

6.2 Presenting the Data

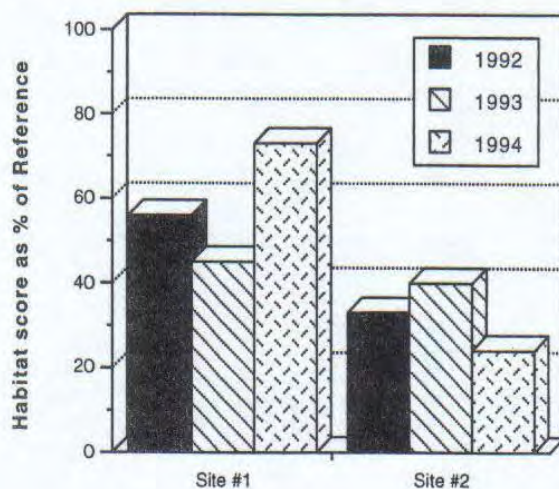
When presenting numerical data, one of your chief goals should be to maintain the attention and interest of your audience. This is very difficult using tables filled with numbers. Most people will not be interested in the absolute values of each parameter at each sampling site. Rather, they will want to know the bottom line for each site (e.g., is it good or bad) and seasonal and year to year trends.

Graphs and charts, therefore, are typically the best way to present volunteer data. Take care, however, that your graphs "fit" your audience and are neither too technical nor too simplistic.

Fig. 6.1

Example of a bar graph displaying biological data

Habitat scores as a percent of reference condition at sites #1 and #2 for 1992-1994



Graphs and Charts

Graphs can be used to display the summarized results of large data sets and to simplify complicated issues and findings. The three basic types of graphs that are typically used to present volunteer monitoring data are:

- Bar graph
- Line graph
- Pie chart

Bar and line graphs are typically used to show results, such as bioassessment scores, along a vertical or y-axis for a corresponding variable (such as sampling date or site) which is marked along the horizontal or x-axis. These types of graphs can also have two vertical axes, one on each side, with two sets of results shown in relation to each other and to the variable along the x-axis.

Bar Graph

A bar graph uses columns with heights that represent the value of the data point for the parameter being plotted. Fig. 6.1 is an example using fictional data from Volunteer Creek.

Line graph

A line graph is constructed by connecting the data points with a line. It can be effectively used for depicting changes over time or space. This type of graph places more emphasis on trends and the relationship among data points and less emphasis on any particular data point.

Fig. 6.2 is an example of a line graph again using fictional data from Volunteer Creek.

Pie chart

Pie charts are used to compare categories within the data set to the whole. The proportion of each category is represented by the size of the wedge. Pie charts are popular due to their simplicity and clarity. (See Fig. 6.3)

**June phosphorus concentrations
at Sites #1 and #2 from 1991 - 1997**

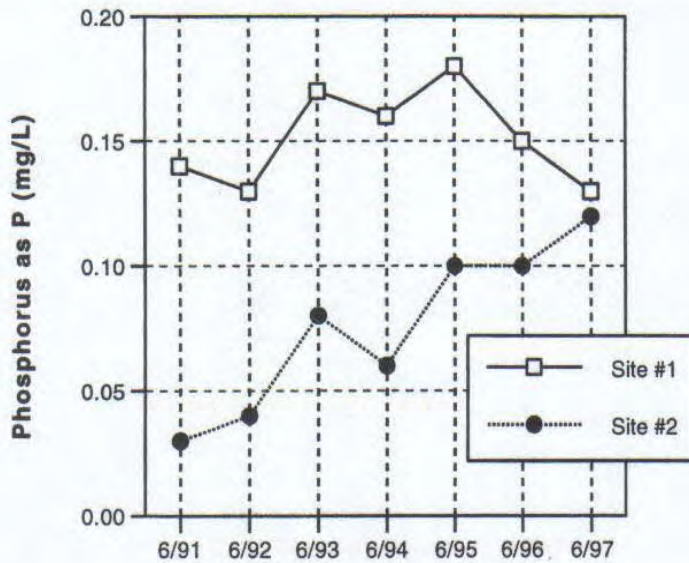


Fig. 6.2

Example of a line graph depicting trends in phosphorus data

Graphing Tips

Regardless of which graphic style you choose, follow these rules to ensure you use them most effectively.

