

Veracio Tutorial 4:

Understanding the details of Weighting

Veracio is an online tool that helps people create surveys and understand them in an accurate manner by using weights to summarize survey responses. These weights are used to ensure the survey results more closely reflect the underlying population. There are three basic steps to using Veracio:

Step 1: Create a Survey

Step 2: Collect Responses

Step 3: Analyze Results

This tutorial will walk you through the more technical details of Step 3, analyzing your results. We will talk more about how the Veracio tool weights your data and what that means.

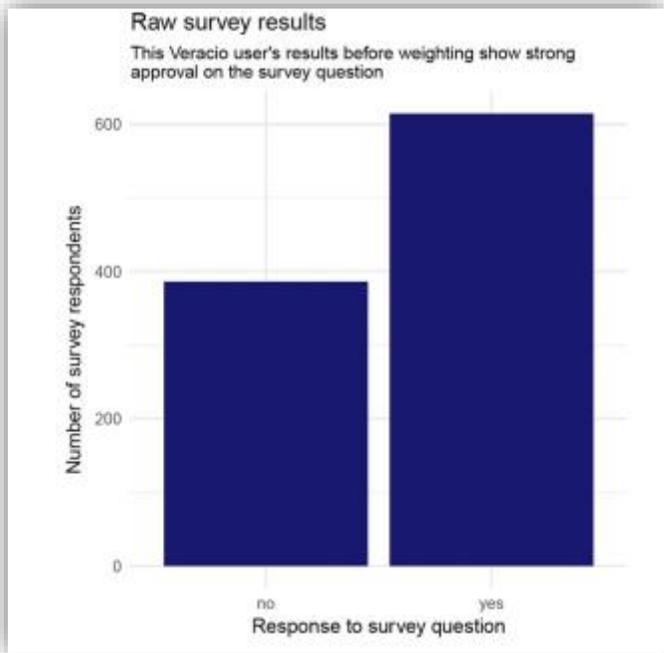
When you set up your survey, you have designed a survey with questions relevant to your organization, the population served, customers, audience, or constituents. You set up this survey using the [Veracio website](#), but the questions that you want answers to are not the only questions on the survey. The survey will also always include demographic questions, about such factors as sex, race/ethnicity, age, and/or education; the Veracio tool uses these demographic questions to statistically weight the survey responses.

When you set up the survey, you must specify what geographic area the survey respondents reside in. This geographic area can be a United States county, state, or the entire U.S. and more than one can be specified. The Veracio tool uses U.S. Census data to implement the statistical weighting, so it is important to accurately identify the appropriate geographic location of your survey respondents. It is not better to err on the side of larger or smaller geographic regions; carefully choose the most appropriate region for your population of interest.

When setting up your survey, you will also have to specify which demographic questions to ask on the survey. The choices include sex, race/ethnicity, age, and education, but not all of these can be used at the same time because of the way Census data is organized into tables. Also, unless your survey is very large, it is unlikely that weighting by many of these indicators will improve the survey results. The Veracio tool is able to test which of the weighting indicators gives the best result, so it does not hurt to err on the side of including more demographic questions. Possible sets of demographic questions to ask include sex, race/ethnicity, and age, and sex, race/ethnicity, and education. You cannot include both age and education questions on the same survey because of the way the U.S. Census tabulates data on these characteristics.

After you have designed the survey, including the appropriate geography/geographies, the demographic questions, and the actual questions of interest, the survey is released and shared with survey respondents. During the survey period, responses are stored. At the end of the

survey period, all the survey responses are ready to be analyzed. For this guide, let's imagine that a Veracio user designs a survey for her state with one question on it that has possible answers of "yes" and "no"; perhaps she works for a social service organization and wants to ask residents in her state, "Do you approve of the way our organization is handling our job?"



Furthermore, let's imagine that the survey respondents are biased compared to the real population in the state; they do not share the same demographic characteristics. Let's say that more women than men took the survey, and more white community members took the survey than people of color.

Choosing the best set of weighting indicators

When Veracio users design a survey like this example survey, they will not know ahead of time which set of demographic factors will be the best set of weighting indicators to use for their individual surveys, because they cannot predict how their survey respondents will be biased. For example, maybe many more men than women will be respondents to your survey. Or maybe many more young people than older adults will answer your survey. We can't know ahead of time. Once the survey results have been collected, the Veracio tool uses bootstrap resampling¹ of the survey itself to find the best set of demographic factors for weighting. After extensive testing, we have found that the best set of weighting indicators to use for the final weights in your survey is the combination of indicators that produce the consistently lowest variance. In other words, when we try all possible combinations of weighting factors, the ones who combine most often to give us the most precise estimate of your survey's results are the combination of weighting factors that is most likely to make your sample represent your community.

