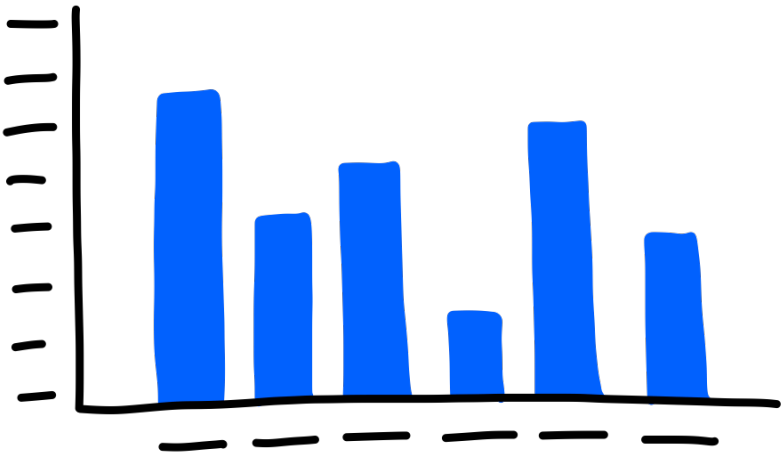
Bar Graph



Should Be Used For

Data which represent discrete things on a continuous dimension, which it would make sense to group into bins along that dimension

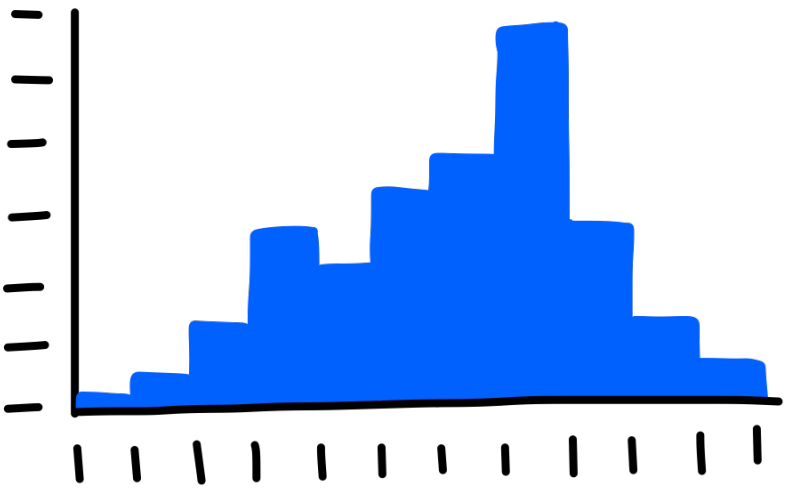
Examples

- Money made in a period
- Durations of events, binned by duration segment

Pitfalls

- Be careful that your data is actually in bins, not continuous
- Make sure the vertical axis starts at 0, or your are being misleading
- Have a bar chart for when adding up amounts of a thing in a bin, use a histogram for adding up the number of things in a bin

Histogram



Should Be Used For

Data which represent discrete things on a continuous dimension, which it would make sense to group into bins along that dimension.

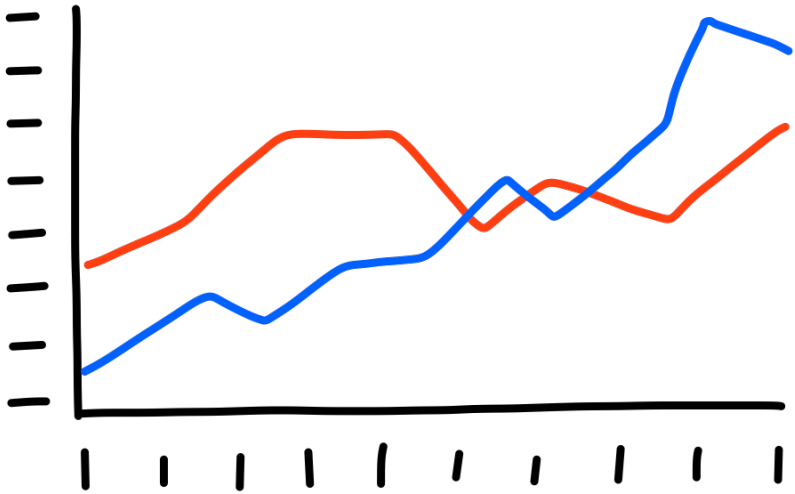
Examples

- Events on a timeline, binned by time segment
- Accuracy of shots, binned by distance segment

Pitfalls

- Setting improper bin sizes
- Setting an improper bin offset
- Have a bar chart for when adding up amounts of a thing in a bin, use a histogram for adding up the number of things in a bin

Line Graph



Should Be Used For

Data which has a single value for each value over a continuous dimension, where it makes sense to interpolate between data points.

