

STEP 5: DETERMINE HOW TO UNDERSTAND THE INFORMATION (ANALYZE DATA)

Now that you have collected data, what does it mean? Making sense of this information is arguably one of the hardest parts of a community assessment. However, **acquainting yourself with the process allows you to become a good “user” of data**. This means being able to understand the basics of how data is analyzed, so that you can include it effectively as you plan and conduct your community assessment. Planning for your data analyses:

- Determines if the data you collect is practical for analysis.
- Reviews the appropriateness of your data collection methods.
- Informs the data collection instrument you create. How you ask a question and structure the responses affects how that data is analyzed later.
- Lets you know if you will need help with analysis. Do community partners have the needed skills, knowledge and resources? Do you need to find someone else to do it?

Data analysis may seem like a daunting task, especially if you plan to collect a lot of data. But every data analysis plan involves the same six basic stages, which are outlined here. Discuss these six stages with your partnership during planning. Determine who is responsible for each stage.

⇒ Step 5 has 6 stages, which will be reflected in the *Step 5 Planning Worksheet: Data Analysis Plan* (following page 4-57):

- 5.1 Check the Data
- 5.2 Go Back to the Primary Assessment Questions
- 5.3 Reduce the Amount of Data
- 5.4 Analyze the Data
- 5.5 Verify Findings
- 5.6 Interpret Findings and Draw Conclusions

5.1 Check the Data

First, take a look at all of your compiled data to ensure it's all there and that it all makes sense. **You want complete and quality data**. You are looking for any mistakes people might have made when filling out a survey, or an interviewer or focus group facilitator might have made when asking questions or taking notes. Correct

these common mistakes before data analysis. **Perform quality checks throughout your data collection process.** Start from the very beginning.

Common mistakes include:

- A **missing or incomplete response**. Any missing or incomplete responses will exclude that respondent from your analysis of that particular question. This may change the total number of respondents for that question, so you should be careful of this change when it occurs. You may want to consider completely excluding a questionnaire or qualitative interview with many missing or incomplete responses. If there are many missing or incomplete responses to a particular question, you may ask your data collector or respondents why this occurred. Some common reasons include: a lack of relevant responses, a question that respondents don't know how to answer, or a question that makes many respondents feel uncomfortable and so they don't want to reveal their answer. If you find out that either of these reasons are behind the missing data from a particular question, you may consider eliminating that particular question from any later analyses.*
- An **impossible response** (i.e. age answered as "168"). If it is possible to determine the reason for this data collection mistake, and to accurately fix it, then you should do so. If the mistake is less clear, then you may want to consult with the data collector to see what might have happened. If you are unable to determine why the mistake happened and how to fix it, then you can treat it the same as missing data.
- A **response unrelated to the question** (i.e. age answered as "female", or a qualitative interview question response included under the wrong question). Again, if it is possible to easily fix this data collection mistake, you should. However, this may mean a mistake occurred when asking or answering the questions during data collection, and thus you may want to consult with the data collector. If it is appropriate to contact the respondent, then you may want to clarify this mistake with that person. If neither of these options are available for you, or the lost data isn't critical, then can treat these mistakes the same as missing data.
- A **response that contradicts an earlier response** (i.e. answered as non-smoker in one question and answered as smoking one pack a day in a later question). Again, if appropriate, it may be possible to check with your data collector or respondent to see what might have happened. If this is not possible or critical, then you will have to eliminate those particular responses and treat

* If you are collecting the data yourself, and are still in the middle of data collection, another way to solve this problem is to rephrase the way the question is asked, or the way the responses are given. This allows you to change the question in time to collect the appropriate information before data collection ends.

them like missing data. If you believe these mistakes occur throughout a particular survey or interview, you may want to consider removing this respondent from later analyses.

- An **unclear response**. If a discussion occurs in a focus group that is unclear, you may want to discuss with the moderator how to better understand what is being said at that point in the focus group. If a qualitative interview response is unclear, you can also check with the interviewer to clarify what was said. If you still are unable to understand the response, and if it is appropriate to do so, then you may want to follow up with that respondent and ask him or her to help you better understand the response.

5.2 Go Back to Your Primary Assessment Questions

Before analyzing your data, return to your primary assessment questions and use them as your starting point.

- Be sure to only analyze the data that helps you answer the primary assessment questions.
- This ensures you don't get sidetracked into analyzing any other pieces of data with your limited time and resources.
- It is normal to collect more data than is needed for your community assessment. You can return to any other interesting data after completing your Community Assessment Plan.

