

## Visualise data

Data visualisation is the process of representing data graphically in order to identify trends and patterns that would otherwise be unclear or difficult to discern.

Data visualisation serves two purposes: to bring clarity during analysis and to communicate.

The choice of what type of graph or visualisation to use depends greatly on the nature of the variables you have, such as relational, comparative, time-based, etc. Here we have adopted and modified the categorization system used by [ManyEyes](#) (archived link, IBM closed this service in 2015).

That said, sometimes graphing data with an inappropriate visualisation can lead to insights during analysis that would have remained hidden. Experimentation with visualisations during analysis is okay, but when communicating a visualisation, use the graph types listed under the proper methods below. Incorrect visualisation leads to confusion, errors, and abandonment among viewers.

The methods listed here can support both purposes of analysis and communication. You may want to graph data during analysis to see, for example, spikes in website traffic related to your social media campaigns. Visualisation, in this instance, eases data analysis. When communicating that data, however, the visualisation may need to be simplified and key areas may need emphasis in order to call the attention of readers and stakeholders. See the discussion under Report and Support Use for more information about how you may want to repackage a data visualisation for communication purposes.

Each main method below contains several visualisation possibilities. Click on each to see examples and read advice on using and choosing that visualisation method.

This graphic by Andrew Abela from [Extreme Presentations](#) provides a good representation of different types charts that can be used to visualise data.

Diagram showing four categories of charts to choose from

(c) 2006 A. Abela, used with permission. [www.ExtremePresentation.com](http://www.ExtremePresentation.com). [View this chart as a pdf](#).

## Methods

### See relationships among data points

- [Scatterplot](#)

A Scatterplot is used to display the relationship between two quantitative variables plotted along two axes. A series of dots represent the position of observations from the data set.

- [Matrix chart](#)

A matrix chart shows relationships between two or more variables in a data set in grid format.

- [Network diagram](#)

A network diagram uses a set of nodes and connecting lines to display of how people (or other elements) in a network are connected.

It is usually a product of social network analysis.

## Compare a set of values

- [Bar chart](#)

A bar chart plots the number of times a particular value or category occurs in a data set, with the length of the bar representing the number of observations with that score or in that category.

- [Block histogram](#)

A histogram is a graphical way of presenting a frequency distribution of quantitative data organised into a number equally spaced intervals or bins (e.g. 1-10, 11-20...).

- [Bubble chart](#)

Commonly used on maps, and x/y-axis plots, or no plot at all, bubble charts communicate the raw count, frequency, or proportion of some variable where the size of the bubble reflects the quantity.

- [Bullet graph](#)

Bullet graphs encode a single variable as a bar.

- [Deviation bar graph](#)

Deviation bar graphs are simply two bar charts aligned, where one of the charts runs right to left rather than left to right.

- [Dot plot](#)

Dot plots encode single data points with circles, often on a line.

- [Small multiples](#)

Small multiples are an array of graphs on the same scale that are grouped together in a row or grid and are often used to simplify a data display.

## Changes over time

- [Line graph](#)

A line graph is commonly used to display change over time as a series of data points connected by straight line segments on two axes.

- [Slopegraph](#)

A slopegraph is a lot like a line graph, in that it plots change between points however, a slopegraph plots the change between only two points, without any kind of regard for the points in between.

- [Split axis bar graph](#)

While many graph types geared toward comparisons ask the viewer to subtract the difference between the heights of two bars or the space between two points on a line, a deviation bar graph simply graphs the difference.

- [Stacked graph](#)

Stacked graphs depict items stacked one on top (column) of the other or side-by-side (bar), differentiated by coloured bars or strips.

## See the parts of a whole

- [Icon array](#)

An icon array is a display in which one shape is repeated a specific number of times (usually 10, 100 or 1,000) and then some of the shapes are altered in some way (usually by colour) to represent a proportion.

- [Pie chart](#)

A pie chart is a divided circle, in which each slice of the pie represents a part of the whole.

The categories that each slice represents are mutually exclusive and exhaustive. Data with negative values cannot be displayed as a pie chart.

- [Treemap](#)

A treemap displays hierarchical relationships through a set of rectangles, sized proportionately to each data point, clustered together into one large rectangle.

## Analyse a text

- [Phrase net](#)

Phrasenets are useful for exploring how words are linked in a text and, like word clouds and word trees, can be informative for early data analysis.

- [Word cloud](#)

Word clouds or tag clouds are graphical representations of word frequency that give greater prominence to words that appear more frequently in a source text.

- [Word tree](#)

Word trees use a visual branching structure to show how a pre-selected word(s) is connected to other words.

## See the world

- [Demographic mapping](#)

Demographic mapping is a way of using GIS (global information system) mapping technology to show data on population characteristics by region or geographic area.

- [Geo-tagging](#)

Geo-tagging is the process of adding geographic information about digital content, within “metadata” tags - including latitude and longitude coordinates, place names and/or other positional data.

- [GIS mapping](#)

GIS mapping will typically display one data variable or indicator, often using colour coding to indicate the density, frequency, or percentage in a given region, allowing quick comparison between regions.

- [Interactive mapping](#)

Interactive mapping involves using maps that allow zooming in and out, panning around, identifying specific features, querying underlying data such as by topic or a specific indicator (e.g., socioeconomic status), generating reports and other means of u