**FOURTEEN Quantitative Data Analysis**

On the one hand, data analysis is a core competence of market research professionals. Thus, any graduate-level market research text will devote hundreds of pages to this topic. On the other hand, the material in those chapters is virtually indistinguishable from what might be found in any text on applied statistics or statistics for the social and behavioral sciences. In that sense, there is little that is unique about the data analyses performed for market research. In essence, the same statistical procedures are used across all the social sciences, and many of the practitioners of data analysis in market research will have a PhD in some related discipline like psychology, sociology, or economics. In fact, there are three specific statistical procedures, widely used across the social sciences, that account for the bulk of the data analyses actually performed in day-to-day commercial market research. The limited goal of this chapter is to give the general manager and the engineering manager a handle on these procedures so that you know what to expect.

Given the brevity of this chapter, its goals must be especially limited. Think of it as a briefing. I can’t teach you how to do data analysis in this space: that requires at least the hundreds of pages found in the standard market research text or, more properly, the years of practice designing and conducting statistical analyses that form the core of most contemporary social science PhD programs. What I can do is give you the names of things and put these names in a context. If this reduces your bewilderment or increases your composure the next time a market research study is discussed, then you should be able to ask more critical questions and ensure that the proposed study meets your needs as a decision maker. In addition, I hope to make you a more critical reader of secondary research and completed market research reports.

**Procedure**

First, an overview of how quantitative “data” come into existence and get analyzed and reported.

1. In most cases, somehow, some way, an Excel spreadsheet containing numbers corresponding to respondents’ answers gets created. If the procedure was a web-based survey, involved computer-assisted telephone interviewing, or was administered on a computer (as in the case of many conjoint studies), then the Excel spreadsheet is created automatically as part of data collection. If paper forms were used, then someone entered the responses, represented numerically, into Excel. The Excel data are generally in matrix form (“flat file,” in computerese) where the rows correspond to individual respondents and the columns contain a numerical representation of each respondent’s answer to each of the questions administered. For example, if a 5-point scale was used to measure purchase intentions, and that respondent indicated he would definitely purchase the new product when available, a “5” might be entered in the corresponding cell of the spreadsheet.

2. From Excel, the data are typically imported to a specialized statistical analysis program. SPSS ([www.spss.com](http://www.spss.com/)) and SAS ([www.sas.com](http://www.sas.com/)) are two comprehensive packages used by many academics; numerous other such packages exist, including free versions available over the web. Within the statistical package, raw data can be recorded, transformed, aggregated, or disaggregated at will, and virtually any statistical analysis can be performed simply by pulling down a menu and selecting a few options.

3. The data analyst often has a PhD, but routinized analyses on standardized instruments, such as a satisfaction questionnaire, may be performed by MBAs or other people who have accumulated hands-on experience. If the study is a one-off affair, then the analyst will probably generate a variety of tentative analyses never seen by the client as the analyst assimilates the data and decides on a reporting approach. If the study is more routinized, then analysis may be as simple as pushing a button to trigger a canned series of tests. In this case, the analyst sees roughly the same output as the client.

4. The results of the analyses are formatted as tables and embedded into a narrative (which, in routinized cases, such as the *n*th satisfaction study done within the banking category, may be largely boilerplate).

5. Results are presented to the client and discussed. Depending on the contract and what was paid for, the research firm may attempt to add quite a bit of interpretation to the data, to the point of recommending specific courses of action based on the data. Alternatively, the research firm may confine itself to explaining, clarifying, and defending the validity of the results and the procedures used, leaving substantive interpretation to client management.

**Types of Data Analysis in Market Research**

Most data analysis in market research consists of one of the following:

1. Tabulating and cross-tabulating proportions; an example would be agreement with an opinion item cross-tabulated with some other factor such as brand owned or education, in the consumer case, or size of business in the B2B case.
2. Comparing means (averages) across items, groups of customers, or time periods; an example would be total annual expenditure on the product category for males versus females, in summer vs. winter, on the West Coast vs. the East Coast, or by homeowners versus renters.
3. Predicting an outcome as a function of antecedent variables, as when level of satisfaction is shown to covary with expenditure, length of relationship with the vendor, number of changes in account team personnel, size of customer, and type of service contract.