⇒ **Here are some examples of how assessment questions can guide your analysis:**

1. If your question asks — *What public health issues concern community members?* You need to pick out all of the concerns mentioned by your survey, interview or focus group respondents. Count the number and percentage of people who identified each concern, then rank the concerns into the order they were most mentioned so you can rank their importance to your community members.
2. If your question asks – *How do residents think about their neighborhoods?* You need to identify the different ways that survey, interview or focus group respondents discussed their neighborhoods. Then you can list them and provide some description and numbers of times each were mentioned.
3. If your question asks – *How are women's perceptions different than men's?* Then you can summarize survey, interview or focus group data separately for women, and then for men. Once you understand the responses from each group, you can compare them to see how similar or different they are from each other.

5.3 Reduce the Amount of Data

Examine your compiled data and figure out ways to reduce its sheer magnitude in order to analyze and interpret it.

- Look at small pieces of the data at a time. Start with one data collection method or health issue, and then move on to another when you are done.
- Don't be afraid to eliminate from your analysis any irrelevant data—data that doesn't answer your primary questions or is unrelated to what you would like to include in your assessment findings.
- Be sure to save your data in its original form so that—if needed—you can return to it later.

At this point it is helpful to create a master spreadsheet or document, which allows you to compile your data so you can look more easily for interesting findings.

1. Numeric Spreadsheet

A spreadsheet is a grid of rows and columns, which can be used to enter and look at your quantitative data all at once. This is sometimes called “eyeballing” your data, and it allows you to look for data patterns, see what comparisons might be interesting to make, and determine what might be important to examine further.

Create a spreadsheet with your questions listed across the top and your respondents listed down the left-hand side. You can indicate each question and respondent with a number, which makes it easier to fit your data into a smaller number of pages, as well as protects the confidentiality of your respondents. Your questions are organized into columns and your respondents organized into rows, then the response to each question can be placed in the row that corresponds with that respondent.

You can total up the number of responses to each question at the bottom of that column. To make it easier to enter your data into the spreadsheet and to calculate the totals, you may create a number corresponding to each response. This is called **coding**. That is, you have created a number code for each question response. You should keep track of what the codes are for each question. Some common codes are:

- “Yes” or “True” = 1
- “No” or “False” = 0
- Multiple choice responses numbered 1 through 5
- Scaled responses (strongly disagree, disagree, neither disagree nor agree, agree, strongly agree) = 1 through 5
- “Don't know” = 8

- No response/ missing data = 9

Mistakes may be made during data entry. In order to avoid any mistakes, make sure one person enters the data from the survey and a second person double-checks the data entry.

See *Appendix B: Computer Software to Compile and Analyze Data* (page 5-63) for instructions and resources on using computer software to create a spreadsheet.

2. Qualitative Summary Document

It is more difficult to compile qualitative data in a way that enables you to “eyeball” it all at once. By its very nature, qualitative data is lengthy. However, you can help yourself in analyzing the data by creating a summary document that contains the gist of what discussion and responses were recorded for each focus group discussion question or qualitative interview question asked.

Create a new document and list each question as a heading. Under each heading, summarize the relevant discussion or responses underneath each. This may include some cutting and pasting from your source documents—your interview and focus group notes—to your summary document. However, you should be sure to keep your source documents intact. This is your original data, and you may find that you need to go back and re-read the context in which the discussion or response occurred.

If you have only a few interviews or short focus group discussions, this summary may be all you need to “eyeball” the data. However, if you have a lot of qualitative data, you may need to do some further summary of the data in order to be helpful to your analysis. You can do this by summarizing the discussions or responses into concise bullet points under each of the question headings. You can always refer to the source documents to get additional information.

If you think it might be helpful, you can also create a table of your questions and summarized discussion points for easier viewing. This is an excellent way to organize qualitative data so you can look for relationships between discussion points, and across questions.

5.4 Analyze the Data

Once you have accomplished phases 1-3, the specific techniques for analyzing the data look very different for qualitative and quantitative data. What follows is a description and an exercise for each data analysis technique. For quantitative data, the technique is **Looking for Patterns in the Data**, and for qualitative data the technique is **Looking for Themes in the Data**.