- *Each graph should have a clear purpose.* The graph should be easy to interpret and should relate directly to the content of the text of a document or the script of a presentation.
- *The data points on a graph should be proportional to the actual values so as not to distort the meaning of the graph.* Labeling should be clear and accurate and the data values should be easily interpreted from the scales. Do not overcrowd the points or values along the axes. If there is a possibility of misinterpretation, accompany the graph with a table of the data.
- *Keep it simple.* The more complex the graph, the greater the possibility for misinterpretation.

Summary of water quality ratings for Volunteer Creek



(Total no. of stations= 52)

Fig. 6.3

Example of a pie chart summarizing water quality ratings

- *Limit the number of elements.* Pie charts should be limited to five or six wedges, the bars in a bar graph should fit easily, and the lines in a line graph should be limited to three or less.
- *Consider the proportions of the graph and expand the elements to fill the dimensions, thereby creating a balanced effect.* Often, a horizontal format is more visually appealing and makes labeling easier. Try not to use abbreviations that are not obvious to someone who is unfamiliar with the program.
- *Create titles that are simple, yet adequately describe the information portrayed in the graph.*
- *Use a legend if one is necessary to describe the categories within the graph.* Accompanying captions may also be needed to provide an adequate description of the elements.

Summary Statistics

Summary statistics can reduce a very large data set to a few numerical values that can then be easily described and analyzed. Such statistics include the mean and standard deviation—two of the most frequently used descriptors of environmental data.

Textbook statistics commonly assume that if a parameter is measured many times under the same conditions, then the measurement values will be randomly distributed around the average with more values clustering near the average than further away. In this ideal situation, a graph of the frequency of each measure plotted against its magnitude should yield a bell-shaped or normal curve. The *mean* and the *standard deviation* determine the height and breadth of this curve, respectively.

The mean is simply the sum of all the measurement values divided by the number

of measurements. This statistic is a measure of location and in a normal curve marks the highest point at the center of the bell.

The standard deviation, on the other hand, describes the variability of the data points around the mean. Very similar measurement values will have a small standard deviation while widely scattered data will have a much larger standard deviation.

While both the mean and standard deviation are quite useful in describing stream data, often the actual measures do not fit a normal distribution. Other statistics often come into play to describe the data. Some data are skewed in one direction or the other. Other data may have a flattened bell shape.

It is important to note that biological information often does not follow normal, bell-shaped distribution. This is because biological communities are dynamic, complex, and interdependent systems; many factors influence them, and these cannot be statistically predicted. For example, bioassessment scores plotted against habitat assessment scores will be at their best when habitat quality is at its best. For data that is nonnormally distributed, the mean and the standard deviation are not appropriate summary statistics.

For describing nonnormally distributed data, it is best to use statistics that can convey the information for a variety of conditions and which are not overly influenced by the data points at the extremes of the distribution. The median and the interquartile range are two statistics that are commonly used to describe the central tendency and the spread around the median, respectively. These statistics are derived by placing the data points in order of value from lowest to highest. The median is simply the value that is in the middle of the data set. The interquartile range is the difference between the value at the 75 percent level and the value at the 25 percent level.

The best method for presenting this type of data is called a box and whisker plot. One simple box and whisker plot will graphically display the following information:

- Median
- Variability of the data around the median
- Skew of the data
- Range of the data
- Size of the data set

Statistical software packages for computers will easily construct box and whisker plots. You can construct these plots by following procedure shown below:

1. Order the data from the lowest to the highest.
2. Plot the lowest and highest values on the graph as short horizontal lines. These are the extreme values of the data set and represent the data range.
3. Determine the 75 percent value and the 25 percent value of the data set. These values define the interquartile range and are represented by the location of the top and bottom lines of the box.
4. The horizontal length of the lines that define the top and bottom lines of the box (the box width) can be used as a relative indication of the size of the data set. For example, the box width that describes a data set of 20 values can be displayed twice as wide as a data set of 10 values. Any proportional scheme can be used as long as it is consistently applied.
5. Close the box by drawing vertical lines that connect to the ends of the horizontal lines.
6. Plot the median inside the box.