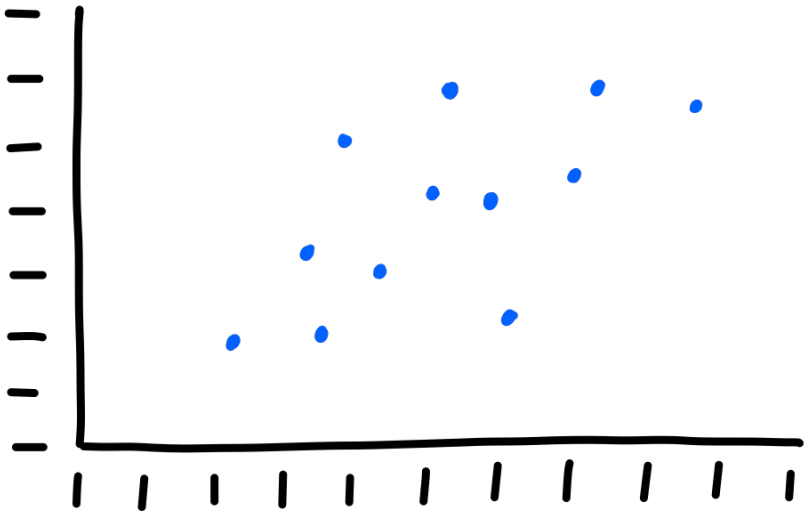
Examples

- Values over time
- Mathematical function plot

Pitfalls

- Displaying data where it doesn't make sense to interpolate
- Be careful which type of interpolation you use
- There may only be one value for each value on the continuous dimension

Scatter Plot



Should Be Used For

Data that is to be compared along 2 dimensions.

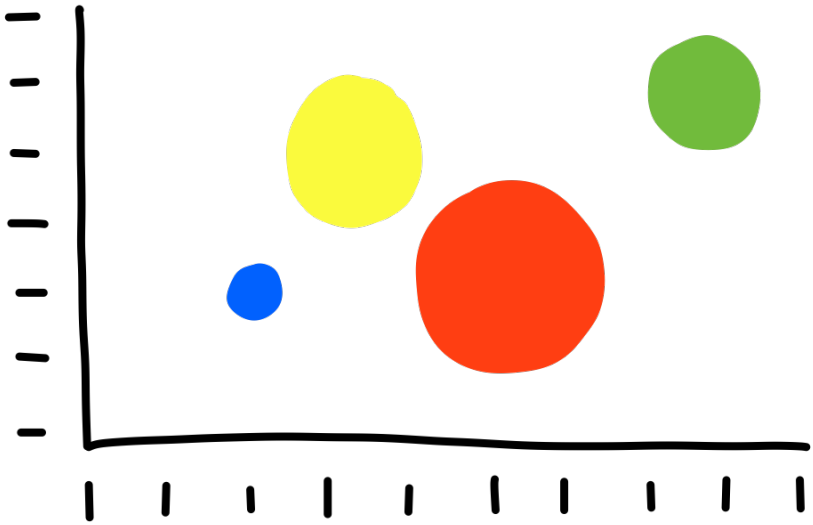
Examples

- Petal length vs. petal width of a sample of flowers

Pitfalls

- If the variables have any other known rules (like with time or percentage) then consider using a graph that better fits those rules

Bubble Chart



Should Be Used For

Correlation between two variables, when relative size is another important factor

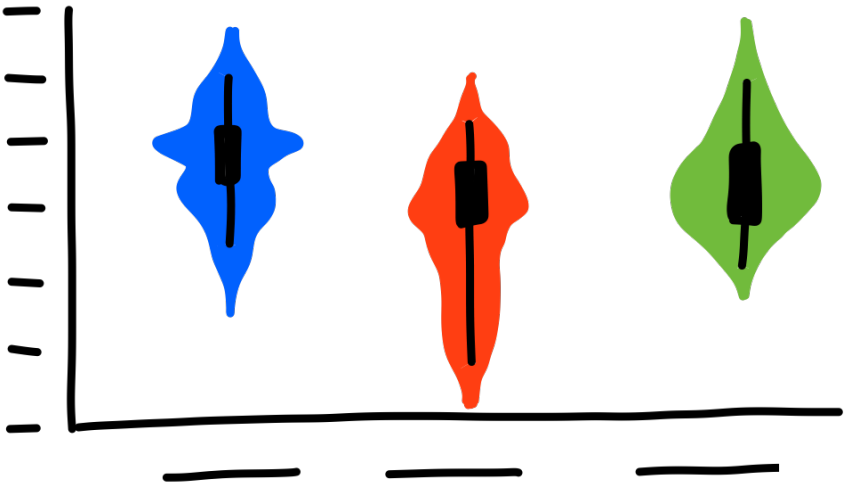
Examples

- Countries compared by GDP, average life expectancy, and populations

Pitfalls

- not making appropriate use of the size or color
- sizing the bubbles improperly

Violin Plot



Should Be Used For

Don’t use this. Use a histogram or a box/candlestick plot instead.

