IPDET

Handbook

Module 10

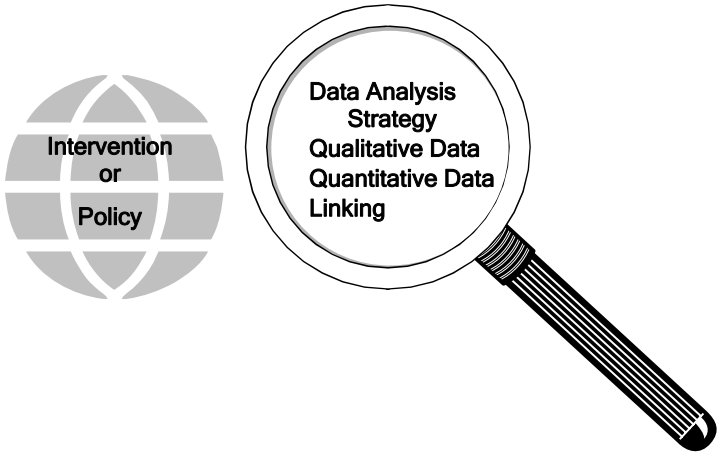
Data Analysis and Interpretation

Introduction

Now that you understand how evaluation questions begin the process, which then continues through design, data collection, and sampling, you will be able to think about the next stage – data analysis and interpretation. In this module, you will look at the ways to analyze and interpret your data so that you can answer the questions about the intervention or policy that you are evaluating.

This module has four topics. They are:

- Data Analysis Strategy
- Analyzing Qualitative Data
- Analyzing Quantitative Data
- Linking Quantitative Data and Qualitative Data.



Learning Objectives

By the end of the module, you should be able to:

- discuss when to use qualitative and quantitative data
- discuss how to analyze qualitative data. including: making good notes, drawing out themes and patterns, content analysis, summarizing qualitative data, controlling for bias, and affinity diagrams
- discuss how to analyze quantitative data
- discuss measures of dispersion including standard deviation
- discuss ways to analyze survey results
- define commonly used descriptive statistics, including: frequency, percent, mean, median, mode, money, percent change over two points in time, ratio and comparisons
- discuss measures of relationship
- discuss inferential statistics.

Key Words

You will find the following key words or phrases in this module. Watch for these and make sure that you understand what they mean and how they are used in the course.

qualitative analysis
quantitative analysis
themes and patterns
content analysis
affinity diagrams
descriptive statistics
measures of central tendency
mode
median
mean
nominal data
ordinal data
interval/ratio data
measures of dispersion
range
standard deviation
normal distribution
descriptive statistics
frequencies
percent (proportion) distributions
percent change over two points in time (rate of change)
comparisons
measures of association
measures of correlation
direct relationship
inverse relationship
correlation
independent variable
dependent variable
crosstabs
inferential statistics
statistical significance tests

Data Analysis Strategy

Developing the data analysis strategy is an important part of the planning process. It helps to know the options for data analysis, with the various strengths and weaknesses, as you plan your research. This is very important. In the design matrix, specify so that you know how you will use the information collected. A common mistake is collecting vast amounts of data that are never used. This overview will provide you with the big picture issues.

Whether you choose qualitative data or quantitative data, you will find your data collection and data analysis will overlap. Figure 10.1 shows a graph plotting data collection and data analysis over the evaluation. Notice that at the start of data collection, a small amount of time is spent in data analysis., especially if you do a pilot test first As the evaluation continues, more time is spent in data analysis and less in data collection.

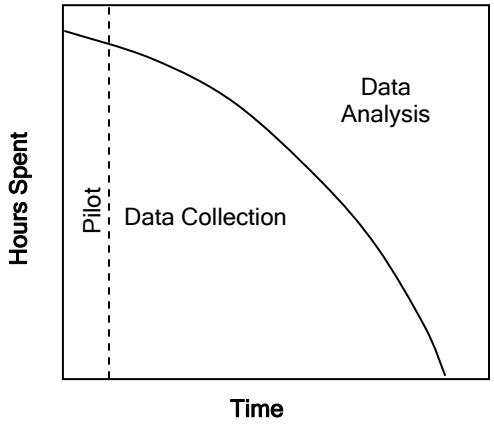


Fig. 10.1: Data Collection vs. Data Analysis over Time.

There are two key types of data analysis, qualitative and quantitative.

Qualitative analysis is best used in situations where we need a fairly in-depth understanding of the intervention, including cases where we are evaluating something relatively new, and can be used to answer questions like:

- Is the intervention being implemented according to plan?
- What are some of the difficulties faced by staff?
- Why did some participants drop out early?
- What is the experience like for participants?
- Are there any unexpected impacts on families and communities?

Quantitative analysis can be used to answer questions like:

- What is the percent distribution?
- What is the average?
- How do participants rate the usefulness and relevance of the intervention?
- How much variability is there in the data?
- What is the relationship between program objectives and its outcomes? How can we measure the strength of that relationship?
- Are the results statistically significant?

Analyzing Qualitative Data

Qualitative data analysis is used for any non-numerical data collected as part of the evaluation. Unstructured observations, open-ended interviews, analysis of written documents, and focus groups transcripts all require the use of qualitative techniques. Analyzing qualitative data is challenging, although many people find it interesting. Great care has to be taken in accurately capturing and interpreting qualitative data.

Making Good Notes

When you collect qualitative data, you want to capture as much information as possible. It is important to accurately capture your observations; good notes are essential. This means paying close attention to language: what people say and how they say it. Try not to interpret what they say as you write your notes. Write down anything thing you observed, any body language, or anything that happened while you were collecting data (for example, many interruptions during the interview). You may also want to capture your immediate thoughts, reactions, and interpretations. Keep them in a separate section of your notes.

As mentioned in an earlier module, it is extremely important to provide time immediately after an interview, observation, or focus group to review your preliminary notes and make additions, and write up your notes so they will make sense when you look at them later on. It is surprising how difficult it is to make sense of notes taken in an interview, focus group, or observation session – even from just a day earlier.

Even if you have tape-recorded the session, a small amount of time invested in a preliminary write-up while it is fresh in your mind will save hours and hours of listening to or watching tapes or poring over transcripts later on.

As you probably recall, **triangulation** is the use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment by cross-checking results. Triangulation is useful in qualitative data analysis. For example, you might use the following examples of three mixed sources of data:

- interviews, focus groups, and questionnaires
- questionnaires, available data, and expert panels
- observations, program records, and interviews
- interviews, diaries, and available data.

Table 10.1 is a summary of the key considerations in the early phase of qualitative data analysis.

Table 10.1: Key Considerations in the Early Phase of Qualitative Data Analysis

While collecting data	• Keep good records • Write up interviews, impressions, notes from focus groups • Make constant comparisons as you progress • Meet with the team regularly to compare notes and make adjustments
Write contact summary report	• Write a one page summary immediately after each major interview or focus group • Include all the main issues • Include any major information obtained • What was the most interesting, illuminating or important issue discussed or information obtained? • What new questions need to be explored?
Use tools to help you	• Create a separate file for your own reactions during the study, including your feelings, hunches, and reactions • File your ideas as they emerge • Keep a file of quotations from the collection process for use in bringing your narrative to life when you write your report
Get ready for the data analysis	• Make sure all of the information is in one place • Make copies and place originals in a central file • Use copies to write on, cut and paste as needed

Suggestions for Analysis

Once you have collected all your data, you may have pages of notes, tape recordings, observations, and/or other items. Gather your data together and find a place to work. You will probably need a large table to give you room to sort your information.

Nancy Porteous[1] et al suggest the following process for analyzing qualitative data (outlined in the next three pages).

Have the following materials available…:

- several highlighters (a different colour for each evaluation question)
- a worksheet for each evaluation question
- all data including notes, transcripts and/or tapes from all interviews or focus groups
- all collection tools for self-completed questionnaires, registration forms, observations, or chart reviews.

Figure 10.1A shows an example of a blank qualitative data analysis worksheet:

Evaluation Question:

Color, code, or symbol:

Topics	Quotes	Findings

Fig 10.1A: Qualitative Data Analysis Worksheet (example, blank).

[1] Nancy L. Porteous, Barbara J. Sheldrick, and Paula J. Stewart (1997). *Program evaluation tool kit: A blueprint for public health management.* Ottawa, Canada: Ottawa-Carleton Health Department. Pp 66-68. Used here with the permission of the author.

You will need at least one worksheet for each evaluation question. Then follow the following process.

- Write each evaluation question in the space provided at the top of each worksheet.
- Choose a code to identify your data. It might the color of a pen, pencil, or highlighter, or it might be a symbol. Record your color or symbol in the second space at the top of each worksheet.

Now that you have your worksheets identified and ready, you can begin going through your notes and materials and recoding the information on the worksheets.

