

# Presenting Data

Thinking beyond the surface.

Yesoda Bhargava

A drug advertisement:

“Antibiotic 100% effective in chronic urinary infections”

What do you think?

What if I told you something more?

What is it that you want to know?

Sample size.

What if sample size is 8?

Are you still impressed?

# Tribal Tanzania study

A deadly disease exists in Tanzania and the national incidence rate is 0.1 per 10000.

A researcher performs the study in a tribal area and reports a rate of 1.3 per 10000.

There is a huge difference in the rates.

Until?

Wait! You ask, what is the sample size.

Glad, you asked. 8000.

For the rate to be 1.3 per 10000, how many deaths are required in sample of 8000?

1.

Are you impressed?

# What is the lesson?

- Rates and proportions are powerful tools, but we must beware of them becoming detached from the original data.
- Next time, look at percentage and sample both when you read news, research paper or even in friendly discussions!
- Quote both in your papers, the rate and sample size.

# Decimal places vs Significant Figures.

- What is the difference?
- First few non zero digits of the number are important and we call these as **significant digits** or **significant figures**.
- There is usually little point in quoting statistical data to more than three significant figures.
- 0.001096 has how many significant digits? How many decimal places?
- What if I want three significant digits in this number? How will I write?
- 0.00110. The leading zeroes are not significant.

- The number of significant digits to put depend on the use to which the number is to be put and on how accurate it is anyway.
- Eg. If I am measuring temperature to the nearest half degree, then there is little point in quoting mean to more than 3 significant figures.
- NEVER ROUND NUMBERS TO FEW SIGNIFICANT DIGITS BEFORE CALCULATIONS HAVE BEEN COMPLETED.
- Finish all your calculations then round.

