

Lesson

Share

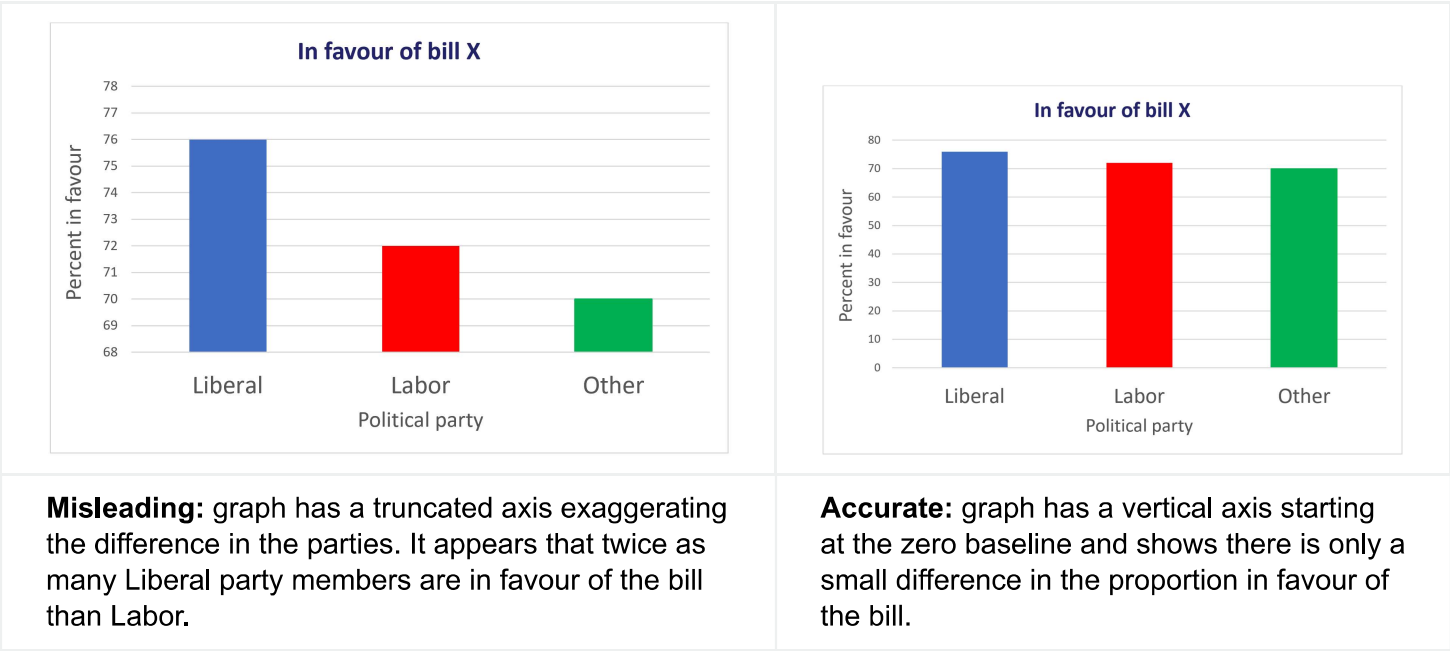
Graphs are an important tool to convey statistics, however, a poorly constructed graph may lead a reader to an incorrect conclusion. A misleading graph may be created to intentionally bias the reader to a particular interpretation or accidentally by someone unfamiliar with creating an effective way to display the given data. Some common features of graphs that may lead to incorrect interpretations are:

- 1. Omitting the baseline
- 2. Showing an inappropriate or irregular scale
- 3. Scale or labels not clearly given
- 4. Leaving data out
- 5. Using pictures or three-dimensional graphics that distort differences
- 6. Using the wrong graph for a given data type

Let's look at an example of each characteristic above.

