

1

Understanding number

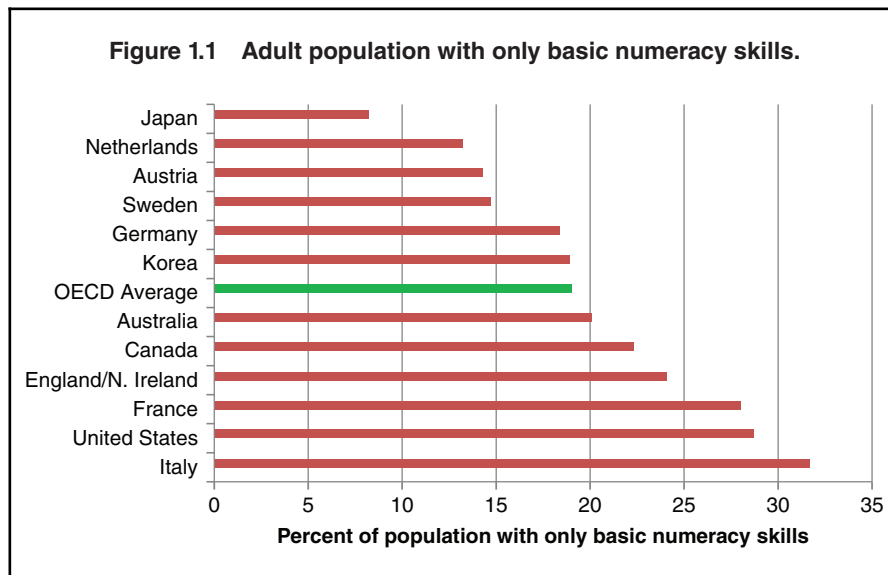
This chapter deals with the basics of how users perceive numbers and how they should be presented to allow maximum uptake of the meaning of the data.

A false assumption of many who seek to communicate numerical information is that their audience is as able to handle the information as they are. The available data in for the majority of OECD countries is that numeracy skills are significantly below those for literacy.

Numeracy skills in the general population in England are poor – and are not improving. Information from a skills survey in England in 2011¹ noted that Numeracy skills had declined slightly since the last survey in 2003. Seventeen million adults in England in 2011 (just under half the working-age population) were at ‘Entry Levels’ in numeracy – roughly equivalent to the standards expected in primary school. Further, the survey showed that 78% of the working-age population were at or below level 2 numeracy. These people may not be able to compare products and services for the best buy, or work out a household budget; essentially they would not achieve a good mark in a mathematics examination at age 16.

More generally, the OECD published the results of a survey in October 2013² showing that 19% of adults in 21 countries had mathematics skills at the level of a 10-year-old. These adults could only manage one-step tasks with sums, sorting numbers or reading graphs: many could only perform sums with money or whole numbers. For individual countries, the percentage at this level of numeracy ranged from 8% in Japan to 32% in Italy. A selection of the results is shown in Figure 1.1. For England and Northern Ireland, the estimate is that around 8.5 million adults are at this level.

2 PRESENTING DATA



1.1 Thousands separator

When considering a number with more than three digits, most of us will not take in the whole number but will try to understand it. If the number is just a set of digits together without any separators, most will try to find the place for thousands separators to form the basis of understanding the number. Then, depending on the size of the number, will try to do some rounding to understand the number and put it into our mental understanding range.

Example 1.1

Let us consider a number which represents the population of a country.

12573981

How do we try to understand the number? None could honestly say they could read the digits as given and fully understand it without any other process going on in our brains.

So what do we do?

First our brains need a little help. We can introduce separator characters to split the digits into groups of three, starting from the right-hand side. These separators are known as ‘thousands separators’. In some countries (United States and United Kingdom), the separator is a comma; for much of Europe, the separator is a space but in some European countries a full stop is used. I visited one country’s statistical

UNDERSTANDING NUMBER 3

office website and found three different practices: comma, space and nothing used for the separator! Whichever is chosen, it must be consistently used throughout an organisation's output.

The number, with these separators would then appear as either

12,573,981 or 12 573 981 or 12.573.981

This process here helps us to identify that we have 12 million, 573 thousands and 981 units.

But how does our brain handle all this information? It doesn't. The brain will not retain all of the information but introduce some abbreviation of the data. This could either be by truncating the number (i.e. throwing some digits away): for example, after the millions – giving 12 million. Many will argue that a lot of information has been thrown away: but, does it matter? Only, one could argue, if it affects the understanding. Here the number is actually nearer to 13 millions: so would it be better to think of the number in this way? Truncating is easy but may not always be helpful in aiding our understanding. Without another number to compare this one with, it is probably acceptable.

