

# How to display data badly

---

Karl W Broman

Biostatistics & Medical Informatics  
University of Wisconsin – Madison

<http://www.biostat.wisc.edu/~kbroman>

# Using Microsoft Excel to obscure your data and annoy your readers

---

Karl W Broman

Biostatistics & Medical Informatics  
University of Wisconsin – Madison

<http://www.biostat.wisc.edu/~kbroman>

# Inspiration

This lecture was inspired by

H Wainer (1984) How to display data badly. American Statistician  
38(2):137-147

Dr. Wainer was the first to elucidate the principles of the bad display of data.

The now widespread use of Microsoft Excel has resulted in remarkable advances in the field.

3

# General principles

The aim of good data graphics:

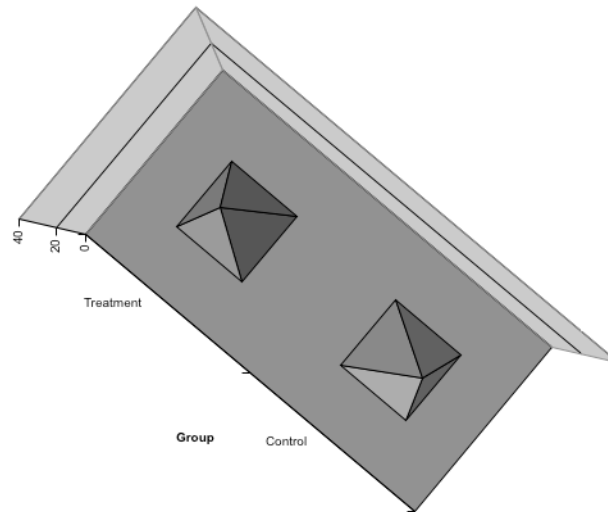
Display data accurately and clearly.

Some rules for displaying data badly:

- Display as little information as possible.
- Obscure what you do show (with chart junk).
- Use pseudo-3d and color gratuitously.
- Make a pie chart (preferably in color and 3d).
- Use a poorly chosen scale.
- Ignore sig figs.