Omitting the baseline

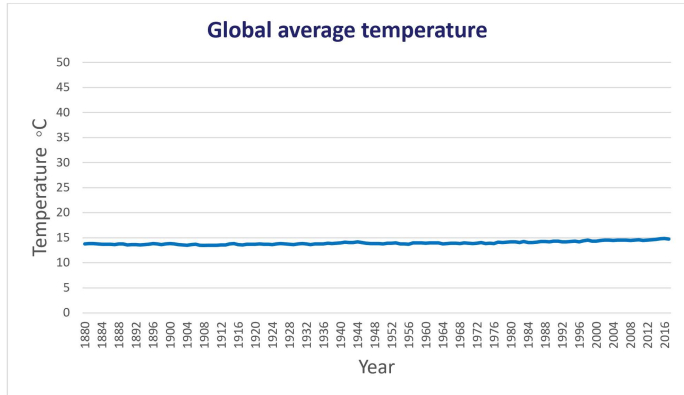
Not starting the vertical axis at zero can give the impression that there is significant difference between values when in fact there may be very little change. This is referred to as a truncated-graph.



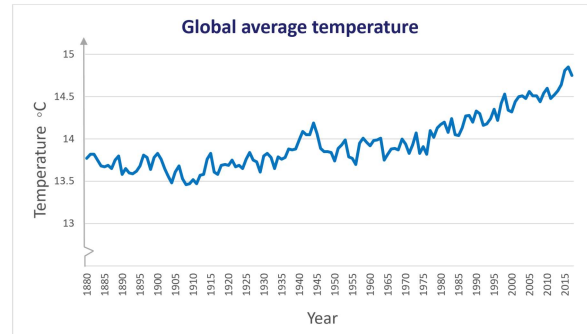
This is not to say the vertical axis always has to start at zero, as we can see in our next example. However, caution should be taken to not exaggerate differences. When not starting the vertical axis at zero a clear indication of a broken scale should be given and it is not recommended for graphs such as column graphs where the viewer takes in the comparative areas of the different categories.

Showing an inappropriate or irregular scale

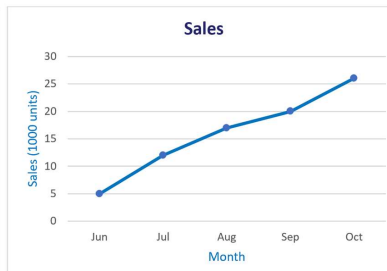
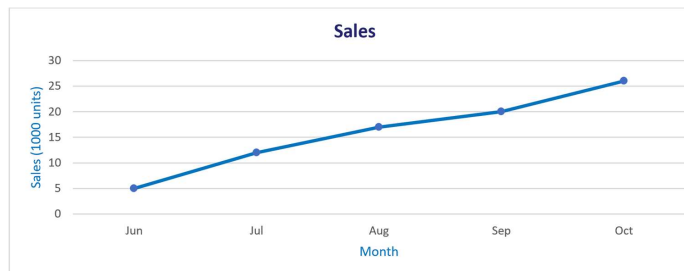
The scale of given for the graph should be even and in proportion to the data. The scale should not be compressed or expanded to exaggerate or diminish change.



Misleading: the range of values shown on the vertical axis is disproportionate to the data. This makes it appear as though the temperature is constant - no variation or overall increasing trend.



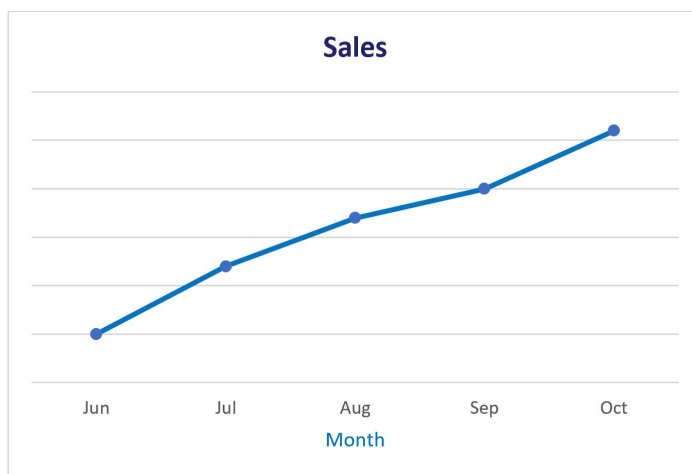
Accurate: the range of values shown on the vertical axis is in proportion to the data. We can now see the data has high variability and an overall increasing trend.



