Tables as graphs: the Ramanujan principle

Andrew Gelman says that there is more visual information in a table than you might realise, so it is worth presenting them well.

Tables are commonly read as crude graphs: what you notice in a table of numbers is (a) the minus signs, and thus which values are positive and which are negative, and (b) the length of each number, that is, its order of magnitude. The most famous example of such a read might be when the mathematician Srinivasa Ramanujan supposedly conjectured the asymptotic form of the partition function from simply looking at a table of the first several partition numbers (see box): he was essentially looking at a graph on the logarithmic scale.

Table 1 shows a simple example, a list of the populations of the five largest countries of the world and a selection of smaller ones.

This table is full of numbers (and would be even more cluttered had we not rounded the higher populations to the nearest million), but the most natural way to read it is as a graph: China and India have four digits (which corresponds to having more than a billion people), the next few countries have three digits, then come the two-digit countries and finally the single-digit ones, whose populations are less than 10 million each.

A bit more information is conveyed by the size of the leading digit. As Howard Wainer has noted, numbers also convey some information within an order of magnitude: the digit “1” takes up less “ink” than any other digit. And this is generally relevant: from Benford’s law we know that approximately 30% of numbers begin with 1. (For more on Benford’s law, see Chris Weir’s article on page 164, or Significance, June 2007. Benford’s law can be seen in operation in Table 1, where 5 of the numbers begin with 1, two begin with 2, two begin with 3, and none at all begin with 4,5,7,8 or 9.) At the other extreme, the numerals 6, 8 and 9 are physically large. The physical size of the leading digit gives a clue – an imperfect clue, but a clue nonetheless – to a number’s magnitude. Figure 1 shows the pattern for the digits 1–9 as they appear on a liquid crystal display calculator. Larger digits take up more ink: the correlation between the digits 1–9 and their display ink is a stunning 0.58!

In a table of statistical results, the reader might also note the boldface type or stars that indicate statistical significance. In addition, the physical placement of the numbers in a table points towards possible views: it is much easier to compare two numbers aligned vertically than to make a horizontal comparison, and other cues such as font size and colour can guide the reader even more (sometimes in a way perhaps not intended by the table’s creator).

The Ramanujan principle supports the recommendation to scale up numbers so they are generally larger than 1 in absolute value. For example, 58% and 9% are easier to tell apart (based on the length-of-number cue) than 0.58 and 0.09. If we really must display numbers in tables with many significant figures, it would probably generally be better to display them like this: 3.1416, so as not to distract the readers with those later unimportant digits.

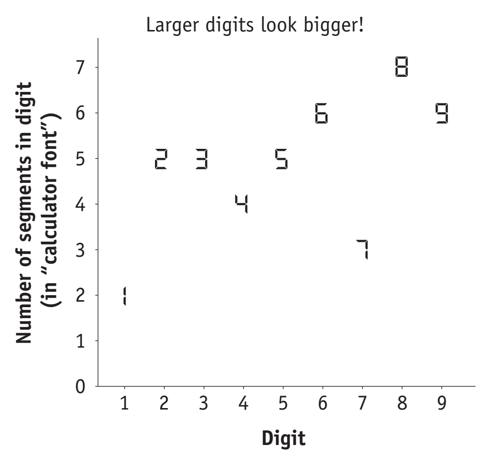

Figure 1. Physical size of the digits 1–9

I still prefer to display numerical information in graphs – and I have expressed this preference in published research in topics ranging from congressional elections to arsenic in Bangladesh, from toxicology to opinions on health care. But if you do present tables, it is good to understand how they might be viewed. It is naïve to consider a table as a simple data dump; rather, it is crude graphical display.

Andrew Gelman is a professor of statistics and political science and director of the Applied Statistics Center at Columbia University.

Table 1

Country	*Population (millions)*
China	1339
India	1210
USA	311
Indonesia	237
Brazil	190
Mexico	112
Thailand	67
Canada	34
Guatemala	14
Jordan	6.2
Jamaica	2.7

The partition function

The partition function gives the number of ways of writing an integer as a sum of smaller positive integers. Thus 4 can be written in five different ways, as:

4, 3+1, 2+2, 2+1+1, 1+1+1+1

The partition number of 4 is therefore 5. The partition numbers of the numbers 1–10 are:

Number	1	2	3	4	5	6	7	8	9	10
Partition number	1	2	3	5	7	11	15	22	30	42

In other words, there are 42 ways of summing smaller integers to make 10. The partition number of 100 is 190 569 292. The partition number of 1000 is 24 061 467 864 032 622 473 692 149 727 991. It was from looking at a table of such numbers that Ramanujan made his conjecture that as n becomes very large, the partition number of n tends to

$$p(n) \to \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$