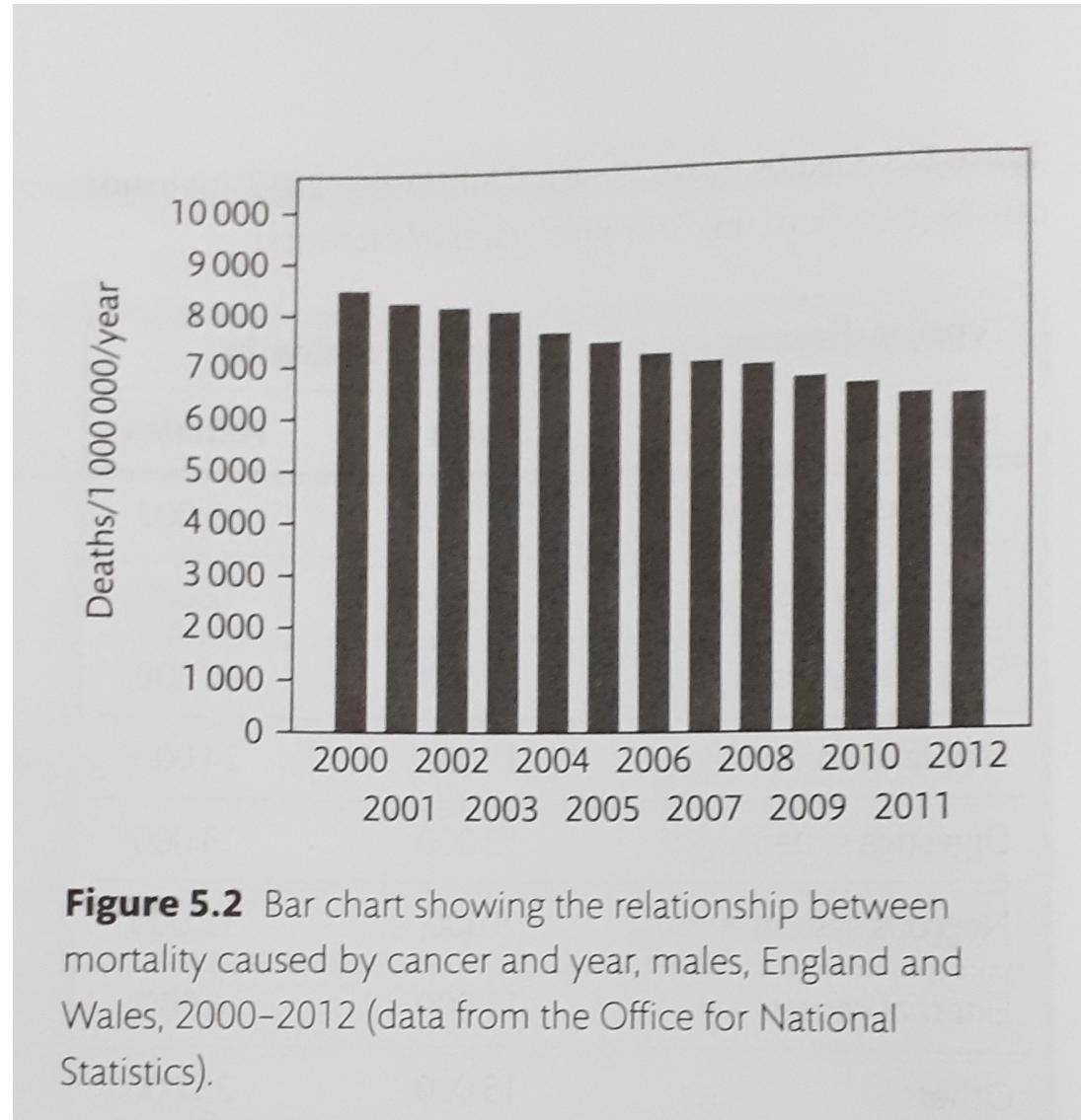
# Presenting tables

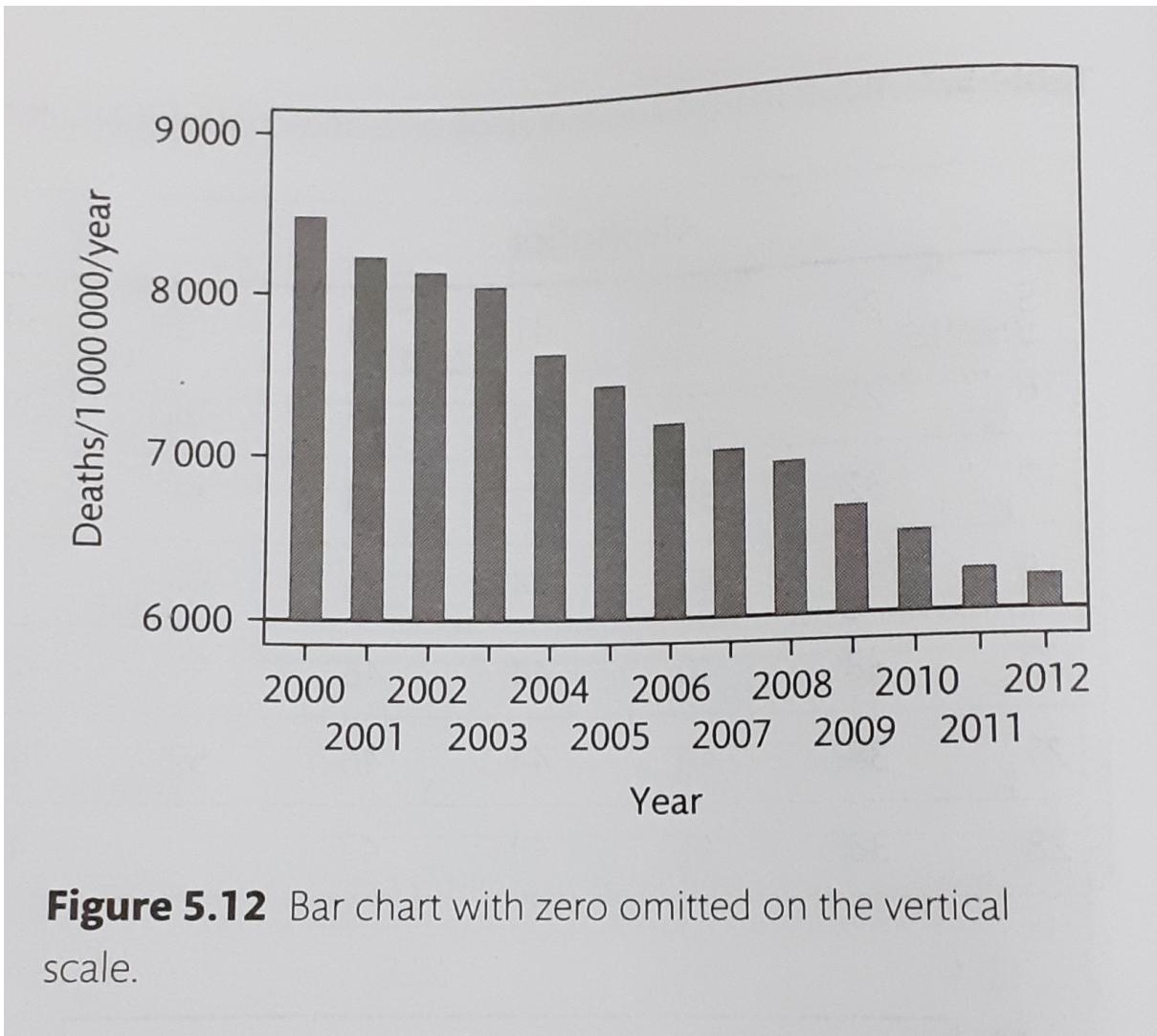
- A table is intended to communicate information, so it should **be easy to read and understand.**
- Table should have clear title, stating clearly and unambiguously what the table represents. Rows and columns must be labelled clearly.
- Sometimes in research paper we choose to represent results using table but in conference for audience we may use graphics/figures.
- Why? In a conference , the most important thing is to convey an outline of the analysis quickly.