Another approach would be to round the number. Rounding is the process where we reduce the number of digits and say that the number is 'nearly' or 'about' something. The difficulty with rounding is that few know how to do it well!

Should the number be rounded to millions? Hundreds of thousands? Or thousands?

Looking independently, just at this one number, most would round the number to 12.6 million. Is this right? It may be. Let us leave the question for a little while.

Principle 1.1: Choose one symbol for the thousands separator and use consistently.

.....

1.2 Decimal separator

Just as differences of symbol exist for thousands separators, some differences exist for the decimal separator. In the United States and the United Kingdom, a full stop is always used as a decimal separator; in Europe, some countries use the full stop and some a comma. Whichever is used, again, it must be consistently used throughout an organisation's output. One example from a national statistical office showed the use of both a comma and a full stop in the same table! An extract of the table is given in Table 1.1. Note also, in this example, that different rounding has been applied to data in lines two to four.

4 PRESENTING DATA

Table 1.1 Different decimal separators used in the same table.

Demographic indicators	Basic social indicators			
	1970–75	1980–85	1993–99	2005
Life expectancy at birth (years)	66.6	69.7	66.9	70.6
Infant mortality (per thousand live births)	80.8	46.8	47	65,1
Birth rate (per thousand people)	37	40	32.6	31,4
Death rate (per thousand people)	8.1	7	5.3	4,4

Principle 1.2: Choose one symbol for the decimal separator and use consistently.

1.3 Level of detail in comparisons

In seeking to communicate data, it is important to have the same denomination and rounding of data in any data to be compared. So it is more difficult for the user to compare £1.8 billion of turnover with a profit of £232.9 million: it should be phrased as £1.82 billion of turnover and £0.23 billion of profit. Similarly, comparing profit across years, an increase from £246.4 million to £386 million would be better as ‘an increase from £246 million to £386 million.’ An even worse example of presentation would be to compare the number of homes in the stock, 1.8 million, with a new build figure of 9,768 in the last year: this means nothing to most readers and would be better either comparing the number of this year’s newly built homes with that of the previous year or noting a percentage increase (though this would be very small).

In Table 1.1, some of the data is given to one decimal place and some to whole numbers. This may be a result of using Excel which would normally drop trailing zeros unless forced to use the same number of decimal places in the data.

In one press release I found the following statement:

11.048 million NHS sight tests, an increase of 563 thousand (5.4%) on the previous year.

Here the user is expected to convert the second number to millions. We have spurious accuracy and difficulty in understanding. This would be better as:

The number of NHS sight tests given increased by 0.6 million (5%) to 11.0 million.

Principle 1.3: Always provide data for comparisons in the same units.

1.4 Justification of data

When we did our first mathematics lessons in school where we had to add numbers of a magnitude over 100, we were taught that we should start by putting three letters in a row as:

h t u

These letters represent hundreds, tens and units. We were encouraged to write the numbers neatly in the correct columns below the letters. This places the digits of equal value in the same column.

The use of different typefaces in the preparation of tables is easy today. Personal computers have over a hundred preinstalled typefaces which can be used in basic word processing and spreadsheet packages. Choosing the wrong typeface can interfere with our understanding of the presented data: many examples can be seen in simple tables in magazines or newspapers.

Example 1.2

Consider the following data:

20,011
111,111
88,444

Here, using the Constantia typeface, the proportional spacing for the digits has confounded the positioning of the columns of data. Also here the greater height of the '8' draws attention to them. Why is this important? When asked to identify the largest number in a column of data, we have been taught to look to the left of the column of numbers: the one that is further left (if the data are right justified) is, we believe, the largest number.

This, however, relies on a principle of presentation that more and more people are ignoring. In visits to many official government websites, I can find data centred or left-justified in columns of tables (a simple one-button operation in Excel): such presentation is definitely wrong. Let me explain why.

In this example, using this principle, we would conclude that the bottom number is the largest as its first digit is further to the left than the first digit of the first number. Examining the data clearly indicates that this conclusion is wrong as the number in the middle has an extra digit and even that is too far to the right.

.....

Principle 1.4: Ensure typeface used for data presentation has the same width for all digits.

6 PRESENTING DATA

Principle 1.5: Right justify data in columns, ensuring digits of equal value are under similar valued digits in the other numbers.

