SITUATIONAL USE OF DATA WEIGHTING

TRC
Consumer Insights.
Market Innovations.

White Paper Series

Introduction

The culmination of a quantitative market research project is some type of reporting of study findings. The validity of such findings is predicated on any number of elements including questionnaire design, proper sample frames, respondent selection, data collection methods and data cleaning. However, the most important criterion in determining the validity of findings is the extent to which the data collected represents the actual population under study.

While every effort is made to obtain a representative sample of respondents, it is understood that often the final data set is not entirely representative of the population of interest. This may occur due to both reasons beyond the researcher's control (non-response bias, deficiencies in sample availability) as well as due to the researcher's intervention (quota sampling). Our final recourse before reporting findings is to take stock of the data set and determine if it needs to be altered in order to make it representative (or more representative). This alteration of the data is typically referred to as weighting.

Given the importance of ensuring the most representative sample possible, it is critical that the researcher understand when weighting might be required and the various methods for providing viable weights. The purpose of this paper is to provide a guideline for identifying those situations where data weighting is to be considered as well as the considerations that go into the decision to alter the data.

Brief Review of Weighting

Before considering various situations where data weighting is appropriate please consider a brief review of the weighting concept. If we assume our data set is representative, then analysis proceeds under the concept that the respondents in the sample represent the members of the population in proper proportion (for example, the percentage of males, females, customers, non-customers, etc. are nearly equivalent in the sample and the population). Having achieved proportional representation in our sample, respondents are grouped according to various characteristics and attitudes and tabulated accordingly; with each respondent counting as one person.

If a data set contains specific groups of respondents that are either over-represented or under-represented, then the sample is not indicative of the population and analyzing the data as collected is not appropriate. Instead, the data should be redistributed (or weighted) so that we achieve proportional representation in our sample. Specifically, each data point will carry a weight and rather than each respondent counting equally as one sample member, will thereafter represent either more or less than one sample member when results are tabulated.

This process will be illustrated more clearly by example in the following section. The important concept to remember is that the goal of any study is to obtain a representative sample. If that is not achieved naturally, redistribution of the data is required to yield one.

Scenarios Where Data Weighting May Be Required

Intentional Bias - Quota Sampling

One of the most common circumstances where weighting is necessary is when quota sampling is employed. Typically, studies employ quota sampling in order to obtain readable sample sizes for various sub-segments of interest from the population. This ensures adequate margin of error for segment-to-segment comparisons. However, this stratified sampling approach does not typically yield a viable total sample. Should the researcher desire analysis of the total sample as well as segment comparisons, redistribution of the data is required.

Let us consider a simple example of a study where stratified sampling is utilized.

In Example 1 (the next page), the weighted data must be used when analyzing results that combine the two quota cells. When tabulating the weighted data each Economy respondent will count as 1.8 persons and each First/Business respondent will count as .2 persons.

While this example is very simple, the processes employed are replicated in more complicated weighting schemes. This technique can be applied to multiple cells and across intersections of various quota designations (gender by region, age by race). In almost all instances where quota sampling is utilized weighting is required. Hence, prior to tabulation the researcher needs to consider the appropriate actions

Unintentional Bias

<u>Non-response</u>

In quota sampling, the researcher intentionally introduces bias into the data by establishing a certain number of interviews regardless of a particular population segment. In this next section we look at unintentional data bias. A common form of this is known as non-response bias. This occurs when particular types of respondents are not reached during the study.

Example 1:
A survey of 200 airline travelers was conducted for an airline that wants to measure overall satisfaction among customers as well as compare satisfaction by types of travelers. 100 First/Business Class and 100 Economy Class travelers were surveyed.

We can compare results by the two classes of passengers, however; in fact, only 10% of travelers fly First/Business Class. Since our data as it stands is not representative of the population, redistribution is required.

Data are weighted according to the following scheme:

CELL	Actual # Respondents	Actual Distribution of Cells	Desired # Respondents	Weight
First/Business	100	10%	20	0.20
Economy	100	90%	180	1.80

Historical data tell us that there are certain respondents that are more difficult to reach overall (younger, affluent) and also based on the type of methodology used (no Internet access, call blocking). There are certain sampling techniques that can be utilized to minimize such bias, though it often still exists.

Weighting can be used to help mitigate the effects of non-response bias. In such instances the researcher compares the distribution of key classification variables in the sample to the actual population distribution. If the distribution in the sample is not correct the data would be weighted using the techniques described in the prior section. One drawback here is that when studying artificial populations such as customers and prospects we often do not have reliable distributions for comparison.

Sample Balancing

Let us now consider data sets that are obtained from a particular geographic region – these can include studies that are conducted on the national level, statewide or within a particular city or county. For studies of this type there is a large amount of descriptive data available (via census figures) about the population of interest. As such, when samples of these geographies are obtained it is imperative that the researcher consider the sample distribution of respondent demographics.

When making comparisons to census data there are any number of characteristics that can be used. It is up to the researcher to determine which are to be used in the weighting scheme. Obviously, any variable used must be included in the survey data. Another consideration is missing values in the survey data. Since these are difficult to account for, variables with a high proportion of missing values (such as income) are often excluded. Once the weighting variables are identified the researcher needs to compute the actual weights.

