In Context Toolbox – How to Create Visual Representations of Data

The ***In Context Toolbox*** tip sheets are designed to help middle school and high school researchers prepare a written report. This document will explain how to **create visual representations of data like charts and graphs.**

Need more help with your report? Visit Help in any In Context resources for more **In Context Toolbox Tip Sheets**!

One of the best ways to bring your report to life is to add visual elements. Images do more than just make your paper look pretty - they make it easier for readers to digest the information in your report. This is especially true when your report contains numerical data. If you've taken the trouble to compile some good, old-fashioned hard data, don't just present the raw numbers in text form. Put your data on display! Create graphs to show off your hard work.

Displaying Numbers

Graphs are ideal for communicating the meaning of numerical data. Statistics, which can be difficult to interpret when buried in the text of a report, can be very easy to understand when presented in graphic form. When you use a well-made graph to explain the significance of some stats you've gathered, readers of your report are likely to say, "I get it!" But in order to get this "I get it!" reaction, you have to display your stats in a graph that's appropriate for the kind of data you've collected.

As shown in the table below, the type of values your variables permit determines the best graph to display your data. But how can you tell if your variables have values that are discrete, continuous, or some sort of percentage? Let's consider some examples and see how the process works.

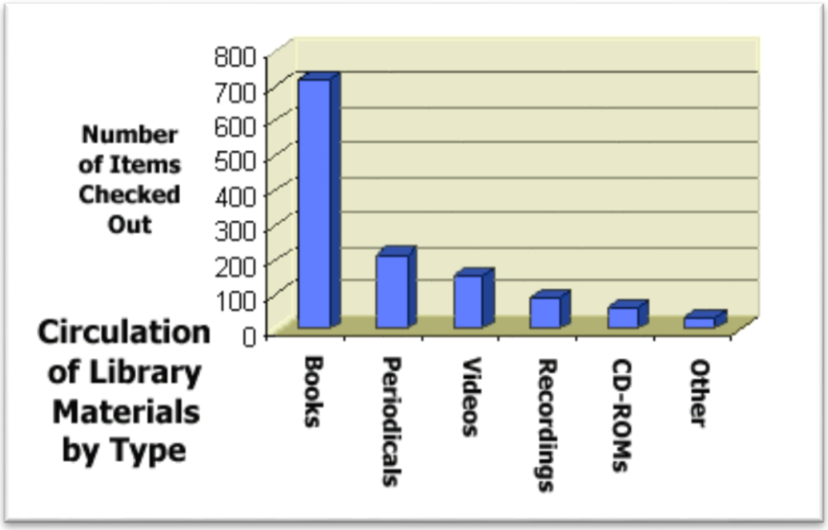
Variable Value Type	**Ideal Graph**
Discrete	**Bar Graph**
Continuous	**Line Graph**
Percentage (portion of a whole)	**Pie Chart**
Percent Completed (rate of progress)	**Line Graph**

Example 1

Say you've been asked to prepare a report about the kinds of materials students check out at your school's library. Assume that you examined your school library's circulation records for four weeks and discovered the following facts:

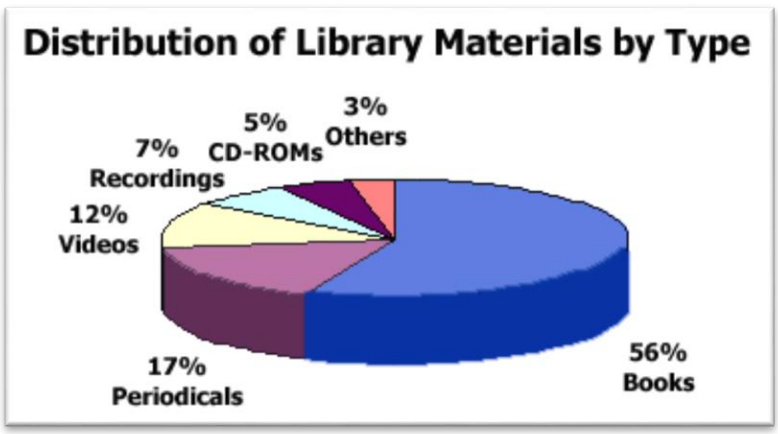
A total of 1,282 materials were checked out from the library during the period. 718 were books (56% of the total). 218 were periodicals (17% of the total). 154 were videos. (12% of the total). 90 were sound recordings. (7% of the total). 64 were CD-ROMs. (5% of the total). 38 were materials that don't fit into any of the above categories. (3% of the total).