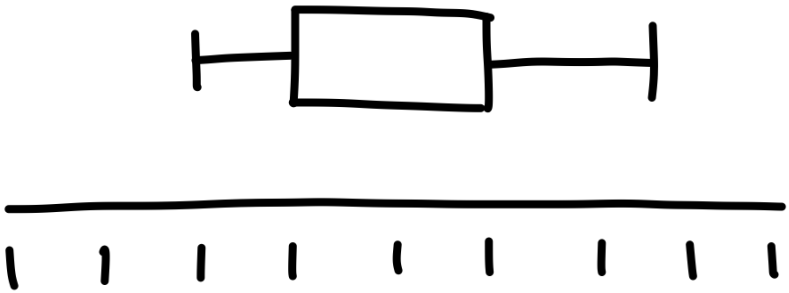
Examples

- No dataset is best shown by this plot

Pitfalls

- Might be confusing for those who haven’t seen it before
- Either the dataset has an average or the distribution matters
- Doesn’t show units, so you can’t get data from it
- If you use different smoothing values, the data can look completely different

Box Plot



Should Be Used For

Analysis of a distribution over a single dimension

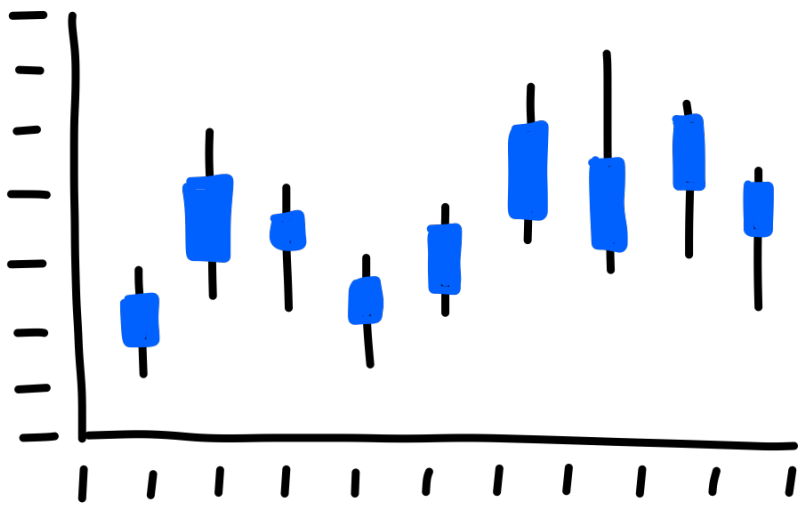
Examples

- life expectancy distribution of a country

Pitfalls

- Showing anything other type of information than the quartiles of a distribution

Candlestick Chart



Should Be Used For

Data that has max, min, and end data over a continuous dimension.

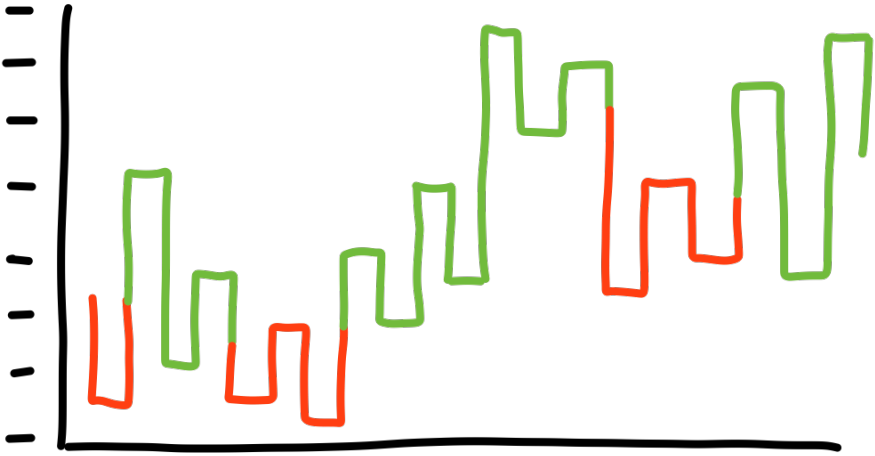
Examples

- Stock prices

Pitfalls

- If not used for stocks, people might be confused

Kagi Chart



Should Be Used For

Getting the trend of an Up/Down motion in a variable without the noise.

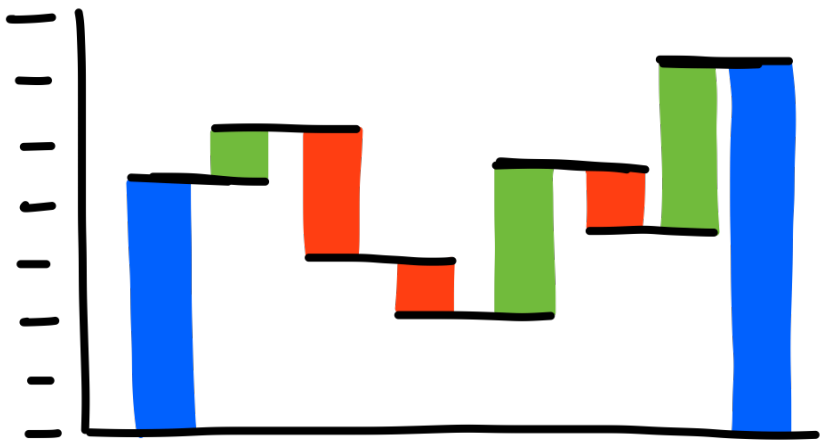
Examples

- Stock prices

Pitfalls

- It has so far been used exclusively in Japan showing stock trends, therefore it is not universally understood
- Does not show detail of up/down movement, just the trend and major shifts
- Does not have a horizontal axis - one step is one major shift

Waterfall Chart



Should Be Used For

Show how an initial value is affected up and down by various factors, or over time.

