LUISS

Degree Program in Economics and Business

Course of Statistics

Manipulation of Statistical Data in Media Campaigns

Prof. Marco Perone Pacifico

SUPERVISOR

Marco Astrologo - 284411

CANDIDATE

Academic Year 2024/2025

Abstract

In our increasingly digitized society, statistics are now a part of decision-making in all aspects of society. From economic policy-making to guiding personal consumer choices, statistical data profoundly influences the manner in which we perceive and interact with the world around us. But the easy accessibility of information has also made manipulation of it astonishingly prevalent, leading to disinformation that can distort public opinion and mislead important decisions. This thesis responds to the subtle dynamic between statistical data and society and how it is so central to daily life.

In Part One, the thesis begins by examining the existing problems of gathering good quality statistical data on the Internet and moving on to discuss how the public struggle to interpret and meaningfully interact with statistical data. It goes on to analyze shared tactics used to deceive or mislead individuals, pointing out the devious way in which numbers could be manipulated for the sake of making misleading arguments.

Part Two turns attention to real-world cases, illustrating the manipulation of statistics used in advertising to influence consumer behavior and in politics to shape public opinion. These cases enlighten as much about the strategies as about the consequences of statistical deception, informing more of its potential effect on society.

Finally, Part Three discusses the responses to the problem of misinformation. This section examines the ethical responsibilities of scientists and public institutions, means of carrying out public education to make individuals data literate, and new challenges brought about by artificial intelligence and procedurally generated content. In grappling with these issues, the thesis seeks to offer a balanced perspective on how society can be protected from misinformation in the age of data-informed decision-making.

Table of Contents

Introduction

If one were to engage in a thought exercise and arbitrarily select the pillars upon which modern society rests, most responses would include science, technology, politics, and commerce. The utilization of statistical data plays a fundamental role in each of these domains: from evaluating the impact of an experimental medical treatment, through analytic engagement data collected by social networks, to analyzing independent reports that influence specific economic policies, or leveraging data to market a new product launch. Therefore, it is of crucial importance, as a society, to ensure that the use of statistical data across all fields is conducted in good faith, consistently, and, most importantly, in a scientifically accepted manner recognized by the international community.

This is not always the case. Firstly, it can be genuinely challenging, even when one intends to, to locate reliable and independently verified sources. Secondly, even when such sources are found, they are often aimed at a limited audience of experts rather than the general population, making consultation and comprehension particularly difficult for the average individual. Furthermore, there is the issue of the incorrect utilization of statistical data itself: through various techniques that will be later detailed, it is alarmingly easy to decontextualize data, misinterpret correlations as causal relationships, or downplay the effects of certain agents.

One might argue that everyone shares some responsibility: communicators exploit short-form content and the limited attention spans of new generations, shaped by social media, by presenting isolated, out-of-context data; the general public struggles to keep pace with scientific literature, search engines, and tools to access quality data; and finally, experts rarely recognize the responsibilities they hold and rarely take steps to effectively meet the communicative needs of their audiences.

How can we educate the public to critical thinking when looking at numbers? In what way the scientific community can be more responsible towards their audience? What are the next challenges we are going to phase in this new era of technology?

Part I – Our world in data: The importance of statistics in modern society

I. A NEEDLE IN A HAYSTACK: INTERNET AND RELIABLE DATA

Whether it's healthcare, economy, social media, or scientific breakthroughs, numbers are at the heart of how we interpret and engage with our reality. Yet, as digital content has exploded in the last few decades, the availability of clear, accessible, and reliable statistical information hasn't kept pace. Even though there's more data than ever before, figuring out where to find trustworthy statistics, and how to understand and trust them, has only gotten more challenging. Technical barriers, confusing designs, misinformation, and a general lack of awareness all combine to make the process more frustrating and complex than it might seem at first.

The current information landscape is characterized by constant exposure to news, commentary, and data from a wide range of sources including social media, websites, television, and podcasts. This overabundance can lead to what is often referred to as "news fatigue" or "data fatigue," where individuals feel overwhelmed and disengaged. In fact, when people are exposed to too much information, they could potentially ignore statistics completely or rely on simplified narratives that confirm their pre-existing views. This environment makes it challenging for users to identify and engage with reliable statistical content.