4

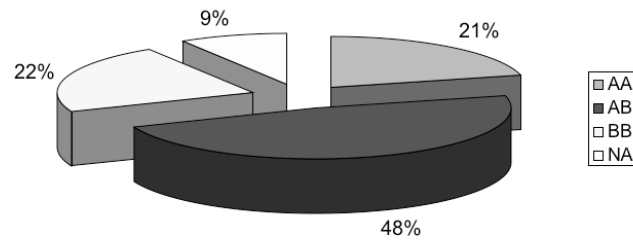
# Example 1



5

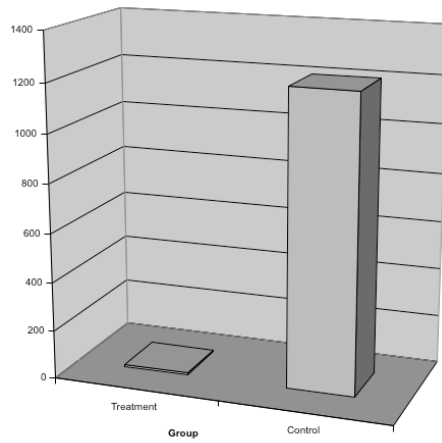
# Example 2

Distribution of genotypes



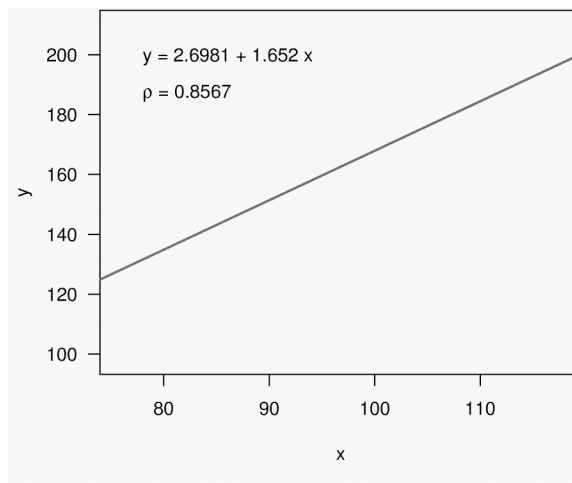
6

### Example 3



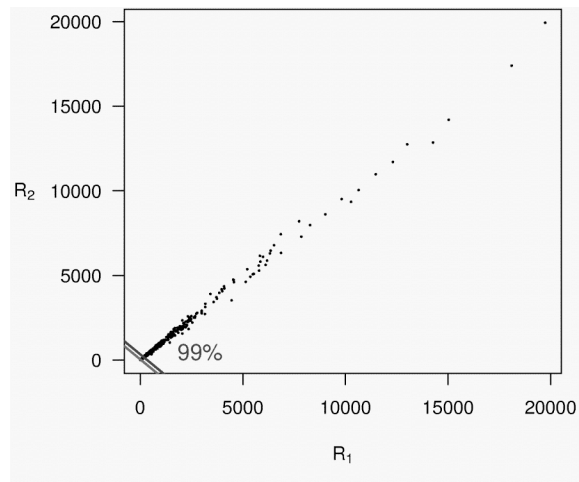
7

### Example 4



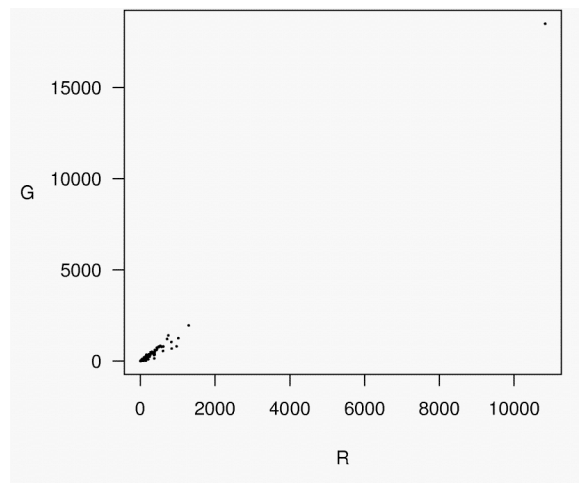
8

## Example 5



9

## Example 6



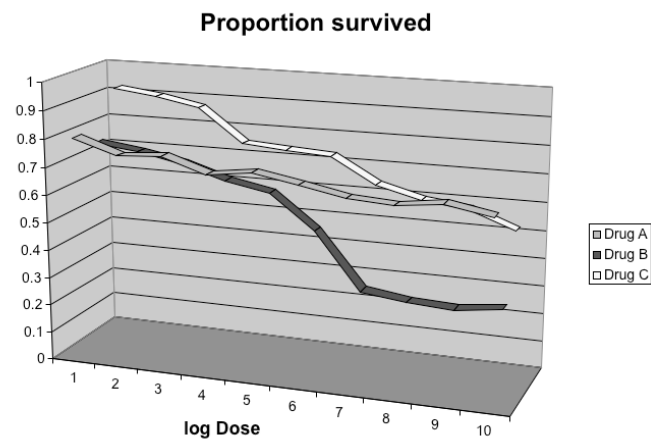
10

## Example 7

$N$	$b/c = 10.0$		$b/c = 10.0$		$b/c = 100.0$	
	$r^*$	$G$	$r^*$	$G$	$r^*$	$G$
3	2	0.2	2	2.225	2	22.47499
4	2	0.26333	2	2.88833	2	29.13832
5	2	0.32333	3	3.54167	3	35.79166
6	3	0.38267	3	4.23767	3	42.78764
7	3	0.446	3	4.901	3	49.45097
8	3	0.50743	4	5.5765	4	56.33005
9	3	0.56743	4	6.26025	4	63.20129
10	4	0.62948	4	6.92358	4	69.86462

11

## Example 8



12

## Displaying data well

- Be accurate and clear.
- Let the data speak.
  - Show as much information as possible, taking care not to obscure the message.
- Science not sales.
  - Avoid unnecessary frills — esp. gratuitous 3d.
- In tables, every digit should be meaningful. Don't drop ending 0's.

13

## Further reading

- ER Tufte (1983) The visual display of quantitative information. Graphics Press.
- ER Tufte (1990) Envisioning information. Graphics Press.
- ER Tufte (1997) Visual explanations. Graphics Press.
- WS Cleveland (1993) Visualizing data. Hobart Press.
- WS Cleveland (1994) The elements of graphing data. CRC Press.

14

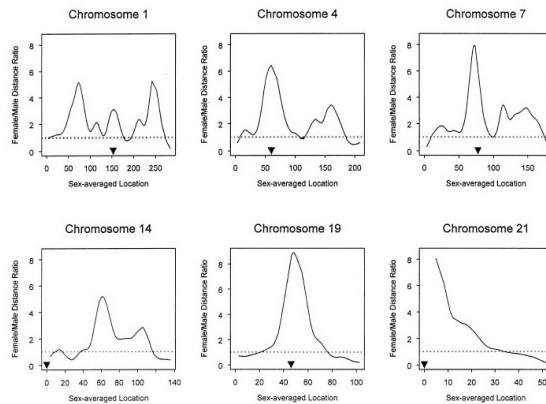
# The top ten worst graphs

With apologies to the authors, we provide the following list of the top ten worst graphs in the scientific literature.

As these examples indicate, good scientists can make mistakes.