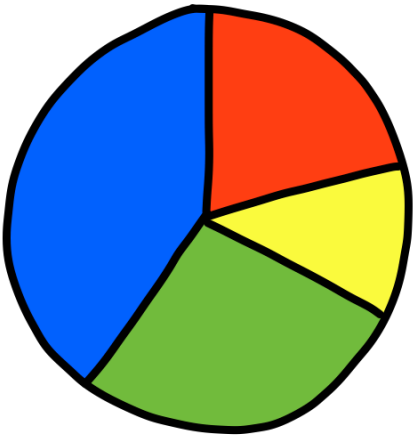
Examples

- Company balance, effects by segment over a time period

Pitfalls

- There is an emphasis on the change, if you need more accuracy (like something on an actual horizontal axis) then use another chart

Pie Chart



Should Be Used For

Relative quantities of categories of a set

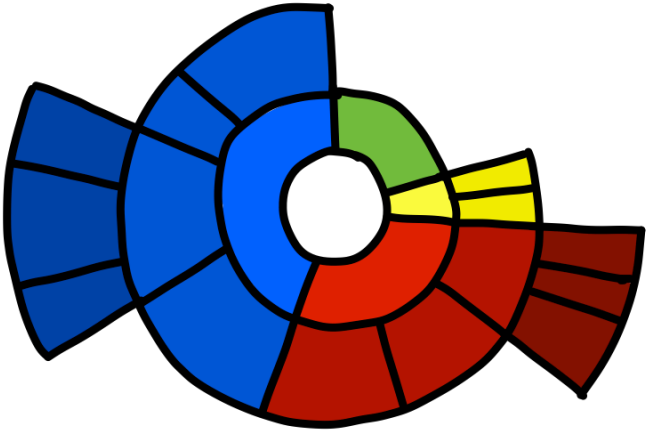
Examples

- Percentage of revenue by business segment
- Multiple choice form responses

Pitfalls

- Try not to print digits along with the pie
- Don’t Represent something that can have more than 100%
- Showing divisions that are smaller than can be seen (< 1% usually)

Sunburst Chart



Should Be Used For

Relative categories of a whole, including nested categories

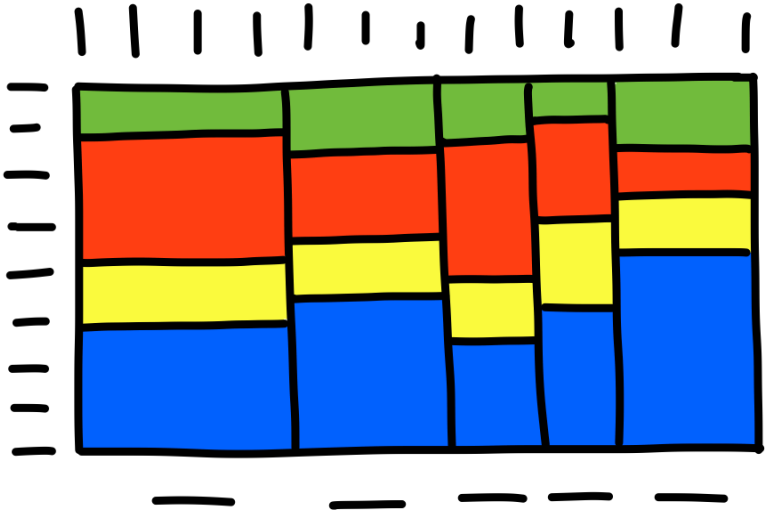
Examples

- Business Division revenues, broken out by subdivisions

Pitfalls

- If you want to compare the divisions, use a Mosaic chart

Marimekko / Mosaic Chart



Should Be Used For

Comparison of relative percentages between different sized bins

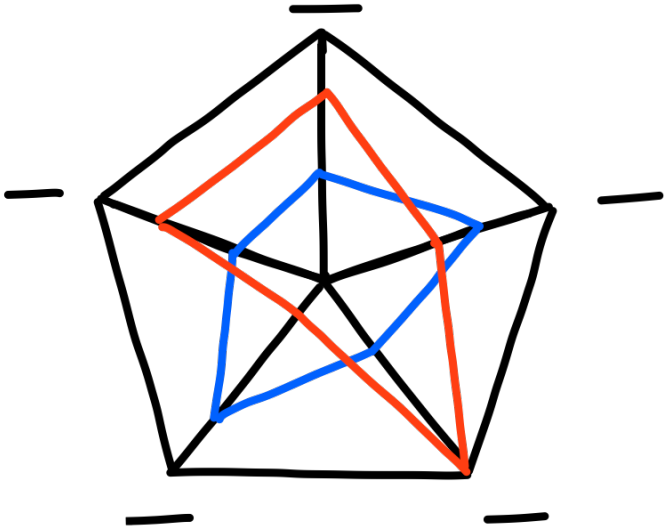
Examples

- Revenue of company by market segment

Pitfalls

- Be sure the bins have the same pieces, or at least the piece you want to compare amongst all bins
- Be sure to Sort percentage by the same order across all bins, or by a consistent ordering scheme
- If the bins are purely qualitative (they don't have a comparable value themselves), use stacked bar charts instead

Radar / Spider / Star Chart



Should Be Used For

Showing relative coverage of a thing over multiple different categories by scores in each category.