⇒ Quantitative Data Analysis — Looking for Patterns in the Data

Analysis of quantitative data usually includes the following numeric calculations and comparisons. For information on how to conduct higher-level statistical analyses, please see *Appendix E: Resources* (page 5-89). If you are ready to learn more about statistics, enroll in a basic statistics course at your local community college.

1. Calculate averages:

An average is calculated by adding up numeric data and dividing that sum by the total number of persons from whom the data was collected. An average is also called a “mean.” For example: $2 + 4 + 6 + 10 = 22$. $22/4=5.5$, where 5.5 is the average or mean.

Averages are calculated from data that makes sense to be averaged. Such data is called **continuous data**, which means it has a potentially infinite number of possible values. An example is personal income. If personal income is collected by asking how much money a person makes per year, and then recording the specific number given, there are an infinite number of specific responses that could be collected. To average this data, you would add up all of the income responses, and divide them by the number of persons from whom this income data was collected. The average you get is the average income of the respondents, which is a meaningful number.

Some other examples of data that could be collected in a continuous way include: age, height, weight, blood pressure, number of years living in the United States, number of years working for an employer, number of visits to the doctor and number of children, to name a few.

2. Count frequencies:

Some data cannot be averaged, because the average you get is not meaningful. That is because this is **discrete data**, and not continuous data. Discrete data is numeric data that has been arranged into sets of values that create breaks in continuity. Let’s take the income example again. If that same question were collected in a discrete way, then instead of being able to give a specific number, the respondent would have to pick from some choices that were set up ahead of time. These look something like this:

1. Less than \$15,000 per year

2. \$15,000 to \$24,000 per year
3. \$25,000 to \$50,000 per year
4. Over \$50,000 per year

The income question could also collect a “yes” or “no” response if participants were asked if their annual income last year was above or below \$15,000 (this might be interesting to you if \$15,000 was the cut off for eligibility to a specific public assistance program). The responses would then look like:

1. Yes
2. No

You can see from this example how the data would look very different, as it will be a series of 1’s and 2’s, or 1’s through 4’s. Averaging those numbers is not as meaningful as averaging the continuous income data.

Instead you will want to use a different calculation technique, where you count the frequency of responses. A frequency is the number of times each response was given. In the multiple choice example above, count the frequency of each of the four responses. In the yes/no example, count the number of times “yes” was chosen, and each time “no” was chosen.

3. Calculate proportions (percentages)

Once you have counted frequencies, this number can be expressed as a proportion, also called a “percentage”. You do this by creating a fraction with your data, so that the frequency count is in the numerator (the number on top) and the total number of respondents is in the denominator (the number on the bottom). To calculate the proportion, the numerator is divided by the denominator, and then the answer is multiplied by 100. Proportions are calculated like this:

$$\text{Proportion} = \frac{\text{number of responses}}{\text{total number of respondents}} = \text{result} \times 100$$

Expressing data frequencies as a proportion is helpful because it standardizes the data. What you get as a result of your calculation in each case is the proportion of total respondents who chose that particular response, from 0 to 100% - even if the number of responses and the number of respondents look very different for each calculation. This makes it easier to make comparisons between different data.

Let’s look at the income example again. If 45 of your respondents selected “yes” to the question asking if their annual income is less than \$25,000 per year, and there were a total of 100 respondents, then your percentage would equal:

$$\text{Proportion} = \frac{45 \text{ answered "yes"}}{100 \text{ respondents}} = 0.45 \times 100 = 45\%$$

If there were 165 total respondents and still 45 who selected “yes”, the proportion of the total who answered yes would get smaller. We can see this when we calculate the new proportion:

$$\text{Proportion} = \frac{45 \text{ answered "yes"}}{165 \text{ respondents}} = 0.27 \times 100 = 27\%$$

4. Calculate rates:

Rates are similar to proportions, in that they are standardized calculations allowing for easier comparisons of different kinds of data. But instead of being expressed as a percentage, a rate is expressed as the frequency of a given event per some unit of the population. This is helpful when an event is uncommon, such as the occurrence of a rare disease. Rates are calculated much like proportions, although the fraction you calculate is multiplied by a larger number than 100. This number is usually 1,000, 10,000 or 100,000, and is equivalent to the standardized number of persons the rate will be expressed as. The equation is:

$$\text{rate} = \frac{\text{number of some event in population}}{\text{total population}} = \text{result} \times 100,000$$

