3.30343           0.49965
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate       SE Estimate
--------------------------------------------------------
VRREL2
   Total                    1.95748           0.35481
   1                        1.95748           0.35481
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate       SE Estimate
--------------------------------------------------------

VRREL3
   Total                    6.91000           0.77793
   1                        6.91000           0.77793
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate       SE Estimate
--------------------------------------------------------
VRREL4
   Total                    8.33380           0.68670
   1                        8.33380           0.68670
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate       SE Estimate
--------------------------------------------------------
VRREL5
   Total                    0.93560           0.12568
   1                        0.93560           0.12568
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.


--------------------------------------------------------
Variable
```

```
    SUDAAN Reserved
      Variable One          Estimate      SE Estimate
--------------------------------------------------------
VRREL6
   Total                     1.12888          0.21522
   1                         1.12888          0.21522
--------------------------------------------------------
```

Example 6: Rate of crime, aggregated years

Example 6A is an estimate of a personal victimization rate with person-level domain characteristics of interest. *Example 6B* is an estimate of a property victimization rate with the inclusion of a household-level domain characteristic. Both estimates are computed with variables available on the modified person and household files.

Example 6A - Rate of violent crimes reported to police by sex, age category, race/ethnicity, and marital status, 2009–2011

Year(s): 2009-2011
Crime Type(s): violent victimizations reported to police (*rsa_rpt1+rob_rpt1+aast_rpt1+sast_rpt1*)
Domain(s): sex (*sex*); age category (*agecat*); race/ethnicity (*race_eth*); marital status (*marital*)
Weight: *wgtpercy*
Subpopulation: 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))
Calculated Directly from Adjusted Files?: yes

Code:
```
data example6a_per;
 set adjper;
  *aggregate violent crimes into a single measure for outcome of
   interest;
 viol_rpt1=(rsa_rpt1+rob_rpt1+aast_rpt1+sast_rpt1);
run;

proc vargen data=example6a_per design=wr;
 subpopn (2009 <= year) and (year <=2011);
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgtpercy;
 xmean vrviolrpt: adjinc_wt*viol_rpt1*1000;
 class sex agecat race_eth marital / nofreqs;
 tables sex agecat race_eth marital;
 setenv colwidth=15 decwidth=5;
 print estim seestim;
 output / filename=ex6a replace estimfmt=f10.6 seestimfmt=f10.6;
run;
```

Output:

```
                          S U D A A N
          Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                          Release 11.0.0



DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
     Sample Weight: WGTPERCY
     Stratification Variables(s): YR_GRP V2117
     Primary Sampling Unit: V2118



Number of observations read    : 708493    Weighted count :1270197175
Number of observations skipped : 106187
(WEIGHT variable nonpositive)
Observations in subpopulation  : 427018    Weighted count :767609782
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SEX.

--------------------------------------------------------

Variable
   SEX                   Estimate     SE Estimate
--------------------------------------------------------

VRVIOLRPT
   Total                 10.25480        0.49316
   Male                   9.68295        0.53535
   Female                10.80133        0.75009
--------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, AGECAT.

--------------------------------------------------------
Variable
   AGECAT                Estimate     SE Estimate
--------------------------------------------------------

VRVIOLRPT
   Total                 10.25480        0.49316
   12 to 14              10.14255        1.52829
   15 to 17              13.04128        1.60744
   18 to 20              22.57700        4.63942
   21 to 24              16.93357        1.89590
   25 to 34              14.07592        1.17349
   35 to 49              11.41399        0.89398
   50 to 64               6.59324        0.58231
   65 or older            1.90274        0.37196
--------------------------------------------------------



Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, RACE_ETH.