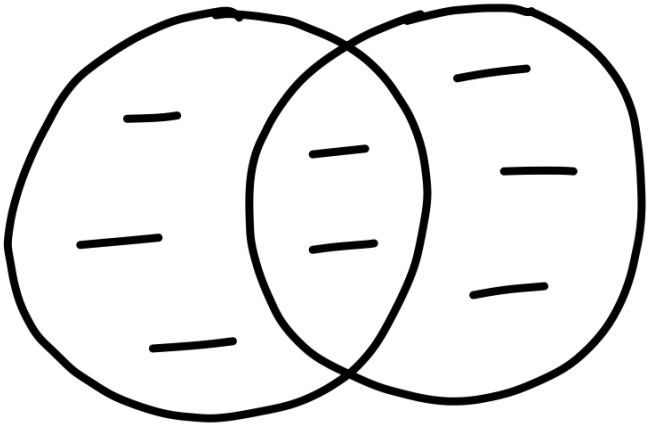
Examples

- Score/Grade in multiple disciplines of a person or company

Pitfalls

- If getting the general sense of coverage is not the focus, but the focus is on the variables themselves, try a stacked bar chart
- Too many categories (keep to ≤ 8)
- Too many things compared (keep to ≤ 3) - usually just show one

Venn Diagram



Should Be Used For

When highlighting the differences between two categories or specifically showing that something has properties of two categories.

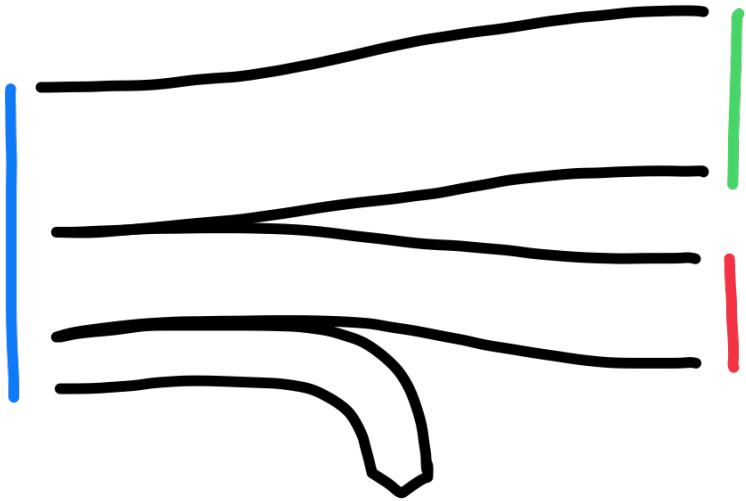
Examples

- numbers belonging to sets

Pitfalls

- If anything else apart from the existence of examples of something in both sets is the focus, compare it with respect to that focus
- Don’t show more than 3 categories

Sankey Diagram



Should Be Used For

When showing amount of flow between states.

Examples

- Energy through a system
- Money through a company
- People through a process

Pitfalls

- Be careful of quantities that are created or destroyed, you'll need to mark those explicitly
- If it's flow between categories, consider a Chord diagram

