Section 1.8 - Misleading Graphs and Statistics

It is a well-known fact that statistics can be misleading. They are often used to prove a point, and can easily be twisted in favor of that point! The purpose of this section is to learn how to recognize common statistical deception in order to avoid being misled.

Bad Sampling

When you use a sample to represent a larger group, you must make sure that the people in the sample are fairly representative of the larger group.

Example 31.1

Decide whether a mall is a good place to find a sample for a survey about the amount of allowance received by people ages 10 to 15.

Solution.

The mall is probably not a representative place to find a fair sample of people in this age range. Taking a sample at the mall might not represent fairly those people who receive a small allowance, or none

Misleading Graphs

Good graphs are extremely powerful tools for displaying large quantities of complex data; they help turn the realms of information available today into knowledge. But, unfortunately, some graphs deceive or mislead. This may happen because the designer chooses to give readers the impression of better performance or results than is actually the situation. In other cases, the person who prepares the graph may want to be accurate and honest, but may mislead the reader by a poor choice of a graph form or poor graph construction.

The following things are important to consider when looking at a graph:

1. Title
2. Labels on both axes of a line or bar chart and on all sections of a pie chart
3. Source of the data
4. Key to a pictograph
5. Uniform size of a symbol in a pictograph
6. Scale: Does it start with zero? If not, is there a break shown
7. Scale: Are the numbers equally spaced?

Scaling and Axis Manipulation

A graph can be altered by changing the scale of the graph. For example, data in the two graphs of Figure 31.1 are identical, but scaling of the Y-axis changes the impression of the magnitude of differences.

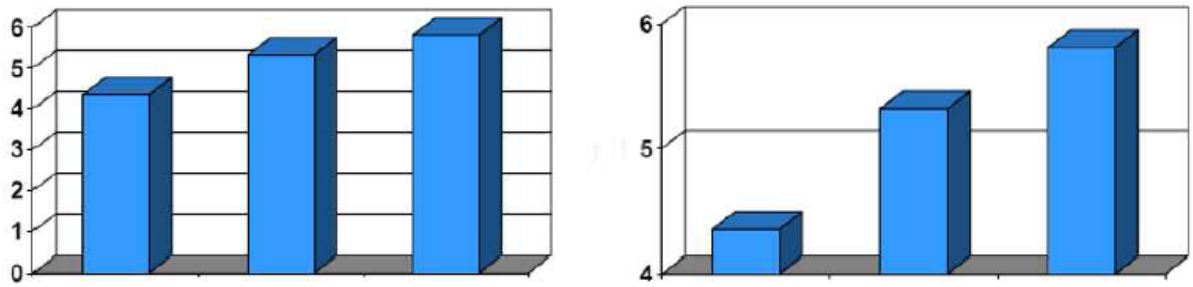


Figure 31.1

Example 31.2
Why is the bar chart below misleading? How should the information be represented?

