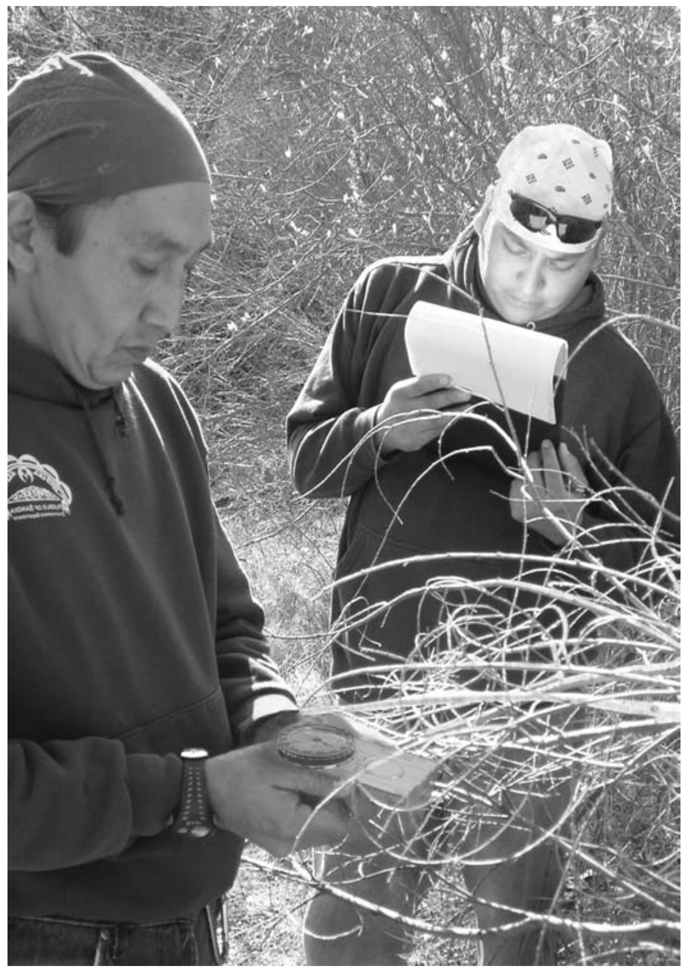
Handbook SIX

Analyzing and interpreting monitoring data



Collaborative Forest Restoration Program

January 2005

The multiparty monitoring handbook series

This multiparty monitoring handbook is part of a series of guides to monitoring collaborative forest restoration projects. The series was written specifically for projects funded through the USDA Forest Service's Collaborative Forest Restoration Program (CFRP). The Handbooks in the series are:

Handbook 1—What is multiparty monitoring?
Handbook 2—Developing a multiparty monitoring plan
Handbook 3—Budgeting for monitoring projects
Handbook 4—Monitoring ecological effects
Handbook 5—Monitoring social and economic effects
Handbook 6—Analyzing and interpreting monitoring data

Multiparty monitoring is required of all CFRP grantees; however, the methods and approaches presented in these workbooks are to serve as guides and references only. The specific methods are NOT required. Because there is a wide diversity of projects funded through the CFRP, many grantees will have different requirements for monitoring and/or monitoring assistance.

The content of these handbooks was largely conceived at a series of workshops held in 2003 that were sponsored by the following: Ecological Restoration Institute (ERI), Forest Trust, Four Corners Institute, National Forest Foundation, Pinchot Institute for Conservation, USDA Forest Service—Collaborative Forest Restoration Program.

These handbooks are updated periodically and the latest versions will be available on the Collaborative Forest Restoration Program Web site at www.fs.fed.us/r3/spf/cfrp/monitoring. For more information on this series, contact the Ecological Restoration Institute, Box 15017, Flagstaff AZ 86011-5017.

CFRP grantees are also eligible for multiparty monitoring training workshops and technical assistance from the CFRP monitoring team. This free service will be provided through September 2006. Call 866.614.8424 for details.

Handbook 6 authors: Kimberly Lowe, Ann Moote, Tori Derr, Melissa Savage, Jim Ranger-Moore, and Jesse Abrams. Handbook series authors/editors: Tori Derr, Ann Moote, Melissa Savage, Martha Schumann, Jesse Abrams, Laura McCarthy, and Kimberly Lowe.

Design, copy-edit, and production Joel Viers, ERI; cover photo Ann Moote, ERI, photos on page 47 ERI and Ann Moote. 01.06.05

This Handbook series is largely funded by the USDA Forest Service and is published by the Ecological Restoration Institute at Northern Arizona University

Table of contents

Appendices

List of Tables

List of Figures

Guide to this Handbook

Why analyze monitoring data?

The basic purpose of data analysis it to identify patterns of change in your indicator over time, and to evaluate these changes. Without doing some kind of analysis, it will be difficult for you to know the effect your project is actually having.

Data analysis methods

An indicator is something that can be measured over time to document changes in a specific condition.

Your data may consist of numbers, descriptions, or images. Numbers are *quantitative data* and are analyzed using mathematics. Descriptions and images are *qualitative data* and can be analyzed using tables, maps, or figures. For both qualitative and quantitative data there are four basic steps to data analysis, but the methods used at each step are different. In the boxes below you will find two routes for completing your data analysis, one for qualitative data and one for quantitative data. Depending on your indicators and the type of data you collected, you may only need to follow one route to complete your data analysis.

What kind of data do you have?

numbers = quantitative data

1. Organize data by combining all the measurements for each indicator into one table (page 7)
2. Analyze data by calculating averages, percentages, and other values (page 8)
3. Interpret analysis results by calculating change over time and comparing results to the original goals for each indicator (page 11)
4. Verify analysis results by looking for statistically significant values using the sign test (page 64) or by calculating the confidence interval (page 68)

words or images = qualitative data

1. Organize data by grouping similar kinds of information (coding) (page 32)
2. Analyze data using tables, maps, or figures (page 34)
3. Interpret analysis results by asking what they mean in terms of meeting your project goals (page 39)
4. Verify analysis results by comparing data collected from different sources or through different methods, examining alternative interpretations, looking for negative cases, or asking your data sources to review your results (page 41)

When to analyze your data

It may seem that data analysis is one of the last steps in the monitoring process, but we strongly suggest analyzing your data as soon as possible after you have collected it. This is desirable for a number of reasons. First of all, looking at your baseline data will help your group decide what level of change would represent success and set appropriate target values for your indicators. Secondly, if you analyze your baseline data before you begin your project you may identify additional questions that you want to monitor or even refine your treatment methods to better meet your goals. Later on, looking at changes in your indicators while the project is in progress will help your group decide if changes need to be made to the management prescriptions, or the management process, to better meet the project goals.

Where to get help

The data analysis techniques presented in this handbook are not difficult. Most of them can be easily done using little more than a calculator and scratch paper. However, you may want may want additional assistance analyzing your data.

The CFRP Multiparty Monitoring Technical Assistance Team is available through September 2006 to provide FREE assistance with monitoring and data analysis to any CFRP grantee. You can reach us by calling 866.614.8424.

This handbook may also be used by project partners working together with people who have experience with monitoring analysis. Is there someone in the community who is willing to help, perhaps someone who recently took a course in statistics? Nearby community colleges and universities usually have faculty members who can help. Assistance might also be found at local, state, or federal government natural resource or community development agencies.

Understanding your indicator data

For many of the ecological, social, and economic indicators listed in Handbook 4 (*Monitoring ecological effects*) and Handbook 5 (*Monitoring social and economic effects*), you will have collected numerical data. Data represented by numbers are also called *quantitative data*. Quantitative data are analyzed using mathematics. Some of the more common methods for analysis, including some basic statistical tests, are on pages 13–24 and in Appendix 5 of this handbook. Quantitative data analysis can seem intimidating at first, but the majority of the techniques presented in this handbook can be easily done using little more than a calculator and scratch paper.

Quantitative data is represented by numbers.

—

Qualitative data is represented by words or images.

You also may have gathered data through interviews, maps, or other methods that consist of words or pictures rather than numbers. These data are called *qualitative data*. Qualitative data can be analyzed in two ways. One way includes systematically reducing, organizing, interpreting, and verifying the data you have collected. This type of analysis does not use mathematics. Another way includes reducing, describing, and then giving numbers to qualitative categories. This type of analysis will give you both descriptive information as well as numbers (quantitative data) that can then be analyzed using quantitative methods. We will discuss how to do this later in this handbook, on pages 32–33.

Indicators that will produce quantitative data

Below, we list indicators (from Handbook 4 data sheets) that should be analyzed quantitatively. For instructions on analyzing these indicators, see pages 24–30 of this handbook.

- Bird and butterfly abundance and species composition
- Tree density and size
- Surface fuels
- Canopy cover
- Height from ground to tree crown
- Understory plant composition

- Understory cover
- Extent of bare soil
- Total number of workers employed by the project
- Amount paid to project workers
- Number of acres protected from fire through creation of defensible space, fuelbreaks, or other fuel reduction projects

Indicators that will produce qualitative data

You may also have gathered qualitative information through interviews, maps, photo points, or other methods that produce data more easily described with words or pictures. Indicators that can be monitored qualitatively are listed below. Instructions for analyzing qualitative data are covered in pages 31–47 of this handbook. Indicators that will produce qualitative data include:

- Landscape openings
- Extent that traditional forest users' knowledge and practices are used
- Extent that stakeholders previously in conflict are now working together on this project
- Quality and timeliness of communication among all project partners
- Existence and use of written organizational procedures, rules, and operational guidelines
- Existence and use of financial management tools or system

Indicators that can produce both quantitative and qualitative data

If you have gathered both descriptive and numerical data for a single indicator, you will want to do a combination of qualitative and quantitative data analysis. For example, you may have collected data on your project team's level of commitment to communication and group learning by asking each team member to estimate the number of hours they spend on various activities related to communication and group learning. This would give you quantitative data—number of hours. However, you may also have asked project participants to describe, in their own words, how different project activities contribute to group learning, or to give their perspectives on how communication is going overall. This would give you qualitative data—verbal descriptions from project participants. Indicators that can be monitored with some combination of quantitative and qualitative data are listed below.

- Classification of riparian plant community structure
- Number and diversity of wood products that can be processed locally
- Extent that project workers receive health benefits
- Extent that project workers can participate in family and community life
- Extent that project workers are trained to use and do use appropriate safety gear
- Opportunities for local families to recreate in the forest
- Availability of and access to medicinal, food, heating, or building materials from the forest
- Location of acres of defensible space created by the project in relation to areas considered to be at highest risk from wildfire
- Level of and commitment to communication and group learning
- Extent that different perspectives are represented on the project team and in project activities

See Appendix 6 for an example of presenting qualitative and quantitative data together.

Transforming qualitative data to quantitative data

Coding is a process of classifying data using terms ("codes") that describe the characteristic you are interested in measuring.

In some cases, particularly when your data is in the form of verbal responses from a survey or interviews, you may want to transform qualitative data into quantitative data by *coding* it. Once qualitative data has been coded, it can be transformed into numbers and analyzed using statistics. Coding is described in Step 2 under *Analyzing qualitative data*, page 32.

Target values for your indicators

For each indicator that will be monitored, your multiparty monitoring team should come to agreement on a target value or desired direction and level of change for that indicator. For example, you may have decided that you want to decrease tree density to 60 trees per acre on your project site, or that you want to increase understory cover by at least 20 percent. These values can then be used as benchmarks to help you decide whether or not you are achieving your goal.

It is difficult to prescribe target values for restoration project indicators in a general way. For one thing, all forest stands and all communities are different, and one restoration prescription would not fit all sites. Every restoration project team will define success in a different way, but many specific values for desired conditions will be similar to values in other projects. For example, many projects will want an increase in the number of jobs in the community and a decrease in the number of exotic plant species in the forest.

An example of a set of target values for ecological indicators is presented in Table 1 on the facing page. These are examples only. Your target values may vary, depending on the goals of your project. Some version of this set of values might be suitable in ponderosa pine forest restoration projects in the Southwest. Note that Table 1 shows sample target values for ecological indicators *one year after* treatment. Target values for different time periods, such as two or four years after treatment, may be different.

For some qualitative indicators, your targets may simply indicate the desired direction of change (e.g., "more communication between community members and agency staff"). It is also possible that you will not have a target value for an indicator. For example, your multiparty monitoring team may want to measure the number of sales resulting from different marketing mechanisms to compare approaches and select the ones that work best. In

Table 1 — Examples of targets for ecological indicators

Indicator	Sample targets (one year post-treatment)
Bird abundance and species composition	At least 10% increase in average abundance of native species
Butterfly abundance and species composition	At least 10% increase in average abundance of native species
Seedling density	40%–50% reduction in average density
Surface fuels	At least 70% reduction in average weight
Height from ground to tree crown	Average distance of eight feet or more
Understory plant composition	At least 10% increase in number of native species
Understory plant cover	At least 20% increase in percent cover
Extent of bare soil	No more than 30% bare soil
Permanent photo points	Increased grasses and other understory plants
Landscape openings	More small meadows between patches of trees
Classification of riparian plant community structure	More mature cottonwood trees with native understory (Type 1 and Type 2 stands)

this case, you will be comparing the number of sales from each approach against the number of sales from other approaches, not against a pre-determined target value. Examples of target values for some of the social and economic indicators described in Handbook 5 are presented in Table 2.

