|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Guideline** | **Rating** | | | |
| **Category** | **Description** | **2** | **1** | **0** | **n/a** |
| **Text –** minimize text | **Short descriptive title is left-justified in upper left corner**  Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. |  |  |  |  |
|  | **Text size is hierarchical and readable**  Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen. |  |  |  |  |
|  | **Text is horizontal**  Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points. Consider switching graph orientation (e.g., from column to bar chart) to make text horizontal. |  |  |  |  |
|  | **Data are labeled directly**  Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars and next to lines). Eliminate/embed legends when possible because eye movement back and forth between the legend and the data can interrupt the brain’s attempts to interpret the graph. |  |  |  |  |
|  | **Labels are used sparingly**  Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis. Do not add numeric labels \*and\* use a y-axis scale, since this is redundant. |  |  |  |  |
| **Arrangement –** thoughtful arrangement of visual makes it easier for audience to interpret | **Proportions are accurate**  A viewer should be able measure the length or area of the graph with a ruler and find that it matches the relationship in the underlying data. Y-axis scales should be appropriate. Bar charts start axes at 0. Other graphs can have a minimum and maximum scale that reflects what should be an accurate interpretation of the data (e.g., the stock market ticker should not start at 0 or we won’t see a meaningful pattern). |  |  |  |  |
|  | **Data are intentionally ordered**  Data should be displayed in an order that makes logical sense to the viewer. Data may be ordered by frequency counts (e.g., from greatest to least for nominal categories), by groupings or bins (e.g., histograms), by time period (e.g., line charts), alphabetically, etc. Use an order that supports interpretation of the data. |  |  |  |  |
|  | **Graph is two-dimensional**  Avoid three-dimensional displays, bevels, and other distortions. |  |  |  |  |
| **Color –** thoughtful use of colors | **Color scheme is intentional**  Colors should represent brand or other intentional choice, not default color schemes. Use your organization’s colors or your client’s colors. Work with online tools to identify brand colors and others that are compatible. |  |  |  |  |
|  | **Color is used to highlight key patterns**  Action colors should guide the viewer to key parts of the display. Less important, supporting, or comparison data should be a muted color, like gray. |  |  |  |  |
|  | **Color is legible when printed in black and white**  When printed or photocopied in black and white, the viewer should still be able to see patterns in the data. |  |  |  |  |
|  | **Color is legible for people with colorblindness**  Avoid red-green and yellow-blue combinations when those colors touch one another. Avoid using red to mean bad and green to mean good in the same chart. |  |  |  |  |
|  | **Text sufficiently contrasts background**  Black/very dark text against a white/transparent background is easiest to read. |  |  |  |  |
| **Lines –** reduce chart junk | **Excessive lines— gridlines, borders, tick marks, and axes**  can add clutter or noise to a graph, so eliminate them whenever they aren’t useful for interpreting the data. |  |  |  |  |
|  | **Gridlines, if present, are muted**  Color should be faint gray, not black. Full points if no gridlines are used. Gridlines, even muted, should not be used when the graph includes numeric labels on each data point. |  |  |  |  |
|  | **Graph does not have border line**  Graph should bleed into the surrounding page or slide rather than being contained by a border. |  |  |  |  |
|  | **Axes do not have unnecessary tick marks or axis lines**  Tick marks can be useful in line graphs (to demarcate each point in time along the y-axis) but are unnecessary in most other graph types. Remove axes lines whenever possible. |  |  |  |  |
|  | **Graph has one horizontal and one vertical axis**  Viewers can best interpret one x- and one y-axis. Don’t add a second y-axis. Try a connected scatter plot or two graphs, side by side, instead. (A secondary axis used to hack new graph types is ok, so long as viewers aren’t being asked to interpret a second y-axis.) |  |  |  |  |
| **Overall –** only include visuals that are necessary | **Graph highlights significant finding or conclusion**  Graphs should have a "so what?" – either a practical or statistical significance (or both) to warrant their presence. For example, contextualized or comparison data help the viewer understand the significance of the data and give the graph more interpretive power. |  |  |  |  |
|  | **The type of graph is appropriate for data**  Data are displayed using a graph type appropriate for the relationship within the data. For example, change over time is displayed as a line graph, area chart, slope graph, or dot plot. |  |  |  |  |
|  | **Graph has appropriate level of precision**  Use a level of precision that meets your audiences’ needs. Few numeric labels need decimal places, unless you are speaking with academic peers. Charts intended for public consumption rarely need p values listed. |  |  |  |  |
|  | **Individual chart elements work together to reinforce the overarching takeaway message**  Choices about graph type, text, arrangement, color, and lines should reinforce the same takeaway message. |  |  |  |  |
| **OVERALL SCORE** |  |  |  |  |  |