Follow this procedure to record information:

- Read all completed tools or notes and transcripts in one sitting.
- Use your highlighters to mark the parts that deal with each evaluation question.
- Go back and carefully read all of the data that pertains to the first evaluation question.
- In the "Topics" column of the worksheet, write down each opinion, idea, or feeling that pertains to the expectations for that evaluation question.
- Even if a topic is expressed in a different way or using different terms during the session, do not re-write it.
- Leave a space between each topic, allowing you to keep track of how frequently each point is raised.
- Keep a tally of the number of times an opinion, idea or feeling is mentioned.

Figure 10.1B shows the worksheet after this step is completed.

Evaluation Questions: Were participants satisfied with the series?

Color, code, or symbol: Yellow

Topics	Quotes	Findings
Parents decide on topics		
~~///~~ ~~///~~ ~~///~~ ~~///~~ ~~///~~ ~~///~~ ////		
Cover a couple of topics per session		
~~///~~ ~~///~~ ~~///~~ ///		
Not enough time spent on each topic		
~~///~~ ~~///~~ ~~///~~ ///		

Fig. 10.1B: Qualitative Analysis Worksheet (example showing topics recorded)

Once you complete the "topics" column, complete the rest of the worksheet in the following way:

- From your notes, extract and insert quotes that best represent each topic.
- Make your conclusions about specific points and write them in the "Findings" column.
- Organize your findings by type or category
- Use numbers of responses to give precision and a sense of magnitude:

Occasionally, the minority view is important and you will need to report it. Use your judgement but always make it clear that only one or a few respondents expressed that opinion.

Figure 10.1C shows the completed worksheet.

Evaluation Questions: Were participants satisfied with the series?

Color, code, or symbol: Yellow

Topics	Quotes	Findings
Parents decide on topics 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 ////	I think the process of deciding would be valuable.	There was a strong feeling that parents should be more involved in the choice of topics.
Cover a couple of topics per session 𝍸 𝍸 𝍸 ///	Sometimes we just got into a topic and then it was time to leave or move to something else. We need more time to discuss	Many participants (38 or 52 interviewed) thought there should be more time for discussion.
Not enough time spent on each topic 𝍸 𝍸 𝍸 ///		

Fig. 10.1C: Qualitative Data Analysis Worksheet (example showing quotes and findings)

Drawing-out Themes and Patterns

When analyzing qualitative data, your goal is to summarize what you have seen or heard in terms of common words, phrases, themes, or patterns.

As you review the material, you will begin to make notes. It helps to read a few to get a sense of what is there and to develop a general framework for analyzing the rest of the data. However, you may discover other themes as you go along and may have to re-read earlier material.

As you identify the words, issues, themes, or patterns, identify where they are located so you can find them again if you need to verify exact quotes or context. This will be very tedious the first time you do it; as you gain experience, you will find you can locate potentially important information much more quickly.

Some people find it helpful to use a spreadsheet that identifies the common themes and where they are located in their notes. Other people use note cards to sort through qualitative data.

Case 10.1; ORET/MILIEV Programme in China

> "The Development and Environment Related Export Related Transactions (ORET/MILIEV) Program is a program to finance certain types of projects through a combination of a development cooperation grants and commercial loans. The program is designed to help general employment, boost trade and industry, and improve environmental quality in developing countries.[2]"

Table 10.2 shows the structure of the database of initial evidence for the 35 projects visited during the evaluation.

Table 10.2: Structure of the database of initial evidence for the visited projects[3].

Evaluation Criteria	Issues	Project 1	Project 2	…
Policy Relevance	Listed as a priority in the local/ministries' development plans			
	In line with sector development strategies at national level			
	O/M info-sources			
	Main reasons for applying for O/M			
	Without support, would the project be implemented			
	Who initiated the project			
	Effect on poverty alleviation (during appraisal)			
	Effect on W&D (during appraisal)			
	Effect on environment (during appraisal)			
	Whether the project/programme objectives be adjusted			
Efficiency	State of the project			
	Achieved the objectives on schedule and within budget			
	Delays during implementation comparing with fact sheet			
	Delays during implementation according to end user			
	Main causes of delays			
	Price of the equipment			
	Main causes resulting in higher prices			
	Period of appraisal by Dutch side			
	Duration of application procedures in Dutch			
	Duration of (whole) application procedures in China			
	On-lending procedure (time-consuming)			
	Acquiring tariff exemption			
	Cooperation between end user and supplier			
	Comments on efficiency by end user			
Effectiveness	Achievement of the short-term objectives			
	Achievement of the long-term objectives			
	Successfully trained personnel			
	Quality of delivered goods/services			
	Spare parts supply			
	Prices of spare parts			
	After sales service (degree of satisfaction)			
	Number of jobs created in the project			
	Increasing indirect employment			
	Equipment still functioning x years after FCC signed			
	Project sustainability			
	Repayment of loans			
	New supplier representative offices/joint venture/agents, etc.			
	China			

continued on next page

[2] Chinese National Centre for Science and Technology Evaluation (NCSTE) (China) and Policy and Operations Evaluation Department (IOB) (the Netherlands) (2006). *Country-led Joint Evaluation of the ORET/MILIEV Programme in China*. Amsterdam: Aksant Academic Publishers. p.1`5.

[3] Ibid. pp. 36-37.

Case 10-2 continued

Impacts	Overall impacts on end user Additional Dutch exports to China Demonstration and replication effect Poverty alleviation Gender and development Environmental impacts Local economic promotion Future cooperation opportunities
Suggestions	On appraisal by the Dutch side On application procedures on the Dutch/Chinese side On on-lending procedures On lowering prices On 60% content On working capital/matching funds On qualification of Dutch suppliers On bidding On purchasing procedure On evaluation/supervising

Content Analysis

Content analysis is a systematic approach to qualitative data analysis that identifies and summarizes the messages that the data are sending. We usually use the term content analysis to refer to the analysis of such things as books, brochures, written documents, speeches, transcripts, news reports, and visual media.

A typical content analysis might be to examine the content of children's textbooks to see whether they cover the necessary material for learning a particular subject, presented in a way that is reading level appropriate and fits the context in which the children live and study. A deeper analysis might examine whether the textbooks convey a specific political agenda or biased interpretation of history.

Sometimes content analysis is used to when working with narratives such as diaries or journals, or to analyze qualitative responses to open-ended questions on surveys, interviews, or focus groups.

Computer Help for Qualitative Data Analysis

There are software packages to help you organize data derived from individual interviews and focus group interviews. These include text-oriented database managers, word processors, or automatic-indexing software. They are specifically developed for working with text applications. When you enter transcripts into a word processor, they can be organized, indexed, and coded.

Qualitative analysis software can be a powerful tool for organizing vast amounts of data produced through focus groups or individual interviews. These software programs are called Computer-Assisted Qualitative Data Analysis Software (CAQDAS) — also sometimes simply called Qualitative Data Analysis Software (QDAS or QDA software.)

CAQDAS programs search, organize, categorize, and annotate textual and visual data. Programs of this type allow you visualize the relationships between and among data and/or theoretical constructs, and help you to build theories.

Some packages on the market include Ethnograph[4], Qualpro[5], Hyperqual[6], Atlas-ti[7], QSR's N6 (formerly NUD*IST)[8] and other variations.

CAQDAS can help you code data. The following is a list of the strategies to help you: [9]

- **Memos** are the most basic way to annotate your data. Like small electronic stick-up notes, you can attach memos to all sorts of data bits.
- **Free coding** allows you to mark sections of data and attach a code to these sections.

---

[4] Ethonograph website at: http://www.qualisresearch.com/default.htm

[5] Qualpro website at: http://www.qualproinc.com/

[6] Hperqual website at: http://home.satx.rr.com/hyperqual/

[7] Atlas-ti website at: http://www.atlasti.com/

[8] QSR Software, for N^ (formerly NUD*IST) website at: http://www.qsr.com.au/

[9] Loughborough University Department of Social Sciences. *New methods for the analysis of media content. CAQDAS- A primer.* http://www.lboro.ac.uk/research/mmethods/research/software/caqdas_primer.html#what

- **Automatic coding** procedures work in various ways. The most common way is the automatic coding of search results, but other procedures, such as the automatic recoding of data according to previously specified queries, so-called "*supercodes*" in N6 also exist.
- Software generated **coding suggestions** are a novel feature of Qualrus, in which an algorithm suggests codes on the basis of previously occurring codes.
- **Multimedia coding** is offered by N6, HyperRESEARCH and Qualrus. These programs allow you to code sequences of audio or video files and parts of pictures. Some other CAQDAS allow you to link to external multimedia files.

Just as with the search functions, the range of coding functions available varies with the specific CAQDAS.

Summarizing Qualitative Data

Generally, you will report your qualitative data in terms of "common themes" or "a number of people said…." Sometimes, however, there is an isolated idea or perspective that you may want to highlight even though it is not a common theme. There is no single rule to apply, and you will need to tailor your reporting to the specific context and data.

