

VISUALIZING DATA

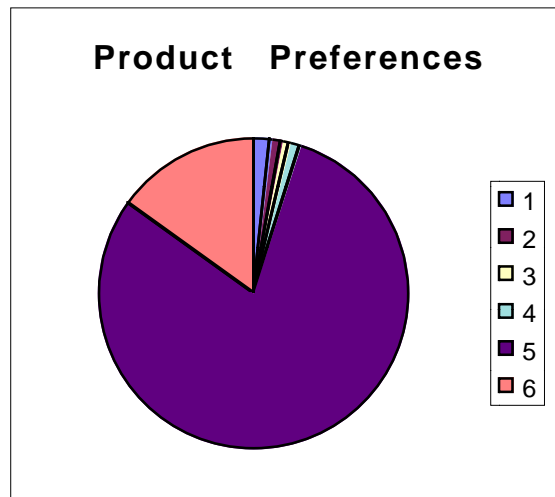
By Senior Associate Jeff Johnson, McREL Eisenhower High Plains Consortium for Mathematics and Science

The mathematics and science education communities have recognized the importance of students' abilities to visualize data and to make inferences from various forms of graphical representation. But the choices that are made in how data is to be displayed will influence the graphical appearance, and as a result, such choices can have a direct impact on the viewer's interpretations.

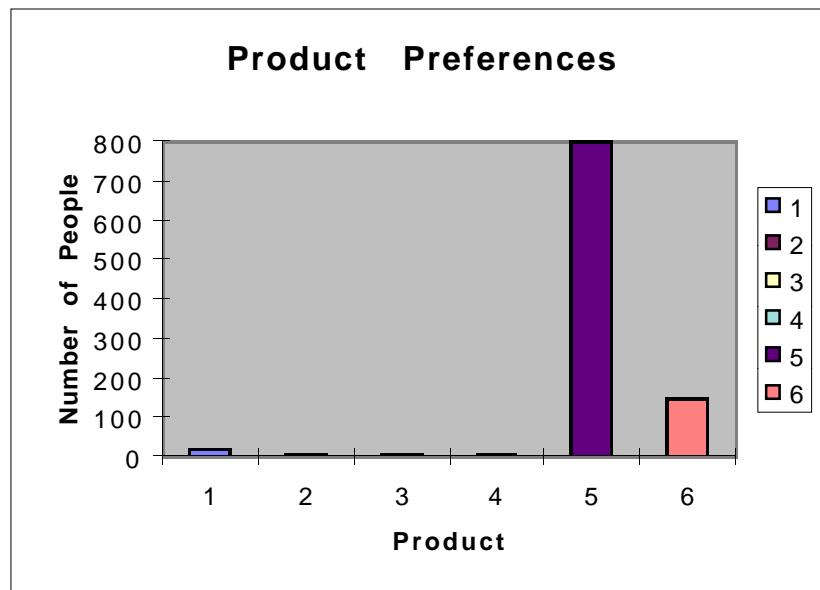
The Graph Decision

The decision to use a particular type of graph, such as a pie graph or a histogram, is influenced by the purpose for graphing the data and the insights that can be gained from various representations of the data. Viewing the data from other perspectives can sometimes suggest the need for further investigation or to support a particular conclusion.

A pie graph is useful for representing percents of a whole. For instance, the percentages of people preferring each of six different products have been summarized in the pie graph below. While we can make gross observations with respect to the products having larger percentages, it would be difficult to draw any conclusions about the less preferred products if the percentages are not indicated.