--------------------------------------------------------
Variable
   RACE_ETH              Estimate     SE Estimate
--------------------------------------------------------

VRVIOLRPT
   Total                 10.25480        0.49316
   Non-Hispanic
```

```
    White               9.59106          0.55619
 Non-Hispanic
    Black              16.32449          2.15038
 Hispanic               9.11081          0.90575
 Non-Hispanic
    Other               6.21558          1.43177
 Non-Hispanic More
    than One Race      20.78554          3.67528
---------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: (2009 <= YEAR) AND (YEAR <= 2011)
by: Variable, MARITAL.

```
---------------------------------------------------------
Variable
   MARITAL            Estimate      SE Estimate
---------------------------------------------------------
VRVIOLRPT
    Total              10.28470         0.49491
    Never Married      15.24882         1.03449
    Married             5.03580         0.38311
    Widowed             3.73865         0.91618
    Divorced           18.37270         1.86821
    Separated          38.60338         5.88557
---------------------------------------------------------
```

Example 6B - Rate of property crimes reported to police by household income, 2009–2011

Year(s): 2009-2011
Crime Type(s): property crimes reported to police (*hburg_rpt1+tft_rpt1+mvtft_rpt1*)
Domain(s): household income (*hincome*)
Weight: *wgthhcy*
Subpopulation: 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))
Calculated Directly from Adjusted Files?: yes

Code:
```
data example6b_hh;
 set adjhh;
 *aggregate property crimes into a single measure for outcome of
  interest;
 prop_rpt1=(hburg_rpt1+tft_rpt1+mvtft_rpt1);
run;

proc vargen data=example6b_hh design=wr;
 subpopn (2009 <= year) and (year <=2011);
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgthhcy;
 class hincome / nofreqs;
 tables hincome;
 xmean vrproprpt: adjinc_wt*prop_rpt1*1000;
 setenv colwidth=15 decwidth=5;
 print estim seestim;

run;
```
Output:

```
                          S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                          Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WGTHHCY
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    : 398238    Weighted count :608895975
Number of observations skipped : 129435
(WEIGHT variable nonpositive)
Observations in subpopulation  : 239205    Weighted count :368251383
Denominator degrees of freedom :    160




Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, HINCOME.

---------------------------------------------------------
```

```
Variable
  HINCOME               Estimate      SE Estimate
---------------------------------------------------------
VRPROPRPT
  Total                 50.41720        0.98907
  Less than $7,500      64.05274        4.36642
  $7,500 to $14,999     60.84082        3.67851
  $15,000 to
    $24,999             54.83199        2.79486
  $25,000 to
    $34,999             54.83207        3.22249
  $35,000 to
    $49,999             55.92131        3.03029
  $50,000 to
    $74,999             49.01113        2.02956
  $75,000 or more       49.94612        1.49055
  Unknown               43.61699        1.54434
---------------------------------------------------------
```

Example 7: Computing victimization rates based on incident characteristics not included on the provided file

The two examples below estimate personal and property victimization rates for incident-level characteristics not contained on the modified person and household files, and thus require the pre-processing steps outlined in ***Section 3.2***. SAS is used to calculate victimization summaries from the incident-level file and merge incident summaries onto the household and person files. Victimization rates are then calculated from the modified person-level file (for personal crimes) or the modified household-level file (for property crimes). *Example 7A* is an estimate of a personal victimization rate and *Example 7B* is an estimate of a property victimization rate. *Example 7A* is for a single year, and *Example 7B* is for aggregated years.

Example 7A - Rate of violent crimes by location of crime, 2011

Year(s): 2011
Crime Type(s): violent victimizations by location of crime (*violent_home*; *violent_other*)
Domain(s): n/a
Weight: *wgtpercy*
Subpopulation: 2011 (*year*=**2011**)
Calculated Directly from Adjusted Files?: no