15

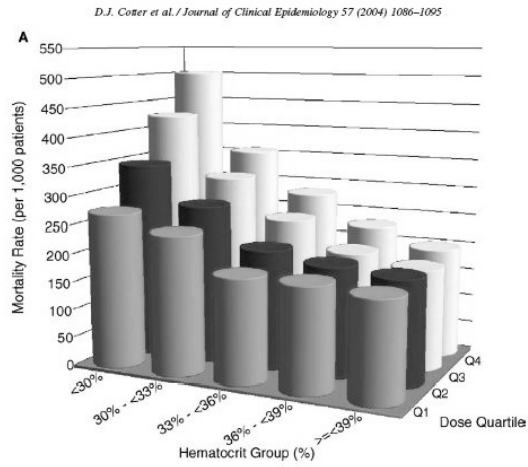
10



Broman et al., Am J Hum Genet 63:861-869, 1998, Fig. 1

16

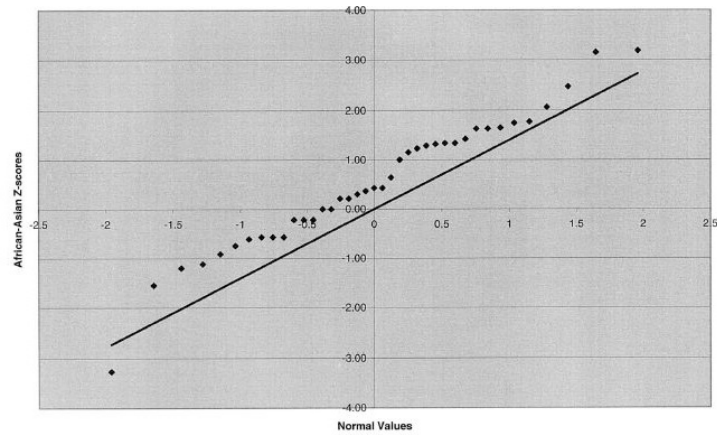
9



Cotter et al., J Clin Epidemiol 57:1086-1095, 2004, Fig 2

17

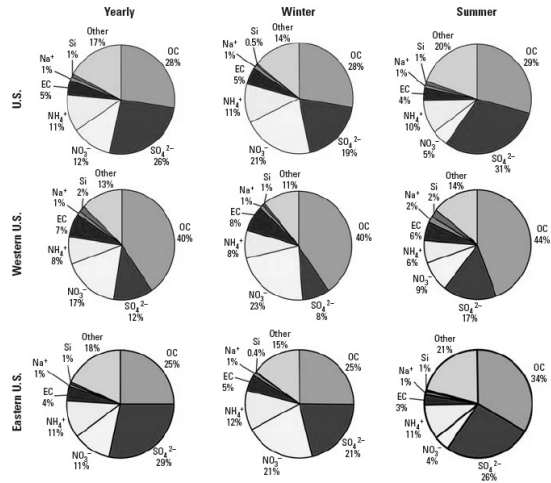
8



Jorgenson et al., Am J Hum Genet 76:276-290, 2005, Fig 2

18

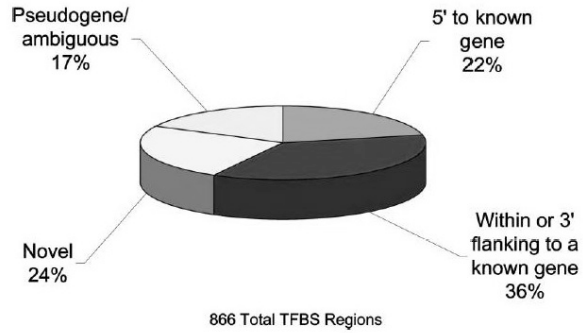
# 7



Bell et al., Env Health Persp 115:989-995, 2007, Fig 3

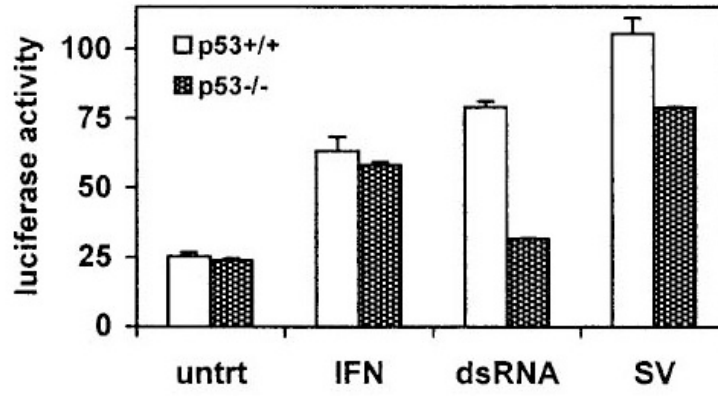
# 6

## Distribution of All TFBS Regions



Cawley et al., Cell 116:499-509, 2004, Fig 1

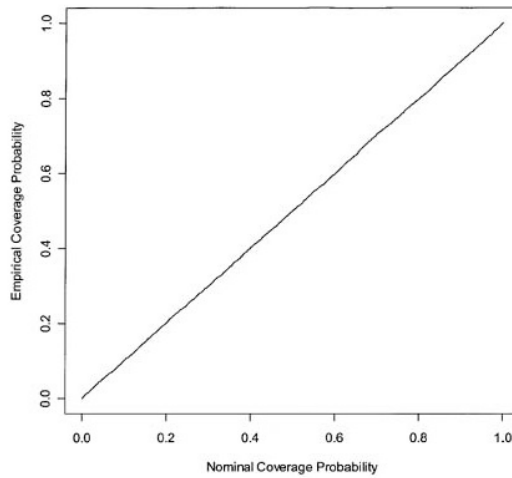
5



Hummer et al., J Virol 75:7774-7777, 2001, Fig 4

21

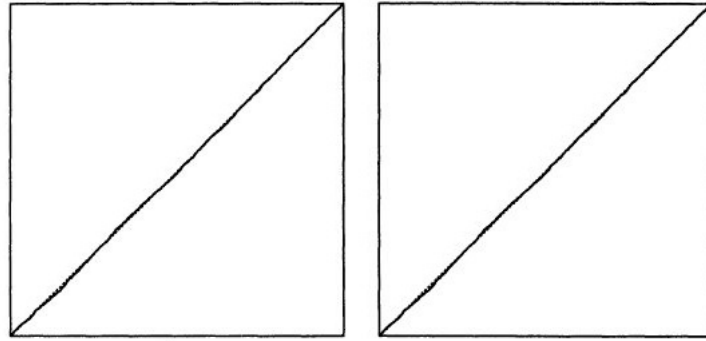
4



Epstein and Satten, Am J Hum Genet 73:1316-1329, 2003, Fig 1

22