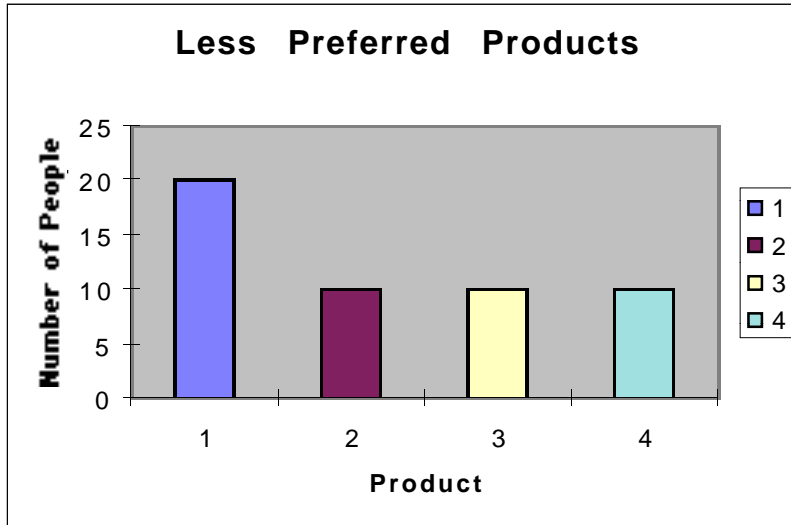
The histogram is another common graphical representation to allow for visual comparisons of data, but as in the case of pie graphs, there are limitations in the comparisons to be made if the bars of the histogram are very small. In the histogram below, it is difficult to tell the difference between the preferences of products 1, 2, 3 and 4.



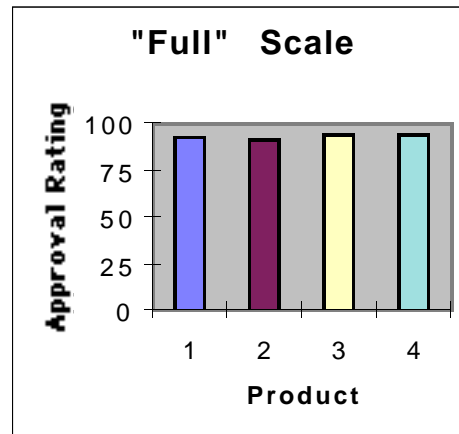
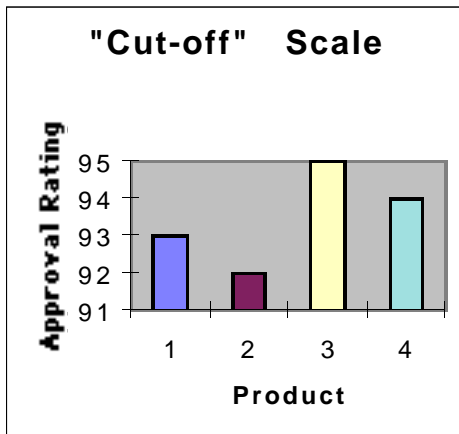
The Scale Decision

Given the problem above, it may be necessary to select an alternative representation to allow for a closer examination of the data. The choice of an “appropriate” scale can be used to highlight certain data and consequently influence the reader’s interpretation. Sometimes small differences in data can be significant, but in an effort to highlight the differences, there is the danger of misrepresenting the data. This may result in exaggerated differences in some cases, and in other cases, significant differences may be overlooked.

One solution to the problem above would be to separate the histogram into two parts, using a smaller scale to allow for a closer examination of the less preferred products.



However, care should be taken to avoid possible misinterpretations when using histograms. For instance, advertisements often show a type of histogram with a “cut-off” scale. In the example below, the approval ratings of four products are indicated on a scale of 0 to 100. If a cut-off scale is used (left), the viewer can be led to believe that the differences in the four product ratings are greater than they are. In essence, the scale is not consistent, and the distance from 0 to 91 in the vertical scale is ignored. Furthermore, some readers may incorrectly conclude that product 3 has a rating four times that of product 2. But when the same data is displayed using a “full” scale (right), the differences are put into perspective.



Conclusion

In summary, decisions about the type of graph and the choice of scale should reflect the benefits to be gained from the graphical representation of data. The graphs need to be both accurate and effective to allow for visually enhanced understanding.