For example, if there were 15 cases of lung cancer in your community, and 3,500 community members, the rate of lung cancer could be calculated as:

$$\text{Lung cancer rate} = \frac{15}{3,500} = 0.00428 \times 100,000 = 428 \text{ cases per } 100,000 \text{ persons}$$

Disease data is often calculated in this way, which makes it easier to compare the rates of specific diseases in your community with the rates from other communities, counties, or the state. You can see from this calculation that multiplying by 100,000 makes the number more meaningful: 0.00428 doesn't mean much, but 428 cases of lung cancer for every 100,000 persons is a useful way of determining how common a disease is in your community. The unit of population used may differ from 1,000 to 100,000, depending on who is doing the calculations. You can redo the calculation if the rate you would like to compare it to (such as the rate for all Californians) is calculated in a different way.

5. Compare averages, frequencies, proportions and rates:

If you are interested in whether the data looks different between different groups, then you can calculate the averages, frequencies or proportions within those specific groups, and then compare them. Some groups you may want to compare are men and women, different age groups, different racial or ethnic groups, individuals with insurance and those without, for example. Sometimes it's helpful to compare your data in many different ways to see if any patterns emerge.

In the income example, you may want to compare the incomes of men and women:

- If **comparing averages**, you want to first average income for women, and then average income for men. This comparison is fairly easy, as your results will look something like this: the average income for women in our sample was \$24,000 per year, and the average income for men was \$32,000.
- In **comparing frequencies** between men and women, you might find that 37 women have incomes less than \$15,000 per year (responded “yes”), and 27 men have incomes under \$15,000. This is a meaningful comparison if there were approximately the same number of men and women included in the analysis.
- If there are different numbers of men and women included in the analysis, then comparing frequencies is not very meaningful. This is when you want to calculate and **compare proportions**. Using the number above, if there were 200 women included in the analysis and 350 men, then your proportions of respondents with income less than \$15,000 per year would be 19% among women, and 8% among men. Comparing the data in this way shows a bigger difference between men’s and women’s incomes, than just comparing frequencies.
- If **comparing rates**, make sure that the rate you calculate and the rate you compare it to are calculated and expressed in the same way, using the same unit of population. Using the lung cancer example, your community’s lung cancer rate is 428 cases per 100,000 persons. If you find that the state rate is 475 cases per 100,000 persons, then your community fares better, on average, than the rest of the state. If the state rate is 102 cases per 100,000, then you know that your community suffers from lung cancer more frequently than the rest of the state, alerting you to a major health problem that should be addressed.

6. Present the data:

Using tables or graphs provides a useful visual presentation of the data.

Tables:

- Suitable for providing simple or more complicated numeric or percentage information
- Best used for side-by-side comparison of data for various variables or groups
- Important to use when you want to show the exact numeric or percentage values
- Here is an example of a table that was used to show results in a research report about diabetes from the California Health Interview Survey (AL Diamant, SH Babey, ER Brown, N Chawla. *Diabetes in California: Findings from the 2001 California Health Interview Survey*. Los Angeles: UCLA Center for Health Policy Research, 2003.) The full report is available at:
<http://www.healthpolicy.ucla.edu/pubs/publication.asp?pubID=68>

EXHIBIT 10. HEALTH INSURANCE COVERAGE OF NONELDERLY ADULTS BY DIABETES DIAGNOSIS, AGES 18-64, CALIFORNIA, 2001		
	ADULTS DIAGNOSED WITH DIABETES (N=915,000) %	ADULTS NOT DIAGNOSED WITH DIABETES (N=19,488,000) %
EMPLOYMENT-BASED	57.7	63.8
MEDI-CAL	22.0	9.7
PRIVATELY PURCHASED	3.9	6.7
OTHER PUBLIC	2.6	1.3
UNINSURED	13.9	18.4
TOTAL	100	100

Note: Totals may not add to 100% due to rounding.

