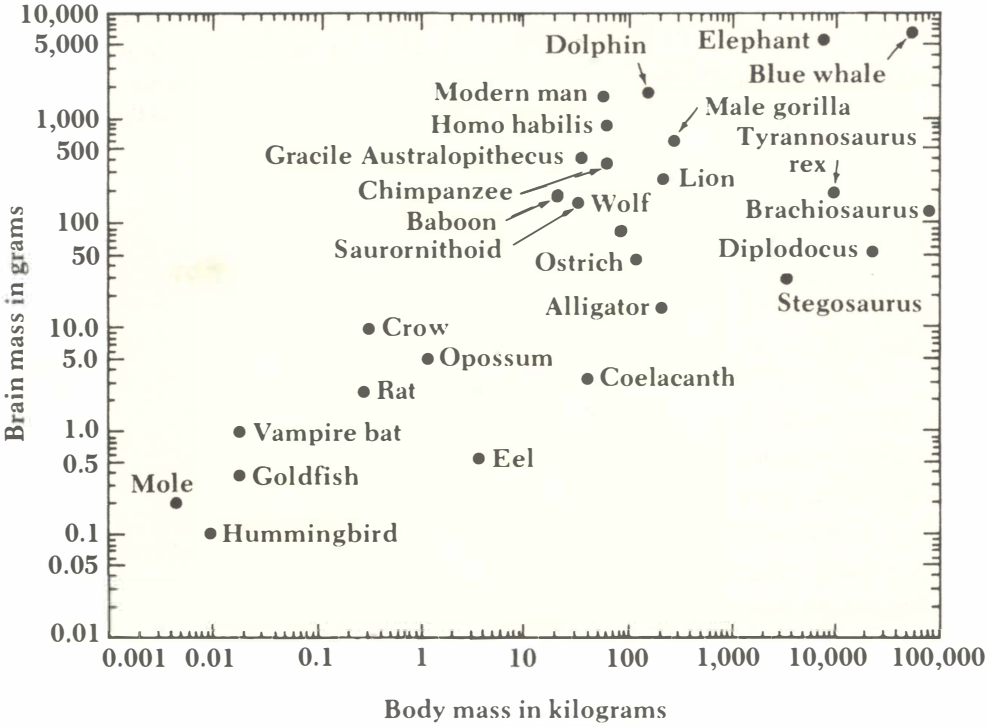
WHEN and how should data points in statistical graphics be labeled with words? In a classic book, *The Elements of Graphing Data,* William Cleveland suggests that word-labels on data may well "interfere with our assessment of the overall pattern of quantitative data."[19] Several examples then show interfering labels within data fields, including this

[19] William S. Cleveland, *The Elements of Graphing Data* (Monterey, California, 1985), 46.



Graph from Carl Sagan, *The Dragons of Eden: Speculations on the Evolution of Human Intelligence* (New York, 1977), 39, based on Harry J. Jerison, *Evolution of the Brain and Intelligence* (New York, 1973), 42-45.

noisy, cluttered scatterplot of the empirical relationship between body mass and brain mass for 26 animals. Cleveland's analysis suggests these imperatives for putting words on data points:

> Do not allow data labels in the data region to interfere with the quantitative data or to clutter the graphs. Avoid putting notes, keys, and markers in the data region. Put keys and markers just outside the data region and put notes in the legend or in the text.[20]

[20] Cleveland, *The Elements of Graphing Data*, 44-47.

Conflicting with the idea of integrating evidence regardless of its mode, these guidelines provoke several issues:

First, labels *are* data, even intriguing data. For example, among the really big animals, relatively smaller brains are found in the prehistoric tyrannosaurus rex, brachiosaurus, diplodocus, stegosaurus—a result that emerges from seeing data dots linked to their names. Or, why is the hummingbird shown as heavier than the mole, the wolf than humans? Such plotting errors can be more easily detected when data points are named. And where would the gnat, mosquito, cat, hammerhead shark, or centaur appear on the graph? Just like numbers, nouns are evidence.