Table 2 — Examples of targets for social and economic indicators

Indicator	Sample targets (one year post-treatment)
Amount and diversity of wood products that can be processed locally	Have two products, want three (50% increase) Make vigas and latillas, also want to produce wood pellets
Number of sales resulting from specific marketing mechanisms	This indicator may be used to compare and evaluate different marketing methods, in which case there would be no target value
Number and type of trainings completed by project workers	All employees receive safety training, ten employees receive prescribed fire Red Card training, and all monitoring team members receive plant identification training.
Types of equipment used	Mechanized, low-impact equipment
Extent that project workers can participate in family and community life	Fewer than one night per week spent away from home because of work Fewer than one weekend per month worked At least one community activity attended each month Able to regularly collect medicinal, food, heating, or building materials from the forest
Number of acres protected from fire through creation of defensible space, fuelbreaks, or other fuel reduction projects	750 acres
Number of different perspectives or stakeholders represented on the project team and in project activities	Six (project employees, traditional forest users, community leaders, environmental organizations, agencies, and general community)
Level of commitment to communication and group learning	Evidence of increased quality of communication among project partners

Analyzing quantitative data

Quantitative data analysis includes four steps:

Step 1 – Organize your data
Step 2 – Analyze your data
Step 3 – Interpret your data
Step 4 – Validate your data

Materials needed for quantitative data analysis:

Necessary
Pencil
Paper
Calculator

Desirable
Computer
Software (e.g., Microsoft Excel)

The first two steps—organization and analysis—are important to accurately interpret your monitoring data. The third step—interpretation—will tell you what changes have occurred with your indicators. The fourth step—verification—will help you know whether your data are accurate enough to base management decisions on them. Data verification is also important when there are different points of view about the best way to reach your goals or about the potential outcomes of your project.

Step 1– Organize your data

Organizing in one place all of the data your group has collected is the first step in the analysis process. If your group has used the Handbook 4 data sheets for the ecological indicators, it is likely you had to make multiple copies of the data sheets for each indicator. You now need to combine all of the data for each indicator from every data sheet into a single table. If you have access to a computer, it is best to put this information into an electronic spreadsheet (see Appendix 3 for help in making a spreadsheet).

Make sure all data sheets for each indicator are stored together, perhaps in a folder or binder, so you have easy access to them once your group starts adding the numbers together.

Similarly, for social and economic indicators, you will also want to compile all of your data into a single table or spreadsheet. For example, if you are measuring the number, type and duration of educational and training opportunities for youth, you might create something similar to Table 3 (next page).

After you have compiled your data, check to be sure that you did not make an error when you copied the numbers from your data sheets. Common errors include writing a number in the wrong row or column, misplacing a decimal point (e.g., 2.45 has a very

Table 3 — Example of compiled data

Type of training	Length of training (days)	How many times offered per year?
Prescribed fire (Red Card)	3	2
Chainsaw safety	1	4
Plant identification	2	1

different meaning than 24.5), or transposing numbers (e.g., 46 when you mean 64), and accidentally skipping a line on the data sheet so that values collected for one specimen are actually written on the line for another.

If a data point is clearly incorrect and you do not know the correct value, you will need to throw that data point out. If this happens, you will need to remove that data point from every data sheet, even if it has been accurately recorded in the past. This is also true if you have a missing data point on one of your datasheets but have that point on all of the other datasheets. Do not guess about data or make up numbers, it is better just to remove it from the analysis. Always be careful when recording data onto the datasheet and when entering them into the computer.

One individual value within a larger data set is a data point, for example, 26.6;

Plot #	Species	Diameter	BA
1	PP	17.9	26.6
2	PP	21.1	23.5
3	aspen	5.5	6.3

all of the data in this table makes up the data set.

Step 2 – Analyze your data

Quantitative data is analyzed using mathematical calculations. To know which calculations are appropriate for analyzing your indicator data, you must first know what kind of data you have.

Continuous versus nominal data

You'll notice some data were collected by taking actual measurements in inches or feet, such as diameter at breast height. These are called *continuous data*. Other data you have collected will just be tally marks, for example, indicating how many of each species you counted in a particular plot. These are called *nominal data*. Qualitative data that you have transformed into quantitative data (pages 32–33) are nominal.

It is important to understand the difference between continuous data and nominal data, because this is what determines which statistical test to use. You can do many more types of analysis with continuous data than you can with nominal data; however analysis of nominal data can also provide valuable results. Look at the list below to see which type of analysis is right for the type of data you have collected. Later in the handbook we will walk through the step-by-step process of doing each of these tests. You will use these statistical calculations frequently during the analysis process.

Statistics that can be used for:

continuous data	**nominal data**
• Average	• Frequency distribution
• Median	• Percentage
• Range	
• Standard deviation	
• Percentage	

Skewed data

Sometimes you will collect continuous data with a lot of numbers that are big or a lot of numbers that are small. These are called skewed data. This is very common with ecological data. Your data are less likely to be skewed if your sample size is large (over 30 individuals).

Some statistical tests are more appropriate than others for skewed data. For example, to find the average of a skewed data set, it is better to calculate a median than a mean. These tests are discussed in the following pages.

Types of calculations

Mean: Calculating the mean, or average value, for your data set gives you an idea of the "typical" value for that indicator. You will probably want to calculate a mean for each of your indicators. Mean is calculated only for continuous data.

Median: The median is another way of measuring average, in this case representing the "middle" point in an ordered data set. If your data are skewed, you should calculate the median, not the mean. Median is calculated only for continuous data.

Range: Range is the maximum spread of your data, from the lowest value to the highest value. Calculating range is a quick way to determine how variable your data are. Range is calculated only for continuous data.

Standard deviation: Standard deviation is another way to measure the spread of all of the data points in your sample. While range measures the *maximum* spread of your data points, the standard deviation tells you their *average* spread. A small standard deviation means there is not a lot of variability within your sample—meaning most of the indicators you measured in that sample are very similar—while a large standard deviation indicates there is a large variability within your sample—meaning the indicators you measured in that sample are very different. You will also need to calculate standard deviation in order to calculate the standard error of your data set. Standard deviation is calculated only for continuous data.

Frequency: Frequency is a measure of how often a certain response or value is found in your data. For example, you may want to know whether ponderosa pine trees are the dominant tree species on your site. To answer this kind of question you will want to count the *frequency* that you found ponderosa pine among all of the tree species you observed. Frequency is calculated only for nominal data.

Frequency answers the question "how many ponderosa pine trees did we find?"

—

Percentage answers the question "what proportion of all the trees we recorded were ponderosa pine trees?"

Frequency distribution: A frequency distribution is a way of showing how many times each value was observed. Frequency distributions are only calculated for nominal data. Creating a frequency distribution is a quick way to determine whether or not your data are skewed.

Percentage: Percentage is a way to determine how often a certain response or data value was observed compared to the total number of observations. Percent change over time is often used to interpret indicator data. Percentages can be calculated for both continuous and nominal data.

Explanations of how you calculate each of these values are provided in the section, *How to do the calculations* on page 13–24. In the section, *How to analyze your indicator data* (page 24–30, you will find specific instructions for analyzing each of the indicators presented in Handbooks 4 and 5.

Step 3 – Interpret your data

Most monitoring consists of comparing two sets of data that were collected at different points in time. This means that you measured the *same indicators*, using the *exact same methods*, at two different times. For instance, you probably collected baseline data before the project began, and then repeated your measurements the next year, or after your project was completed.

By doing the same calculations on your initial baseline data and on the data you collected at a later time, you can identify changes in your indicators over time. Look for patterns of change within your data set for each indicator. Have the average values or percentages gone up or down? Is this what you were expecting? Compare your results to the target values you set at the start of your project (pages 4–6).

Most target objectives state that a specific number or a range of specific numbers is acceptable. Look at the actual change you measured and compare it to your target to decide whether or not you are achieving your goal.

For example, say your community wants the forest canopy cover to be reduced by at least 20 percent but no more than 40 percent. If, after restoration treatments, a canopy cover reduction of 50 percent is found, then the partners should revisit the thinning prescription and revise it so that fewer trees are cut in the futures. Future thinning is then more likely to succeed in reducing canopy cover to the desired level.

If you only have pre-treatment data, you can still interpret your statistics and use that information to adapt your management protocol in order to better meet your project goals. Analysis of baseline data can be very helpful for adapting your management prescriptions and setting future monitoring goals. Looking at baseline data early in the project will help your group set appropriate target values for your indicators. Instructions for analyzing baseline data alone are provided on pages 23–24.

Step 4 – Verify your data

In most case, you will have collected data on a sample of your study area (for example, trees measured within sample plots), not the whole population (for example, all the trees contained within a your treatment area). Data verification is a way to determine if the change you observed in the sample is representative of change in the entire *population*.

Mathematical data verification only works when *samples* are representative of the population. Otherwise the data are said to be *non-representative* and cannot be used to draw conclusions about the population. The best way to assure a representative sample is to use a random sampling method and to collect a large number of samples (at least 30). Random sampling is discussed in Handbooks 4 and 5.

Verifying your data will help you know whether your data are accurate enough to base management decisions on them. Data verification is also important when there are different points of view about the best way to reach your goals or about the potential outcomes of your project.

Statistical methods such as the *sign test* (Appendix 5) can tell you whether the changes you observed are statistically significant. Methods such as constructing *confidence intervals* (Appendix 5) can give you a range of values that is likely to contain the true mean of the population you are sampling from. While the methods described in Appendix 5 can be completed using a calculator or computer spreadsheet program, many people will want to consult a statistician for help with data verification.

How to do the mathematical calculations for continuous data

In this section we will explain how to calculate the average (mean and median) value for your indicator, the range and standard deviation of your data, percentages, and change over time for

Table 4 — Mock data for Data Sheet I (partial sheet shown)

Adult living trees larger than 5 inches in diameter		
Species	Diameter (inches)	Height to tree crown (feet)
DF	12.1	3.1
PP	11.2	2.3
JN	5.3	2.1
JN	6.2	2.5
PP	13.5	6.4
PP	12.1	6.8

Table 4 uses the following abbreviations to record tree species:
DF = Douglas fir, PP = ponderosa pine, JN = juniper

continuous data. Remember, continuous data means that each of your data points is a measurement, such as inches or feet—it is not a count, such as how many species are on each plot. We show examples of each calculation using mock data for Data Sheet I from Handbook 4, *Density and size of live and dead trees*, and *Height from ground to tree crown* (Table 4). Mark the previous page because we will use the data in this table for many examples throughout the handbook.

If you are using a calculator, the following steps will guide you through the process of calculating these statistics. All of these tests can also be done using computer spreadsheets (directions provided in Appendix 3.)

Mean

To calculate a mean, add all of the individuals and then divide by the total number of individuals.

Example of calculating mean diameter using data from Table 4:

(1) 12.1 + 11.2 + 5.3 + 6.2 + 13.5 + 12.1 = 60.4

(2) 60.4/6 (total number of individuals) = 10.067

(3) Answer: mean = 10.07″

Explanation: The mean diameter at breast height for all of the trees sampled is 10.07 inches.

Use mean to calculate the average unless your data are skewed.

—

If your data are skewed, use median to calculate the average.

Median

To calculate the median, simply list all of the data points in numerical order, from smallest to largest. Then select the middle-most number. If you have an even number of data points, add the two middle-most numbers together and divide that total by two to find the median.

Example of calculating median diameter using data from Table 4:

(1) Ordered list of values: 5.3, 6.2, 11.2, 12.1, 12.1, 13.5

(2) Since we have an even number of data points, we will need to add the two middle-most numbers together and then divide that total by two:

11.2 + 12.1 = 23.3/2 = 11.65

(3) Answer: median = 11.65″

Explanation: The median value for diameter at breast height is 11.65 inches. Notice that this value is larger than the mean. That is because there were more large numbers than small numbers in this data set.

Range

Range tells you the maximum spread of your data, from the lowest value to the highest.

—

Standard deviation (page 17) gives you the average spread of all data points in your sample.

To calculate the range of a data set, subtract the lowest value from the highest value for all data points.

Example of calculating range of diameters using data from Table 4:

(1) 13.5 - 5.3 = 8.2

(2) Answer: range = 8.2″

Explanation: The difference between the largest tree diameter and the smallest tree diameter is 8.2 inches.

Percentage

You can calculate the percentage of your data set represented by a particular value or response using the following steps:

(1) Count the number of times you recorded a value or response.

(2) Divide that number by the total number of data points in the set. This will give you a number less than or equal to one (e.g., 0.50).

(3) Multiply that number by 100 to get percent (e.g., 0.50 x 100 = 50%).

Example of calculating percent species composition using data from Table 4:

Using the data from Table 4 on page13, we can determine the tree species composition using the steps described above.

(1) From Table 4 we see there were one Douglas fir, three ponderosa pine, and two juniper trees.

(2) There were six trees in the data set. Therefore, we divide the frequency of each species by six and then multiply by 100 to get the percentage.

a. One Douglas fir tree/6 trees in plot = 0.17 x 100 = 17%

b. Three ponderosa pine trees/6 trees in plot = 0.50 x 100 = 50%

c. Two juniper trees/6 trees in plot = 0.33 x 100 = 33%

Explanation: The tree composition on our plot is 17% Douglas fir, 50% ponderosa pine, and 33% juniper.

Example of calculating percentage of survey responses:

Finding percentages for closed-ended survey questions is very similar to finding percentages of species composition (discussed above). Basically, you are just looking for a part of the whole.

