

10 ESSENTIAL PREREQUISITES FOR SURVEY DATA ANALYSIS

By John Sevec



mTab, 1100 E. Orangethorpe Ave, Suite #251,
Anaheim, CA 92801
Phone: 1.800.693.3110
Email: info@mtab.com
Web: www.mtab.com

BACKGROUND

While there is broad demand for access to survey data results, a relative few have intimate involvement in shaping the manner of survey administration and the collection of survey data. The availability of easy to use survey analysis software tools promotes the research analyst's ability to jump right into the analysis of survey data without the critical perspective of the design of the survey, resulting in an unfortunate disconnect between the analysis of the survey data and the survey methodology / questionnaire. This disconnect all too often leads to flawed analysis, improper conclusions, and misinformed insights.

The “Ten Essentials” presented in this paper provide the basic steps that a researcher should follow when analyzing the results of any survey to ensure that the insights and conclusions produced are valid and based on a sound understanding of the underlying survey and its responses.

10 Essential Prerequisites for Survey Data Analysis:

- 1) Understand the Questionnaire
- 2) Develop a Survey Analysis Plan
- 3) Understand the Survey Methodology
- 4) Check Survey Data Weighting
- 5) Match the Database Implementation to the Questionnaire
- 6) Don't Assume you can Conduct A Regional Analysis
- 7) Identify the Key Survey Questions
- 8) Identify the Key Survey Segmenting Questions
- 9) Eliminate Noise
- 10) Cross Check Your Results

1) Understand the Questionnaire

You can't ask a meaningful question of the survey data without a full understanding of the questions asked of the respondents. For example, do not assume that you can determine customer loyalty or advertising awareness just because your firm has recently conducted a survey of your customers; you first need to understand if the questions underlying the survey will support your intended analysis.

While examining the questionnaire, put yourself in the position of a typical respondent. Give thought to how a respondent would interpret the survey questions. Are there any ambiguous survey questions that will require careful consideration of their result?

Make note of the implications of the questionnaire, such as skip patterns identifying questions that have responses based upon the result of a previous question. For example, you would need to know that the rating question "Overall opinion?" was asked only of those answering YES to the question "Did you experience any serious problems?" prior to presenting the results of your analysis.

2) Develop a Survey Analysis Plan

While you're reviewing the questionnaire, make notes of reports, charts and overall project needs. Your subsequent analysis of the data will benefit if you start with several hypotheses to prove or disprove prior to diving into the data. Your survey analysis plan will provide structure and focus and prevent you from becoming overwhelmed and lost within the survey results. If your survey analysis software allows for the convenient trending of survey data, be sure to make note of the questions that would be particularly interesting from the standpoint of longitudinal analysis.

3) Understand the Survey Methodology

Once you understand the survey questions, you need to understand who provided the answers! Would it make a difference if the survey was fielded on Facebook or from the AARP's website? Similarly, you'd probably want to know that the product survey you are about to analyze was only fielded among respondents rejecting your project.

If the survey project was purchased from a syndicated market research provider, you should be able to obtain this information from that organization's website. If this survey was administered internally, then you may need to dig deeper.

Here are a series of questions to ask to help you understand the methodology and objectives of the survey project:

- a) How did respondents qualify to take this survey?
- b) How was the survey conducted (mail, phone, e-mail, web, personal interviews, or some combination of the above)
- c) Was the survey response triggered by some event, such as a recent product purchase or a service occasion?
- d) Was there a screening process involved? Is the data reflective of pre-screened respondents only, or will I need to identify a question or variable that identifies respondents that passed the screening process?
- e) Were there quotas established for specific target groups of respondents?
- f) How many respondents are there overall and for each quota group?
- g) Over what time period was the survey conducted?
- h) Does the survey data reflect an individual wave or multiple waves of data?
- i) Was the same respondent asked to take the survey multiple times (i.e., a “diary” type survey) or is each survey response unique to an individual respondent?
- j) Was the same respondent asked a series of questions about various occasions? For example, a set of questions rating each service occasion the respondent experienced, or each separate occasion during which the respondent used or consumed a product?
- k) Was this survey designed to obtain results by region or by key markets?
- l) Is this a competitive benchmarking survey project? If so, who were the key competitors, or said another way, were any competitors specifically excluded for some reason?
- m) Was a cleaning process employed post-collection of the data? Simple example: were respondents that answered “Extremely Important” to every Likert scale question removed from the sample? Were cross-check questions employed within the survey design to identify consistency within the respondent’s answers?