Code:
```
*step1: identify records with victimization characteristic(s) of
 interest on the modified incident-level file ;

data example7a_inc;
 set adjincident;
 vcrime=(newoff in (1 2 3 4)); *a;

 *identify location of crime;
 if v4024 in (1 2 3 4) then place_inc=1; *b;
 else if v4024 in (5 6 7) then place_inc=2; *c;
 else if v4024 in (8 9 10 11) then place_inc=3; *d;
 else if v4024 in (12 13 14 24 25 26 27) then place_inc=4; *e;
 else if v4024 in (15 16 17) then place_inc=5; *f;
 else if v4024 in (18 19) then place_inc=6; *g;
 else if v4024 in (20 21 22) then place_inc=7; *h;
 else if v4024 in (23) then place_inc=8; *i;
 else if v4024=. then place_inc=9; *j;

*collapse location of crime;
 if place_inc in (1 2) then place=1; *k;
 else place=2; *l;
```

```
 *create binary classifications for each location, excluding
  dummy records and crimes that occurred outside of the united
  states;
 violent_home=(vcrime=1 and place=1 and exclude_outus=0 and
              dummy=0);
 violent_other=(vcrime=1 and place=2 and exclude_outus=0 and
              dummy=0);
run;
```

Code Comment(s):
```
    a) Identify violent crimes
    b) Respondent's home
    c) Near respondents home
    d) Friend's home
    e) Commercial
    f) Parking lot-garage
    g) School
    h) Open areas, on street or public transportation
    i) Other
    j) Missing
    k) Respondent's home/near respondent's home
    l) Other location
```

```
*step2: create victimization summary file ;

proc sort data=example7a_inc(keep=yearq idper serieswgt
                             violent_home violent_other);
 by yearq idhh idper;
run;

*calculate the number of victimizations of each type by person
 and reporting period ;
proc means data=example7a_inc noprint;
 by yearq idper;
 weight serieswgt;
 var violent_home violent_other;
 output out=violent_place_sums(drop=_type_ _freq_)
       sum=;
run;

*step3: merge victimization summary file onto the person-level
        file ;

proc sort data=adjper; by yearq idper; run;
```

```
data example7a_per;
 merge adjper(in=a)
       violent_place_sums(in=b);
 by yearq idper;
 if a;
 if violent_home=. then violent_home=0; *a;
 if violent_other=. then violent_other=0;
run;
```

Code Comment(s):
    a) The two incident count variables are missing for persons not
       included on the victimization summary file, so they are set to '0'
       (no victimizations of this type)

```
*calculate victimization rates ;
proc sort data=example7a_per; by yr_grp v2117 v2118; run;

proc vargen data=example7a_per design=wr;
 subpopn year=2011;
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgtpercy;
 xmean vrviolrhome: adjinc_wt*violent_home*1000; *a;
 xmean vrviolother: adjinc_wt*violent_other*1000;

 setenv colwidth=15 decwidth=5;
 print estim seestim;
run;
```

Code Comment(s):
    a) The final step outlined in section 3.2 was to calculate the
       victimization adjustment factor, but it is already contained on the
       modified person file.

Output:

```
                         S U D A A N
           Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      August 2012
                        Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WGTPERCY
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    : 708493    Weighted count :1270197175
Number of observations skipped : 106187
(WEIGHT variable nonpositive)
Observations in subpopulation  : 143122    Weighted count :257542238
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.
```

```
---------------------------------------------------------
Variable
  SUDAAN Reserved
    Variable One          Estimate        SE Estimate
---------------------------------------------------------
VRVIOLRHOME
  Total                    8.33448            0.63137
  1                        8.33448            0.63137
---------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SUDAAN Reserved Variable One.