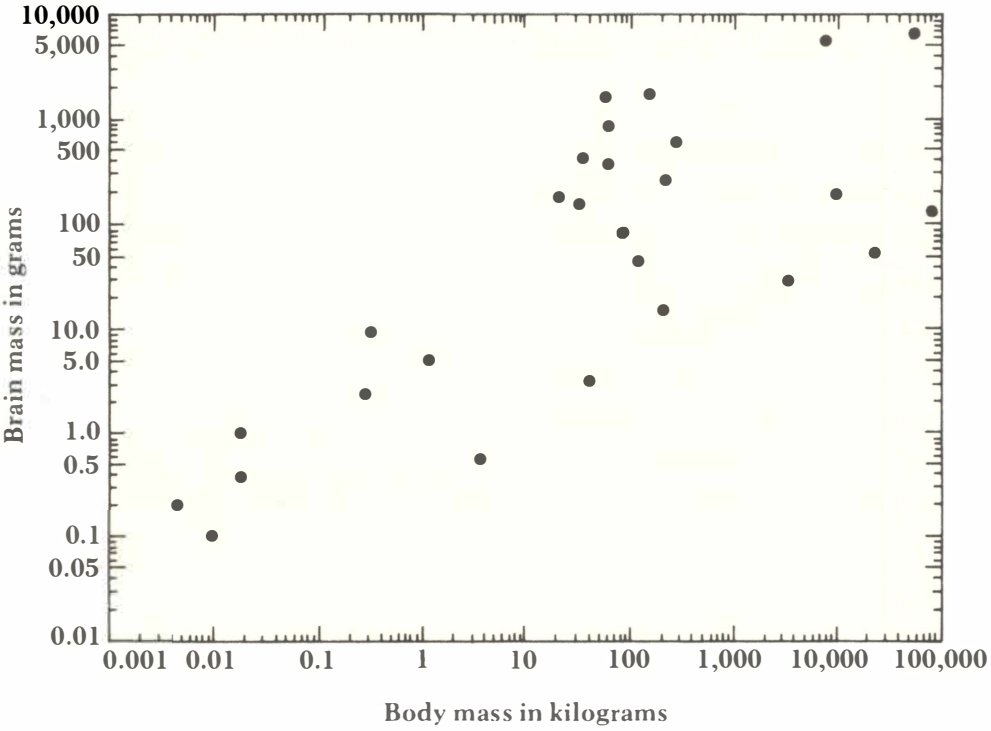
Second, when labels abandon the data points, then a code is often needed to relink names to numbers. Such codes, keys, and legends are impediments to learning, causing the reader's brow to furrow.

Third, segregating nouns from data-dots breaks up evidence on the basis of mode (verbal vs. nonverbal), a distinction lacking substantive relevance. Such separation is uncartographic; contradicting the methods of map design often causes trouble for any type of graphical display.

Fourth, design strategies that reduce data-resolution take evidence displays in the wrong direction.

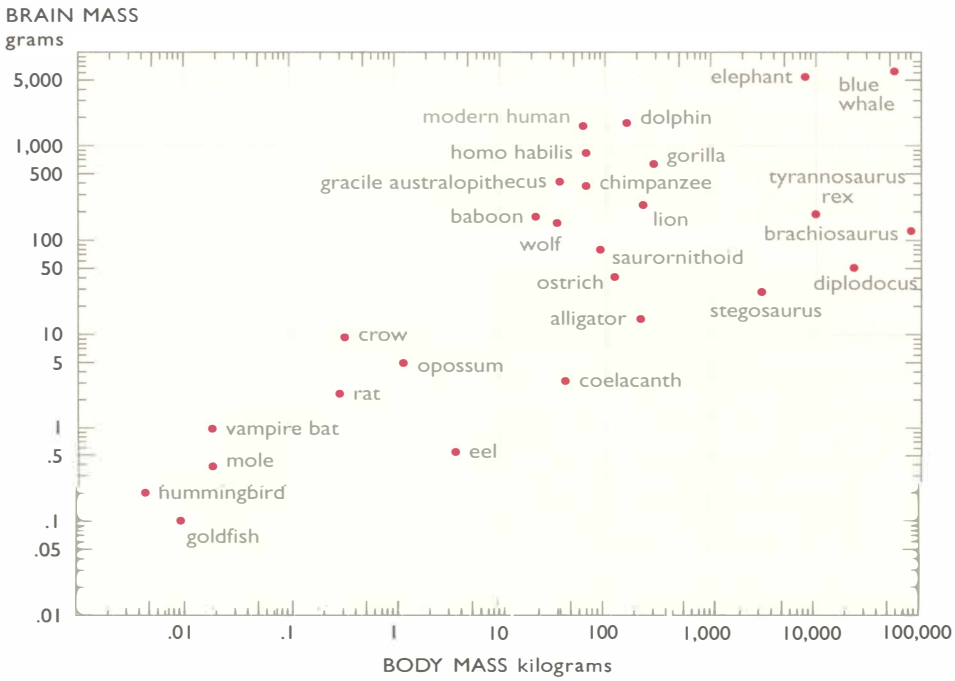
Fifth, what clutter? Even this supposedly cluttered graph clearly shows the main ideas: brain and body mass are roughly linear in logarithms, and as both variables increase, this linearity becomes less tight.