Given consideration of your survey analysis plan, do you need to rethink your report objectives based upon your understanding of the survey project methodology?

4) Check Survey Data Weighting

Survey data can be weighted for many different reasons. Most commonly, weighting is applied to adjust for sampling quota discrepancies. In this case, the weighting is ideally mild, meaning that any single respondent is unlikely to be weighted more than twice the weighting of any other respondent.

For a basic understanding of the data, you need to know if the survey data was weighted, and if so, how. Check the implications of the weighting by performing a frequency analysis using un-weighted and weighted data of the questions or variables included within the weighting criteria. For example, if the survey data was mildly weighted by age, compare the un-weighted and weighted frequency distribution of the respondent age question to determine the impact of the weighting.

Often studies are weighted for other reasons. A competitive buyer study, for example, may incorporate sales weighting in which respondents are weighted as a function of the products purchased. The weights are based upon product sales over a known time period. In this case, the impact of the weighting will be enormous but by design; the analysis of the survey data on a product basis was designed to reflect the relevant differences in product sales.

It is quite possible that you'll find your survey data has multiple weights, and the alternative weights may be required for the appropriate analysis of selected survey questions. If your study has multiple weights, you will need to know if your survey analysis software will automatically employ the appropriate weight as required. If your survey analysis software is not up to this task, you will need to know when and how to incorporate the appropriate weighting for any given analysis.

5) Match the Database Implementation to the Questionnaire

Congratulations, you're now ready to crack open the study with your survey analysis software. Prior to analyzing the survey data, take a moment to compare the survey database questions with the questionnaire and your survey analysis plan. Is each survey question represented within data? Are there additional questions within the survey data that you cannot identify on the questionnaire?

It is quite common for survey data to include respondent information that may have been available in advance of the survey if the questionnaire was conducted from an existing respondent panel, or other pre-determined source.

Additional questions may also represent added value to the survey data that your analysis will benefit from. For example, if your questionnaire included open-end responses (e.g., "What did you like about our product/service?") the data may include the open-end question responses (i.e., verbatim responses) as well as quantitative questions representing coded categories of the open-end respondents.

Other examples of survey data questions incremental to the questionnaire include the results of cluster or factor analyses. Cluster and factor analysis are statistical techniques that summarize the responses of multiple survey questions into a single new segmenting question. Be sure to include these questions within the “key segmenting questions” which we will discuss in more detail below.

Similarly, the survey data may include additional questions such as indices, or variables specific to the survey project; for example, reach and frequency questions included within an advertising awareness survey database. When compiling mTab survey databases, we often work with our customers to incorporate external data, such as customer support or CRM data which can greatly increase the value of the analysis.

Don't look past the additional questions within your survey database just because you couldn't square them away with the questionnaire; do a little digging to understand their full meaning. Chances are, the questions were added to enhance the analysis of the survey data, and it will be to your benefit to understand their meaning and intent.

6) Don't Assume You Can Conduct a Regional Analysis

Congratulations, you're now ready to crack open the study with your survey analysis software. Prior to analyzing the survey data, take a moment to compare the survey database questions with the questionnaire and your survey analysis plan. Is each survey question represented within data? Are there additional questions within the survey data that you cannot identify on the questionnaire?

It is quite common for survey data to include respondent information that may have been available in advance of the survey if the questionnaire was conducted from an existing respondent panel, or other pre-determined source.

Additional questions may also represent added value to the survey data that your analysis will benefit from. For example, if your questionnaire included open-end responses (e.g., “What did you like about our product/service?”) the data may include the open-end question responses (i.e., verbatim responses) as well as quantitative questions representing coded categories of the open-end respondents.

7) Identify the Key Survey Question(s)

Based upon your review of the questionnaire and methodology steps, you should be able to identify the key survey question(s). For example, in the case of a competitive buyer study, the question representing the product purchased would be a key study question. In other cases, the sampling quota variables (e.g., age, income, market) would be considered key questions.

Once identified, you need to explore the key questions by running weighted and un-weighted relative frequencies, including an analysis of the counts of respondents that elected not to answer the key questions. Here you are looking for small sample quotas which may require data aggregation prior to analysis. You will want to be sure that your survey analysis software supports on-the-fly aggregation of question responses that will facilitate this process. In the worst case, you may find that you need to revise your survey analysis plan to eliminate analyses you've planned to cover. It's better to eliminate the analysis now rather than completing your analysis based upon insufficient sample or invalid methodology.