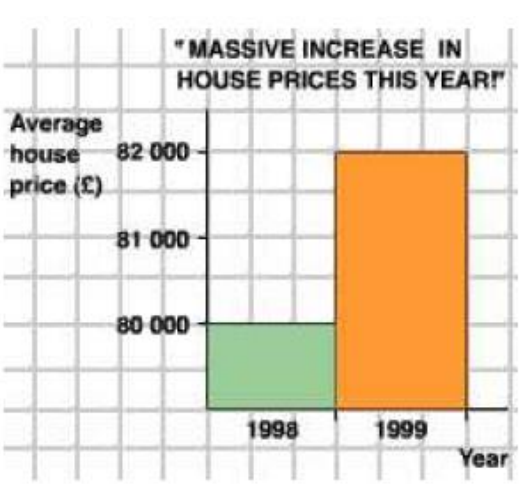


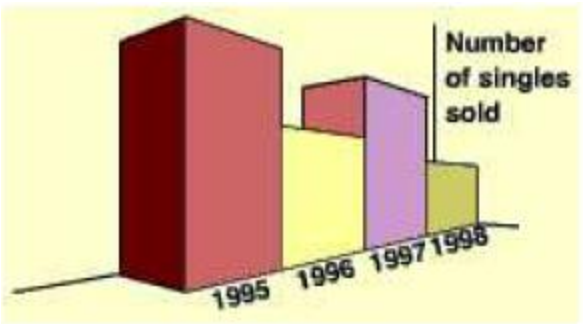
Figure 31.2

Solution.
The bar chart indicates that house prices have tripled in one year. The scale of vertical must start at 0 and that's not the case. A less misleading graph would look like the one in Figure 31.2. This gives a much more accurate picture of what has happened.

Example 31.3
What is wrong with the information represented on this graph?

Solution.
Although the vertical scale starts at 0, it does not go up in even steps. This distorts the graph, and makes it look as though the biggest jump is between 1 and 2 rather than 3 and 4. Also, there are no labels on the axes so we have no idea what this graph represents!

Three Dimensional Effects
Example 31.4
What is wrong with this 3D bar chart?



Solution.
This 3D bar chart might look very attractive, but it is also very misleading. There is no scale on the vertical axis, and because of the perspective it looks as though the sales for 1995 were far greater than those for any other year. In fact they were identical to those for 1997. It would be much better to draw a 2D bar chart like the one shown in Figure 31.3 with the appropriate labelling on each axis:

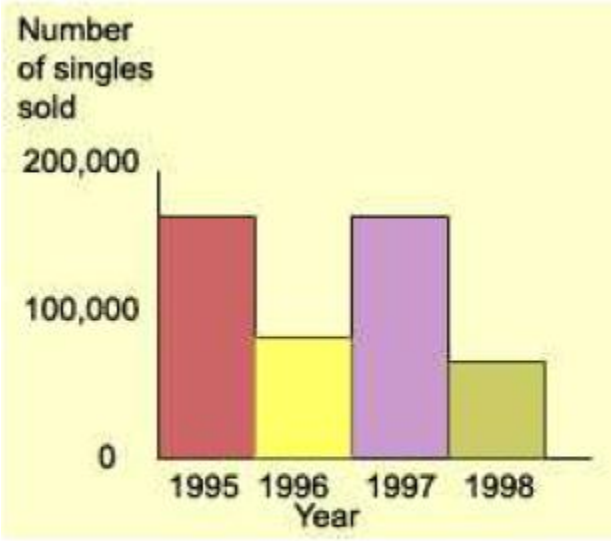
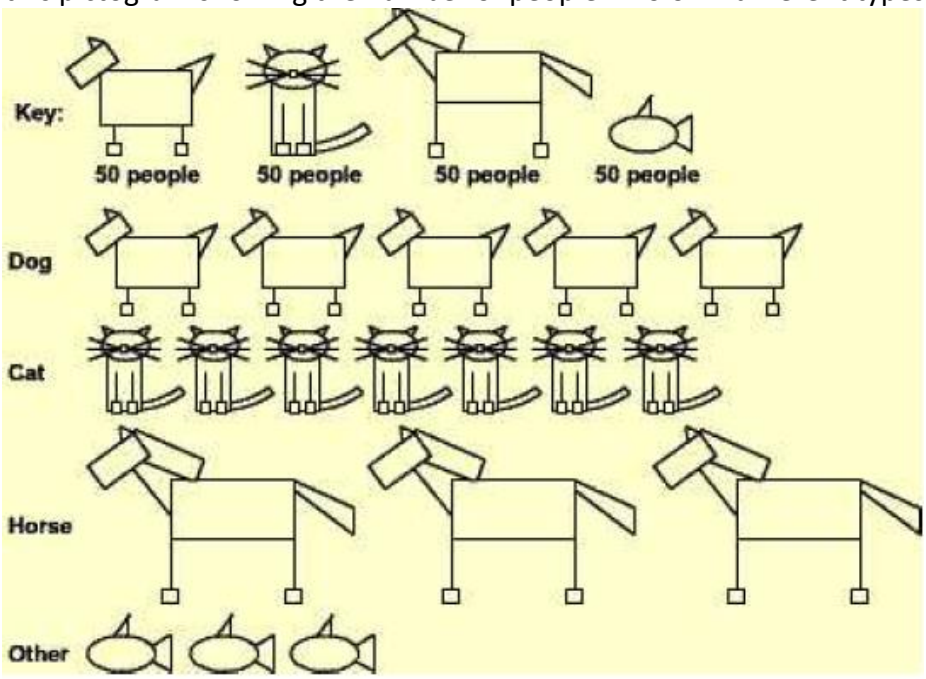


