Chart type

Advice on which chart types are most appropriate to visualise the major relationships seen in data.

Table of contents

There are eight common relationships that charts display. Prioritise what you want to highlight in the data and choose the chart type accordingly.

The eight common relationships within data are the following:

- comparisons of magnitude (size)
- time series
- ranking
- part-to-whole
- deviation
- distribution
- correlation
- spatial (maps will be covered separately in phase 4)

Consider the message you want to communicate and choose your data accordingly. Your message might be better conveyed by deriving variables.

Start data that are likely to be compared from the same point on a chart – a shared horizon. Use a clustered chart to

compare values; only the first category is easily comparable in stacked bar charts.

EXAMPLE

Do

Adoptions by sex
England and Wales, 1998 to 2012

EXAMPLE

Don't

Adoptions by sex
England and Wales, 1998 to 2012

To show:

- X is bigger than Y
- A is almost twice the size of B

Comparisons of size are shown most effectively as horizontal or vertical bars. Always begin the y-axis at zero.

Small differences in magnitude, starting the y-axis at a non-zero value

If there are small differences between values sometimes it is necessary to start the y-axis at a non-zero value.

Always put a break in the y-axis if you don’t start at zero.

EXAMPLE

Use a dot (or other symbol) plot to make comparisons between values. The size of the visual element representing the data (dot position) is representative of the data value itself.

EXAMPLE

Do

You can also show small differences between data by adjusting the deviation. This is changing what data can be seen from a chosen value (the deviation section has more information).

EXAMPLE

Rather than over-emphasising month-to-month or point-to-point comparisons of estimates a time series can show:

- change
- trend
- fluctuation
- growth
- decline
- increase
- decrease

Time series axes

Time should always run from left to right along the horizontal axis.

Time series charts

A time series with regular intervals can be presented using line charts, bar charts or a combination of both.

Bars should be used to emphasize individual values at distinct points in time. Use them when data points are at equal intervals.

EXAMPLE

A line chart will emphasise the overall pattern of the data and highlight trends. Use them when you have lots of data points or just a few. Multiple times series should always use line charts.

EXAMPLE

Use a dot plot with a line when there are lots of data points or the interval between data points is not equal. Show if data are irregular.

EXAMPLE

Do

EXAMPLE

Don't

Multiple time series shouldn't be presented using bar charts. Use a line chart to make sure the trends in the series are clear. Use points on a line to highlight individual data points, to read specific values or highlight when the data were sampled.

EXAMPLE

Do

EXAMPLE

Don't

Time series charts don't have to begin at zero, if a chart doesn't start at zero this must be indicated by breaking the y-axis in an obvious way.

A chart can tell a very different story depending on the scale of the axes.

This chart gives the impression that the measles, mumps and rubella (MMR) vaccination level has remained high and fairly stable.

EXAMPLE

An overview

MMR vaccination uptake at age 1
UK, 1992 to 2012

When the y-axis is altered a different picture emerges showing that the measles, mumps and rubella vaccination has dropped considerably since 1997.

EXAMPLE

A more focused view

MMR and Diphtheria vaccination uptake at age 1
UK, 1992 to 2012

You can use two charts with different axes scales to ensure that the data are represented without bias whilst highlighting the important message.

To show:

- greater than
- less than
- equal to
- from lowest to highest

Ranking Charts

Use bar charts to show data that are ranked, in either ascending or descending order. Horizontal bars should be used.

A bar chart should always be ranked by value, unless there is a natural order to the data (for example, age or time).

EXAMPLE

Don't

To highlight the highest values the largest value should be at the top of the chart.

EXAMPLE

Do

Descending, largest values highlighted

To highlight the lowest values the smallest value should be at the top of the chart.

EXAMPLE

Do

Ascending, smallest values highlighted

If you are talking about data in terms of first, second or third or “the top 10” they should always be in descending order.

EXAMPLE

Do

Top 10 girls’ baby names
England, 2013

EXAMPLE

Don't

Top 10 girls’ baby names
England, 2013

Plotting a change in rank

Use a slope chart to highlight a change in rank.

EXAMPLE

Ranking multiple series

Rank the most important or recent data if there are multiple series and the other data sets should be ordered correspondingly.

To show:

- ratio
- percentage
- proportion
- share
- breakdown
- make up
- hierarchy

Bar charts and pie charts should be used to show part to whole relationships.

Pie charts should only be used when there are less than six categories, otherwise use a bar chart or, if appropriate, combine categories.

Rank the categories in a pie chart and start the first segment at the 12 o'clock position.

Segments of a pie chart must sum to 100%. If the categories do not sum to a meaningful whole, don't use a pie chart. Where appropriate categories can be combined to highlight a certain message but should never be removed.

EXAMPLE

Do

All main categories included

Religion categories combined

EXAMPLE

Don't

No religion and not stated categories removed

If no categories are dominant use a bar chart to illustrate your data.

EXAMPLE

Do

EXAMPLE

Don't

Multiple part to whole

Use bars to enable comparisons to be made across multiple part to whole charts.

EXAMPLE

To enable comparisons within sub-categories

To enable comparisons across sub-categories

To show:

- number of times more than the average
- the difference from

Use a bar chart to plot deviation from a fixed value, or series of values.

Deviation where the value of data is most important

EXAMPLE

GDHI per head (£)
England, 2011

Deviation where the amount of change is most important

EXAMPLE

GDHI per head index comparison with England average
England, 2011

Deviation where the amount of change is most important

Use small multiples to plot deviation for multiple series. The axes should be identical for each small multiple.

EXAMPLE

To show:

- frequency
- distribution
- profile
- range
- concentration

- normal curve
- population pyramid
- shape

For one variable

Use a histogram to show a distribution of data. Use small gaps between the bars to emphasise the profile of the data.

EXAMPLE

Usually resident population aged 0 to 21
UK, 2013

For two variables

Use a population pyramid to show the distribution of comparable data sets and highlight differences in the profile of the data.

EXAMPLE

For more than two variables

To compare four variables population pyramids can be overlaid, with the least important data set displayed using an outline pyramid instead of bars.

EXAMPLE

Small multiple charts can also be used for multiple distributions. Use the same scale to enable comparison across charts..

EXAMPLE

Box-plots can also be used to compare distributions with two or more variables.

EXAMPLE

Correlation charts are often associated with causality and they should be used with caution.

Correlation can show:

- increases with
- relates to
- changes with
- varies with
- caused by

Anscombe's Quartet

Anscombe's quartet is a powerful illustration of the drawback of relying solely on basic descriptive statistics to summarise data. The data in all four of the graphs in the quartet are virtually identical when using standard descriptive methods. Looking at your data before analysing it is something that Anscombe was passionate about.

EXAMPLE