For example assume that you had the following data from a survey of project workers:

Monthly pay	Employees
Less than $500	2
$500–$100	4
$1001–$2000	11
$2001–$3000	2
More than $3000	1
Total number of employees	20

Suppose you are interested in knowing what percentage of your project personnel make over $1000 per month. Using the given data, you could calculate the following:

(number of employees paid >$1000 per month)/(total number of project employees) = 14/20 = 0.7 x 100% = 70%

Explanation: From these results we see that 70% of the people working on the project earn over $1000 per month.

Change over time

Monitoring frequently involves calculating change in a specific indicator between two time intervals such as before treatment and after. Most often, this is expressed as percent change over time. For example, assume we measure canopy cover before treatment and we see we have 80 percent cover of our plot. We measure it five years after the treatments and we see we have 64 percent cover on our plot. We can calculate the percent change between two time intervals using the following steps:

(1) Calculate the percent canopy cover before treatment, as instructed above (let us assume it is 80%)

(2) Calculate the percent canopy cover after treatment (assume it is 64%)

(3) Subtract the pre-treatment percent from the post-treatment percent:

64% - 80% = -16% difference in treatment

(4) Divide the difference between treatments by the percent canopy cover before treatment to get the percent change:

Percent change = (-16/80) x 100% = - 20%

Explanation: From these results we see there was a 20% reduction in canopy cover between the pre-treatment and post-treatment measurements.

Standard deviation

Calculating standard deviation is easily done using a computer program that can do statistics, such as *Microsoft Excel* (see Appendix

3 for help). If you don't have access to a computer, standard deviation can also be calculated by hand. To calculate standard deviation by hand, make a table with four columns (as seen in Table 5 below) and then work through the following steps:

(1) List all of the data points in column A

(2) Put the average of the data points in every row in column B.

(3) Subtract the average from each data point (column A - column B), and record the difference in column C.

(4) Square each value in column C and place the result in column D (squaring a number is multiplying it by itself).

(5) Add all of the values found in column D to get the total *sum of squares*.

(6) Count the total number of individuals in the sample (call this total "n"). Now subtract one from that number (n - 1). Always subtract one, no matter how large your sample size is.

(7) Divide the sum of squares by the value calculated in step 6. This is the variance for this set of data.

Table 5 — Steps for calculating standard deviation

column A (data points)	column B (average of the data points)	column C (column A - column B)	column D (values in column C)^2
	all values will be the same		
		sum of the squares (sum of column D)	
		sum of squares/(n-1)	
		standard deviation = $\sqrt{\text{sum of the squares}/(n-1)}$	

(8) Take the square root of the result from step 7 (the variance) to find the standard deviation.

Example of calculating standard deviation using data from Table 4:

To calculate the standard deviation of the tree diameters shown in Table 4, we created Table 6 with four columns, as described above (Table 5).

Table 6 — Example of calculating standard deviation

column A (data points from Table 4, second column)	column B (average of the data points)	column C (column A - column B)	column D (values in column C)^2
12.1	10.07	2.03	4.12
11.2	10.07	1.13	1.28
5.3	10.07	-4.77	22.75
6.2	10.07	-3.87	14.98
13.5	10.07	3.43	11.76
12.1	10.07	2.03	4.12
		sum of the squares (sum of column D)	59.01
		sum of the squares/(n-1)	59.01/5 = 11.80
		standard deviation = $\sqrt{\text{sum of the squares/(n-1)}}$	3.435

(1) The numbers in column A are the same as those in the second column of Table 4, and represent the diameter at breast height of each tree.

(2) Column B shows the average diameter at breast height that we calculated on page 14.

(3) Column C shows how much each tree diameter listed in column A differs from the average tree diameter shown in column B.

(4) Column D lists the values in column C, squared.

(5) The sum of the squares = 4.12 + 1.28 + 22.75 + 14.98 + 11.76 + 4.12 = 59.01

(6) In this case we have six individuals in the sample: 6 - 1 = 5.

(7) $\frac{\text{Sum of squares}}{(n-1)} = \frac{59.01}{5} = 11.80$ = diameter variance

(8) $\sqrt{11.80} = 3.435$

(9) Answer: standard deviation = 3.44 (by rounding up).

Explanation: Most of the data—about two-thirds of it—will fall within one standard deviation from the mean. In our example, the majority of the tree diameters measured are within 3.44 inches of the mean (or 10.07 +/- 3.44). This means there is not a lot of variability within your data.

How to do the mathematical calculations for nominal data

In this section we will explain how to calculate the frequency distribution, percentages, and change over time for nominal data. You cannot calculate mean, median, range, or standard deviation for nominal data, because your data is simply a count, such as number of species, not a value, like feet or inches. We show

Table 7 – Mock data for Data Sheet E (partial sheet shown)

Seedlings		
Species	Number of individuals (tally marks)	Total
DF	III	3
PP	II	2
JN	I	1
AJ	I	1
GO	III	3

Table 7 uses the following abbreviations to record tree species:
DF = Douglas fir, PP = ponderosa pine, JN = juniper,
AJ = alligator juniper, GO = Gambel oak

examples of each calculation using mock data for *Data Sheet E: Seedling density* (Table 7).

If you are using a calculator, the following steps will guide you through the process of calculating these statistics. All of these tests can also be done in computer spreadsheets (directions provided in Appendix 3.)

Frequency distribution

Once your group has identified what particular species of trees, plants, or animals are in your plots, it will be useful and easy to count how many individuals trees, plants or animals there actually are of each species. This count of individuals of each type is a frequency distribution. It is basically just counting how often something occurs.

Using mock data from Table 7, previous page, we can make the following frequency distribution by looking at the total number of each species in column C:

Species	Frequency
DF	3
PP	2
JN	1
AJ	1
GO	3

When you are making a frequency distribution, it will be important to make sure you are counting all of the data from each data sheet that relates to that indicator. As we noted earlier, it is likely your group will need to make copies of the datasheets, so just be sure you have all of those copies handy when you actually start counting how many individuals there are within each species. Once you have made the frequency distribution, it is very easy to turn this data into a bar chart (see Appendix 4). Bar charts are a useful way to display data.

Percentage

Calculating percentages with nominal data is exactly the same as for continuous data. You can calculate the percentage of your data set represented by a particular value or response using the following steps:

(1) Count the number of times you recorded a value or response (the frequency).

(2) Divide that number by the total number of data points in the set. This will give you a number less than or equal to one (e.g., 0.50).

(3) Multiply that number by 100 to get percent (e.g., 0.50 x 100 = 50%).

Using the data from Table 7, we can determine the tree species composition using the steps described above.

(3) From the frequency distribution table, we see there were three Douglas fir, two ponderosa pine, one juniper tree, one alligator juniper and three Gambel oak.

(4) There were ten trees in the data set. Therefore, we divide the frequency of each species by ten:

a. Three Douglas fir/10 trees in plot = 0.30 x 100 = 30%
b. Two ponderosa pine/10 trees in plot = 0.20 x 100 = 20%
c. One juniper/10 trees in plot = 0.10 x 100 = 10%
d. One alligator juniper/10 trees in the plot = 0.10 x 100 = 10%
e. Three Gambel oak/10 trees in the plot = 0.30 x 100 = 30%

<u>Explanation</u>: The tree composition on our plot is 30% Douglas fir, 20% ponderosa pine, 10% juniper, 10% alligator juniper, and 30% Gambel oak.

Change over time

The procedure for finding percent change over time for nominal data is identical to the method used for continuous data (page 17).

Interpreting baseline data alone

If you only have pre-treatment data, you can still interpret your statistics and use that information to adapt your management protocol in order to better meet your project goals. Analysis of existing data can be very helpful for adapting your management prescriptions and setting future monitoring goals.

For example, we know how to calculate the percent change in an indicator value (e.g., tree density) between two times using baseline data and then data collected after baseline:

(1) Calculate the percent canopy cover before treatment, as instructed above (let us assume it is 80%)

(2) Calculate the percent canopy cover after treatment (assume it is 64%)

(3) Subtract the pre-treatment percent from the post-treatment percent:

64% - 80% = -16% difference in treatment

(4) Divide the difference in treatments by the percent canopy cover before treatment to get the percent change:

Percent change = (-16/80) x 100% = -20%

Explanation: From these results we see there was a 20% reduction in canopy cover between the pre-treatment and post-treatment measurements. However, if the group has collected baseline data and already selected a specific tree density as one of its project goals, you could work this equation in the opposite direction to determine the exact difference the treatment must make to meet this specific goal. Knowing the desired result of your treatment can be an asset when planning treatment strategies or determining how to allot limited funds among multiple projects.

For example, if your group would like to see tree density reduced by 30 percent:

(1) Calculate the tree density before treatment (let us assume average tree density across all your plots is 80%).

(2) Determine the desired change between pre-treatment and post-treatment (let us assume it is 30%).

(3) Multiply the before-treatment value by the desired percent change: 0.80 x 0.30 = 0.24

(4) Subtract the value obtained in step 3 from the pre-treatment percent to find what the post-treatment values must be to meet the monitoring goal of a 30% change: 0.80 - 0.24 = 0.56 x 100 = 56% canopy cover

Explanation: To meet the monitoring goal of a 30% reduction in tree density, the monitoring group must bring the post-treatment tree density value from 80% to 56%.

How to analyze your indicator data

You may be wondering how much data analysis is “enough” to have useful information. This is a difficult question to answer because it depends on how your group intends to use the monitoring data. The most important part of the analysis process is collecting good data—meaning large sample sizes and data that were selected randomly. With this type of data, any of the statistical tests will produce valuable information. If your sample sizes are small or your data are skewed you can still calculate several of the statistics (e.g., mean, standard deviation, percent, change over time), but you will not be able to verify your data using confidence intervals. Below you will find step-by-step instructions for analyzing quantitative data for indicators discussed in Handbooks 4 and 5. Page numbers in parentheses direct you to instructions for performing the specific calculation. See pages 44–47 for examples of presenting results.

Ecological indicator data

Bird abundance and species composition:

(1) From your data sheets, you have a list of bird species. From this list you can count the abundance of birds of each species.

(2) The first measurement, bird species composition, is simply a list of birds found in the area. Make two species composition lists, one for birds identified in the breeding season and one for the non-breeding season. Within these two categories, separate native from non-native bird species.

(3) Use a frequency distribution to see the abundance of bird species (page 21). Track the change in bird species composition by counting the total number of species at each time interval those data are collected

(4) Calculate the percent of each species at each time interval (page 22).

(5) For species of concern, such as rare or problematic species, you may also want to calculate percent change in abundance between time periods (e.g., before and after project) (page 17).

Butterfly abundance and species composition:

(1) You will have two kinds of data on butterflies. You will have a list of butterfly species, and a series of abundance values for each species.

(2) The first measurement, butterfly species composition, needs no calculation. It is simply a list of butterflies found in the area.

(3) Use a frequency distribution to see the abundance of butterfly species (page 21). Track the change in butterfly species composition by counting the total number of species at each time interval that data is collected

(4) You can also calculate percent change in abundance over time, for each species, for all species, or for a select group of interest (page 17).

Seedling density:

(1) The data you collected are the number of seedlings found in each transect.

(2) Use a frequency distribution to see the abundance of tree species (page 21). Track the change in seedling density by counting the total number of species at each time interval that data are collected

(3) Calculate the pre-treatment (mean or median) number of seedlings across all transects for each species (page 14).

(4) If desired, you can also calculate the percentage of seedling density for each species across all transects (page 15).

(5) Calculate the percent change in total density of all seedlings between two time periods (page 17).

Surface fuels:

(1) From your data sheets, you have information on the amount of woody fuel by size category for two sampled points on each transect.

(2) Combine the information from these data sheets and enter the totals into the worksheet shown in Appendix 2.

(3) Calculate the amount of dead downed fuel in each size class using the methods given in Appendix 2.

(4) Calculate the total amount of downed dead fuel in the sampled area using the methods in Appendix 2.

(5) Calculate the percentage of wood fuel in each size class (page 15).

(6) Calculate the percent change in fuel load over time (page 17).

Extent of canopy cover:

(1) From your data sheets, you have a series of values for the number of points with canopy cover overhead for each transect.

(2) For each transect, find the percent canopy cover (Handbook 4, page 32).

(3) Calculate the average (mean or median) percent canopy cover across all of the transects (page 14).

(4) Calculate percent change over time in canopy cover (page 17).

Classification of riparian plant community structure:

(1) Find the percentage of each structure type on every transect (page 22).

(2) Find the percent change in each structure type between two time intervals for every transect (page 17).

Density and size of live and dead trees:

(1) From your data sheets, you have the number of live and dead trees per plot, and the diameter in inches of those trees.

(2) Calculate the average (mean or median) number of trees per plot, keeping track of live and dead trees separately (page 14).

(3) Calculate separate densities for live and dead trees per acre by using the area conversion shown in Appendix 1.

(4) Use a frequency distribution to see the abundance of tree species (page 21). Track the change in tree species composition by counting the total number of species at each time interval that data is collected

(5) Calculate the percent of each species (page 22).

(6) Calculate the percent change in species between two time intervals (page 17).

(7) Average size of trees (i.e., average tree diameter) can be calculated using the same methods. Calculate either the mean or median tree size and range of tree sizes (in diameter) for all sampled plots (page 14).

(8) Repeat the same steps for dead trees.

Height from ground to tree crown:

(1) From your data sheets, you have a series of distances, in feet, from the ground to the crown.

(2) For each plot, calculate the average (mean or median) distance and the range of distances from the ground to the bottom of the crown for all trees sampled (page 14).