Figure 31.3

Deceptive Pictographs
Example 31.5
What is wrong with this pictogram showing the number of people who own different types of pets?



Solution.
On this pictogram there isn't a category for those people who do not own a pet. The pictures are different sizes and it appears that more people own a horse than any other animal. An improvement would be to redraw the pictogram with each of the animals the same size and aligned with one another as shown in Figure 31.4.

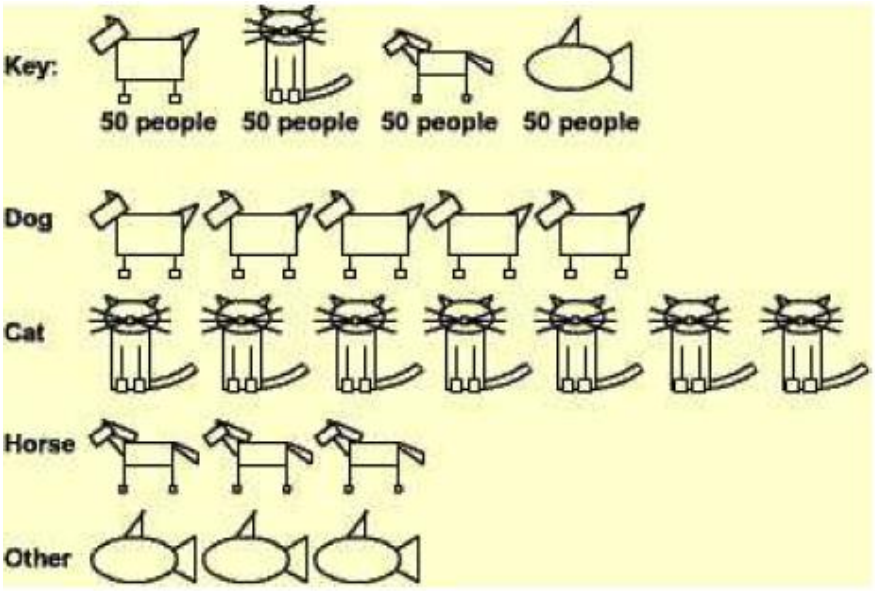
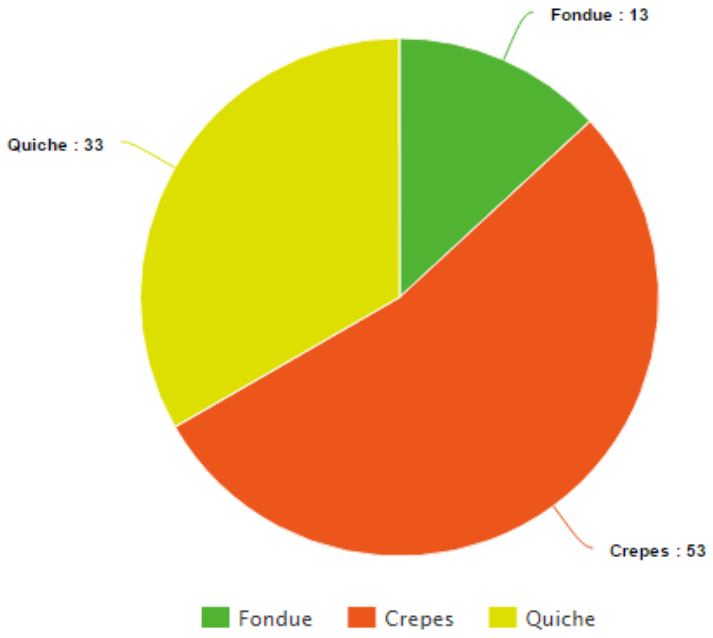


