IDEA Data Center

Survey Response Analysis App

Reference Guide

Westat®

The IDEA Data Center (IDC) created this publication under U.S. Department of Education, Office of Special Education Programs grant number H373Y240002. Amy Bae and Alexis Lessans serve as the project officers.

Westat is the lead organization for IDC. For more information about the center's work and its partners, see www.ideadata.org.

April 2025

Suggested Citation:

Nimkoff, T., & Schneider, B. (2025, April). Survey Response Analysis App *Reference Guide*. IDEA Data Center. Rockville, MD: Westat.

Table of Contents

Tables

Figures

Survey Response Analysis App Reference Guide

Introduction: Why This App?

The U.S. Department of Education's Office of Special Education Programs (OSEP) allows states to collect data for certain indicators of the State Performance Plan/Annual Performance Report (SPP/APR) using survey methods. Each year, in addition to reporting on progress against set performance targets for these indicators, states must report on the quality of that year's respondent data, including survey response rates, representativeness, and—beginning with the 2022 APR submission—nonresponse bias.[1] Specifically, for Indicators C4. Family Involvement, B8. Parent Involvement, and B14. Post-School Outcomes, states must address the following:

- overall response rate and response rates by key respondent subgroups (including race/ethnicity and at least one additional, stakeholder-approved variable);
- extent to which the demographics of respondents are representative of the target population, including a description of the metric the state used to assess representativeness;
- analysis of the response rate, including any nonresponse bias the state identified, and the steps the state has taken to reduce nonresponse bias and promote responses from a broad cross section of the target population; and
- strategies the state will implement to increase the response rate year over year, particularly for underrepresented subgroups, and to ensure that future response data are representative of those subgroups.

Response rates, data representativeness, and nonresponse bias are distinct, but related, concepts. **Response rates** reflect the percentage of people asked to respond to a survey who actually do respond. The degree to which those survey respondents proportionally replicate the target population indicates the **representativeness** of the responses. **Nonresponse bias** is a type of statistical bias—that is, a systematic error in a survey that causes the survey estimate to be too high or too low—that results from nonresponse to the survey. When certain subgroups are less likely to respond to a survey, resulting in their systematic underrepresentation in the survey data, and those underrepresented subgroups differ from other subgroups in what the survey is trying to measure, nonresponse bias affects the data. Note especially that statistical bias is a form of

[1] U.S. Department of Education. (n.d.). *Part B Sate Performance Plan (SPP) and Annual Performance Report (APR) Indicator Measurement Table: FFYs 2020–2025. (FFY 2023 Submission).* Washington, D.C.: Author. https://sites.ed.gov/idea/files/FFY2023-Part-B-SPP-APR-Reformatted-Measurement-Table.pdf

systematic error, rather than random error. This means states can expect nonresponse bias to occur again if they repeat the survey with a different random sample or in another year.

IDEA Data Center (IDC) developed the *Survey Response Analysis App* (SRA App) to help states examine the quality of their survey data. As an interactive application powered by state-of-the-art statistical software, the SRA App allows users to conduct reproducible analyses of response rates, representativeness, and nonresponse bias, tailored to their survey's data collection method.

States can use the SRA App to answer questions such as the following:

- What are our response rates, and do they differ across subgroups?
- Are some subgroups in the population overrepresented or underrepresented in our respondent data?
- How do survey outcomes differ across subgroups?
- Can statistical adjustments reduce nonresponse bias?

The SRA App provides a point-and-click user interface for analyzing data within your preferred web browser, such as Google Chrome or Microsoft Edge. Although you interact with the application through your web browser, your data remains on your local machine; the app will not transmit your data elsewhere. Thus, you can be assured that confidential data are securely stored and analyzed only within your organization's systems.

Finally, note that while IDC designed this reference guide to help users, and while we intend for the application itself to be as user-friendly as possible, the SRA App employs sophisticated statistical methods that require an intermediate-to-advanced understanding of underlying concepts. All screenshots in this document were taken from the SRA App. For technical assistance using this tool, please contact your IDC State Liaison or email IDEAdata@westat.com.

The remainder of this guide provides instructions for using the app, including information on the statistical concepts behind selecting and interpreting each analysis option. This guide does not provide recommendations for constructing surveys or include practice-based strategies for increasing response rates or representativeness.

Getting Ready to Use the SRA App: Preparing Your Dataset

To get the most from the SRA App, IDC recommends that your dataset includes certain elements for use within the application. You also will need certain variables in your dataset and may add additional variables if you wish to perform certain analyses.

In the companion document to this reference guide, *Getting Ready to Use the* Survey Response Analysis App*: Preparing Your Dataset*, you will find a detailed guide for ensuring that the dataset you use includes the necessary variables to make the most of the app's analysis options. We recommend reviewing this document prior to using the app.

Using the SRA App: Setup

The SRA App opens to the Welcome tab, which provides an overview of the tool. When you are ready to use the app, you will move from the Welcome tab to the Setup module. The instructions in this section of this reference guide cover the steps you will take to load a prepared dataset into the application in preparation for statistical analyses.

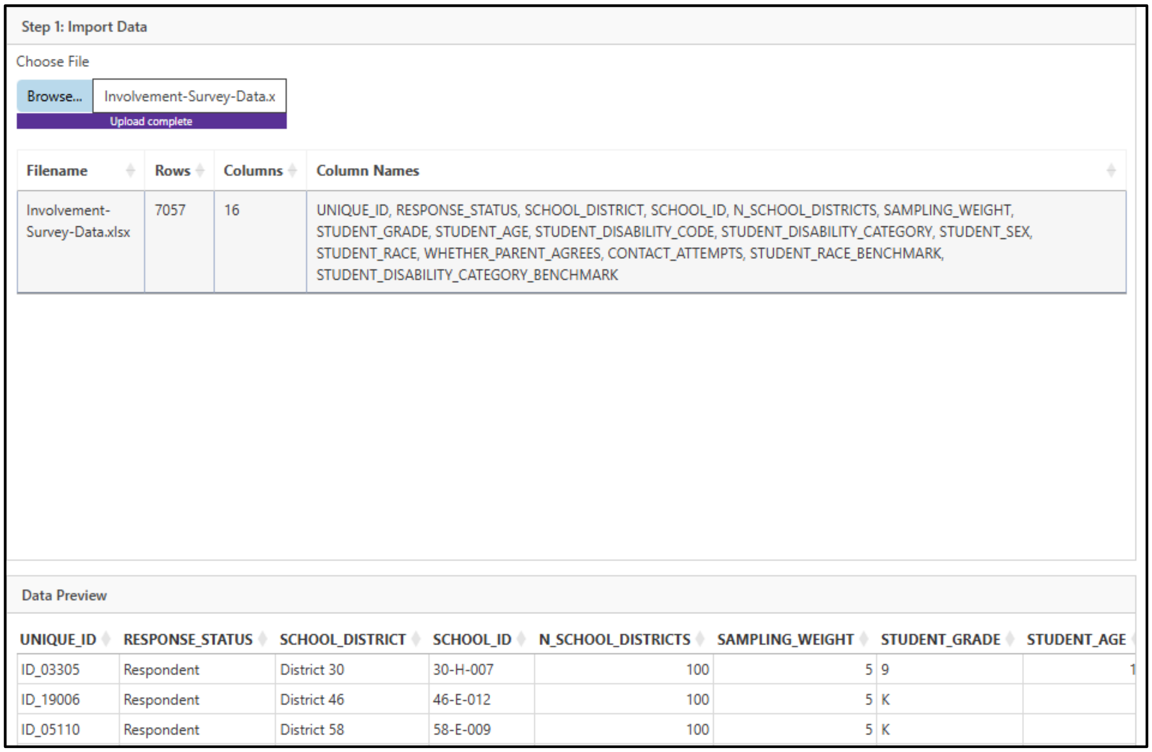
Note that the application also includes written instructions covering the steps of each module. In addition, built-in tool tips provide more information and definitions of key terms. You can access these tool tips by hovering over the text within the application.

Step 1: Import Data

The first step in the Setup module is to load a prepared dataset into the app. Click the Browse button to locate the dataset (in CSV, Excel, SPSS, or SAS format) located on your local computer. Then, select the file you want to import. See the companion document to this reference guide, *Getting Ready to Use the Survey Response Analysis App: Preparing Your Dataset*, for guidelines on setting up your dataset for optimal performance.

Next, as shown in figure 1, the application will provide a summary of the contents of the dataset along with a preview of the first few rows and columns of the dataset.

Figure 1. Screenshot of dataset summary within the SRA App



Note that importing your data into the SRA App does not upload those data to the web. Although you interact with the application through your web browser, your data remain solely on your local machine; the application does not transmit your data elsewhere. Thus, you can be assured that confidential data are securely stored and analyzed only within your organization's systems.

> **Keep in Mind**
>
> Because the application does not save the information you enter during each web session, the next time you launch the app to begin a new session, you will need to complete the Setup module again.

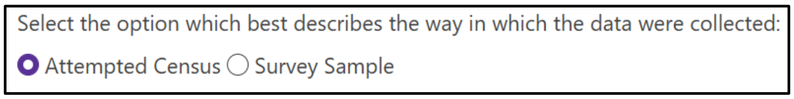
Step 2: Describe the Dataset

After loading your desired dataset into the app, you must answer a series of prompts that identify how the state collected those data. This will allow the application to properly conduct the analyses that you select. Read on to learn what different answers to each of these prompts mean in the context of your data and the application.

Identify Whether the Data Were Collected Using an Attempted Census or a Survey Sample

First, indicate whether the data come from an attempted census or a survey sample, as illustrated in figure 2. In an attempted census, you seek responses from every member of the target population; for example, you sent a survey invitation to the parent of every child receiving special education services in the state. While only a fraction of the population might respond, the key feature of an attempted census is that you sought responses from 100 percent of the target population. In contrast, with a survey sample, you seek responses from a subset of the population, such as the parents of a subset of randomly selected students from the full list of students receiving special education services in the state.

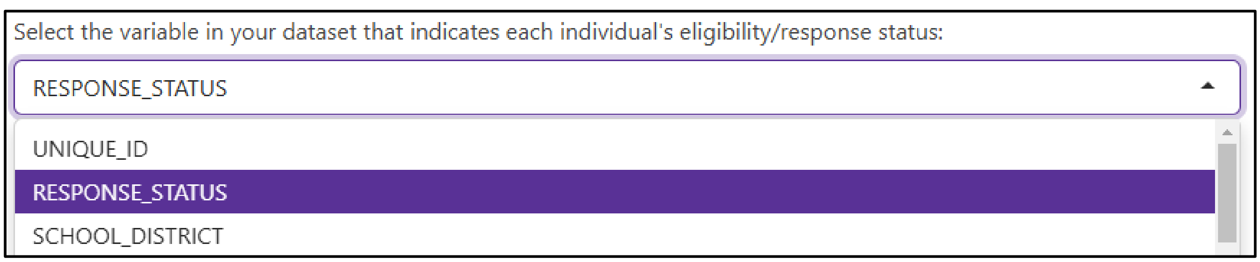
Figure 2. Screenshot of survey data collection prompt within the SRA App



Identify the Response and Eligibility Status Variable in Your Dataset

Next, you must identify the variable in your dataset that indicates each person's response and eligibility status for the survey, as illustrated in figure 3. This variable classifies each person in the data based on whether the person responded to the survey and was eligible to do so. This variable should have at most four distinct categories: eligible respondents, eligible nonrespondents, cases known to be ineligible, and cases whose eligibility status is unknown. (See the companion resource titled *Getting Ready to Use the* Survey Response Analysis App*: Preparing Your Dataset* for more information on how to construct a variable for response and eligibility status.)

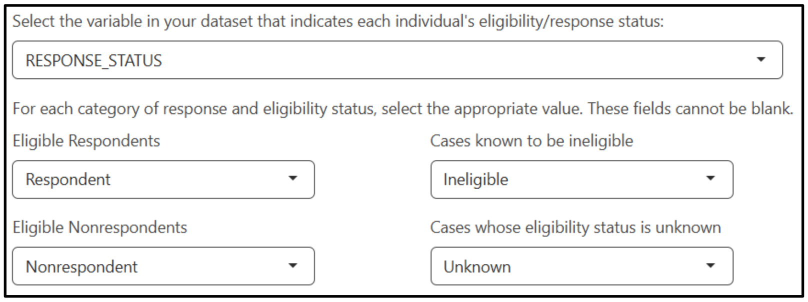
Figure 3. Screenshot of response and eligibility status variable prompt within the SRA App



Once you identify which variable within your dataset indicates response and eligibility status, indicate how the application should interpret each category of that response and eligibility status variable. The application lists four possible categories that the variable may contain. For each category, use the drop-down list to indicate the value of the response and eligibility status variable corresponding to that category, as shown in figure 4.

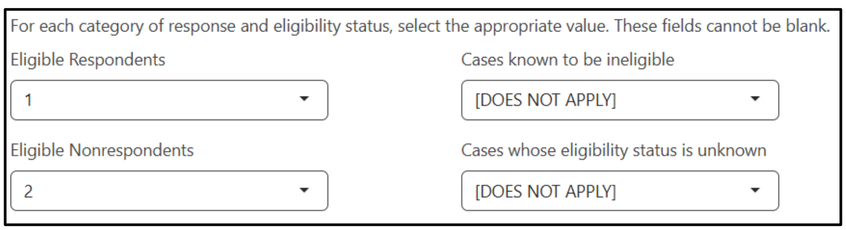
If you do not have cases known to be ineligible or cases with unknown eligibility in your dataset, then your response and eligibility status variable will have values only for eligible respondents or eligible nonrespondents, and you will simply select does not apply for the other two categories. Do not leave the fields blank.

Figure 4. Screenshot of drop-down response and eligibility category prompts within the SRA App



Note that for response and eligibility status, as well as all other variables in your dataset, you can use numeric values or unique abbreviations to reflect each value if app user(s) know what those numbers or abbreviations represent. Figure 5 illustrates the use of numeric values within a variable.

