

2019 National Survey of Organ Donation Attitudes and Practices (NSODAP)

Data Operations Manual

Table of Contents

2019 National Survey of Organ Donation Attitudes and Practices (NSODAP)	1
Data Operations Manual.....	1
Fielding Period	2
Data Completeness	2
Data Accuracy	4
NSODAP Data Files	4
Data Dictionary	4
NSODAP Weights	4
Analysis in SAS.....	5
Interpreting Results	5
Documentation for Analytical Procedures.....	9
Analysis in Stata	10

This Data Operations Manual will guide the user to understand and use the 2019 NSODAP survey data.

Fielding Period

The NSODAP data collection plan specified the collection of 2,000 completed cases using telephone interviewing and an additional 8,000 completed cases using an online web panel. Telephone survey data collection began with a pilot data collection of 109 cases between November 12, 2018, and December 19, 2018. The remaining 1,891 telephone cases were collected between January 2, 2019, and March 9, 2019. Web panel survey data collection of 8,000 cases was conducted between March 1, 2019, and April 1, 2019.

Telephone interviewers received proprietary training from American Directions Research Group, including a guide for administering the organ donation questionnaire, including a phonetic pronunciation guide. A toll-free helpdesk line was used for this project (1-844-688-3737) to answer any respondent questions. Out of the approximately 30,000 calls placed, a total of 11 calls were received by the helpdesk for this project from 11/12/2019 to 3/10/2019. A full sample disposition for all phone numbers sampled is available in the file NSODAPSampleDisposition.xlsx.

Data Completeness

NSODAP respondents were free to answer or decline to answer any question. Even though respondents could decline to answer any question, the NSODAP experienced a high degree of completion. Telephone respondents answered approximately 95% of the questions, and web respondents answered 99% of the questions; respondents overall answered 98% of the questions. No imputation was conducted on missing data. Missing data are stored as blank values (“.”) in variables ending with “_miss” and as the code “99” in variables ending in “no_miss.” This allows responses of “99” to indicate “Don’t Know” or that the respondent refused to answer, and for percentages to sum to 100%, including missing values. To address data completeness, and meet the needs of different analysis types, the NSODAP dataset contains several variations of each variable, as shown in Table 1.

Access to legacy data for 2012 and 2005 is available by contacting donation@hrsa.gov.

Table 1. Variable Naming Convention Guide

Variable Naming Rule	Coding	Usage
VariableName Example: Q4	Original raw variable from survey software	Not advised except for quality check of other variables.
VariableName_miss Example: Q4_miss	Missing values are coded as blank	When a dummy missing code would interfere with results, such as for parametric statistics (regression, correlation).
VariableName_nomiss Example: Q4_nomiss	Missing values are recoded to "99" <i>If the question is based off of skip logic, and the question was not shown, the question is coded as blank rather than "99"</i>	When missing values should be included in the total frequencies, such as in a response table. Do not use for parametric statistics.
VariableName_bin Example: Q4_bin	4-point scale recoded to binary 2-point scale Missing values are coded as blank	When a binary variable is desired, such as a condensed frequency table, or a logistic regression, and dummy missing codes are not desired.
VariableName_bin_nomiss Example: Q4_bin_nomiss	4-point scale recoded to binary 2-point scale Missing values are recoded to "99" <i>If the question is based off of skip logic, and the question was not shown, the question is coded as blank rather than "99"</i>	When a binary variable is desired with dummy missing codes, such as condensed frequencies.

Data Accuracy

All survey responses are accurately recorded in the raw data files. Several quality assurance procedures were taken to ensure accuracy. These quality assurance steps include:

- Interviewer training with phonetic pronunciation guide
- Cognitive interviews to ensure questionnaire flow and respondent comprehension
- Test cases collected with computer-assisted telephone interviewing (CATI) software and web survey software and examined for accuracy
- Dummy-data generated by survey software to ensure accurate skip logic paths
- Survey software enforces valid inputs to survey questions.
- Thorough review of the data by survey research analysts
- All variable recoding, calculations, and analysis are conducted by SAS script to ensure data accuracy and reproducibility of the cleaned data file

NSODAP Data Files

The **nsodap_puf.sas7bdat** is the clean SAS data file available for public use.