(3) Calculate the average (mean or median) distance and range of distances from ground to crown across all plots.

Understory cover and extent of bare soil:

(1) For each cover type (grasses, forbs, litter and bare soil), calculate the average (mean) cover and the range in cover for all of the small plots in a transect (page 14).

(2) To calculate the percent cover of vegetation, add the percent cover values for grasses and forbs.

(3) For each cover type, calculate percent change in average cover over time (page 17).

Understory plant species composition:

(1) From your data sheets, you will have lists of plant species and the abundance of each species.

(2) Use a frequency distribution to see the abundance of each plant species (page 21). Track the change in plant species composition by counting the total number of species at each time interval that data is collected. It is also important to keep track of native and non-native species composition. Make two separate lists for native and non-native plant species.

(3) You can calculate percent change in plant composition over time, for each species, for all species, or for a select group of interest. See page 17 for instructions on calculating percent change over time.

Social and economic indicator data

The statistics you perform on your social and economic indicator data will depend on the type of data you gathered.

For several of the social and economic indicators, you will probably only have one value. Examples of indicators for which you are likely to only have one value include:

- Total number of workers employed by the project
- Amount paid to project workers each month
- Number of community members or households regularly gathering medicinal, food, heating, or building materials from the forest
- Number of stakeholder groups involved in project activities
- Number of acres protected from fire through creation of defensible space, fuelbreaks, or other fuel reduction projects

If you only have one number from each time you gathered indicator data, the only statistical test you can perform is percent change over time (see page 17). In some cases you will have several numbers for the same indicator. This is particularly common if you used surveys, interviews, or focus groups to collect your data. For example, you may have asked several different people the same question and gotten a range of different responses. Indicators for which you are likely to have multiple values include:

- Level of commitment to group communication and learning
- Quality and timeliness of communication among all project partners
- Extent of community participation in project activities
- Extent of agency participation in project activities

If you have more than one value for an indicator, you will want to calculate the average of all the values. You can also calculate standard deviation and standard error to determine the accuracy of the average value. To analyze these indicators, do the following:

1 – Calculate the mean value of each indicator (page 14)
2 – Calculate the standard deviation from the average (page 17)
3 – Calculate percent change in the average value over time (page 17).

Analyzing qualitative data

Qualitative data analysis includes four steps:

Step 1 – Organize your data
Step 2 – Analyze your data
Step 3 – Interpret your analysis
Step 4 – Verify your data and your interpretation

Step 1– Organize your data

Materials needed for data analysis:

Necessary
Pencil
Paper

Desirable
Colored pencils or markers
Computer to organize data

If you have collected qualitative data, you probably have gathered extensive written notes, transcripts, or other non-numeric data from interviews, observations, maps, diagrams, or photos. When gathering qualitative data it is common to end up with interesting information that may not be relevant to your project. The first thing you will want to do is reduce the quantity of data you have collected to focus just on the information that is useful to your monitoring project. We recommend reconvening your entire multiparty monitoring team to review the data collected and to identify which data to analyze and which to leave out of the analysis.

Begin this process by revisiting your original monitoring indicators. What were the questions about your project that your multiparty monitoring team wanted to answer? You will want to work with data that address these questions. You may also want to analyze data that seem important to your project, even if you did not identify them at the outset. Remember, you want to keep all data that you collected in a safe place, even if you do not wish to analyze it for this project.

If you have gathered qualitative data from interviews, focus groups, surveys, and perhaps document review, you will want to further organize your data by coding it. If you are using data gathered from photo points or community mapping, you do not need to code your data and can proceed to *Step 2 – Analyze your data*.

Coding qualitative data

Organizing qualitative data involves categorizing them according to different topics and subtopics. This is also known as coding the data. The simplest way to code your data is to identify and group information related to each indicator that you are measuring.

There are three steps to coding data:

1 – Read through all responses to a question
2 – Identify themes and common responses
3 – Code all responses according to themes

Items 1 and 2 are ideally completed with a multiparty team. Each person in the team reads responses to a question, and identifies themes or common responses to these answers. As a group, the multiparty team discusses these themes or responses and agrees on which will be used to code the data.

For example, imagine your monitoring team used interviews to identify issues related to small diameter wood utilization. Let's say your team interviewed a total of 30 people, including wood products business managers, mill operators, cutting crews, end-product purchasers, and Forest Service staff. Once the interviews were completed, each person from the monitoring team reads all of the interview responses and identifies themes that came up in the interviews. Themes can be analyzed either qualitatively or quantitatively according to the project's needs and desires.

You can use numbers to code, or count, the number of times a particular issue was discussed, as shown in Table 7. This is just one way of organizing your data. Another way is to count the number of times these issues were brought up by people in their different roles as mill operators, cutting crews, Forest Service staff, etc.

When you use numbers to code qualitative data (as we do in Table 8, next page, by counting each time a theme was present in people's interview responses) you are *transforming* qualitative

Table 8 — Coding to transform qualitative data into quantitative data

Key issues in small-diameter utilization	Number of times each issue was discussed
Timing of supply	24
Size of wood	14
Quality of end-products	8
Milling issues	12

responses into numbers. Using the coded responses in Table 8, your monitoring team could say that "timing of supply" was the most important issue, followed by "size of wood" and "milling issues."

You could also follow the steps described on pages 15–16 to calculate the percentage of respondents who selected a particular answer and other values, or use simple statistics, as described earlier. Just beware of drawing any conclusions about the larger population from a sample which is skewed, not random, or has a sample size smaller than 30 (see Handbooks 4 and 5 for information on random sampling).

It is also possible to use coding as an initial way to organize data into tables, maps, or diagrams for qualitative analysis, as described below. For example, rather than counting the number of times each issue was raised, your team might instead pull out examples of what different project partners said for each issue. This kind of table allows you to see what different people said about the same issue and will help you interpret qualitative data. Using tables to organize descriptive data is described in the next section and shown in Table 10.

When analyzing qualitative data from focus groups, interviews, surveys, or other methods, it can be difficult to determine what information is relevant and how to group your data into meaningful categories. For instance, you may have dozens of pages of transcripts from interviews you performed to determine issues related to the extent of community participation in project activities. Look for patterns in the data, such as frequency, size, structures, causes, or consequences, as shown here:

Frequency (how often something occurs)

- How often are community partners invited to take part in project activities?
- How often do they participate?

Size (how large or how many)

- How many people attend project activities?
- How long do the meetings last?

Structures (different forms of the indicator)

- What are the different types of project activities (e.g., field tours, meetings, ...)?

Causes (factors leading to a result)

- Why do people choose to get involved?
- Why do some stakeholders choose not to participate?

Consequences (results of causes)

- How does participation in project activities affect people's attitudes toward the project?

Organizing your data in this way will help you to pull out key issues as well as important differences among your responses.

Step 2 – Analyze your data

Coding (described above) is really the first step in analyzing qualitative data from interviews, surveys, focus groups, and some document analysis. The next step, described below, is to visually display your coded data in a way that makes it easier to interpret. You can do this by displaying the coded data in tables, maps, or diagrams.

Tables

Assume your multiparty monitoring group was particularly interested in improving collaboration among groups and individuals with different perspectives on forest restoration. The group chose to hold a series of focus groups to gather data on the following indicators:

- Number of different perspectives represented in project activities
- Extent of the environmental community's participation in project activities

Table 9 — Example of a table for organizing qualitative data

	Group 1 – Project employees	Group 2 – Traditional forest users	Group 3 – Community leaders	Group 4 – Environmental organizations	Group 5 – General community members
Different perspectives represented in project activities (name them)					
Ways that agency representatives have participated					
Ways that traditional forest users have participated					
Ways that environmental group representatives have participated					

- Extent that traditional forest users participate in project activities
- Extent of Forest Service or other agency participation in project activities

You might then choose to organize your data as in Table 9 (previous page). You would then fill in the table with the data you had already gathered from interviews or surveys. You could combine numeric and descriptive data in this table or use a separate table to organize the quantitative and qualitative data related to participation. An example of how it might look is shown in Table 10. Here we see the way people from three of the five groups responded to "ways that traditional forest users have participated, and how often." In practice, you would fill in all the responses for each of the categories in the table above using quotes from your interviewees. You should choose quotes, or interview excerpts, that capture the issues of importance to your interviewees or to your group.

Table 10 — Example of organizing focus group responses

	Group 1 – Project employees	Group 2 – Traditional forest users	Group 5 – Forest Service
Ways that traditional forest users have participated	"We haven't had as much luck getting traditional forest users to participate as we would have liked. Some came to our initial meetings, and we have had some of the fuelwood gatherers coming into the project site to get wood. We'd like to see more involvement and get more input into how they see the project over the long term."	"At first I didn't know the project was happening, but since I found out I've been up there a few times to get wood. Last season wasn't very good for pinyon nuts, so we hope this year will be better. My grandmother said she was interested in checking the area for medicinals she collects every fall. I need to get up there with her sometime soon."	"In the past, people used this area for pinyon and juniper. I've also seen some women out there collecting herbs. I think participation of these groups has been pretty good for fuel wood collection, but I'm not sure how much other users have been involved. We issued a total of 42 permits for fuelwood, but we only gave out one permit for collecting herbs."

Maps

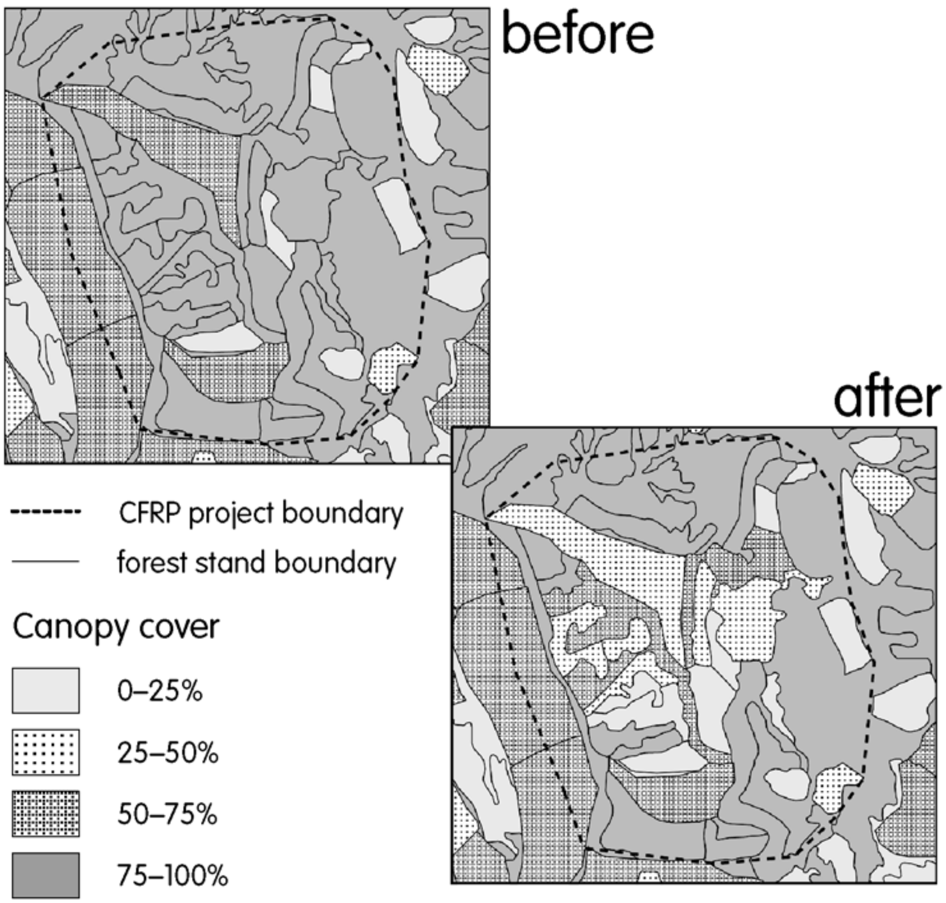
Maps can be used to display interconnections between different kinds of data. For example, if your group is monitoring how your project is affecting wildfire risk in your community, you may have gathered data on areas that New Mexico State Forestry, USDA Forest Service, or perhaps local fire departments consider to be at highest risk from wildfire. You could show these areas on a map, along with any areas where your project has conducted fuels reduction work. You could also include other important information such as the location of structures, important wildlife habitat, and other areas or points of interest.

Similarly, if your community chose to gather data on different opportunities for local families to recreate in the forest, you might have interviewed households to find out where, when, and how different families recreate in the forest. You could then create a series of maps showing recreational uses at different times of the year on the same area of land. By overlaying these maps, you can begin to see overlapping patterns of use that may help you interpret the data.

You could also use these methods to organize ecological data. For example, if you collected data for tree size and density, understory cover, and canopy cover, you could code data collected along transects in plots so that different colors represent different categories. You might choose to color 0–25% canopy cover light green; 26–50% canopy cover medium green; 51–75% canopy cover darker green; and 76% canopy cover very dark green. You could also use different patterns as in the map on the next page.

You could make similar maps for tree size and density, and for understory cover. By overlaying these maps, you begin to get a simple, visual representation of the effects your project may have had or may be having.