Moreover, the visibility of content online is shaped more by engagement than by accuracy. Social media platforms are designed to amplify information that attracts attention, regardless of its quality. As a result, misleading or decontextualized statistical claims can quickly go viral. For example, in 2023, a certain crime-related statistic was widely circulating online. Although the numbers used were technically correct, they were framed misleadingly to support a distorted narrative.[1] These kinds of distortions reduce public trust in statistics and contribute to the overall confusion surrounding data in digital spaces.

Even when one would want to genuinely inform themselves, most general users turn to search generalized search engines like Google when looking for statistics, but these kinds of portals are optimized for web pages, not for datasets: users may receive blog posts, news articles, or irrelevant results instead of direct access to raw data. Even Google's dedicated tool, *Dataset Search*, despite being a significant step in the right direction, it has shown to have its good amount of flaws. A recent

[1] Sarah E. Norman, *How Data Literacy Can Keep America Safe*, TIME, June 29, 2023, https://time.com/6290961/data-literacy-statistics-

study found that users struggled with both searching for and interpreting the datasets available on the platform.[2] This shows that search tools remain insufficiently adapted to the needs of those looking for reliable, structured data.

In addition to navigational difficulties, many people are simply unaware of where credible data can be found. National statistical agencies, such as the *UK's Office for National Statistics* (ONS), publish large volumes of high-quality public data. Yet a 2021 study by *NatCen Social Research* found that while 75% of UK adults had heard of the ONS, only 36% had ever knowingly used its data.[3] This awareness gap significantly limits the public's engagement with trustworthy sources and emphasizes the need for broader data education.

Even when individuals manage to find official statistics, they often encounter further obstacles. For instance, many government and institutional data portals are not designed with non-expert users in mind. They frequently require technical knowledge, use scientific labels, or lack intuitive interfaces. A comparative study of 41 national open data portals found that most of them failed to meet the needs of diverse user groups, particularly general users who lack specialized training.[4] This usability gap discourages engagement and renders even open-access data effectively inaccessible to a large segment of the population.

Another common barrier is the paywall. Much of the most accurate and in-depth statistical research is locked behind academic journals or commercial data providers. These paywalls prevent independent researchers and the general public from accessing high-quality information unless they are affiliated with an institution or can afford subscription fees. As a result, people often resort to secondary summaries or simplified media interpretations, sources that may distort or selectively report the original data. This contributes to a two-tiered system of knowledge access and deepens the divide between those who can critically engage with statistics and those who cannot.

All these factors contributes to the fact that many datasets published by public institutions or researchers remain largely unused, a phenomenon sometimes described as the presence of "data graveyards." These are datasets that are technically accessible but are buried in poorly organized portals, lack clear indexing, or are not promoted effectively. According to *Open Data Watch*, the gap between data availability and actual use is significant, and addressing it requires improving both the

---

[2] K. Sostek, D. M. Russell, et al., *Discovering Datasets on the Web Scale: Challenges and Recommendations for Google Dataset Search*, Harvard Data Science Review (Special Issue), 2024, https://hdsr.mitpress.mit.edu/pub/dataset-search/release/1.

[3] S. Butt et al., *Public Confidence in Official Statistics 2021*, NatCen Social Research (UK), 2022, https://uksa.statisticsauthority.gov.uk/publication/public-confidence-in-official-statistics-2021/.

[4] A. Nikiforova and K. McBride, *Open Government Data Portal Usability: A User-Centred Usability Analysis of 41 Open Government Data Portals*, Telematics and Informatics 58 (2021): 101539, https://www.sciencedirect.com/science/article/pii/S0736585321000542.

structure and visibility of data sources.[5] Without effective tools and strategies for discovery, the potential value of statistical data remains untapped.