But verbal arguments do not resolve design questions. Visual evidence decides visual issues. And it turns out that Cleveland has a strong point. The 26 labels do in fact clutter up the graph, obscuring relations among the data. Perhaps something will show up if all the words disappear:



Without the dark typography of the labels, we see very differently: the big blob of words in the top half of the original graph inflates the visual variability of body mass for heavier animals. Thus one possible solution for label-clutter, particularly in exploratory data analysis, is to examine *both* scatterplots, with and without labels.

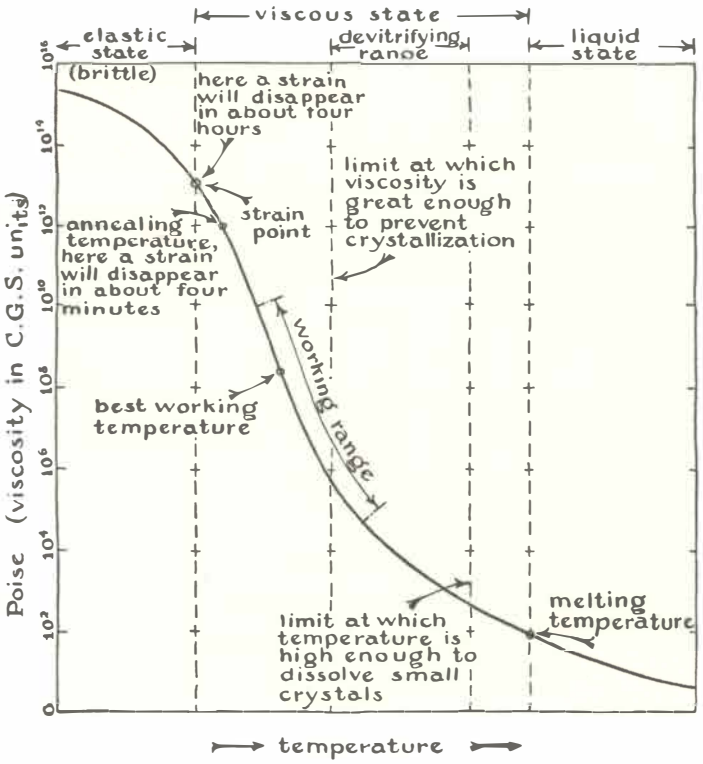
Good design, however, can dispose of clutter *and* show all the data points and their names. To repair this graphic, the data-dots need to gather themselves together on a *different visual level* from their labels. And the labels need to calm down. Like good maps, statistical graphics should have a layered depth of reading. Not a hierarchy of importance for verbal versus quantitative information, but rather a pluralism of distinctions. This suggests a redesign.