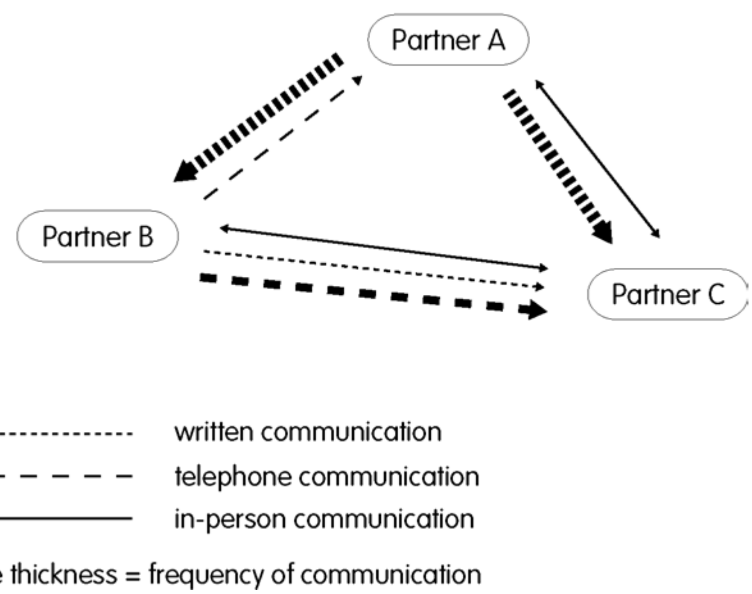
Figure 1 — Data mapping



Diagrams

Diagrams can be useful ways to depict networks of communication, flow of materials, or relationships. For example, assume your multiparty monitoring team chose to measure the quality and timeliness of communication among all project partners, and surveyed project partners to determine how often they communicate in person, over the phone, or in writing with other project partners. You could then create a diagram with different-looking lines (e.g., different colors or dashed and solid lines) linking project partners who communicate. You could use arrows on the ends of the lines to indicate the direction of communication (i.e., one way or two-way communication) and show frequency of communication by the thickness of the line (or by drawing multiple lines), as shown in Figure 2:

Figure 2 — Sample communication network



Step 3 – Interpret your data

Once you have organized and described the data, you can move to the next step and begin to ask what is causing the observed change. When interpreting monitoring data, you will always be comparing one set of data with a previously gathered set of data and asking why there was or was not a change.

This is the stage where your multiparty monitoring team will make comparisons and consider causes and relationships that could be responsible for the observed data changes. Which things appear to lead to other things? Which activities appear to lead to observed outcomes? Having a team, rather than a single person, answer these questions will help to remove personal bias from the interpretation of qualitative data.

With some types of qualitative data, the focus is on describing people's perceptions, rather than determining whether something is objectively "true" as in statistics. You may be measuring, for instance, things like the quality and timeliness of communication or extent of participation in project activities. In these cases, you may want to report on the variety of opinions expressed rather than trying to lump them into a single "result."

When interpreting your data, be careful not to jump to conclusions about cause and effect. An observed relationship between an action and an outcome does not necessarily mean that the action caused the observed outcome. Examine your data carefully, looking for different explanations for apparent patterns. Discuss your interpretations among the entire multiparty monitoring team. Ask yourselves whether your interpretation is consistent with known research and past experience. Interpretations of qualitative data should always take into account the possibility that there could be other explanations for the observed results.

The multiparty approach is critical to the interpretation of qualitative data. It provides one of the best ways to improve the reliability of your results by having several different people independently interpret the data and then compare their conclusions.

When interpreting data, your group should be asking yourselves the following questions:

- How do I know what I know? How might my perceptions, experiences, and background be affecting my analysis of the data?
- How do my sources (the people I interviewed, surveyed, or observed) know what they know? What shapes their views of this project? How do they perceive this project and this monitoring effort? How might these perspectives have affected their responses?
- What outside factors might be affecting the outcomes?

Step 4 – Verify your data and your interpretation

Verification of qualitative data means checking your data and your interpretation for biases. Verifying your data will help you know whether your data are accurate enough to base management decisions on them. Data verification is also important when there are different points of view about the best way to reach your goals or about the potential outcomes of your project.

With qualitative data, verification can be particularly important when you are measuring issues that may be somewhat controversial, such as different groups' perspectives on collaboration. The most common means of verifying qualitative data and analysis is triangulation. Other methods of verifying qualitative data and interpretation include examining alternatives, looking for negative cases, and checking results with respondents.

Triangulation

Collecting data from multiple sources or with multiple methods is called *triangulation* and it helps to verify qualitative data. If data from several different methods point to a common result, then that result has more validity than a result arising from only a single method. For example, you can triangulate your data by having different focus groups answer the same questions, as was shown in Tables 8 and 9. The sample questionnaire on the next page (Figure 3) shows how you can triangulate your data by asking the same question in different ways. Section A of this sample questionnaire asks for a count—quantitative data, while Section B asks for narrative, qualitative data. You can also triangulate your data by using two different data collection methods, for example, both surveys and interviews, to gain answers to the same questions. An example of using qualitative and quantitative data together is given in Appendix 6.

You can also triangulate your data analysis by having multiple analysts interpret the data and then comparing their results. In a multiparty group it can be particularly useful to have people from different stakeholder groups involved in interpreting data from their perspectives.

Figure 3 — Sample questionnaire

Section A					
	Once per week or more often	Less than once per week but more than once per month	Less than once per month but more than once per year	Less than once per year	Never
1. How often does someone in your household use the forest for recreation?					
2. How often does someone in your household use the forest to hunt or gather plants for food or medicine?					
3. How often does someone in your household use the forest to gather wood for heating, cooking, or building materials?					

Section B
4. Please describe all the different ways that people in your household use the forest:
5. What kinds of forest materials (plants, animals, wood, etc.) does your household collect from the forest?

When triangulating data, methods, or analysis, the goal is not to look for *exactly the same data or results*, but to look for *consistency* in the data or results. If there is consistency in results from different methods, then you can be fairly confident your results are accurate.

If, on the other hand, you find inconsistencies between the results of different methods, or between data from a single method, you may need to re-examine the issue. For example, say you used open-ended survey questions to assess the extent to which stakeholders previously in conflict are now working together, and you find that most respondents say there is no conflict on the project while a few respondents report very high levels of conflict. You should not disregard the few who report experiencing conflict. Instead, take a closer look at the differences between those who report no conflict and those who report high levels of conflict: is it a certain kind of stakeholder that reports conflict? Why is the conflict not reported by others? If your monitoring gives you contrasting results, report on the inconsistencies. They could be evidence of important dynamics that will need further attention, or may indicate a problem in the data collection or analysis process.

Examining alternatives

Examining alternatives means looking for alternative patterns, themes, and explanations for your data. To do this, try organizing your data differently and see if it suggests different findings. Think of other logical explanations for the results you observed and see if they can be supported by the data. Examining alternate explanations improves the credibility of your data analysis by strengthening confidence in your data interpretation.

Looking for negative cases

Looking for negative cases means examining your data *outliers*—data points that do not fit the pattern, theme, or explanation that you are using to explain the data set as a whole. For example, look at interview or survey responses that contradict the majority of the responses, and discuss whether there may be other explanations that may support these responses. There may be logical reasons why these data points do not fit the overall pattern (e.g., the respondents may be from a particular sub-set of the population that has had different experiences than the majority). If you cannot come up with any reasons for the data outliers, ask yourselves whether these outliers might make your conclusions less valid. On the other hand, outliers may suggest changes in your methods.

Asking your data sources

Asking your data sources to review and confirm your interpretations is also an appropriate method of verifying qualitative data analysis. You may want to re-convene a focus group or mail a summary of your data analysis to people you interviewed or surveyed and ask them for feedback on your group's interpretation of the data. This is an important step that will add credibility to your findings.

Displaying and communicating results

Effectively communicating results to the monitoring group or members of the community is an important part of your monitoring project. Depending on who you want to reach, you may want to share your monitoring results through oral presentations; posters or short leaflets that give a colorful, easy-to-understand overview of what you learned; or more detailed technical reports. Table 11 (next page) lists different ways to communicate results and appropriate audiences for each method. When displaying and communicating your results, it is important to remember your obligation to protect the confidentiality of any interviewees or focus group participants. You can give people pseudonyms (false names) if you want to refer to them in a report or presentation, or simply refer to them as "interviewee #1," "focus group participant #2," etc.

Table 11 — Ways to communicate results for different audiences

Ways to communicate results	Appropriate audiences for each method
Oral presentations in a meeting setting or on radio or television	General public
Posters visualizing central findings with pictures, charts, and few words	General public, youth, less knowledgeable audiences
Short leaflets or brochures that highlight general findings and recommended actions	General public and community groups
Written descriptions of what was monitored, why, how, and what was learned (e.g., newspaper article, or executive summary)	Informed publics, decisionmakers, advisory committees, etc.
Technical report (including description of methods and figures or charts showing findings)	Researchers, outside groups, and others interested in details

Visual tools like graphs, maps, diagrams, and figures are very effective ways to communicate monitoring results. For example, the familiar "thermometer" used by the United Way is a simple but effective way of using a bar chart to display the level of funding it has received to date. Instructions for making bar charts and pie charts are provided in Appendix 4. Storytelling, or verbal description, is a valid way to report some data, and can be the most effective way of explaining results.

Use visual tools effectively and efficiently

Two common mistakes in using tables, figures, and other visual tools are making *too many* and making them *too confusing*. Each visual tool should be easy for your audience to understand, and presented in a way that makes sense to them. Instead of presenting a large number of visual tools, choose a few that will have the most meaning to your audience. If possible, consolidate information from several tables or figures into a single visual tool. Just be careful not to include so much information that it becomes cluttered and hard to follow. For examples of using visual tools, see Appendix 4 and Appendix 6. Also see the discussion of tables, maps, and diagrams on pages 35–39.

Know your audience and know what you want to say to them

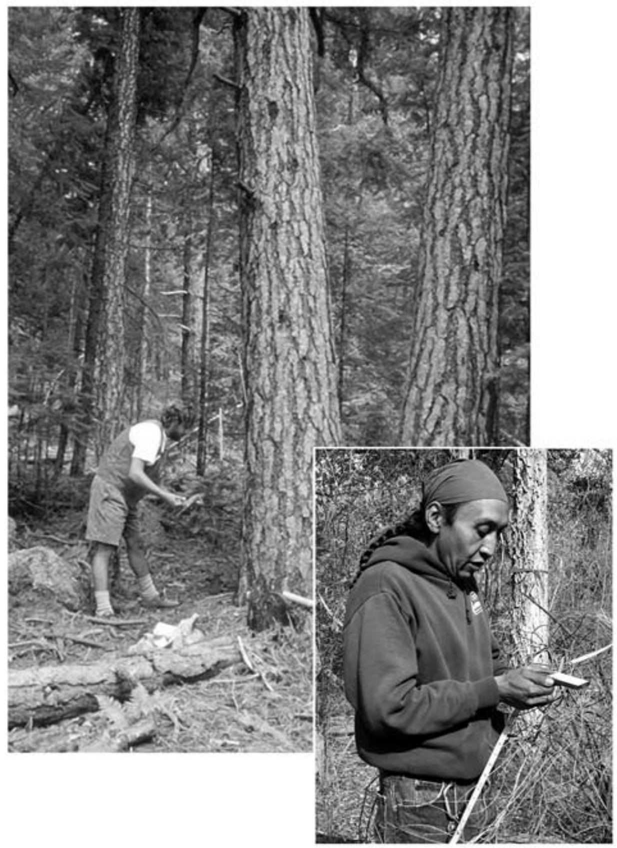
Whatever methods or tools you choose, focus your presentation on information that is new, relevant to the project and to the audience, and useful. Do not waste time or space on things people already know or findings that will not inform the project or the community. Make sure the person developing these materials *fully* understands the data and how they were analyzed, including reasons why specific statistical tests were performed and the specific methodologies for each analysis. This is important to ensure the presenter does not misrepresent information and can answer questions related to the data analysis.

Provide a balance of data and interpretation

Conclusions and abstract concepts are more powerful when they are combined with data summaries, whether they are numerical or qualitative. Do not overwhelm your audience with data, rather provide enough data that they can understand how you arrived at your conclusions. Any recommendations you make need to be substantiated by your findings. For an example of using both qualitative and quantitative data to arrive at a finding, see Appendix 6.

Prepare for the critics

In addition to your findings, be able to succinctly describe how data was gathered, the data itself, and how the data was analyzed. Be able to answer the following questions about your monitoring results: How solid is the evidence? How will this information be heard and understood? How might this information be used, or misused? To avoid misinterpretation and misuse of your monitoring results, report your data gathering and analysis methods as completely and truthfully as possible. If you need to include large amounts of raw data in your report, consider putting these in an appendix. Transparency is the key to credibility. Also, larger sample sizes will always make the results of your statistical analysis better.



Appendix 1 – Conversions

Metric to English conversions:

1 centimeter	= 0.393 inches
1 meter	= 3.28 feet
1 square meter	= 10.76 square feet
1 hectare	= 2.47 acres

Area Conversion:
To convert a value per sampled area to a value per acre:

(1) Calculate the area sampled:

a. For the area of a plot, multiply the length times the width of each plot.

Example: 30 x 30 = 900 square feet.

b. Add together all plot areas (for total area sampled).

Example: 30 plots x 900 square feet = 27,000 square feet.

(2) Divide the total area (in square feet) by the number of square feet per acre. (The area of an acre is 43,560 square feet.) This will give you the number of acres sampled.

Example: 27,000 square feet/43,560 square feet per acre is equal to 0.62 acres sampled.