Source: 2001 California Health Interview Survey

Pie Charts:

- Best when you have simple percentages and the “slices” of the pie are not too numerous
- Ideal for depicting the size of each part as a percentage of the whole
- Avoid dividing the pie graph into too many “slices”. It can lead to confusion when interpreting it
- Important to make sure the grayscale or patterns you use to represent the different “slices” are clear and distinguishable from one another. The best option is to display in color, if possible.
- Here is an example of a pie chart that was used to show results from a research report on hunger in Los Angeles County from the California Health Interview Survey (CA DiSogra, W Yen, M Flood, and A Ramirez. *Hunger In Los Angeles County Affects Over 200,000 Low-Income Adults, Another 560,000 At Risk*. Los Angeles: UCLA Center for Health Policy Research, 2003.) The full report is available at: <http://www.healthpolicy.ucla.edu/pubs/publication.asp?pubID=92>

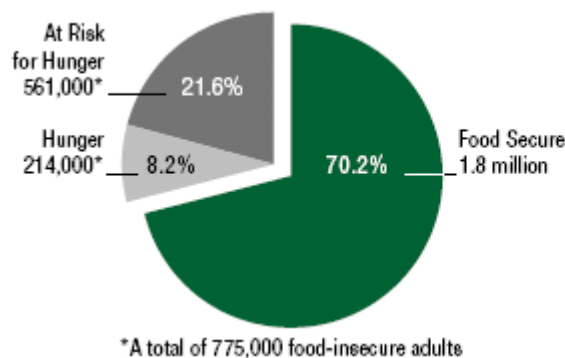
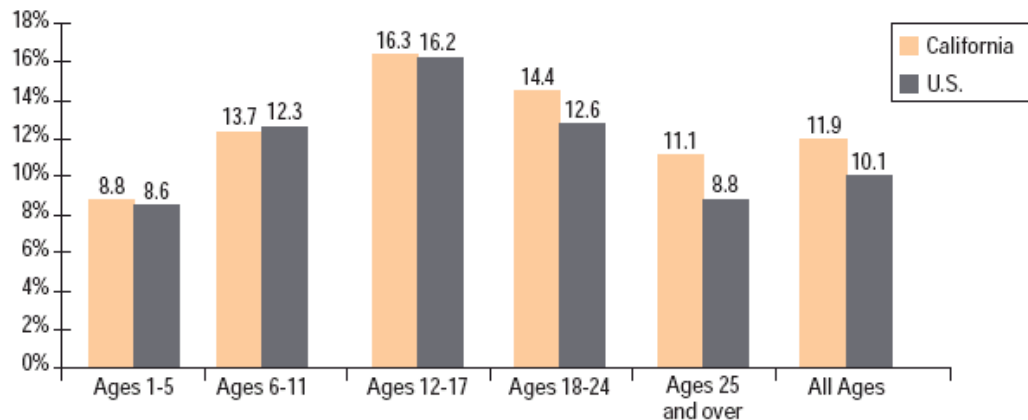


Exhibit 1:
*Food Insecurity among
 2.6 Million Adults with
 Family Incomes Less
 Than 200% FPL,
 Los Angeles County*
 Source: 2001 California
 Health Interview Survey

Bar Graphs:

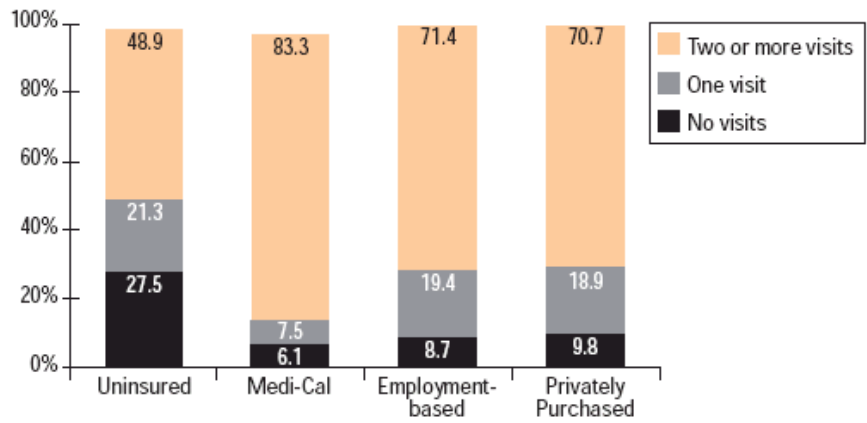
- Good for comparing quantities – simple bar lines are easy to read and compare
- Avoid comparing things that are on different scales—uneven scales can lead to confusion when interpreting the graph
- Important to make sure the grayscale or patterns you use to represent the different "bars" are clear and distinguishable from each other. The best option is to display in color, if possible.
- Here are two examples of bar graphs used to show results in a research report about asthma in California from the California Health Interview Survey (YY Meng, SH Babey, E Malcolm, ER Brown, and N Chawla. *Asthma in California: Findings from the 2001 California Health Interview Survey*. Los Angeles: UCLA Center for Health Policy Research, 2003.) The full report can be found here: <http://www.healthpolicy.ucla.edu/pubs/publication.asp?pubID=83>