Sometimes it is useful to count the incidence of specific themes to give some sense of how often a particular line of thinking is encountered among respondents. For example, you might want to specify that X% of the news stories had a liberal bias as compared to Y% that had a conservative bias.

If you are using a **mixed method** (qualitative and quantitative) data collection approach, you will want to find themes and comments that help clarify and illuminate some of the qualitative data. For example, if 55% of the respondents were dissatisfied with the accessibility of the intervention, it helps to have a representative mix of comments that help illustrate what kinds of problems people were experiencing.

You will want to capture some of the "quotable quotes." These are the actual statements of the participants and are chosen because they clearly present a theme or an important point you want to emphasize. There is power with these words, so select them carefully. You may find that your audience will be more likely to remember a quote than a page of description. Be careful that you do not introduce bias here. You may want to present several different quotes that show the range of issues and perspectives about the same theme.

Table 10.3 summarizes suggestions for interpreting qualitative data.

Table 10.3: Summary of Suggestions for Interpreting Qualitative Data

Develop Categories	• Use recurrent themes, ideas, words, and phrases. • Use categories that are large enough to capture a range of views but not so large as to be meaningless. • Make categories distinct from each other.
Code the Data	• Develop a coding scheme. • Develop decision rules for coding; they must be exhaustive and unambiguous. • Train your coders to use the coding scheme.
Check for Reliability	• Do a pre-test with a small sample of qualitative data. • Check for inter-rater reliability – do people measuring the same thing, in the same way, get the same results? • If problems exist, fix them, then pre-test again.
Analyze the Data	• Bring order to the data. • Consider placing data on cards. • Consider placing data on a spreadsheet. • Consider using a computer to assist with data analysis. • Sort the data to reveal patterns and themes.
Interpret the Data	• Teams of at least two people, when possible, should review and categorize the material. They should: - compare their findings - if findings are different, review and revise. • Look for meaning and significance in the data. • Link themes and categories to processes of the program and/or to the outcomes. - are some themes more prevalent with discussing process issues? - are some themes more relevant when discussing outcome issues? - look for alternative explanations and other ways of understanding the data.
Share and Review	• Share information early and often with key informants. • Have others review early drafts with the intention of obtaining information, questions, other ways of interpreting the data, and other possible sources of data.
Write the Report	• Describe major themes. • Highlight interesting perspectives even if only said by one or two people. • Stay focused, with so much data; it is easy to get lost. • Include only important information that is. Ask yourself: - does this information answer the evaluation questions? - is this information useful to the stakeholders?

Case 10.2 – an experiment in decentralizing schools in Nicaragua – is an example of an evaluation that used primarily qualitative methodology. The evaluation collected the views of teachers, parents, and school directors, and one of the themes that emerged was individual school autonomy. This theme was summarized in the evaluation report, which illustrates the use of quotes to bring the narrative to life.

Case 10-2: Presenting Qualitative Results:

An Experiment to Decentralize Schools in Nicaragua: Views of Parents, Teachers and Directors

Among the many issues explored in this study, one was individual school autonomy. An excerpt from the report based on interviews and focus groups follows.

> The director has a participatory view of how the school should operate: "This style of management… makes the school more open to the community and more attractive to parents." For the director, individual school autonomy gives her more flexibility to involve teachers and parents in decision making.
>
> At the same time, the director talks as if she is following the desires of the parish priest in detailing her optimistic feelings about decentralization: "It is an improvement. We know the Ministry is trying to do the best thing. Nicaragua is catching up with the developed countries. It is costly, but we must walk down this path."
>
> Critical readers might claim this director is simply putting a positive spin on decentralization, or following a careful script enacted for visitors from Managua. This may in fact be the case.. Yet members of teacher and parent focus groups, with some contradictions, articulated similar meanings and understandings of what autonomy means and implied actions. One teacher, when asked what autonomy meant to her, said: "With autonomy, a teacher needs to be more accountable, to be on time, to plan her classes better, to teach better." Another teacher agrees with the director that changes linked to autonomy will be more social than economic in character: "Autonomy means becoming more independent of the Ministry in how to manage the school, with [resources] from the parents and the community. But here the contributions of parents are very low, they are poor."
>
> Turning to parents, one mother in our focus group stated: "Autonomy means to share responsibility with the Ministry of Education." …. Another parent argued in a positive tone; "To be able to make our own decisions. Not be told this or that."

Controlling for Bias

There is some risk of bias in working with qualitative data (if not using software) in particular; we often see what we want to see and genuinely miss things that do not conform to our expectations. It helps (but does not always completely remedy the situation) to have another person analyze the data. By comparing the two analyses, new themes or different ways of understanding the data may emerge. When reporting qualitative data, you sometimes will not be able to present a count of how many or what percent said or did something. Since all participants were not asked the same question, you really do not know how everyone felt about that question.

When conducting a content analysis, for example, evaluators review documents and code them in terms of themes. The coders must be trained. Having two people read and code the same set of documents helps better control for individual differences in perceptions. If your evaluators are well trained and the operational definitions and rating systems are clear and agreed upon in advance, both evaluators would have a high rate of agreement in their ratings of the material. Analysts call this **inter-rater reliability** and a high rate would be an indicator of credibility. A low rate of agreement between raters indicates that you need to revise your operational definitions and/or rating systems.

Affinity Diagram Process

If you are working with several people on the evaluation project, you may find it helpful to have each person identify what they believe are the common themes or interesting points they would like to report. An affinity diagram is a good strategy to use here (see Table 10.4). Have people write their ideas on file cards or sticky notes (one idea per card/note, but as many cards per person as they think is important), and than have everyone place their cards or notes on a wall. As a group, you can then sort them into similar ideas and themes. By using this process, everyone's ideas are considered and there is less ownership of a single idea. It is also a very quick way to develop an organizing structure for the analysis and final report.

Table 10.4: Affinity Diagram Process

Step	Process
1.	In silence, have all team members identify ideas or themes that they observed and consider interesting or important, and write them on a card or sticky note – only one idea or theme per card, but as many ideas as a member considers important.
2.	In silence, the team members place their ideas on a wall.
3.	In silence, each person on the team begins to sort the cards so that similar ideas are placed together.
4.	Once it appears that the cards have been sorted in terms of general themes, the team discusses the groupings.
5.	The initial groups are not fixed: new groupings can be made and ideas can be shifted.
6.	The team discusses names for each of the large themes.

Challenges to Qualitative Data Analysis

Qualitative data analysis can be time consuming. It may also be difficult to develop a coding scheme.

If you use more that one person to assign codes, reliability between the coders is essential.

Concluding Thoughts on Qualitative Data Analysis

For a number of reasons, many people are afraid of using statistics. As a consequence, there is a strong tendency to think that using qualitative methods is somehow the easier option. But as we have seen in this section, there is a lot more to doing good qualitative data analysis than meets the eye of the casual observer. Analyzing qualitative data is labor intensive and time consuming, but can reveal some of your most valuable information. Be sure to plan enough time to do this well.

As noted in an earlier module, qualitative methods can be powerful tools for looking at causality – whether observed changes are due to the intervention or to something else. An excellent resource with step-by-step guides for how to do systematic qualitative data analysis (descriptive, causal, and other) is *Qualitative Data Analysis: An Expanded Sourcebook, 2nd edition (Miles, M.B. and Huberman, A.M., editors (1994) Thousand Oaks, CA: Sage Publications).*

Analyzing Quantitative Data

Quantitative data are analyzed using statistics. This section will introduce you to some of the most important statistical concepts you need to know as a user and conductor of development evaluations.

Statisticians divide statistics into two large categories:

- **descriptive statistics**, which (in its narrowest definition) is typically used to analyze census or non-random sample data by summarizing data collected from a sample about a qualitative or qualitative variable.
- **inferential statistics**, which is typically used to analyze random sample data by predicting, a range of population values for a quantitative or qualitative variable, based on information for that variable from the random sample. Part of the prediction includes a reliability statement which states the probability that the true population value lies within the range of values.

While there are some data analysis techniques that are used only with inferential statistics, many can be used with both kinds of data. This overview will start with the most common data analysis techniques used for descriptive data, and then it will focus on commonly used data analysis techniques for data obtained using random samples.

Elements of Descriptive Statistics

Distribution statistics describe the frequency and/or percentage distribution of a single variable within the sample (gender, for example, or marital status, or citizenship). It tells you how many and what percent, as in: 33% of the respondents are male and 67% are female (see Table 10.5).

Table 10.5: Distribution of Respondents by Gender

Male		Female		Total
Number	Percent	Number	Percent	Number
100	33%	200	67%	300

Source: Fabricated Data

There are two kinds of measures used to summarize distribution:

- If you want to know how similar the characteristics are, use a **measure of central tendency**.
- If you want to know how different the characteristics are, use a **measure of dispersion**.