Many more esoteric kinds of analyses, such as multidimensional scaling or structural equations models with latent variables (to convey the flavor of the jargon), also play a role in market research, but these three types of analyses are the workhorses employed every day. In the case of each of these procedures, individual numbers are compared in an attempt to detect meaningful or real differences. How do customers of Brand A differ from customers of Brand B? Which attributes are worth a lot of money to customers, and which attributes are not worth any money? Which factors serve to increase customer satisfaction, and which have no effect?

Sometimes the reality of the difference between two numbers is deemed to be obvious, as in the following cross-tabulation:

If the sample is large, further statistical analysis of this comparison only confirms what we see at a glance—owners of Brands A and B hold very different opinions. On the other hand, much of the time the data looks more like this breakdown of the customer base for each of three brands:

Or maybe like this:

As we move from 2 × 2 cross-tabulations to 3 × 5 cross-tabulations, and as the different results move closer and as the numbers become more abstract, as in the preference ratings, it becomes more and more difficult to say, with confidence, that customers of Brand A tend to have lower incomes than customers of Brand B, or that different segments place different importance weights on key features.

If you recall the sampling chapter, you know that all these numbers obtained on the sample are only fallible estimates of the true population values in any case. The question becomes, When does a difference between two numbers obtained in the sample represent a reliable, actionable difference, and when is the difference only apparent and dismissible as an artifact of the random variation inherent in all sample data? (Sometimes the question is more naturally thought of as whether two items have a real *association*; but this reduces to the question of whether the degree of association is different from zero. Hence, I shall refer to “difference” throughout this discussion.)

Modern statistical data analysis arose to address precisely this question: Which apparent differences in data are real, and which are not? It is important to understand that statistical analysis gives us no direct access to the truth; it simply indicates whether the apparent difference is probably real or, alternatively, how probable it is that the apparent difference in the sample reflects a real difference in the population, so that it is signal rather than noise.

By convention, an apparent difference is accepted as a real difference if, when the appropriate test statistic is applied, it indicates that the difference in question would arise by chance in fewer than 5 of 100 cases. Now, let us unpack that unwieldy sentence. Every kind of data difference has associated with it one or more statistical procedures, and each such statistical procedure allows a computation of a number known as a test statistic. These test statistics are computed using the same sorts of assumptions about probability and probability distributions as underlie the discussion of sample size.

Suffice to say that when you calculate a test statistic, you acknowledge that if the study were repeated with a new probability sample, you would not get exactly the same results each time. This variability follows from the fact that you are drawing a limited sample from a very large population. The mathematics underlying the test statistic then envisions repeating the study an infinite number of times and uses the data in the present sample to estimate how often, in that infinite series of repetitions, you would get a difference this big *by chance alone*. If the answer is fewer than 5 in 100, such results are generally given an asterisk and a footnote that reads something like “*p* < .05.”

The important thing to retain from this discussion is that statistical analysis is simply a set of mathematically based conventions for determining when you can accept an apparent difference as a real difference. More pointedly, if no statistical analysis has been applied, and the difference is not of the order 80–20 versus 20–80, you ought not to assume that the apparent difference is a real one. If you are reading a report that makes much of certain apparent differences—say, the 52 percent agreement in segment 1 versus the 43 percent agreement in segment 2—but that includes no statistical analysis, then you should become very suspicious. Worse, if there is a plethora of such statistical analyses, but the sample is not a probability sample, then you should question how much weight can be placed on the results; strictly speaking, in the absence of a probability sample, the results of the significance tests are not correct. More exactly, the results of conducting a statistical test on a nonprobability sample are of unknown validity. You simply don’t know whether to accept them or not.

**Managerial Perspective on Data Analysis**

As a general manager receiving the results of data analyses, your primary responsibility is to understand that apparent differences need not be real differences. It behooves a general manager to have a deep humility concerning the ability of any research endeavor to produce a picture of the world as it truly is. Given this stance, you know to be skeptical as soon as you are intrigued. That is, when you see a difference that seems actionable or would resolve an uncertainty, your next response should always be, Is it real? Your second responsibility, then, is to understand the role played by statistical analysis in vetting apparent differences. You must have the discipline not to accept reported differences at face value, absent an appropriate test. Your final responsibility is to accept that even statistical analysis only provides an estimate of the odds that a difference is real rather than apparent. It takes quite a bit of tough-mindedness to accept that some proportion of statistical judgments indicating that a difference is real will be wrong—about 5 in 100, actually. Because general managers may encounter hundreds of data comparisons in a year, they can be virtually certain that some differences vetted as real are not. As I said at the outset, market research reduces uncertainty but cannot eliminate it.