Bar Graph Using Example 1



The circulation figures lend themselves to display via a bar graph. Notice that the dependent variable, the one that you'd plot along the vertical axis in a graph, is the number of items checked out. The independent variable, the one that you'd plot along the horizontal axis, is the type of material (books, periodicals, etc.) Notice further that there are six distinct variations for the independent variable: books, periodicals, videos, sound recordings, CD-ROMs, and other. These variations are entirely separate and indivisible - there are no gradations involved. You can't divide the variation "book" into measurable portions - either an item is a book or it isn't. Variables of this kind are known as discrete variables. With **discrete variables**, there are no "in-between" values. Since there are no "in-between" values, discrete variable data should be displayed in separate chunks - one chunk for each of the variable's variations. The bars in a bar graph display data in exactly this way, as you can see in the graph at right.

Pie Chart Using Example 1



The percentage figures lend themselves to display via a pie chart. This is because the percentages we computed in this example represent relative portions of the total. That is, they represent the parts of the whole. If you add the six individual portions together, you get "the whole," or 100%.

Pie charts are the perfect way to show how the parts of a whole are distributed. It's easy to see why. Each slice of the pie represents a relative share of the whole. The size of each pie slice is determined by the percentage figure: the higher the percentage, the bigger the slice. For example, the **Books** slice in the graph at left accounts for the lion's share of the pie. This is because, at 56% of the total, books account for more than half of the items checked out of the library. Really drives home the point, doesn't it? So, whenever you want to display data that's broken down into percentages that add up to 100%, use a pie chart.

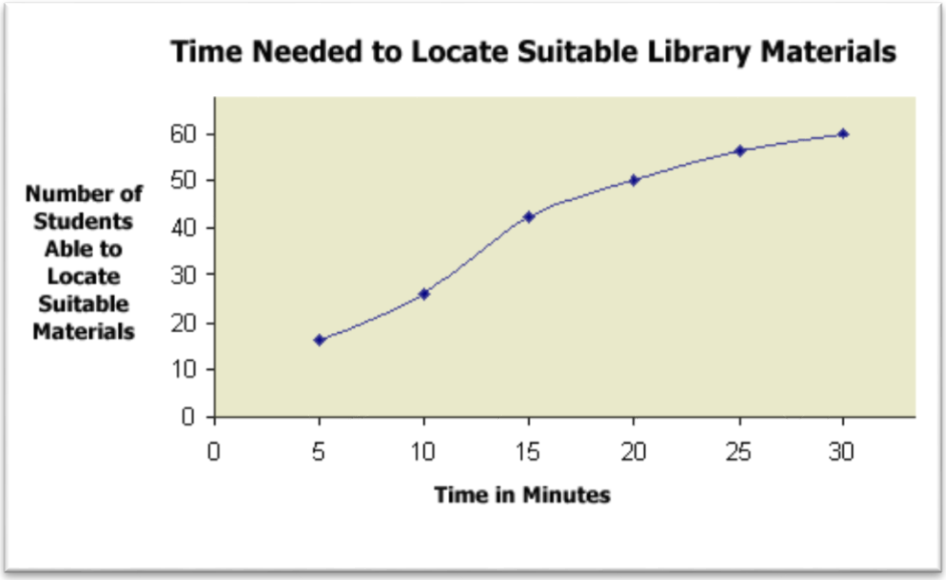
Example 2

Let's imagine you are asked to do a second report about your school library. This time, you are invited to observe a group of 60 students doing library research. Your task is to find out how long it takes them to find the materials they need. You decide to track the students' progress at five minute intervals. Here's what you discover:

After five minutes, 16 students (27%) had found the materials they needed. At ten minutes, 26 students (43%) had found the materials they needed. At fifteen minutes, 42 students (70%) had found the materials they needed. At twenty minutes, 50 students (83%) had found the materials they needed. At twenty-five minutes, 56 students (93%) had found the materials they needed. By thirty minutes, all 60 students (100%) had found the materials they needed.