Measures of Central Tendency, the 3-M's

The three measures of central tendency that are most often used are sometimes called the three M's:

Mode: Most frequent response.

Median: Mid-point or middle value in a distribution; half the values are larger, half are smaller. Note that in even-numbered data sets, there will be no identifiable single case that represents the mid-point. In such situations, the median is defined as the average of the two middle cases (the sum of the two middle cases divided by 2.)

Mean: Average – the sum of all collected values divided by the number of values collected (sample size), calculated in the following way: $\text{mean} = \Sigma(X_i) \div n$

The Mean and the Median[10]

The two most commonly used statistics are the mean and the median.

Examples:

Table 10.6 gives random samples of six countries, selected from a larger population. Data are presented for two variables (percent of the population living in cities and geographic region):

Table 10.6: Urban Percent Populations, Sample Data

Country	% Urban
Bolivia	65
Algeria	60
Central Africa Republic	41
Georgia	61
Panama	58
Turkey	75

Source: Fabricated data, 2005 survey.

[10]Gene Swimmer, *Qualitative Data Analysis, Part I.* IPDET Handbook 2005. p. 2-3. Used with the permission of the author.

Suppose we wanted to summarize the information for the urban (quantitative) variable.

> The mean would be (65+60+41+61+58+75) ÷ 6 = 60.
>
> The two middle cases are 60 and 61; therefore the median would be (61 + 60) ÷ 2 = 60.5

Notice that the mean is greatly affected by extreme values in the sample, while the median is not. Suppose the urban percentage for Turkey had been 87 (instead of 75.) The mean would increase to 62, but the median would be unaffected. For this reason, the median is the preferred measure for summarizing variables that are potentially distorted by extremely high or low values. For example, median income usually gives a clearer picture of the center of the income distribution than the mean, because income is spread over a huge range in many countries.) On the other hand, it can be argued that expressing data in terms of the median wastes information (in not being affected by extremes.) For that reason, the sample mean might be considered a better predictor of the center of the population than the sample median.

Types of Data

Which measure of central tendency to use depends on the type of data you have: nominal data, ordinal data, and interval or ratio data.

- **Nominal data** belongs to named categories such as gender (male, female), religion (Buddhist, Christian, Jewish, Muslim), or country of origin (Burma, China, Ethiopia, Kenya, Peru, Zimbabwe).
- **Ordinal data** are data can be placed on a scale that has an order to it but the “distance” between consecutive responses is not necessarily the same. Scales that go from "most important" to "least important," or "strongly agree" to" strongly disagree" are examples of ordinal data. Ordinal data lack a zero point, such as I.Q. scores.
- **Interval/Ratio data** are real numbers. Ratio numbers have a zero point and can be divided and compared to other ratio numbers.

Table 10.7 describes when to use each type of data.

Table 10.7: When to Use Types of Data.

If you have this type of data:	Choose this measure of central tendency:
Nominal Data	mode
Ordinal Data	mode or median
Interval/Ratio	mode, median or mean

For interval/ratio data, the choice will also depend on the distribution itself. If it is a normal distribution, the mean, median, and mode should be very close. The mean would be the best description of central tendency. However, if you have a few very high scores or a few very low scores, the mean will no longer be close to the center. In this situation, the median will be a better descriptor of where the center of the distribution is.

Measures of Dispersion[11]

There are two measures which are commonly used to measure the spread of quantitative variables: the range and the standard deviation.

Range

The **range** is defined as the difference between the highest and lowest value of variable. Using the data in Table 10.6, the range for the percent (%) urban population is 75 – 41 = 34. Although it is simple to calculate, the range is not very revealing, because it is determined exclusively by two observations, with all other cases ignored.

Standard Deviation

The most commonly used measure of dispersion for interval or ratio data is **standard deviation**. Standard deviation is a measure of the spread of the scores on either side of the mean. The more the scores differ from the mean, the larger the standard deviation will be. One way to think of the standard deviation is as a kind of "mean of the mean."

[11] G. Swimmer "Qualitative Data Analysis, Part I," IPDET 2005 course notes, pages 4-5. Used with the permission of the author.

To better understand standard deviation, you will need to understand the **normal distribution**. This is sometimes called the "bell curve" because it resembles a bell shape. In a normal distribution the majority of the data occur in the middle of the range of data. There are fewer and fewer data at either side of the middle. That is, most of the examples in a set of data are close to the mean, while fewer examples tend to one extreme or the other.

For example, you have data about the height of a large group of people. The frequency of each height will probably fall into a normal distribution. Most of the people will be towards the middle (mean) with fewer people being shorter and taller. At the extreme ends there will be the very shortest person and at the other end, the very tallest person. Most people's height will be close to the mean because most people are somewhere in between short and tall.

Figure 10.1 shows a normal distribution. The "y" axis shows the frequency of the data, the x axis shows the value of the data. For example if this were a normal distribution of the height of a population, the "x" axis would show the value (measure) of the height, the "y" axis would show the number of people with this measure.

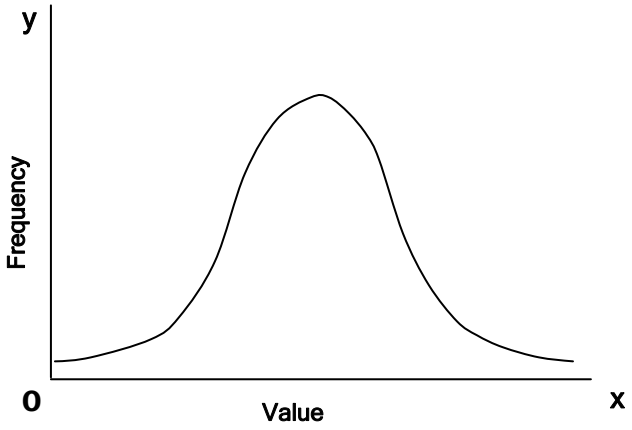


Fig. 10.1: A Normal Distribution.

Not all sets of data will have data that match the normal distribution. Some will have steeper curves; others will have more flat curves, while others will have the curve closer to one end than the other. Figure 10.2 shows examples of some of these differences. But all normally distributed data will look very similar to the normal distribution.

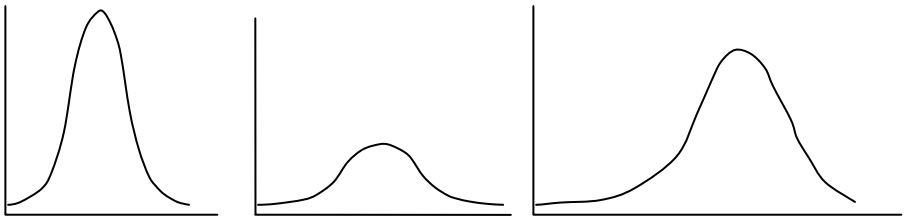


Fig. 10.2: Examples of Data Not Matching the Normal Distribution.

The standard deviation is a statistic that measures how closely all the data in a set of data are clustered around the mean. When the data in a data set closely match the normal distribution, the standard deviation is small. When the data in a data set are spread differently than the normal distribution, the standard deviation gets larger.

The computation of standard deviation can be complicated, but you can understand it you look at it visually.

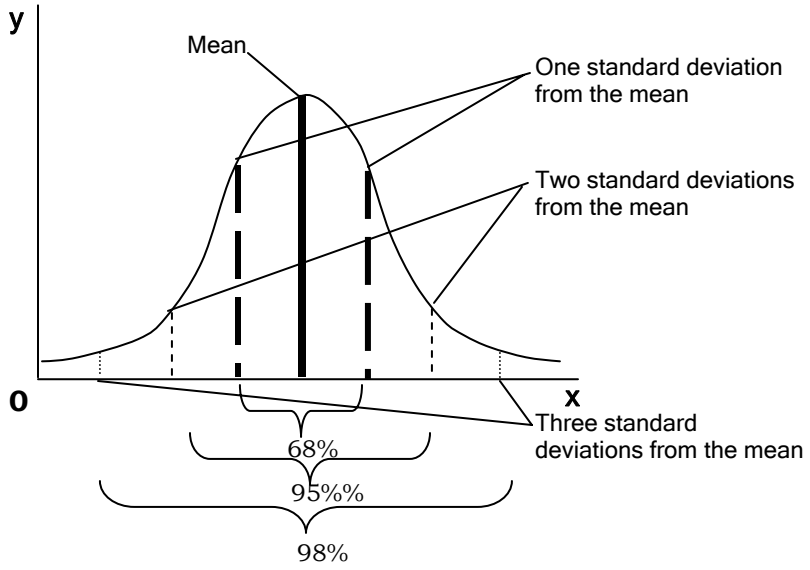


Fig. 10.3: Standard Deviation.

