Section E: Analyzing and Interpreting Data

Program evaluation requires analyzing and interpreting data in order to make decisions about programs such as what program activities to fund, which activities to improve, and how to improve them. Up to this point in the evaluation process:

- You focused your evaluation with a select set of questions about program implementation and results in order to make appropriate decisions about your program.
- You also selected practical and affordable methods for gathering reliable and credible data.

The next step is to make sense of the data you collected so that you can answer specific evaluation questions about the quality of your program, what it accomplished, and the extent to which different categories of participants benefited.

How you analyze, organize, or summarize your data depends on the type of data you collected:

- Quantitative data, or numerical information, is analyzed or summarized using descriptive and inferential statistics.
- Qualitative data, or information gathered in narrative form, is analyzed or organized into categories or themes, which can in turn be summarized by percentages.

In this section of the *Program Evaluation Toolkit*, there is a description of how to analyze and interpret quantitative data using a selection of descriptive statistics, and how to analyze and interpret qualitative data by organizing it into meaningful categories. Data storage options that facilitate subsequent analyses are also discussed.

E.1 Analyzing Quantitative Data

Numerical data are used to summarize useful and important information about large groups of people. In order to get the most out of this information, it is helpful to know something about the different types of data elements, or variables, which can be analyzed using quantitative methods.

Categorical variables describe people and things in discrete categories. For example, gender is a categorical variable that has two categories – male and female.

- Migrant status is a categorical variable that can have two categories (migrant and non-migrant) or three categories (Priority for Services [PFS], other migrant, and non-migrant).

Continuous variables are expressed as numbers on a continuous scale. For example, age is a continuous variable that can be measured with different levels of precision (e.g., years, months, days, hours).

- A scale score is a continuous variable representing student achievement, but proficiency level is a categorical variable that is also used to describe student achievement.

Even though categorical variables are represented by names or labels and continuous variables are represented by numbers, both types of data can be analyzed using quantitative methods.

In the broadest sense, analyzing data means looking for patterns that are relevant to what you are studying. In the case of evaluation, which relies on making comparisons to assess value, patterns are observed in order to compare two or more groups. For example:

- Those who participated in a program and those who did not
- PFS and other migrant and non-migrant students
- Students who achieved success as defined by measurable program outcomes and those who did not

The following descriptive statistics can help reveal relevant patterns when comparing for two or more groups.

The two branches of statistics are descriptive statistics and inferential statistics.

Descriptive statistics describe a specific sample of people or things.

Inferential statistics, including the use of statistical significance tests, are used to understand or draw conclusions about the population from which your specific sample was obtained.

For example, you could use descriptive statistics to identify changes in academic performance among migrant students who received a specific instructional service (e.g., one-on-one tutoring) through the local migrant education program, you could then use inferential statistics to determine whether these findings could be generalized to the population of migrant students in your school district, as long as they also received the same instructional service.

E.1.1 Descriptive Statistics for Categorical Variables

Frequencies are counts calculated for categorical variables to describe the composition of a group such as a group of people, things, or survey responses. These counts express amounts and are usually translated into percentages for easy monitoring of changes over time. For example, changes in the demographic composition of migrant students in your state may explain an increase in demand for certain types of services or a decline in participation in longstanding services.

A rate is a percentage that expresses the extent of an action or condition. You and your colleagues routinely compare academic success rates for migrant and non-migrant students when you calculate the percentage of students in each group who have achieved proficiency, and you look for changes in the pattern of differences between the two groups to see whether instructional services for migrant students are helping to close the gap with non-migrant students.

A cross-tabulation, or crosstab, is the calculation of a set of frequencies using two or more categorical variables to describe a group. The result of this calculation is a contingency table that shows how the members of your group are sorted into subgroups depending on the categories they inhabit for each variable. Figure E.1 is an example of a simple 2x2 contingency table, in which frequencies are calculated for two variables at the same time and both variables have only two categories:

Figure E.1 Example of 2x2 Contingency Table

	Migrant Status of Limited English Proficient (LEP) Students		
English as a Second Language (ESL) Program Participation Status	**# Migrant LEP Students**	**# Non-Migrant LEP Students**	**Row Totals**
# Who participated in the after-school ESL program	50	275	325
# Who did not participate in the after-school ESL program	100	75	175
Column Totals	150	350	500

In this example, a total of 500 Limited English Proficient (LEP) students, some of whom were migrant, some of whom were non-migrant, were identified as being able to benefit from a district-wide after-school English as a Second Language (ESL) program. Space was budgeted for all 500 eligible students. The row totals show that 325 of the eligible students participated. The remaining 175 eligible students did not participate. For eligible LEP students as a whole, the participation rate was 65% (325 ÷ 500). This result is respectable for a new program.

To identify ways to improve participation, consider reasons why 35% of eligible students did not participate. Knowing that the reasons for non-participation may be different for migrant and non-migrant students, examine the cells in each column and note that the participation rate for migrant students (50 ÷ 150, or 33%) was much lower than the participation rate for non-migrant students (275 ÷ 350, or 79%).

Based on this finding, make it your first priority to identify the reasons why migrant LEP students chose not to participate, or were not able to participate, in the ESL program.

E.1.2 Descriptive Statistics for Continuous Variables

The mean for the distribution of a continuous variable is the same as the average for a set of continuous, numerical values. A mean is one of several "measures of central tendency" used to calculate a "typical" number to describe a whole group of numbers or the people they represent and compare it to the typical value for another group, or for the same group at another point in time.

Consider the case of a state assessment system that defines cut points that divide continuous scale scores into four levels of proficiency: (1) well-below proficiency, (2) below proficiency, (3) proficient, and (4) advanced. Your state performance target is to increase the percentage of PFS students and other migrant students who are proficient in reading by 10% this year in each grade.

- In spite of all your targeted and evidence-based efforts, you are discouraged to find that only small percentages of students in each grade have gained enough to move into the next highest proficiency level. In order to get a more precise view of what has occurred in the lower two proficiency levels, you calculate mean scale scores for migrant students in each of these groups and find positive change in the mean scale scores for each group in every grade. So, even though students did not meet their state performance targets, you can see that they did achieve growth.
- Some of the differences in the grade-by-grade mean scale scores are bigger than others. You compare strategies used in the most successful grades with those used in less successful grades to identify programs that may have made a difference.

A median is another measure of central tendency and is often used to describe a set of numbers that includes some extreme values, or outliers. When the distribution includes outliers, the mean can be skewed and either overestimate or underestimate the middle or "typical" value. While the mean is the computed middle, the median is literally the middle-most number in a set of numbers when you arrange them in order.

For more information about analyzing quantitative data, go to:

Quantitative Data Analysis: An Introduction, U.S. General Accounting Office, May 1992

Taylor-Powell, E. (1996). Analyzing Quantitative Data. University of Wisconsin-Extension.

Leahy, J. (2004). Using Excel for Analyzing Survey Questionnaires. University of Wisconsin-Extension, 2004.

E.2 Analyzing Qualitative Data

While quantitative data answer questions related to magnitude (how many, how much, to what extent), qualitative data can help us understand more complex issues related to "how" and "why." For this reason, many evaluators and program directors prefer working with qualitative data.

Just like quantitative data, qualitative data are analyzed by systematically looking for patterns. If instead of studying numbers, researchers carefully comb through interview transcripts or open-ended survey responses to find the themes and patterns that emerge in people's answers to our questions. Categories and subcategories are then defined to describe different kinds of answers. In order to convey the relative strength of the themes that are uncovered, the qualitative analysis findings are then summarized by calculating the percentages of respondents who provided answers in the different categories and displaying them in a table.

Step 1: Organizing individual responses into broad categories. Qualitative researchers create categories, or codes, of increasing specificity as they review data. Level 1 codes are the broadest categories that can be used to describe your data.

- In the case of a focus group or interview that includes multiple questions, you can use your interview questions as your Level 1 categories.
- In the case of an open-ended survey question that was answered by people in different roles, you might use the roles of the respondents (student, parent, teacher, Migrant Education Program [MEP] staff) as Level 1 categories.
- You can begin with a set of Level 1 categories or themes that you expect to find in the data based on your involvement in the data collection, past experience, or knowledge of the subject area being discussed.

Consider the example of an open-ended survey question that was answered by a group of migrant students in 10th grade and by the parents of these students. Seventy migrant students and 85 parents attended a four-hour Saturday morning information session to learn about preparing for and succeeding in college. Students and parents attended the same sessions and received the same information. Therefore, you might begin with two Level 1 categories: one category for answers given by the students [coded "STU"] and a second Level 1 category for answers given by parents [coded "PAR"].

You have invited students and parents to a brief follow-up session to provide additional information they requested. Forty-nine students and 62 parents attend the follow-up session. Before they leave, you give each of them a half-page questionnaire with two questions. The first question asks them to use a five-point scale (from 1 = "not at all useful" 5 = "very useful") to rate the usefulness of the information they have received for planning their next steps toward college. The second question reads, "What specific pieces of information, if any, were useful to you?" There is plenty of blank space on the questionnaire to write their answers.

Step 2: Within each broad Level 1 category, organize individual responses into related subcategories. The process of sorting responses in Level 1 categories into narrower and more informative sub-categories is known as Level 2 coding.

- Going back to our example, you have collected surveys from all of the students and parents and are ready to begin sorting their answers into more meaningful categories. You begin reading through each group's individual answers to identify Level 2 sub-categories.
- The information that was most useful to students fell about equally into three sub-categories – academic, extracurricular, and financial. You code their answers using the abbreviations "ACAD," EXTRA," and "FINAN," respectively. Parents' answers also fell into the same sub-categories, but most of the information they found useful fell into the "FINAN" sub-category.

Step 3: Summarize results in a table showing percentages of responses in each Level 2 sub-category and for each Level 1 category.

Level 2 Categories	**Level 1 Categories**	
Categories of useful information	**Students (n=49)**	**Parents (n=62)**
Academic	39%	37%
Extra-curricular	42%	12%
Financial	29%	77%
Non-specific, positive (e.g., "It was all useful")	4%	6%
Non-specific, negative (e.g., "Nothing was useful")	2%	---

*Percentages add to more than 100% for each group because answers fall into more than one sub-category.