The above three graphs show the same data but by compressing or expanding the horizontal axis it can appear as if the trend is gradual or quite abrupt.

Scale or labels not clearly given

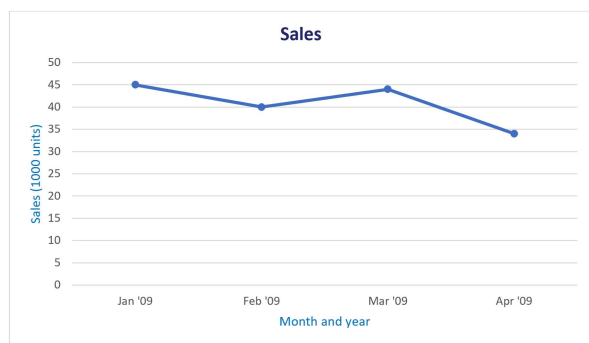
If a scale is omitted or missing units then the reader can not interpret if the trend seen is significant or not.



Without a scale and units on the vertical axis we cannot tell if the sales are in single units sold, 100 units sold, dollars, or so forth, this means we have no means to tell if the difference and upward trend seen is significant. There may also be a truncated axis.

Leaving data out

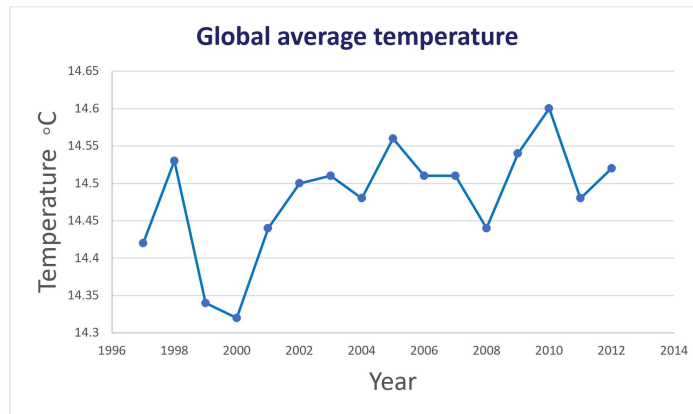
A common way to mislead the audience is to cherry-pick the data shown, so it only includes statistics that supports a particular conclusion. Whenever a broad range of information exists, appearances can be manipulated by highlighting some facts and ignoring others.



This graph only shows a few months data and displays an overall negative trend.



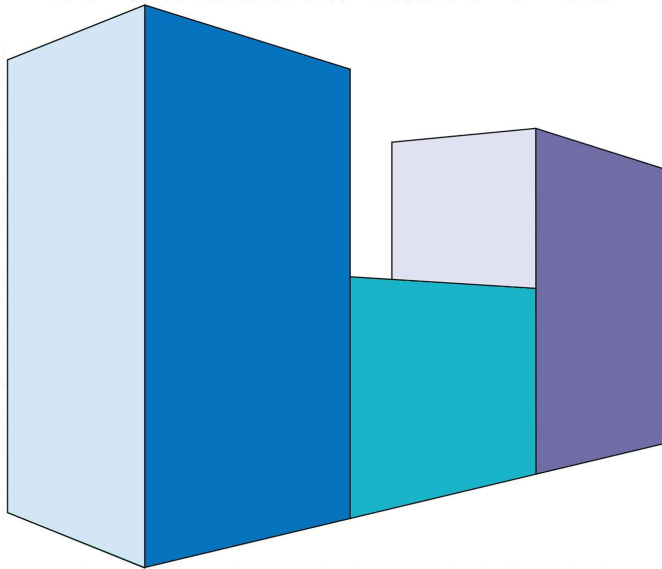
Expanding the data to three years, we see a bit more variability and an overall upward trend.