```
---------------------------------------------------------
Variable
  SUDAAN Reserved
    Variable One          Estimate        SE Estimate
---------------------------------------------------------
VRVIOLOTHER
  Total                   14.23472            1.15399
  1                       14.23472            1.15399
---------------------------------------------------------
```

Example 7B - Rate of property crimes by time of day, 2009-2011

Year(s): 2009-2011
Crime Type(s): property crimes by time of day (*property_day*; *property_night*; *property_missing*)
Domain(s): n/a
Weight: *wgthhcy*
Subpopulation: 2009-2011 ((**2009** <= *year*) and (*year* <=**2011**))
Calculated Directly from Adjusted Files?: no

```
*step1: identify records with victimization characteristic(s) of
 interest on the modified incident-level file ;
data example7b_inc;
 set adjincident;

 pcrime=(newoff in (6 7 8)); *a;

 *identify time of incident;
 if v4021b in (1 2 3 4) then time_day=1; *b;
 else if v4021b in (5 6 7 8) then time_day=2; *c;
 else if v4021b in (9 98) then time_day=3; *d;
 *create binary classifications for each time, excluding dummy
  records and crimes that occurred outside of the united states;
 property_day=(pcrime=1 and time_day=1 and exclude_outus=0 and
               dummy=0);
 property_night=(pcrime=1 and time_day=2 and exclude_outus=0 and
               dummy=0);
 property_missing=(pcrime=1 and time_day=3 and exclude_outus=0
               and dummy=0);

run;
```

Code Comment(s):
```
a)    identify property crimes
b)    day
c)    night
d)    missing
```

```
*step2: create victimization summary file ;
proc sort data=example7b_inc(keep=yearq idhh serieswgt
        property_day property_night property_missing);
 by yearq idhh;
run;
```

```
*calculate the number of victimizations of each type by household and
reporting period ;
proc means data=example7b_inc noprint;
 by yearq idhh;
 weight serieswgt;
 var property_day property_night property_missing;
 output out=property_time_sums(drop=_type_ _freq_)
        sum=;
run;


*step3: merge victimization summary file onto the HH-level
        file ;
proc sort data=adjhh; by yearq idhh; run;
data example7b_hh;
 merge adjhh(in=a)
       property_time_sums(in=b);
 by yearq idhh;
 if a;
 if property_day=. then property_day=0; *a;
 if property_night=. then property_night=0;
 if property_missing=. then property_missing=0;
run;
```

Code Comment(s):
   a) The three incident count variables are missing for households not
      included on the victimization summary file, so they are set to '0'
      (no victimizations of this type)

```
*calculate victimization rates ;

proc sort data=example7b_hh; by yr_grp v2117 v2118; run;

proc vargen data=example7b_hh design=wr;
 subpopn (2009 <= year) and (year <=2011);
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgtthhcy;
 xmean vrpropday: adjinc_wt*property_day*1000; *a;
 xmean vrpropnight: adjinc_wt*property_night*1000;
 xmean vrpropmiss: adjinc_wt*property_missing*1000;
 setenv colwidth=15 decwidth=5;
 print estim seestim;
run;
```

Code Comment(s):
   a) The final step outlined in section 3.2 was to calculate the
      victimization adjustment factor, but it is already contained on the
      modified person file.

Output:

```
                          S U D A A N
          Software for the Statistical Analysis of Correlated Data
         Copyright     Research Triangle Institute     August 2012
                         Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
     Sample Weight: WGTHHCY
     Stratification Variables(s): YR_GRP V2117
     Primary Sampling Unit: V2118


Number of observations read    : 398238    Weighted count :608895975
Number of observations skipped : 129435
(WEIGHT variable nonpositive)
Observations in subpopulation  : 239205    Weighted count :368251383
Denominator degrees of freedom :    160


Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SUDAAN Reserved Variable One.

---------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate      SE Estimate
---------------------------------------------------------

VRPROPDAY
   Total                  51.62050         1.22708
   1                      51.62050         1.22708
---------------------------------------------------------

Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SUDAAN Reserved Variable One.

---------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate      SE Estimate
---------------------------------------------------------
VRPROPNIGHT
   Total                  56.47746         1.33153
   1                      56.47746         1.33153
---------------------------------------------------------




Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: ( 2009 <= YEAR ) AND ( YEAR <= 2011 )
by: Variable, SUDAAN Reserved Variable One.