For example, if the survey included demographic questions on sex, race/ethnicity, and age, the Veracio tool will calculate weights for bootstrap resamplings of the survey using all possible combinations of these indicators. In this case, those combinations will be as shown in this table.

Possible Weighting Factors
sex
race/ethnicity
age
sex, race/ethnicity
sex, age
race/ethnicity, age
sex, race/ethnicity, age

The Veracio tool takes each one of these options in turn and weights many bootstrap resamplings of the original survey using that set of weighting factors. For example, first we will take only sex and weight 500-1000 bootstrap resamplings of the survey using just sex. We can calculate the mean and variance of the weighted survey results after using only sex as the weighting factor. Next, we will take only race/ethnicity, weight 500-1000 bootstrap resamplings using just race/ethnicity, then calculate the mean and variance of these weighted survey results. We continue through all the possible combinations of weighting factors.

¹ Bootstrap resampling is essentially taking many repeated small samples from a dataset.

At the end of this process, we have a mean and variance for each possible way to weight the survey. We cannot know ahead of time which mean weighting is the correct weighting, but our extensive simulations show that the weighting choice that results in the lowest variance is most likely one that matches the real result in the survey most closely. For this example, that is the combination of sex and race/ethnicity.

Reporting post-stratification weights

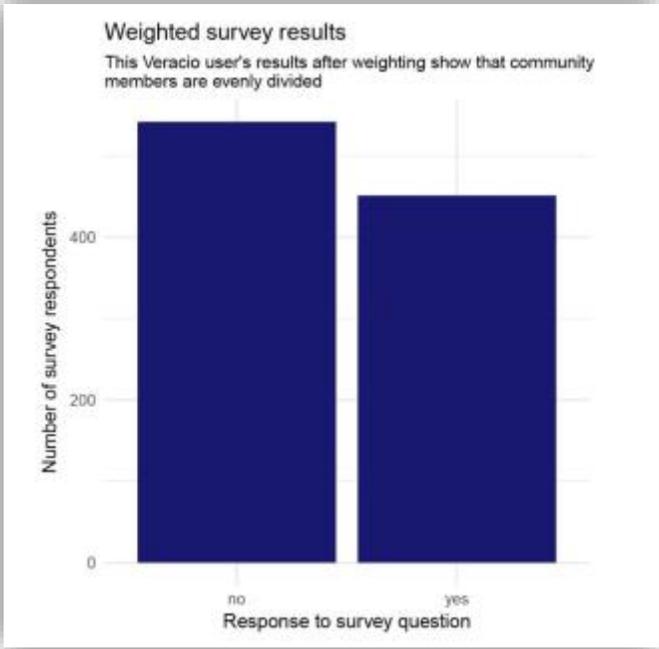
Once the Veracio tool has identified the best set of weighting indicators to apply to a specific survey, we then can weight the real survey one final time using this set and report these weights to the user. The Veracio tool using post-stratification weighting. What this means is that a weight (a number we compare to 1, but possibly bigger or smaller than 1) is calculated for each survey respondent so that the known population characteristics for the demographic factors in the real population can be reproduced in the survey. The Veracio tool uses 5-year American Community Survey population estimates from the United States Census for measurements of real population characteristics. These are called post-stratification weights because they are calculated after the survey is finished. If a survey respondent ends up with a weight of 2, that means their answer should be weighted (multiplied) by 2; similarly, if a survey respondent ends up with a weight of 0.5, their response should be weighted by 0.5.

Here is an example of a survey that has been weighted. Each row represents one respondent. The survey collects the data for the columns for Sex, Ethnicity, Age, and the Survey Question (Agree Disagree in this case). Then, after the data has been collected, Veracio adds a column called "Weight." The first row has a weight of 1.2 this means that this person will represent 1.2 people in the population. The second row has a weight of 0.5. This means that this person will represent half a person in the population. A person with a weight over 1 means that this type of person is under-represented in your survey's sample. A person with a weight under 1 means that this type of person is over-represented in your survey's sample relative to your population of interest.