EXHIBIT 1. LIFETIME ASTHMA PREVALENCE BY AGE IN CALIFORNIA AND THE UNITED STATES



Source: 2001 California Health Interview Survey and
2000 National Health Interview Survey

EXHIBIT 12. VISITS TO A MEDICAL DOCTOR IN THE PAST YEAR BY TYPE OF INSURANCE, NONELDERLY ADULTS WITH ASTHMA, AGES 18-64, CALIFORNIA, 2001



Note: Totals may not add to 100% because a small percentage of respondents reported they did not know how many times they had seen a doctor in the past 12 months.

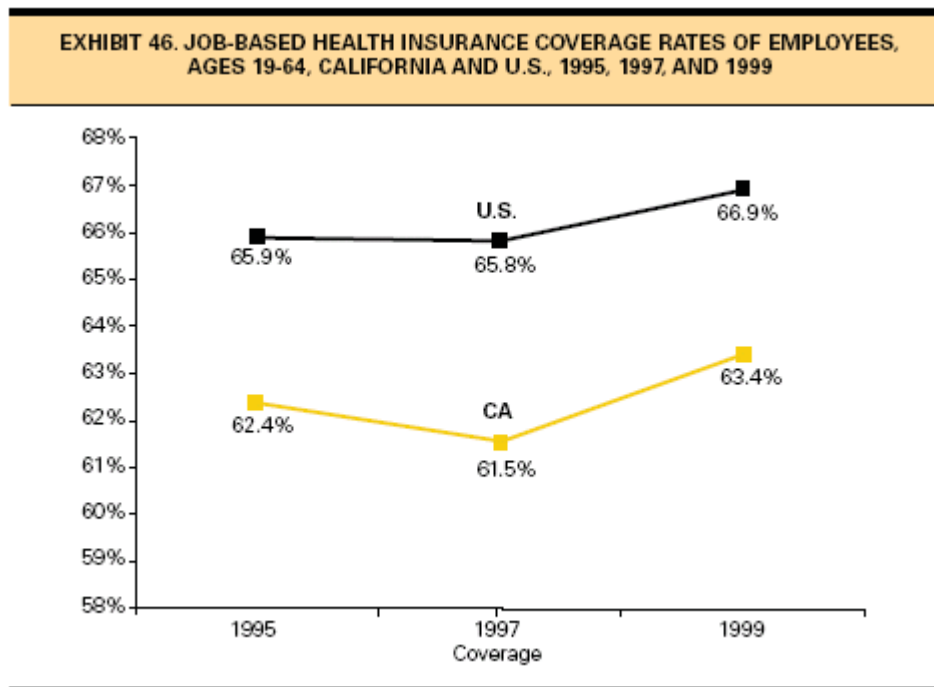
Source: 2001 California Health Interview Survey

Line Graphs:

- Excellent choice when illustrating trends over time
- Line movement, up and down, is easy to understand and interpret

Here is an example of a line graph that was used to show results from a research report on health insurance (ER Brown, N Ponce, T Rice. *The State of Health Insurance in California: Recent Trends, Future Prospects*. Los Angeles, CA: UCLA Center for Health Policy Research, 2001.) The full report can be found at:

<http://www.healthpolicy.ucla.edu/pubs/publication.asp?pubID=78>



Source: February 1995, 1997, and 1999 Current Population Surveys

7. Determine your findings:

Did you discover interesting results from your data? Did you find interesting patterns? These are your findings.

EXERCISE 5.4A

Please go to *Appendix C: Materials for Step 5 Data Analysis Exercises* (page 5-73) to find the materials for this exercise. Included there is a sample quantitative survey about childhood asthma, and an example of a numeric spreadsheet created to enter the data collected through this survey. The columns represent the questions, numbered 1 through 8, and the rows represent the respondents, numbered 1 through 20. You'll notice that only the respondents who answered "yes" to survey question four— "Has a doctor ever told your child that he or she has asthma" —were included in this spreadsheet. The numbers next to the response check boxes correspond to the way the data was coded and entered into the spreadsheet.

Use the data in this spreadsheet to answer the following questions:

- 1) What is the **average** age of the survey participants' children?
- 2) **How many** children are male? **How many** are female? What **proportion** of the children are male and female?
- 3) **How many** children are:
 - Latino?
 - Pacific Islander?
 - American Indian/ Alaskan Native?
 - Asian?
 - African American?
 - White?
 - Other?