- The findings summarized in the table reinforce the generally high ratings given for usefulness by students and parents who attended the follow-up session. The vast majority of participants in both categories could name at least one piece of useful information they had received in the original four-hour information session.
- Given the different kinds of information that students and parents found useful, you and your colleagues decide to offer the four-hour information session again next year, but you plan to separate students and parents and provide information that will be tailored to what each group finds more useful for college planning.

E.3 Summary of Key Points

- Analyzing and interpreting the data help identify themes or patterns that emerge from the information collected. These themes can be used to answer specific evaluation questions about the quality and impact of programs.
- Quantitative data, or numerical information, is analyzed or summarized using descriptive and inferential statistics. Quantitative data can be used to summarize useful information about large groups of people.
- Qualitative data, or information gathered in narrative form, is analyzed or organized into categories or themes, which can in turn be summarized by percentages. Qualitative data can help us understand more complex program issues related to "how" and "why."
- Both kinds of data – quantitative and qualitative – can be summarized in tabular form to reveal patterns, show comparisons or demonstrate impact.

E.4 Reflection Questions

1. Have we considered the appropriate variables in the interpretation of the data?
2. Have we identified all of the major themes and subthemes?
3. Have the data been analyzed in a way that answers specific evaluation questions?
4. Are there other ways of interpreting the data that we have overlooked?

E.5 Resources and Tools in Appendix E

Appendix E.1 Using Excel to Analyze Quantitative Data
Appendix E.2 Using Access to Analyze Qualitative Data
Appendix E.3 Using Inferential Statistics

Appendix E.1 Using Excel to Analyze Quantitative Data

Microsoft Excel, which is widely available, is able to perform many calculations that are useful in analyzing data. The following instructions are based in large part on a document called *Using Excel for Analyzing Survey Questionnaires* published by the University of Wisconsin Cooperative Extension in 2004.

While this publication assumes the use of Microsoft Excel 2002, we have updated the instructions for those using Microsoft Excel 2010, and we have included examples that are relevant to education in general or to Migrant Education Programs (MEPs) specifically.

The instructions below assume that the user is proficient in the use of databases and familiar with Microsoft Excel. *Using Excel for Analyzing Survey Questionnaires* is an excellent resource for users who need more detailed instructions on how to use Excel. The document is available on the web at https://learningstore.uwex.edu/assets/pdfs/G3658-14.pdf.

As we walk through some instructions for using Excel to analyze data, we will use a very simple survey as an example. There are two questions in our survey:

Q1. Does the student have a Priority for Services (PFS)?

- "Yes" responses are coded "1," and
- "No" responses are coded "2."

Q2. Did the student participate in MEP-funded instructional programs this year?

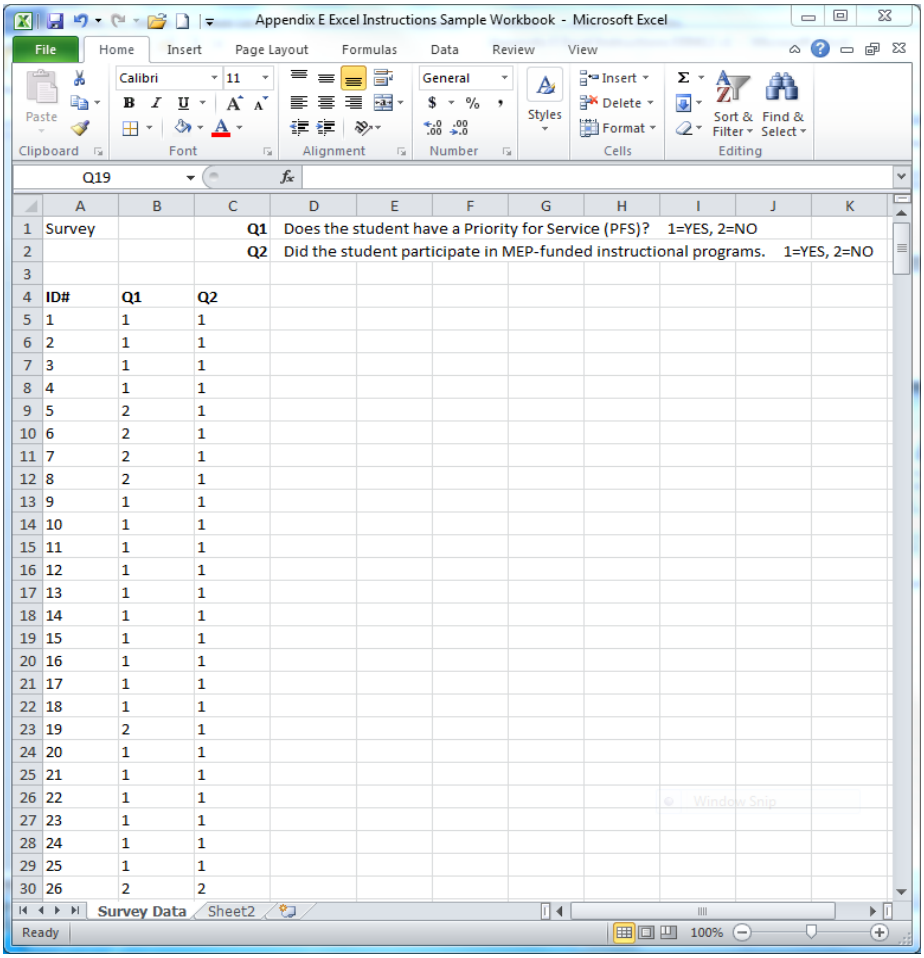
- "Yes" responses are coded "1," and
- "No" responses are coded "2."

Assume that surveys for 250 students were completed and returned to us. Responses were as follows:

1. Does the student have a Priority for Services?	**Yes**		**No**		**Total**	
	200		50		250	
2. Did the student participate in MEP-funded instructional programs this year?	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
	190	10	15	35	205	45

Figure 1 depicts how this simple survey database might be set up in Excel.

Figure 1.1 Example Survey Database in Excel



	A	B	C	D	E	F	G	H	I	J	K
1	Survey			Q1	Does the student have a Priority for Service (PFS)?	1=YES, 2=NO					
2				Q2	Did the student participate in MEP-funded instructional programs.	1=YES, 2=NO					
3											
4	**ID#**	**Q1**	**Q2**								
5	1	1	1								
6	2	1	1								
7	3	1	1								
8	4	1	1								
9	5	2	1								
10	6	2	1								
11	7	2	1								
12	8	2	1								
13	9	1	1								
14	10	1	1								
15	11	1	1								
16	12	1	1								
17	13	1	1								
18	14	1	1								
19	15	1	1								
20	16	1	1								
21	17	1	1								
22	18	1	1								
23	19	2	1								
24	20	1	1								
25	21	1	1								
26	22	1	1								
27	23	1	1								
28	24	1	1								
29	25	1	1								
30	26	2	2								

E.1.1 Calculating Frequencies

Frequencies are percentages used to describe the composition of a group, and there are a couple of different ways to calculate frequencies in Excel.

One method of calculating frequencies in Excel is to use the COUNTIF function to determine how many times a response occurs, and then use a formula to determine what percentage that number represents of the whole survey pool. In a blank cell, type the formula:

=COUNTIF(range, criteria)

Where:

- "range" includes the cells that hold your data, with the first cell and the last cell separated by a colon (:), and
- "criteria" is the specific response code for which you are calculating the frequency.

For the sample survey data above, the range for Question 1 is B5:B254. The criteria will be 1 (Yes) or 2 (No). We will create a separate COUNTIF calculation for each of the criteria.

- First, we'll calculate how many students had a Priority for Services (PFS). The formula is:

  =COUNTIF(B5:B254, 1)

  This formula tells Excel to look at the survey responses in cells B5 through B254 and count how many people responded "Yes" to Question 1.
- To calculate how many students did not have a PFS, the formula is

  =COUNTIF(B5:B254, 2)

  This formula tells Excel to look at the survey responses in cells B5 through B254 and tell us how many people responded "No" to Question 1.

To convert these absolute numbers into percentages, we'll use simple formulas to divide the number of "Yes" responses by the total number of survey responses and the number of "No" responses by the total number of survey responses.

Figure 2 displays the results of our frequency calculations in Excel.

For "Yes" responses, the formula is:

=D277/D279

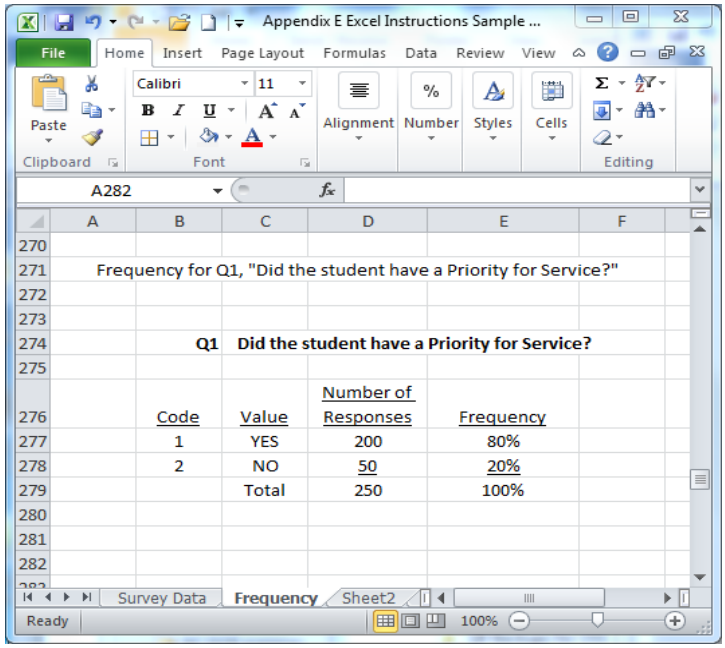
200/250 = 80%

For "No" responses, the formula is:

=D278/D279

50/250 = 20%

Figure E.1.2 Results of Frequency Calculations



E.1.2 Calculating Rates or Percentages

While frequencies represent the *amount* of something, rates convey the *extent* of something.

To calculate rates in Excel, use the COUNTIF function to determine how many survey respondents indicated a certain answer for each response available on the survey, and then simply divide the number for each response by the total number of responses to convert the answers into percentages.