One other poor presentation of data is often seen in the investment sector. Performance data for various funds are given to differing numbers of decimal places. In two cases recently, not only were the data so presented but, in addition, the data were centred within columns as shown in Example 1.3.

Example 1.3

(a) Centre-justified and different rounding

Fund	Performance over 1 year
A	-46.2%
B	138%
C	125.92%
D	111.365%
E	-31.7%
F	104.25%

If the numbers are right justified, we have:

(b) Right-justified and different rounding

Fund	Performance over 1 year
A	-46.2%
B	138%
C	125.92%
D	111.365%
E	-31.7%
F	104.25%

Now the investor is looking for the largest return. Using the principle of looking to the left of the column, fund D is the highest performance. But looking carefully at all of the data, this is clearly wrong. Here one needs either to round the data to the same level or align the data on the decimal point. Given the variation in the data, the preference here would be rounding to whole numbers as in c. Note the removal of the % symbol from each data item and the change to the column heading.

(c) Right-justified and similarly rounded

Fund	Performance over 1 year (%)
A	-46
B	138
C	126
D	111
E	-32
F	104

.....

Principle 1.6: Where possible, round data to be compared to the same level.

1.5 Basic rounding

The main purpose of rounding data is to put the numbers into a memorable and understandable range. Nobody would quote the population of the world to the last digit as it is increasing every second. Also, to write a number with 10 digits is beyond most people's comprehension. According to the United Nations³, the world's population reached 7 billion on 31 October 2011 and by 22 October 2012, it had reached 7,074,173,556. Now taking Principle 1.6 here, we would have 7.0 billion in 2011 and 7.1 billion in 2012.

Computer programs can handle numbers less than or greater than 0.5 easily: those less than 0.5 are rounded down; those greater than 0.5 are rounded up to the next whole number. However, most computer programs would round any number with an exact 0.5 up to the next whole number. But what is correct? If we start with a line from 0 to 1 and divide it into equal small chunks from either end, we will find that 0.5 is exactly in the middle – neither nearer to the 0 or the 1. So which way should it be rounded? To avoid bias, numbers ending in exact '.5' should be rounded to an even digit.

3 divided by 2 = 1.5 (exactly); rounded to a whole number gives 2

5 divided by 2 = 2.5 (exactly); rounded to a whole number gives 2

7 divided by 2 = 3.5 (exactly); rounded to a whole number gives 4

This principle is easily extended to other numbers: 0 to 4 would be rounded down; 6 to 9 would be rounded up. Similarly, if rounding to the nearest thousand, 0 to 499 would be rounded down and 501 to 999 would be rounded up. So,

14,245 is rounded to 14,000

14,675 is rounded to 15,000

8 PRESENTING DATA

But what happens when we have an exact 500? Technically, it is neither nearer the lower thousand nor the higher one. Again, the principle to be employed should be to round to an even digit. Thus

14,500 is rounded to 14,000

123,500,000 is rounded to 124,000,000 or written as 124 million.

One of the most frequently asked questions with regard to rounding is ‘To what level should I round?’ In Example 1.1, it was suggested that most would round the number to 1 decimal place of millions. But that is just one number on its own.

Now, instead of just considering one population number, let’s consider two.

Example 1.4

**12573981 and
111894397**

When most people see numbers like this, they ignore them – as they are outside their mental understanding. These numbers also start at the same position on the page but, using Principle 1.5, we need to move the bottom number one digit position to the left.

**12573981 and
111894397**

Then it is clear that the bottom number is the larger of the two as it has a digit further to the left than the first number.

The first task in trying to understand the data is to add thousands separators. So, introducing these, we have:

**12,573,981 and
111,894,397**

If we round these numbers as in Example 1.1, we would have:

**12.6 million and
111.9 million**

We can now ask if we need the digit representing the hundreds of thousands: probably not. These would then round to

**13 million and
112 million**

.....

Such rounding puts the numbers within the boundaries of our understanding and they are much more memorable than the raw numbers. But what is the appropriate level of rounding for numbers? Before answering that, let me give another example.

Example 1.5

The population of Wales in 2012 was estimated to be 3,074,100; the number of sheep in Wales was estimated to be 8,898,200.

Neither of these numbers is particularly memorable as written in the above paragraph. However, if we said:

The population of Wales is	3 million
and	
The number of sheep in Wales is	9 million

These numbers would be in our minds for some time. If we added another statement, we would communicate even better – and make the information totally memorable:

For every person in Wales, there are 3 sheep!

.....