Put in more positive terms, if a good sample was obtained—that is, a correctly sized probability sample—and if two key numbers appear to be really different (*p* < .05), and you have no countervailing data or experience, then you should feel comfortable acting on these results. More than 95 times out of 100, the data will be vindicated. But remember: It is never 100 percent certain.

**Dos and Don’ts**

* ***Don’t*** be intimidated by statistical tests, no matter how scanty your education in this area. From a managerial standpoint, every test, no matter how esoteric, is just an effort to mark out certain findings as real and noteworthy.
* ***Do*** be suspicious of any quantitative research report that doesn’t include statistical tests. Without a test, how are you to know which nominal differences can support a decision?
* ***Don’t*** accept a statistical test at face value in the absence of information about the sampling procedure and a judgment as to whether it was a probability sample.
* ***Do*** acquire a habit of skepticism when presented with table after table of numbers, and slide after slide of full-color charts and graphs. Which differences are real, and which are just random variation?

**Discussion Questions**

1. Although many market research efforts use only some combination of (1) comparison of means, (2) cross-tabulations, and (3) regression, there are many other statistical tools available to be applied when specialized circumstances demand.
	1. Test your skills as a web searcher: Identify half a dozen other statistical tools that are applied with some frequency in market research. For each one, explain what it does for the manager that tests of means, cross-tabulations, or regressions cannot do. Pinpoint the unique insight available through this tool, and the kinds of marketing decisions it might support.
	2. How many of these tools are applicable to: (1) mostly B2C, (2) mostly B2B, or (3) equally applicable? Discuss.
2. In the prior chapter the claim was made that all statistical tests presume a probability sample. Explain why this is so. (You will need to recall your instruction in prior statistics classes.)
3. You read in a report, or hear in a presentation, a test result. For instance, a cross-tabulation is tested and men and women are shown to hold different opinions “at *p* < .05.” You determine that the research did not use a probability sample. It appears to be a bona fide sample from the population of interest, but it is certainly not a probability sample.
	1. What stance would you take toward the finding that “men and women hold different opinions”? Your answer could be anything from “the truth value of this finding is unknown” to “men and women probably do differ, but maybe not by quite as much as the cross-tabulation shows.” Explain and defend your answer.
	2. How does your answer change, if at all, if I change the example in each of the following ways:
		1. The finding was instead that men and women do *not* hold different opinions, as seen by the test result, which was “*p* > .20.”
		2. The original finding of a difference was actually “*p* < .01.”
		3. The sample was twice as large as you initially thought, say, 800 versus 400.
		4. The difference between men and women was lopsided: 2:1 versus 1:2 in terms of the ratio of agreement to disagreement in men versus women.
		5. *Extra Credit*: Find the reverse split, men versus women, in a sample of 400, that would produce a chi-square value of exactly 3.84 (corresponding to an alpha of exactly .05).
		6. *More Extra Credit*: How big would the sample have to be for a reverse 2:1 split across men and women to produce a chi-square of exactly 3.84?
		7. If you do the extra credit, does this change your initial answer to (a)? Discuss.
4. A senior colleague, apprised that a probability sample was imperative if statistical tests were to be performed, shrugs and says “too bad—we’ve always included statistical tests in these reports and management expects to see them. I’m not going to drop the tests and I’m certainly not going to include some mealy-mouthed footnote about probability samples—I don’t need the grief.” Discuss how much weight should be placed on the tests in this report in terms of guiding real world decisions.
	1. You express your qualms more forcefully in a second conversation, and your colleague says, “Look—I understand the technical issue. This test of means I’m flagging as ‘significant at *p* < .05’ is probably not significant at exactly that level. If I collected a real probability sample over and over, sometimes I’d discover the probability to be only .09 (borderline significant, as some would say), and others it would be even better, maybe .01, strongly significant. The fact remains that the results are going to be roughly the same as in a probability sample, meaning that those differences that are significant at .05 are worth more attention than those differences where the odds of a chance result are .50. You see, the significance tests are just a way to flag certain differences and direct the attention of management to those. This is a business, not a classroom.”
	2. Now what would you say?

**Suggested Readings**

**Cases for Part IV**

**Suggested Case: Xerox #1 (Harvard 9-591-055)**

**Synopsis**

[This case has two foci: a set of satisfaction surveys, and a small experiment to test guarantees. This synopsis only applies to the satisfaction surveys; see below for a synopsis of the portion of the case that presents an experiment.][1](https://jigsaw.vitalsource.com/books/9781483387246/epub/OEBPS/s9781483398228.i985.xhtml#s9781483398228.i1032)

In the late 1980s, Xerox pioneered the development of ongoing surveys of customer satisfaction. At the time of the case, management had put in place four different surveys, and the questions below ask you to assess the value of these efforts.