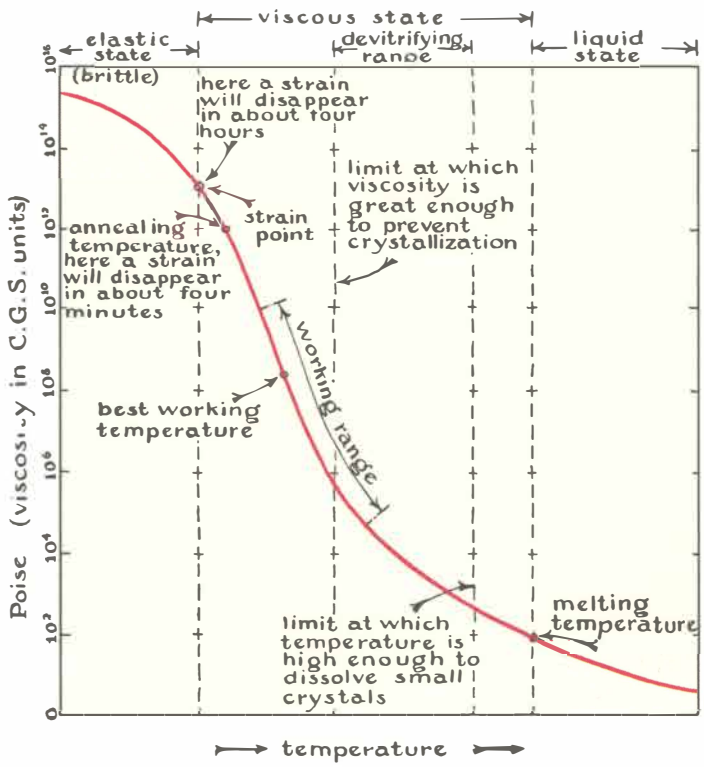


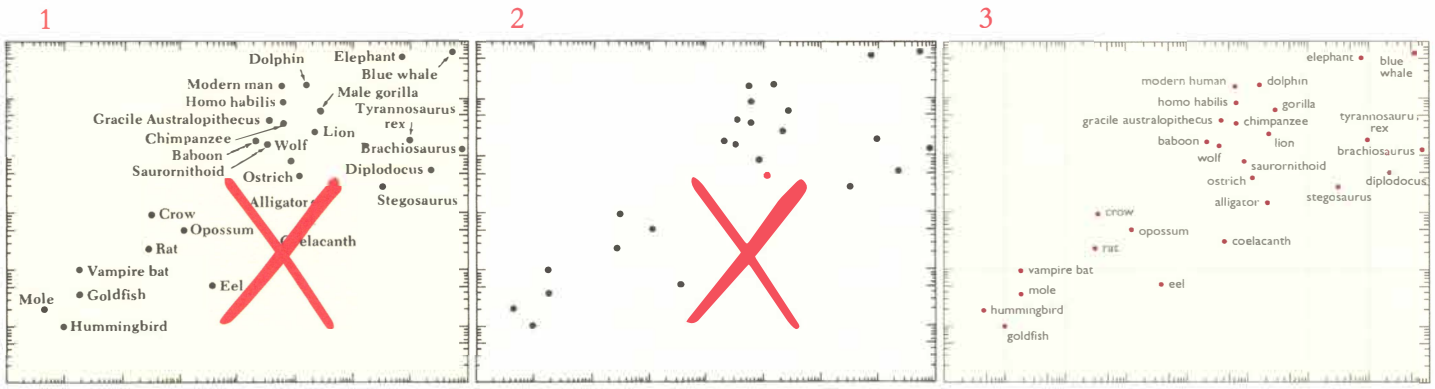
In this revised graph, *red* helps to cluster 26 data-dots now placed in a quiet field of grayed-down words. *Label clutter has vanished, but the labels are still there.* Clutter calls for a design solution, not a content reduction.

At lower left, this chart on the varying viscosity of glass in relation to temperature has an overall sameness of texture and color. Administrative elements (frames, grids, pointer lines, tick marks) are as visually active as the evidence curve itself. At right, the red color pulls out the curve from the graphic debris, while maintaining a unity of text and linework by means of the cartographic strategy of layering and separation.

Roger Hayward in John Strong, *Procedures in Experimental Physics* (New York, 1938), 6.







The redesign sequence reflects these fundamental principles: 1. Clutter is a failure of design, not an attribute of information. 2. Visual problems should not be fixed by reducing content-resolution (such as, for example, discarding words that label data). 3. Instead, fix the design.

Words and data-dots are abstracted representations of actual animals and body/brain masses. In a spirit of seeking visual solutions for visual problems, let each animal represent *itself* at its two-space location in the scatterplot below. Image sizes are proportional only to space available—except for the big brachiosaurus and tiny humans, shown together here at right, whose amazing relative sizes are approximately correct. Other details below repay study.