



HOW TO LIE WITH STATISTICS

There are three kinds of lies:
"Lies, damned lies and statistics."

Benjamin Disraeli (1804 - 1881), writer and politician, Prime Minister of England in 1868 and from 1874 to 1880, i.e. when under Queen Victoria British colonialism was in full swing.



The following examples are either out of Darrel Huff's "How to Lie with Statistics" (Pelican, 1972) or Walter Krämer's "So lügt man mit Statistik" (Campus Verlag, 1992).

1. THE LYING RESPONDENTS

A house-to-house survey purporting to study magazine readership was once made in which a key question was: What magazines does your household read? When the results were tabulated and analysed it appeared that a great many people loved "Harper's" (zweitälteste wiss.-kult. Zeitschrift der USA, Anm. d. Verf.) and not very many read "True Story" (US-amerikanisches Magazin mit "wahren Geschichten" vor allem romantischen Inhalts, Anm. d. Verf.). Now there were publishers' figures around at the time that showed very clearly that "True Story" had a circulation more in the millions in contrast to that of "Harper's" in the hundred thousands. Perhaps we asked the wrong kind of people, the designers of the survey said to themselves. But no, the questions had been asked in all sorts of neighbourhoods all around the country. The only reasonable conclusion then was that a good many of the respondents, as people are called when they answer such questions, had not told the truth. The only thing the survey had uncovered was snobbery.

In the end it was found that if you wanted to know what certain people read it was no use asking them. You could learn a good deal more by going to their houses and saying you wanted to buy old magazines and seeing what they had. Then all you had to do was to count the "Yale Review" and the "Love Romances" Even that dubious method, of course, does not tell you what people read, only what they have been exposed to.

Similarly, the next time you learn from reading that the average American brushes his teeth 1.02 times a day ask yourself a question. How can anyone have found out such a thing? Is a woman who has read in countless advertisements that no-brushers are social offenders going to confess to a stranger that she does not brush her teeth regularly? The statistic may have meaning to one who wants to know only what people say about tooth-brushing but it does not tell us a great deal about the frequency with which bristle is applied to incisor.

Reason 1 why to mistrust any statistics:

Respondents lie. They don't tell you what they think is the truth, they tell you what they think you want to hear.

2. THE WRONG FIGURES

What is safer to travel with, a plane or a train? Killer number one, the car, is left aside for the moment.

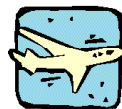
There are two answers, a rational one and an intuitional one.

Reason tells us that flying is safer. The same newspapers and magazines that report all catastrophes and all accidents in its entirety, and in colour in order to bump up their sales figures keep repeating that statistics show that flying is safer than taking the train. On average, they say, less people die in plane crashes than in train accidents. But how is this average evaluated?

Normally the number of casualties is divided by the number of all kilometres covered. Thus they get:



0.9 people killed per
1 billion passenger·km



0.3 people killed per
1 billion passenger·km

So three times as many people die in train accidents than do in plane accidents! Then why do we break out in cold sweat when we enter a plane but don't when we hop on a train? Our intuition tells us that the train is safer! And our intuition is not always wrong.

Instead of dividing the number of people travelling by the number of kilometres covered it could just as well be divided by the number of hours spent in danger. To me this figure seems to be even more relevant than the number of kilometres covered. I am not afraid of going to bed each evening even if I know that 99 % of all people die there. Thus the averages will be:



0.07 people killed per
1 billion passenger·hour



0.24 people killed per
1 billion passenger·hour

So the former advantage of flying has been inverted. More than three times as many people die in plane crashes than in train accidents per hour spent in the corresponding vehicle!

Reason 2 why to mistrust any statistics:

Not always are the figures used the important ones.

3. THE PRE-SELECTED SAMPLE

A psychiatrist wrote in an article that in his opinion the whole world has gone crazy. When asked how he arrived at this belief he said: "Look at all the people in my practice!"