**Questions**

1. These surveys were justified in part by a commitment to Total Quality Management. Explain why it is important to survey customers as part of such an effort. Why not just count defects and complaints, which would be arguably simpler and cheaper, and might pinpoint more exactly any manufacturing quality problems?
2. What is the business rationale for investments to achieve higher levels of customer satisfaction? What sort of financial payoff can a firm like Xerox reasonably expect?
3. At the time of the case, a quite massive survey effort is described. In any such instance, now or then, we may suppose that sooner or later, budgets will tighten. If you had to jettison one of the four survey efforts as least important, which one and why? Conversely, which one is so valuable that you would hold on to it until the very last?
4. Evaluate the sample size used in the Periodic Survey—is it too small, too big, or about right? How can you tell?
5. Consider the example survey instrument included as an exhibit. Critique this from the standpoint of good questionnaire design. Are any important types of questions missing? Are there questions that don’t really belong?
6. A lot has changed in both survey research and the business of measuring customer satisfaction in the decades since the Xerox effort described in this case. What, exactly, do you suppose a firm like Xerox does differently in 2015? What elements remain more or less the same?
	1. Hint: Answer separately in terms of what has changed in survey implementation, versus what has changed in the nature of the questionnaire used. Change may (or may not) be much greater in one versus the other.
	2. The easy answer here is “now they do it over the Web instead of with paper.” So what? Does that difference in administration make a difference to anything else?
		* i. Is the change to Web administration the biggest change, or are there any more fundamental changes?
	3. Regardless of changes in how satisfaction surveys are done or in the content of satisfaction questionnaires, has there been any change in the basic rationale for surveying customers about their satisfaction?
		* ii. Equivalently, would your answers to questions #1 and #2 above have been any different in 1990 than today? Why, or why not?

**Suggested Case: Xerox #2 (Harvard 9-591-055)**

**Synopsis**

[This case has two foci: a set of satisfaction surveys, and a small experiment to test guarantees. This synopsis only applies to the experiment; see above for a synopsis of the satisfaction survey portion, and see footnote to Xerox #1 re ordering information.]

Having improved its measured customer satisfaction, Xerox seeks competitive advantage based on the superior quality of its copiers. The suggestion under discussion at the top levels of management is whether to offer some kind of guarantee. This proves not to be a simple yes or no decision—there are multiple ways of constructing a guarantee, and no consensus as to the best approach. Management proceeds to conduct some focus groups along with a small experimental test, delivered by phone and supplemented with a brief questionnaire. Your task is to interpret the findings of this quantitative research to make a recommendation about guarantees.

**Questions**

1. Were the focus groups’ money well spent? What role did they play in supporting the experimental test?
2. Based on the results, which is the best guarantee to offer? What evidence supports your judgment?
	1. Your instructor may supply you with a mock-up of the underlying data. If so, select the crucial result(s), and apply the appropriate statistical test(s).
	2. Are the statistical results sufficiently stark to drive the decision, or will a judgment call be involved as well? Explain.
3. Could Xerox have just as well done a conjoint study rather than the simple laboratory experiment described? Why, or why not?

**Suggested Case: Star Digital (Harvard #M-347)**

**Synopsis**

Star Digital is described as a video rental company that advertises through banner or display ads on the Web. They have run an experiment in an attempt to quantify the ROI of their Web advertising. The experiment had a test and a control group, and a sample of the data is in hand for analysis of the results.

[Note: You must download the spreadsheet with the results of the experiment from the Harvard site, and load it into a statistics program of your choosing. (The Excel data pack does not contain the necessary statistical procedures.)]