What **proportion** of children are:

- Latino?
- Pacific Islander?
- American Indian/ Alaskan Native?
- Asian?
- African American?
- White?
- Other?

Create a table that allows you to compare the frequencies and percentages for each race/ethnicity category. Put numbers and percentages in the columns, and the different categories of race and ethnicity in the rows.

- 4) If the total population of this community is 54,786, and there were a total of 600 diagnosed cases of childhood asthma, calculate the **rate** of childhood asthma per 1,000 people.

- 5) If the state rate of childhood asthma is 65 cases per 1,000 people, then how does this community compare to the entire state of California?
- 6) Calculate the frequency of each of the responses to survey question five, "How often does your child's asthma limit his or her physical activity?", among the African American children.
- 7) Calculate the frequency of the responses to survey question five among the white children.
- 8) How do these frequencies among African Americans and whites compare? What conclusions could you draw from these results?

⇒ **Qualitative Data Analysis – Looking for Themes in the Data**

The process for analyzing qualitative data is very different than the process for analyzing quantitative data. Although there are fewer steps involved, qualitative data analysis can often be trickier—and just as time-consuming—as quantitative data analysis. Qualitative data is always more voluminous than numbers, and determining data findings is more subjective than doing calculations with numeric data, as it requires more abstract thinking and interpretation. Even with the best of intentions, readers and analysts of qualitative data bring with them their own set of values, interests and assumptions, which may bias the way they interpret a focus group discussion or an interview response. You want to minimize the amount of bias that occurs in the analysis process, as releasing biased findings is not helpful in addressing the real issues of concern to your community. It also allows potential opponents or critics to discredit your work.

There are two very important steps you can take to avoid this bias in your qualitative data analysis:

- Involve at least two, but preferably three persons in every step of the data analysis process. This ensures that at least two different viewpoints and interpretations are captured when determining the findings. Qualitative data analysis is more about consensus, and less about calculations.
- Create as structured a data analysis process as is appropriate, to ensure that the data is processed in a thorough and efficient way. It can be tempting to quickly summarize the data, highlighting topics you heard raised in one focus group or you “know” are important ones to your community. However, if you systematically read and summarize the data, you may find that those topics were not raised as much as other topics that were discussed. Sometimes the data can be surprising, so let it surprise you.

1. Read data:

Read through all of your qualitative data. This may include transcripts or notes from focus group discussions or qualitative interviews. As you read through the data, note common topics or ideas that emerge several different times.

2. Create categories:

A category is a major topic area into which your qualitative data naturally falls. The categories you use in your qualitative data analysis can be either the main health issues on which your assessment has decided to focus, or they can be the particular questions of your interview or focus group. Once you have identified and listed the main categories, you should reorganize the data to fit under these categories, so you can re-read the discussions related to each category.

Examples of broad categories include: health status, access to health care, demographic issues, and opinions about community life.

3. List emerging themes:

When you read through the data again, reorganized by category, note the different sub-categories—or themes—that emerge throughout the discussion. For example, if the category is access to health care, the themes that may emerge in the discussions could include: lack of insurance, lack of transportation to get to health facilities, inconvenience of hours of operation of health facilities, lack of child care and lack of time.

4. Rank order:

After re-reading the data and noting the emerging themes, create a list of each theme that emerged under each category. Arrange these themes in order of how frequently they appear in the data. What you will have is a separate list for each category, ordered from the most frequent to the least frequent themes.

Compare these lists to see if any of the same themes emerged under multiple categories. This reveals relationships between ideas discussed under one topic to ideas discussed under another topic.

Compare the lists created by your two (or three) analysts, and discuss the reason for these differences. All the analysts should then discuss the themes they found until they are able to agree and create one list from which to guide the rest of the analysis.

5. Summarize:

Summarize the discussion that occurred around each of your themes. Organize this summary by category heading, starting with the most common themes and ending with the least common. You may decide to include only the themes raised by more than one person, so you can conclude that these were more representative of the community experience than a lone voice. You may decide to highlight certain findings, as you may want to highlight certain categories or themes that were discussed.

Summarize those themes that emerged under two different categories separately.

If you find that different kinds of focus group participants or interviewees discussed the topics in distinctly different ways, then summarize these differences separately. For example, if some focus groups were conducted among men, and others among women, you may want to summarize the discussions separately on topics where the opinions expressed were different. Compare and contrast these differences in your summary.

