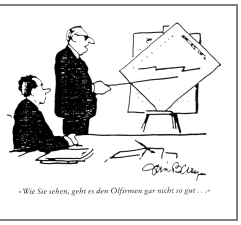
How to lie with statistics

EST III, 17.12.2001

I. Introduction

Why do people fake statistics?

Where are statistics (mis)used?	**Why do people fake statistics?**	
• engineering sciences • medicine & social sciences • advertisement • politics & economy	• because of their incompetence • intentionally: – to persuade or shock others – to influence people	

Basic steps in using statistics

Defining the problem	Collecting the data	Analysing the data	Reporting the results

II. How to talk back to a statistic

5 basic questions to detect errors and manipulations in statistical information:

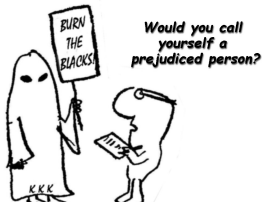
1. Who says so?

• An interested person?	• selection of favourable data, suppression of unfavourable • selection of favourable methods
• O.k. names?	• Dr. title or name of a scientific laboratory ⇒ people have more confidence in the repo

2. How does he know?

• Survey	• is the sample big enough? • a real random sample?	• leading questions? • did people answer honestly?
• Methods of data analyse	• correlation and causality confused?	• are the methods really applicable to this problem?

Example: ***Should anyone have lice?***
People on the New Hebrides observed that ill people did not have lice.
⇒ conclusion: "lice make a man healthy" (correlation → causality)
This is false! Explanation: having lice is normal, and when people get a fever, the body temperature rises and the lice leave this "uncomfortable" person.



3. What's missing?

• Used definitions	• are they documented?
• Used methods	• average meaningless without standard deviatic • sample size big enough?

Example: ***On the average…***
In mathematics there are many definitions for average values; in statistics, the following three are used above all: mean (arithmetic mean), median (central value), mode (value occurring the most frequently)

4. Has the subject been changed?

• Changed definitions	• shifting base of percentage values…
• "Firsters"	• synthetic superlatives
• Is there a relation between raw figures and conclusions?	• Unfortunately, there is not always a relation!
• Correlation & causality confused?	

Example: *The shifting base*
A price decreases 30% in one year and then increases 30% in the following year.
People might think that over all the price has not changed, but:
basic price: £100 → first year (-30%): £70 → second year (+30%): £91 !!

5. Does it make sense?

• Lack of thinking?	• occurs quite often!
• Are long-term trends reliable?	
• Are precise numbers really correct? • Estimation of errors?	• Precise numbers given are often not as precise! (Example: Adding numbers with different accuracies)

Example: *In 1997, researchers found that the wolf has been domesticated much earlier than believed. They gave their estimation: 135,000 years ago ± 300% !!*
(Source: SCIENCE Magazine, Vol. 276, 13 June 1997 pages 1647-1648)

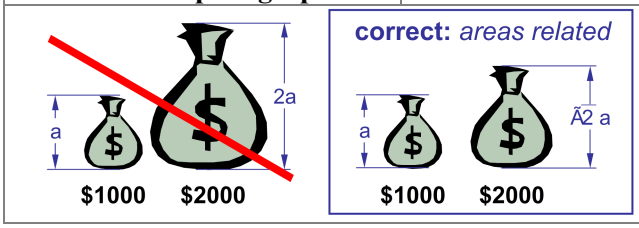
III. Graphical presentation

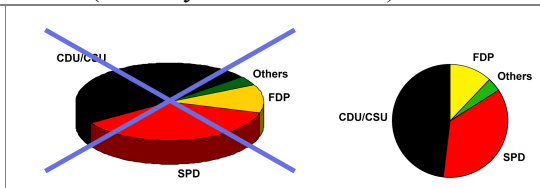
Types of diagrams

curve, bar chart, pie chart, pictograph, scatterplot, flow chart, map …

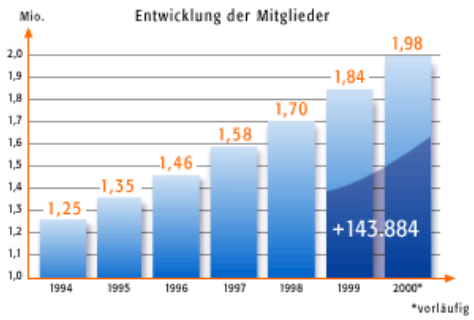
Optical illusions

• **False proportions**	• distorted axis (missing zero, missing values, perspective, …)
• **Faked pictographs**	• areas must be related (not only one dimension!)

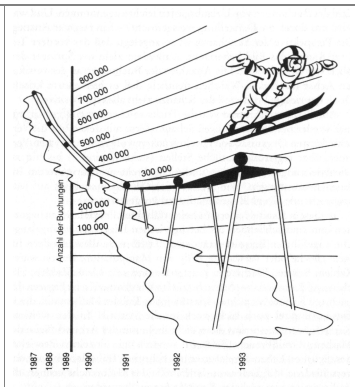




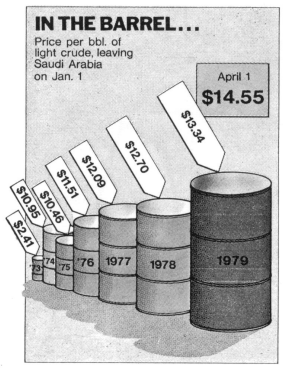
Examples



missing zero



distorted axis



false proportions (lie factor 10!)

IV. Conclusion

Statistics are often misused (by lack of knowledge or even intentionally), but not all statistics are faked!

Bibliography

Almer, Ennis C.: *Statistical Tricks and Traps: An Illustrated Guide to the Misuses of Statistics*. Los Angeles: Pyrczak Publ., 2000.

Dewdney, Alexander K.: *200 Prozent von nichts: die geheimen Tricks der Statistik und andere Schwindeleien mit Zahlen*. Basel: Birkhäuser, 1994.

Huff, Darrell: *How to Lie with Statistics*. New York: W. W. Norton, 1982.

Krämer, Walter: *So lügt man mit Statistik*. Frankfurt/Main: Campus Verlag, 1991.

Krämer, Walter: *So überzeugt man mit Statistik*. Frankfurt/Main: Campus Verlag, 1994.

Krämer, Walter: *Statistik verstehen: eine Gebrauchsanweisung*. 4. Aufl., Frankfurt/Main: Campus Verlag, 1999.

Paulos, John Allen: *INNUMERACY: Mathematical Illiteracy and Its Consequences*. New York: Hill and Wang, 1988.

Spirer, Herbert F. / Spirer, Louise / Jaffe, A.J.: *Misused Statistics: Second Edition, Revised and Expanded*. New York: Dekker, 1997.

Tufte, Edward R.: *The Visual Display of Quantitative Information*. Ceshire: Graphics Press, 1983.

Zelazny, Gene: *Wie aus Zahlen Bilder werden: Wirtschaftsdaten überzeugend präsentiert*. Wiesbaden: Gabler, 1986.