Finally, the increasing use of generative artificial intelligence tools such as language models adds another layer of complexity. While these tools can assist with summarizing or explaining data, they are also known to fabricate statistics or cite non-existent sources. For instance, research from the *University of Oxford* has focused on developing techniques to predict when language models are likely to "hallucinate" or produce false information.[6] As these systems become more integrated into daily information-seeking behavior, users must remain aware of their limitations, particularly when dealing with statistical or scientific content. Blind reliance on AI-generated information, without cross-checking sources, risks introducing even more confusion into an already crowded and chaotic data landscape.

II. LOST IN THE NUMBERS: THE PUBLIC'S STRUGGLE WITH STATISTICAL DATA

In the digital age, many individuals turn to social media and opinion-based blogs for information, often at the expense of more reliable sources. This trend is particularly concerning as social media platforms amplify misinformation and reinforce echo chambers. The rapid spread of information on these platforms, coupled with algorithms designed to maximize engagement over accuracy, facilitates the spread of misleading or false statistical content. One factor contributing to this dynamic is the "easiness effect", a cognitive bias where simplified presentations of complex data, often seen in social media posts, give individuals the illusion of understanding and reduce their perceived need for expert input.[7]

Distinguishing between real and misleading statistics becomes increasingly difficult in this context. Many people lack a basic understanding of statistical concepts such as margins of error, statistical significance, and data provenance. This knowledge gap can significantly affect their ability to interpret data accurately. The *U.S. Census Bureau* emphasizes the importance of these concepts,

---

[5] Open Data Watch, *Overcoming Data Graveyards in Official Statistics*, 2022, https://opendatawatch.com/publications/overcoming-data-graveyards-in-official-statistics/.
[6] University of Oxford, *Major Research on Hallucinating Generative Models Advances Reliability of Artificial Intelligence*, June 20, 2024, https://www.ox.ac.uk/news/2024-06-20-major-research-hallucinating-generative-models-advances-reliability-artificial.
[7] Lisa Scharrer et al., *When Science Becomes Too Easy: Science Popularization Inclines Laypeople to Underrate Their Dependence on Experts*, Public Understanding of Science 26, no. 8 (2016): 1003–1018, https://journals.sagepub.com/doi/abs/10.1177/0963662516646047.

noting that a failure to understand them can result in misguided conclusions or decisions based on misinterpreted results.[8]

Even when data is technically accurate, the conclusions drawn from it may be incorrect. A common example is the confusion between correlation and causation, where a statistical relationship is wrongly assumed to imply a causal link. This misinterpretation can have serious implications, particularly in public health or policy settings. As discussed in *The New Atlantis*, confusing correlation with causation can lead to ineffective responses and incorrect reasoning.[9] The complexity of statistical language aggravates the issue: one study found that 70% of postgraduate students struggle to understand statistical content in academic literature, comparing the challenge to learning a new language.[10] This difficulty extends to the general public, making it harder to engage critically with numerical information.

Moreover, even reliable data can be misused. Statistics may be selectively presented or framed in ways that support a specific argument while omitting important contextual information. During the COVID-19 pandemic, for instance, some groups used legitimate data visualizations to advance unscientific views, often by stripping the numbers of proper interpretation.[11] This shows how the presentation of data can shape public perception, sometimes misleadingly, even when the numbers themselves are accurate.

Finally, a particularly insidious trap involves the blend of true data with false claims. By anchoring misleading narratives in real statistics, communicators can create an illusion of credibility that is difficult to challenge. A 2023 study published in *Nature* found that when individuals used search engines to fact-check deceptive news articles, their likelihood of believing the misinformation actually increased.[12] This paradox reveals how difficult it is to disentangle truth from distortion when falsehoods are embedded in factual content.

---

[8] U.S. Census Bureau, *Understanding and Using American Community Survey Data: What All Data Users Need to Know*, U.S. Department of Commerce, 2018, https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch07.pdf.

[9] The New Atlantis, *Correlation, Causation, and Confusion*, n.d., https://www.thenewatlantis.com/publications/correlation-causation-and-confusion.