(3) Divide the number of trees sampled by the number of acres sampled. This will give you the number of trees per acre:

There were 50 trees in the 30 plots, 0.62 acres were sampled, therefore 50 trees/0.62 acres = 80.6 trees per acre.

Appendix 2 – Calculating amount of downed woody fuels

Downed woody material computation summary (from Brown 1974)____

Forest: __________ Compartment: __________ Stand: __________

Size class (inches)	(1) Constant	(2) n	(3) Diameter²	(4) Specific gravity	(5) Secant	(6) Slope factor	(7) Total transect length	(8) Total tons per acre
0–0.25	11.64							I:
0.25–1	11.64							II:
1–3	11.64							III:
3+	11.64	1						IV:
							Total: I + II + III + IV =	V:

To find tons per acre for each size class, multiply (1) x (2) x (3) x (4) x (5) x (6) and divide the resulting number by (7). See text to find values for these columns.

To find total tons per acre, add I + II + III + IV to get total V.

Fill in columns (1)–(8) using the following methods:

(1) Constant: This is 11.64 for all size classes

(2) n: For sizes less than 3″, this is the number of pieces encountered in each size class. For sizes over 3″, this value is 1.

(3) Diameter^2: For sizes less than 3”, use the table below to determine diameter^2 based on the dominant forest type. For example, in the case of ponderosa pine nonslash in the 0.25–1″size class, simply enter 0.2366.

Size class (inches)	Forest type	(3) Diameter²		(4) Specific gravity	
		nonslash	slash	nonslash	slash
0–0.25	ponderosa pine	0.0378	0.0471	0.463	0.540
	others	0.0151	0.0151	0.48	0.48
0.25–1	ponderosa pine	0.2366	0.2467	0.464	0.543
	others	0.289	0.289	0.48	0.48
1–3	ponderosa pine	2.967	3.574	0.392	0.502
	others	2.76	2.76	0.40	0.40
3+	ponderosa pine	Square the diameter of each piece; add these together to get the value for column 3		0.40	0.40
	others			0.40	0.40

For sizes over 3″, you do not need to use the table to determine diameter^2. Instead, square the diameter for each piece encountered and add these squares together. Enter this "sum of squares" into the space in column (3).

For example, if you encountered three pieces of debris, measuring 3.8″, 6.1″, and 12.0″, you would compute the following:

(3.8 x 3.8) + (6.1 x 6.1) + (12.0 x 12.0) =

14.44 + 37.21 + 144.0 = 195.65

Enter 195.65 into the table.

(4) Specific gravity: Use the table above to determine specific gravity based on the dominant forest type.

(5) Secant: This is a number which corrects for the fact that pieces of debris do not lie at perfectly perpendicular angles to the transect line. The number you use will depend on the size and whether the pieces are slash or nonslash fuels.

Nonslash: for pieces less than 3″ the secant is 1.13; for pieces 3″ and larger the secant is 1.00.

Slash:

Size class (inches)	Species	Secant	
		Fresh slash	Slash 1 year and older
0–0.25	ponderosa pine	1.25	1.25
	others	1.40	1.15
0.25–1	ponderosa pine	1.25	1.25
	others	1.13	1.13
1–3	ponderosa pine	1.22	1.22
	others	1.10	1.10
3+	all species	1.00	1.00

(6) Slope factor: This is a number which corrects for the average slope of your transects. Use the following table to determine what value to enter into column (6):

Slope (%)	Correction factor
0	1.00
10	1.00
20	1.02
30	1.04
40	1.08
50	1.12
60	1.17
70	1.22
80	1.28
90	1.35
100	1.41
110	1.49

(7) Total transect length: This is the total length of transect line, calculated for each size class. To get this value, multiply the number of transects by the length of each transect. This value will be different for different size classes, since the length of transect varies for each size class. For example, if you had six transects of six feet long for measuring the 0–0.25″ size class, your total transect length would be 6 x 6 = 36. In these same six transects, if each transect was 35 feet long for the 3″ + size class, your total transect length for that row would be 6 x 35 = 210.

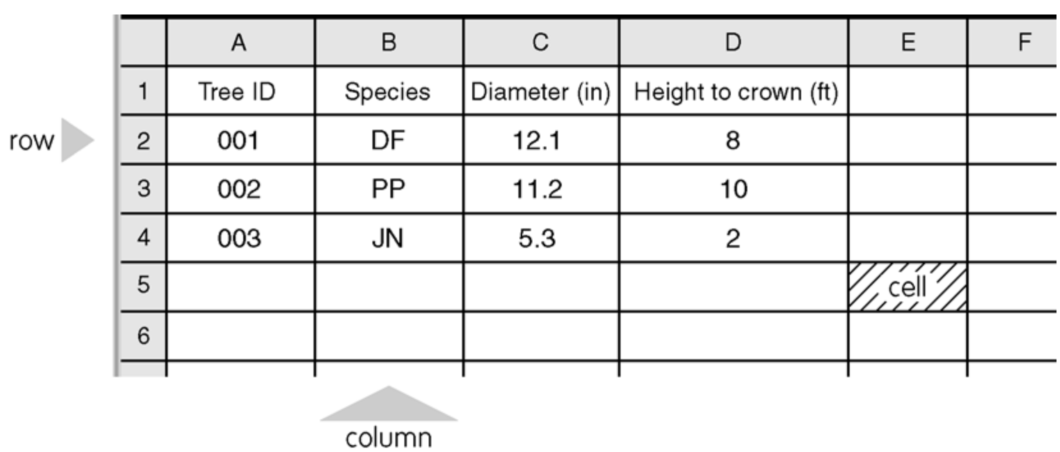
(8) Calculating fuel load: For each size class, multiply the values in columns: (1) x (2) x (3) x (4) x (5) x (6). Divide the product of these six columns by the value in column (7). This will give you tons per acre for each size class. To calculate total fuel load, simply add the values in column (8) for each size class: I + II + III + IV. The sum of these values is the total fuel load of your sampled area, measured in tons per acre.

Appendix 3 – Using computer spreadsheets

Instructions for setting up a spreadsheet and calculating descriptive statistics using a computer.

Setting up a spreadsheet

When you open a computerized spreadsheet, such as *Microsoft Excel*, you will see a screen of blank data cells. There will be capital letters running left to right along the top. Each one of these letters represents a *column*. There will also be numbers running from top to bottom along the left side of the screen. Each one of these letters represents a *row*. Each box is called a *cell*.



(1) <u>Define data worksheets</u>

The data collected during a monitoring project will likely represent many different indicators. If you look towards the bottom of the screen, you can see tabs that say "Sheet 1…Sheet 2…Sheet 3." If you choose, each indicator can be analyzed and stored on its own worksheet, indicated by these tabs. This is an easy way to organize the data. To do this, simply place the cursor on the "Sheet 1" tab and then click twice. When the words "Sheet 1" are highlighted, you can rename the tab by typing in a label for a particular indicator. For example, you may want to have one sheet for adult living trees, one for dead standing trees, and one for saplings.

- Inserting more worksheets: You can insert more worksheets by clicking on the Insert button in the toolbar and then clicking on Worksheet.
- Reordering the worksheets: You can change what order the worksheets are in by clicking on a specific worksheet tab and dragging it left or right, depending on which way you would like it reordered.

(2) Label the columns

Once you have set up the worksheets in an organized and easy-to-understand way, you can now begin entering the data. It is a good idea to put a title at the top of every column to label what the data in that column represents. As seen above, we have used *Data Sheet A: Tree stem density and size*, as an example.

- Adjusting the width of the column: If the wording in your label is wider than the column, you can adjust the width of the column by putting the cursor near the capital letter of the cell you would like to enlarge. Once the cursor has changed shape into a cross, place it on the line at the side of the box, click, and then drag the box to the desired width.

(3) Enter the data and save your work

It is important to make sure all of the data from the data sheets are accurately entered into the worksheets.

- To change the format of the cell: If there are certain data you would like to draw attention to, simply click on the top cell of that set of data, click and drag your cursor to the bottom cell of the set of data. You will notice now your data is surrounded with a thick black line. Now click on your right mouse button to see many options for changing the cells. If you click on Format Cells, you will see options for changing the color of these cells, how the number is represented, etc.
- It is a good idea to save your work often when entering data. Do this by clicking on the small disk symbol in the toolbar or by clicking on the word File in the toolbar and

then Save As. The first time you save, you will have to enter a file name. Once you have set the file name, the program will automatically update the file every time you save.

Using spreadsheets to calculate descriptive statistics

The mean, median, mode, standard deviation, standard error, range, and variance can all be easily calculated in a computer spreadsheet. The following instructions are specific to *Microsoft Excel*; however most spreadsheet programs work basically the same way:

(1) Make sure all data have been entered from each data sheet and checked for errors.

(2) Click on Tools in the toolbar.

(3) Click on Data Analysis in the drop-down menu.

This function is one that has to be installed. If you do not see it in the drop down menu, you will have to install this function.

(4) Find Descriptive Statistics in the menu of choices and click on it.

(5) Input the data you would like to analyze.

Only put in the data you would like to analyze for one indicator at a time. Generally that data will only be in one column. Do not try to calculate the descriptive statistics for more than one variable at a time.

- Input: Put your cursor in the *input range* box and click. Now move your cursor over and highlight the cells containing data you would like to analyze. You will notice the data you have selected now has a flashing line around it and the input box now contains the cell numbers it will do the analysis on. If you have a label in the first row, check that cell.

(6) Output options: If you would like to have the results of the statistical analysis reported in the same worksheet as your data, then click on the *output range* box and then click on the cell where you would like to have the results

reported (e.g., the next empty column). If you would like to have the results reported on a different worksheet, choose that option. Make sure you select the box for Summary Statistics. The results of your analysis will look similar to the table below. Ignore the statistics that are not relevant to your analysis (such as kurtosis).

	A	B	C
1	*Column1*	*Column2*	
2			
3	Mean	5.33333333	
4	Standard Error	0.464621368	
5	Median	5	
6	Mode	5	
7	Standard Deviation	2.12916259	
8	Sample Variance	4.533333333	
9	Kurtosis	-0.968573575	
10	Skewness	-0.03689731	
11	Range	7	
12	Minimum	2	
13	Maximum	9	
14	Sum	112	
15	Count	21	
16			

Appendix 4 – Bar charts and pie charts

A simple but effective way to display your monitoring results is to make bar charts and pie charts. This appendix describes how to do that by hand and using *Microsoft Excel*.

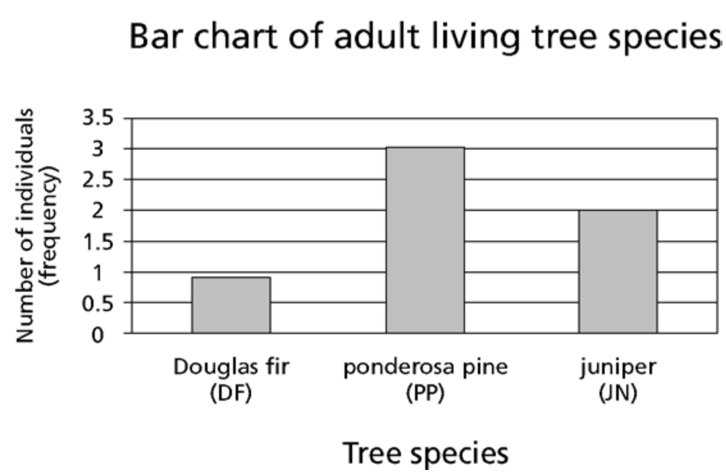
Bar charts

Bar charts illustrate the frequency of different values or responses with bars or lines. They are an excellent way to visually display data your group has collected. You can easily make a bar chart to display relative frequency of your data points using the following steps:

(1) Put the number of individuals that have been counted (frequency) on the y-axis (up and down axis).

(2) Put the name of the item being measured on the x-axis (left to right axis).

Example: We can display tree species data from Table 4 on page 13 to make a bar chart:



Pie charts

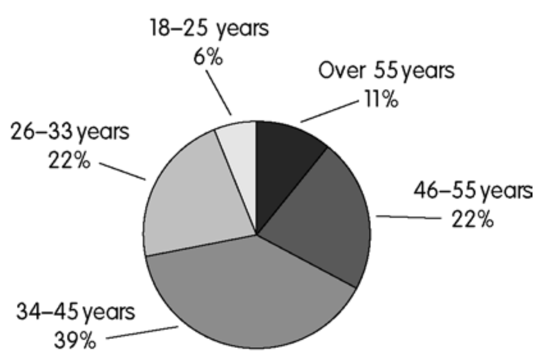
Pie charts are a good way to display percentages in visual form if you are going to be presenting your data in front of a group, in a report, or even if your group would like to "see" an overall representation of the data you have collected. Pie charts illustrate the percentages of the data visually. The whole pie represents 100 percent of the data.

Example: If you wanted to collect age data on a questionnaire, it would be easy to display your results in a pie chart.

Age:

18–25 years	2
26–35 years	8
35–45 years	14
46–45 years	8
Over 55 years	4

Pie chart representing age of questionnaire respondents



Using computer spreadsheets to make bar charts and pie charts

Before your group is ready to make a bar chart or pie chart in a computerized spreadsheet, it must first be comfortable with the terminology and methods explained in Appendix 2. The following instructions are specific to *Microsoft Excel*; however, they will work for most spreadsheet programs:

(1) List the *independent variables* in column A (remember the independent variables are those that are being measured, such as tree species or the answer choices for the close-ended survey questions). These will be on the x-axis of your bar chart (the left to right axis) or they will be used to label the “slices” of your pie in a pie chart.