For our sample survey, we might want to know whether the rate of participation in Migrant Education Program (MEP)-funded instructional programs differs between students who have a PFS and those who do not.

Recalling the responses on our surveys, we know from the 250 surveys returned to us:

- 200 students have a Priority for Services (PFS)
  - 190 of these students participated in MEP-funded instructional programs;
  - 190 divided by 200 = 95%; therefore,
  - 95% of students who have a PFS participated in MEP-funded instructional programs.
- 50 students do not have a PFS
  - 15 of these students participated in MEP-funded instructional programs;
  - 15 divided by 50 = 30%; therefore,
  - 30% of non-PFS students participated in MEP-funded instructional programs.

In this very simple example, calculating the percentages manually is easy. However, for a question that has more possible combinations of responses, this could be a tedious and time-consuming task. Instead, we can use the COUNTIFS function in Excel to do these calculations for us. [Notice the "S" in the formula when we are calculating a frequency based on more than one criterion. This is not a typo!]

In essence, the COUNTIFS function calculates frequencies for each possible combination of responses. We can then convert these absolute numbers into percentages in order to compare results across groups.

Using Excel for our calculations, the formula would be:

=COUNTIFS (criteria_range1, criteria1, criteria_range2, criteria2, …)

Where:

- "Criteria_range1" includes the cells that hold data for the first criterion you wish to count;
- "Criteria1" is the first response code for which you are calculating the frequency;
- "Criteria_range2" includes the cells that hold data for the second criterion;
- "Criteria2" is the second response code to count, etc.

Let us return to our simple survey example to illustrate the use of the COUNTIFS function. Our example involves two questions, each with its own corresponding range of data and two criteria (or response codes) for each question. However, you can use the COUNTIFS function to calculate rates for more than two questions and more than two criteria per question (e.g., Yes, Maybe, No).

As a reminder, we want to know if the rate of participation in MEP-funded instructional programs differs between students who have a PFS and those who do not. First let us look at students who do have a PFS and who did participate in MEP-funded instructional programs:

- The range for Question 1 responses is B5:B254, and the code for "Yes" is 1 (criteria1 = 1);
- The range for Question 2 responses is C5:C254, and the code for "Yes" is again 1 (criteria2 = 1).

The formula, therefore, will be:

=COUNTIFS(B5:B254, 1, C5:C254, 1)

This formula tells Excel to give us the number of students who answered "Yes" to Question 1 AND "Yes" to Question 2. There are 190 students in our survey who answered "Yes" to both Question 1 and Question 2.

We can do the same to students who do not have a PFS but who did participate in MEP-funded instructional programs:

- The range for Question 1 responses is B5:B254, and the code for "No" is 2 (criteria1 = 2);
- The range for Question 2 responses is C5:C254, and the code for "Yes" is 1 (criteria2 = 1).

The formula, therefore, will be:

=COUNTIFS(B5:B254, 2, C5:C254, 1)

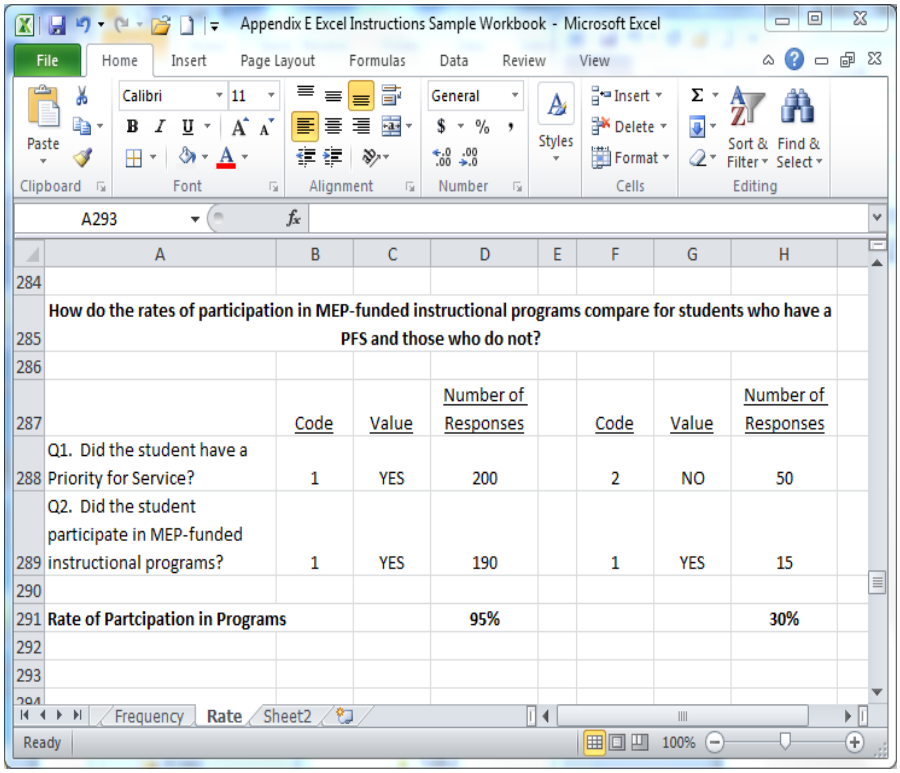
Figure 3 shows what this exercise might look like on our Excel spreadsheet.

To convert the absolute numbers into percentages, we'll use simple formulas to divide the number of responses that meet our criteria by the total number of students in that group:

- For students who do have a PFS and did participate in instructional programs,
  =D291/D290 = 190 divided by 200 = 95%.
- For students who do not have a PFS and did participate in instructional programs,
  =H291/H290 = 35 divided by 50 = 30%.

In answer to our question, we can say that the rate of participation in MEP-funded instructional programs in our survey is higher for students who do have a PFS (95%) than for students who do not have a PFS (30%).

Figure E.1.3 Example Exercise in Excel



E.1.3 Creating Cross-tabulations

Another approach to comparing data across groups involves cross-tabulations, or crosstabs. Excel uses the Pivot Table Wizard to create crosstabs.

We will work with the same question we used earlier: Does the rate of participation in Migrant Education Program (MEP)-funded instructional programs differ between students who have a Priority for Services (PFS) and students who do not? We will also use the database we have already set up (see Figure E.1.1).

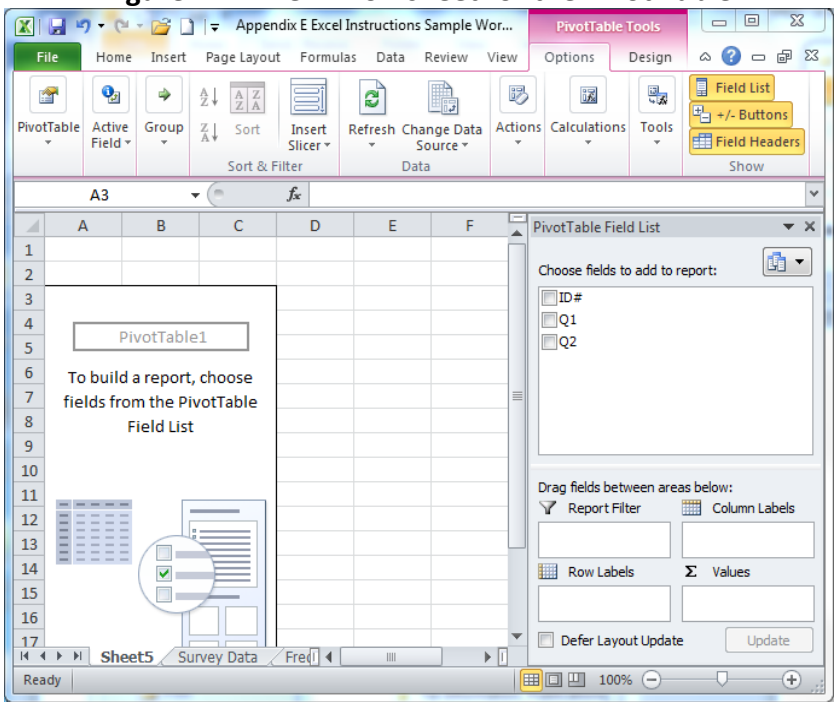
To start creating a crosstab, go to the Insert tab in Microsoft Excel 2010, click on PivotTable, and choose PivotTable.

The first dialogue box will ask you to select a data range. Our data range is $A$4:$C$254. You can type that in or simply highlight the range on your screen. Notice:

- We use "$" to anchor the columns (as in, $A and $C) and rows (as in, $4 and $254), and
- In the range, we also now include the cells with the column headers, which is necessary in order to identify the variables we are using.

Excel also wants to know where to put the pivot table, so choose New Worksheet, and click OK. A PivotTable Layout window will appear in the new worksheet (see Figure 4).

Figure E.1.4 New Worksheet for the Pivot Table

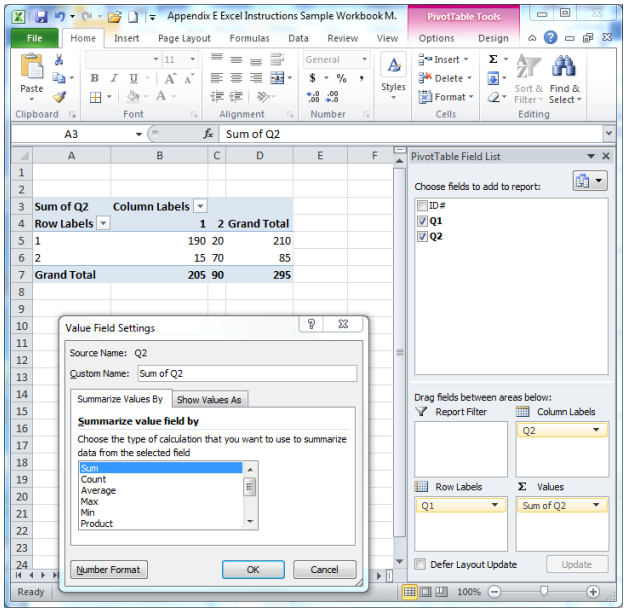


For the Pivot Table, we need to designate an independent variable and a dependent variable. For this example, the independent variable will be Question 1 (*Does the student have a Priority for Services?)*, and the dependent variable will be Question 2 (*Did the student participate in instructional programs this year?*). Since independent variables are used to explain something about dependent variables, we are in effect asking whether participation rates differ by PFS status. Therefore, we are using PFS status (the independent variable) to explain participation rates (the dependent variable).