[10] Dani A. and E. Al Quraan, "Investigating Research Students' Perceptions About Statistics and Its Impact on Their Choice of Research Approach," Heliyon 9, no. 10 (September 29, 2023): e20423, https://doi.org/10.1016/j.heliyon.2023.e20423

[11] Matt Motta, Dominik Stecula, and Christina Farhart, *How Visualizing Data Legitimizes Unorthodox Science*, arXiv, 2021, https://arxiv.org/

[12] Gordon Pennycook and David G. Rand, *How Online Search Can Increase Belief in Misinformation*, Nature 621 (2023): 822–827, https://www.nature.com/articles/s41586-023-06883-y.

III. FIGURES DON'T LIE, BUT LIARS FIGURE: COMMON STRATEGIES FOR MISLEADING WITH STATISTICS

While many members of the public struggle to fully understand statistical data, this widespread lack of literacy also creates opportunities for deliberate manipulation. Data is often viewed by people as being objective and factual, yet how numbers are being chosen, defined, and presented can actually misguide public opinion. Technical accuracy in many cases may be used to mislead or to build appealing but partial narratives. Awareness of these tactics is necessary to critically approach statistical arguments, especially in a world where data-driven arguments are heavily utilized in media, advertisement, and politics. This section examines some of the most popular techniques utilized to manipulate statistical perceptions apart from straight-up falsification.

Cherry-picking, or the fallacy of incomplete evidence, refers to the practice of selectively presenting only the data that supports one's argument while ignoring information that might contradict it.[13] The term "cherry-picking" originates from the act of harvesting only the healthiest fruits, therefore creating, for an outside observer, the false impression that the entire tree's yield is of similarly high quality. This selective sampling misleads by presenting a partial view as if it were representative of the whole, thus distorting the true condition of the broader reality.[14] In communication, the selected facts may individually be accurate but are stripped of the broader context necessary to form a balanced understanding. By emphasizing favorable statistics and omitting contradictory evidence, one could construct persuasive yet fundamentally misleading narratives.

For example, a political figure might highlight a decline in unemployment in a specific city while omitting the fact that national unemployment rates are simultaneously rising. This technique exploits the public's trust in numbers, encouraging the mistaken belief that isolated examples reflect general trends. Cherry-picking is widely recognized as a propaganda technique in political discourse and is often associated with a broader strategy known as card stacking. This method manipulates audience perception by emphasizing one side of an issue while repressing opposing views.[15] As noted by communication scholars, card stacking is "commonly used in persuasive speeches by political candidates to discredit their opponents and to make themselves seem more worthy".[16] The term "stacking the deck" itself derives from a magician's trick, where a seemingly shuffled deck of cards

---

[13] 12. "Cherry Picking in Data Analytics: How to Avoid Common Pitfalls," Institute of Data, accessed April 2025, https://www.institutedata.com/blog/cherry-picking-in-data-analytics/.
[14] "Cherry picking," Wikipedia, last modified July 2023, https://en.wikipedia.org/wiki/Cherry_picking.
[15] Institute for Propaganda Analysis, *The Fine Art of Propaganda: A Study of Father Coughlin's Speeches* (New York: Harcourt Brace and Company, 1939), 95–101.
[16] John C. S. Kim, *The Art of Creative Critical Thinking* (Lanham, MD: University Press of America, 1993), 317–318.

is in fact arranged in a predetermined order, allowing the magician to control the outcome while maintaining the illusion of randomness.[17]

Another subtle but highly effective method of manipulating perceptions involves framing changes using percentages (relative terms) instead of absolute values, or vice versa, to alter the perceived significance of an effect. Percentages often make small changes seem dramatic, while absolute numbers can downplay major shifts. For example, if an experimental drug reduces illness from 2 cases per 100 people to 1 case per 100, this can be described as either an “absolute risk reduction of 1 percentage point” (from 2% to 1%) or a “50% reduction in risk.” Both statements are factually correct, yet the relative figure sounds far more impressive despite the minor absolute improvement.[18] Similarly, in a heart attack prevention study, if 1 out of 100 patients had a heart attack on a drug compared to 2 out of 100 without it, marketers might advertise a "50% reduction in heart attacks," without mentioning the absolute risk dropped only by 1%.