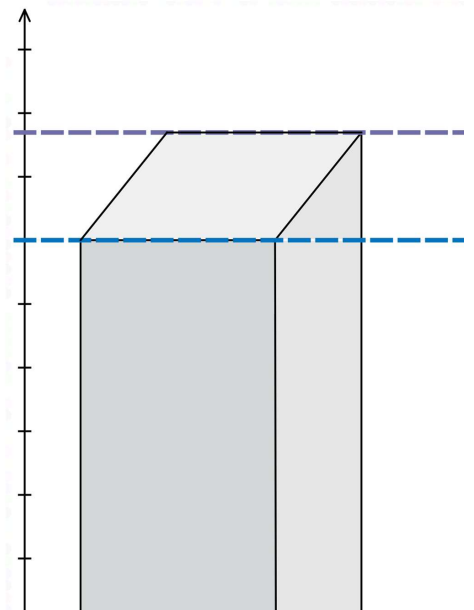
Selecting temperature data from the years 1997 - 2012 we see relatively stable temperature with no obvious upward trend. Compare this to the graph given earlier.

Using pictures or three-dimensional graphics that distort differences

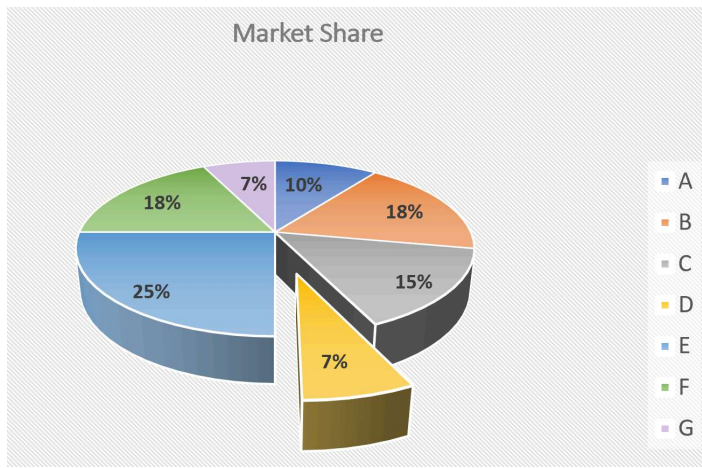
Using graphics in perspective can make it appear as though the section of the graph at the front is larger in comparison to sections further back even when they are the same size. Improperly scaling graphics can also distort the difference in categories significantly.



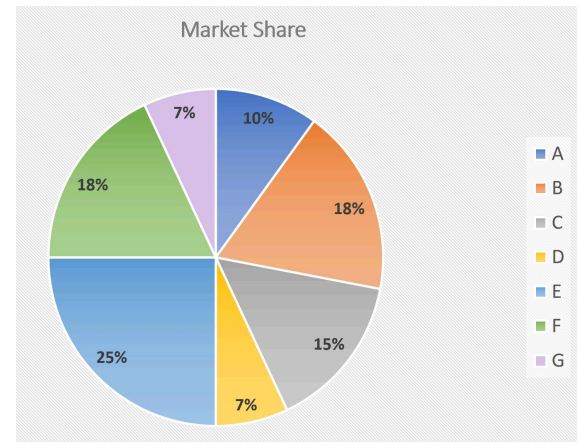
In this column graph the front bar looks far larger than the bar at the back due to perspective. However, both columns represent the same value.