On the PivotTable worksheet, click on Q1 (PFS status, the independent variable) in the Pivot Table Field List, and drag it down to the Row Labels section. Click on Q2 (participation, the dependent variable) and drag it down to the Column Labels section. Click on Q2 again (in the Pivot Table Field List), and drag it down to the Values section.

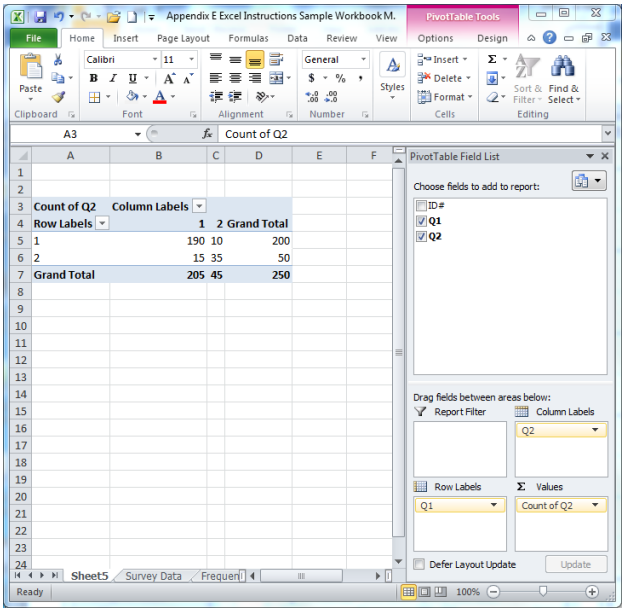
In the Values section, you can choose what you want Excel to do with the data. In our example, we want Excel to count the number of responses for each value. However, the default in Values is to calculate the sum of the responses (see Figure 5.)

Figure E.1.5 Sum of Values (Default)



To change the default operation in Values, click on the drop down menu in Values, choose Value Field Settings, and choose Count. You will notice that the label in the Values box changes from “Sum of Q2” to “Count of Q2” (see Figure 6).

Figure E.1.6 Count of Values



While some of the numbers in the PivotTable might look familiar (190, 200, 250), we do not know what we are looking at without our variable (or field) names. We will have to type them in ourselves:

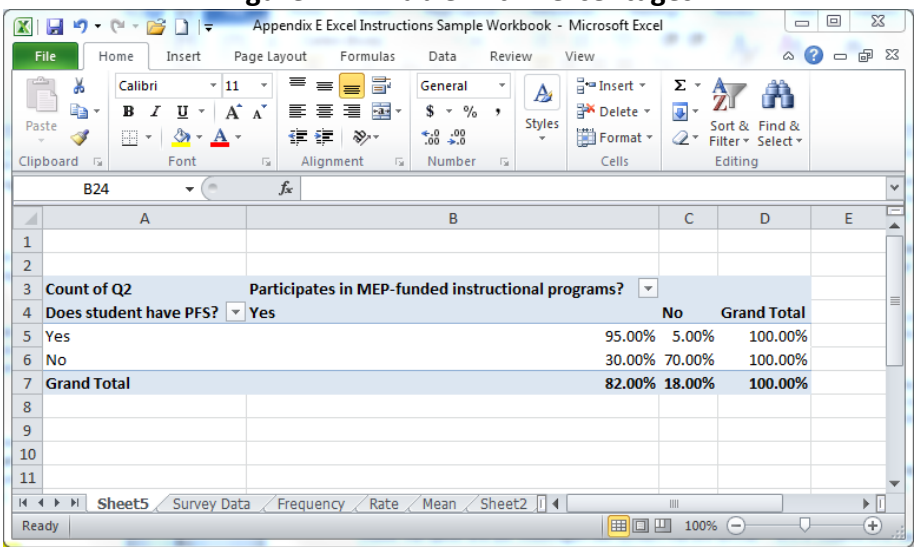
- In the PivotTable itself, click on “Row Labels” (which we know represents Q1) and type “Does student have PFS?” Next, let’s define the values in the rows: Change “1” to “Yes” and change “2” to “No.”
- Now click on “Column Labels” (which we know represents Q2) and type “Participated in MEP-funded instructional programs?” Next, define the values in the columns: Change “1” to “Yes” and change “2” to “No.”

Now we can now see that the number of students who have a PFS and participate in instructional programs is 190. We can also see that the number of students who do not have a PFS but do participate in instructional programs is 15. But, again, percentages are more helpful in order to compare participation rates for the two groups.

To display percentages in the pivot table, go back to the Values Section and click on the drop down box next to “Count of Q2” (the column variable), choose Value Field Settings, and click on the tab called “Show Values As.” Then, click on the drop down menu next to “No Calculation,” and select “% of Row Total.” Now you should see the familiar result that 95% of PFS students participated in instructional programs this year compared to 30% of non-PFS students.

In this case, we select “% of Row Total” because the independent variable Q1 is the row variable, and we are interested in how Q1 explains the dependent variable Q2, participation. The results in the Pivot Table show us that participation rates are very different depending on whether the migrant student does or does not have a PFS. In this case, PFS status does indicate something important about participation!

Figure E.1.7 Table with Percentages

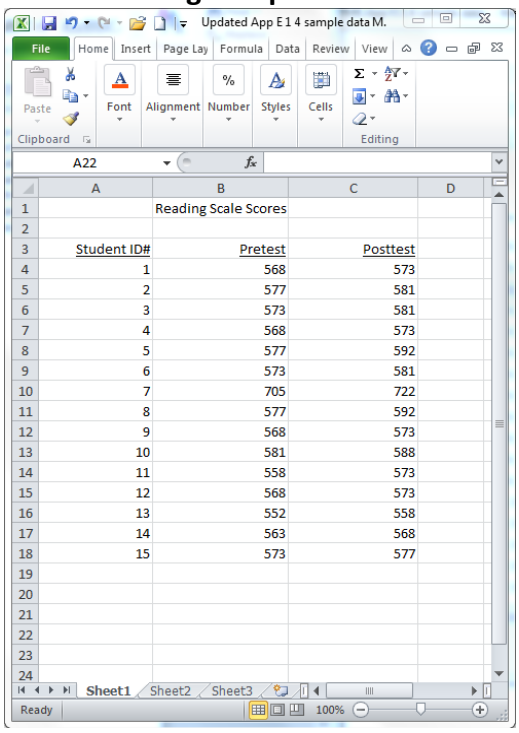


E.1.4 Calculating Means and Medians

As mentioned earlier in this section of the *Toolkit*, means and medians can be useful when individuals in a group are described by numbers on a continuous scale. Excel contains functions for calculating both means and medians.

Consider the example of a group of migrant students who are rising 9th graders. They participated in an eight-week reading and math enrichment program at the high school they will be attending in the fall. On the first day of the program, every student was tested in Reading Comprehension and Mathematics Problem Solving using the SAT-10 Advanced 2 test booklet. The teacher did not go over the correct answers with the students, but he did emphasize these skills throughout the enrichment program. On the next-to-last day of the program, students took the same two sections of the SAT-10 again. Their raw scores were converted to scale scores. Results for reading comprehension are displayed in Figure 8.

Figure E.1.8 Reading Comprehension Scale Scores



	A	B	C	D
1		Reading Scale Scores		
2				
3	Student ID#	Pretest	Posttest	
4	1	568	573	
5	2	577	581	
6	3	573	581	
7	4	568	573	
8	5	577	592	
9	6	573	581	
10	7	705	722	
11	8	577	592	
12	9	568	573	
13	10	581	588	
14	11	558	573	
15	12	568	573	
16	13	552	558	
17	14	563	568	
18	15	573	577	
19				
20				
21				
22				
23				
24				

We want to know whether the average performance for the group at the end of the summer was better than their average performance at the beginning of the summer. This can be determined this by comparing the means of the two groups of scores.

To calculate the mean in Excel, use the formula

=AVERAGE(range)

Where:

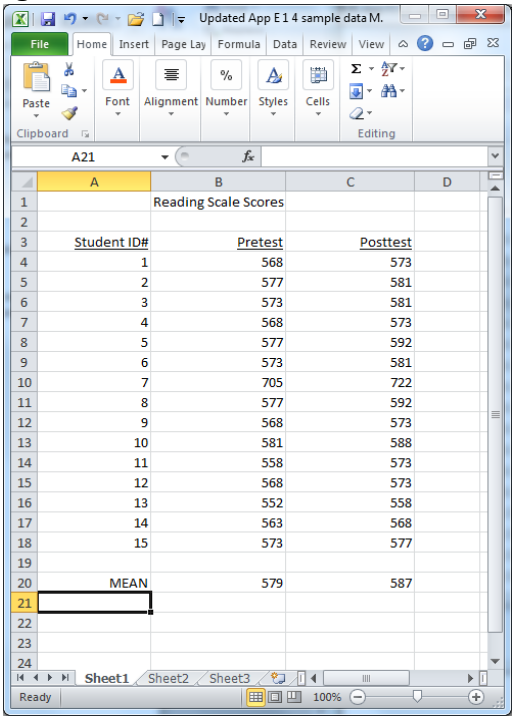
- "Range" indicates the first and last cells that define the range of data for which you want to calculate the mean.

Figure 9 shows the means of the pretest and posttest scores, which we calculated using the following formulas:

=AVERAGE(B4:B18) for the pretest, and
=AVERAGE(C4:C18) for the posttest

As you can see, the mean of the posttest scores is 587 compared to the pretest mean of 579, an increase of 8 points.

Figure E.1.9 Pretest and Posttest Means



	A	B	C	D
1		Reading Scale Scores		
2				
3	Student ID#	Pretest	Posttest	
4	1	568	573	
5	2	577	581	
6	3	573	581	
7	4	568	573	
8	5	577	592	
9	6	573	581	
10	7	705	722	
11	8	577	592	
12	9	568	573	
13	10	581	588	
14	11	558	573	
15	12	568	573	
16	13	552	558	
17	14	563	568	
18	15	573	577	
19				
20	MEAN	579	587	
21				
22				
23				
24				

Reviewing the data, you notice that both the pre- and post-scale scores include an outlier (highlighted in Figure E.1.10). You can see that the outliers inflate the average pretest and posttest scores. You could just remove that student's scores from the data and analyze it again. Instead, you decide to compare the median pretest and posttest scores in order to get a more realistic idea of the middle or "typical" value in each set of scores.