Figure 5. Screenshot illustrating use of numeric variable values within the SRA App



Indicate How the Application Should Handle Persons With Unknown Eligibility

Next, if you indicated that your data include cases whose eligibility to complete the survey is unknown, the application will ask, “Should cases with unknown eligibility be grouped with nonrespondents for all analysis types other than response rates?” In other words, when conducting analysis types other than the calculation of response rates, you must indicate whether you want to group cases with unknown eligibility together with eligible nonrespondents.

If most cases with unknown eligibility are likely to be eligible for the survey, then it is useful to treat unknown eligibility cases as eligible nonrespondents for most analysis types. However, when calculating response rates, it is still helpful to distinguish between eligible nonrespondents and unknown eligibility cases. If you believe that most cases with unknown eligibility in your dataset are likely to be eligible for the survey, then IDC recommends you select yes. If you believe that most cases with unknown eligibility are likely to be ineligible for the survey, then IDC recommends you select no.

Step 3: Enter Additional Items for a Survey Sample, if Applicable

If you indicate that your data were collected as a survey sample, the application will provide additional fields where you provide details about your sampling procedures. (If data were collected as an attempted census, proceed to Step 4: Proceed to Analysis.)

Identify Any Population Size Variables in Your Dataset

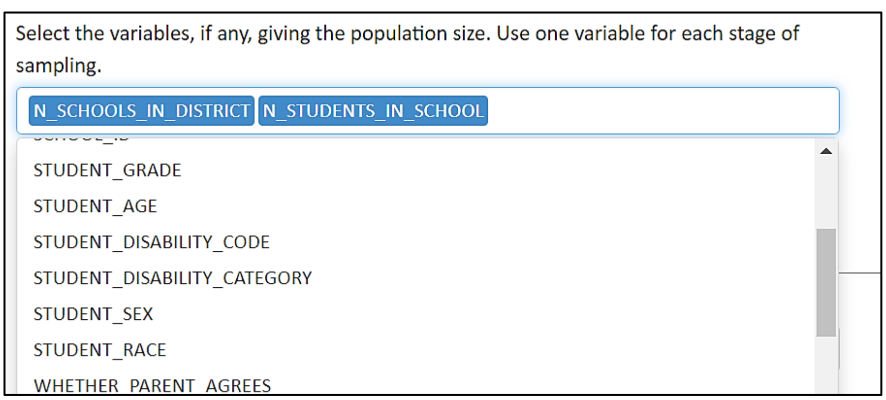
If you sampled a large fraction of the population (>20%) for the survey, then the app can take this into account when calculating statistics such as *p*-values or confidence intervals. You can select variables in your dataset, if any, that identify the population size. When there are no variables in your dataset that provide population size, simply leave the default selection: No Population Size Variables.

Note that your population size variable will vary, depending on the type of sampling strategy you used according to the following:

- If you used stratified sampling, the population size variable should contain the population size for each stratum.
- If you used cluster sampling, the population size should be the number of clusters in the population rather than the number of individuals.
- If you used multistage cluster sampling, you need to select one population size variable for each stage of sampling.

For example, figure 6 shows an example of how to select the population size variables for a multistage cluster sample in which the state first sampled schools from districts and then sampled students from the selected schools.

Figure 6. Screenshot of population size variable prompt within the SRA App

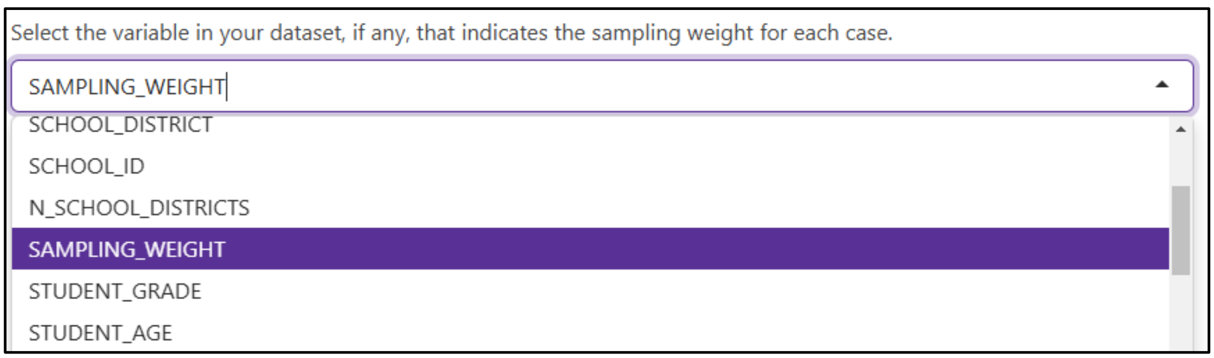


Select the Sampling Weight Variable, if Any, in Your Dataset

The app will allow you to select a variable in your dataset that indicates the sampling weight for each case. When you assume each person had an equal probability of being sampled, simply leave the default selection: no weights.

If the sampling method caused some members of the population to have a larger chance of being sampled compared to other members of the population, you will need to identify which variable from your dataset the application should use to give each person a sampling weight, as illustrated in figure 7. To calculate the sampling weight for a given sampled person, divide 1 by that person's probability of being selected into the sample. The sampling weight variable should not have any missing or negative values.

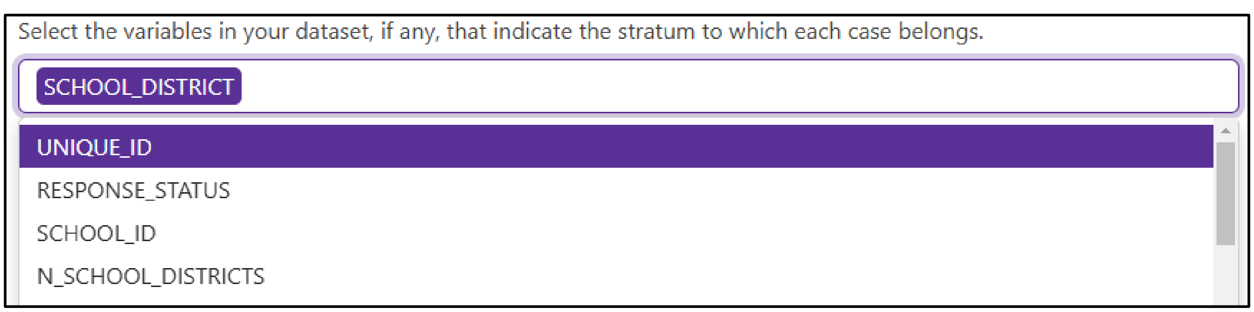
Figure 7. Screenshot of sampling weight variable prompt within the SRA App



Select Any Sampling Strata Variables in Your Dataset

Stratification involves dividing the population into subgroups with predetermined sample sizes. If you do not have strata variables in your dataset, simply leave the default selection: no strata variables. If you used stratified sampling for your survey, select the variable that divides the population into strata, as illustrated in figure 8.

Figure 8. Screenshot of stratification variable prompt within the SRA App

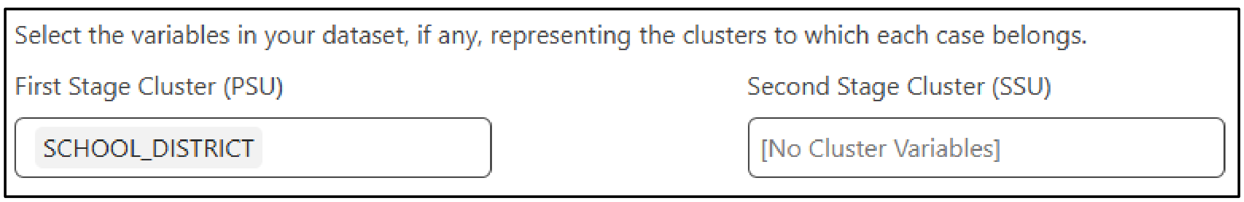


If there are multiple stages of stratified sampling, then select one stratification variable for each stage of sampling. For example, a state may divide its area into geographic regions, randomly select a fixed number of school districts from within each region, and then randomly select a fixed number of students from within each of the sampled school districts. In this example, region and district are the strata the state used for sampling, and you would need to select both stratification variables from the pulldown menu.

Select Any Cluster Sampling Variables in Your Dataset

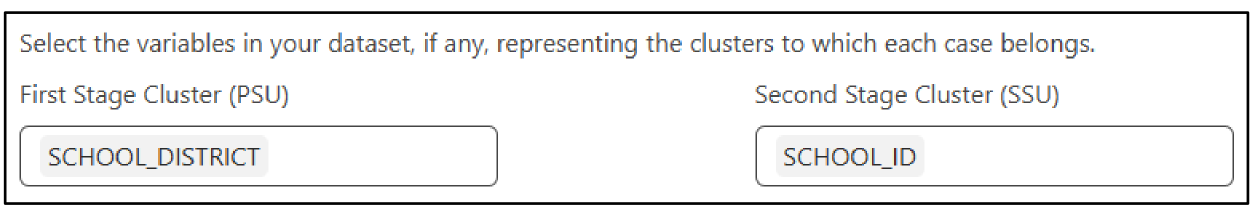
Cluster sampling involves using random sampling to select a group (i.e., cluster) of individuals instead of directly sampling the individuals themselves. If you do not have cluster variables in your dataset, simply leave the default selection: no cluster variables. If you do have cluster variables, select those variables in your dataset representing the clusters to which each case belongs. If you have only one cluster variable, identify it in the field labeled first-stage cluster or primary sampling unit (PSU), as illustrated in figure 9. Leave the second-stage cluster or secondary sampling unit (SSU) field set to the default: no cluster variables.

Figure 9. Screenshot of cluster sampling variable prompt within the SRA App



Cluster sampling also may occur in multiple stages. If this is the case for your dataset, you will also need to select a second-stage cluster (SSU). Figure 10 shows this process, illustrating the selections you would need to make if your state selected a random sample of school districts as the first-stage cluster variable and then selected a random sample of schools from each of the sampled districts as the second-stage cluster variable.

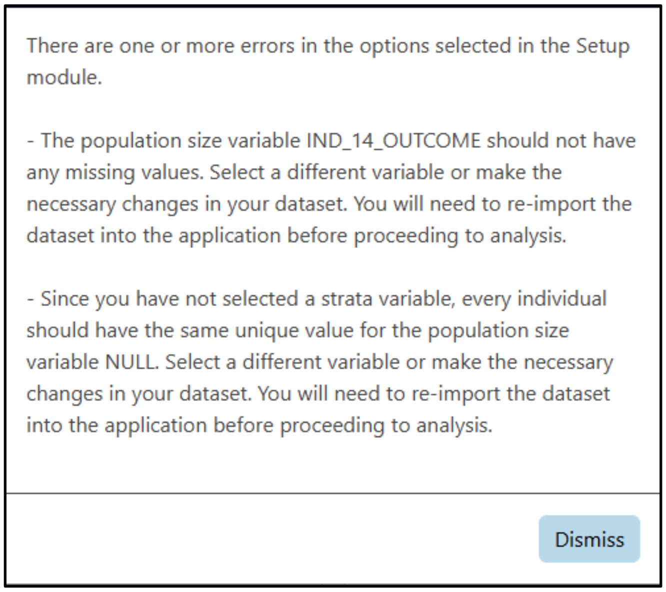
Figure 10. Screenshot illustrating use of second-stage cluster variable within the SRA App



Step 4: Proceed to Analysis

Once you have imported and described your data in the Setup module, select the Proceed to Analysis button. The application will check for errors in the Setup options that you selected before proceeding to the Analysis module. If it detects any errors, it will display an error message that is customized to the field(s) in which the errors occur, as shown in figure 11.

Figure 11. Screenshot illustrating an error message within the SRA App



Once you have reviewed the error message, you can dismiss it to return to the Setup options and correct the errors. After you have made any necessary corrections, select the Proceed to Analysis button again and the application will proceed to the Analysis module.

Remember: the application does not save the information you enter during each web session. The next time you launch the SRA App to begin a new session, you will need to complete the Setup module again for that new session.

Using the SRA App: Analysis and Report

The SRA App allows you to conduct analyses to answer questions about response rates, representativeness, and nonresponse bias in your survey data and to save the results of your analyses in an exportable report. While the Analysis and Report modules are separate tabs in the application, the modules work together, and so we describe them together in this guide.

Once you have proceeded to the Analysis module, you can select from a variety of analysis types based on four key questions:

- What are our response rates, and do they differ across subgroups?
- Are some subgroups in the population overrepresented or underrepresented in our respondent data?
- How do survey outcomes differ across subgroups?
- Can statistical adjustments reduce nonresponse bias?

When you select an analysis type from the menu, a new panel, titled Specify Analysis, will appear with options for conducting that specific analysis. The application will recommend some of these options as defaults. Then, once you submit your options, a pop-up window will appear, showing a table with the resulting statistics for the analysis. Once the analysis is complete, you can add the output to the Report module. Afterwards, you can repeat the same analysis with different options (e.g., calculating response rates separately by race/ethnicity, disability category, and grade level) or select a different analysis type to run and then add those results to the Report module as well.

Keep in Mind

Determining which analysis option(s) to run depends both on the specific questions you want to answer and the data that you have available in your dataset. For example, all analysis options related to calculating response rates require data on nonrespondents, as does comparing subgroup percentages in respondent data to data from respondents and nonrespondents.

Contact your IDC State Liaison for help and advice on determining which analyses to run based on your state's data.

You can also save and view the output tables you added to the Report module as an Excel workbook. To do so, select the Report module. When you do so, a list of the analyses you've added to the report will appear in the Items in Report display. You can remove any items you decide no longer belong in the report and save the remaining output tables by clicking Save to Excel. When you do so, a pop-up window will appear showing the Excel report saved to your local computer. You can now open the Excel file to view your results or save the workbook in a new location and with a new file name for later identification. The Excel file will contain a README tab with basic information about the report, a Setup tab which documents your selections from the Setup module, as well as a tab for each output table that you selected to include in the report.