**Questions**

1. Evaluate the strengths and weaknesses of the experimental design. If your charge had been to determine whether there was a positive ROI for this firm’s Web advertising, what, if anything, would you have done differently?
	1. It’s all right to ratify the design exactly as implemented; but you have to give reasons why there is no better or more optimal design.
2. Select an appropriate statistical procedure for this sort of data, and conduct the analyses needed to answer the key question: Did the Web advertising produce revenue in excess of its costs?
	1. Were the two advertising networks both effective?
	2. If so, were these equally effective?
3. You may find it helpful to return to the Europet case at the end of [Part II](https://jigsaw.vitalsource.com/books/9781483387246/epub/OEBPS/s9781483398228.i545.xhtml#s9781483398228.i545). Star Digital exemplifies the difference between online and offline advertising: Unlike Europet, here there is a direct connection between ad exposure and purchase.
	1. But, there is the same need for a two-step calculation to calculate ROI: (1) quantifying the statistical association between ad exposure and purchase and (2) combining that information with the cost of the ads and the lifetime profit contribution of converting a prospect into a customer.
4. In most ad campaigns, online or offline, some prospects will be exposed to the advertising multiple times; there is a general sense that up to a certain point, multiple exposures are helpful in converting prospects, and a necessary investment. Put another way, effective frequency is generally greater than one; few ads are so powerful as to make the sale after a single exposure. Frequency information was collected in this case; devise a statistical test to determine whether more frequent exposure was in fact associated with higher rates of conversion for the Star Digital ad campaign.
5. As with Europet, only a sample of the total results is provided. Discuss the gains and losses of running analyses on the sample, rather than the (huge) total dataset, and the implications of your answer for the hullabaloo surrounding Big Data.

**Suggested Case: Portland Trailblazers (Harvard #UV2971)**

**Synopsis**

A major league sports team, suffering from an attendance slump and multiple other ills, considers whether to redesign its ticket package. A conjoint study that includes price is conducted to assess the appeal of different possible ticket packages, and key results are available for interpretation.

[Note: This is a much more straightforward case than MSA, and suitable as the first case in a conjoint module].

**Questions**

1. Do any of the seats appear to be mispriced? What does “mispriced” mean in this context, and what operation could you perform on the conjoint results to answer this question?
2. Assuming there is an opportunity to increase (or decrease) the price of at least one seat type, use the conjoint results to calculate the exact price change to make the seat fairly priced relative to its value as perceived by customers.
3. Are any of the promotional offers cost-effective? By cost-effective, I mean that the value delivered to the customer exceeds the cost to provide it.
4. An alternative to re-pricing a seat that appears over-priced is to add promotional offers that bring additional value. This only makes sense, of course, if the promotional offer itself is “under-priced,” in terms of its cost to implement, versus the value it contributes in the eyes of customers.
	1. Is there a seat where this would make sense? If so, which promotional offer would you add?

**Suggested Case: MSA, Planning the AMAPS Product Line (Harvard #9-590-069)**

**Synopsis**

This is an enterprise software case, a hard-core, heavy-duty B2B technology case. It will be most successful in an MBA class consisting of working professionals, and where many students have experience with B2B markets.

At the time of the case, MSA has a problematic product offering within the defense contracting market, which itself has been in turmoil. The product consists of software to help defense manufacturers do resource planning. To help decide what to do about the problematic software offering, market research, including focus groups and a conjoint study, has been commissioned. The results of the research are given, and the task is to interpret these results.

**Questions**

1. State the decision problem that led to the market research reported.
2. If no market research had been done, could you predict management’s decision? Explain your answer with respect to what you know about organizational behavior, post-merger and acquisition, and anything you might know about the management of large-scale software projects.
3. Is there a big enough market to justify investing funds in resuscitating the AMAPS/g product? Provide a financial calculation based on the conjoint results.
4. How about the various enhancements to the AMAPS/g—do the conjoint results support investing in any of the enhancements to the core product? Again, provide a financial calculation based on the conjoint results.
5. How confident are you in the market size estimates derived from the conjoint? More exactly, drawing on all the research, including the focus groups and the survey questions that accompanied the conjoint, would you argue that the size estimates are biased upward, or probably conservative, or not obviously biased either way?
6. Why was it important (or was it not) to do focus groups before the conjoint? The focus groups probably added several tens of thousands of dollars, and a month or two, to an already expensive and time-consuming effort. If you think the groups were worthwhile, be specific about the value added, relative to just proceeding directly to the conjoint.
7. Could a survey, or an experiment, have replaced the conjoint? If not, what is it about the MSA situation that mandates a conjoint analysis?
8. This sort of B2B technology product category is sometimes presented as the kind of situation where only customer visits really make sense. If MSA had done only customer visits, with results directionally or impressionistically consistent with the actual conjoint results, what do you suppose management’s decision with respect to the AMAPS/g product would have been? Explain.

1I use both parts of this case and order it under the number given; if you are only going to use one part, Harvard provides a split A and B version, allowing you to order only the part you need.