Data Dictionary

All variables in the dataset are described in the accompanying data dictionary spreadsheet. This includes the variable order number, name, source, data type, length, SAS format, and label.

NSODAP Weights

There are three sets of weights for NSODAP that adjust the sample to national Census characteristics. Different weights are used for different purposes.

WTFINAL adjusts all 10,000 cases to national Census characteristics, and includes both telephone and web cases together. This is used for most results shown in the Report of Findings.

WTPHONE adjusts the 2,000 telephone cases to national Census characteristics, and does not include the 8,000 web cases for analysis. This is used to compare 2019 telephone results to prior organ donation surveys, which were conducted only by telephone. This is used in the Trends section of the Report of Findings, and should be used when making any comparisons to telephone-only surveys.

WTWEB adjusts the 8,000 web cases to national Census characteristics, and does not include the 2,000 telephone cases for analysis. This is used to compare against telephone cases to identify possible mode effects, and is included in the Trends section of the Report of Findings. If future NSODAP surveys are conducted by web-only, this weight should be used for comparison.

Analysis in SAS

The SAS code needed to create basic weighted frequencies is shown below.

The code below applies the survey weights to create basic frequencies of Q4, with missing values set to 99.

```
proc surveyfreq data=NSODAP ;  
weight WTFINAL ;  
tables q4_nomiss / cl ;  
run ;
```

The code below applies the survey weights to create basic frequencies of Q4, with missing values set to 99, cross-tabulated by Hispanic categories.

```
proc surveyfreq data=NSODAP ;  
weight WTFINAL ;  
tables  
Hispanic*q4_bin_nomiss / row cl chisq ;  
run ;
```

The code below uses the survey weights to conduct a logistic regression predicting Q4 as dichotomous with missing values set to blank, using demographic predictors sex and education level.

```
proc surveylogistic data=NSODAP ;  
weight WTFINAL ;  
CLASS sex educat ;  
model Q4_bin (event='Support/Strongly Support') = sex educat ;  
run ;
```

Interpreting Results

All results include both raw and weighted components. In the example below for Q4 from AllVars_Overall.html, the column “Frequency” indicates the raw number of survey responses. All columns other than “Frequency” are weighted. “Weighted Frequency” shows the frequency

adjusted for population weights, and the “Percent” column indicates weighted percent. In the Figure 1 below, 4,933 respondents “strongly supported” organ donation, with a weighted frequency of 4,806, representing 48.0644% weighted percentage.

Figure 1. Example One-Way Frequencies

Q4. In general, do you strongly support, support, oppose, or strongly oppose the donation of organs for transplantation?							
Q4_nomiss	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
Strongly Support	4933	4806	65.44842	48.0644	0.5949	46.8983	49.2306
Support	4203	4235	65.08128	42.3539	0.5892	41.1989	43.5090
Oppose	431	473.15271	27.59943	4.7315	0.2723	4.1977	5.2654
Strongly Oppose	268	308.90389	22.99949	3.0890	0.2278	2.6426	3.5355
[Don't Know / Refused]	165	176.10965	16.32253	1.7611	0.1626	1.4423	2.0799
Total	10000	10000	64.92062	100.000			

Most results for NSODAP are two-way, cross-tabulating the survey responses by a particular demographic. In Figure 2 on the following page, taken from AllVars_Hispanic_Bin.html, Q4 is condensed to a binary variable, and missing values are coded as 99, labeled “Don’t know/Refuse,” and the survey responses are split by Hispanic ethnicity. Like the one-way frequencies, “Frequency” is the raw result, and all other results are weighted to the U.S. adult population characteristics. Instead of using the “Percent” column like in one-way frequencies, “Row Percent” is the most informative column for two-way frequencies. In the example below, 91.2751% of Non-Hispanic respondents supported organ donation, and 86.4292% of Hispanic respondents supported organ donation.