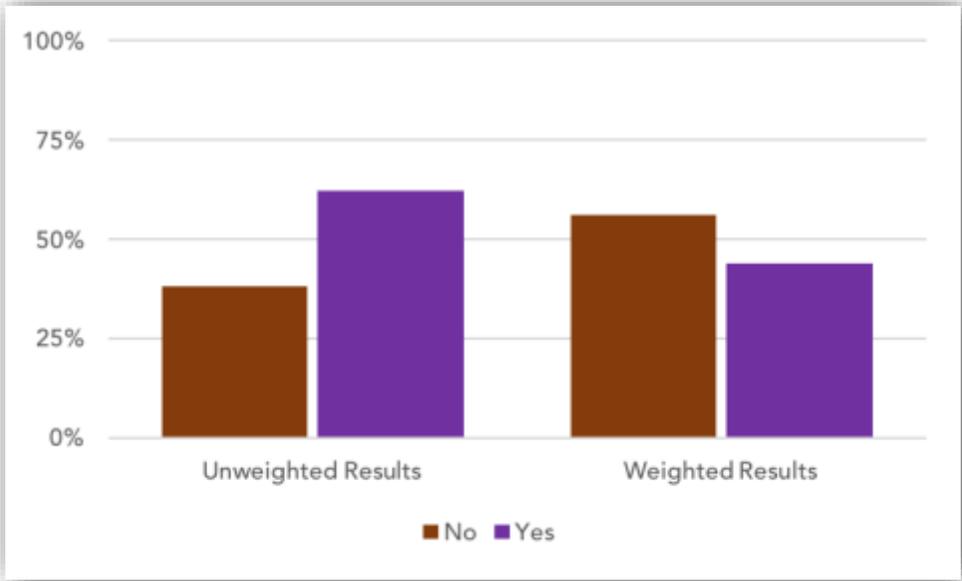
sex	raceethnicity	age	Agree Disagree	Weight
Female	HISPANIC OR LATINO	45 to 64 years	yes	1.2
Female	HISPANIC OR LATINO	65 years and over	yes	0.5
Female	OTHER	65 years and over	no	1.5
Female	OTHER	65 years and over	no	1.5
Female	ASIAN ALONE	25 to 44 years	yes	0.75
Male	HISPANIC OR LATINO	65 years and over	yes	1.9
Male	HISPANIC OR LATINO	25 to 44 years	yes	1.6

When these weights are higher than 1, that individual survey respondent was from a group underrepresented on the survey; when these weights are lower than 1; the survey respondent was from a group overrepresented on the survey. Veracio users can use these weights to

statistically adjust the results of their surveys. For our example survey, this changes how the Veracio user will interpret her survey results.



Weighting changes the results of the survey quite a lot in this case, since it was a highly-biased survey. In the unweighted survey, the results were 38% NO and 62% YES. In the weighted survey, the results are 56% NO and 44% YES.



Reporting estimates of error

Veracio users are also interested in understanding the margin of error of their surveys, and how the weighting procedure implemented by the Veracio tool affects the margin of error and uncertainty in their surveys. The margin of error is also called the confidence interval. This is a measure of the level of uncertainty that surrounds an estimate. Often when you see results from surveys reported it will say, we estimate that such-and-such a percent will vote, plus or minus 2%. That 2% is the margin of error. It tells you the range of probable values for your results.

The raw survey data has two main sources of uncertainty, the margin of error and the bias. The margin of error is the uncertainty we just discussed. The bias is error in your survey results due to the survey sample not being an accurate reflection of the real population in question. The algorithm implemented by the Veracio tool reduces the bias as much as possible.

One result of this is that the margin of error of the weighted survey results is somewhat higher than the margin of error of the raw survey results; we can think of this as an example of the bias-variance tradeoff. For many biased surveys, the overall error is still lower after applying post-stratification weights. This can be a little confusing so let's see how this works for the example survey we have here.

In our raw example survey data, the margin of error is 0.032 (or 3%). In our raw survey data, the bias is 0.18 (or 18%). Taken together, the total error present in our raw survey data is $3\% + 18\% = 21\%$.

Table: Sources of error for the raw, unweighted data

Margin of Error	Bias	Total Error
0.032	0.18	0.212

So, all the results of this survey need to be thought of as representative of the actual results from the community, plus or minus 21%. In our example, this would mean that our raw survey tells us that 38% of the population believes “No”, plus or minus 21%. So somewhere between 17% and 59% of the population believes “No”.

Once Veracio has weighted the data, the margin of error has gone up a little to 0.043 (or 4%) but the bias has been reduced to zero. So, the total error for the weighted results is $4\% + 0\% = 4\%$. So, when looking at our weighted results, we can say that 56% of the population believes “No”, plus or minus 4%. So somewhere between 52% and 60% of the population believes “No”.

Table: Sources of error for the weighted data

Margin of Error	Bias	Total Error
0.043	0	0.043

Our example survey was very biased, so weighting the results has made a substantial improvement in our ability to get accurate results. It is issues exactly like this that the Veracio tool was designed to address.

If you still have questions about creating a survey with Veracio, we’re always here to help. Reach out to our support team via [email](#) or on [Twitter](#).