As long as you keep the SRA App running, you can return to the Setup or Analysis modules to conduct new analyses to add to the Report module. However, remember that if you close the application you will need to repeat the setup process again.

The next section of this guide describes each of the nine analysis options available in the SRA App, grouped by key analysis question, as well as the resulting output from each analysis, which you can save and export in the Report module.

What Are Our Response Rates and Do They Differ Across Subgroups?

The starting point of any nonresponse bias analysis is to calculate response rates, since nonresponse bias can only occur when a survey's response rate is below 100 percent. The analyses in this section address federal fiscal year (FFY) 2020–2025 SPP/APR requirements related to reporting the overall survey response rate and response rates across subgroups, including whether different subgroups are more or less likely to respond to the survey.

Calculate Response Rates by Subgroup

This analysis provides the overall survey response rate as well as a comparison of response rates across subgroups within a selected variable (e.g., race/ethnicity, exit reason), referred to as a grouping variable in the app. The grouping variable should be based on information you have available for the entire population or for the full sample of individuals you invited to complete the survey.

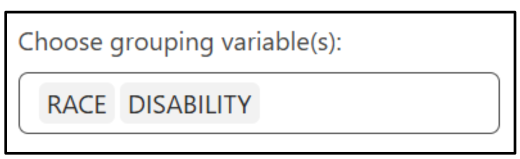
To calculate response rates, your dataset should have one row for every person invited to submit a survey, regardless of whether that person ultimately responded.

Choose the Grouping Variable(s) to Analyze

The application can calculate a single, overall response rate or calculate response rates separately for different subgroups of a selected variable. To calculate an overall response rate, simply leave the Choose grouping variable(s) field empty.

To assess whether different subgroups are more or less likely to respond to the survey by comparing their response rates, select one or more grouping variables from your dataset. Select a single grouping variable (e.g., race/ethnicity) if you want to compare subgroup response rates within that variable. Then, you can repeat the analysis with a different grouping variable. Select multiple grouping variables, such as race/ethnicity subgroup and disability category, as shown in figure 12, to calculate response rates for each combination of values from the grouping variables.

Figure 12. Screenshot of grouping variable prompt for calculating response rates within the SRA App

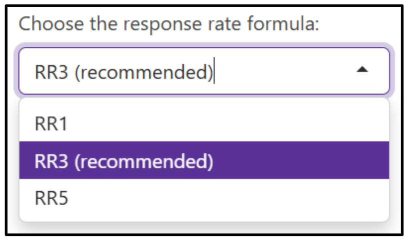


Choose the Response Rate Formula

If you have cases with unknown eligibility in your dataset, you can choose how the application handles those cases for response rate analysis. The American Association for Public Opinion Research (AAPOR) identifies several standard formulas for calculating response rates, three of which are programmed into the application. These formulas differ in how they handle cases whose eligibility status is unknown.

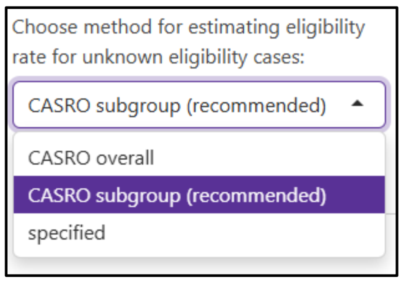
As seen in figure 13, the application recommends RR3 for most purposes, which uses an estimate of the eligibility rate among cases with unknown eligibility. Alternatively, you can choose RR1, which assumes that all cases with unknown eligibility are eligible, or RR5, which assumes that all cases with unknown eligibility are ineligible.

Figure 13. Screenshot of response rate formula prompt within the SRA App



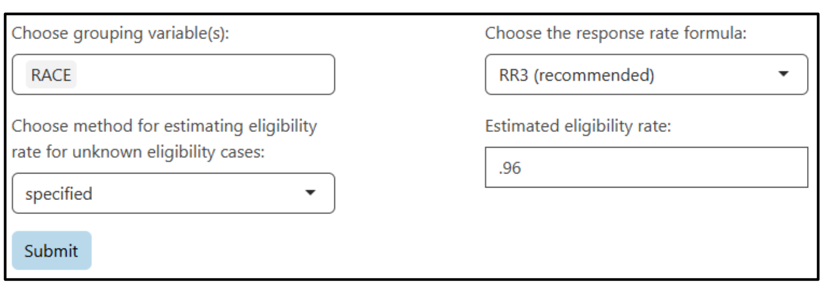
If you use RR3, you can also choose the method for estimating eligibility rate, as seen in figure 14. The application recommends a method proposed by the Council of American Survey Research Organizations (CASRO), referred to as CASRO subgroup, which uses the eligibility rate among cases with known eligibility status to calculate estimates separately for each subgroup.

Figure 14. Screenshot of method for estimating eligibility rate prompt within the SRA App



As also shown in figure 14, there are additional options for estimating eligibility rate: CASRO overall and specified. If you select CASRO overall, the app will assume the eligibility rate is constant across all subgroups and use this single estimate when calculating response rates for subgroups. If you instead select specified, you can specify an estimated eligibility rate for the application to use. To do so, first select specified, then enter a number between 0 and 1 to use as the estimated eligibility rate for cases with unknown eligibility status, as seen in figure 15.

Figure 15. Screenshot illustrating use of a specified eligibility rate within the SRA App



Note that if your dataset does not include cases with unknown eligibility status, you do not need to choose the response rate formula or the method for estimating eligibility rate for unknown cases. Simply leave the default selections RR3 (recommended) and CARSO subgroup (recommended) in place, as they will not affect the analyses.

When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis, as shown in figure 16.

Figure 16. Screenshot of an output table produced for calculating response rates within the SRA App

	RACE	Response Rate (Unweighted)	Total sample size	Number of eligible respondents	Number of eligible nonrespondents	Number of ineligible cases	Number of unknown eligibility cases	Estimated eligibility rate for unknown eligibility cases (unweighted)
1	American Indian or Native Alaskan	**88.89%**	20	16	2	2	0	96.00%
2	Asian	**93.75%**	18	15	1	2	0	96.00%
3	Black	**82.19%**	78	60	13	5	0	96.00%
4	Hawaiian or other Pacific Islander	**100.00%**	14	12	0	2	0	96.00%
5	Hispanic	**70.45%**	46	31	13	2	0	96.00%
6	Two or more races	**86.05%**	46	37	6	3	0	96.00%
7	White	**82.26%**	278	218	47	13	0	96.00%

+ Add to Report   Close

Add Your Output to the Report

Once you have submitted an analysis, you can add the analysis output to the Report module and then close the pop-up table when finished, or you can simply close the table if you do not want to add it to the Report module. You can repeat the same response rate analysis with a different grouping variable or move to a different analysis option.

Keep in Mind

You can save and view the output tables you added to the Report module as an Excel workbook. Just select the Report module, and a list of the analyses you've added to the report will appear in the Items in Report display. You can remove any items you decide no longer belong in the report and save the remaining output tables by clicking Save to Excel.

Table 1 shows an example of an output table once you have added it to the Report module and saved it to Excel. Please note that this sample table, and all tables in this guide, make use of fictitious data for the purpose of illustrating the functionality of the SRA App. This output table illustrates the output results from calculating overall response rate within the SRA App.

Table 1 shows the unweighted overall response rate (refer to first column, Response rate [Unweighted]), as well as the total number of cases in the dataset (refer to second column, Total sample size), broken down into the numbers of eligible respondents, eligible nonrespondents, ineligible cases, and cases with unknown eligibility (refer to third through sixth columns, respectively). If applicable, the output table also provides the percentage of cases with unknown eligibility that the application estimates are eligible for the survey, based on the response rate formula you specified during analysis (refer to seventh column, Estimated eligibility rate for unknown eligibility cases). Note that E*MAPS*, the SPP/APR data submission system, will auto-calculate overall response rate based on your state's entry of eligible respondents and total target population or sample size. You can use results from the SRA App to confirm that calculation.

Table 2 illustrates the type of output table that results when the application evaluates whether different subgroups (in this case, student race/ethnicity; refer to first column) are more or less likely to respond to the survey by comparing their unweighted response rates (refer to second column, Response Rate [Unweighted]). In this example, students who are Hispanic or Latino have a lower number of eligible respondents compared with their total sample size than other subgroups. Specifically, Hispanic or Latino students have a lower response rate (31.90%) compared to other subgroups, whose response rates range from 66.47 percent (students who are Asian) to 87.50 percent (students who are Native Hawaiian or Other Pacific Islander). This output may be informative because subgroups with lower response rates are at risk of being underrepresented in the survey data.

As a final example, table 3 uses additional data to show the comparison of response rates across subgroups within the student disability category variable. Response rates range from 52.66 percent (students with hearing impairments) to 69.97 percent (students with multiple disabilities).

Table 1. Sample output table—Calculate overall response rate

Response rate (unweighted)	Total sample size	Number of eligible respondents	Number of eligible nonrespondents	Number of ineligible cases	Number of unknown eligibility cases	Estimated eligibility rate for unknown eligibility cases (unweighted)
63.39%	7,057	4,255	2,111	327	364	95.11%

Table 2. Sample output table—Calculate response rates by subgroup (student race/ethnicity)

Student race/ethnicity	Response rate (unweighted)	Total sample size	Number of eligible respondents	Number of eligible non-respondents	Number of ineligible cases	Number of unknown eligibility cases	Estimated eligibility rate for unknown eligibility cases (unweighted)
AM (American Indian or Alaska Native)	**68.55%**	64	39	17	7	1	88.89%
AS (Asian)	**66.47%**	70	39	18	11	2	83.82%
BL (Black or African American)	**69.63%**	958	632	216	47	63	94.75%
HI (Hispanic or Latino)	**31.90%**	1,023	309	601	51	62	94.69%
MU (Two or More Races)	**74.78%**	176	124	39	10	3	94.22%
PI (Native Hawaiian or Other Pacific Islander)	**87.50%**	35	28	4	3	0	91.43%
WH (White)	**68.19%**	4,731	3,084	1,216	198	233	95.60%

Table 3. Sample output table—Calculate response rates by subgroup (student disability category)

Student disability category	Response rate (un-weighted)	Total sample size	Number of eligible respondents	Number of eligible non-respondents	Number of ineligible cases	Number of unknown eligibility cases	Estimated eligibility rate for unknown eligibility cases (unweighted)
Autism	**63.88%**	301	183	93	14	11	95.17%
Developmental delay	**66.77%**	484	307	130	23	24	95.00%
Emotional disturbance	**64.54%**	166	100	40	10	16	93.33%
Hearing impairment	**52.66%**	58	30	25	1	2	98.21%
Intellectual disability	**58.86%**	386	218	137	15	16	95.95%
Multiple disabilities	**69.97%**	144	95	37	8	4	94.29%
Orthopedic impairment	**58.33%**	36	21	10	0	5	100.00%
Other health impairment	**63.70%**	617	378	176	22	41	96.18%
Specific learning disability	**63.04%**	2,805	1,687	847	122	149	95.41%
Speech or language impairment	**63.61%**	1,984	1,190	595	108	91	94.29%
Traumatic brain injury	**64.09%**	38	23	11	2	2	94.44%
Visual impairment	**64.19%**	38	23	10	2	3	94.29%

Test Whether Subgroups Differ in Likelihood of Responding

When you find differences in response rates across subgroups, you can test whether subgroups differ in their likelihood of responding to the survey. Differences in response rates across subgroups may be due to randomness or to systematic differences between subgroups in their likelihood of responding to the survey.

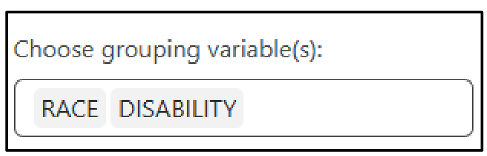
This analysis uses a chi-squared test of independence to evaluate the likelihood that any observed difference in response rates between the subgroups is due to systematic differences between groups rather than randomness of the sample being analyzed (see statistical significance in the Glossary for more information). The analysis assesses if there is a statistically significant association between variables—for example, between race/ethnicity and likelihood of responding to the survey.

For this analysis, your dataset should have one row for every person invited to submit a survey, regardless of whether that person ultimately responded.

Choose the Grouping Variable(s) to Analyze

To begin, select one or more variables in the data that divide the sample into different subgroups (e.g., race/ethnicity), and whose values you know for both respondents and nonrespondents, as seen in figure 17. Grouping variables can be either numeric (e.g., age in years) or categorical (e.g., primary disability category).

Figure 17. Screenshot of grouping variable prompt for testing whether subgroups differ in likelihood of responding within the SRA App



For each grouping variable, the application will conduct a chi-squared test to assess whether the subgroups defined by that variable have significantly different likelihoods of responding to the survey. When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

Once you have submitted an analysis, you can add the analysis output to the Report module and then close the pop-up table when finished, or you can simply close the table if you do not want to add it to the Report module. You can then repeat the same analysis with a different grouping variable to see if other subgroups differ in their likelihood of responding to the survey, or you can select a different analysis type to run.

Table 4 shows an example of the resulting output table for this analysis type once you have added it to the Report module. For each grouping variable you have entered into the analysis (refer to first column, Auxiliary variable), the application will provide the resulting *p*-value (refer to fifth column, *p*-value). The output table also will include supplemental information about the specific statistical test the application used in the analysis.