8) Identify the Key Survey Segmenting Questions

Similar to the key study questions, while examining the questionnaire, an experienced research analyst would take note of the questions that would likely yield interesting variations between respondents. If you are new to survey analysis, respondent age is usually a safe bet as a segmenting question, so include age at a minimum.

Once again, run weighted and un-weighted frequencies of the segmenting questions, being sure to examine the respondents that elected not to answer. Use this analysis to determine if you need to aggregate the responses of the segmenting questions or if you should expand your search and look for other segmenting questions.

The key study and segmenting questions you've established in the steps above will serve as column banners for the subsequent cross tabulation analysis of the survey data.

9) Eliminate Noise

After the consideration of the steps listed above, we're now ready to dive into a detailed analysis of the survey data. If every survey question yielded new and exciting information, we could simply start cross tabulating each individual question against our previously selected key question and segmenting question column banners. In reality, identifying insight is more like finding a needle in a haystack. You will benefit by approaching the analysis in an orderly and consistent fashion and by leveraging the data mining features of your survey analysis software.

Here are a few tips for sifting through the survey data:

- With your key questions and segmenting questions selected as column banners, select all other questions concurrently as rows, ranking the results in a manner that identifies the key variations in segmenting questions.

- Summarize the data by formatting Likert scale questions as either means (weighted averages) or “top box” questions.
- Use graphical displays like radar charts and scatter charts which allow you to identify outliers at a glance.
- Use significance testing to identify the statistical significance of differences in key and segmenting questions, ranking the results using the test statistics to conveniently seek out the survey questions that yield significant differences.

The objective is to narrow down your analysis to the subset of survey questions that yield interesting variations within the key and segmenting questions. Compare the remaining questions to your survey analysis plan. Did you find additional questions of interest that were not included within your survey analysis plan? The additional questions deserve additional analysis as they may represent new insights.

10) Cross Check Your Results

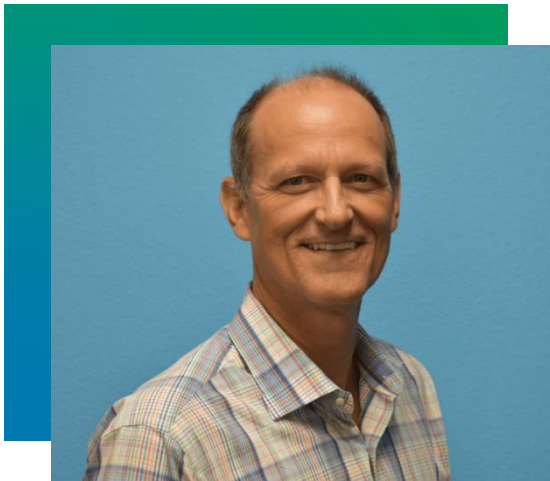
In many cases, the analysis of survey data stems from the initial presentation or summary report of a survey project that left the impression “we need to look into this in more detail”. If this is your situation, you should make every attempt to replicate the key tables and charts provided within the presentation materials. If you can tie back to these results, you will greatly enhance your confidence in extending the analysis beyond the presentation’s top line measures. However, don’t be surprised if you find that you can’t immediately match the presentation reports. This doesn’t mean that your analysis is wrong; more likely, the presentation reports have additional qualifying conditions that are not fully documented. For example, the reports may include a respondent filtering criterion that you did not include within your analysis. You should approach these differences as new and interesting data and look to understanding the implications of the differences between your analysis and the reports. If there were indeed qualifying conditions, why were they employed?

Always keep in mind that different does not mean wrong! With the complete understanding of the survey project as outlined in the steps above, you should feel confident in your analysis of the survey data and your conclusions will be equally valid. As you present your results, be careful to clearly spell out the underlying assumptions to make it painfully clear how you’ve constructed your analysis.

Conclusions

As research analysts, we strive to obtain the most interesting, insightful and accurate information possible in a timely manner. Our companies and customers rely on our ability to distill a very complicated subject into actionable information. The analysis of survey data is based upon the bedrock of the survey project's methodology. You are at risk of providing misleading conclusions if you do not take the time to obtain a basic understanding of the survey project's underpinnings. With proper consideration of the "Ten Essentials" listed above, your analysis will be sound and lead to insights that can shape the direction of your business.

About the Author



John is the SVP of Business Development at mTab, where he is responsible for overseeing the organic growth of the business. John joined mTab in 1990.

John has over 30 years of marketing information and research experience, 27 of which have been at mTab.