Chord Diagram

Should Be Used For

Diagramming the relationships amongst segments of a whole

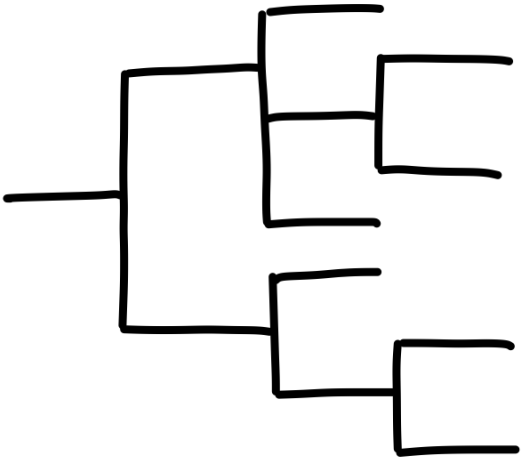
Examples

- Showing connections of characters

Pitfalls

- when no relationship, use pie chart

Tree Diagram



Should Be Used For

Showing Hierarchical Relationships

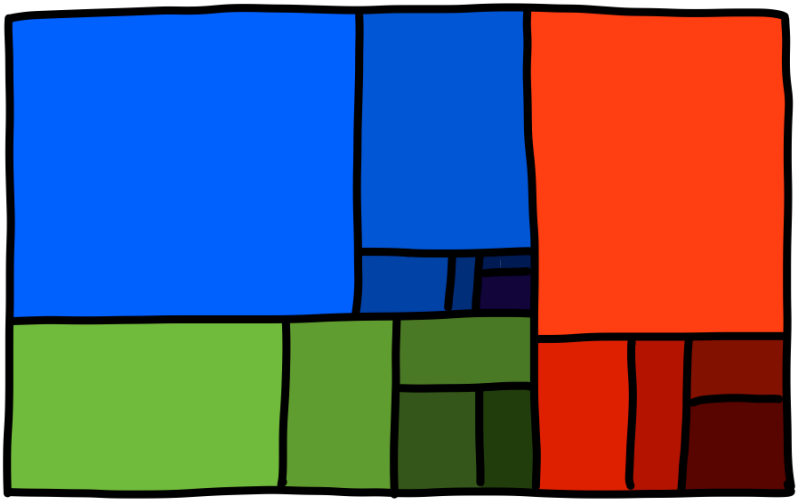
Examples

- Family Trees

Pitfalls

- Be careful with the way you draw the tree if some leaf nodes appear under multiple branch nodes (like with incestuous royal family trees)

Treemap



Should Be Used For

Showing the relative size of nodes in a tree

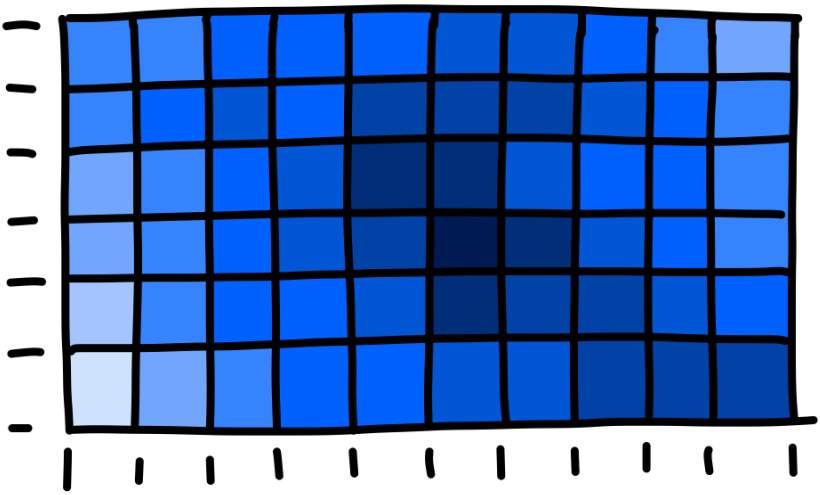
Examples

- File system file & folder sizes
- Biomass of various types of living organisms

Pitfalls

- Do not show data that is not hierarchical or not in a tree format
- Only use if you want to show the relative sizes of branch nodes too, not just the leaf nodes, otherwise choose a bar chart

Heat Map



Should Be Used For

Values over 2 qualitative (or binned) dimensions

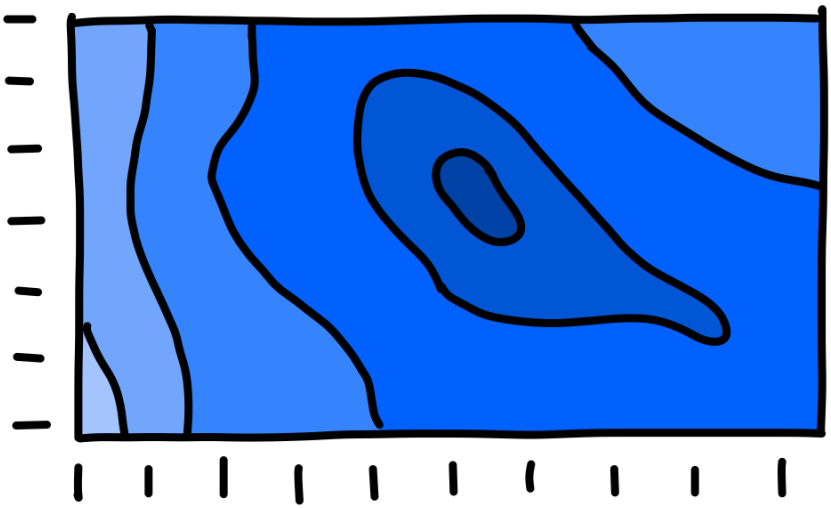
Examples

- Winning rates of different chess engines against each other
- Average temperature for each month over years

Pitfalls

- If putting linear data in a matrix, make sure the wrapping point makes sense
- Usually, the two axes should have the same dimension (like distance), if not, be sure to make it obvious
- The emphasis should be in the patterns along the 2 dimensions