(2) List the *dependent variables* in column B (remember the dependent variables are those that have been counted, either as a frequency or as a percentage). These will be on the y-axis of your bar chart (the up and down axis) or they will be used to describe the size of the “slices” of your pie in the pie chart.

Your worksheet should look like this:

	A	B	C	D
1	Douglas fir	1		
2	ponderosa pine	3		
3	juniper	2		
4				
5				

independent variable (column A) — dependent variable (column B)

(3) Click, drag, and highlight all the cells that contain your information. The cells that will be included in the graph should now be changed to a different color.

(4) Go to the toolbar, look for the graph symbol, and click it.

(5) Choose which type of graph you would like to make (in this function, bar charts may be called "bar graphs" or "column charts"). Click Next.

(6) Work your way through the steps that will allow you to change many aspects of your graph. This may include changing the color of the bars or "slices" of pie, having the labels displayed, adding a title, etc.

(7) When you are finished making changes, click Finish and save your work.

Appendix 5 – Data verification

There are many types of statistical tests that can be used to verify your monitoring data. Providing instructions for all of the possible tests is beyond the scope of this handbook. The purpose of this appendix is to present some of the possibilities for verifying your monitoring data should you want to use more rigorous statistical tests.

Non-representative data

Before you choose a verification (or inferential statistics) test, it is important to know if your data are likely to be representative of the whole population. *Non-representative data* result from sampling methods that are not random or from a very small sample size. For example, let's say you wanted to monitor the extent to which project workers receive health benefits. If your sample only included workers employed by established businesses and didn't include self-employed contractors, your data would be *non-representative* of the population of workers. Your results might show that 90 percent of project workers have health benefits when in fact, when the self-employed are included, the real number is more like 50 percent. Similarly, if you hand-picked which plots to measure for ecological data your data may be biased

If your data are non-representative, you can still calculate several statistics that will be useful for comparing change over time on your sample plots. However, you cannot use the sign test or confidence intervals to verify your data and you should be cautious about drawing conclusions about the larger population or area based on non-representative data.

Did my treatment make a difference?

If you have small sample sizes of skewed continuous data you can use the *sign test* (page 64) to measure if the changes between the pre-treatment data and the post-treatment data were significant (not due to chance).

If you have nominal data and your data are not skewed, you can answer this question by calculating a *confidence interval* (page 68). (Remember, skewed data means a lot of your numbers are big or a lot of your numbers are small.)

Sign test: This is a good way to measure whether or not the changes between the pre-treatment data and the post-treatment data are due only to chance or if the treatment really did have an effect. The sign test is only used for continuous data and can be used for skewed data.

Confidence interval: Information from the confidence interval will tell you what kind of results you would expect to see in the population at large if you had done the treatment on the entire population, instead of just the sample. The confidence interval can be used for nominal data, but is not valid for skewed data.

Other common tests used to verify the type of data you have collected during your monitoring project include: t-tests, chi-square, analysis of variance, and linear regression. Providing instructions for all of the possible tests is beyond the scope of this handbook. See the references listed in Appendix 8 for additional resources.

Sign test

If you have a sample that has less than 30 data points or is skewed, the sign test is an easy and appropriate way to assess the results of your project. The sign test is only used for *paired data*—that is, data collected before a treatment and data collected after a treatment from the same places. The sign test is used for continuous data.

Calculating the sign test

(1) Make a four-column table like the one on page 66. List all of your data points from the pre-treatment measurements in column A.

(2) List all of the data points from the post-treatment measurements in column B.

(3) Subtract the post-treatment measurements from the pre-treatment measurements and put that value in column C.

(4) If the difference between the pre-treatment data and the post-treatment data is positive put a plus sign in that row in column D. If the difference in negative, put a negative sign in that row in column D.

(5) Count the number of positive signs. In the equation we will use, the number of positive signs will be labeled D.

(6) Count the number of data points in column A (this will be labeled as n) and divide that number by two. This will be labeled as n/2 in the equation.

(7) Divide the total number of data points in column A (n) by 4 and then take the square root of that number. This will be labeled as $\sqrt{n/4}$ in the equation.

(8) Plug the values you calculated in steps 4 through 7 in the following equation to get the z+score.

$$z+ = \frac{D - (n/2)}{\sqrt{n/4}}$$

(9) Use the z+ score to look up the associated value in the z-table (page 67).

(10) Once you have found the value associated with the z+ score, multiply that value by 2.

(11) If this number is smaller than 0.05 you can say there are statistically significant differences in the pre-treatment and post-treatment data. This is very favorable.

(12) If this number is larger than 0.05 then you know there were not any statistically significant differences in the pre-treatment and post-treatment data.

column A (data points for pre-treatment measurements	column B (data points for post-treatment measurements	column C ((pre-treatment data) - (post-treatment data))	column D (positive or negative signs)
34.5	37	-2.5	-
23.5	21.2	2.3	+
10.3	8.2	2.1	+
45.2	47.2	-2.0	-
5.2	4.2	1.0	+
6.1	6.0	0.1	+

<u>Example of calculating the sign test</u>:

From the table above, n = 6 and the total number of positive signs (column D) = 4

Number of total data points (n) divided by 2:

$(n/2) = 6/2 = 3$

The square root of the number of data points divided by 4:

$$4\sqrt{n/4} = \sqrt{6/4} = 1.22$$

$$z+ = \frac{4 - 3}{1.22} = 0.82$$

Using the z-table (next page; see Table instructions), we see the associated area for the z+ score of 0.82 is 0.2939. We multiply that value by 2 to get 0.5878. Since 0.5878 is larger than 0.05 we can conclude there were not any statistically significant differences between our pre treatment data and our post-treatment data. This means the treatment did not have a measurable effect on the plot.

The z-table

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Example of how to use this table:

1 Locate the first column of numbers on the left hand side of the table. These numbers are in bold. If you have a z+ score of 0.82, as we did in our example, locate the row that begins with 0.8. Keep your finger on the 0.8.

2 Locate the first row of numbers on the top of the table. These numbers are in bold. If you have a z+ score of 0.82, locate the column that begins with 0.02 (because 0.8 + 0.02 = 0.82). Keep your other finger on the 0.02.

3 To find the value of the area we are looking for, trace your fingers from 0.8 to the right and from 0.02 down until they meet. This value is 0.2939.

4 Multiply this value by 2 to get the true area of z+ score. In this case 2 x 0.2939 is 0.5878.

5 Compare this value to 0.05. If larger than 0.05, then the difference is not statistically significant. If it is smaller than 0.05, then the difference is statistically significant and that means your treatment had a measurable effect on the plot.

Confidence intervals

A *confidence interval* is a range between two numbers that tells you how closely the average value in your sample reflects the average value for the entire population. The confidence interval for a statistic is always calculated for a pre-determined confidence level using the standard error for your sample data. We will always use 95 percent because it is the most commonly used value among natural resource researchers.

Remember, do not use a confidence interval if your data are skewed, were not randomly selected, or if you have a small sample size. If any of those are true about your data, then the results of a confidence interval would not accurately represent the population as a whole. If you have less than 30 samples in your dataset, there are other statistical tests we can help you do to analyze your data other than the confidence interval.

Calculating the confidence interval for continuous data

A confidence interval includes two numbers: one to represent the *upper limits* and one to represent the *lower limits*. To calculate the confidence interval's upper and lower limits, you will need to know:

- the sample mean (average),
- the z-score (a constant, based on the confidence level; for a 95% confidence level, the z-score will always be 1.960),*
- the sample's standard deviation (see page 17 for instructions on determining your sample's standard deviation) and the total sample size (number of data points or n).

---

* For data such as yours, where we do not have information about the actual population, we could use a t-score instead of a z-score. We will be using the z-score because it is easier to calculate and the differences between the t-score and the z-score are insignificant.

Then simply plug those numbers into the following equations:

$$\text{Upper limits} = \text{average} + \left(1.960 \times \frac{\text{standard deviation}}{\sqrt{n}}\right)$$

$$\text{Lower limits} = \text{average} - \left(1.960 \times \frac{\text{standard deviation}}{\sqrt{n}}\right)$$

Example: Suppose your sample of tree diameters had a mean of 10.15 inches, a standard deviation of 3.44 inches, and your sample size is 6.**

$$\text{Upper limits} = 10.15 + (1.96 \times 3.44/\sqrt{6})$$

$$= 10.15 + 2.75 = 12.90 \text{ inches}$$

$$\text{Lower limits} = 10.15 - (1.96 \times 3.44/\sqrt{6})$$

$$= 10.15 - 2.75 = 7.40 \text{ inches}$$

Explanation: In this case, we can say we are 95% confident that the mean diameter of all the trees on the site (our total population) is between 7.40 inches and 12.90 inches.

Confidence interval for nominal data

A confidence interval is a range around your average value tells you where the average value for your population is likely to fall. If the average for your post-treatment data falls outside of the confidence interval of your pre-treatment data, then the change that occurred is said to be statistically significant. For nominal data you can calculatethe confidence interval for a percentage (e.g., species composition at a site) and for percent change between two time intervals, such as pre-treatment and post-treatment.

** In this example we have a sample size of only six. When calculating a confidence interval with your real data, it is important to have a sample size larger than 30.

Calculating confidence interval for a percentage

To calculate the upper and lower limits for the percentage's confidence interval, you simply need to know what your percentage is. For natural resources data we use a 95% confidence level, which has a corresponding *z-score* of 1.96. Plug those values into the following equation:

$$\text{Upper limit} = \text{percentage} + \text{z-score}\sqrt{\frac{(\text{percentage}\ (1\text{-percentage})}{n}}$$

$$\text{Lower limit} = \text{percentage} - \text{z-score}\sqrt{\frac{(\text{percentage}\ (1\text{-percentage})}{n}}$$

<u>Example</u>: Suppose you were measuring the percentage of tree mortality from bark beetles. Your group counted six trees out of 100 that had died, making your percentage 6% (i.e., 6/100). The z-score for a 95% confidence interval is 1.96. The confidence interval for these values is calculated as:

$$\text{Upper limit} = 0.06 + 1.96\ \ 0.06\sqrt{\frac{(1\text{-}0.06)}{100}} = 0.107$$

$$\text{Lower limit} = 0.06 - 1.96\ \ 0.06\sqrt{\frac{(1\text{-}0.06)}{100}} = 0.013$$

<u>Explanation</u>: In this case, we can say we are 95% confident that the proportion of the tree population that has died because of bark beetles is between 0.013 and 0.107 of the whole. So, we can assume with 95% confidence the percentage of bark beetle mortality within the entire population is 1.3% to 10.7%.

Calculating the confidence interval for change over time

If you have found the percentage change of a specific species between two time intervals you can calculate the confidence interval for that value as follows. For this calculation, you will need to know the *percentage* from each time you sampled (e.g., before and after treatment) and the total size of the sample for each sample. Calculate the upper and lower limits of this confidence interval following these steps:

(1) Subtract the percentage from the second time you sampled (P_2) from the first time you sampled (P_1) to find the difference (D) between the two:

$$P_1 - P_2 = D$$

(2) Find the standard error for the difference between the two sample times using the following formula:

$$\text{SE of difference} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

(3) Plug the values you got from steps 1 and 2 into the following formulas to find the upper and lower limits for the confidence interval:

Upper limit: D + (z-score x SE of difference)

Lower limit: D - (z-score x SE of difference)

<u>Example</u>: Suppose your group is curious to see if your restoration efforts attracted more songbirds to the restoration site. The first time you sampled the plots you found 150 birds (this is your n_1) and 25 of them were songbirds. To find the percentage of songbirds in the plots you calculated 25/150, or 0.17 (P_1) = 17%. Your group performed some treatments and went back to take measurements the next year. This time you found 160 birds (this is your n_2) and 52 of them were songbirds. This means your percentage of songbirds for the second year was 52/160, or 0.33 (P_2) = 33%. Your group will use a confidence level of 95% (corresponding z-score is 1.96). Now simply follow steps we discussed above:

(1) Subtract the percentage in second time interval (P_2) from the first time interval (P_1) to find the difference (D) between the two:

$P_1 - P_2 = D$

$0.33 - 0.17 = 0.16$

33% - 17% = 16% change in songbirds from year 1 to year 2

(2) Find the standard error for the difference between the two time intervals using the following formula:

$$\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

In this case:

$$\sqrt{\frac{0.17(1-0.17)}{150} + \frac{0.33(1-0.33)}{160}} =$$

$$\sqrt{0.0009 + 0.0014} = \sqrt{0.0023} = 0.05$$

(3) Plug the values you got from steps 1 and 2 into the following formulas to find the upper and lower limits for the confidence interval:

Upper limit: D + (z-score x SE of difference)

Lower limit: D - (z-score x SE of difference)

$0.16+(1.96 \times 0.05) = 0.26$

$0.16 - (1.96 \times 0.05) = 0.06$

Explanation: With this information, we can assume that if the group performed the same treatments on the entire population (the whole project area) that it had performed on the sample plots, the increase in songbirds would be between 0.06 and 0.26 of all of the birds, or between 6% and 26% of the total bird population.

Appendix 6 – Presenting qualitative and quantitative data together

The following is an example of how to present findings from qualitative data (such as from interview transcripts) and quantitative data (such as numbers of groups participating, from document review) together. Excerpts from the interview transcripts are used to support claims made in the text, and these claims are compared against quantitative data to produce a more complete picture of the patterns of interest.