**Data Visualization Checklist**

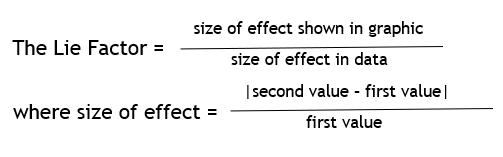
Adapted my checklist from <https://stephanieevergreen.com/updated-data-visualization-checklist/>

I found Tufte’s Principles to be very helpful throughout the course and plan to utilize them in the future. I’ve included some highlights from the course lectures as well as Tufte’s website for future reference.

Principles include:

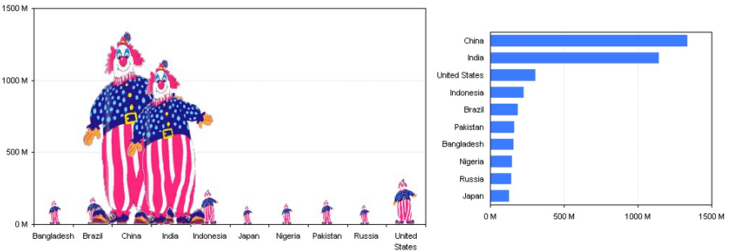
* Tell the truth – Graphical integrity principles
* Do it effectively with clarity and precision – Design principles

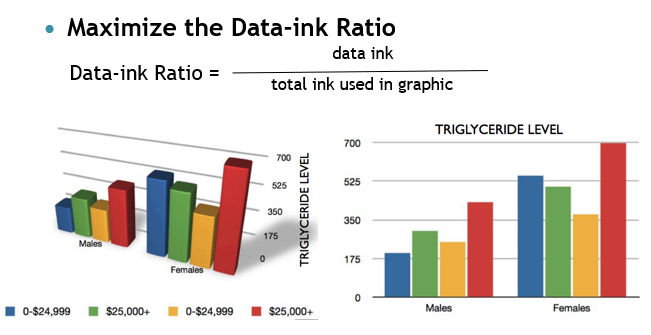
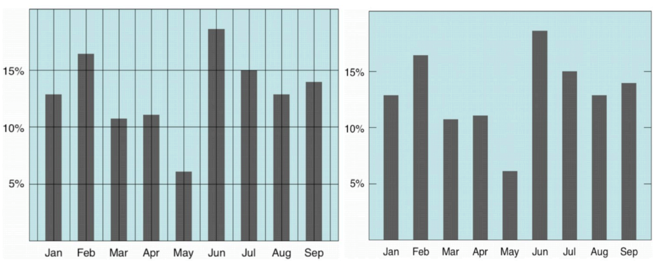
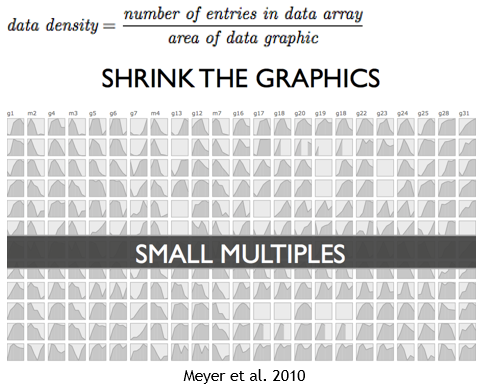
Integrity principles

* Missing scales
* Scale distortion
  + Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity.
  + The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
  + Unintended size coding

Large China and India clowns inadvertently make the size difference appear larger.