If only two characteristics are to be used in weighting the data the researcher might employ a technique identical to the one portrayed in our airline study example. Let's say we want to weight on gender (2 groups) and region (9 census designations). This would yield a total of 18 cells. The process would be the same as described for two groups (in Example 1), where the researcher would determine the desired number of respondents in each cell and weight accordingly.

However, with this type of data there are often more than two characteristics of interest – for example gender (2 groups), region (9 groups), age (5 groups) and race (4 groups) – all of which yields 360 individual cells to populate. This poses a number of *difficulties. First, while distributions* for individual variables might be available, the distribution for each combination might not. Second each individual cell might not have survey respondents populating the cell, making weighting impossible. To combat this, we can employ a technique called sample balancing.

In sample balancing the weighting variables are redistributed individually, rather than computed cross variable cells (example males; 18-24; Northeast). The process is iterative with weights being applied on variable #1, then that new data set is weighted on variable #2; then that new data set is weighted on variable #3. This process is repeated again and again in order to achieve distributions on all the weighting variables that are close to the population.

Because non-response bias is difficult to measure, the researcher should apply some type of sample balancing on the data regardless of how it looks unweighted. Whether to employ sample balancing or simply compute weights for individual cells is dependent on the population information available, the condition of the survey data and the number of respondents available in each cell.

Comparing Two Samples

Aside from making sure that data sets are representative, data weighting can also be a useful tool in comparing two samples. This can occur in any number of instances; some of the most common being Test vs. Control studies, Wave-to-Wave studies and studies that mix methodologies. In this section we will examine how data weighting might be utilized.

A basic premise of Test vs. Control studies is that there are two (or more) comparable populations in every respect but one – which is exposure to some type of stimuli. The goal of the study is to see if such exposure changes attitudes and/or behaviors. Because of this, it is imperative that the researcher is confident that

differences seen between test and control groups be attributable to the stimuli and not inherent differences in the group composition. As such, before making comparisons across test and control cells, the researcher needs to compare the data sets on key demographic and behaviors that they would expect to be similar.

Example #2

A cable company wishes to determine the impact of a new movie service. A random sample of customers is selected to receive the service. After three months, two groups are surveyed; those with the new service (test) and those without (control). Satisfaction, loyalty and viewership levels are to be compared.

Upon completion we find a disparity in older respondents in the test cell (35% age 65+ vs. 20% in the control). It is known from past research that satisfaction and viewership varies by age. How can this data be used to compare the groups?

The researcher can redistribute the data from one of the data sets to match the other. This would be analogous to redistributing sample data to match the population. Here, we are less concerned with the distribution of age in the population than we are in having two comparable data sets. In this instance the researcher might take the Control Group distribution and weight the Test data to wind up with a comparable 20% of respondents age 65+. This would then allow comparisons of key variables across the cells.

This same process applies to wave studies. Again, the researcher is looking to compare groups, in this instance across time. The underlying assumption is that in each wave comparable, representative samples of some population are being drawn. Of course, it may occur that in one wave a bias is introduced. As such, before analyzing wave-to-wave data the researcher should compare key demographic and behavioral variables to ensure there is no change in the composition of the data sets. Should there be any changes, the researcher should consider weighting the data prior to analysis in order to ensure that any differences are real and not due to a sampling bias.

Emphasizing Key Characteristics

A slightly different take on weighting occurs when we want to emphasize a specific characteristic of the respondent. This technique is often used when there are certain respondents we want to count more heavily due to their importance. For example, a client may want to overemphasize the opinions and behaviors of customers who spend the most money with them.

The approach in this scenario remains basically the same as has been described previously. The analysis moves away from the idea that each respondent counts as one and utilizes redistribution of the data. In the case where a client wanted to look at survey data based upon dollars spent, the researcher would assign each respondent a weight based on this variable. So if a customer spends $100,000 with the client they would get a weight of 100 while someone who spends $5,000 would get a weight of 5. When the data are weighted and then analyzed, responses are distributed so that instead of looking at opinions based upon people, the data would be showing responses based upon dollars spent.

Some Considerations

Regardless of the scenario that dictates the use of data weighting, there are a few things the researcher should always consider. Most important is whether weighting is actually required. While weighting is a useful tool, the preference is always not to weight. Statistical formulas such as standard deviation and margin of error (among others) are all based on the proportional relationship between sample and population. Once the data are redistributed, the computation of these measures becomes more complicated. If unsure if weighting is necessary the researcher should compare key measures on weighted and unweighted data. If the two data sets are generally comparable, the recommendation would be to leave the data as is.

Secondly, consideration should be made as to the number of cases in each of the cells that is going to be weighted. Because margin of error is dependent on the base size the researcher wants to avoid too high a weight placed on too small a sample size. For example, if you only interviewed 50 males of 1,000 but you want them to represent ½ the population; this is a problem. In this instance each respondent is going to count as 10 respondents; because we only have a sample of 50 there is a large margin of error, meaning the chances that our 50 are not representative is high and any skewness in the original 50 is only going to be magnified.

Of course, each study is a unique case and consideration to weight the data must be given to any number of factors such as how skewed is the sample, what is the data to be used for, the cost of interviewing more respondents per cell among others. The key is that the researcher be cognizant of the extent to which their data sets are representative, investigate the use of data redistribution and proceed accordingly.