


Statistics & graphs send a very powerful message to people.

Graphs have the capability to strengthen suggestions about the data based on the type of graph, the colours used and other tools.

Just because you see a graph does not mean you should believe it.

Examine carefully where the data came from and what it is telling you.




Sources of Bias

Bias is any influence that unreasonably affects or sways the results of a sample survey or census.

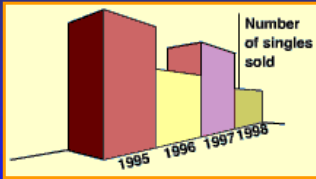
This includes bad & small samples, loaded questions, non-response bias, interviewer bias, order of questions, processing errors etc.

e.g. A survey on the effects of passive smoking, sponsored by a major tobacco manufacturer, is hardly likely to be impartial.



Misleading Graphs

Graphs can be misleading in a number of ways
- scale, labels, size or missing data etc.



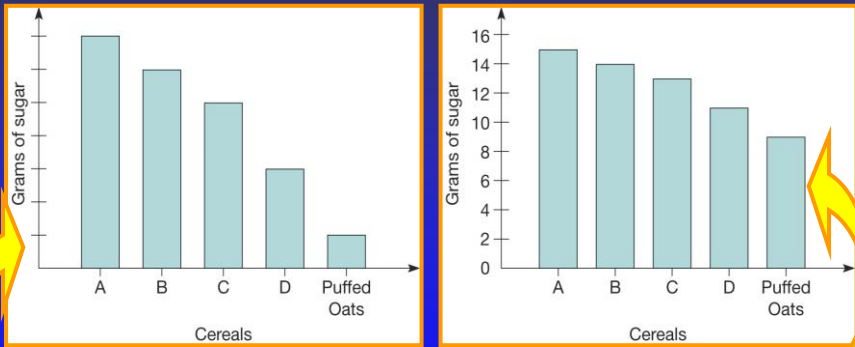
Year	Number of singles sold
1995	100
1996	25
1997	25
1998	10

This 3D bar chart looks attractive, but it is very misleading. Why?

There is no scale on the vertical axis. This perspective makes it look as though the sales for 1995 were far more than those for any other year. In fact they were identical to those for 1997.

What is this graph below trying to tell us?

Puffed oats has almost no sugar compared to other cereals.




Cereal	Grams of sugar (Left Graph)	Grams of sugar (Right Graph)
A	~15	15
B	~12	14
C	~9	13
D	~5	11
Puffed Oats	~2	9

Why is the graph a poor graph?


What do you now notice?

No scale on axis

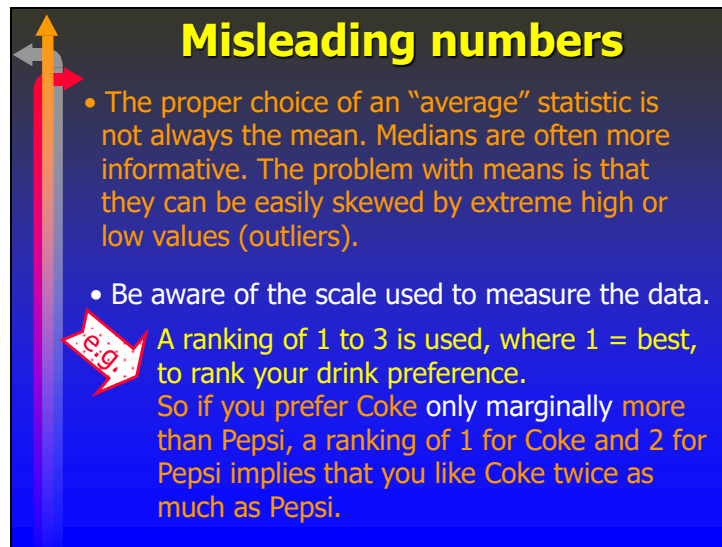
Although less, it still has a lot of sugar.



The purpose of using graphs is usually to compare trends or general magnitudes rather than provide precise data points. Use a table to show precise data points.




So to correctly interpret a graph, we should analyze the **numerical** information given in the graph instead of being misled by its general shape.



Misleading numbers


- The proper choice of an "average" statistic is not always the mean. Medians are often more informative. The problem with means is that they can be easily skewed by extreme high or low values (outliers).
- Be aware of the scale used to measure the data.

 A ranking of 1 to 3 is used, where 1 = best, to rank your drink preference. So if you prefer Coke only marginally more than Pepsi, a ranking of 1 for Coke and 2 for Pepsi implies that you like Coke twice as much as Pepsi.

• A company newsletter states that the annual bonus for a full-time employee of the Company is R31,417. Use the data in the table below to determine if this is an accurate statement.

Annual Income (R)			
25,000	18,000	32,000	28,000
22,000	19,000	34,000	92,000
29,000	31,000	21,000	26,000

The news article used the **mean** of the bonuses for the "average" annual bonus. All but 3 of the bonuses are below this value. The bonus of R92,000 has **distorted** the value of the mean. A better measure for the annual bonus would be the **median**, R27,000.




Source: The Mercury, 10 November, 1997

Di: 98% say it was a plot

British Sunday newspaper "The People" revealed yesterday that nearly everyone who responded to a survey organised by the paper believe Diana, Princess of Wales, was murdered. It says more than 98 per cent of 5,600 people who rang a hotline answered "yes" to the question: "Were Diana and Dodi killed as part of a secret operation?" Only 93 said they believed the crash was an accident.

1. Why would you question the results of this survey?



**T
a
s
k**

Find three newspaper or magazine articles that contain incorrect or misleading analysis or presentation of the statistics.

For each article, hand in a copy of the article and a short critique of the incorrect or misleading analysis or presentation of the statistics in the article.