* + Show data variation, not design variation

Design Principle

* Maximize Data-ink ratio
* Avoid chart junk
* Maximize Data Density

From Tufte website:

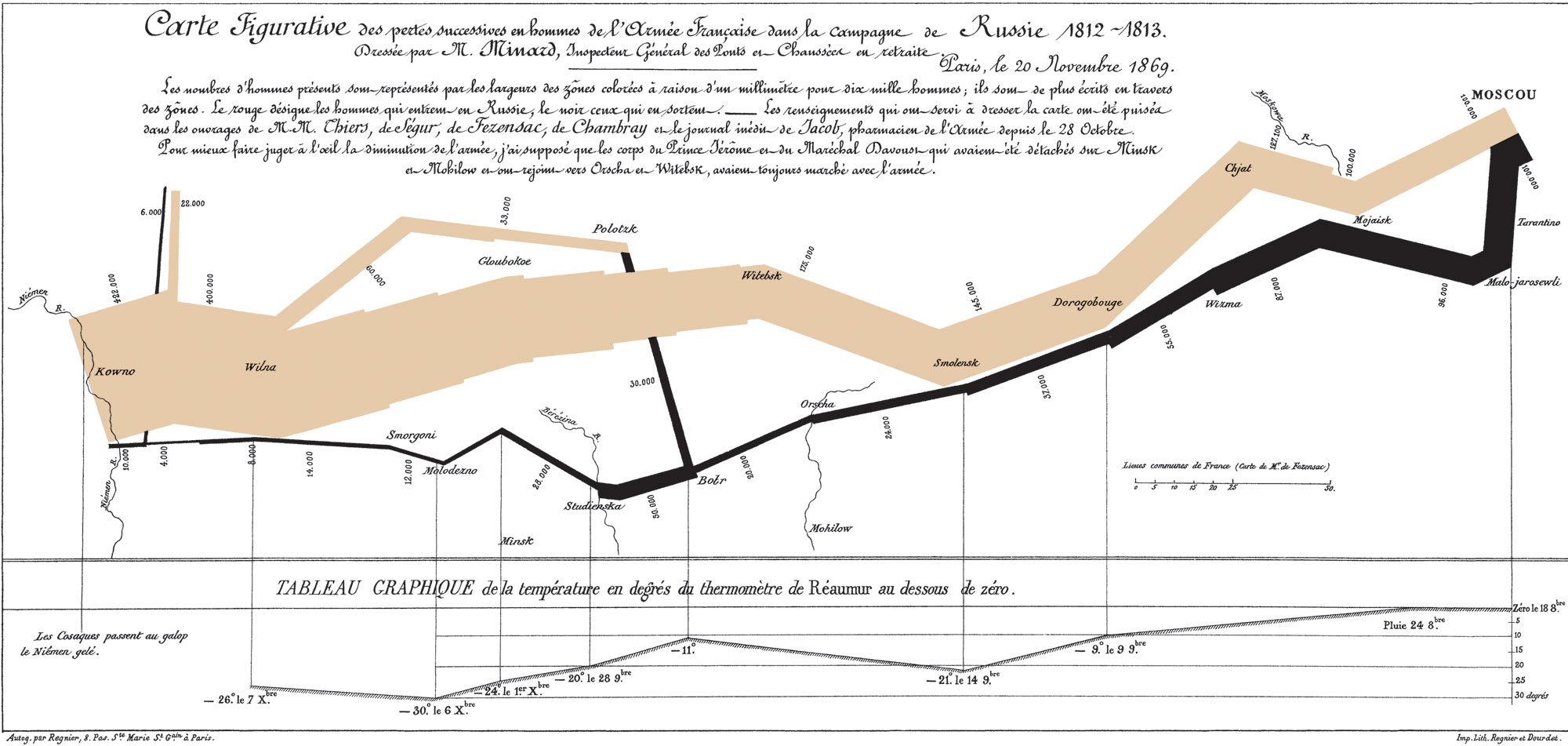
<https://sites.google.com/site/tufteondesign/home/six-fundamental-principles-of-design>

Tufte on Design and Data

Tufte suggests six fundamental principles of design: show comparisons, show causality, use multivariate data, completely integrate modes (like text, images, numbers), establish credibility, and focus on content. Each should be geared towards fully embracing the goals you defined for a given data display. For each principle, we outline examples of how to apply it to improve your visualizations.