Table 4. Test whether subgroups differ in likelihood of responding, by student race/ethnicity and student disability category

Auxiliary variable	Test statistic	Numerator degrees of freedom	Denominator degrees of freedom	*p*-value	Name of test	Method of variance estimation
Student race/ ethnicity	87.5711	6	38,190	**0**	Rao-Scott chi-square test (second-order adjustment)	linearization
Student disability category	1.3121	11	70,015	**0.20998**	Rao-Scott chi-square test (second-order adjustment)	linearization

In a statistical hypothesis test, a result is statistically significant if a *p*-value falls below a certain predetermined threshold, such as 0.05. In table 4, the result for the variable student race/ethnicity is statistically significant ($p < 0.05$), while the result for student disability category is not statistically significant ($p > 0.05$). These results indicate subgroups within the student race/ethnicity variable have systematic differences in their likelihood of responding to the survey. That is, one or more race/ethnicity subgroups included in the analysis showed significantly different response rates compared to the other race/ethnicity subgroups. In interpreting the results of a statistical significance test, it is important to note that a statistically significant difference may not be practically significant; that is, the magnitude of differences between subgroups may be small despite being statistically significant. For this reason, you should also consider the actual response rates across subgroups in your data.

FYI: Interpreting the Results of a Significance Test

The term **statistically significant** does not necessarily mean that an observed pattern is of practical importance, but simply that the pattern is unlikely to be a result of chance. For example, for large datasets where differences in groups can be precisely measured, a small difference between groups may be statistically but not practically significant.

The SRA App includes several analyses that report the results of a significance test, such as the chi-squared test of independence and the *t* test. The primary output of a significance test is a *p*-value, which is a statistic whose value ranges between 0 and 1. A high *p*-value indicates that an observed pattern in the data could easily have occurred purely due to random chance, while a low *p*-value indicates that an observed pattern would rarely occur purely due to random chance. Many organizations use a standard for interpreting *p*-values, known as statistical significance, where they describe *p*-values below a certain threshold (e.g., below 0.05) as statistically significant.

Identify Variables That Predict Likelihood of Responding

Certain variables may be highly correlated with the likelihood of responding to a survey and therefore can serve as good predictors when modeling survey response. When you find differences in response rates across subgroups within a variable, you can also test whether that variable predicts—or is related to—whether an individual is a respondent rather than a nonrespondent.

This analysis uses logistic regression to help identify variables in your dataset that predict the response indicator. If a grouping variable in your dataset effectively predicts response to the survey, subgroups within that variable may be at risk for being underrepresented among respondents, which is one factor in the potential for nonresponse bias with respect to that variable.

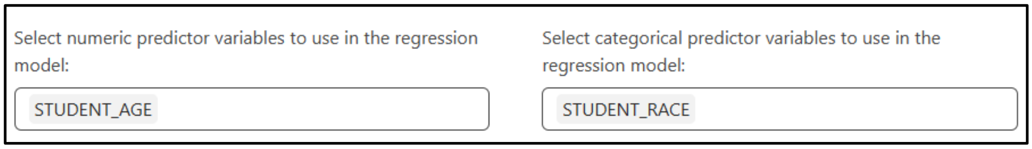
In contrast to the analysis based on a chi-squared test, the regression analysis can examine multiple variables simultaneously, so you can assess whether a variable’s association with response to the survey is independent of other variables in the model. For example, you can assess whether student race/ethnicity is predictive of whether a parent responds independently from student age.

For this analysis, your dataset must have one row for every person invited to submit a survey, regardless of whether that person ultimately responded.

Select the Grouping Variable(s) to Analyze

To begin, select the grouping variables to use as predictors in the regression model. As seen in figure 18, you can select multiple predictor variables and the variables can be either numeric (i.e., continuous) or categorical, but you must specify the numeric and categorical variables separately.

Figure 18. Screenshot of grouping variable prompt for identifying variables that predict likelihood of responding within the SRA App

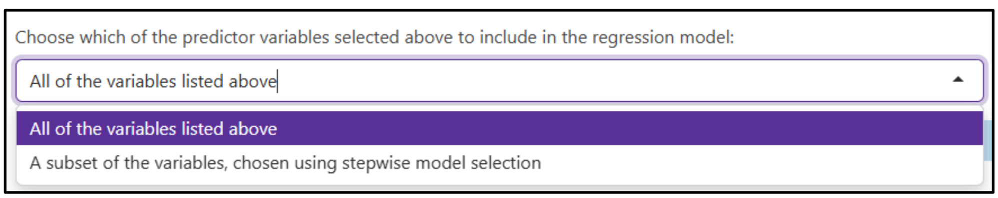


For each categorical variable of *n* categories, the application uses the first category that appears in the data as the reference level to create *n*-1 dummy variables for each of the remaining categories and uses the dummy variables as the predictors in the logistic model. (Note: you can control which category the application uses as the reference variable by alphabetizing the categories.) The output table for this analysis type will indicate which category the application used as the reference level for each categorical variable.

Choose Predictor Variables to Include in the Regression Model

After choosing your grouping variables, indicate how the app should determine which grouping variables to include in the regression model. You can include all the variables listed above in the model, or you can use a stepwise model selection method, which may reduce your list to a subset of the variables, as seen in figure 19.

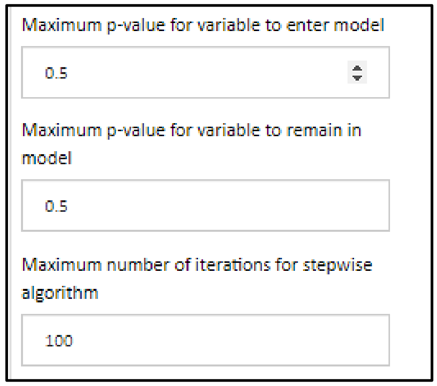
Figure 19. Screenshot of likelihood of responding predictor variables prompt within the SRA App



If you use the stepwise model selection, then the application will add and remove variables in the model using an iterative procedure. At each iteration, the application will add a variable to the model if the variable has a sufficiently small *p*-value, and it will remove a variable if the variable has too large a *p*-value. Therefore, you can specify the maximum *p*-value a predictor variable can have in order to enter the model at a given iteration of the stepwise model selection algorithm. Next, you can specify the maximum *p*-value a predictor variable can have in order to remain in the model at a given iteration of the stepwise model selection algorithm. Finally, you can specify the maximum number of iterations of the procedure.

As seen in figure 20, the application provides default values for this purpose; however, you may also specify how you would like the application to conduct the procedure by changing those defaults.

Figure 20. Screenshot of likelihood of responding stepwise model selection prompts within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

Once complete, you can add the analysis output to the Report module and then close the table when finished. Next, you can repeat the same analysis with different variables and parameters, or you can select a different analysis type to run.

Table 5 shows an example of the resulting output table for this analysis type once you have added it to the Report module. For each grouping variable you have included as predictors in the regression model (refer to first column, Predictor variable), the app will provide the resulting variable-level *p*-values (refer to second column, Variable-level *p*-value), as well as the estimated coefficient (refer to fourth column4, Estimated coefficient) and coefficient *p*-values (refer to ninth column, Coefficient *p*-value from *t* test). For categorical predictor variables, the output table includes a note at the bottom identifying the specific categories that the application used as reference levels for the regression. (Note: you can control which category the app uses as the reference variable by alphabetizing the categories.) The output table also will include supplemental information about the specific statistical test the app used in the analysis.

A statistically significant variable-level *p*-value indicates that the specific grouping variable effectively predicts response to the survey, given the other variables included in the regression model. This implies that subgroups within that variable are at risk for being underrepresented among respondents, which is one factor in the potential for nonresponse bias with respect to that variable.

In table 5, for example, the *p*-value for the variable student age is not statistically significant, indicating student age is not predictive of whether an individual is a respondent rather than a nonrespondent in this dataset, whereas the *p*-value for the variable student race is statistically significant ($p < 0.05$), indicating student race is predictive of whether an individual is a respondent rather than a nonrespondent for this survey. Specifically, coefficient *p*-values from the *t* test indicate that the HI (Hispanic or Latino) subgroup within the student race grouping variable effectively predicts the likelihood of responding to the survey. Note that while the output table includes the intercept, an intercept *p*-value is not meaningful in interpretation of the results.

Table 5. Prediction of survey response status, by student age and student race

Predictor variable	Variable-level *p*-value	Category of categorical variable	Estimated coefficient	Standard error	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval	Coefficient *p*-value from *t* test	Likelihood ratio test chi-squared statistic	Likelihood ratio test DEff	Likelihood ratio test numerator df	Likelihood ratio test denominator df
(Intercept)			**0.8039**	0.2961	0.2234	1.3843	**0.00665**				
Student age	**0.67691**		**-0.0028**	0.0069	-0.0163	0.0107	**0.68257**	0.1685	1.0074	1	6,722
Student race	0	AS	**-0.1073**	0.3963	-0.8842	0.6696	**0.78656**	475.9552	1.0017	6	6,722
Student race	0	BL	**0.0460**	0.2942	-0.5307	0.6226	**0.87582**	475.9552	1.0017	6	6,722
Student race	0	HI	**-1.5379**	0.2935	-2.1133	-0.9625	**0.00000**	475.9552	1.0017	6	6,722
Student race	0	MU	**0.3051**	0.3369	-0.3552	0.9654	**0.36511**	475.9552	1.0017	6	6,722
Student race	0	PI	**1.1817**	0.6076	-0.0094	2.3728	**0.05184**	475.9552	1.0017	6	6,722
Student race	0	WH	**-0.0184**	0.2871	-0.5812	0.5443	**0.94877**	475.9552	1.0017	6	6,722

NOTE: For categorical predictor variables, the application used following categories as reference levels for the regression—Student race: AM (American Indian or Alaska Native).

Are Some Subgroups in the Population Overrepresented or Underrepresented in Our Respondent Data?

For Indicators 8 and 14, the FFY 2020–2025 SPP/APR requires states to report on the extent to which respondent data are representative of their target population across key demographic variables, including race/ethnicity and at least one additional stakeholder-informed variable. The analyses in this section compare subgroup percentages among your survey respondents to subgroup percentages in the target population.

You can determine whether observed differences between subgroup percentages among respondents and subgroup percentages in the target population are indicative of some subgroups being overrepresented or underrepresented in the survey based on your state's metric. When subgroups, such as students who are Hispanic or Latino, are systematically underrepresented in your survey data, there is a potential for nonresponse bias with respect to that variable. The application can compare subgroup percentages in your respondent data to (a) subgroup percentages from all eligible cases in your dataset (i.e., respondents and nonrespondents) or (b) subgroup percentages from an external dataset reflecting population data.

Compare Subgroup Percentages in Respondent Data to Data From Respondents and Nonrespondents

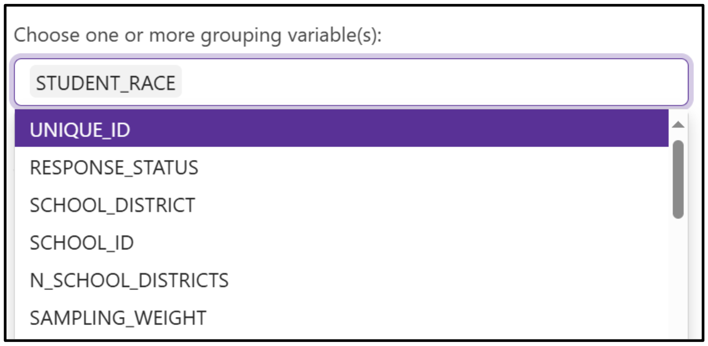
This analysis calculates subgroup percentages in your respondent data, subgroup percentages from all eligible cases, and the resulting percentage difference. The analysis also uses a *t* test to assess whether observed differences between subgroup percentages among respondents (i.e., estimates) and subgroup percentages in the target population are indicative of some subgroups being overrepresented or underrepresented in the survey.

For this analysis, your dataset should have one row for every person invited to submit a survey, regardless of whether that person ultimately responded. You will indicate the specific parameters of the *t* test.

Choose the Grouping Variable(s) to Analyze

To begin, select one or more grouping variables from your dataset to use as the comparison variables, as seen in figure 21. The grouping variables can be either categorical (e.g., race/ethnicity) or continuous (e.g., age in years), but they must be available for both respondents and nonrespondents.

Figure 21. Screenshot of grouping variable prompt for comparing subgroup percentages in respondent data to data from respondents and nonrespondents within the SRA App

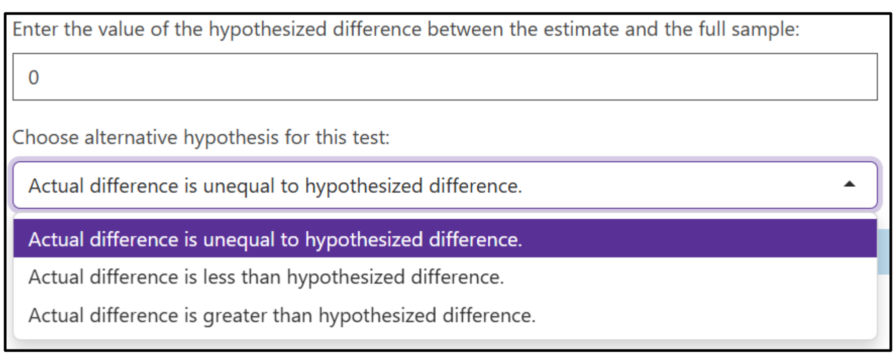


Indicate the Hypothesis for the t Test

First, you will identify the value of the hypothesized difference between the respondent data estimate and the estimate based on data from both respondents and nonrespondents, as seen in figure 22. The test evaluates whether the difference between the two groups is equal to this value. You can choose to leave the application's default setting as is, which tests whether there is any difference between the two groups, using a value of zero, or you can enter a specific value of interest.

Next, you must choose an alternative hypothesis for the test, as also seen in figure 22. You can choose to leave the application's default settings as is, which tests whether the actual difference between the two groups is unequal to the hypothesized difference. Alternatively, you can test whether the actual difference is less than the hypothesized difference or greater than the hypothesized difference.

Figure 22. Screenshot of *t* test hypothesis prompts for comparing subgroup percentages in respondent data to data from respondents and nonrespondents within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

Once complete, you can add the analysis output to the Report module. Close the table when you are finished. Afterwards, you can repeat the same analysis with other grouping variables, or you can select a different analysis type to run.