To calculate the median in Excel, use the formula

=MEDIAN(range)

Where:

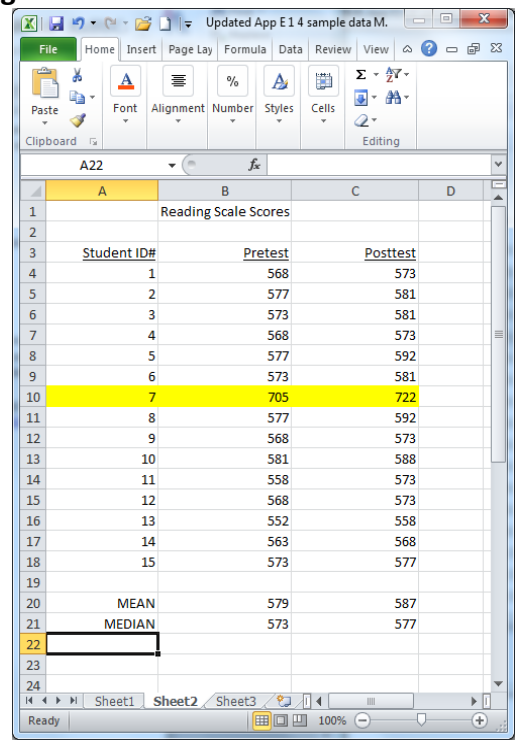
- "range" indicates the first and last cells that define the range of data for which you want to calculate the median.

Figure 10 shows the medians of the pretest and posttest scores, which we calculated using the following formulas:

=MEDIAN (B4:B18) for the pretest, and
=MEDIAN (C4:C18) for the posttest

The median pretest score is 573 (compared to a pretest mean of 579). The median posttest score is 577 (compared to a posttest mean of 587). And the pre-post difference between the medians is a more modest 4 points (compared to 8 points for the pre-post means). In addition to illustrating how to calculate means and medians, this example also demonstrates how outliers can influence your analysis and conclusions.

Figure E.1.10 Pretest and Posttest Medians



	A	B	C	D
1		Reading Scale Scores		
2				
3	Student ID#	Pretest	Posttest	
4	1	568	573	
5	2	577	581	
6	3	573	581	
7	4	568	573	
8	5	577	592	
9	6	573	581	
10	7	705	722	
11	8	577	592	
12	9	568	573	
13	10	581	588	
14	11	558	573	
15	12	568	573	
16	13	552	558	
17	14	563	568	
18	15	573	577	
19				
20	MEAN	579	587	
21	MEDIAN	573	577	
22				
23				
24				

For more information about using Excel to analyze quantitative data, go to: *Using Excel for Analyzing Survey Questionnaires* at https://learningstore.uwex.edu/assets/pdfs/G3658-14.pdf.

Appendix E.2 Using Access to Analyze Qualitative Data

In *Section E.2 Analyzing Qualitative Data* of the *Program Evaluation Toolkit*, we explained how to analyze qualitative data by organizing focus group responses to open-ended survey questions into categories and assigning codes to each category. Level 1 codes are given to responses that fall into the broadest categories you see in your data. Level 2 codes are used to organize data in Level 1 categories into narrower and more descriptive categories.

In the past, we might have taken one of the following approaches to analyzing qualitative data:

- Cutting responses out of the paper on which they were printed and physically re-organizing them into groups that fit together in some relevant or interesting way, then counting the number of scraps of paper we included in each category grouping; or
- Reading through the responses, generating a list of the categories we found, creating a code for each category, then penciling codes in next to each response, and counting the number of times we assigned each code.

Between the paper cuts and eraser shavings, these approaches could be very messy indeed!

E.2.1 Microsoft Access as Qualitative Analysis Software

Today you can purchase software (or download freeware) to code, sort, and summarize qualitative data electronically (e.g., NVivo, Atlas.ti, QDA Miner, etc.). These packages have a lot to offer the serious qualitative researcher, but this may be more firepower than you actually need. Before you purchase or download something new, consider a tool that you may already have in your Microsoft Office Professional arsenal, the database software, Access.

> NOTE: For the purpose of this discussion, we assume that you have some familiarity with Access. If you are not familiar with it, and you have it on your computer, we encourage you to learn more about this flexible data management tool. One way to do this is by taking an online short course through a university computing center. For example, the Computer Training Unit in the McKimmon Center at North Carolina State University offers self-paced online courses in Access for users at three levels of experience. Learn more by visiting: https://onece.ncsu.edu/mckimmon/divisionUnits/ctu/index.jsp

We routinely create databases in Access to store both quantitative and qualitative data. We export the quantitative data into Excel or a statistical software package (e.g., SPSS, SAS, R) for further data analysis. However, we keep our qualitative data in Access and create forms, queries, and reports to help us code and summarize the data.

All data are entered into tables in Access. It can either be entered directly into the table in DataSheet View, or in a data entry "form." Forms are often more visually appealing than

datasheets. In Form View, you can create data entry "windows" that are large enough for you to see your entire text block (e.g., an excerpt from a focus group interview or a single open-ended survey response).

E.2.2 Qualitative Data Example

Before we show you how we use Access forms, queries, and reports to analyze qualitative data, consider the following example:

- Your Migrant Education Program (MEP) office serves as a "broker" for all non-instructional, support services to migrant students, families, and out-of-school youth (OSY) in your area. The recruiters in your office do their best to make sure the people they identify are given information about all available services.
- Once a year, you go out into the migrant community to find out what your target population thinks about the services you offer and what services they still need. To do this, you set up focus groups at several local churches that have sizable numbers of migrant parishioners. You will use the information gathered in the focus groups to improve communication about the services you offer, improve the services themselves, and find out what other services are needed in the community.
- With the help of the pastor or priest and church staff, you schedule two types of focus groups: one for parents of students who are enrolled in school, and another for older teens and young adults who are out of school and living on their own while they work in your community (e.g., OSY, ages 17–21). Each focus group has 8–12 participants.
- At the beginning of each focus group, you give participants a one-page survey that lists all of the services provided through your office. They are asked to check all of the services they have used during the past year. In addition to asking for some basic demographic information (e.g., age, gender, OSY or parent; if parent, number and ages of children; etc.), the survey also includes two open-ended questions:
  - Of all the services we offer, which ones have been the most helpful to you? How have they been helpful to you?
  - Which services have been the least helpful? Why do you think they were not helpful?
- After everyone has completed the survey, you collect them (to analyze later) and begin the focus group. Over the course of 45 minutes to an hour, you ask the following questions:
  - After seeing our list of services, were there any that you did not know about? If so, what services are you just hearing about? Are these services you think you can use?
  - What other services do you think are needed in this community? Why are they needed?
- Each focus group has a facilitator who asks the questions, makes sure everyone gets a chance to contribute to the conversation, and generally keeps the discussion on track to end on time. Each group also has an observer who takes detailed notes and records the

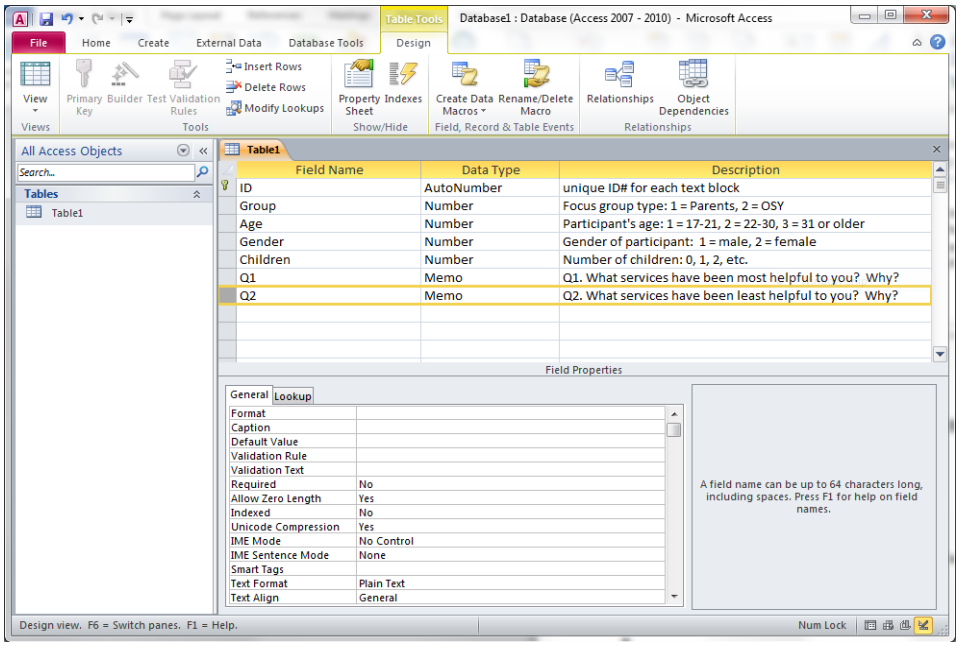
discussion (using a digital or tape recorder) in case parts of it need to be reviewed later for more details or longer quotes.

- When all of the focus groups are done, you will have a focus group dataset that includes responses from all or most participants to your two discussion questions. For coding purposes, you will need to separate each discussion into distinct chunks or text blocks. If we have a transcript of the entire focus group, we try to excerpt an individual's entire answer to a question to create one text block. If we are working from notes, we summarize a single individual's response from our notes to create a text block.
- In addition to the focus group dataset, you will also have a survey dataset with information about all of the participants, including open-ended responses with their thoughts about the most and least helpful services they have received.

Creating and coding text blocks from focus group data is complicated by the fact that participants may repeat themselves, or interrupt each other, or reconsider their answers. This does not make for nice, neat text blocks! We encourage you to try it and see what we mean. But for now, we will continue with the simpler example of creating text blocks from written responses to open-ended survey questions.