One standard deviation away from the mean in either direction on the horizontal axis (between the bold dashed lines) accounts for somewhere around 68 percent of the people in this group. Two standard deviations away from the mean (between the lighter dashed lines) account for roughly 95 percent of the people. And three standard deviations (between the dotted lines) account for about 99 percent of the people.

If the curve from a data set is flatter and more spread out, the standard deviation would be larger in order to account for those 68 percent or so of the people. The value of standard deviation tells you how spread out the examples in a set are from the mean.

If everyone scored 75 on a test, the mean would be 75 and the standard deviation would be 0. If everyone scores between 70-80 (also giving a mean of 75), the standard deviation would be smaller than if everyone scored between 40-90 (still a mean of 75). Put another way:

Small standard deviation = not much dispersion.

Large standard deviation = lots of dispersion.

Standard deviation is superior to range, because it allows every case to have an impact on its value. It is based on the idea that you will try to measure the average distance of the measurements from the mean of the variable.

Even with a small sample, calculating the standard deviation is time consuming. Thankfully, most statistical programs, including SPSS for Windows, and Excel can do the calculations for you.

Analyzing Quantitative Survey Results

Non-open-ended survey results can be reported in terms of percent answering (52% women, 48% men, for example). Sometimes, the questions ask for specific counts ("Were you employed in the past week?" or "How many goats do you own?" These would also be reported in terms of percents, absolute numbers, or both.

Other times, people are asked to give opinions along a scale. For example, one may ask whether the respondents have been able to apply what they have learned, and are given a five point scale ranging from "not at all" to "a lot." When analyzing this type of data, you would want to establish a decision rule: will you focus on the percent who answered at the extreme ends of the scale, or will you focus on those who answered on either side of the middle category? Some guidelines might be helpful but there are no firm rules here (see Table 10.8).

Table 10.8: Guidelines for Analyzing Quantitative Survey Data

	Guideline
1.	Choose a standard way to analyze the data and apply it consistently.
2.	Do not combine the middle category with categories at each end of the scale: focus on only one category.
3.	Do not report an "agree" or "disagree" category without also reporting the "strongly agree" or "strongly disagree" category.
4.	Analyze and report percentages (or numbers).
5.	Provide the number of respondents as a point of reference.
6.	If there is little difference your results, raise the benchmark: what do the results look like when you focus on the questions that received a majority saying "very satisfied" or "strongly disagree"?
7.	Remember that data analysis is an art and a skill; it gets easier with training and practice.

To gain further understanding, let us look at an example: a survey of clients of a health center. The data are shown in Table 10.9.

Table10.9: Client Views on Health Care Services at the Local Clinic

1. Considering your experiences with the local health clinic, do you agree or disagree with the following statements?

	Strongly disagree	Disagree	Neither	Agree	Strongly agree
I wait a long time before being seen.	10%	20%	10%	35%	25%
They are willing to answer my questions	5	10	30	30	25
I receive good health care at the clinic	15	25	10	25	25

N=36
Source: Fabricated Data, 1999 Survey.

One way to analyze these data is to report that half the respondents agree or strongly agree that they receive good health care and 55% agree or strongly agree that clinic staff is willing to answer questions. However, 60% agree or strongly agree that they wait a long time before being seen. In this analysis, the decision rule was to report the combined percentages of agree and strongly agree.

If the results were different, one might use a different strategy. For example, consider if the results looked like those in Table 10.10.

Table 10.10: Client Views on Health Care Services at the Local Clinic

1. Considering your experiences with the local health clinic, do you agree or disagree with the following statements?

	Strongly disagree	Disagree	Neither	Agree	Strongly agree
I wait a long time before being seen.	50%	20%	10%	15%	5%
The staff are willing to answer my questions	0	5	0	30	65
I receive good health care at the clinic	0	20	0	55	25

N=36
Source: Fabricated Data, 1999 Survey.

The analysis in this case might read that 80% (combination of agree and strongly agree) of the respondents agree or strongly agree that they receive good health care and that staff is willing to answer questions. The greatest strength appears to be the willingness to answer questions, with 65% reporting they strongly agree.

Commonly Used Descriptive Statistics

A range of statistics can be used. Some of the more frequently used statistics are listed in Table 10.11.

Table 10.11: Commonly Used Descriptive Statistics (with illustrative examples using fake data from a new university)

Descriptive Statistic	Example
Frequencies (numbers, a count of how many).	50 students graduated from the university last year.
Percent (proportion) distributions	20 % of the students at the university are women.
Mean (average)	the average age of the students was 25 years of age
Median (mid-point)	ages ranged from 18 to 40, with the mid-point at 24 years of age
Mode (the most frequent value)	the most frequently reported age was 22.
Money: costs, revenues, expenditures, etc., total or average amount	• Costs of running the program increased by $100K (or 50% over the past 5 years) • The average cost of treating a patient fell by $22, or 8%
Percent change over two points in time (sometimes called rate of change)	the university increased its enrollment by 40% over the past year.
Ratio: the number of students per faculty member.	the student/faculty ratio is 15:1
Comparisons: could be numbers, percents, means	• the average salary for graduates from the new university was 20% higher than they had been receiving in their previous jobs • 90% of the employers of the university graduates report being very satisfied with their employees compared to 75% for those who employed graduates from other universities.

Which One to Use?

Figure 10.4 lists some of the most frequently used descriptive analyses, what information they provide, and some examples of when they might be used.

Frequently Used Descriptive Analyses

Frequency Distributions: Number and Percent

Describing Parts of a Whole (100%)

- Percent: parts of a whole expressed as a percent, for example 75%
- Proportion: parts of a whole, expressed as a decimal, not as a percent: .75, for example

Rates: number of occurrences that are standardized; allows for comparison, for example:

- deaths of infants/1000,000 births
- crop yield per acre

Ratio: another way to show the relationship between two numerical variables; shows relative proportions, for example:

- student to teacher ratio is 15:1

Rates of Change or Percentage Change:

- shows change over time, for comparing two items
- [(new time - older Time) / (older time)] x 100 gives percent rate of change, for example:
  - the rate of change from 1980 to 1985 is calculated:

    = [(12,000 - 10,000) / 10,000] x 100

    = 20%

Rates of Change from Prior Year:

Year	Acres Made Available	Rate of Change between measures
1980	10,000	Baseline
1985	12,000	20%
1990	19,000	58%
1995	28,000	47%

Therefore, acres made available increased 20% from 1980 to 1985.

Take care when using the lower and upper case “n/N” with statistics.

- n = an indefinite number
- N = a set of ranked data (ordinal scale) “

Fig. 10.4: Frequently Used Descriptive Analyses

Describing Two Variables at the Same Time

Sometimes you want to describe two variables at the same time. For example, suppose you want to describe the composition of the hands-on and lecture classes. For each class, you want to know what percent were boys and what percent were girls. Analysis of the data shows that the hands-on classes consist of 55% boys and 45% girls, while the traditional lecture classes consist of 55% girls and 45% boys (see Table 10.12).

Table 10.12: Class Participants – Demographics

	Hands-on Classes Number	Hands-on Classes Percent	Traditional Lecture Classes Number	Tradition Lecture Classes Percent
Boys	28	55%	34	45%
Girls	22	45%	41	55%
Total 125	N=50	100%	N=75	100%

Source: Fabricated Data, 2004 survey.

A cross tabulation (often abbreviated as **crosstab**) displays the joint distribution of two or more variables. Crosstabs are usually presented as a contingency table in a matrix format. Whereas a frequency distribution provides the distribution of one variable, a contingency table describes the distribution of two or more variables *simultaneously*. Each cell shows the number of respondents that gave a specific combination of responses, that is, each cell contains a single cross tabulation. Crosstabs are used when you are working with nominal and ordinal data, or when you have categorized interval/ratio data.

Table 10.13 shows an example of crosstab results applied to the data in the previous table, above.

Table 10.13: Crosstab Results (based on Table 10.11)

	Hands on Classes	Traditional Classes	Total %
Boys	45%	55%	100%
Girls	35%	65%	100%

N=125 Source: Fabricated data, 2004 survey.

The following might be an interpretation of the above crosstab results: in this sample, boys are somewhat more likely (45%) to take the hands-on classes as compared to girls (35%).

There appears to be some relationship but how strong is it?

When you look at **measures of association** you need to also understand the concepts of independent and dependant variables.

Independent variables are variables which you believe explain a change in the dependent variable. For example, in a program evaluation, the variables in the program itself are the independent variables.

Dependent variables are the variables you want to explain. For example, in a program evaluation, the dependent variables are the outcome measures.

You can also compare the means of data. For example, if you were looking to determine if South African women have a lower per capita income then men you could compare the mean income of women to the mean income of men, as shown in Table 10.14.

Table 10.14: Comparison of Mean Incomes by Gender

	mean income
Women	27,800 South African Rand
Men	32,400 South African Rand

N=1149 Source: Fabricated data, 2002 survey.

