2025 National Survey of Organ Donation Attitudes and Practices (NSODAP)

Data Operations Manual

Table of Contents

This Data Operations Manual will guide the user to better understand and use the 2025 NSODAP survey data. The data is described in terms of its timeliness, completeness, and accuracy. The data files, cleaning and recoding process, and analysis are described here.

Fielding Period

The NSODAP data collection plan specified the collection of 2,001 completed cases using telephone interviewing and an additional 8,007 completed cases using an online web panel. Telephone survey data was collected between September 2024 and January 2025. Web panel survey data was collected in January and February 2025.

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.

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 datasets have a high degree of completeness. 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 prior survey years 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" *If the question is based off of skip logic, and the question was not shown, the question is coded as blank rather than "99"*	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" *If the question is based off of skip logic, and the question was not shown, the question is coded as blank rather than "99"*	When a binary variable is desired with dummy missing codes, such as condensed frequencies.
VariableName_misszero Example: Q4_misszero	4-point scale recoded to binary 2-point scale Missing values are recoded to "0" *If the question is based off of skip logic, and the question was not shown, the question is coded as blank rather than "0"*	To calculate significant differences between group scores, where missing codes may generate errors for Chi-Squares, such as zero-cell errors. For consistency in comparing to 2019, missing values are a part of proportion scores and are included in the denominator.

Data Accuracy

All survey responses are accurately recorded in the 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

In the final dataset, raw data is cleaned, labeled, and recodes create new calculated variables. Population weights are applied. The file NSODAP_2025_final is the cleaned and final dataset.

SAS File Locations

To function correctly, the NSODAP SAS files must be structured in four directories, the locations of which can be defined by the user at the very beginning of the master SAS script NSODAP_Analysis.sas:

Data all data files (ending in .sas7dbat) go in this directory

Formats all data format files (ending in .sas7bcat) go in this directory

Output all HTML output files appear in this directory

Programs all scripting files (ending in .sas) go in this directory

In the example below, these four directories are located in "C:\NSODAP\SAS\..." but the user can relocate these directories as needed, and update the location text to the correct locations.

```
/* UPDATE LOCAL FILE LOCATIONS HERE */

%LET DataLocation       =C:\NSODAP\SAS\Data\;
%LET FormatsLocation    =C:\NSODAP\SAS\Formats\;
%LET OutputLocation     =C:\NSODAP\SAS\Output\;
%LET ProgramLocation    =C:\NSODAP\SAS\Programs\;
```

NSODAP Data Files

NSODAP_2025_final_codes.xlsx is the raw data file of all cases, including partial and incomplete cases, in Excel format. Values are recorded as numeric codes identical to the SAS datafiles. A value of "#NULL!" indicates missing data. The variable FinalComplete indicates whether a case is complete, which is defined as finishing the telephone interview or clicking the "submit" button on the web survey.

NSODAP_2025_final_labels.xlsx is the excel data file of all final cases, including partial and incomplete cases, in Excel format. Values are recorded as response labels rather than numeric codes. This dataset is identical to the other Excel file, except it uses response labels instead of the numeric values.

NSODAP_2025_final.sas7bdat is the clean data file for the 2025 NSODAP. Labels and recodes are already applied.

NSODAP_2025_only_verbatims.sas7bdat is the data file of open text responses with respondent ID.

NSODAP_2025_with_verbatims.sas7bdat is the data file of open text responses merged with the clean survey data. This file is large size is not recommended for non-qualitative analysis.

NSODAP SAS Programs

NSODAP_Frequencies.sas creates all of the one-way and two-way frequencies for all survey and analytical questions. Grouping variables include gender, age, race, ethnicity, education, geographic region, state, religion, health status, and health insurance status. Additional grouping variables can be added by invoking the FrequenciesBy Macro at the end of the script. This will create a set of frequencies including all response options and another set of frequencies collapsed to a binary set of response options. For example, the invocation below is used to create frequencies of all survey questions by race categories:

```
%FrequenciesBy(ByVar=RaceCat);
```

To create frequencies by income groups, the user would add this line to the end of the script:

```
%FrequenciesBy(ByVar=D9IncomeLevel);
```

Note that a grouping variable cannot also be included as an analytical variable. If the user wishes to split by support for organ donation (Q4), the user would need to remove that item from the macro temporarily or create a new variable that represent Q4. The output is saved as a series of HTML files, with one output file per grouping variable. These files can also be opened in Excel. This script should be included if the user wants to generate frequency results.

NSODAP_Regressions.sas calculates the regression models featured in the report using PROC SURVEYLOGISTIC. The output is saved as a series of HTML files, which can also be opened in Excel. This script should be included if the user wants to generate regression models.