3



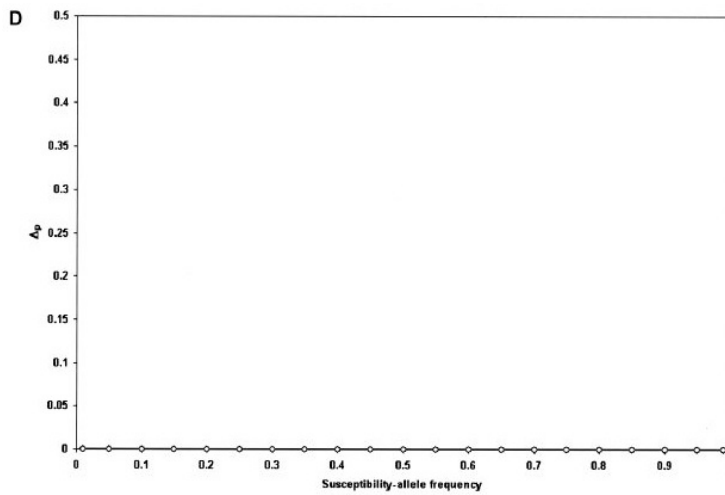
(a)

(b)

Mykland et al., J Am Stat Asso 90:233-241, 1995, Fig 1

23

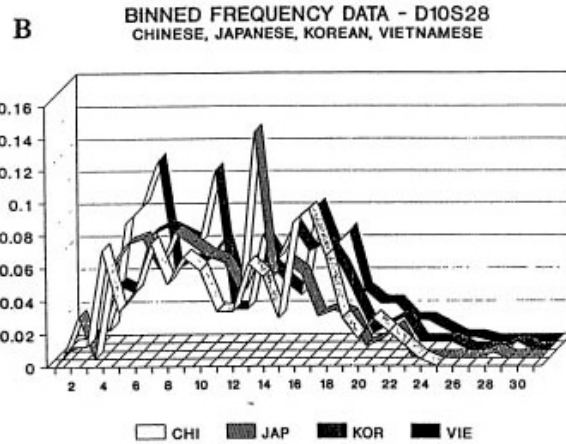
2



Witke-Thompson et al., Am J Hum Genet 76:967-986, Fig 1

24

1



Roeder, Stat Sci 9:222-278, 1994, Fig 4

25

## The new iWork '08.

Streamlined word processing, stunning presentations and now spreadsheets. Work wonders, without the work.

**Pages '08**  
Writing comes naturally when you're using Pages '08, the streamlined word processor for the Mac.  
[Learn More](#)

**Keynote '08**  
Create absolutely stunning, cinema-quality presentations more easily than ever before.  
[Learn More](#)

**Numbers '08**  
Introducing Numbers, the familiar yet revolutionary new spreadsheet application.  
[Learn More](#)

26