In this example[12], the dependent variable is annual income, and the independent variable is gender

In many project evaluations, we are interested in whether there is a difference in the average values of a quantitative or qualitative variable for a pair of samples. For example, you might be interested in these questions:

- As a result of an irrigation project, are the average crop yields higher than before the project (quantitative variable)?
- Is there a difference in the proportion of patients who are dissatisfied with their care, comparing an existing hospital and a new hospital built under a development project?

[12] Swimmer, Gene. Quantitative Data Analysis, Part I," IPDET 2005 course notes, p. 22. This example and the text following used with the permission of the author.

The big question is whether the difference indicates a difference in the population means (or proportions) or did the result occur by chance variation in the samples taken from the two populations? In this statistical test, it is commonly assumed that there is no difference between the two population means (or proportions.) This issue will be addressed in the section dealing with inferential statistics.

Measures of Relationship

Measures of relationship (or association) tell you how strongly variables are related. Association never proves cause, but it can suggest a causal relationship if (and only if) there is a strong measure of association.

While there are many kinds of measures of association, they are usually reported in terms of a zero to 1 scale to indicate the strength of the relationship. A perfect relationship would score 1. A relationship showing no association at all would score zero. In other words, the closer the measure is to zero, the weaker the relationship and the closer the measure is to 1, the stronger the relationship.

Measures of correlation are calculated on a scale of -1 to +1. These measures show the direction of the relationship through the sign (positive or negative). A measure with a positive sign means that as the variables change in the same direction: both go up or both go down. This is called a **direct relationship**: for example, as the years of education increases, individual income increases. A negative sign indicates that the variables have an **inverse relationship** meaning that they move in the opposite directions: for example, as age increases, health decreases. A measure of association of -1 would therefore mean a perfect inverse relationship. A measure of association of -.1 would be close to zero and is a very weak inverse relationship.

You can do a correlation between two interval or ratio variables or use a multiple regression technique to estimate the impact of several variables simultaneously on the dependent variable. This technique works with interval and ratio level data. There are other techniques that can be used for ordinal and nominal data but they are less commonly used.

According to Swimmer[13],

> ...most social science research focuses on the relationship between two or more variables. In fact, many social science theories are based on the idea that changes in one variable cause the other variable to change in a similar or opposite direction, for example: increases in nutrition cause infant mortality rates to fall and increases in exports lead to higher levels of economic growth. This is important for evaluation research, because a development project is sometimes aimed at changing a variable which will in turn generate a benefit by changing another factor – a project aimed at increasing literacy rates in order to reduce the spread of AIDS, for example. Although statistical analysis can never prove causality (i.e., that nutrition *causes* a decline in infant mortality), it can determine whether the data supports the theory: that is, countries with better nutrition usually have lower infant mortality rates.

Inferential Statistics

Inferential statistics enable you to make an estimate about a population based on a **random sample** selected from that population. Whenever you are using sample data, your major concern is whether the results are a function of some quirkiness of your sample rather than a true picture of the population. If you had picked a different sample, would your results be fairly similar?

Statisticians have developed tests to estimate this. These are called **statistical significance tests** and do a very simple thing:

> *Statistical significance tests allow you to estimate how likely it is that you have gotten the results you see in your analysis by chance alone.*

[13] Swimmer, G. "Qualitative Data Analysis, Part I;" IPDET 2005 course notes, p. 29.

Statistical tests come in 100+ varieties. You may have heard of some of the more common statistical tests, such as Chi Square and the t-test. The good news is that all the different statistical tests are interpreted using the same guidelines. Evaluators typically set the benchmark for statistical significance at the .05 level. This is sometimes called the alpha level or the p value (for probability of error). That is, we set the benchmark so that we are at least 95% certain that the sample results are not the result of random chance. If we want to raise the bar, we would set the level at .01 to be 99% certain that the sample results are not due to chance alone.

All tests of statistical significance are partly based on sample size. If the sample is very large, small differences are likely to be statistically significant. You still need to decide whether the differences are important, given the nature of your research. Importance is always a judgment call.

Chi Square

Chi-square is one of the most popular statistical tests because it is easy to calculate and interpret. The purpose of chi square is to determine whether the observed frequencies (counts) markedly differ from the frequencies that we would expect by chance. Chi square is used when you want to compare two nominal values (for example, marital status and religious affiliation). Or you want to compare two ordinal variables (scaled responses) or a combination of nominal and ordinal variables.

The chi square statistic is the sum of the contributions from each of the individual cells in a data table. Every cell in the table contributes something to the overall chi square statistic. If a given cell differs markedly from the expected frequency, then the contribution of that cell to the overall chi square is large. If a cell is close to the expected frequency for that cell, then the contribution of that cell to the overall chi square is low. A large chi square statistic indicates that somewhere in the table, the observed frequencies differ markedly from the expected frequencies. It does not tell which cell (or cells) is causing the high chi-square, only that they are there.

Chi square measures whether two variables are independent of one another based on observed data.

t-Test

A t-Test is a statistical technique that can determine whether one group of numerical scores is statistically higher or lower than another group of scores.

This analysis is appropriate whenever you want to compare the means of two groups. It is especially appropriate for the analysis of a project where you want to compare the mean scores of the group affected by the project with the mean scores for the control group (unaffected by the project). This leads us to a very important conclusion: when we are looking at the differences between scores for two groups, we have to judge the difference between their means *relative* to the spread or variability of their scores. The t-test does just this.

The Logic of Statistical Significance Testing

These tests are set up to measure the probability of getting the same results if there really was no difference in the population as a whole. Researchers call this the **null hypothesis,** and it is always based on zero difference in the population...

Suppose a survey based on a random sample of people in Pakistan shows that there was a 5,000 rupee difference in annual income between men and women. Our test might be expressed in this way: if there really is no difference in the population, what is the probability of finding a 5,000 rupee difference in income between the men and women in a random sample? If there is a 5% chance (.05) or less (that's our benchmark), then we will conclude that the sample results are an accurate estimate of the population.[14] We would conclude that there is indeed a difference of about 5,000 rupees, and that difference is **statistically significant**.

Most reports do not go beyond a benchmark of .05 or 5%. This means that we are 95% certain that our sample results are *not* due to chance or that the results are statistically significant at the .05 level.

[14] Technically, social scientists would say that "the null hypothesis of no difference is rejected". If, on the other hand, the benchmark for statistical significance is .01 (which is less than .05), social scientists would say that the test failed to reject the null hypothesis: these results are not statistically significant.

Table 10.15 shows common tests for statistical significance and what kind of data they test.

Table 10.15: Common Tests for Statistical Significance

Statistical Test	Kind of Data Needed
chi square	a test using nominal and ordinal data
t-test	dependent variable, ratio data
	independent variables, 2 categories

Linking Qualitative Data and Quantitative Data

Miles and Huberman discuss how qualitative and quantitative data may be linked in a study[15]. They begin their discussion with a quote from Fred Kerlinger, a highly regarded quantitative researcher: "There's no such thing as qualitative data. Everything is either 1 or 0." They then offer an opposing view: that all data are basically qualitative. They also discuss the concept that all research has a qualitative grounding.

The argument defending quantitative or qualitative data has been going on for some time and will probably continue long into the future. In development evaluation, you will need *both* quantitative and qualitative data to understand the world; that is . "…quantities are *of* qualities, and a measured quality *has* just the magnitude expressed in its measure".[16]

Miles and Huberman ask the question "should the two sorts of data and associated methods be linked during study design, and, if so, how can it be done, and, for what purposes?" These researchers identify others' views on these questions. The following is a summary of information from other authors on this topic.

Linking two sorts of data allows the evaluator to:

- enable confirmation or corroboration of each other via triangulation
- elaborate or develop analysis, providing richer detail
- to initiate new lines of thinking through attention to surprises or paradoxes, "turning ideas around" and providing fresh insight. (Rossman and Wilson) [17]:

---

[15] Matthew B. Miles and A. Michael Huberman (1994), *Qualitative data analysis: An expanded sourcebook*, 2nd Edition (M.B. Miles and A.M. Huberman, editors) Thousand Oaks, CA: Sage Publications pp. 40 – 43.

[16] Ibid.

[17] G. B. Rossman and B. L. Wilson (1994). Numbers and words revisited: Being "shamelessly methodologically eclectic." *Quality and Quantity, 28,* 315–327.

Green, Caracelli and Graham[18] note that:

- such studies can help in sequential testing: that is, the results of the first method inform the second's sampling, instrumentation, etc.
- such studies can expand the scope and breadth of a study by using different methods in different components.

Firestone[19] argues that:

- on the one hand, quantitative studies "persuade" the reader through de-emphasizing individual judgment and stressing the use of established procedures. This can lead to more precise results that can be more easily generalized
- on the other hand, qualitative research persuades through rich depiction and strategic comparison across cases. In this way, it overcomes the "abstraction inherent in quantitative studies."