Fig. 6.4 is an example depicting the extreme values, interquartile range, and median of biosurvey metric scores from 52

sites sampled in Volunteer Creek in June, 1995.

Maps

Displaying the results of your monitoring data on a map can be a very effective way of showing the data and helping people understand what it means. A map shows the location of sample sites in relation to land features, such as cities, wastewater treatment plants, farmland, and tributaries that may have an effect on water quality. Because a map also displays the stream's relationship to neighborhoods, parks and recreational areas, it can help to develop concern for the stream and strengthens interest in protecting it.

Choosing a Map

It is best to have two types of maps. One should be a working map with a lot of detail. The other should be used for display

Box Plot of Total Metric Scores from June, 1995
(No. of sites = 52)

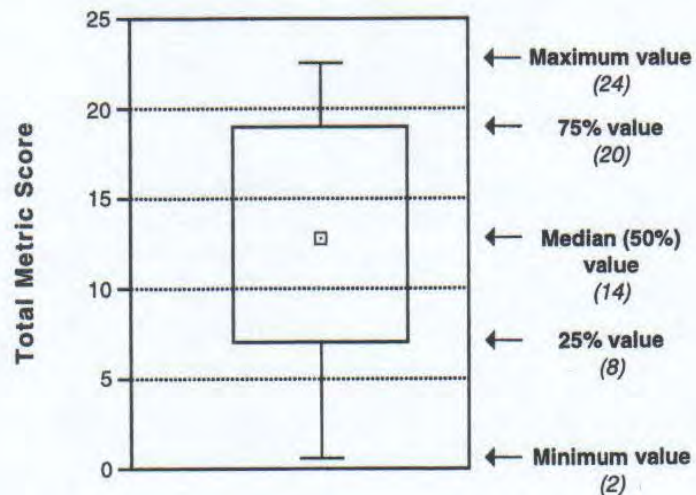


Fig. 6.4
Example of a box plot

purposes. The working map should include important features such as:

- Stream and its tributaries
- Wetlands
- Lakes and ponds
- Cultural features such as roads
- Rail and power lines; municipal boundaries
- Some indication of land use patterns and vegetation.

The map should be of a scale large enough to add the location of sample sites.

U.S. Geological Survey (USGS) 7.5 minute quads (scale of 1:24,000; 1 in. = 2,000 ft) are available with and without topographic contours (elevation markings). These maps are available for the most of the United States.

The USGS maps are particularly useful if your information will be incorporated into a geographic information system (GIS), since many of these systems use the USGS maps as base maps. For your data to be used in a GIS, it is likely that you will have to provide the latitude and longitude of your sample sites, which can be obtained by using the grid markings on the USGS topographic maps. Several different coordinate systems are marked, including standard latitude/longitude and the Universal Transmercator coordinates. For assistance in learning how to use these coordinate markings, talk to the local USGS office or someone in the geography department at a university. It may also be possible for the GIS office you work with you to “digitize” the maps, thus saving you the trouble of trying to calculate the coordinates.

The display map is best used to illustrate your program results at public meetings or in reports. This map should be simpler than the detailed map and show only principal features such as roads, municipal boundaries, and waterways. It should have sufficient detail and scale to show the location of sample sites, and have

space for summary information about each of the sample sites. Commercial road atlases and county or town road maps available from state transportation departments are examples of the types of maps that can be used for display purposes (See Fig. 6.5).

Creating a Display Map

Some suggestions for using a map to display your data include:

- Keep the amount of information presented on each map to a minimum. Do not try to put so much on one map that it becomes visually complicated and difficult to read or understand. Use another map to display a different layer or “view” of the data. For example, if there are several dates for which you wish to display sampling results, use one map for each date.
- Clearly label the map and provide an explanation of how to interpret it. If you need a long and complicated explanation, you may want to present the data differently. If you have reached a clear conclusion, state the conclusion on the map. For example, if a map shows that tributaries are cleaner than the mainstem, use that information as the subtitle of the map.
- Provide a key to the symbols that are used on the map.
- Rather than packing lots of information into a small area of the map, use a “blow-up” or enlargement of the area elsewhere on the map to adequately display the information.
- Use symbols that vary in size and pattern to represent the magnitude of results. For example, a site with a fecal coliform level of 10 per 100 milliliters could be a light gray circle one-sixteenth inch diameter while a site with a level of 200 per 100