Contour Plot



Should Be Used For

Showing regions of similar values along 2 continuous axes

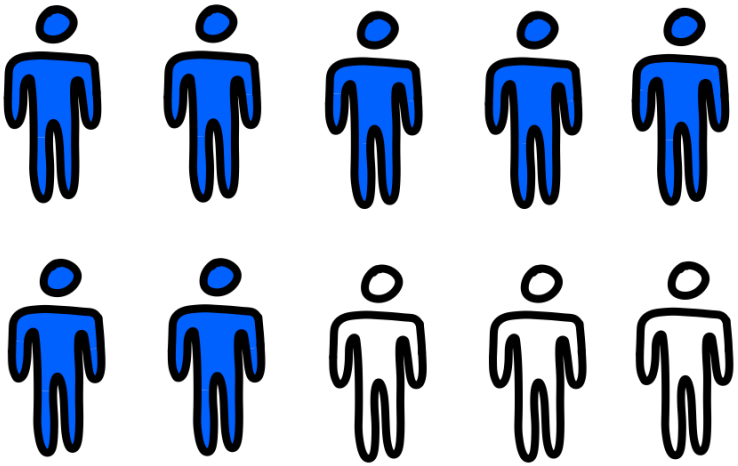
Examples

- Topographical maps
- Temperature maps

Pitfalls

- Make sure color or shading aligns with region order

Pictograph



Should Be Used For

Showing an amount or percentage in an intuitive way

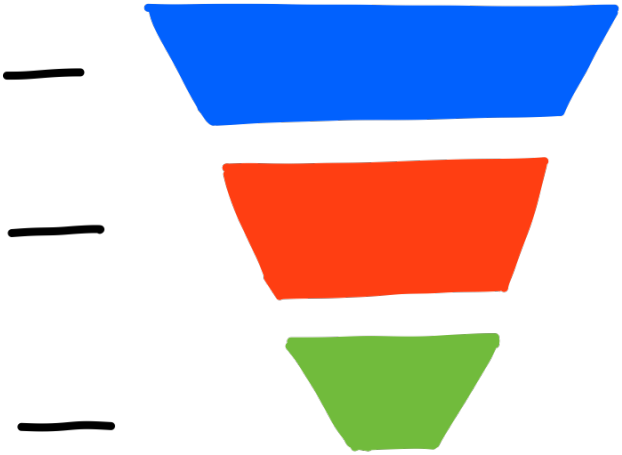
Examples

- Percentage of population with a certain trait
- Ratings

Pitfalls

- Must have an emphasis that these are tangible things
- Should be able to count them (or if a percentage, be able to say “five of twenty peope…” or something)
- Worrying too much about precision lessens the impact, feel free to round

Funnel Chart



Should Be Used For

Showing a progressive reduction of something

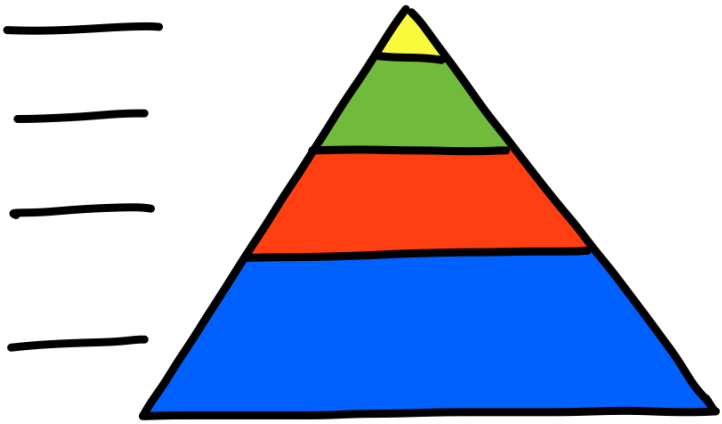
Examples

- Ad viewership to customer conversion quantities

Pitfalls

- Not focusing on reduction

Pyramid Charts



Should Be Used For

Showing the proportions of strictly hierarchical amounts

Examples

- Amount of people at various levels of an organization

Pitfalls

- Don’t use to describe flow, use a funnel chart instead
- Don’t use to show the hierarchical nature of a tree structure, use a tree diagram instead
- Don’t use to show proportion of non-hierarchical data, use a pie chart instead

Timeline

Should Be Used For

Showing events relative to each other across time

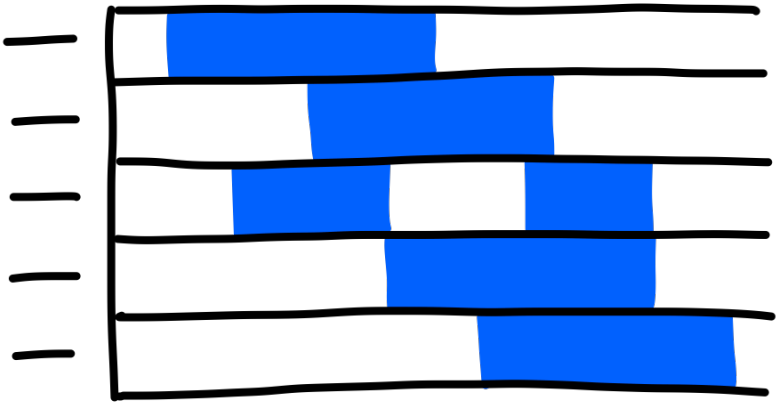
Examples

- Major historical events

Pitfalls

- If you don’t care as much about showing the time between events accurately, be sure to note that
- If using multiple categories, or causally related events, use a gantt chart instead