E.2.3 Creating Qualitative Data Analysis Forms in Access

When you are ready to analyze the surveys you collected before the focus groups, you can start by entering all of the data into an Access table. First, create a table in Design View.



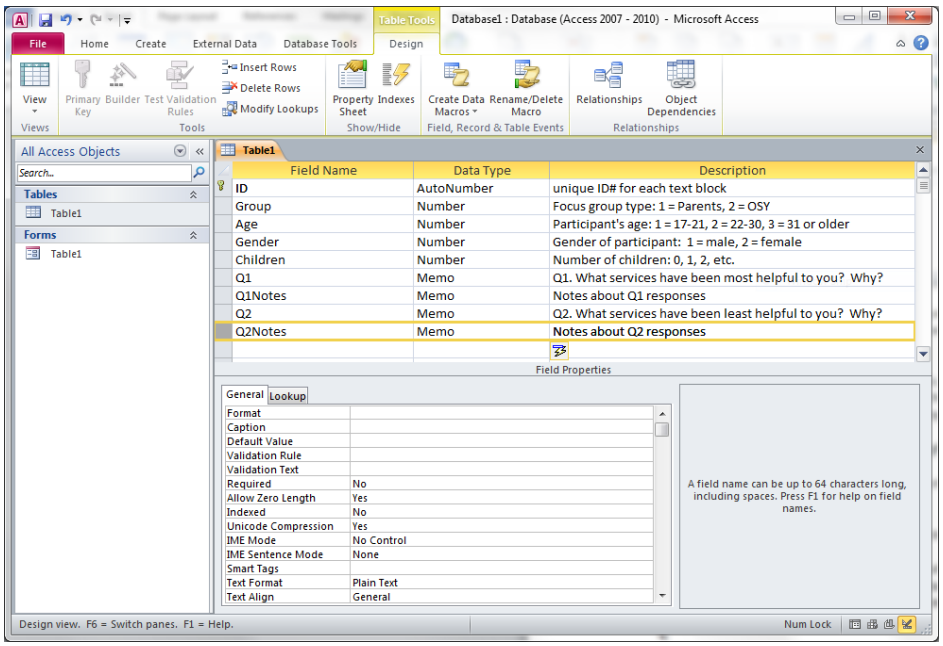
Next, enter your survey data. You can enter data directly into the table in DataSheet View, or you can create a form for your table (e.g., using the Form Wizard) and enter your data into the form.

No matter how we enter the survey data in the first place, we like to create separate forms just for analyzing answers to open-ended questions. We start with a very simple form and add to it as we identify different themes or categories in the survey responses.

Our qualitative data analysis forms include the following features:

- A unique survey number, so we can refer back to the original written response as needed
- A data entry "window" for the memo field that contains each text block in our dataset
- Another data entry window for the memo field where we will write notes about possible categories for each answer, the main idea, and quotable bits (e.g., specialized or colorful vocabulary, descriptive phrases, and sentences or passages that are typical for a particular category)
- Multiple "Yes/No" check boxes, one for each code (Level 1, Level 2, etc.) that we assign to the open-ended responses

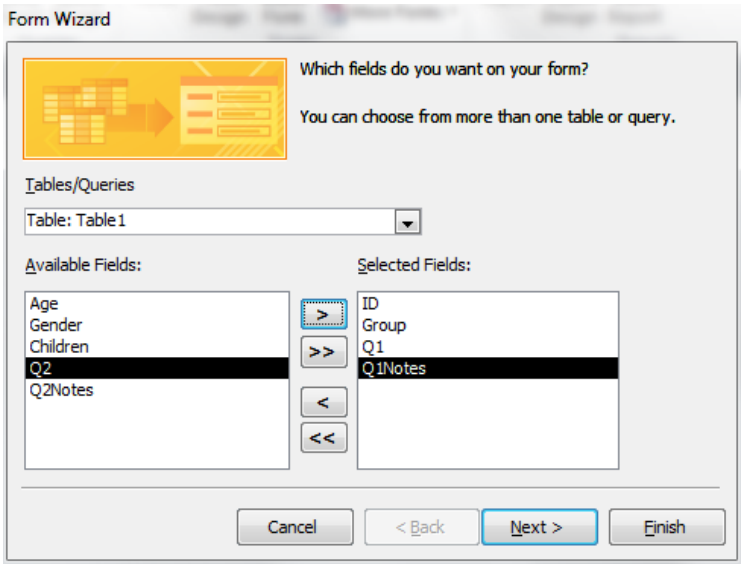
We always begin our analysis by reading through survey answers and taking notes as we go. So, before we start creating a form for our first question, add fields in the table for notes about the answers to each question. Now, save and close the table (Table 1).



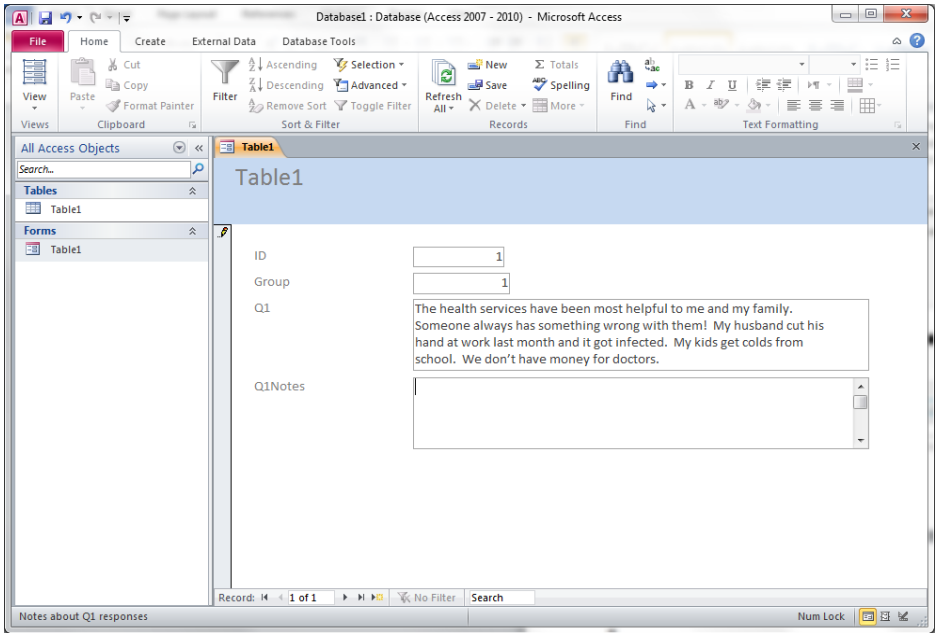
To begin creating a form for Question 1:

- Select Create > Form Wizard.
- In the Tables/Queries drop-down menu, select the table with your survey data (Table 1).
- Next, click on each variable you want to include in your analysis, including Q1.

- Click on the Field Name in the Available Fields window to select, then click the single ">" to move the variables you want into the Selected Fields window.



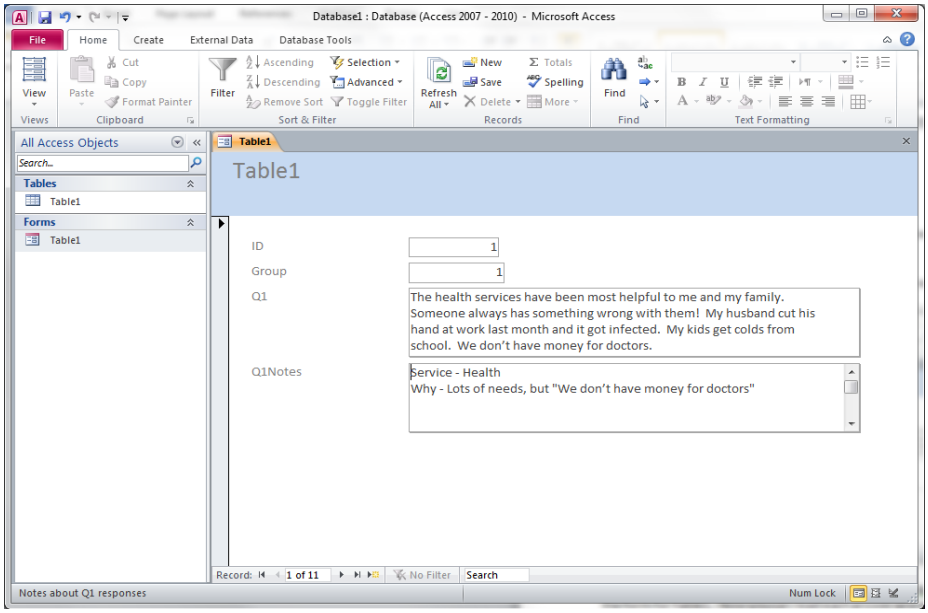
- Click Next > to select a pre-designed layout for your form. In this example, we will use the Columnar layout.
- Click Next > and rename the form, if you wish.
- Then click Finish to see what your form looks like so far.



E.2.4 Selecting Level 1 Coding Categories

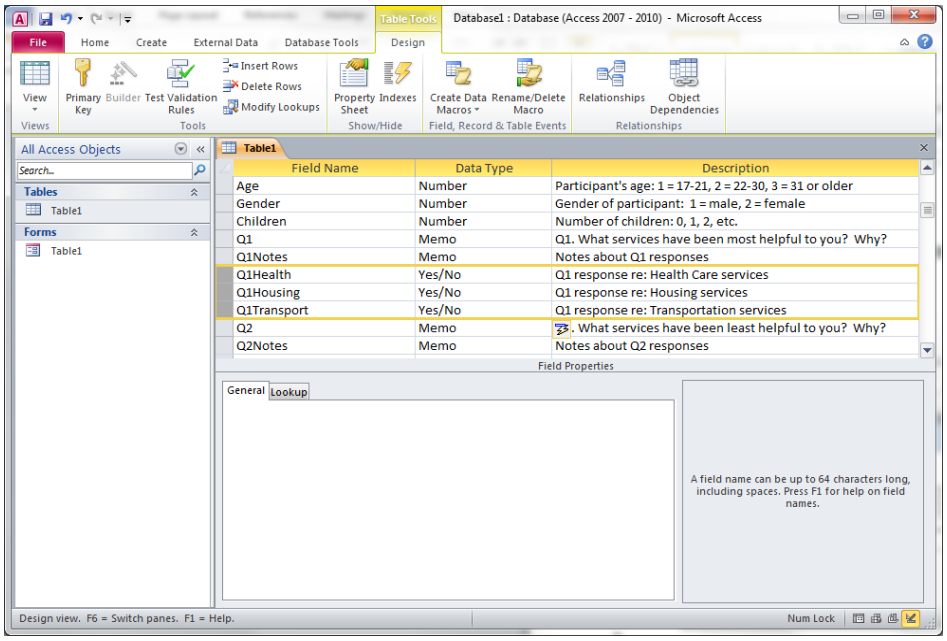
Notice that the answers to Question 1, which were already entered in Table 1, are now visible in the form for Table 1. Now you can read each answer and add some notes, while you get familiar with what participants wrote and begin to identify possible coding categories.