Table 6 shows an example of the resulting output table from this analysis type after you have added it to the Report module. For each grouping variable you have included (refer to first column, Auxiliary variable), the application produces the subgroup mean or percent among respondents (refer to third column, Mean/percent among respondents), subgroup mean or percent among all eligible cases (refer to fourth column, Mean/percent among all eligible cases), the resulting difference between those percentages (refer to fifth column, Difference), and the *p*-value resulting from the *t* test (refer to seventh column, *p*-value). The output table will also include supplemental information about the specific statistical test the app used in the analysis.

In table 6, response rates examined by student race show that 7.26 percent of survey respondents were Hispanic/Latino students compared with 14.44 percent of students in the target population. This results in a difference of 7.18 percentage points. The results of the *t* test are statistically significant ($p < 0.05$), meaning that the observed differences between respondents and the target population are indicative of this subgroup (i.e., students who are Hispanic/Latino) being underrepresented in the survey data.

When subgroups, such as students in a specific race/ethnicity category, are systematically underrepresented in your survey data, there is also a potential for nonresponse bias with respect to that variable. To see if nonresponse bias exists with respect to this variable, you will need to examine survey outcomes across subgroups to determine if the underrepresented subgroup also differs from other subgroups in what the survey is trying to measure.

Table 6. Comparison of respondent data subgroup percentages to subgroup percentages in all eligible cases, by student race

Auxiliary variable	Category of auxiliary variable	Mean/percent among respondents	Mean/percent among all eligible cases	Difference	Standard error of the difference	*p*-value	Test statistic	Degrees of freedom	Standard error of the mean/percent among respondents	Standard error of the mean/percent among all eligible	Covariance between the two means/ percents
Student race	AM	**0.92%**	**0.85%**	**0.07%**	0.08%	**0.39976**	0.8421	7,055	0.15%	0.11%	0.00%
Student race	AS	**0.92%**	**0.88%**	**0.04%**	0.09%	**0.63981**	0.468	7,055	0.15%	0.11%	0.00%
Student race	BL	**14.85%**	**13.54%**	**1.32%**	0.31%	**0.00002**	4.265	7,055	0.55%	0.42%	0.00%
Student race	HI	**7.26%**	**14.44%**	**-7.18%**	0.38%	**0.00000**	-19.0714	7,055	0.40%	0.43%	0.00%
Student race	MU	**2.91%**	**2.47%**	**0.45%**	0.13%	**0.00090**	3.3217	7,055	0.26%	0.19%	0.00%
Student race	PI	**0.66%**	**0.48%**	**0.18%**	0.05%	**0.00081**	3.3508	7,055	0.12%	0.08%	0.00%
Student race	WH	**72.48%**	**67.36%**	**5.12%**	0.45%	**0.00000**	11.3801	7,055	0.68%	0.57%	0.00%

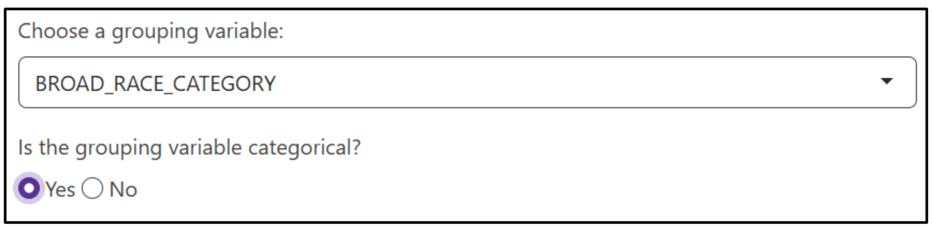
Compare Subgroup Percentages in Respondent Data to External Data

If you do not have subgroup percentages for nonrespondents in your dataset, you can compare respondent subgroup percentages to data from a reliable external source (often referred to as benchmark data). The SRA App can use a chi-squared goodness-of-fit test or a *t* test to assess differences between respondent data and data from an external benchmark. Differences between survey data and the external data source may indicate that some subgroups in the target population are overrepresented or underrepresented in the survey.

Choose and Describe the Grouping Variable to Analyze

To begin, select a single grouping variable from your dataset to use as the comparison variable. Next, indicate if the variable is categorical or is not categorical (i.e., is continuous), as seen in figure 23.

Figure 23. Screenshot of grouping variable prompts for comparing subgroup percentages in respondent data to external data within the SRA App



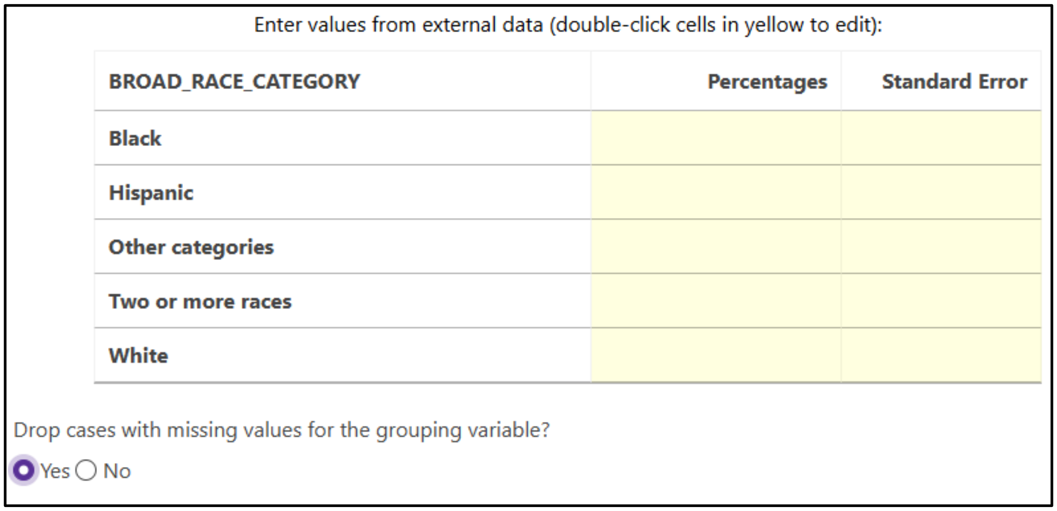
Categorical Grouping Variable: Specify the Analysis

If the grouping variable is categorical, the application will produce a pop-up table that is prefilled with the categories of the variable you selected. You will need to enter the corresponding values (as percentages) of the grouping variable from the external data into the table. If the provided percentages for benchmark values do not sum to 100, the app will rescale them to sum to 100.

You also have the option to enter the standard errors of the percentages, as seen in figure 24. IDC recommends you enter these values if the external data are survey estimates and the standard errors are large (e.g., if an estimated percentage is 10% and the standard error is 5%). You also can leave the standard errors blank; in that case, the application will treat them as zeros when calculating the test statistics.

Next, as also shown in figure 24, indicate whether the application should exclude cases with missing values for the grouping variable. If the grouping variable has missing values in the data, you can only make an estimate by removing rows of data with missing values.

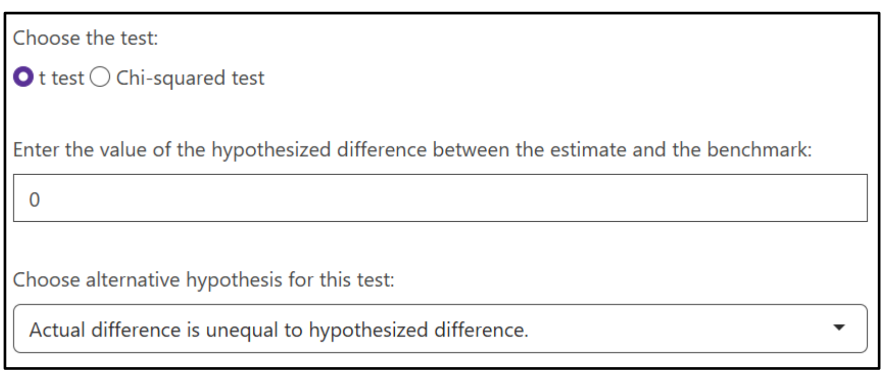
Figure 24. Screenshot of categorical grouping variable prompts for comparing subgroup percentages in respondent data to external data within the SRA App



Then, you will choose whether to compare the subgroup percentages using *t* tests or a chi-squared test. If you choose *t* tests, the app will provide tests of differences for specific categories. If you choose a chi-squared test, it will provide a single overall test of whether any of the survey estimates significantly differ from benchmarks.

To conduct the *t* test, you will first enter the value of the hypothesized difference between the respondent data estimate and the estimate based on the benchmark, as shown in figure 25. The test evaluates whether the difference between the two groups is equal to this value. You can choose to leave the default settings as is, which tests whether there is any difference between the two groups, using a value of zero. Next, you will choose an alternative hypothesis for the test. You can choose to leave the default setting, which tests that the actual difference between the two groups is unequal to the hypothesized difference, or you can test whether the difference is less than the hypothesized difference or greater than the hypothesized difference.

Figure 25. Screenshot of categorical variable *t* test hypothesis prompts for comparing subgroup percentages in respondent data to external data within the SRA App



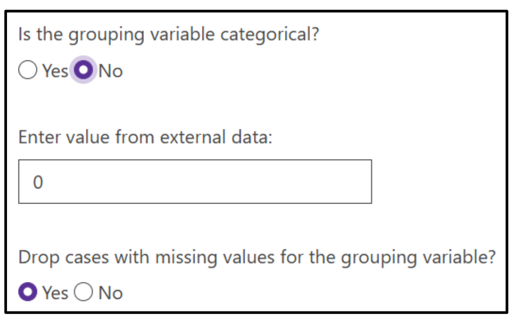
For categorical comparison variables only, you can conduct a chi-squared goodness-of-fit test to compare the distribution of the survey estimates to the distribution of the external estimates. Simply select chi-squared test rather than *t* test in the first prompt shown in figure 25.

When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis. You can add the analysis output to the Report module, then close the table when finished. You can repeat the same analysis with other grouping variables, or you can select a different analysis type to run.

Continuous Grouping Variable: Specify the Analysis

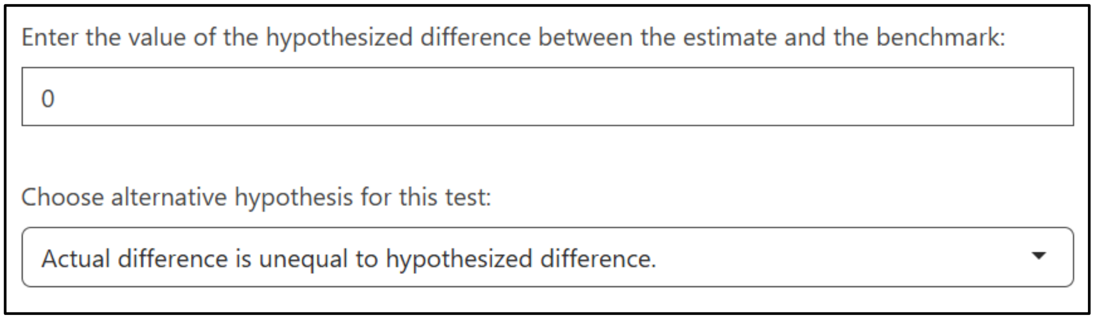
If the grouping variable that you want to use as the comparison variable is continuous (e.g., age in years), you will indicate this by answering no to the question, “Is the grouping variable categorical?” as shown in figure 26. The application will now automatically conduct a *t* test to assess the difference between the survey estimate mean (calculated using sampling weights, if you provided them) and the external benchmark mean. Enter the mean from the corresponding external variable (by replacing the default zero) and indicate if the application should exclude cases with missing values for the comparison variable, also seen in figure 26. If the grouping variable has missing values in the data, you can only make an estimate by removing rows of data with missing values.

Figure 26. Screenshot of continuous grouping variable prompts for comparing subgroup percentages in respondent data to external data within the SRA App



You also will need to enter the value of the hypothesized difference between the respondent data estimate and the estimate based on the benchmark, as seen in figure 27. This test evaluates whether the difference between the two groups is equal to this value. You can choose to leave the default settings as is, which tests whether there is any difference between the two groups, using a value of zero. Next, choose an alternative hypothesis for the test, as also seen in figure 27. You can choose to leave the default setting, which tests that the actual difference between the two groups is unequal to the hypothesized difference, or you can test whether the difference is less than the hypothesized difference or greater than the hypothesized difference.

Figure 27. Screenshot of continuous variable *t* test hypothesis prompts for comparing subgroup percentages in respondent data to external data within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

You can add the analysis output to the Report module and then close the table when finished. Next, you can repeat the same analysis with other grouping variables, or you can select a different analysis type to run.

Table 7 shows an example of the resulting output table for a comparison of respondent subgroup percentages to external data using a *t* test, after you have added the output to the Report module. For each value of the selected grouping variable (refer to first column, Category), the application produces the subgroup mean or percent among respondents (refer to second column, Estimate from respondents), subgroup mean or percent based on the benchmark (refer to third column, External benchmark estimate), the resulting difference between those percentages (refer to fourth column, Difference), and the resulting *p*-value from the selected statistical test (refer to sixth column, *p*-value). The output table also will include supplemental information about the specific statistical test used in the analysis.

Table 7. Comparison of respondent data subgroup percentages to subgroup percentages in external data, by exit reason, using *t* test

Category	Estimate from respondents	External benchmark estimate	Difference	Standard error of the difference	*p*-value	Test statistic	Degrees of freedom
Dropped out	**13.11%**	**17.20%**	**-4.09%**	1.71%	**0.01748**	-2.3866	387
Graduated with a regular high school diploma	**79.69%**	**76.40%**	**3.29%**	2.04%	**0.10786**	1.6116	387
Received a certificate	**7.20%**	**6.40%**	**0.80%**	1.31%	**0.54345**	0.6081	387

In table 7, the user has compared respondent subgroup percentages for the categorical variable exit reason using *t* tests. For example, results show that students who dropped out were 13.11 percent of survey respondents, compared with 17.20 percent of the target population, resulting in a difference of 4.09 percentage points. Results of the *t* test are statistically significant ($p < 0.05$), indicating that the observed differences between respondents and the benchmark data are indicative of this subgroup being systematically underrepresented in the survey data.