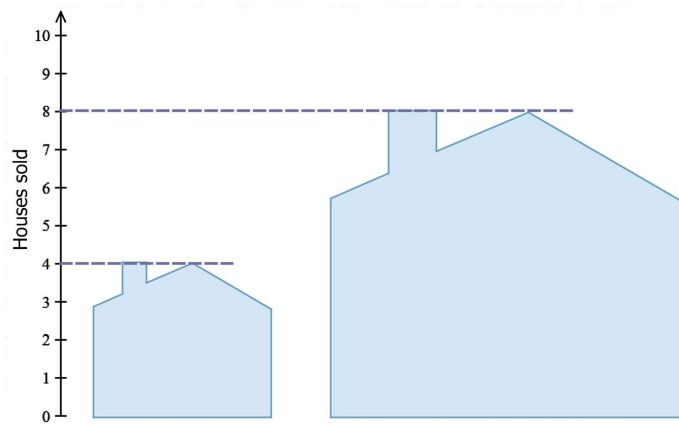
Three dimensional bars can also lead to confusion when reading the scale.



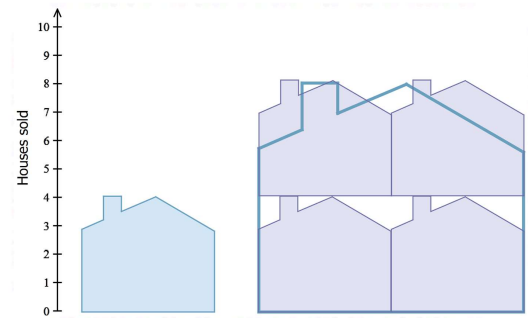
Here again we see a three dimensional graphic in perspective with one piece brought forward from the chart. The perspective makes the front piece of 7% look far larger than the equal valued slice at the rear.



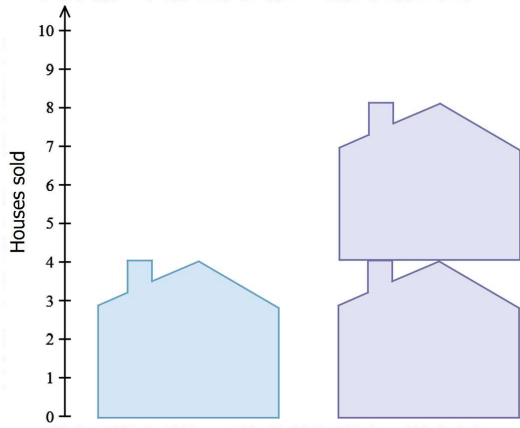
Here is the same data without the perspective and we can now clearly compare the size of the slices. The two equal valued slices now appear the same size.



The vertical axis indicates that twice as many houses were sold. However, by scaling both dimensions the house on the right appears far larger than twice the house on the left.

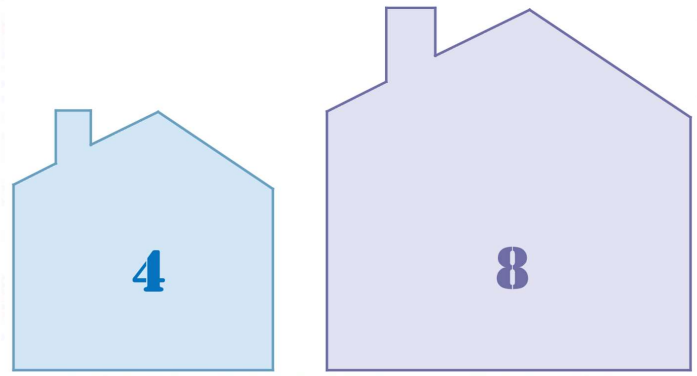


This image shows comparatively the house on the right is in fact four times the area of the house on the left.