For example:



Eventually, you will read through and write notes for all of the answers to Question 1, or you may read many of the responses and find that you are not seeing any new categories. Now you can decide which broad categories you will use to organize the answers to Question 1.

Categories of services are the obvious choice for this question. All the answers you have read have had to do with health, housing, and transportation services. So, save and close the form, then go back to Table 1 to add fields for these categories.

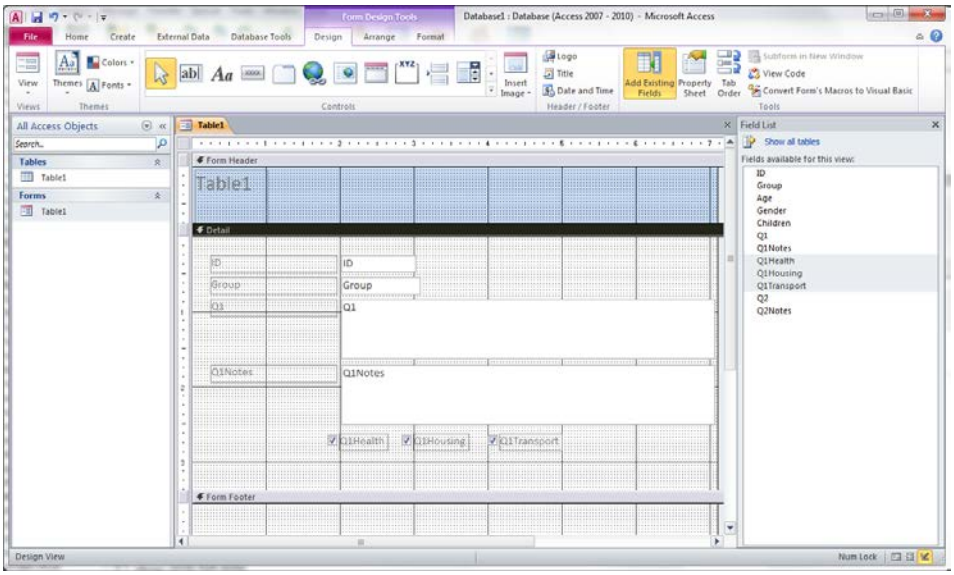


We just created a set of Level 1 codes! We can add more Level 1 codes for categories of services later if we run across some that we missed in our first pass through the data. All we have to do is save and close the data entry form, then open Table 1 in Design View to insert rows and add fields.

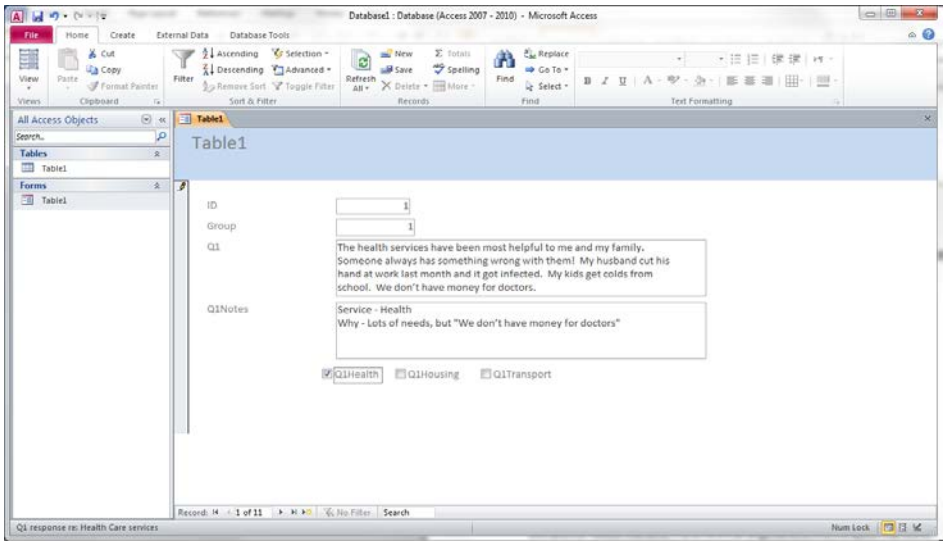
We use the "Yes/No" data type for all of our coding fields. Basically, they are simple check boxes. They are quick and easy to use as we read through all responses and select the category that is the best fit.

To add our Level 1 coding check boxes to the form, save and close Table 1. Then:

- Open the form for Table 1 and switch to Design View.
- Click on "Add Existing Fields" in the Toolbar at the top of the screen.
- Click each coding variable and drag it down to the form, which now appears in grid format.
- Place each coding variable check box wherever you like, as long as it does not overlap with another field on the form. Try different arrangements to find the layout that works best for you.



Save the new layout of your form and switch back to Form View. Once again, you will see the responses that have been entered into Table 1 as well as the notes you have already entered. Now you can just click on the box next to the appropriate coding category, and voilà – you have just coded an entry!

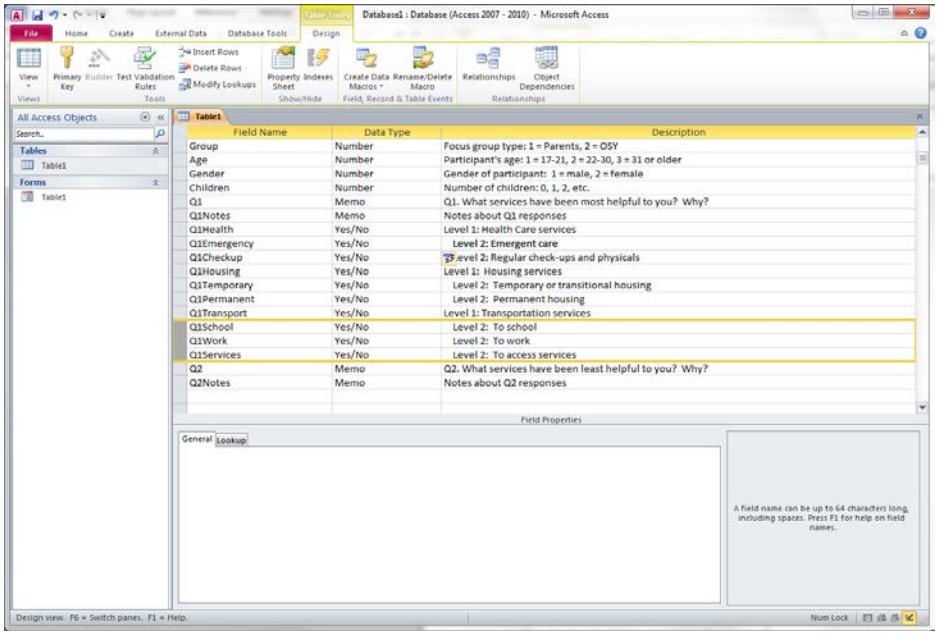


E.2.5 Selecting Level 2 Coding Categories

At some point, you will begin to see more narrow categories within your broad Level 1 categories. As you decide what these next categories are, you can add them to your table and data analysis form the same way you added the Level 1 coding categories and check boxes:

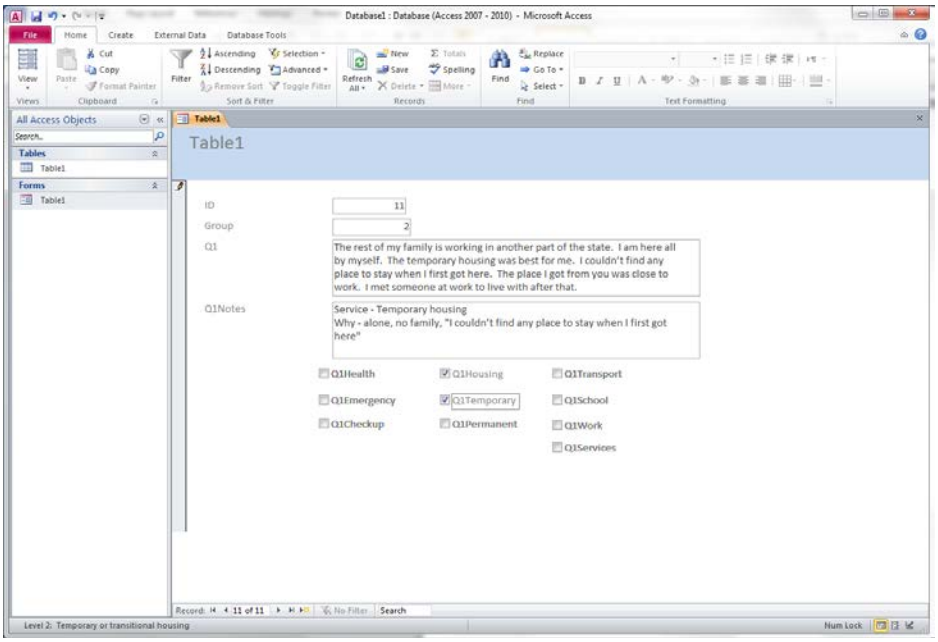
- Save and close the form.
- Open Table 1 in Design View.

- Insert rows, as needed.
- Add Field Names, select “Yes/No” data type, and write a brief description for each new coding category.
- Revise Level 1 descriptions as needed.