So how should the numbers be rounded to be as helpful to the recipient as possible? When comparing numbers, it is very difficult for the brain to deal with three digits. The recommended level is two digits: in education-speak, this is taking advantage of the number bonds up to 100. Purchasing a chocolate bar for 65 pence and offering a £1 coin to the shopkeeper, most can work out that the change should be 35 pence. This use of decimal-based currency builds up the number bonds to 100 in our memory so that the differencing of two numbers with two digits is relatively easy.

Beyond the two digits of comparison, most would take out a calculator! When such comparisons are demanded through reports, on company profits for example, many people either ignore the numbers or misunderstand them. Hence a communication failure has occurred.

One concept that has been developed to overcome this issue is effective rounding.

1.6 Effective rounding

Using this concept, numbers to be communicated are presented in such a way that the user can quickly appreciate the message and understand the information. Sufficient digits are given so that the magnitude of the numbers is retained but the information inconsequential to the comparison is discarded.

Let us take two numbers representing the profits of a company in successive years:

Year 1: £13,484.2 million Year 2: £15,857.4 million.

10 PRESENTING DATA

The user has to concentrate very hard to identify the difference. The first reason this is the case is that the numbers are across the page (and in a report are usually separated by text). The second reason is that the numbers are outside the mental range of understanding of most: so, if time is spent on doing the comparison, both will get rounded in the user's mind before the comparison is done: £13.5 million and £15.9 million – from which we can easily judge which is the larger.

So let us make this a little easier and put the second number under the first.

Year 1: £13,484.2 million
Year 2: £15,857.4 million

With this presentation, we start looking at the left-hand side of the numbers and start comparing digits. The second digit identifies that the Year 2 profits are larger. But still, we have to identify how much larger and retain the information.

Is the decimal part of the numbers important? Not really: dropping them does not make a difference to our understanding.

What about the units of millions? The tens of millions? The hundreds of millions?

As discussed, we need to have two digits for a discriminating comparison: obviously, where the digits are the same, no discrimination is possible.

Taking these numbers, we see that the first digit in both is the same: not effective therefore in discriminating between the numbers. The second digits are different and thus effective in discriminating between the numbers: so this digit is the first effective digit in the comparison. Diagrammatically, we have:

Year 1:	£1	3	,	4	8	4	.	2	million
Year 2:	£1	5	,	8	5	7	.	4	million.
		⊗	⊗						

Where signifies not effective in discrimination and signifies effective in discrimination.

The principle is that we are considering number differences (in schools, these are termed number bonds) up to 100; so the next digit will give us that. This is the same whether or not the digit is the same value or not. Here we have a 4 and an 8; they are

UNDERSTANDING NUMBER 11

different. That is to say, once a column of discriminating digits is found, the next column is counted as the second effective column regardless of the digits in the column.

Year 1:	£1,3,4,84.2 million
Year 2:	£1,5,8,57.4 million.
	<div style="display: flex; justify-content: center; gap: 10px;"> × ✓ ✓ </div>

Having found our two effective digits for comparison, we round to that level

Year 1:	£1,3,4,84.2 million
Year 2:	£1,5,8,57.4 million

and have:

Year 1: £1 3, 5 0 0 million
Year 2: £1 5, 9 0 0 million.

This gives us a difference to calculate of

$$59 - 35 = 24$$

And, into the context, a difference of

£2,400 million.

Alternatively, the numbers could be presented as:

£13.5 and £15.9 billion
with a difference of £2.4 billion!

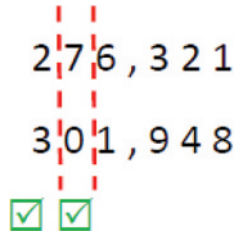
Another example:

People in a city on two census dates:

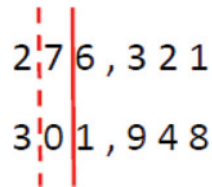
276,321 and 301,948

12 PRESENTING DATA

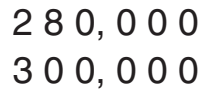
Putting them in a column and introducing the digit discrimination:



So we round to the second digit:



Which gives:

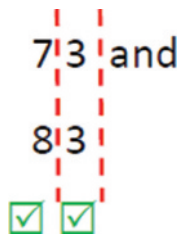


And then anyone can easily see the change of 20,000 and appreciate how much of a change from the first figure ($20/280 = 1/14$).