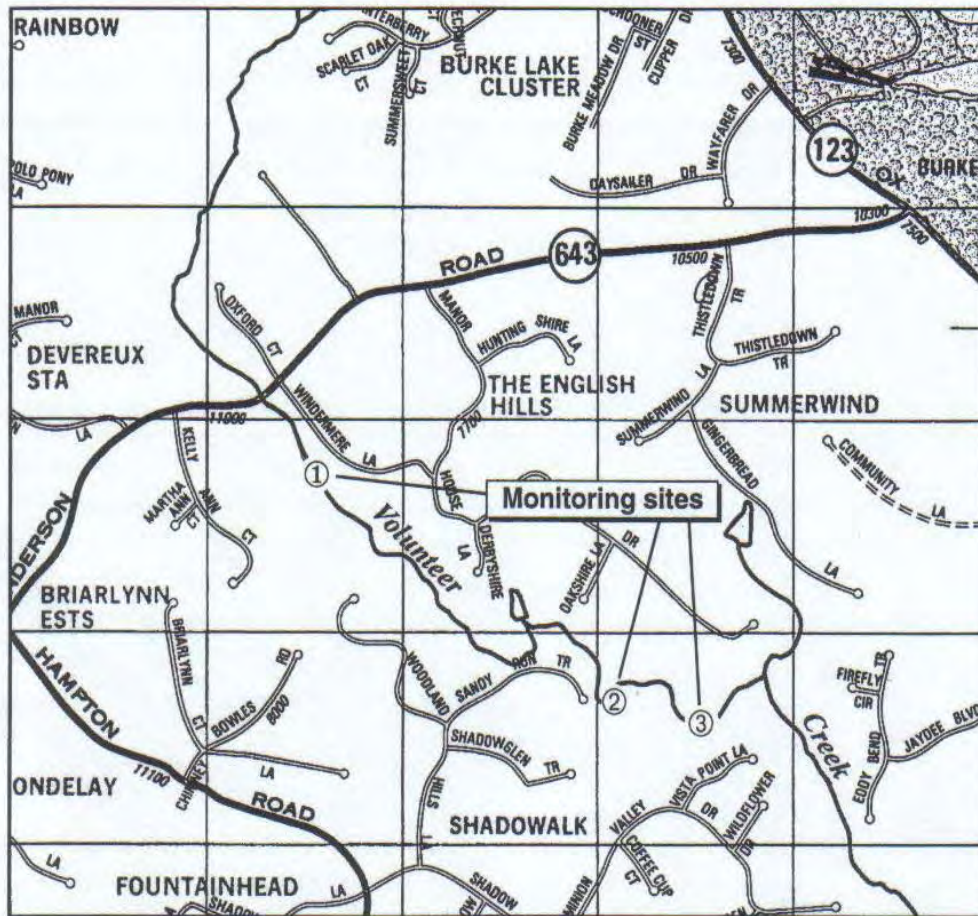


Figure 6.5

A road map is useful for displaying station locations.

millimeters would be a dark gray circle one-quarter inch diameter. Start by finding the highest and lowest values, assign diameters and patterns to those and then fill in steps along the way. For the above example you might have four ranges: 0 to 99, 100 to 199, 200 to 500 and 500 +.

Maps on Demand

EPA provides a World Wide Web service known as Maps on Demand that allows users to generate maps displaying environmental information for anywhere in the U.S. (except Hawaii, Puerto Rico, and the Virgin Islands). Types of information that can be mapped include EPA-regulated facilities, demographic information, roads, streams, and drinking water sources. Maps of varying scales can be generated on the site (latitude and longitude), zip code, county, and basin levels. Submit your request and email address, and after a brief wait, you will be able to view your map on-line or download it. Maps on Demand can be reached through EPA's *Surf Your Watershed* homepage at www.epa.gov/surf/info.html.