

EDUCATION

There is more to data and numbers than meets the eye

BIPRA PRASUN DAS

How often do we pause to think about a statistic on a report or a figure someone mentions during a debate on the talk shows that we regularly consume?

Seventy percent of people support this or one in five people suffer from that, these numbers shape how we think about the world, and help us form opinions. Nonetheless, how often do we ask ourselves: *Who collected this data? How was it measured? What's missing from the picture?*

We, generally, trust numbers and charts, assuming that data must be the result of rigorous analysis, however, numbers carry an element of subjectivity due to the techniques used to collect, frame, and display data, often revealing biases.

How data is presented visually can play a huge role in determining the narrative the data intends to convey. One of the simplest yet most effective tricks is the manipulation of graph axes. When a chart's vertical axis does not start at zero, small changes can be made to look dramatic. Data can also lie by omission. For instance, presenting a smaller time frame, for a statistic of long-term decline, can make it seem like the phenomena is a short-term surge.

Similarly, political campaigns often highlight month-to-month swings in approval ratings or economic

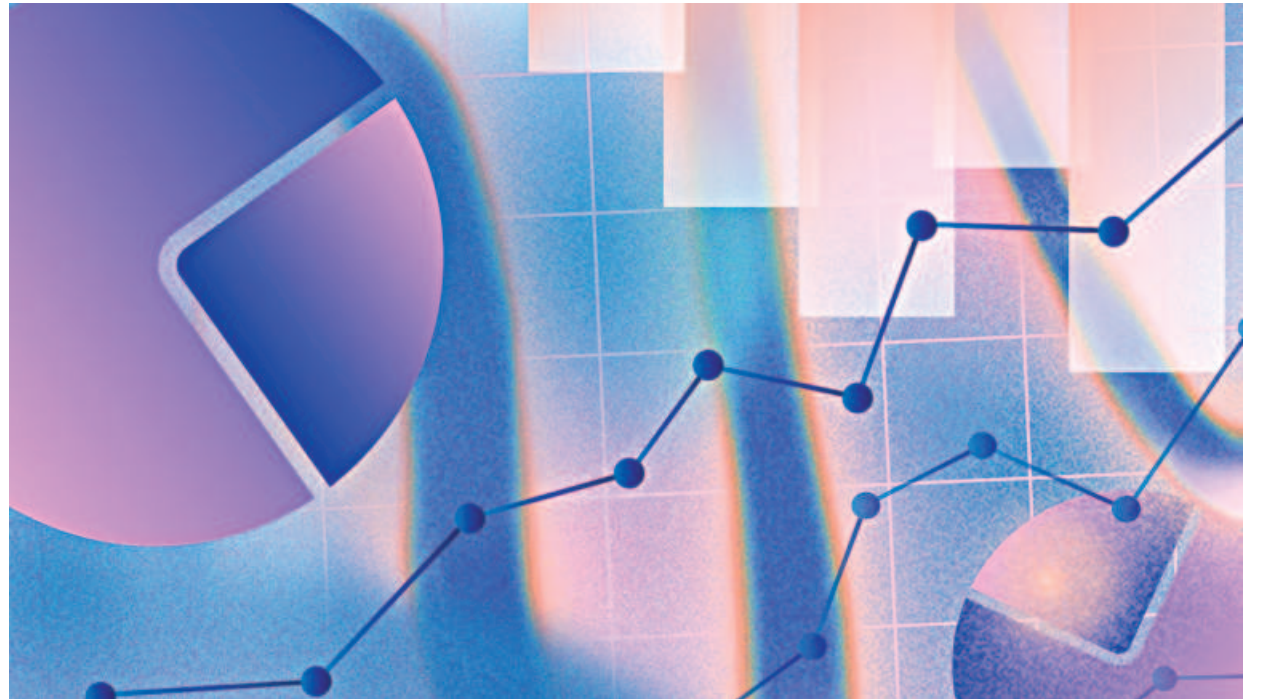


ILLUSTRATION: ABIR HOSSAIN

In a world where data is everywhere, on the television, in our feeds, in the mouths of politicians and pundits, it is easy to mistake numbers for truth. However, statistics are not just neutral facts; they're shaped by how they're gathered, framed, and shared.

indicators without acknowledging that those fluctuations fall within normal statistical noise. By choosing the "right" start and end points, a relatively flat trend can be recast as evidence of sudden success or failure. In addition, usage of an inappropriate graph type can be deceiving. A notorious case is drawing a line chart (which implies continuous data) for unrelated categories, connecting points that shouldn't be connected. For qualitative categories (like regions or sectors), a bar chart is usually correct. A stacked pie chart or area chart with too many slices can also exaggerate differences, since human eyes struggle to comprehend the true scale of the data with slices alone.