Then:

- Open the form for Table 1 and switch to Design View.
- Click on “Add Existing Fields” in the Toolbar at the top of the screen.
- Click each new Level 2 coding variable and drag it down to the form.
- Arrange Level 2 coding boxes in close proximity to their corresponding Level 1 coding boxes. Re-arrange Level 1 check boxes as needed.



E.2.6 Organizing Data in Categories

By using Access to create and apply codes to your data, you will have the flexibility to look at your data in a variety of ways. Once you have coded all of your data, you can sort it into coding categories for further analysis. You can do this by:

- Setting up queries for each coding category (Create > Query Wizard > Simply Query Wizard > OK, etc.); and
- Using queries to create reports with all of the responses that were coded into that category (Create > Report Wizard, etc.)

You may see interesting patterns or differences in why or how different services in the same category were especially helpful (or not helpful) to participants. From there you could decide to create Level 3 coding categories in order to include these patterns in your analysis.

You can also export the data in your check boxes to Excel and then run frequencies on these fields to create summary tables. We encourage experienced Access users to give this coding method a try, and we encourage new Access users to learn more!

Appendix E.3 Using Inferential Statistics

In *Section E.1 Analyzing Quantitative Data* of the *Program Evaluation Toolkit*, we explained how to use descriptive statistics to summarize information about migrant students, their participation in programs, their attitudes, their achievements, etc. Descriptive statistics such as frequencies, rates, means, and medians are relatively easy to calculate and display. Using descriptive statistics to compare two or more groups may reveal important differences or instructive patterns.

We often make programmatic decisions based on available descriptive information. However, the results obtained from descriptive statistics are at best suggestive and cannot be used to draw valid conclusions about programs. In order to make more definitive and convincing statements about program impacts, we analyze data using inferential statistics that test whether meaningful or significant differences have been achieved and whether these findings can be generalized beyond the specific group of people we studied.

> NOTE: The information provided in this Appendix is no substitute for the expertise of a qualified statistician who understands the complexities of inferential statistics. Nevertheless, we have included information about this topic for MEP directors and staff who are curious about how to go beyond merely describing the patterns in their data to testing for more definitive conclusions about these patterns.

The intention of this Appendix is to explain some basic concepts and provide examples for using specific inferential statistics to evaluate Migrant Education Program services and activities. Because we cannot adequately cover the complexity of this topic here, we have included links to several online resources (searchable textbooks, tutorials, and applets) at the end of this Appendix in case you want to learn more or refresh your memory.

E.3.1 Basic Concepts in Inferential Statistics

> NOTE: This section of the Appendix includes a number of technical terms. If you are not interested in the underlying principles of inferential statistics, you can skip this section. However, if you would like to learn more, we encourage you to use the terms that appear in italics to search the references provided at the end of this Appendix.

Inferential statistics are used to draw conclusions about a *population* of people based on information from a *sample* of that population.

For the purposes of statistical analysis, a population is the entire group of people about whom we are interested. In order to draw the strongest possible conclusions about the population of interest, we would collect and analyze relevant or interesting data from every person in that

population (e.g., data about race/ethnicity, gender, migrant status, education level, attitudes, behaviors, etc.)

When it is not possible to collect data from every subject in the population of interest, we collect data from a subset or sample of the population instead. For example, we might select a sample of migrant students from the high schools in our school district and use the information we collect from them to make general statements about the population of all migrant high school students in the district.

The appropriate use of inferential statistics involves collecting a *representative sample* of adequate size to detect outcomes if, in fact, they have been achieved. Traditional inferential statistics are based on the measurement of numerical characteristics of a sample, which in turn are assumed to be the same as the characteristics of the population from which the sample was drawn.

- Examples of numerical characteristics include measures of *central tendency* (mean, median, mode) and *spread* or *dispersion* of the data (variance, standard deviation).
- These numerical characteristics are also known as *parameters* and are calculated for continuous variables. *Parametric statistics* is the more traditional and purist branch of inferential statistics.
- *Non-parametric statistics*, which are not based on the numerical characteristics of a sample, can be used to analyze categorical and continuous variables. Instead of using measures of central tendency and dispersion, non-parametric statistics are based on the rank order of data values.

In order for parametric statistics to produce valid and unbiased estimates of population parameters, certain conditions must be met. A representative sample, specifically a *random sample* of the population, is a fundamental condition for correctly using parametric statistics.

- If the sample does not represent the population, then the statistics you calculate for the sample will not represent the true parameters of the population. They will only represent the sample.
- A *simple random sample* is a single sample drawn at random from a population (e.g., all migrant students in your district or in your state). Every subject in the population has an equal chance of being selected. This type of sample is representative of the population as a whole.
- A *stratified random sample* is comprised of simple random samples drawn from specific subgroups (or strata) in the population (e.g., Limited English Proficient (LEP) migrant students and non-LEP migrant students). This type of sample is representative of specific subgroups in the population.

The size and "shape" of a sample must also be considered when deciding whether you can correctly use parametric statistics.

- Random selection alone does not guarantee that a sample will be representative of the population. The size of the sample is also important. A sample of 10 books in a library of 3,000 is very unlikely to represent the depth and breadth of fiction vs. non-fiction or the range of genres and topics included in the entire collection. The larger the sample, the more representative it becomes.
- We use *power analysis* to determine how large a sample we need in order to be very confident (e.g., 95-99% confident) that any estimates we calculate from the sample will have small *margins of error* (e.g. 1-5%).
- Another condition for using parametric statistics is that the population from which you draw your sample must have a *normal distribution*. In a normal distribution, most of the observations in the sample are clustered fairly close to the mean or average value in the sample. The number of increasingly larger and smaller values tapers away on either side of the mean, which results in a bell-shaped curve when you graph the distribution.
- Fortunately, according to the *Central Limit Theorem*, a sample of at least 30 observations approximates a normal distribution enough to allow the use of parametric statistics. Nevertheless, we still strive for larger samples in order to increase confidence that the margins of error will be small.

E.3.2 Using Inferential Statistics in Program Evaluation

In the context of program evaluation, there are several common challenges to using parametric inferential statistics:

- It is rarely possible to randomly select people to participate in programs. There are ethical issues involved in deliberately withholding a potential benefit from people whose lives may improve with it. There are also practical barriers to identifying and recruiting every person who could benefit just so you can draw a random sample.
- Suppose you are able to randomly select participants. The simplest form of experimental design involves collecting baseline data, exposing participants to the program, and then collecting follow-up data to see whether participants achieved the desired outcome. Unfortunately, between the baseline and follow-up data collections (or the pre-test and the post-test), some participants will inevitably drop out of the program. So the random sample you had at the time of the pre-test is no longer random at the time of the post-test.
- Because of budget constraints and/or eligibility requirements for targeted interventions, activities and services may be available or appropriate for only a small group of participants. Small samples sizes (e.g., fewer than 30 subjects) are not uncommon in program evaluation.

The bottom line is that program evaluation data rarely meet the conditions needed to correctly use and draw conclusions from parametric statistics. Fortunately, these limitations are not fatal to the cause of program evaluation.

- The most common question asked in program evaluation is, "Did the program have the desired effect on the participants?" Therefore, the population of interest is all program participants. In a small program, you can collect data from every participant. If you need to collect a sample of participants in a large program, all participants (e.g., all members of the population) are either known to you or at least identifiable and therefore may be easier to sample at random.
- When you lose participants to attrition from the program or have a low response rate from your data collection activities, look for ways to compare the posttest or low-response sample to what you know about the pretest sample or program participants as a group. The point is to determine how similar the former is to the latter. If the groups continue to be similar on characteristics that matter in your program, then you have some basis for arguing that the posttest or low response sample is still representative of the original sample.
- Non-parametric statistics do not require the same conditions as parametric statistics in order to draw valid conclusions. They can be used to analyze small samples of categorical and continuous variables with non-normal or skewed distributions.

The following table provides some examples of evaluation questions and the appropriate tests for analyzing related data.

Table E.3.1 Examples of Evaluation Questions and Related Tests

Type of Analysis	Sample Evaluation Question	Parametric Statistic	Non-parametric Statistic
1. Compare means between two distinct (i.e., independent) groups.	Is the mean reading scale score for LEP migrant students in a language enrichment class significantly different from mean reading scale scores for LEP migrant students who did not take part in the class?	Two-sample t-test	Wilcoxon rank-sum test
2. Compare two measurements of a continuous variable taken from the same individual.	For a sample of migrant students in a summer enrichment program, was there a statistically significant difference between mean math pretest and posttest scores?	Paired t-test	Wilcoxon signed-ranks test
3. Compare means between three or more independent groups.	Comparing LEP migrant students in three supplementary language instruction groups (pull-out group instruction, one-on-one tutoring, summer language immersion program), are there statistically significant differences in their mean pretest English Language Proficiency scores? Are there significant differences in their mean	Analysis of variance (ANOVA)	Kruskal-Wallis test

Type of Analysis	Sample Evaluation Question	Parametric Statistic	Non-parametric Statistic
	posttest ELP scores?		
4. Estimate the degree of association between two continuous variables.	Is there a statistically significant association between the amount of time spent in supplementary language instruction and scores on an English Language Proficiency test?	Pearson coefficient of correlation	Spearman's rank correlation
5. Estimate the degree of association between two categorical variables.	Is there a statistically significant association between participation in specific non-instructional support programs (e.g., transportation, nutrition, medical care, counseling) and proficiency levels (I, II, III, IV; or I-II and III-IV) on the state reading and mathematics assessments?	[Not appropriate for categorical data]	Chi-squared test

NOTE: Table adapted from Hoskin, T. *Parametric and Nonparametric: Demystifying the Terms*

E.3.4 Online Resources about Inferential Statistics

In this Appendix, we tried to provide enough basic information about inferential statistics to convey the underlying importance of representativeness, sample size, and a normal distribution. Another important topic that we have not discussed is *hypothesis testing*. For more information about all of these topics, we encourage you to check out the following resources.

Online Statistics Education: A Multimedia Course of Study [http://onlinestatbook.com/]
HyperStat Online Statistics Textbook [http://davidmlane.com/hyperstat/]
StatSoft Electronic Statistics Textbook [http://www.statsoft.com/textbook/]
Sample size calculator [http://www.raosoft.com/samplesize.html]
Random number generator [http://stattrek.com/statistics/random-number-generator.aspx]