Show comparisons

One of the major examples Tufte uses in showing comparisons looks at Charles Joseph Minard's map of Napoleon's march to and return from Russia. Minard shows the size of the French army by the width of the line. This makes plain how the army rapidly reduced in size as it marched east, and also makes the comparison between the size of the army that departed France and the size of the army that returned clear.

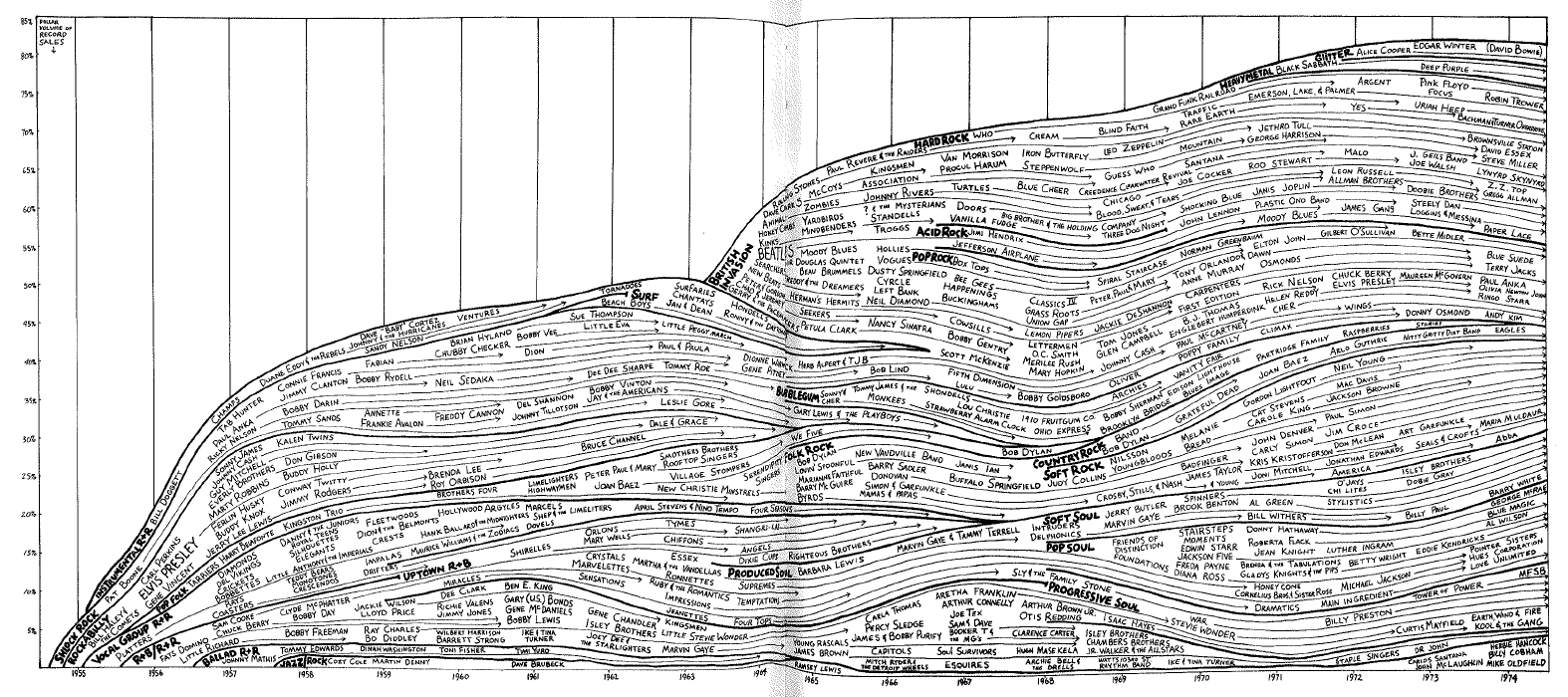


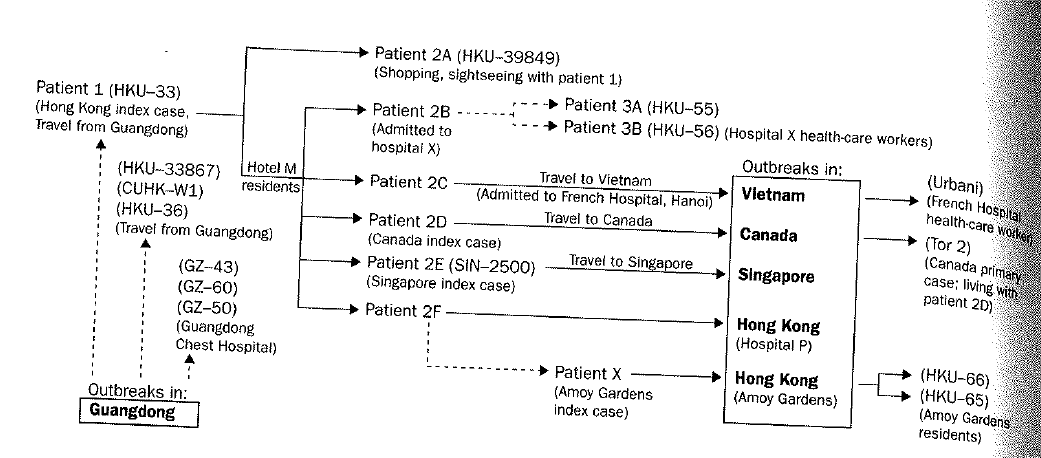
Edward Tufte, *Beautiful Evidence*(2006), 122

Show causality

In the example above, the goal was to show comparisons. Another possible goal is to show causality. In the above example, we can see causaulity in the black returning line of soldiers and the graph of temperature at the bottom of the chart. Minard shows how without even engaging in battle, the march itself killed thousands due to freezing temperatures.

Another example shows the influence of music groups on one another over a twenty year period, while a second shows the transmission of SARS.