The process of collecting data itself can have flaws in them too. For example, surveys can be misleading if their design is not transparent or inclusive. The sample of a survey is important. If it over-represents certain groups (like young, urban, tech-savvy users in online polls) and under-represents others (like rural or older populations), the results will be skewed. Privilege bias also plays a role: those with more time, education, or internet access are more

likely to participate, meaning marginalised voices often go unheard. Even the way questions are worded in a survey can subtly skew the results in the favour of a particular outcome.

During the Covid-19 pandemic, misleading graphs and statistics flooded social media, some of which was fuelled by a mix of public anxiety and a desperate search for answers. Platforms like Facebook, which rewards content based on likes, shares, and comments, were a part of the problem.

Research from Yale School of Management has identified how social media platforms inadvertently encourage the spread of statistical misinformation through their reward systems. This research also shows that frequent users, driven by habit rather than critical thinking, shared false and true headlines at almost the same rate. This shows the issue isn't just individual gullibility, but a deeper flaw in how distribution channels such as social media are designed. Posts about unproven remedies went viral not because they were accurate, but because they were engaging. In dire situations, the compounding effects of fear and flawed tech design become a perfect storm for spreading statistical misinformation.

Bangladesh itself has seen several troubling cases where official statistics were manipulated for political gain. Reports revealed that during the Sheikh Hasina regime, key economic data were distorted. One major example came from the Bangladesh Bureau of Statistics (BBS), which reported that unemployment fell from 2.61 million in 2022 to 2.43 million in 2023, which was the lowest rate since 2002. This was surprising because the country's economy was clearly struggling at the time, with falling industrial output, imports, exports, and remittances. Economists later discovered that the BBS had changed how they measured unemployment. Instead of counting people who lost their jobs during the Covid-19 pandemic as unemployed, they were labelled "out of the labour market."

So, how can we protect ourselves from being misled by statistics? It starts with asking the right questions. Checking

for the source is a good place to start. Reliable statistics should come with clear citations or links to the original data. If a claim doesn't say where the numbers came from, that is a red flag. Next, when there's a chart, it is useful to examine them carefully. Ask yourself: do the axes start at zero? Are the time periods complete and consistent?

A bar chart that starts mid-way up the scale or skips years can completely change the story. It also helps to understand whether or not the variables are correlated. Just because two lines on a graph rise together doesn't mean one caused the other. And when it comes to polls, it is important to ask: *Who was surveyed? How many people? Was it a random sample or just a group of internet users?* Without taking these factors into account, numbers can easily be manipulated.

In a world where data is everywhere, on the television, in our feeds, in the mouths of politicians and pundits, it is easy to mistake numbers for truth. However, statistics are not just neutral facts; they're shaped by how they're gathered, framed, and shared. This doesn't mean we should stop trusting data altogether, rather it shows how important it is to engage with data thoughtfully. With the current saturation of information, we only stand a better chance of understanding it by slowing down, asking questions, and critically engaging with the information presented to us.

References:

1. The Daily Star (September 2, 2024). *Statistics given by Hasina regime need urgent correction.*
2. Yale Insights (March 31, 2023). *How Social Media Rewards Misinformation.*
3. The Financial Express (December 7, 2024). *Data distortion should be a thing of past.*
4. Statistics How To (n.d.). *Misleading Graphs: Real Life Examples.*
5. Rho, Y., & Zhu, J. (2024). *Various Misleading Visual Features in Misleading Graphs: Do they truly deceive us?*

Bipra Prasun Das studies Computer Science and Engineering at North South University.