Nobody, of course, is stupid enough to apply results from such a small survey to the whole of the population. Or is there somebody?

Such distorted sample surveys cause more prejudices and false reports than anything else in the field of statistics.

At the ETH Zurich the postgraduates from China are very much favoured by the professors because they work hard, never complain, are very clever and always finish their job. The Swiss postgraduates are lazier, less obedient and less smart and more likely to fail. The figures alone could lead to the conjecture that Chinese people are just cleverer than the Swiss.

It may be like that, but it is not supported by this example. Only the very best out of millions of Chinese students are sent abroad to do their postgraduate and then it is absolutely normal that they outperform the more average Swiss students. Most probably there are as many idiots and geniuses in China as there are anywhere else.

Every now and then, it can be read that the figures prove beyond every doubt that a birth at home is much safer than a birth in a hospital. Even if the figures are correct, again the sample is pre-selected. Whenever any difficulty should occur a pregnant woman will not want to take the risk and stay at home. So all critical cases end up in hospital, while only the unproblematic ones stay at home. Thus, it does not come as a surprise that home births appear to be safer.

Another survey result can be found in the "Times" (30 March 1990) with the message that 60 % of all civil aviation pilots die before the age of 65! The chairman of the pilot's trade union is said to be deeply concerned and that the "International Federation of Airline Pilots Association" is planning a thorough survey of 70,000 pilots concerning their eating and drinking habits and their exposure to stress and sexual tension to reveal the reason for the mysterious passing away of the pilots. Instead of an expensive survey the chairman should have given a thought to how the survey in the "Times" had been created.

The "Times" writes that the figures have come from pension funds and life insurance companies in Great Britain, South America and Canada. Even if the figures were correct and representative for all pilots in the world one might ask what exactly is meant by "60 % of all civil aviation pilots die before the age of 65". Most probably it just says that of all active and retired pilots who have died in the last year 60 % were younger than 65. But then the mystery is solved! Civil aviation has expanded enormously during the last three decades, so a big part of all pilots today is quite young. In twenty years the situation will have changed completely, but today only very few are veterans and older than 65. So if a pilot dies today it is very likely that he is younger than 65.

Reason 3 why to mistrust any statistics:

It is hard to find a survey sample that is not pre-selected.

4. THE MANIPULATED DIAGRAM

The eye is a wonderful organ that catches loads of information very quickly. In addition, the brain immediately decides what information is important enough to become conscious and what is not. It even tries to put everything into a context and tends to fix missing parts of information by making it up. All this comes in very handy to those people who want us to see things their way. It opens the gate to all kinds of cheatings.

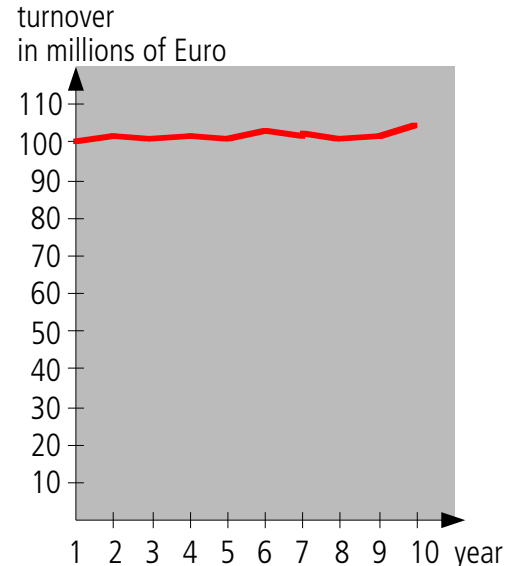
Let's take a company that is preparing its 10th anniversary with a publication that should attract new shareholders. Unfortunately business was bad, not one year with profit, no market share gain, no increase in investment or number of employees, only the turnover improved a little during the last ten years. Of course, these are exactly the figures that are taken to represent the company's success.