When subgroups, such as students who dropped out, are systematically underrepresented in your survey data, there is also a potential for nonresponse bias with respect to that variable. To see if nonresponse bias exists with respect to this variable, you will need to examine survey outcomes across subgroups to determine if the underrepresented subgroup also differs from other subgroups in what the survey is trying to measure.

Keep in Mind

Differences between the survey estimates and the external benchmark data may be attributable to a variety of factors besides nonresponse, such as measurement error or differences between when the external data were collected and when the survey was conducted. Therefore, you should interpret results from this analysis as an indication of the *potential* for nonresponse bias, rather than as an indication of the presence of nonresponse bias.

You can also choose to conduct a chi-squared goodness-of-fit test for a categorical comparison variable to obtain a single overall test of whether any of the subgroup estimates significantly differ from subgroup benchmarks. Table 8 shows an example of the resulting output table for this analysis type, after you have added it to the Report module. The app produces the resulting *p*-value from the chi-squared test (refer to fourth column, *p*-value), as well as supplemental information about the specific statistical test used in the analysis.

Table 8. Comparison of respondent data subgroup percentages to subgroup percentages in external data, by race/ethnicity, using chi-squared test

Test statistic	Degrees of freedom	Scale parameter	*p*-value	Name of test	Method of variance estimation
4.7211	1.951	0.9879	**0.0874**	Rao-Scott chi-square goodness-of-fit test	linearization

As shown in table 8, the resulting *p*-value is not statistically significant ($p > 0.05$), indicating that the percentages of subgroup categories among respondents do not systematically differ from the external benchmark.

How Do Survey Outcomes Differ Across Subgroups?

For Indicators 8 and 14, the FFY 2020–2025 SPP/APR requires states to report on the analysis of response rates, including any nonresponse bias the state identified in the data. When response rates are below 100 percent, nonresponse bias will arise when two conditions occur: (1) certain subgroups are less likely to respond to a survey, resulting in their systematic underrepresentation in the survey data, and (2) the underrepresented subgroups differ from other subgroups in what the survey is trying to measure (e.g., parent involvement, post-school outcomes). The potential impact of the nonresponse bias depends upon the data and its content. For example,

- Does the underrepresented subgroup report worse outcomes? If so, the overall outcome percentage the state is reporting could be overestimated.
- How large was the underrepresented subgroup relative to other subgroups? If the underrepresented subgroup was a relatively large proportion of the target population, nonresponse bias would have a larger effect than if the subgroup was a relatively small proportion of the target population.
- If the underrepresented subgroup was perfectly represented in the survey data, how many fewer students or parents of students would report positive outcomes?

When an analysis of the response data indicates that subgroups are not representative, you can examine the data to determine whether those subgroups differ in terms of their values on the outcome of interest. The analyses in this section address FFY 2020–2025 SPP/APR requirements related to identifying potential nonresponse bias in respondent data.

Compare Outcomes Across Subgroups

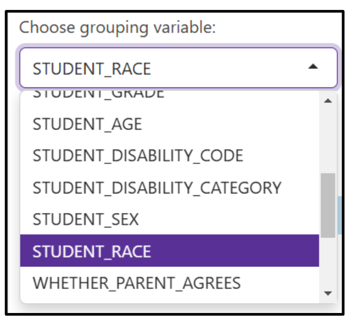
This analysis compares survey outcomes across subgroups by calculating percentages for each category of an outcome variable separately for each subgroup. The application uses a chi-squared test to assess whether observed differences among subgroups in outcome percentages may simply be due to randomness rather than actual population differences.

You can conduct this analysis even if your dataset does not include data from nonrespondents because it examines data only from respondents. However, your dataset must have a variable that captures the outcome(s) measured in the survey. For example, an Indicator 8 dataset may have a parent involvement variable with a value of agree or disagree for each respondent. An Indicator 14 dataset may have a post-school outcome variable that, for a given respondent, has a value of higher education, competitively employed, other school/work, or not engaged.

Choose the Grouping Variable to Analyze

To begin, select a variable to use for grouping summaries of the outcome variable, as seen in figure 28. For example, if analyses of response rates indicated that your respondent data were not representative with respect to student race, you can select the student race variable for further analysis.

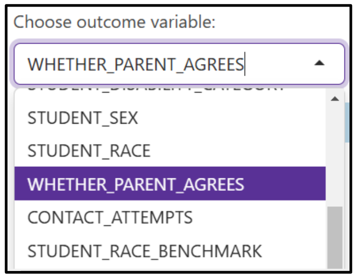
Figure 28. Screenshot of grouping variable prompt for comparing outcomes across subgroups within the SRA App



Choose the Survey Outcome Variable

Next, you must choose the outcome variable for the comparison, as shown in figure 29.

Figure 29. Screenshot of outcome variable prompt for comparing outcomes across subgroups within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

You can add the analysis output to the Report module and then close the table when finished. Next, you can repeat the same analysis with a different grouping variable, or you can select a different analysis type to run.

Table 9 shows an example of the resulting output table for this analysis type after you have added it to the Report module. For each value of the selected grouping variable (refer to first column, Student race), the application produces the percentages (refer to third column, Percent) for each category of an outcome variable (refer to second column, Whether parent agrees), as well as a summary statement of the resulting *p*-value at the bottom of the output table (see table note). The output table will also include supplemental information about the specific statistical test the application used in the analysis

In table 9, agreement among responding parents, examined by student race, is shown to differ among subgroups. A statistically significant result ($p < 0.001$) indicates that observed differences in outcome percentages within the student race variable are likely not simply due to chance, and that one or more subgroups showed significantly different rates of agreement from the rest of the subgroups. Note that the chi-squared test of independence only assesses associations between categorical variables; it cannot provide any inferences about causation. Further investigation of percentages across subgroups shows ranges of parent agreement from 25.24 percent (students who are Hispanic/Latino) to 87.18 percent (students who are Asian). The 95 percent confidence interval for agreement among parents of Hispanic/Latino students ranges between 20.71 percent and 30.39 percent, and the upper bound of this confidence interval is lower than the lower bounds of the other subgroups' confidence intervals. Thus, there is strong evidence that parents of students who are Hispanic/Latino were less likely to agree than parents of other student subgroups. If parents of students who are Hispanic/Latino were also shown to be underrepresented in the respondent data, it would indicate nonresponse bias in the survey data with respect to student race. In this instance, the overall outcome percentage the state is reporting could be overestimated.

Table 9. Comparison of parental agreement across subgroups, by student race

Student race	Whether parent agrees	Percent	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval	Weighted count	Unweighted count
AM	AGREE	**53.85%**	**38.33%**	**68.65%**	21	21
AS	AGREE	**87.18%**	**72.67%**	**94.56%**	34	34
BL	AGREE	**45.89%**	**42.03%**	**49.79%**	290	290
HI	AGREE	**25.24%**	**20.71%**	**30.39%**	78	78
MU	AGREE	**45.16%**	**36.63%**	**53.98%**	56	56
PI	AGREE	**71.43%**	**52.40%**	**85.02%**	20	20
WH	AGREE	**56.84%**	**55.09%**	**58.58%**	1,753	1,753
AM	DISAGREE	**46.15%**	**31.35%**	**61.67%**	18	18
AS	DISAGREE	**12.82%**	**5.44%**	**27.33%**	5	5
BL	DISAGREE	**54.11%**	**50.21%**	**57.97%**	342	342
HI	DISAGREE	**74.76%**	**69.61%**	**79.29%**	231	231
MU	DISAGREE	**54.84%**	**46.02%**	**63.37%**	68	68
PI	DISAGREE	**28.57%**	**14.98%**	**47.60%**	8	8
WH	DISAGREE	**43.16%**	**41.42%**	**44.91%**	1,331	1,331

NOTE: The test of whether the survey outcome (i.e., whether parent agrees) differs among subgroups defined by student race has a *p*-value of < 0.001, based on a chi-squared test of independence.

Identify Variables That Are Predictive of Survey Outcomes

Certain variables may be highly correlated with the likelihood of responding to a survey and therefore can serve as good predictors when modeling survey response. When you find differences in survey outcomes across subgroups within a variable, you also can test whether that variable predicts—or is related to—responses on what your survey is trying to measure. This analysis uses logistic regression to help identify grouping variables in your dataset that are predictive of survey outcomes.

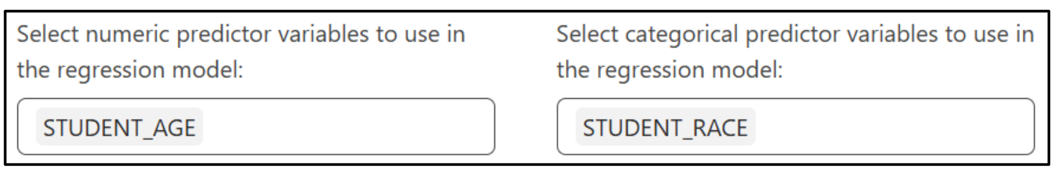
In contrast to the bivariate analysis based on comparing survey outcomes across subgroups, the regression analysis tests multiple variables simultaneously, so you can assess whether a variable's association with survey outcomes is independent of other variables in the model. For example, you can assess whether student race/ethnicity is predictive of post-school outcomes, independently from student exit reason. If a grouping variable effectively predicts response to the survey, and it also effectively predicts survey outcomes, there is risk of nonresponse bias in your respondent data with respect to that variable.

You can conduct regression analyses even if your dataset does not include data from nonrespondents. However, your dataset must have a variable that captures the outcome measured in the survey. That variable may be a numeric variable or a binary categorical variable (i.e., a variable that has only two values). For example, an Indicator 8 dataset may have a parent involvement variable with a value of either agree or disagree for each respondent. An Indicator 14 dataset may have a post-school outcome variable that has a value of either engaged (which combines higher education, competitively employed, other school/work) or not engaged for each respondent.

Select the Grouping Variable(s) for the Analysis

To begin, choose one or more grouping variable(s) to use as predictors in the regression model, as seen in figure 30. You can select multiple predictor variables, and the variables can be either numeric or categorical, but you must specify the numeric and categorical variables separately.

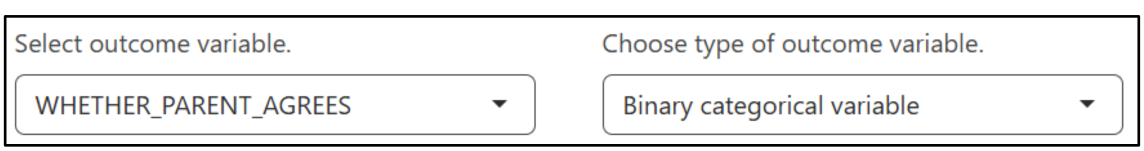
Figure 30. Screenshot of grouping variable prompts for identifying variables that predict survey outcomes within the SRA App



Select the Outcome Variable

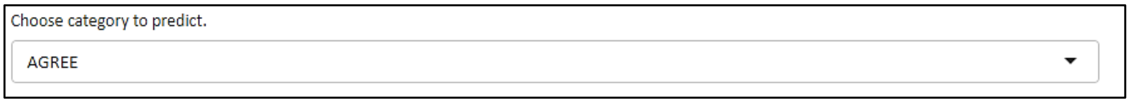
Next, choose the outcome variable whose value the application will predict using the regression model and indicate if that outcome variable is a numeric variable or a binary categorical variable (i.e., a variable that has only two values), as seen in figure 31. The app will use a linear regression for a numeric outcome variable and logistic regression for a binary categorical outcome.

Figure 31. Screenshot of outcome variable prompts for identifying variables that predict survey outcomes within the SRA App



If you choose a binary outcome variable, you must also choose which value of the outcome variable to predict, as seen in figure 32.

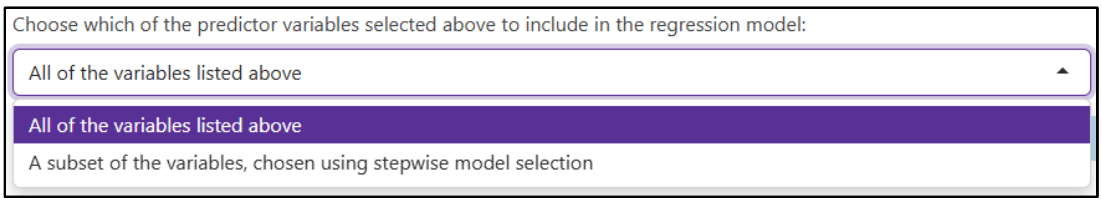
Figure 32. Screenshot of binary outcome variable prompt for identifying variables that predict survey outcomes within the SRA App



Indicate How the Application Should Run the Regression Model

Finally, as seen in figure 33, you can include either all the variables that you selected to use in the regression model, or you can potentially reduce this list by using a stepwise model selection procedure to select a subset of variables that can significantly predict the outcome.

Figure 33. Screenshot of regression model prompt for identifying variables that predict survey outcomes within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

You can add the analysis output to the Report module and then close the table when finished. You can then repeat the same analysis with different variables and parameters, or you can select a different analysis type to run.

Table 10 shows an example of the resulting output table for this analysis type after you have added it to the Report module. For each grouping variable you have included as predictors in the regression model (refer to first column, Predictor variable), the application will provide the resulting variable-level *p*-values (refer to second column, Variable-level *p*-value), as well as the estimated coefficient (refer to fourth column, Estimated coefficient) and coefficient *p*-values (refer to ninth column, Coefficient *p*-value from *t* test). For categorical predictor variables, the output table includes a note at the bottom identifying the specific categories that the app used as reference levels for the regression. (Note: you can control which category the app uses as the reference variable by alphabetizing the categories.) The output table will also include supplemental information about the specific statistical test the app used in the analysis.