**Table 2.10** Assessment of radiological appearance at six months as compared with appearance on admission (data from MRC 1948)

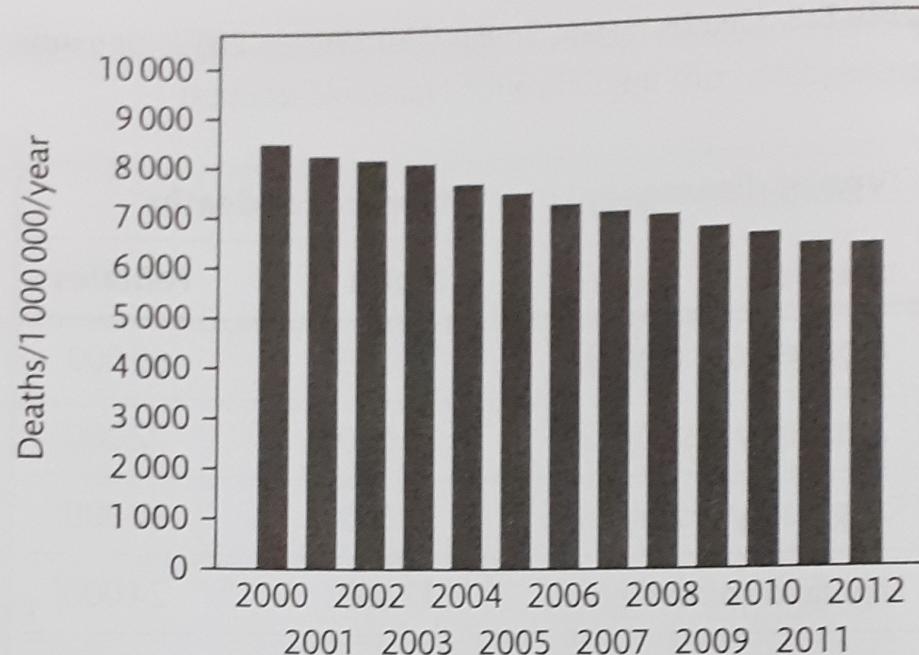
Radiological assessment	S Group	C Group
Considerable improvement	28 51%	4 8%
Moderate or slight improvement	10 18%	13 25%
No material change	2 4%	3 6%
Moderate or slight deterioration	5 9%	12 23%
Considerable deterioration	6 11%	6 11%
Deaths	4 7%	14 27%
<b>Total</b>	<b>55 100%</b>	<b>52 100%</b>