Figure 31.4

Example 31.6
A survey was conducted to determine what food would be served at the French club party. Explain how the graph misrepresents the data.



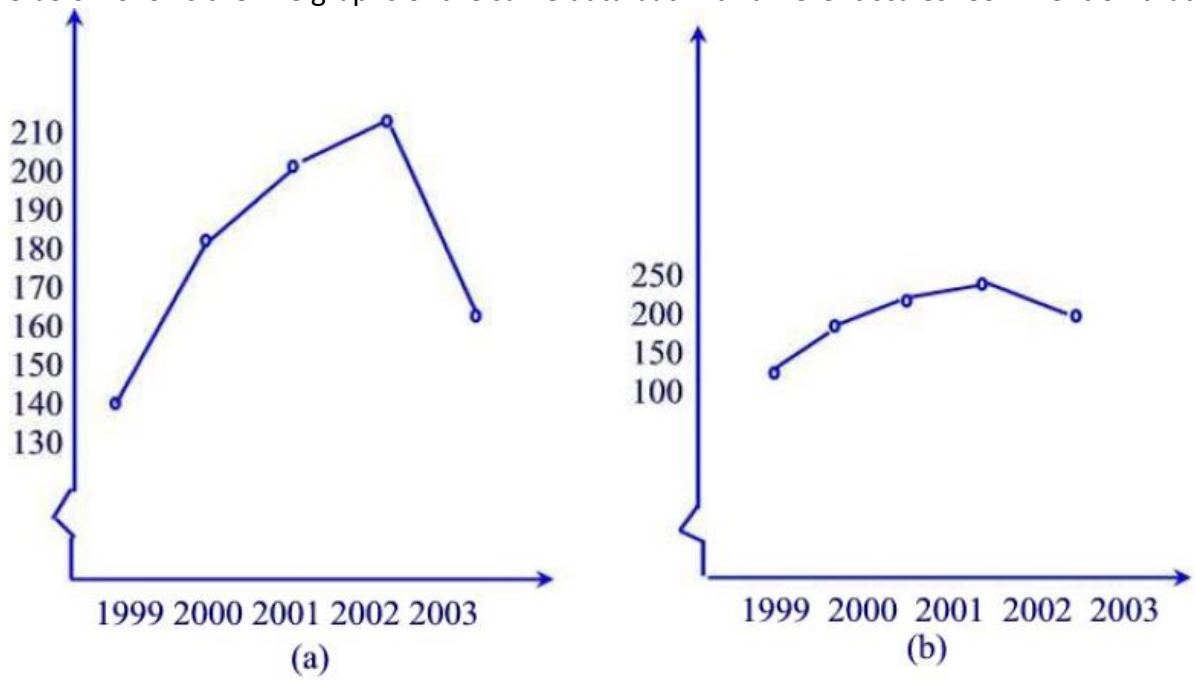
Solution.
The percentages on the circle graph do not add up to 100.

Example 31.7

The number of graduates from a community college for the years 1999 through 2003 is given in the following table:

Year	1999	2000	2001	2002	2003
# of Graduates	140	180	200	210	160

The figure below shows the line graphs of the same data but with different scales. Comment on that.



Solution.

The two graphs do not convey the same message. In Figure (b) the spacing of the years on the horizontal axis is more spread out and that for the numbers on the vertical axis is more condensed than Figure (a). A college administrator might use a graph like Figure (b) to convince people that the college was not in serious enrollment trouble.