Interpreting Results

All results include both raw and weighted components. In the example below for Q4, 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, 9,256 respondents "supported" or "strongly supported" organ donation, with a weighted frequency of 9,274, representing 92.6700% weighted percentage. The proportion supporting organ donation and the confidence interval for this support are highlighted in Figure 1.

Figure 1. Example One-Way Frequencies

The SURVEYFREQ Procedure

q4. In general, do you strongly support, support, oppose, or strongly oppose the donation of organs for transplantation?

Q4_bin_nomiss	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
Do not support	653	613.74639	33.92582	6.1326	0.3334	5.4790	6.7861
Support	9256	9274	93.51564	92.6700	0.3656	91.9533	93.3867
Missing	99	119.84469	16.14433	1.1975	0.1606	0.8826	1.5124
Total	10008	10008	93.71394	100.0000			

Most results for NSODAP are two-way, cross-tabulating the survey responses by a particular demographic. In Figure 2 on the following page, 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

The SURVEYFREQ Procedure

Table of Income by Q4_bin_nomiss

Income	Q4_bin_nomiss	Frequency	Weighted Frequency	Std Err 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	
Income: Less than $30,000	Do not support	265	293.15625	26.05190	3.2022	0.2810	2.6514	3.7530	12.3992	1.0304	10.3794	14.4189
	Support	1909	2043	61.02079	22.3189	0.6151	21.1132	23.5246	86.4201	1.0772	84.3086	88.5316
	Missing	16	27.91561	8.78087	0.3049	0.0958	0.1172	0.4927	1.1807	0.3689	0.4577	1.9037
	Total	2190	2364	65.83649	25.8260	0.6497	24.5524	27.0997	100.0000			
Income: $30,000 to $59,000	Do not support	143	120.39739	13.80768	1.3151	0.1506	1.0199	1.6104	5.0362	0.5684	3.9221	6.1503
	Support	2141	2240	61.79434	24.4659	0.6236	23.2434	25.6884	93.6905	0.6658	92.3854	94.9955
	Missing	23	30.44106	8.74445	0.3325	0.0954	0.1455	0.5195	1.2733	0.3631	0.5615	1.9852
	Total	2307	2391	63.33370	26.1135	0.6358	24.8673	27.3598	100.0000			
Income: $60,000 to $99,999	Do not support	93	72.26893	10.77207	0.7894	0.1176	0.5590	1.0199	3.5593	0.5238	2.5325	4.5860
	Support	1972	1940	56.78626	21.1919	0.5849	20.0453	22.3384	95.5497	0.6026	94.3685	96.7309
	Missing	12	18.09181	6.27995	0.1976	0.0686	0.0632	0.3320	0.8910	0.3078	0.2876	1.4945
	Total	2077	2030	57.80680	22.1789	0.5939	21.0147	23.3431	100.0000			
Income: $100,000 or more	Do not support	90	65.67408	8.66647	0.7174	0.0948	0.5315	0.9033	2.7718	0.3639	2.0585	3.4851
	Support	2479	2283	56.27547	24.9380	0.5925	23.7765	26.0995	96.3544	0.4252	95.5209	97.1878
	Missing	26	20.70596	5.32366	0.2262	0.0582	0.1122	0.3402	0.8739	0.2241	0.4347	1.3131
	Total	2595	2369	56.80722	25.8816	0.5988	24.7078	27.0553	100.0000			
Total	Do not support	591	551.49665	32.23847	6.0241	0.3464	5.3451	6.7032				
	Support	8501	8506	89.57258	92.9146	0.3779	92.1738	93.6555				
	Missing	77	97.15443	14.85218	1.0612	0.1616	0.7445	1.3780				
	Total	9169	9155	89.82102	100.0000							

Frequency Missing = 839

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 128.1264 with $p < .0001$. This indicates support for organ donation is significantly different between income levels among 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. Chi-squares are calculated using the **_misszero* versions of the variable, with missing values collapsed into the negative response (e.g., missing counts as does not support). This is done to avoid the chi-square testing the frequency of missing values, to be comparable to 2019 results, and to avoid zero-cell count errors.

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

The SURVEYFREQ Procedure

Rao-Scott Chi-Square Test	
Pearson Chi-Square	218.001
Design Correction	1.7015
Rao-Scott Chi-Square	128.1264
DF	3
Pr > ChiSq	<.0001
F Value	42.7088
Num DF	3
Den DF	27504
Pr > F	<.0001
Sample Size = 9169	

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

Data Dictionary

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

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_miss sex educat, or
```