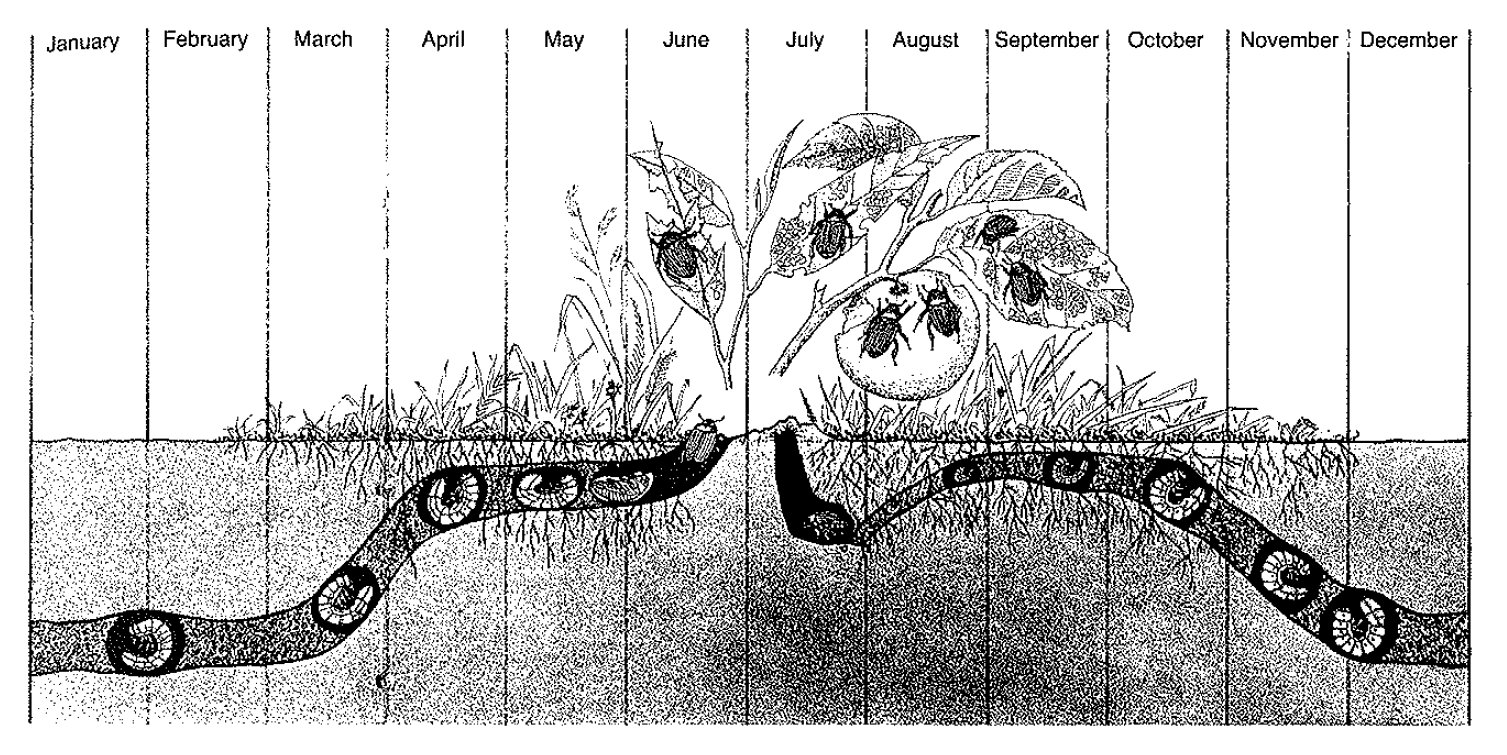




Reebee Garafalo from Edward Tufte, *Visual Explanations* (1997) p 90 and *Beautiful Evidence* (2006) p 78

Use multivariate data

The idea of using multivariate data may seem straightforward, but fully embracing this idea can lead to spectacular results. The example below, which outlines the lifecycle of the Japanese beetle, contains information on time, space, physical appearance, behavior, and interaction with its ecosystem. As this example shows, every relevant type of information should be included.