Table 10. Prediction of parental agreement, by student age and student race

Predictor variable	Variable-level *p*-value	Category of categorical variable	Estimated coefficient	Standard error	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval	Coefficient *p*-value from *t* test	Likelihood ratio test chi-squared statistic	Likelihood ratio test DEff	Likelihood ratio test numerator df	Likelihood ratio test denominator df
(Intercept)			**-0.7978**	0.3338	-1.4523	-0.1433	**0.0169**				
Student age	**0.67691**		**0.0562**	0.0084	0.0397	0.0726	**0**	46.7366	1.0327	1	4,247
Student race	**0**	AS	**-1.6896**	0.5825	-2.8316	-0.5476	**0.0037**	160.8928	1.0038	6	4,247
Student race	**0**	BL	**0.3201**	0.3293	-0.3255	0.9658	**0.3311**	160.8928	1.0038	6	4,247
Student race	**0**	HI	**1.3103**	0.3465	0.6311	1.9896	**0.0002**	160.8928	1.0038	6	4,247
Student race	**0**	MU	**0.47**	0.3679	-0.2512	1.1913	**0.2014**	160.8928	1.0038	6	4,247
Student race	**0**	PI	**-0.9466**	0.5275	-1.9807	0.0876	**0.0728**	160.8928	1.0038	6	4,247
Student race	**0**	WH	**-0.0782**	0.3216	-0.7088	0.5524	**0.8079**	160.8928	1.0038	6	4,247

NOTE: For categorical predictor variables, the application used the following categories as reference levels for the regression—Student race: AM (American Indian or Alaska Native).

An association between a grouping variable and survey outcomes is one factor in the potential for nonresponse bias with respect to that variable. A statistically significant variable-level *p*-value indicates the specific grouping variable effectively predicts survey outcomes, independently from other grouping variables included in the model. For example, if the *p*-value for student race shows a statistical significance ($p < 0.05$), as indicated in table 10, that means that the likelihood of agreeing/disagreeing to the survey varies for parents of students of the same ages in different race categories.

The regression coefficients for specific categories indicate which specific subgroups differ from the reference subgroup; a statistically significant coefficient indicates that there is a statistically significant difference in outcomes between the specific subgroup and the reference subgroup. For example, in table 10 the *p*-value for the HI (Hispanic/Latino) coefficient shows a statistically significant difference ($p < 0.05$) compared to the reference category.

If the results indicate that a grouping variable predicts survey outcomes and also effectively predicts response to the survey (e.g., student race is also a significant predictor of whether an individual is a respondent rather than a nonrespondent in this dataset), then there is risk of nonresponse bias in your data with respect to that variable. For this example, the student race variable was found to be predictive of whether an individual responded to the survey (see table 5). Specifically, parents of students who are Hispanic/Latino were found to be underrepresented among respondents. Table 10 shows that student race/ethnicity is also a significant predictor of survey outcomes ($p < 0.05$). Therefore, nonresponse bias exists in the survey data with respect to student race. When you find variables in your dataset that are predictive of key survey outcomes, as in this example, you can consider using those variables in weighting adjustments for the purpose of reducing nonresponse bias (Kreuter et al. 2010). Note that while the intercept is included in the output table, its results are not used in interpretation of the analysis.

Assess How Outcomes Change as Level-of-Effort Increases

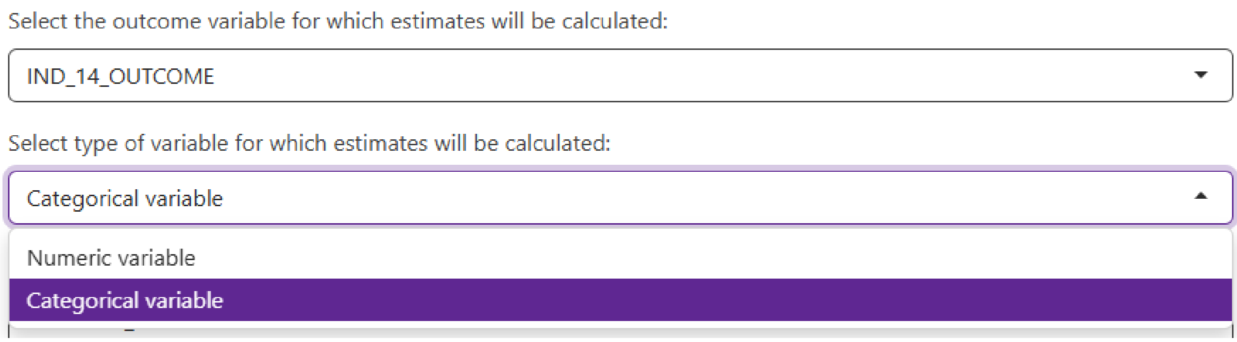
This analysis evaluates how survey outcomes change as you expend additional effort to obtain survey responses, such as by making additional contact attempts or increasing incentives. It calculates the cumulative mean for continuous variables and cumulative proportions for categorical variables. This analysis assumes that respondents who are harder to reach (i.e., require more contact attempts) are more similar to nonrespondents than to respondents who are easier to reach. If the hard-to-reach respondents and easy-to-reach respondents differ with respect to their outcome variable estimates, this could suggest the potential for nonresponse bias within the data (Lin and Schaeffer 1995).

You can conduct this analysis even if your dataset does not include data from nonrespondents. However, your dataset must include one or more variables that reflect level of effort.

Select and Describe the Outcome Variable for the Analysis

To begin, as seen in figure 34, select the outcome variable for which you would like to calculate estimates and identify whether this outcome variable is categorical or numeric. For categorical variables, the application estimates proportions for each category. For numeric variables, the application estimates means.

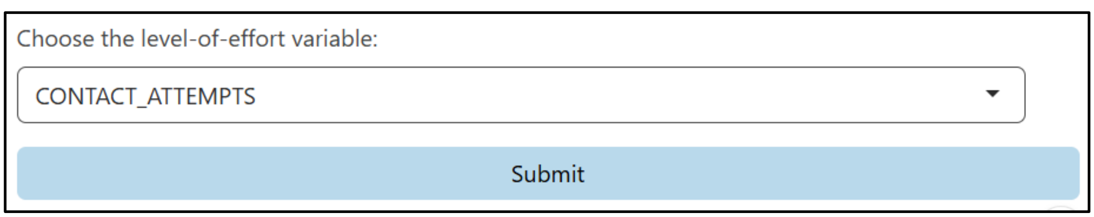
Figure 34. Screenshot of level-of-effort outcome variable prompts within the SRA App



Select the Level-of-Effort Variable

Next, choose the level-of-effort variable for which the application will calculate cumulative estimates, as seen in figure 35. The variable should either be numeric or have categories with labels that, when sorted alphabetically, yield the order you desire.

Figure 35. Screenshot of level-of-effort variable prompt within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

You can add the analysis output to the Report module, then close the table when finished. You can also repeat the same analysis with different variables, or you can select a different analysis type to run.

Table 11 shows an example of the resulting output table for this analysis type after you have added it to the Report module. For each value of the selected level-of-effort variable (refer to first column, Contact attempts), the application provides the cumulative mean or proportion of respondents (refer to fourth column, Estimate from respondents) for each value of the selected outcome variable (refer to second and third columns, Analysis variable and Category of analysis variable, respectively). The output table will also include supplemental information about the specific analysis. If the hard-to-reach respondents and easy-to-reach respondents differ with respect to their outcome variable estimates, this could potentially suggest nonresponse bias within the data.

Table 11. Change in post-school outcomes, by number of contact attempts

Contact attempts	Analysis variable	Category of analysis variable	Estimate from respondents	Standard error of the estimate from respondents	Number of eligible respondents
‘1’	Ind. 14 outcome	Competitive Employment	**54.0%**	3.9%	161
‘1’	Ind. 14 outcome	Higher Education	**11.8%**	2.5%	161
‘1’	Ind. 14 outcome	Not Engaged or Other	**11.2%**	2.5%	161
‘1’	Ind. 14 outcome	Other Employment	**9.9%**	2.4%	161
‘1’	Ind. 14 outcome	Postsecondary Education or Training	**13.0%**	2.7%	161
‘1’ to ‘2’	Ind. 14 outcome	Competitive Employment	**55.9%**	3.3%	229
‘1’ to ‘2’	Ind. 14 outcome	Higher Education	**11.4%**	2.1%	229
‘1’ to ‘2’	Ind. 14 outcome	Not Engaged or Other	**11.4%**	2.1%	229
‘1’ to ‘2’	Ind. 14 outcome	Other Employment	**9.6%**	1.9%	229
‘1’ to ‘2’	Ind. 14 outcome	Postsecondary Education or Training	**11.8%**	2.1%	229
‘1’ to ‘3’	Ind. 14 outcome	Competitive Employment	**55.8%**	3.1%	260
‘1’ to ‘3’	Ind. 14 outcome	Higher Education	**13.1%**	2.1%	260

Table 11. Change in post-school outcomes, by number of contact attempts—Continued

Contact attempts	Analysis variable	Category of analysis variable	Estimate from respondents	Standard error of the estimate from respondents	Number of eligible respondents
‘1’ to ‘3’	Ind. 14 outcome	Not Engaged or Other	**10.0%**	1.9%	260
‘1’ to ‘3’	Ind. 14 outcome	Other Employment	**9.6%**	1.8%	260
‘1’ to ‘3’	Ind. 14 outcome	Postsecondary Education or Training	**11.5%**	2.0%	260
‘1’ to ‘4’	Ind. 14 outcome	Competitive Employment	**55.4%**	2.9%	303
‘1’ to ‘4’	Ind. 14 outcome	Higher Education	**13.9%**	2.0%	303
‘1’ to ‘4’	Ind. 14 outcome	Not Engaged or Other	**10.6%**	1.8%	303
‘1’ to ‘4’	Ind. 14 outcome	Other Employment	**8.9%**	1.6%	303
‘1’ to ‘4’	Ind. 14 outcome	Postsecondary Education or Training	**11.2%**	1.8%	303
‘1’ to ‘5’	Ind. 14 outcome	Competitive Employment	**55.2%**	2.7%	344
‘1’ to ‘5’	Ind. 14 outcome	Higher Education	**13.7%**	1.9%	344
‘1’ to ‘5’	Ind. 14 outcome	Not Engaged or Other	**10.8%**	1.7%	344
‘1’ to ‘5’	Ind. 14 outcome	Other Employment	**8.7%**	1.5%	344
‘1’ to ‘5’	Ind. 14 outcome	Postsecondary Education or Training	**11.6%**	1.7%	344

Table 11 additionally shows how the estimated distribution of post-school outcomes changes when analyzing data from respondents with varying numbers of contact attempts. The first five rows in table 11 show the estimated distribution of post-school outcomes based on respondents with only one contact attempt, the next five rows show the same estimates based on respondents with one to two contact attempts, and so on. If the estimates change substantially as the number of contact attempts increases, this may indicate that harder-to-reach respondents differ from easier-to-reach respondents on the outcome of interest. In table 11, the estimates do not change appreciably when analyzing data from respondents with only a few contact attempts rather than analyzing data from respondents with potentially many contact attempts. Therefore, the level-of-effort analysis does not indicate that increasing contact attempts have an effect on potential nonresponse bias.

Can Statistical Adjustments Reduce Nonresponse Bias in Our Data?

One way of assessing the magnitude of nonresponse bias in your data and the likely effectiveness of statistical adjustments in reducing that bias is to compare survey outcomes you have computed using adjusted weights to those you have computed using unadjusted weights (Krenzke, Van de Kerckhove, and Mohadjer 2005). Weighting is a statistical technique in which you adjust survey data after collecting it to improve the accuracy of the survey estimates. Weighting adjustments can be useful for reducing nonresponse bias when you include variables that are highly predictive of survey response in the weighting adjustment (Brick and Jones 2008). Weighting rebalances the data to reflect the target populations better by counting data from some subgroups more or less than data from other subgroups, which compensates for a lack of representativeness in the original data.

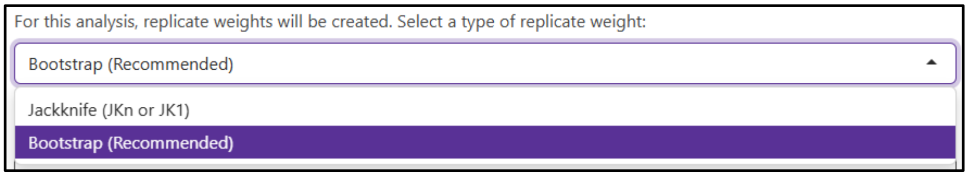
Compare Survey Estimates Based on Respondent Data, Before and After Weighting Adjustments

This analysis uses the weighting technique of raking to increase the influence of respondents from underrepresented subgroups so that these subgroups' influence on estimates is proportional to their share of the target population. The SRA App calibrates initial weights to known totals of one or more auxiliary variables, referred to as grouping variables in the app, and then uses *t* tests to compare estimates of the selected outcome variables before and after weighting adjustments.

Specify the Method for Creating Weights

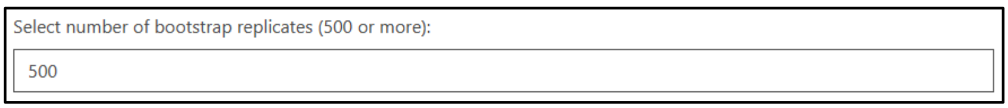
For this analysis, the application will create replicate weights. Therefore, first you need to select either the bootstrap method or jackknife method for the creation of these weights, as seen in figure 36. Both methods are reasonable options. In the app, the bootstrap method is set as default because, while it is less statistically efficient than the jackknife method, it generally takes much less time to run if there are more than 500 sampling units in the dataset.