---------------------------------------------------------
Variable
   SUDAAN Reserved
     Variable One         Estimate      SE Estimate
---------------------------------------------------------
VRPROPMISS
   Total                  24.14251         0.80690
```

```
     1                         24.14251              0.80690
     ----------------------------------------------------------
```

D.4   Identifying Low Quality Estimates

This section demonstrates how to implement the recommendations for identifying low quality estimates and the rounding rules discussed in *Section 4* of the user's guide. Three types of data are needed to identify and flag low quality estimates, regardless of the estimate type: the estimate, the standard error of the estimate, and the unweighted sample size of the estimate. Because *Examples 1-7* provided details for obtaining estimates and standard errors, the examples in this section will focus primarily on calculating the percent relative standard error (RSE), calculating unweighted sample sizes, identifying estimates that should be flagged as unreliable, and verifying that estimates rounding to zero are not presented.

Unweighted sample sizes should take into account the series adjustment. For example, a series victimization with a series count of seven would count as seven victimizations in the unweighted sample size, while a non-series victimization would only count as one. This series adjustment is already incorporated in the victimization counts on the modified person and household-level files, so sample sizes for victimization rates are obtained by taking unweighted sums of victimization counts with the specified characteristic(s) of interest. However, for victimization totals and proportions, the series count must be included in the calculation of the sample size, as demonstrated in the examples below. Example 8A demonstrates implementation of flagging and rounding rules for victimization totals, Example 8B for victimization proportions, and Example 8C for victimization rates.

Example 8A - Total number of personal thefts by sex and race/ethnicity, 2011

Year(s): 2011
Crime Type: personal thefts (*ptft*)
Domain(s): sex (*sex*) and race/ethnicity (*race_eth*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=**0**); non-dummy records (*dummy*=**0**); 2011 (*year*=**2011**)

Code:

```sas
*calculate estimates and standard errors;
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  var ptft;
  class sex race_eth / nofreqs;
  tables sex race_eth;
  print total setotal;
run;
*calculate sample sizes;
proc descript data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011;
  nest yr_grp v2117 v2118 / psulev=3;
  weight serieswgt; *a;
  var ptft;
  class sex race_eth / nofreqs;
  tables sex race_eth;
  print total;
run;
```

Code Comment(s):

a) Weight by the series weight to obtain the series-adjusted count of victimizations

Output:

```
                         S U D A A N
         Software for the Statistical Analysis of Correlated Data
       Copyright      Research Triangle Institute      August 2012
                       Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118
Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :   7255    Weighted count : 23041441
Denominator degrees of freedom :    160
```

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, SEX.

-----------------------------------------------------------------------------
```

```
|                    |          | SEX                                                |
| Variable           |          |----------------------------------------------------|
|                    |          | Total       | Male        | Female      |
-------------------------------------------------------------------------------
|                    |          |             |             |             |
| PTFT               | Total    |   165770.15 |    53583.48 |   112186.67 |
|                    | SE Total |    24276.55 |    18913.90 |    19313.73 |
-------------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

```
-----------------------------------------------------------------------------------
|                    |          | RACE_ETH                                              |
| Variable           |          |-------------------------------------------------------|
|                    |          | Total       | Non-Hispanic | Non-Hispanic | Hispanic  |
|                    |          |             | White        | Black        |           |
-----------------------------------------------------------------------------------
|                    |          |             |              |              |           |
| PTFT               | Total    |   165770.15 |     83955.56 |     29506.98 |  30286.66 |
|                    | SE Total |    24276.55 |     21478.97 |     10989.44 |  10488.60 |
-----------------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

```
---------------------------------------------------------------------
|                    |          | RACE_ETH                          |
| Variable           |          |-----------------------------------|
|                    |          | Non-Hispanic | Non-Hispanic       |
|                    |          | Other        | More than          |
|                    |          |              | One Race           |
---------------------------------------------------------------------
|                    |          |              |                    |
| PTFT               | Total    |     13519.07 |            8501.87 |
|                    | SE Total |      5943.55 |            5281.78 |
---------------------------------------------------------------------
```

```
                        S U D A A N
         Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute      August 2012
                         Release 11.0.0
DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: SERIESWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118

Number of observations read    :  37853    Weighted count :     49229
Observations in subpopulation  :   7255    Weighted count :      8064
Denominator degrees of freedom :    160
Variance Estimation Method: Taylor Series (WR)

For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, SEX.
```

```
---------------------------------------------------------------------------
|               |          | SEX                                          |
| Variable      |          |----------------------------------------------|
|               |          | Total       | Male        | Female          |
---------------------------------------------------------------------------
|               |          |             |             |                 |
| PTFT          | Total    |       51.00 |       16.00 |           35.00 |
---------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