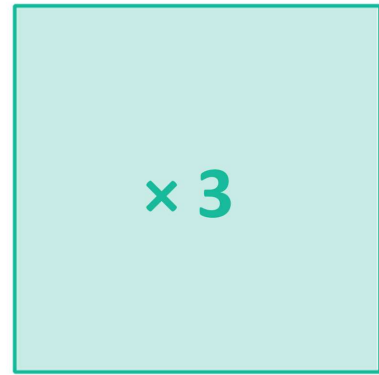
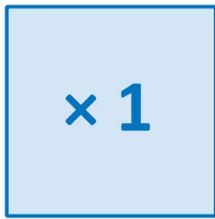


This graph shows the correct use of a pictograph with the icons of equal size. Easy to compare by eye.



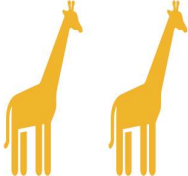


Houses sold



This picture shows the correct scaling if the area of the house represents a given number of sales, then the house on the right has twice the area.








Infographics often use areas of pictures or bubbles to represent values. This is not misleading if correct scaling is used. Above are scaled squares, notice the side lengths of the second and third square are **not** twice and three times the side length of the original square. Mistakenly scaling the side lengths will give an area that is the scale factor squared larger. So if we doubled the side lengths we would get an area four times larger and if we tripled the side lengths we would obtain an area nine times larger.

Animal	Frequency
Alligator	
Elephant	
Giraffe	
Panda	
Tiger	

1 icon = 1 animal

Using different sized icons makes categories appear larger in comparison. It looks as if there are equal numbers of tigers and pandas for instance.

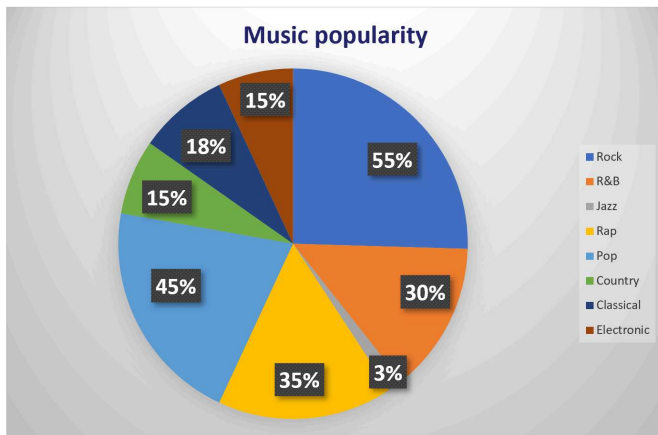
Animal	Frequency
Alligator	
Elephant	
Giraffe	
Panda	
Tiger	

1 icon = 1 animal

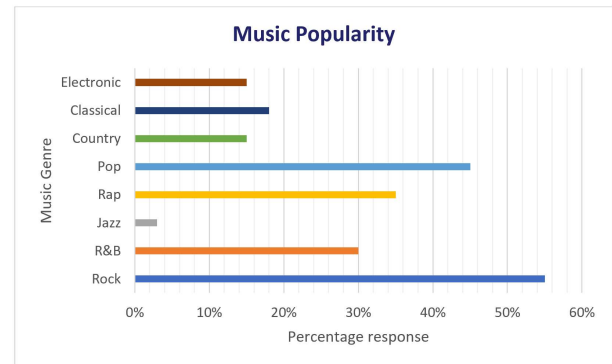
This graph correctly uses the same sized icon for each animal and they are aligned so we can compare numbers at a glance.

Using the wrong graph for a given data type

The type of graph used to visualise the data depends on the type of data you have and the characteristic of the data you wish to highlight. Choosing an inappropriate graph type may lead to the reader to misinterpret the data. We will look in detail at choosing the the most appropriate graphs for different types of data in our next chapter.



Pie charts are used to show the composition of a whole and not comparison across groups. Here we can see the percentages add to more than 100, it is likely that respondents to the survey could select more than one option.



Here the data is represented as a bar chart and we can make clear comparison across all the categories.