In Context Toolbox – How to Create Visual Representations of Data

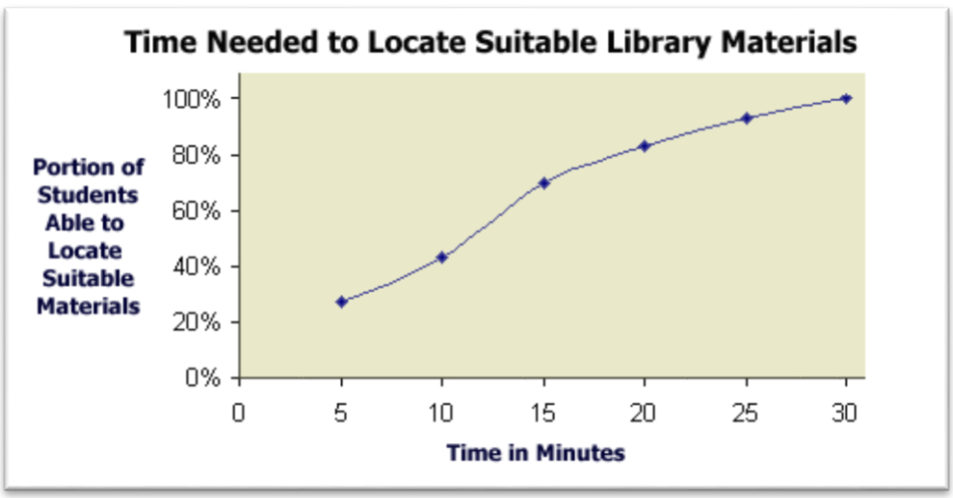
Line Graph Using Example 2



Just as in the first example, you have gathered some useful raw data and some interesting percentage data. Does this mean you should display the raw data in a bar graph and the percentage data in a pie chart? No! There are two crucial differences between the data you collected for your first report and the data you collected for your second report.

First, notice that the independent variable in the second data set is the flow of time. Time is a **continuous variable**. You chose to measure your students' progress at five minute intervals, but it's important to recognize that time has measurable values in between these intervals. Unlike discrete variables, continuous variables always have these "in-between" values. You can't possibly measure all of the "in-between" values of a continuous variable, but you can represent your data in a graph that allows readers to take a guess as to what the "in-between" values might be. The graph to use is a line graph, such as the example below.

In this graph, your actual observations are marked on the graph as dots. These six dots are connected by a smooth line. This smooth line gives readers a sense of how the data change gradually over time. It allows readers to make educated guesses as to what the "in-between" values might be, even though you didn't measure them! This is why a line graph is the ideal way to represent a continuous variable.



The second crucial difference pertains to the percentage figures in example 2. The percentages in these data don't represent parts of the whole; instead, they represent progress towards completion over time. You can tell that this is true by adding the percentages together - they don't add up to 100%. It would be meaningless to display these percentage figures in a pie chart. But a line graph works just fine. Simply plot the percentages on the vertical axis as the dependent variable, and the result is a very meaningful line graph:

Notice that the curve of the second line graph looks quite like the curve in the first line graph. Both graphs show upward momentum toward a goal of 100% completion. However, the use of percentage figures in the second graph makes it clear that the graph is designed to show the rate of progress towards completion. What's more, the use of percentages in the second graph makes it much easier to compare data across multiple graphs. Many readers find graphs with percentage completed figures plotted on the vertical axis easier to interpret than graphs that have raw numbers on the vertical axis.

Need more help? Ask your librarian!