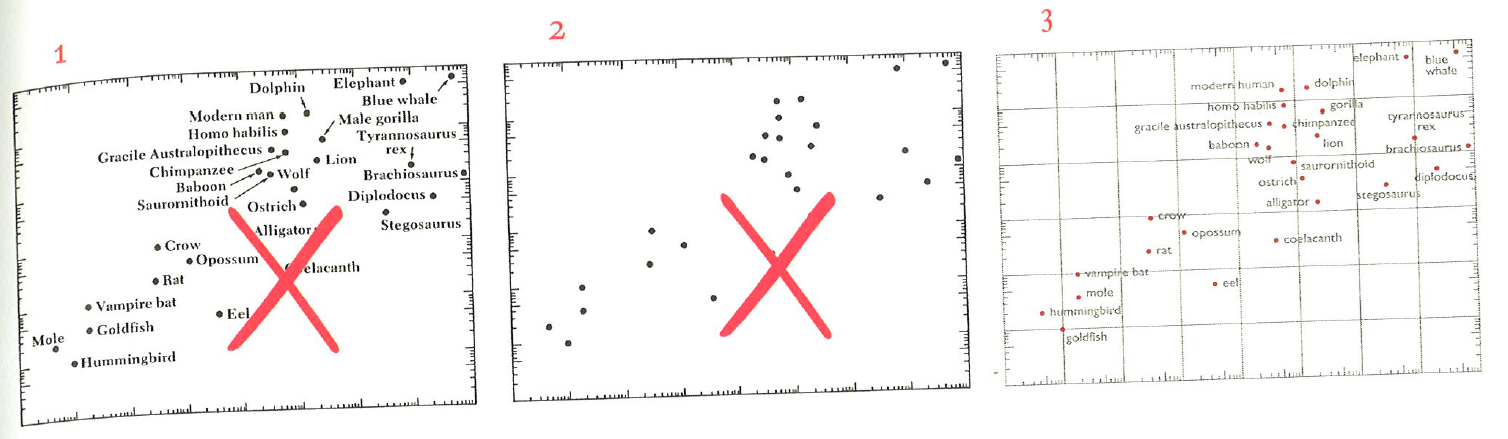


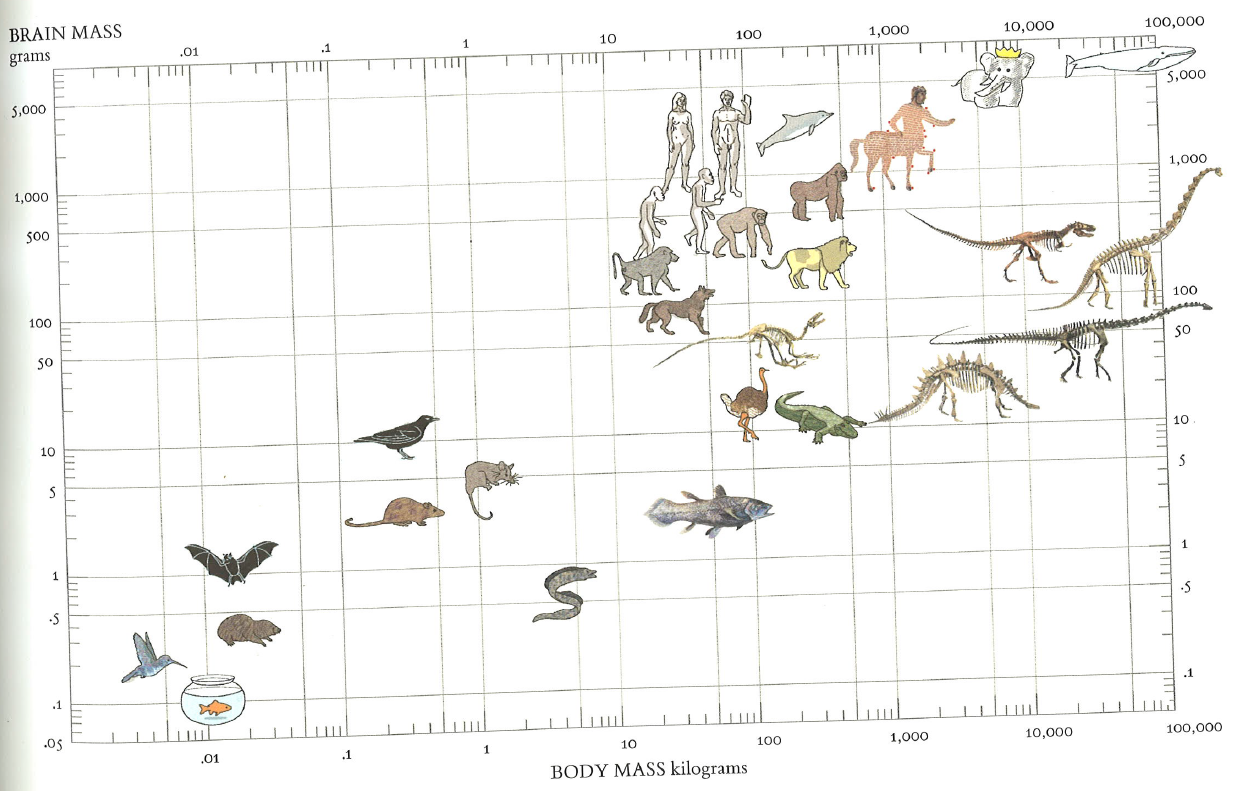
L. Hugh Newman, *Man and Insects* (London, 1965), 104-105

Complete integrate modes

To improve displays of data, it is important to completely integrate text, images and numbers. It is also important to include any information or data that is relevant. This is particularly important in the inclusion of explanatory text on figures. In the Minard example, this means an integration of the chart of the map and army size, but also text information about location, numerical information about the temperature, etc.

In the examples showing various information about animal sizes, we can also see how an effective integration of text can significantly improve readability of the chart, but that integration of pictures can make it even more meaningful.





Edward Tufte, *Beautiful Evidence* (2006) p 121

Establish credibility

Tufte argues repeatedly that including source information is one of the most important aspects of creating a convincing visual display or a convincing presentation. Allowing the viewers to access the source material will give them confidence in your results. The Minard chart is a classic example of this, prominently featuring both the author and his references in the text introduction at the top of the chart.

Focus on content

Above all else, Tufte argues that you must focus on the content to make an effective chart. That means, as outlined in the section on [best practices](https://sites.google.com/site/tufteondesign/home/practical-design-strategies), minimizing the chart architecture wherever possible. It also means that if you find a chart that is more "chartjunk" than content, you should research more and find more valuable content before you create or present your data. Even in the simplest cases, it is possible to display a rich world of information. In the example below, a simple bird book entry on Jays, the writer presents an abundance of information. There is information about common and scientific names, sizes and relative sizes, coloration, appearance in flight, environment and migration patterns, and even the sound of their calls. This kind of display driven by content will make the most "beautiful evidence."