In millions of Euro:

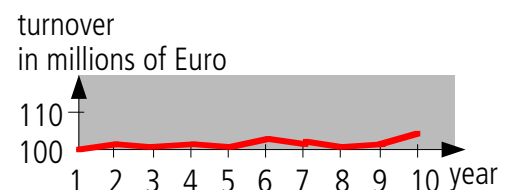
year	1	2	3	4	5	6	7	8	9	10
turnover	100	101	100.5	102	101.5	103	102.5	101.5	103	105

How can these figures be turned into an impressive graph?

The graph on the right shows the sad truth: stagnation rather than thriving dynamics. There is no reason to get carried away with enthusiasm. That is why no such graphs ever show up.

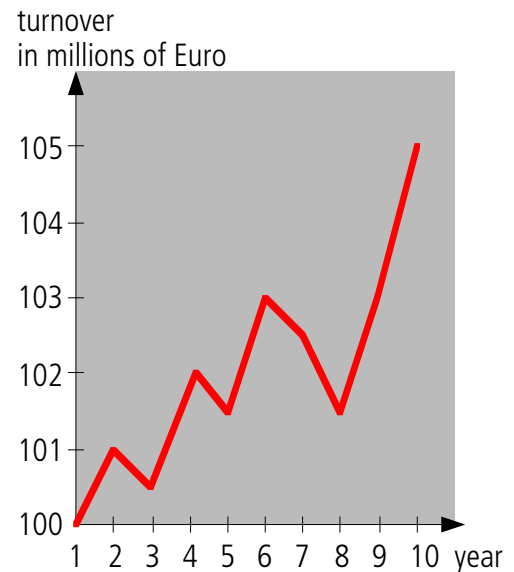


To save space it is common practice to cut off the lower part of the coordinate system. That is all right as long as the inscribing of the y-axis makes clear that this has been done.

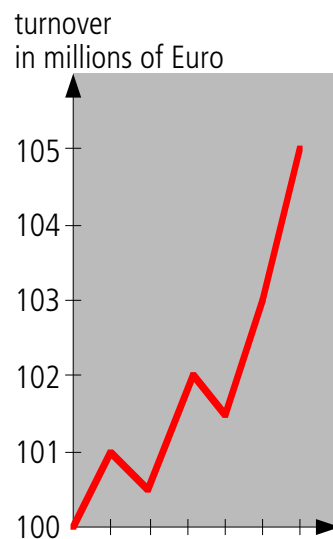


Now the graph is stretched in the direction of the y-axis. This is common practice again and acceptable as long as it is visible from the inscription on the y-axis.

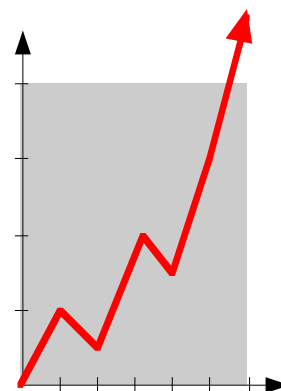
The increase in the turnover has become much more dynamic by now. Only the setback between year 7 and 9 spoil the good impression.



The impression can be jazzed up simply by omitting the corresponding years. To hide the manipulation the label on the x-axis is dropped. That is plain cheating now, but the effect is enormous.



To improve the already good appearance the label on the y-axis is dropped, too. To eradicate every doubt about the development of the company the upper bound of the diagram is lowered and a big arrow is added that shoots through the ceiling up into the sky.



Reason 4 why to mistrust any statistics:

Graphs with missing labels on the axis are very likely to be a big con.

- 1

The "New York Times" reports (5 December 1978) that the average life expectancy of orchestra conductors is 73.4 years, four years more than that of the entire male population at that time. The article ends with the suggestion to become an orchestra conductor in order to live longer.
Assuming the figures are correct, how can that be?

- 2

Next time you read a newspaper check the omnipresent graphs. What kind of manipulations can you detect?