# Misleading graphs

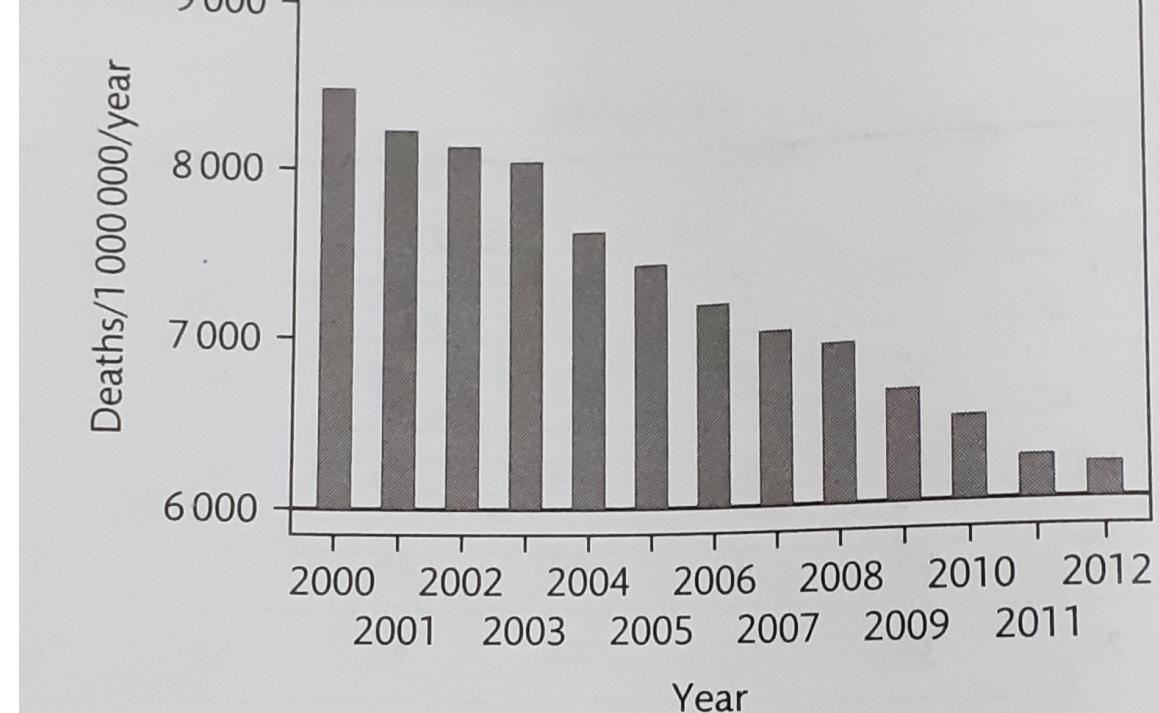




**Figure 5.12** Bar chart with zero omitted on the vertical scale.

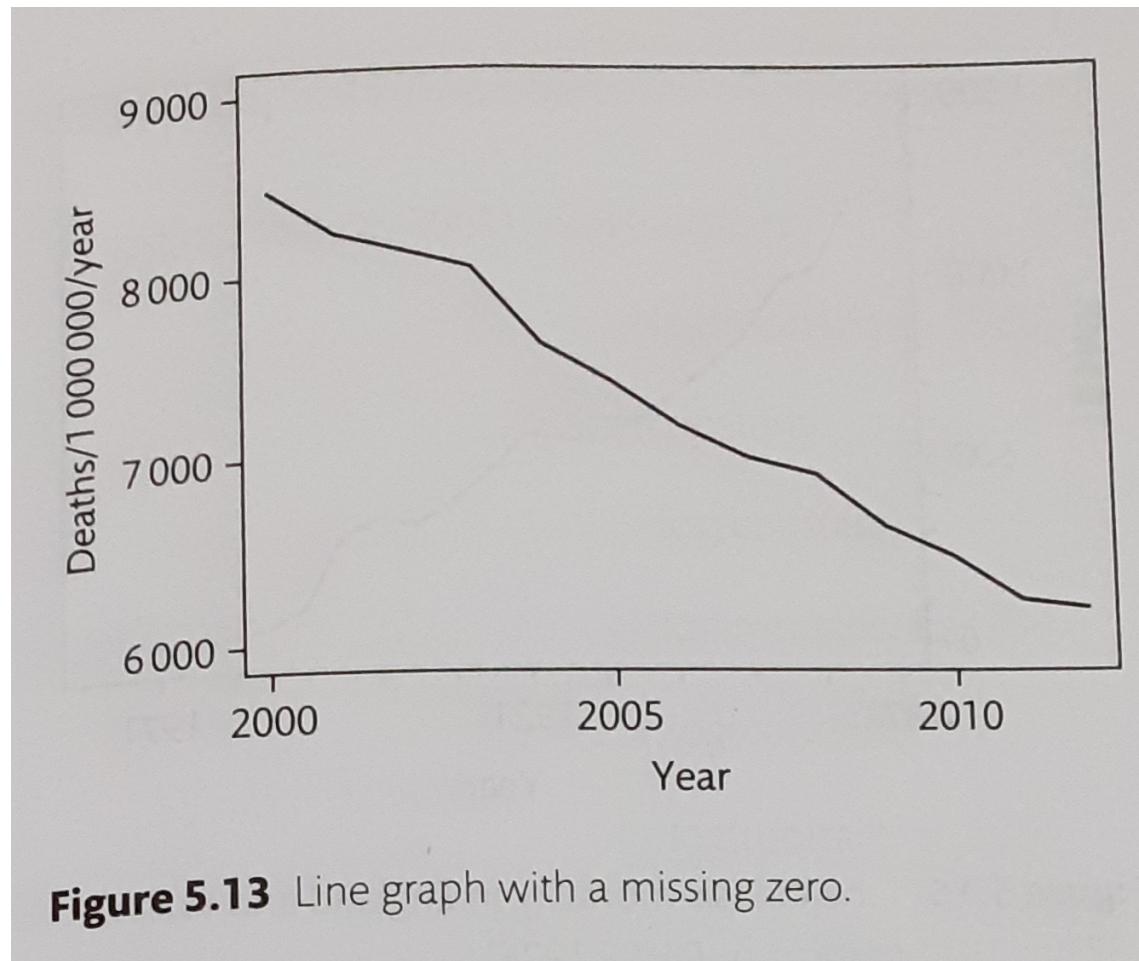


**Figure 5.2** Bar chart showing the relationship between mortality caused by cancer and year, males, England and Wales, 2000–2012 (data from the Office for National Statistics).

