

2018 National Sample Survey of Registered Nurses

Variance Estimation Guidance

The 2018 NSSRN utilized a sampling frame of registered nurses (RNs) built from a list compiled from the National Council of State Board of Nursing and from individual State Boards of Nursing. Sampling was done independently within each of the 50 states and the District of Columbia. There were two sampling strata per state: one for nurses that hold a nurse practitioner (NP) license and another for all other RNs. The sampling rates for RNs and NPs differed across and within each state in order to produce reliable estimates for RNs and NPs within each state.

Weighting

The 2018 NSSRN uses replicate weights to capture the effect of the sampling design on variances. Statistical software designed to work with a simple random sample and a single survey weight generally computes accurate point estimates; however, variances, standard errors, and confidence intervals are too small. Thus, estimates may appear to be statistically significant when they are not. This guide applies to any estimate including mean, median, sum, proportion, count, regression parameter, etc.

Final Adjusted Weight for the Sample (RKRNWGTA)

RKRNWGTA is a person-level full sample weight generated for responding RNs who are eligible to be included in the sample. This weight reflects the original probability of selection of the RN, adjusting for multiple license assignment and noninterview, and benchmarking to known control totals from the frame.

Weight for Replicate Samples (RKRNWGTA1-RKRNWGTA100)

RKRNWGTA1 to RKRNWGTA100 are the replicate weights for subsamples, which are used in the calculation of variances. When data are collected as part of a complex survey, it is often challenging to produce approximate unbiased design consistent with estimates of variance analytically. The variance of survey statistics for simple random sample are inappropriate and generally too small. A technique called **jackknife replication method** is used for estimating variances from the NSSRN.

Using Replicate Weights with Built-In SAS Procedures

- Specifying the VARMETHOD Option

To correctly use the SURVEYxxx procedures, the VARMETHOD= option must be added to the PROC statement to indicate the method used to develop the replicate weights. For the 2018 NSSRN, the option is always VARMETHOD=JACKKNIFE

Example:

```
PROC SURVEYMEANS DATA= NSSRN_2018_PUF_FINAL VARMETHOD= JACKKNIFE;
```

- Declaring the REPWEIGHTS Statement

To properly use the SURVEYxxx procedures, the REPWEIGHTS options must be specified. The REPWEIGHTS statement identifies the variables containing replicate weights. When using 2018 NSSRN data, the statement is

```
REPWEIGHTS RKRNWGTA1 - RKRNWGTA100;
```

- SAS Syntax

```
PROC SURVEYMEANS DATA= NSSRN_2018_PUF_FINAL VARMETHOD= JACKKNIFE;
WEIGHT RKRNWGTA;
REPWEIGHTS RKRNWGTA1 - RKRNWGTA100;
VAR AGE_PUF;
RUN;

PROC SURVEYFREQ NSSRN_2018_PUF_FINAL VARMETHOD= JACKKNIFE;
WEIGHT RKRNWGTA;
REPWEIGHTS RKRNWGTA1 - RKRNWGTA100;
TABLES ED_NDTYPE_PUF;
RUN;
```

Using Replicate Weights with Built-In STATA Procedures

- Declaring the repeated weight

To properly use the svyset command, the jkrweight () option must be specified. The jkrweight () identifies the variables containing replicate weights. When using 2018 NSSRN data, the command is

```
jkrweight(RKRNWGTA1 - RKRNWGTA100)
```

- Specifying the variance

To correctly use the svyset command, the vce() and mse options must be specified. When using the 2018 NSSRN data, the command is

```
vce(jackknife) mse
```

- STATA Syntax

```
svyset [pw= RKRNWGTA ], jkrweight( RKRNWGTA1 - RKRNWGTA100 ) vce(jackknife) mse
```

```
svy: mean AGE_PUF
```

```
svy: tabulate ED_NDTYPE_PUF
```