Figure 2. Example Two-Way Frequencies

Table of Hispanic by Q4_bin_nomiss												
Hispanic	Q4_bin_nomiss	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent		Row Percent	Std Err of Row Percent	95% Confidence Limits for Row Percent	
Not Hispanic	Oppose/Strongly Oppose	579	579.67570	27.84068	5.7968	0.2762	5.2554	6.3381	7.0417	0.3318	6.3914	7.6921
	Support/Strongly Support	8176	7514	60.48061	75.1377	0.5786	74.0034	76.2719	91.2751	0.3652	90.5593	91.9909
	Don't know/Refuse	141	138.55882	13.74810	1.3856	0.1372	1.1166	1.6546	1.6832	0.1663	1.3572	2.0092
	Total	8896	8232	59.38447	82.3200	0.5453	81.2512	83.3888	100.000			
Hispanic	Oppose/Strongly Oppose	120	202.38090	22.57947	2.0238	0.2234	1.5860	2.4617	11.4469	1.1965	9.1014	13.7923
	Support/Strongly Support	960	1528	55.55055	15.2807	0.5135	14.2741	16.2873	86.4292	1.2695	83.9408	88.9176
	Don't know/Refuse	24	37.55083	8.85751	0.3755	0.0884	0.2022	0.5489	2.1239	0.4964	1.1509	3.0969
	Total	1104	1768	59.99415	17.6800	0.5453	16.6112	18.7488	100.000			
Total	Oppose/Strongly Oppose	699	782.05660	35.51718	7.8206	0.3468	7.1408	8.5004				
	Support/Strongly Support	9136	9042	66.68755	90.4183	0.3765	89.6803	91.1564				
	Don't know/Refuse	165	176.10965	16.32253	1.7611	0.1626	1.4423	2.0799				
	Total	10000	10000	64.92062	100.000							

All two-way frequencies include a Chi-Square test immediately following the two-way frequency table. The Rao-Scott Chi-Square accounts for survey design effects and is most appropriate to use on survey data. The Chi-Square example in Figure 3 follows the frequency table in Figure 2. It shows the Chi-Square is 19.9745, with $p < .0001$. This indicates support for organ donation is significantly different between Hispanic and non-Hispanic respondents. Note that Chi-Squares cannot be calculated if any cell in the frequency table has zero cases. To calculate a Chi-Square in these cases, the categories should be recoded to collapse or otherwise eliminate the zero-frequency cell.

Figure 3. Example Chi-Square for Two-Way Frequencies

Rao-Scott Chi-Square Test	
Pearson Chi-Square	41.4988
Design Correction	2.0776
Rao-Scott Chi-Square	19.9745
DF	2
Pr > ChiSq	<.0001
F Value	9.9873
Num DF	2
Den DF	19998
Pr > F	<.0001
Sample Size = 10000	

Documentation for Analytical Procedures

Full SAS documentation can be found online for the two main SAS procedures used to analyze NSODAP survey data:

PROC SURVEYFREQ

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_surveyfreq_sect002.htm

PROC SURVEYLOGISTIC

https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_surveylogistic_sect002.htm

Analysis in Stata

The NSODAP files are written for SAS, but an exported Stata dataset is also provided. The Stata code needed to create basic weighted frequencies is shown below.

The code below applies the survey weights to create basic frequencies of Q4, with missing values set to 99.

```
svyset [pweight=wtfinal]
svy: tab q4_nomiss
```

The code below applies the survey weights to create basic frequencies of Q4, with missing values set to 99, cross-tabulated by race categories.

```
svyset [pweight=wtfinal]
svy, subpop(RaceCat): tab q4_nomiss
```

The code below uses the survey weights to conduct a logistic regression predicting Q4 as dichotomous with missing values set to blank, using demographic predictors sex and education level.

```
svyset [pweight = wtfinal]
svy: logit q4_bin sex educat, or
```