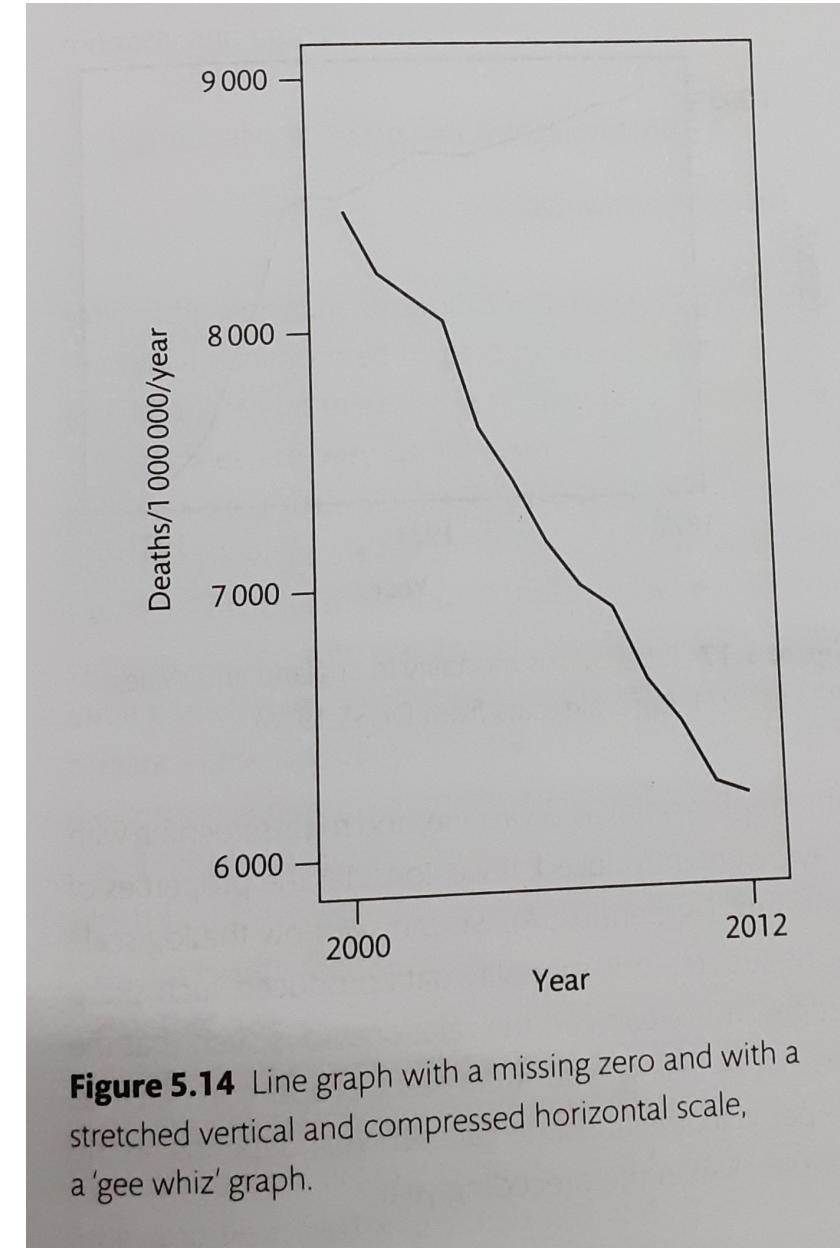


**Figure 5.12** Bar chart with zero omitted on the vertical scale.

Fig. 5.12 omits most of the vertical scale and instead stretches the small part of the scale where the change takes place.



**Figure 5.13** Line graph with a missing zero.



**Figure 5.14** Line graph with a missing zero and with a stretched vertical and compressed horizontal scale, a 'gee whiz' graph.

# Some important points

- Line graphs are particularly at risk of undergoing the distortion of a missing zero.
- In previous figures, we are certainly interested in zero mortality, it is surely what we are aiming for.
- The point is that graphs can so easily mislead the unwary reader so let the reader beware.

Questions?

Next class: Plagiarism, Academic Integrity,  
Predatory Publishing.