From interviews of twelve potential or actual project partners representing nine different organizations, two major obstacles to greater community participation emerged. The first was communication. A primary reason many interviewees gave for not participating more was that they felt uninformed as to when and where meetings and projects were taking place. There does not appear to be a standard, consistent place that community members can find out when and where project activities are occurring. For example, one potential partner explained,

> *By the time I hear about a meeting it's usually already happened …[our organization] doesn't have email, so if that's the way they're communicating then we're out of the loop. We have a phone and get U.S. mail….I think some people assume everyone knows when a meeting is going to happen, but I don't.*

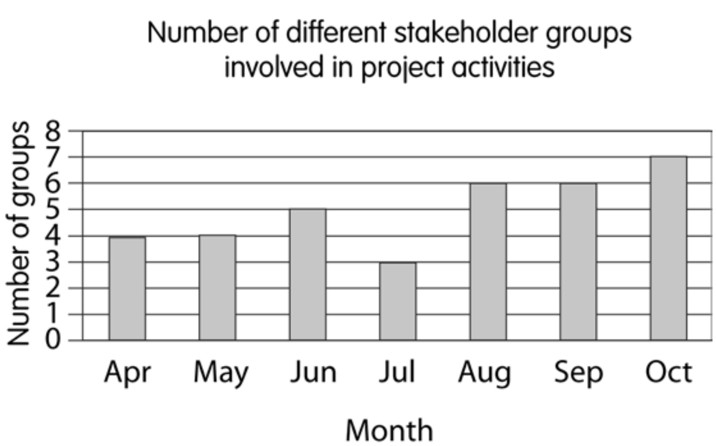
The second major obstacle is the timing and duration of the organizational meetings. Usually held once a month on a Tuesday evening, the meetings have been known to run longer than three hours. Several interviewees expressed hesitation at committing to such a long meeting, as most of the interviewees work at least 40 hours a week and many spend evenings with their children or grandchildren. Even some of the most active project partners identified the timing and length of the organizational meetings as a frustration, and expressed some sympathy with those who did not attend.

According to one project partner:

> *The Tuesday night meetings are probably the most frustrating part. We have way too much to cover in a single night, and most of us are already tired and itching to get home....I know of at least a few individuals who went to one or two of the meetings and then never showed up again. And I understand not wanting to go through that every month, but unfortunately those that don't come on Tuesday nights almost never show up in the field.*

At the same time, eight of the nine organizations (89%) represented in our interviews participated in at least one project activity (organizational meeting, field tour, or restoration activity) and five of the nine organizations (56%) were "regular participants," meaning they took part in two or more organizational meetings, two or more field tours, and two or more restoration projects. This indicates that, despite problems with communication and frustration about long organizational meetings, there was still a high level of interest in, and support for, collaborative restoration activities.

In addition, interest and support for restoration, as measured by the number of different stakeholder groups involved in project activities, appear to be maintaining or even increasing over time, as shown in the bar chart below. Between April and October, stakeholder participation increased from four groups to seven



groups, an increase of 75 percent. Our target value was to have a 50 percent increase in stakeholder group participation, a level that was met and exceeded. In summary, the interest and support for collaborative restoration is high but actual participation by community partners could be improved with better communication and a restructuring of organizational meetings to address the concerns of community partners.

Appendix 7 – Glossary

Average: The central tendency of a data set. See definitions for mean and median. In this handbook, average is synonymous with mean unless othewise specified.

Bar chart: A graphic representations of the frequency of different values using rectangles with heights proportional to the frequencies.

Baseline data: Data collected at the beginning of a project to document the existing situation. These data provide a benchmark that can be used to compare against data collected after the project has been completed.

Bias: An aspect of a measurement or analysis method that tends to cause data or analysis results to be misrepresented. For example, a poorly worded interview or survey question could result in biased results, and personal preferences can bias data interpretation.

Categorical data: Data that has been recorded in groups or categories based on similar characteristics.

Close-ended question: A question with a limited number of pre-determined answers.

Coding: Process of classifying data using terms (or "codes") that represent characteristics you are measuring. Coding can be used to organize qualitative data or to transform qualitative data into quantitative data by counting the number of responses to each code.

Confidence level: A percentage, selected by your multiparty monitoring group, that indicates how confident you are that your sample data is representative of the entire population. Confidence levels are always expressed as a percentage.

Confidence interval: A range between two numbers that tells you how closely the average value in your sample reflects the average value of the entire population. A confidence level is expressed by an upper limit and a lower limit.

Continuous data: Numerical data that can take an infinite number of values.

Data: A set of observations collected through monitoring. Information is derived from data through analysis.

Data analysis: Methods for determining the meaning of data.

Data point: One unit in a data set. For example, a single measurement taken for percent understory cover during pre-treatment collection at transect 2, plot 3.

Data set: A collection of related data points. For example, all the pre-treatment measurements taken for percent understory cover comprise a data set.

Demographics: Data related to the personal aspects of a population, such as age, gender, income level, etc.

Dependent variable: The number of individuals that have been counted. Always represented on the y-axis of a graph.

Descriptive statistics: Statistical tests that are used to describe or summarize the data, regardless if the data represents the population or a sample.

Frequency: The number of times a certain value or response occurs in a set of data points.

Frequency table: A method used to display the number of times each value or response occurs within a data set.

Hypothesis: A formal statement of the proposed outcome between two variables.

Hypothesis testing: Using inferential statistical tests to determine whether the null hypothesis is correct or incorrect.

Implement: To put a plan or agreement into action.

Independent variable: Variables that are being measured such as a tree species. Always represented on the x-axis.

Indicators: A unit of information measured over time that documents changes in a specific condition.

Inferential statistics: Statistical tests used to make inferences or draw conclusions about the entire population when only a sample of the population has been collected. Used to indicate how reliably your sample data reflects the real state of the larger population.

Likert-type: Set of survey or questionnaire response categories (such as "strongly agree," "agree," "disagree," and "strongly disagree") that are used to determine relative intensity.

Lower limits: The lowest number in a confidence interval.

Mean: The value obtained by dividing the sum set of numbers by the total amount of numbers within the set. An average.

Median: Another average, representing the "middle" value in an ordered data set. When data are listed from smallest to biggest, the median is the middle-most number.

Mock data: Sample data used to illustrate a concept or methods for analysis.

Monitoring: The periodic collection and evaluation of data relative to stated project goals, objectives, and activities.

Monitoring plan: An outline for the steps taken to ensure that a project is on track. A complete monitoring plan lists a project's audience, information needs, data collection strategies, indicators, methods that will be used to collect data, and when, by whom, and where data will be collected.

Monitoring team: The group of people involved in developing and implementing a monitoring plan.

Multiparty monitoring: Monitoring involving a diverse group of stakeholders to promote mutual learning and respond to concerns.

Nominal data: Data used make frequency distributions.

Non-representative sample: see "Representative sample."

Null hypothesis: A formal statement that there is no association between two variables.

Open-ended question: A question specifically worded to encourage respondents to provide their own answers. Open-ended questions often begin with words like "why," "what," "how," or "where."

Outlier: In qualitative data, an outlier is a data point that does not fit the predominant pattern or theme being used to describe the data set. In quantitative data, outliers are data points that are outside the range of the data values that we want to describe. Outliers can be due to an error in measurement of the value or simply a rare chance event.

Percentage: A method used to represent the parts of a whole. The relative frequency of a value or response. Calculated by totaling the number of times a particular value or response occurs in a data set, then dividing that number by the total number of data points in the data set.

Pie chart: A method used to graphically represent percentage data.

Population: A clearly defined set of individuals who represent all of the individuals with a specific set of characteristics. For example, all of the trees in plot 1.

Prescription: Any set of management actions designed to achieve a specific goal.

Qualitative data: These are data are described with words rather than having a numerical value.

Quantitative data: These are data that have a numerical value.

Questionnaire: A written survey document listing questions designed to solicit information appropriate for analysis.

Random sampling: A method for selecting a portion of the population for data collection. Every individual has an equal chance of being selected.

Range: The highest value minus the lowest value in a data set. An indication of the total spread of the data.

Representative sample: A sample which has the same distribution of characteristics as the population from which it was selected. Representativeness is enhanced by random sampling and allows you to generalize beyond your sample to the larger population and allows the use of inferential statistics.

Respondent: A person who provided data for analysis by responding to interview or survey questions.

Sample: For statistical purposes a subset of a population.

Significance: The probability that a result is likely to be due to chance alone. By convention, a difference between two groups is usually considered statistically significant if chance could explain it only 5 percent of the time or less.

Skew: A data distribution that is not symmetric, meaning there are more large numbers or more small numbers. If, for example, most respondents on a ten-point scale rated the product a nine or ten, we would describe that distribution as "skewed." Refers to the general shape of a distribution of scores when graphed as a frequency polygon.

Stakeholder: Person who has vested interst in the natural resources or who potentially will be affected by project activities.

Standard deviation: The average amount that all data values for some variable differ from the average (mean). The square root of the variance.

Standard error: A measure of the extent that the mean of your sample data is expected to differ from the mean of the entire population. Calculated by dividing the standard deviation by the square root of the sample size. Often abbreviated SE.

Survey: A set of questions posed to a group of subjects about their attitudes, beliefs, plans, lifestyles, or any other variable of interest. Surveys may be conducted in person, over the phone, or through the use of a written form (a questionnaire).

Transparent: Easily accessed and understood; obvious in structure and meaning. Transparency means that all project information, including goals, actions, and accomplishments, is available to all and clearly understood by everyone.

Treatment: A management action designed to achieve a specific outcome.

Triangulation: Collecting data from multiple sources or with multiple methods to cross-check and validate data. Also, having multiple team members analyze data to cross-check and validate data interpretation.

Upper limits: The highest number in a confidence interval.

Variance: A measure of the average distance that data points in a set spread from the mean of that data set.

x-axis: The left-to-right axis that represents the independent variable.

y-axis: The up-and-down axis that represents the dependent variable.

z-score: A specific value related to each confidence level which is used to determine the confidence interval.

Appendix 8 – References

Babbie, E. 1998. *The Practice of Social Research*, 8th Edition. New York: Wadsworth Publishing Company.

Brown, J.K. 1974. Handbook for inventorying downed woody material. General Technical Report INT-16. Ogden, Utah: USDA Forest Service, Intermountain Forest and Range Experiment Station.

Busch, D.E. and J.C. Trexler. 2003. *Monitoring Ecosystems: Interdisciplinary Approaches for Evaluating Ecoregional Initiatives*. Washington, D.C.: Island Press.

Butler, L.M., C. DePhelps, and R. E. Howell. 1995. Focus Groups: A Tool for Understanding Community Perceptions and Experiences. WREP 128. Accessed online at cru.cahe.wsu/edu/CEPublications/wrep0128.html.

Cooperrider, A.Y., Boyd, R.J., and H.R. Stuart. 1986. *Inventory and Monitoring of Wildlife Habitat*. Washington, D.C.: Government Printing Office.

Elzinga, C.L., D.W. Salzer, J.W. Willoughby, and J.P. Gibbs. 2001. *Monitoring Plant and Animal Populations*. New York: Blackwell Science.

Gibson, H.R. 1994. *Elementary Statistics*. New York: Wm. C. Brown Publishers.

Graziano, A.M. and M.L. Raulin. 1997. *Research Methods: A Process of Inquiry*, 3rd Edition. New York: Addison-Wesley Educational Publishers.

Hayes, A.F., T.L. Gee, I.R. Price, R. Cooksey, and A. Patrech. 1993. Research Methods and Statistics: Lecture and Commentary Notes. School of Psychology, University of New England, Armidale, New South Wales, Australia.

Accessed online at
www.une.edu.au/WebStat/unit_materials/.

Margoluis, R. and N. Salafsky. 1998. *Measures of Success: Designing, Managing, and Monitoring Conservation and Development Projects*. Washington, D.C.: Island Press.

Miles, M.B. and A.M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd Edition. Thousand Oaks, California: Sage Publications.

Morrison, M.L. 2002. *Wildlife Restoration: Techniques for Habitat Analysis and Animal Monitoring*. Washington, D.C.: Island Press.

Motulsky, H. 1995. *Intuitive Biostatistics*. New York: Oxford University Press.

Ott, R. L., and M. Longnecker. 2001. *Statistical Methods and Data Analysis*, 5th Edition. New York: Duxbury Publishing.

Pagano, M. and K. Gauvreau. 2000. *Principles in Biostatistics*, 2nd Edition. New York: Duxbury Publishing.

Patton, M.L. 2002. *Qualitative Research and Evaluation Methods*, 3rd Edition. Thousand Oaks, California: Sage Publications.

Patton, M.L. 2000. *Understanding Research Methods: An Overview of the Essentials*, 2nd Edition Los Angeles, California: Pyrczak Publishing.

Sackett, S.S. 1980. Woody fuel particle size and specific gravity of southwestern tree species. Research Note RM-389. Fort Collins, Colorado: USDA Forest Service, Rocky Mountain Forest and Range Experiment Station.

Thompson, W.L., G.C. White, and C. Gowan. 1998. *Monitoring Vertebrate Populations*. San Diego, California: Academic Press.

USDA. 1999. Wood handbook: Wood as an engineering material. General Technical Report FPL-GTR-113. Madison, Wisconsin: USDA Forest Service, Forest Products Laboratory.