```
---------------------------------------------------------------------------------------
|               |          | RACE_ETH                                                 |
| Variable      |          |----------------------------------------------------------|
|               |          | Total    | Non-Hispanic | Non-Hispanic | Hispanic        |
|               |          |          | White        | Black        |                 |
---------------------------------------------------------------------------------------
|               |          |          |              |              |                 |
| PTFT          | Total    |    51.00 |        24.00 |         9.00 |            9.00 |
---------------------------------------------------------------------------------------
```

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011
by: Variable, RACE_ETH.

```
----------------------------------------------------------------------
|               |          | RACE_ETH                               |
| Variable      |          |----------------------------------------|
|               |          | Non-Hispanic | Non-Hispanic            |
|               |          | Other        | More than               |
|               |          |              | One Race                |
----------------------------------------------------------------------
|               |          |              |                         |
| PTFT          | Total    |         6.00 |                    3.00 |
----------------------------------------------------------------------
```

Identifying Unreliable Estimates:

As outlined in ***Section 4***, it is recommended that estimated totals meeting either of the following criteria be identified as unreliable. In addition, any estimate rounding to zero should not be displayed.

- RSE > 30%

- Count sample size ≤ 10

The percent relative standard error of an estimated total is calculated as:

$$\% \, RSE(Total) = \frac{SE(Total)}{Total} * 100$$

Estimated totals, standard errors, and unweighted sample sizes from the output above are included in ***Table D-1*** (in the "Total," "SE(Total)," and "n" columns, respectively). Percent RSEs are calculated based on the formula above. Based on the recommendations, five of the eight

estimates are flagged as unreliable. The estimate for males is flagged because the percent RSE is greater than 30 percent, while the other four estimates are flagged both because their RSEs exceed 30 percent and because their unweighted sample sizes are less than or equal to 10. All estimates can be displayed because no estimates round to zero.

Table D-1. Identifying Low Quality Victimization Totals

Domain	Total	SE(Total)	n	%RSE (Total)	Flag as Unreliable?	Rounds to Zero?
Overall	165770.15	24276.55	51	14.64		
Male	53583.48	18913.90	16	35.30	√	
Female	112186.67	19313.73	35	17.22		
Non-Hispanic White	83955.56	21478.97	24	25.58		
Non-Hispanic Black	29506.98	10989.44	9	37.24	√	
Hispanic	30286.66	10488.60	9	34.63	√	
Non-Hispanic Other	13519.07	5943.55	6	43.96	√	
Non-Hispanic More than One Race	8501.87	5281.78	3	62.12	√	

Example 8B – Distribution of motor vehicle theft in the Northeast Census Region by report status, 2011

Year(s): 2011
Domain(s): Northeast Census Region (*region*=1); motor vehicle theft (*mvtft*=1)
Covariate(s): report status (*notify*)
Weight: *newwgt*
Subpopulation: within the United States (*exclude_outUS*=0); non-dummy records (*dummy*=0); 2011 (*year*=2011); Northeast Census Region (*region*=1); motor vehicle theft (*mvtft*=1)

Code:
```
*calculate estimates and standard errors;
proc crosstab data=adjincident design=wr;
  subpopn exclude_outus=0 and dummy=0 and year=2011 and region=1 and
        mvtft=1;
  nest yr_grp v2117 v2118 / psulev=3;
  weight newwgt;
  class notify / nofreqs;
  tables notify;
  print rowper serow;
run;