Gantt Chart



Should Be Used For

Time Data With Events & Durations of different independent categories

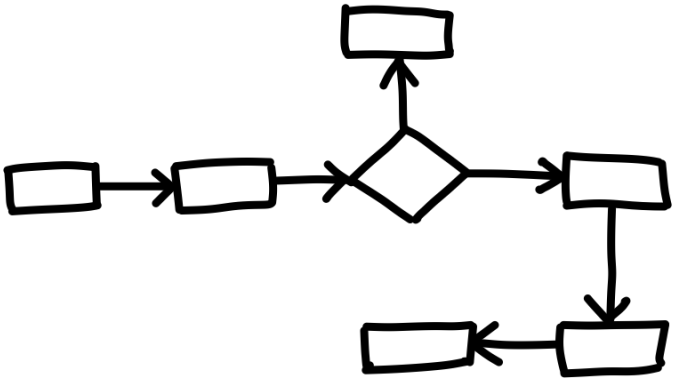
Examples

- Project management

Pitfalls

- Don’t use it to show data with relative sizes on the vertical axis, this makes it difficult to read, instead use another plot to show that comparison information

Flow Chart



Should Be Used For

Showing the steps in a process

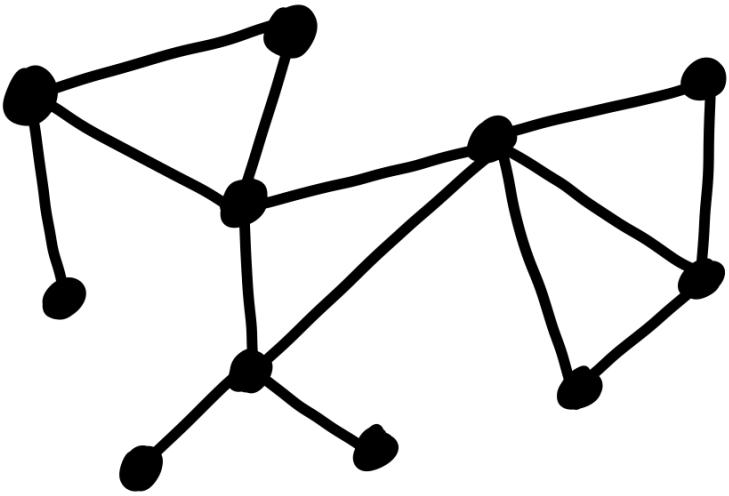
Examples

- Information flow in a program
- Product flow in a value stream map

Pitfalls

- Only shows the discrete events, not any metadata about them
- Don’t annotate too much with metadata, it gets messy and hard to read
- Sometimes there are many small tasks that are less important and a few tasks which are more important. This chart treats them equally, so find a way to highlight the important ones, or encapsulate the less important ones

Network Graph



Should Be Used For

Showing relationships amongst items in a set

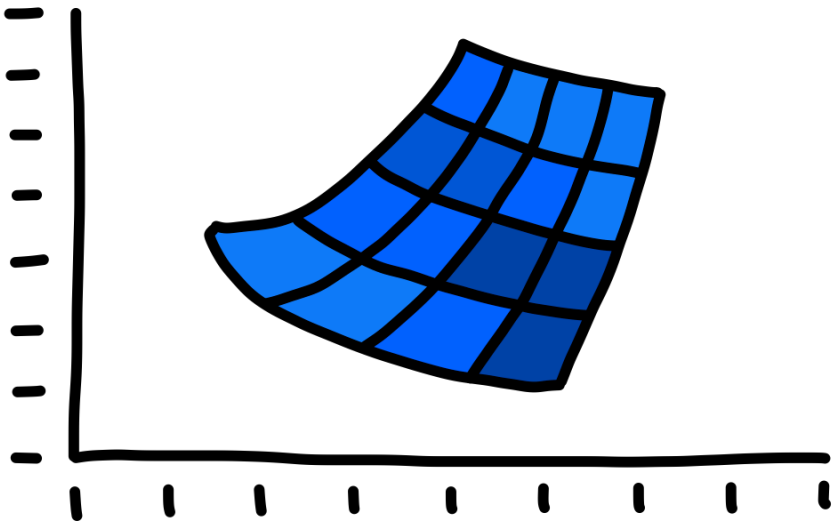
Examples

- Social Network friends
- Neurons
- Locations and distances

Pitfalls

- The items should be independent of each other apart from the connection you're drawing. If not, use a flow chart
- There can only exist one connection between each element (or two if it's a directed graph)

Carpet Plot



Should Be Used For

Showing outcomes of varying 2 independent axes at a time

Examples

- Evaluating multiple options within a range
- Figuring out where to put boundaries across multiple variables

Pitfalls

- The output values you end up with can be overlapping, then split into 2 carpet plots