Sieber[20] offers a detailed list of reasons why linking data are useful, including:

- during the *design phase*: quantitative data can help with the qualitative side of a study. It does this by finding a representative sample and locating deviant cases
- during the *data collection phase*: linking data will supply background data, ensure that information is not overlooked, and help to avoid "elite bias"
- during *analysis*: quantitative data can help by showing the generality of specific observations, correcting the "holistic fallacy" and verifying or casting new light on qualitative findings.

---

[18] J.C. Greene, V.J. Caracelli, and W.F. Graham, (1989). "Prevention effectiveness: A guide to decision designs." *Educational evaluation and policy analysis. 11:*255-274

[19] W. Firestone (1987). Meaning in method. The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*, 16-21.

[20] S. D. Sieber (1973). The integration of fieldwork and survey methods. *American Journal of Sociology, 78*(6), 1335-1359.

Miles and Huberman[21] see qualitative-quantitative linkage at three levels. They are:

- **First level**: the "quantifying" level, where qualitative information can be either counted directly or converted into ranks or scales
- **Second level**: where qualitative information is compared to numerical data, linking the two distinct data types
- **Third level**: where the overall study design is specified, such as the multimethod approaches or more complex ones are identified; all may involve combinations of case studies, surveys, experiments, and unobtrusive-measure studies.

Let us look at examples at each level.

First level (Quantifying level)

Morse[22] cites a study of teenage mothers and the frequency of their use of the word *stuff*, as in "We spent a lot of time together and stuff." Morse calls this "appurtenant counting". The word occurred more than 100 times in 16 open-ended interviews. She suggests that these adolescents were using a "half-child mode of speech" as they coped with adult responsibilities.

In another study, this one on school improvement, by Miles and Huberman, they became aware of the importance of job mobility and looked at people who had changed jobs. They found it was helpful to know, first, how many people had moved (74 for 12 sites). They wanted to know more about why people had moved: how many of them had moved because of their experience with the innovation (83%), and how many were actually upward moves (35%)?

---

[21] Matthew B. Miles and A. Michael Huberman (1994), *Qualitative data analysis: An expanded sourcebook*, 2nd Edition (M.B. Miles and A.M. Huberman, editors) Thousand Oaks, CA: Sage Publications. pp. 40 – 43.

[22] Janice M. Morse and Lyn Richards (2002). "The integrity of qualitative research " In *Read me first for a a user's guide to qualitative methods* by J.M Morse and L. Richards. Thousand Oaks, CA: Sage Publications. P 23-41

Miles and Huberman[23] also converted some interview data into rating scales:

- the degree of pressure teachers had felt to adopt an innovation
- their satisfaction with the assistance they had received
- the ease or difficulty of the implementation process.

Because their cases were schools, not individuals, this conversion involved examining interviews from different people, checking degree of agreement, and arriving at a site-level rating.

They determined that three- to five-point scales seemed easiest and most reliable. In the displays and analysis, they aimed to keep "numbers" like these closely associated with the words from which they drew the judgment, and to keep the words associated with the context.

Linking data types

In a study by Kell[24], researchers looked at the effects of computers on classroom teaching and learning. The research team repeatedly visited classrooms in six school districts over a school year. During the visits, they interviewed teachers and administrators and systematically observed classroom activity. They also asked teachers to complete two different standardized questionnaires at three times during the year. These questionnaires covered not only the teachers' ideas about using computers to enhance the learning of reading skills, but also their concerns about the adoption and use of such innovations. In this way, the two data types were linked, and allowed researchers to study changes occurring over the whole school year.

[23] Matthew B. Miles and A. Michael Huberman (1994), *Qualitative data analysis: An expanded sourcebook*, 2nd Edition (M.B. Miles and A.M. Huberman, editors) Thousand Oaks, CA: Sage Publications. pp. 40 – 43.

[24] Unavailable at time of print.

Mutimethod designs'

In a study of care for the elderly, Wilson and Bolland[25] tested competing theoretical models. They collected data on referrals and goal-oriented actions for 49 organizations and then built a nominated sample of community leaders. They analyzed the referral and goal oriented data to determine which organizations were most central in the network of organizations. They examined the leader data to see which leaders were most central in their network. They found the combined analysis supported one of the theoretical models. They could not have done this with either one or the other data set alone.

[25] Unavailable at time of print.

Hints for Development Evaluators

If you are analyzing qualitative data consider the following:

- start your analysis as you collect the data by making good notes
- draw out themes and patterns
- summarize your data
- control for bias
- if working with several people, consider using affinity diagrams.
- If you are working with quantitative data you will need to use statistics to help you analyze your data. You will use one or more of the following:
- measures of central tendency (how similar) – mode, median, and mean
- measures of dispersion (how different) – standard deviation and variance.

To analyze the data, look at it and use these guidelines:

- choose a standard way to analyze the data and then apply it consistently
- focus on one side of the scale or the other when analyzing ordinal data
- do not combine the middle category with either ends of the scale with ordinal data
- when reporting on one category out of the five, choose an extreme category (strongly agree or strongly disagree)
- do not report an agree (or disagree) category without also reporting the strongly agree or strongly disagree category
- analyze and report percentages (not numbers, and definitely not averages)
- provide the number of respondents
- if there is a little difference in the data, raise the benchmark: what do the results look like when you focus on the questions that received a majority response of “very satisfied”?
- remember that analyzing data is an art and a skill; it gets easier with training and practice.

Hints for Development Evaluation Managers

Mgr

If the evaluation is using qualitative data, check with the evaluator to make sure that the data are answering the important questions:

- Is the intervention being implemented according to plan?
- What are some of the difficulties faced by staff?
- Why did some participants drop out early?
- What is the experience like for participants?
- Is there any unexpected impact on families and communities?

If the evaluation is using quantitative data, check with the evaluator to answer the following questions:

- What is the percent distribution?
- What's the average?
- How do participants rate the usefulness and relevance of the intervention?
- How much variability is there in the data?
- What's the relationship between a program and the outcome measures?
- How strong is the relationship?
- Are the results statistically significant?

Summary

In this module, you learned about how to analyze qualitative and quantitative data. Use the following checklist to check if you have learned the important items. Check those that you know and review those that you do not.

- ❑ describe the guidelines to determine the size of sample that is needed
- ❑ discuss the when to use qualitative and quantitative data
- ❑ discuss how to analyze qualitative data. including:
  - making good notes
  - drawing out themes and patterns
  - content analysis
  - summarizing qualitative data
  - controlling for bias
  - affinity diagrams
- ❑ discuss how to analyze quantitative data
- ❑ discuss measures of dispersion including standard deviation
- ❑ discuss ways to analyze survey results
- ❑ define commonly used descriptive statistics, including:
  - frequency
  - percent
  - mean
  - median
  - mode
  - money vs. non-money costs
  - percent change over two points in time
  - ratio
  - comparisons
- ❑ discuss measures of relationship
- ❑ discuss inferential statistics

Quiz Yourself

Answer the following multiple-choice questions to help test your knowledge of data analysis and interpretation.

You will find the answers to the questions on the last page of this module.

1. Which of the following is the definition of **content analysis**?
   a. a method of summarizing qualitative data
   b. a systematic approach to qualitative data analysis that identifies and summarizes message content.
   c. a systematic approach to quantitative data analysis
   d. a systematic method of taking notes after an interview
2. What is the **affinity diagram** process?
   a. a process of sorting the ideas of a group in order to organize them
   b. a process used to code make good notes
   c. a process used to decide whether to use qualitative or quantitative data
   d. what evaluators use to control for bias
3. Which of the following is the definition of a **mean**?
   a. the numerical average of the responses
   b. the mid-point or half way point of the data
   c. the most frequent response
   d. a measure of the dispersion of data
4. Which of the following is the definition of a **median**?
   a. the numerical average of the responses
   b. the mid-point or half way point of the data
   c. the most frequent response
   d. a measure of the dispersion of data
5. 3. Which of the following is the definition of a **mode**?
   a. the numerical average of the responses
   b. the mid-point or half way point of the data
   c. the most frequent response
   d. a measure of the dispersion of data

6. Which of the following is the definition of **standard deviation**?
   a. the numerical average of the responses
   b. the number, a count of how many
   c. the most frequent response
   d. a measure of the dispersion of data
7. Which of the following is the definition of **frequency**?
   a. the number of, a count of how many
   b. the mid-point or half way point of the data
   c. the most frequent response
   d. a measure of the dispersion of data
8. Which of the following is the reason for testing for **statistical significance**?
   a. the test gives you the probability of getting the results in the population as a whole
   b. it tests for the null hypothesis
   c. the test allows you to estimate how likely it is that you have gotten the results you see in your analysis by chance alone
   d. it tests for how strongly variables are related

Reflection

Think back about previous evaluations with which you have been involved.