*calculate sample sizes;
proc crosstab data=adjincident design=wr;
```

```
  subpopn exclude_outus=0 and dummy=0 and year=2011 and region=1 and
        mvtft=1;
  nest yr_grp v2117 v2118 / psulev=3;
  weight serieswgt; *a;
  class notify / nofreqs;
  tables notify;
  print wsum;
run;
```

Code Comment(s):

b) Weight by the series weight to obtain the series-adjusted count of victimizations. Because flagging rules are based on the denominator sample size, only the overall sample size is needed (not estimates by report status).

Output:

```
                        S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                        Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: NEWWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :115824878
Observations in subpopulation  :     19    Weighted count :    57944
Denominator degrees of freedom :    160

Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND MVTFT = 1 AND REGION = 1
by: NOTIFY.
```

			NOTIFY		
			Total	reported to police	not reported to police
		Row Percent	100.00	92.60	7.40
		SE Row Percent	0.00	5.40	5.40

```
                        S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      August 2012
                        Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: SERIESWGT
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118


Number of observations read    :  37853    Weighted count :    49229
```

```
Observations in subpopulation  :      19    Weighted count :        19
Denominator degrees of freedom :     160
```

```
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: EXCLUDE_OUTUS = 0 AND DUMMY = 0 AND YEAR = 2011 AND MVTFT = 1 AND REGION = 1
by: NOTIFY.
```

```
-------------------------------------------------------------------
|              |               | NOTIFY                            |
|              |               |---------------------------------- |
|              |               | Total   | reported | not          |
|              |               |         | to       | reported     |
|              |               |         | police   | to           |
|              |               |         |          | police       |
-------------------------------------------------------------------
|              |               |         |          |              |
|              | Weighted Size |      19 |       17 |          2   |
-------------------------------------------------------------------
```

Identifying Unreliable Estimates:

As outlined in *Section 4*, it is recommended that victimization percentages (P) meeting either of the following criteria be flagged as unreliable. In addition, any estimate rounding to zero should not be displayed.

- RSE > 30%, where
  - If the percentage is ≤ 50%, use the RSE of log(P)
  - If the percentage is > 50%, use the RSE of log(100-P); or
- Denominator sample size ≤ 10

The percent relative standard errors of the log of the percentages (P) and (1-P) are calculated as:

$$\% \, RSE(\log(P)) = \frac{SE(P)}{P * abs\left(\log\left(\frac{P}{100}\right)\right)} * 100$$

$$\% \, RSE(\log(100 - P)) = \frac{SE(P)}{(100 - P) * abs\left(\log\left(1 - \frac{P}{100}\right)\right)} * 100$$

Estimated percentages, standard errors, and the unweighted denominator sample size from the output above are included in *Table D-2* (in the "P," "SE(P)," and "n (denom)" columns, respectively). Percent RSEs are calculated based on the formulas above. Based on these recommendations, none of the estimates are flagged as unreliable. The denominator sample size is

19, which is greater than the recommended threshold of 10. The percent RSEs for the reported and not reported estimates are 28.03, which are below the 30 percent threshold. However, because no respondents reported that they did not know whether or not the motor vehicle theft was reported to the police, the estimated percent is 0. This estimate should not be presented because it rounds to zero.

Table D-2. Identifying Low Quality Victimization Percentages

Estimate	P	SE(P)	n (denom)	%RSE(P) OR %RSE(100-P)	Flag as Unreliable?	Rounds to Zero?
Reported to Police	92.60	5.40	19	28.03		
Not Reported to Police	7.40	5.40	19	28.03		
Do Not Know	0.00	-	19	-		√

Example 8C – Rate of rape/sexual assault by sex, 2011

Year(s): 2011
Crime Type(s): rape/sexual assault (*rsa*)
Domain(s): sex (*sex*)
Weight: *wgtpercy*
Subpopulation: 2011 (*year=2011*)
Calculated Directly from Adjusted Files?: yes

Code:
```
*calculate estimates and standard errors;
proc vargen data=adjper design=wr;
 subpopn year=2011;
 nest yr_grp v2117 v2118 / psulev=3;
 weight wgtpercy;
 xmean vrrsa: adjinc_wt*rsa*1000;
 class sex  / nofreqs;
 tables sex;
 setenv colwidth=15 decwidth=5;
 print estim seestim;
run;
*calculate sample sizes;
proc vargen data=adjper design=wr;
 subpopn year=2011;
 nest yr_grp v2117 v2118 / psulev=3;
 weight _one_; *a;
 xsum nrsa: rsa;
 class sex  / nofreqs;
 tables sex;
 setenv colwidth=15 decwidth=5;
 print estim;
run;
```
Code Comment(s):
   c) Unweighted because the victimization count RSA already includes the
      series adjustment

Output:

```
                          S U D A A N
           Software for the Statistical Analysis of Correlated Data
           Copyright     Research Triangle Institute     August 2012
                          Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: WGTPERCY
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118
```

```
Number of observations read     : 708493    Weighted count :1270197175
Number of observations skipped : 106187
(WEIGHT variable nonpositive)
Observations in subpopulation  : 143122    Weighted count :257542238
Denominator degrees of freedom :    160


Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
by: Variable, SEX.


---------------------------------------------------------
Variable
   SEX                 Estimate      SE Estimate
---------------------------------------------------------
VRRSA
   Total               0.94815          0.16253
   Male                0.27602          0.08970
   Female              1.59287          0.33234
---------------------------------------------------------


                       S U D A A N
          Software for the Statistical Analysis of Correlated Data
        Copyright     Research Triangle Institute      August 2012
                      Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a With
Replacement (WR) Design
    Sample Weight: _ONE_
    Stratification Variables(s): YR_GRP V2117
    Primary Sampling Unit: V2118
Number of observations read    : 814680    Weighted count :   814680
Observations in subpopulation  : 162867    Weighted count :   162867
Denominator degrees of freedom :    160


Variance Estimation Method: Taylor Series (WR)
For VARGEN Testing: YEAR = 2011
Estimate
by: Variable, SEX.


-----------------------------------------------------------------------------
Variable           SEX
                          Total           Male          Female
-----------------------------------------------------------------------------
NRSA                    79.00000       10.00000        69.00000
-----------------------------------------------------------------------------
```

Identifying Unreliable Estimates:

As outlined in *Section 4*, it is recommended that victimization rates meeting either of the following criteria be flagged as unreliable. In addition, any estimate rounding to zero should not be displayed.

- RSE > 30%

- Numerator sample size ≤ 10

The percent relative standard errors of the victimization rates (VR) are calculated as:

$$\% \, RSE(VR) = \frac{SE(VR)}{VR} * 100$$

Estimates, standard errors, and unweighted sample sizes from the output above are included in *Table D-3* (in the "VR," "SE(VR)," and "n (num)" columns, respectively). Percent RSEs are calculated based on the formula above. Based on these recommendations, the estimated victimization rate for males is flagged as unreliable because the percent RSE is greater than 30 percent and the sample size of the numerator is less than or equal to 10. Because none of the estimates round to zero, all three estimates can be presented.

Table D-3.   Identifying Low Quality Victimization Rates

Domain	VR	SE(VR)	n (num)	%RSE(VR)	Flag as Unreliable?	Rounds to Zero?
Overall	0.94815	0.16253	79	17.14		
Male	0.27602	0.0897	10	32.50	√	
Female	1.59287	0.33234	69	20.86		