6. Determine your findings:

Did you discover interesting results from your data? Did you find interesting patterns? These are your findings.

EXERCISE 5.4B

Please go to *Appendix C* to find the materials for this exercise. This is a sample page from a focus group discussion that was conducted among Women, Infants and Children (WIC) program participants who were asked to discuss the eating habits of their children under five years old. Using this discussion, please discuss and answer the following questions within a small group.

- 1) What themes are being discussed in this focus group? Discuss these within your group until you can express each theme so that it reflects what was said in the focus group, and not any one person's bias.
- 2) What themes occur most frequently in the focus group discussion? Discuss these within your group until you agree on the same set of five most commonly occurring themes.
- 3) Are there any surprising themes? Why are they surprising?
- 4) What conclusions can you draw from these themes?
- 5) What additional information would you like to know to be able draw accurate conclusions?
- 6) Did your group disagree? Why? How can this help to identify themes and draw conclusions?

Once you have selected whether you will collect quantitative or qualitative data, you will want to complete the data analysis plan included as the *Step 4 Planning Worksheet*. Some important issues you will want to consider for this data analysis plan include:

1. What is the nature of the data collected with each method – **is it quantitative, qualitative, or both?**
2. What **specific activities** will need to be accomplished in order to best plan and perform the proposed data analysis? These steps may include some of the data analysis stages outlined for quantitative and qualitative data in this step, as well as activities specific to your community partnership's resources and timeline. These activities will also be specific to the methods you have chosen, so be sure to read more detail about each method you select in *Appendix A: Data Collection Methods*. Some activities may include entering the data, training data entry persons, purchasing computer software, performing calculations or summaries of the data, etc.
3. Determine who in your community partnership has the experience, interest and time to accomplish each activity.
4. Select a due date or timeline for each activity so that this planning worksheet can become an achievable Data Analysis Plan.

5.5 Verify Findings

After looking for patterns in your quantitative data and themes in your qualitative data, you should verify your data findings to ensure that your data analysis was effective in creating reliable findings.

- Based on the results tallied or summarized previously, pull out the main findings from each method.
- Verify these findings by re-tallying and re-summarizing the data to make sure you get the same results.

5.6 Interpret Findings and Draw Conclusions

Discuss the data findings with community partners to discuss how to interpret them.

- Determine what interpretations can be drawn from each of these findings. Invite different perspectives from different community partners, as this strengthens your conclusions.
- Think about whether the results are similar to what you expected. If not, discuss why you think they are different.

- Brainstorm alternative explanations for your results to make sure you have considered all possibilities.
- Make sure the data answers the original assessment questions.
- Draw conclusions that wrap up the main findings and can be shared with external audiences.

EXERCISE 5.6

Please go to *Appendix C* to find the materials for this exercise. These four data tables are from a secondary data source called the California Health Interview Survey (CHIS). Use these tables to answer the following questions in a small group. Discuss your responses.

- 1) What is the rate of asthma among African-American children in Los Angeles?
- 2) What is the rate of asthma among white children in Los Angeles?
- 3) What is the rate of asthma among Latino children in Los Angeles?
- 4) What is the rate of asthma among American Indian/ Alaskan Native children in Los Angeles?
- 5) What is the rate of asthma among African-American children in California?
- 6) What is the rate of asthma among white children in California?
- 7) What is the rate of asthma among Latino children in California?
- 8) What is the rate of asthma among American Indian/ Alaskan Native children in California?
- 9) Do the rates in California differ from the rates in Los Angeles County?
- 10) How would you interpret these findings?
- 11) What conclusions can you draw?

STEP 5 PLANNING WORKSHEET: DATA ANALYSIS PLAN

Data Collection Method(s)	Quantitative or Qualitative Data?	Specific Data Analysis Activities	By Whom?	By When?
Example: Surveys	Quantitative <i>and</i> qualitative	<ul style="list-style-type: none"> • Enter data from the completed survey into a spreadsheet • Calculate means, frequencies and proportions for demographic variables and answers to each question • Calculate cervical cancer rate • Compare answers and rates across different demographic groups to look for differences • Create tables and graphs of results • Compile responses to qualitative questions • Summarize responses and create a summary table • Draw conclusions 	Identify your specific partners here with expertise in these areas	Identify timelines here for each of the activities.