Many will argue that a loss of precision has occurred: that is true – but the increase in understanding and communication of what is happening more than outweighs this. And who can say that the numbers in their full form are accurate to the last digit. This is not to say that the data to the last digit is not necessary: such detail is required for calculations with the data. Further, the full detail of the numbers can be given in a reference table and presented in an Annex.

Finally for this section, a very simple example: two scores on examinations were 73 and 83.

Putting the numbers in columns and looking for the effective digits, we have:



So rounding to the second effective digit

73 and
83

gives

73 and
83

That is, the numbers are left without rounding. Note that here the second digit is the same in each of the numbers. The rule is that once the first column of digits effective for discriminating between the numbers is found, the next column is automatically effective even if it is the same digit.

So, using effective rounding assists the user to understand what message you are trying to get across. It is not a universal panacea to communicating numbers but it certainly helps the provider to consider whether the data are presented in a way they would be understood. Using effective rounding may suggest other issues with the data and may make one wonder whether the data should be included in the presentation.

Let's consider the fairly basic table of population data from 1995 in Table 1.2. The countries included were the countries from which lecturers came to teach on a course in London in 1996.

Table 1.2 Populations of countries, 1996 – 1.

Country	Total population
France	58020376
Denmark	5215732
Germany	81538628
United Kingdom	58491643
Luxembourg	406589
Total	203672968

Now, from the earlier Principles in this chapter, we see that this table has a few problems: the numbers are left-justified; no thousands separators are shown; no rounding of the data has been done.

If we take the issues in order, first let us right-justify the data which gives us Table 1.3.

14 PRESENTING DATA

Table 1.3 Populations of countries, 1996 – 2.

Country	Total population
France	58020376
Denmark	5215732
Germany	81538628
United Kingdom	58491643
Luxembourg	406589
Total	203672968

At least this gives us some sense of the size of the numbers but too much detail is given to really understand the data. So, rounding effectively, we have Table 1.4.

Table 1.4 Populations of countries, 1996 – 3.

Country	Total population (millions)
France	58
Denmark	5
Germany	82
United Kingdom	58
Luxembourg	0
Total	204

The question one would have to ask here is whether it was important to show the data for Luxembourg. In order not to cause an international incident, the answer in this case was that it is important. But still we have a problem: to show Luxembourg as zero is a little of an insult. To overcome this, it is necessary to show the data to 1 decimal place – as in Table 1.5.

As an aside here, it is obvious that the total shown does not equal the addition of the numbers in the table as it is independently rounded from the raw data total. Most people would now understand this but if queries arise, add a footnote to the table indicating that ‘The total shown may not be the sum of the constituent parts because of rounding’.

Table 1.5 Populations of countries, 1996 – 4.

Country	Total population (millions)
France	58.0
Denmark	5.2
Germany	81.5
United Kingdom	58.5
Luxembourg	0.4
Total	203.7

Principle 1.7: Use effective rounding of data where possible to aid communication.

Even this table can be improved by sorting the data by either by country or, better in this case, by size of population. This would give Table 1.6.

Table 1.6 Populations of countries, 1996 – 5.

Country	Total population (millions)
Germany	81.5
United Kingdom	58.5
France	58.0
Denmark	5.2
Luxembourg	0.4
Total	203.7

It is then clear that Germany had the highest population and that of the United Kingdom is just higher than that of France. So a message can be more clearly seen in the rounded and sorted table than in the original Table 1.2.

Principle 1.8: In summary tables, consider sorting the data to elicit a message.

Summary of principles: Understanding number

- Principle 1.1: Choose one symbol for the thousands separator and use consistently.
- Principle 1.2: Choose one symbol for the decimal separator and use consistently.
- Principle 1.3: Always provide data for comparisons in the same units.
- Principle 1.4: Ensure typeface used for data presentation has the same width for all digits.
- Principle 1.5: Right justify data in columns, ensuring digits of equal value are under similar valued digits in the other numbers.
- Principle 1.6: Where possible, round data to be compared to the same level.
- Principle 1.7: Use effective rounding of data where possible to aid communication.
- Principle 1.8: In summary tables, consider sorting the data to elicit a message.

Notes

1. UK's BIS Skills for Life survey 2011. BIS Research Paper 37, December 2011, URN 11/1367 <http://www.nationalnumeracy.org.uk/news/16/index.html>.
2. Table A2.5, page 242 of http://skills.oecd.org/documents/OECD_Skills_Outlook_2013.pdf.
3. 2010 Revision of the World Population Prospects, UN, <http://esa.un.org/unpd/wpp/index.htm>.