- What kinds of data analyses have you used in the past: were they qualitative, quantitative, and/or combination? How did you feel about the analysis you did? How would you do it differently now that you know more about data analysis and interpretation?
- How do you think using descriptive statistics will help you tell your audience more about the results of your evaluation?
- What are your feelings about statistics? Does it make you nervous? If so, try to analyze your fear. Think about how statistics help you communicate to others.

Application Exercise 10-1:

Affinity Diagram Process

Instructions

Obtain three or four fairly substantial newspaper articles about an issue in the area where you work. Read each of the articles to determine the major themes.

If you are working with others on this module, have them do the same. Then together use the affinity diagram process to identify themes. [If you are working alone, please skip this and go on to Exercise 10-2.]

You will be given pieces of paper. Write one theme or idea on each of the pieces of paper. This is to be done in silence. Once you have written all your themes, stick the pieces of paper randomly on the wall. As others add their pieces of paper to the wall, begin to arrange into similar groups—again in silence.

Once the pieces of paper begin to take shape according to themes, people may talk as they continue to re-arrange the post-its.

The large group will then identify the themes.

Application Exercise 10-2

Qualitative Data Coding and Analysis

Instructions

Obtain three or four fairly substantial newspaper articles about an issue in the area where you work. Read each of the articles to determine the major themes.

If you are working by yourself and have access to a computer, set up an Excel grid (or other spreadsheet program, or even a large, handwritten paper grid if you prefer). In the top left cell, write "article," and at the top of the second column, write the word "excerpt." Next, enter each of the comments from the interviews, one per row, in the column labeled "excerpt."

Identify themes from the articles, enter them as column headings, and mark the appropriate cell where a particular article excerpt contains each theme. Finally, write a narrative summarizing the findings from the three articles.

Application Exercise 10-3

Common Mistakes in Interpreting Quantitative Data

1. A survey asked people to report their perceptions and 80% said it was helpful. Which would be the better way to report the findings:
   a. ____"The program is helpful"
   b. ____"The participants found the program helpful."

2. The respondents were asked to identify the barriers and supports to the program. What is the problem with reporting the results in terms of “pros" and "cons"?

3. A survey asked students to rate various components of the course and most rated each of the components positively. What is the problem of writing that “most (70%) of the students felt the course was a success”?

4. 40% of the women and 30% of the men favored curriculum changes. Is it accurate to report the findings as: “a majority of women favored curriculum changes”?

5. 51% favored changing the curriculum. Is it accurate to say "more than half of the respondents said…”?

6. The survey was completed by 5 of the 20 instructors. All five agreed that they were well prepared. Is it accurate to say, “All of the instructors were well-prepared”?

7. In the same survey, is it accurate to report "Eighty percent of the participants said the materials were organized”?

8. Is it accurate to report a 100% increase of women in political office when the actual number increased from 2 to 4, out of 50 elected positions?

9. A training program found that those who participated in the program earned 20% more than those not in the program. Is it accurate to report, “The program caused a 20% increase in salary”?

Application Exercise 10-4

Analyzing Results from a Questionnaire

Instructions:

Please complete this survey, and have two colleagues who are also working through these training modules also complete it. [If no one else is working through these modules, have colleagues complete the first half only.]

Collect the surveys and tally the results. Working by yourself or with others, **summarize all the results**. Include some conclusion about the *overall* findings.

IPDET Survey ID: _____

1. To what extent, if at all, would you say that you currently had the following ability to do each of the following?

	Little or no extent	Some extent	Moderate extent	Great extent	Very great extent
a. Design an evaluation					
b. Analyze data					
c. Develop a survey					
d. Conduct a focus group					
e. Facilitate a stakeholders' meeting					
f. Write an evaluation report					
g. Prepare an oral briefing					

2. At this point in this training program, would you strongly agree or disagree with each of the following statements:

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
a. The material is new to me					
b. The material is interesting					
c. There is sufficient lecture					
d. There is sufficient class discussion					
e. The exercises are helpful					
f. I am learning material I can use					

Any comments?

Further Reading and Resources:

Denzin, N. & Lincoln, Y. (Eds.) (2000). *Handbook of Qualitative Research*, (2nd Ed.) Thousand Oaks, CA: Sage Publications.

Glass, G. & Hopkins, K. (1996). *Statistical methods in education and psychology* (3rd Ed.). Boston: Allyn and Bacon.

Jaeger, R.M. (1990). *Statistics: A spectator sport* (2nd Ed.). Thousand Oaks, CA: Sage Publications.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Porteous, Nancy L., Sheldrick, B.J., and Stewart, P.J. (1997). Program evaluation tool kit: A blueprint for public health management. Ottawa, Canada: Ottawa-Carleton Health Department. (ISBN 0-9694956-3-3.) Available online at http://www.phac-aspc.gc.ca/php-psp/toolkit.html (English) or http://www.phac-aspc.gc.ca/php-psp/toolkit_fr.html (French)

Websites:

Online texts and tutorials:

Lane, D.M. *Hyperstat online textbook.* http://davidmlane.com/hyperstat/index.html

Porteous, Nancy L., Sheldrick, B.J., and Stewart, P.J. (1997). Program evaluation tool kit: A blueprint for public health management. Ottawa, Canada: Ottawa-Carleton Health Department. Available online at http://www.phac-aspc.gc.ca/php-psp/tookit.html (English) or http://www.phac-aspc.gc.ca/php-psp/toolkit_fr.html (French)

StatSoft, Inc. (2001). *Electronic Statistics Textbook.* Tulsa, OK: StatSoft. http://www.statsoft.com/textbook/stathome.html

Statistics at Square One http://bmj.bmjjournals.com/collections/statsbk/index.shtml

Stat Primer http://www2.sjsu.edu/faculty/gerstman/Stat Primer

Computer Software for Qualitative Data

Atlas-ti website at: http://www.atlasti.com/

Ethonograph website at:
http://www.qualisresearch.com/default.htm

Hperqual website at: http://home.satx.rr.com/hyperqual/

Loughborough University CAQDAS – a Primer
http://www.lboro.ac.uk/research/mmethods/research/software/caqdas_primer.html#what

Overview/Summary of Qualitative Software Programs
http://www.quarc.de/software_overview_table.pdf

QSR Software, for N^ (formerly NUD*IST) website at:
http://www.qsr.com.au/

Qualpro website at: http://www.qualproinc.com/

General Site for Statistics:

Interactive Statistical Calculations Pages
http://members.aol.com/johnp71/javastat.html

Statistics Computer Programs

SPSS (free 30 day download)
http://www.spss.com

OpenStat version 4 (similar to SPSS)
http://www.statpages.org/miller/openstat/

Online tutorials for SPSS

Tutorial for SPSS v. 11.5
http://www.datastep.com/SPSSTraining.html/

Getting Started with SPSS for Windows
http://www.indiana.edu/~statmath/stat/spss/win/

Sites with examples

US Census Bureau, International Programs Center
http://www.census.gov/ipc/www/idbnew.html

Carleton University, Canadian Foreign Policy (Journal), The WWW Virtual Library
http://www.carleton.ca/npsia/cfpj

World Health Organization (WHO)
http://www.who.int/health-systems-performance

WHO Statistical Information System (WHOSIS)
http://www.who.int/topics/statistics/en/

The North-South Institute
http://www.nsi-ins.ca/ensi/research/index.html

United Nations Department of Economic and Social Affairs, Statistics Division
http://unstats.un.org/unsd/databases.htm

United Nations Development Program Human Development Report 2002
http://www.undp.org/hdr2002

United Nations Environmental Programme, Division of Early Warning and Assessment (DEWA)
http://www.grid.unep.ch

United Nations High Commission for Refugees, Statistics and Research/Evaluation
http://www.unhcr.ch/cgi-bin/texis/vtx/home

UNESCO Institute for Statistics
http://www.uis.unesco.org/en/stats/stats0.htm

The United States Agency for International Development
http://www.usaid.gov/educ_training/ged.html

The United States Agency for International Development
http://www.dec.org/partners/eval.cfm

International Monetary Fund
http://www.imf.org/external/pubs/res/index.htm

International Monetary Fund
http://www.imf.org/external/np/sta/index.htm

World Bank Data and Statistics
http://www.worldbank.org/data

World Bank World Development Indicators 2005
http://worldbank.org/data/wdi2005

Organisation for Economic Cooperation and Development Development Indicators
http://www1oecd.org/dac/

International Institute for Sustainable Development Measurement and Indicators for Sustainable Development
http://www.iisd.org/measure/default.htm

Answers to Quiz Yourself

1. b
2. a
3. a
4. b
5. c
6. d
7. a
8. c

To continue on to this
Module's Presentation Slides click here

To continue on to the
Next Module click here

To return to the
Table of Contents click here