Figure 36. Screenshot of replicate weight method prompt within the SRA App



If you select the bootstrap method of creating replicate weights, then you also must choose the number of bootstrap replicates. As seen in figure 37, the application's default value is 500, which gives statistically reasonable estimates in a wide variety of settings without being too demanding on the computer. You can choose to increase this number (to 1,000 or 2,000, for example); the calculation will simply take longer to run. Also note that if you have a very large dataset, such as an attempted census of 10,000 cases, you can expect the application to take several minutes to conduct the analysis.

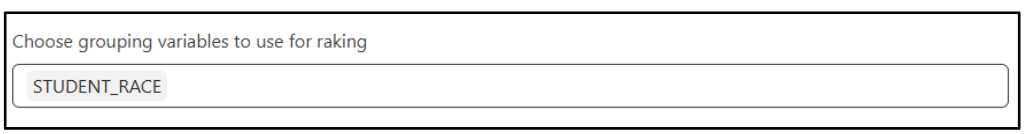
Figure 37. Screenshot of bootstrap replicates prompts within the SRA App



Select the Variables to Use in the Analysis

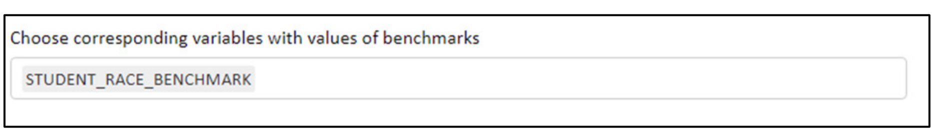
Next, as seen in figure 38, select one or more grouping variables in your dataset which divide the sample into different subpopulations (e.g., race/ethnicity). These variables cannot have any missing values.

Figure 38. Screenshot of grouping variable prompt for weighting adjustments within the SRA App



Then, if you collected your survey data through sampling, you need to select the corresponding variable in your dataset that gives the population benchmark (i.e., population size) for each category of that grouping variable, as seen in figure 39.

Figure 39. Screenshot of benchmark variable prompt for weighting adjustments within the SRA App



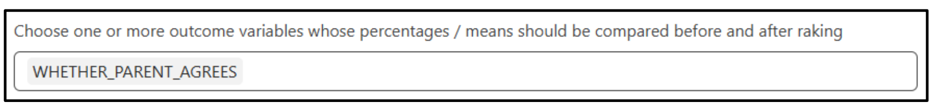
Each respondent and nonrespondent within the same subgroup will have the same benchmark value, as shown in table 12.

Table 12. Example benchmark values

Unique ID	Response status	Student race	Student race benchmark
ID_00029	Respondent	BL	2,823
ID_00666	Nonrespondent	BL	2,823
ID_00116	Respondent	WH	12,386
ID_00668	Respondent	WH	12,386
ID_00247	Respondent	AS	206
ID_00655	Unknown	AS	206
ID_00670	Ineligible	PI	140
ID_00162	Respondent	PI	140

Finally, as seen in figure 40, select the outcome variable for the analysis.

Figure 40. Screenshot of outcome variable prompt for weighting adjustments within the SRA App



When you are ready, submit your options. The application will produce a pop-up table with the resulting statistics for the analysis.

Add Your Output to the Report

You can add the analysis output to the Report module, then close the table when finished. You can repeat the same analysis with different parameters, or you can select a different analysis type to run.

Table 13 shows an example of the resulting output table for this analysis type after you have added it to the Report module. This analysis uses raking weighting adjustments to increase the influence of respondents from underrepresented subgroups so that these subgroups' influence on estimates is proportional to their share of the target population. The application calibrates initial weights to known totals of the selected grouping variables and then uses *t* tests to compare estimates of the selected outcome variables before and after weighting adjustments. For each value of the selected outcomes variable (refer to first and second columns, outcome and outcome category, respectively), the application will provide estimated means of the continuous variables, or the percentage distributions of the categorical variables, before adjustment (refer to third column, Mean/percent before adjustment) and after adjustment (refer to fourth column, Mean/percent after adjustment, as well as fifth column, Differences), as well as the associated *p*-values (refer to seventh column, *p*-value). The output table also will include supplemental information about the specific analysis.

Statistically significant differences between unadjusted estimates (before weighting) and adjusted estimates (after weighting) reflect the potential for nonresponse bias with respect to the specific variable you examined. For example, in table 13, the rate of parental agreement was significantly different with and without weighting of student race ($p < 0.05$). Further, it shows that the agreement rate without the race variable adjustment was 52.93 percent, whereas the agreement rate with the race variable adjustment was 49.44 percent. This means that one or more race/ethnicity subgroups with lower percentage of parental agreement were underrepresented among the survey respondents (i.e., nonresponse bias).

It is important to consider the effect of nonresponse bias on data quality. If a subgroup is underrepresented and reports lower outcomes, the overall percentage reported for that indicator in the SPP/APR might be an overestimation of the true percentage. Conversely, if a subgroup is underrepresented but reports higher outcomes, the overall percentage reported might be an underestimation of the true percentage.

The results from Table 13 indicate that nonresponse bias with respect to race/ethnicity is causing an overestimation in the percent of parental agreement the state would report in the SPP/APR by 3.49 percentage points. Since there is no pre-determined threshold for what an acceptable difference between the unweighted and weighted results is, you can evaluate the effect of nonresponse bias on data quality by considering the following questions:

- Could the identified nonresponse bias affect whether the state met or did not meet the target in its SPP/APR? By how much?
- Could the identified nonresponse bias affect whether the state had slippage? By how much?
- Does the identified nonresponse bias change the direction of the trend over time in meeting the target?
- When comparing the weighted and unweighted outcome percentages, how many students or parents of students are represented by the difference?

You can use weighting adjustments to correct for imbalances between your respondent data and the target population as well as to compensate for subgroups being underrepresented among respondents. If there is not a large change in the estimates, it may indicate that nonresponse bias is not a concern with respect to the specific variable you examined. You may choose to provide both the unweighted and weighted values when describing your results.

Table 13. Comparison of parental agreement before and after weighting, by student race

Outcome	Outcome category	Mean/percent before adjustment	Mean/percent after adjustment	Difference	Standard error of the difference	*p*-value	Test statistic	Degrees of freedom	Standard error of the mean/percent before adjustment	Standard error of the mean/percent after adjustment	Covariance between the two means/percents
Whether parent agrees	AGREE	52.93%	49.44%	3.49%	0.34%	0.00000	10.2377	498	0.76%	0.78%	0.01%
Whether parent agrees	DISAGREE	47.07%	50.56%	-3.49%	0.34%	0.00000	-10.2377	498	0.76%	0.78%	0.01%

NOTE: The application used the following variables weight data from respondents: Student race.

Exiting the SRA App

When you have finished using the SRA App, simply close the browser window with the application to end the session. Recall that the app will not save the information you entered once you close it. Therefore, the next time you launch the app to begin a new session, you will need to complete the Setup tab for that new session, then move to the Analysis tab to select analyses to run, which you can then add to the Report tab for export into Excel.

Glossary

Alternative hypothesis: A statement used in statistical testing that indicates a relationship between variables.

Auxiliary variable: A variable that provides information that is available prior to data collection and which one knows for all units of the population; referred to as a grouping variable in the SRA App.

Benchmark: Data from a reliable external source used for comparison with survey respondent data.

Census: A data collection method where one attempts to collect data from every person in the target population.

Chi-squared test of independence: A statistical hypothesis test used to determine whether there is an association between categorical variables (i.e., whether the variables are independent or related); also known as the chi-squared test of association.

Chi-squared goodness-of-fit test: A statistical hypothesis test used to determine whether the sampling distribution of a statistic matches a specified distribution.

Cluster sampling: A form of sampling from a population where one divides the population into non-overlapping groups, referred to as clusters, and then selects a random sample of the clusters. For example, suppose one divides the population into school districts and randomly selects 10 school districts to participate in the survey. This would be an example of cluster sampling, where the school districts are the clusters.

Confidence interval: An interval of possible values for a population characteristic, based on information from the sample and a specified level of confidence (e.g., 95%) that the interval contains the true population value. For example, when estimating the post-school percent engaged measure, one might have a sample estimate of 55 percent and 95 percent confidence interval of 50–60 percent. The size of the confidence interval reflects how precise the estimate is. For example, if the 95 percent confidence interval for the post-school percent engaged is 25–85 percent, then the sample estimate of 55 percent is imprecise and should be treated with caution.

Degrees of freedom: A technical measure one uses in statistical hypothesis tests, related to both the sample size of the data and the number of groups being compared in a hypothesis test. The SRA App can automatically determine reasonable degrees of freedom to use for hypothesis tests.

Error: The difference between an estimate and the population value of interest. Multiple factors can cause the difference: sampling error, nonresponse bias, etc.

Estimate: A value one calculates using data from a sample instead of using data from the entire population.

Hypothesis test: A statistical method used to determine whether there is enough evidence in a sample of data to infer that a certain condition is true for the entire population. Typically, a hypothesis test is used to determine whether the data provide sufficient evidence to reject a null hypothesis (that there is no difference between two groups or that there is no relationship between two variables) in favor of an alternative hypothesis (that there is a difference between groups or a relationship between variables). See also: statistical significance.

Logistic regression: A type of statistical model that estimates the probability of an event occurring, such as responding or not responding to a survey, based on a given dataset of independent variables. Logistic regression finds the relationships between two data factors, then uses this relationship to predict the value of one of those factors based on the other.

Nonresponse bias: Systematic error that results from nonresponse to a survey. Nonresponse bias arises when two conditions occur: (1) certain subgroups are less likely to respond to a survey, resulting in their systematic underrepresentation in the survey data, and (2) the underrepresented subgroups differ from other subgroups in what the survey is trying to measure.

Null hypothesis: A statement used in statistical testing that indicates there is no difference between two groups or that there is no relationship between two variables.

***p*-value:** A single statistic used to summarize the results of a statistical hypothesis test. A *p*-value ranges from 0 to 1. In a statistical hypothesis test, a result is statistically significant if a *p*-value falls below a certain predetermined threshold (e.g., 0.05). See also: statistical significance.

Representativeness: The degree to which survey respondents proportionally replicate the target population with respect to a specific demographic characteristics, such as age, race/ethnicity, sex, socioeconomic status, etc.

Response rate: The number of responses divided by the number of people asked to respond, usually expressed as a percentage.

Sample: A properly selected subset of the population.

Sampling weight: A value used to correct for sampling procedures that cause individuals to have different probabilities of being randomly sampled for a survey. If individuals have different probabilities of being randomly sampled, then one should generally use a sampling weight. Calculate the sampling weight for a given sampled person as 1 divided by that person's probability of being selected into the sample.

Standard error: A measure of how precise an estimate is. A larger standard error implies a larger margin of error for an estimate. If an estimate has a large standard of error, then the estimate may be very different from the population value.

Statistical bias: Systematic error in a survey estimate that causes it to be too high or too low. Statistical bias is a form of *systematic* error, rather than random error; that is, statistical bias is an error that one would expect to recur if calculating the estimate using a different random sample or if repeating the survey in another year. Systematic error skews the results, making them non-representative of the true population.

Statistical significance: A pattern observed in data is statistically significant if one deems that pattern as unlikely to have arisen in a random sample when the pattern is not actually present in the population from which one draws the sample. Whether one describes a pattern as statistically significant is generally based on the results of a statistical hypothesis test (e.g., a *t* test or chi-squared test). If a *p*-value for the hypothesis test falls below a predetermined threshold (e.g., when a *p*-value is less than 0.05), then the hypothesis test has a statistically significant result.

It is important to note that the statistical significance of a result is completely unrelated to the practical significance of that result. For example, a difference in response rates between two groups (say, 91.3% and 91.5%) might be statistically significant yet still be so small as to be irrelevant for all practical purposes. Describing a result as statistically significant is simply a way to indicate that an observed pattern cannot easily be explained away as a fluke that randomly occurred in the sample of data.

Stratum: See stratified sampling.

Stratified sampling: A method of sampling that involves dividing a population into non-overlapping subgroups, or strata, based on certain characteristics, then randomly sampling from each stratum, with predetermined sample sizes in each stratum. For example, suppose one divides the population into eastern school districts and western school districts and chooses to randomly sample 10 eastern school districts and 20 western school districts. This would be an example of stratified sampling, where the two groups of school districts (eastern and western) are strata.

Subgroup: A subset of participants based on a shared characteristic, such as shared race/ethnicity.

Systematic error: A type of error that one expects to happen again if repeating a survey with a different random sample or in another year.

***t* test:** A statistical test used to compare the means of two groups, sometimes known as a Student's *t*-test.

Weighting adjustment: A procedure that assigns each sampled individual a weight or—if already using sampling weights—modifies a sampling weight. Whereas one uses sampling weights to correct for intentionally sampling individuals with unequal probabilities, one uses weight adjustments to correct for non-sampling aspects of the survey (such as nonresponse) that cause some individuals to have a higher chance of being included in the respondent sample compared to other individuals. Post-stratification and raking are both common weight adjustment methods used to reduce nonresponse bias.

References

The American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (9th Edition). Alexandria, VA: AAPOR.

Brick, J.M., and Jones, M.E. (2008). Propensity to Respond and Nonresponse Bias. *International Journal of Statistics*, *LXVI*(1): 51–73.

Krenzke, T., Van de Kerckhove, W., and Mohadjer, L. (2005). Identifying and Reducing Nonresponse Bias Throughout the Survey Process. *ASA Section on Survey Research Methods*. Alexandria, VA: American Statistical Association. Retrieved January 20, 2023, from http://www.asasrms.org/Proceedings/y2005/files/JSM2005-000572.pdf.

Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., and Raghunathan, T.E. (2010). Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Non-Response: Examples From Multiple Surveys. *Journal of the Royal Statistical Society: Series A: Statistics in Society*, *173*(2): 389–407.

Lin, I.F., and Schaeffer, N.C. (1995). Using Survey Participants to Estimate the Impact of Nonparticipation. *Public Opinion Quarterly*, *59*: 236–258.