Without context, relative framing like these can exaggerate significance and mislead audiences about real-world effects.[19] Politicians and advertisers frequently exploit this framing. A political advertisement might declare, “Violent crime doubled under Opponent X,” citing an increase from 50 incidents to 100. While a 100% rise sounds alarming, the actual numbers might still represent a minor overall rate. On the other hand, problems can be minimized by shifting to absolute figures, such as reframing "only 0.1% of votes were fraudulent" to "thousands of fraudulent votes" to provoke alarm. In marketing, relative percentages are common in product claims: "Customers achieved 3x whiter teeth" or "Battery life increased by 50%," without clarifying whether the improvement is truly meaningful. A "50% longer battery life" might simply mean an increase from 1 hour to 1.5 hours, hardly revolutionary in practice.

Best practice in scientific and ethical communication requires presenting both relative and absolute changes to give a complete picture. When only one format is provided, critical readers should question what might be omitted, and whether the framing is designed more to persuade than to inform.

Surveys and opinion polls are prime tools for shaping political narratives and marketing strategies, but they are also highly susceptible to manipulation. Bias can be introduced at several stages: through the phrasing of questions, the construction of the sample, and the presentation of the results. A common tactic is the use of loaded or leading questions, which nudge respondents toward a particular answer. For instance, contrasting “Do you support our nation's attempt to bring democracy and freedom to other parts of the world?” with “Do you support an unprovoked military incursion by

---

[17]Hyman Ruchlis and Sandra Oddo, *Clear Thinking: A Practical Introduction* (Buffalo, NY: Prometheus Books, 1990), 195–196.

[18] Gareth Grant, "Relative Measures of Effects Can Be Misleading," Students 4 Best Evidence (Cochrane), 16 febbraio 2018, https://s4be.cochrane.org/blog/2018/02/16/relative-measures-effects-can-misleading/

[19] Gerd Gigerenzer, Risk Savvy: How to Make Good Decisions (New York: Viking, 2014)

our country?" demonstrates how framing alone can significantly alter responses.[20] Marketing surveys frequently exploit similar techniques, using emotionally charged phrasing like "How much do you love our new feature?" rather than neutrally asking "How do you feel about our new feature?"[21] Even the order of questions can create bias, with earlier questions influencing responses to later ones, a fact often overlooked by casual readers.[22]

Sampling bias is another crucial vulnerability. A survey's validity depends on how representative its sample is, yet achieving true randomness is very difficult in practice. As *Spirer, Spirer, and Jaffe* note, the concept of a "simple random sample" often breaks down under real-world conditions, with many surveys unintentionally excluding key demographics.[23] Historical failures like the 1936 *Literary Digest* poll, which predicted the wrong outcome of the U.S. presidential election due to a biased sample, illustrate how devastating sampling errors can be. Even today, reliance on outdated methods, such as random house's calls, continues to distort findings by systematically underrepresenting younger or mobile populations. In addition, self-selection bias and non-response bias further weaken survey validity when respondents are not randomly drawn or when participation is voluntary.

Finally, reporting bias occurs after data collection. Organizations often selectively publicize favorable results while ignoring less convenient ones. Politicians may highlight a single poll showing a lead, omitting others that contradict it. Companies may claim "90% customer satisfaction" without disclosing that only a small, selective group of customers responded. As *Weatherburn* highlights in his analysis of crime statistics reporting, selective data presentation, omission of context, and sensationalized headlines are pervasive techniques that undermine public trust.[24] These forms of manipulation, through wording, sampling, and selective reporting, significantly erode the reliability of survey-based statistics, while continuing to fuel sensational media narratives and misleading public perceptions.

Data can often be even more persuasive than words in shaping public opinion, making deceptive graphs a preferred tool in politics, marketing, and media. Although a graph may technically present accurate data, its visual design can subtly but powerfully mislead viewers. Common manipulations include truncated axes, distorted scales, inappropriate graph types, and visual embellishments that distort perception. Unlike written text, visual data carries an "illusion of

---

[20] Gardenier, John, and David Resnik. "The Misuse of Statistics: Concepts, Tools, and a Research Agenda." Accountability in Research 9, no. 2 (2002): 65–74

[21] Fischer, David. Historians' Fallacies: Toward a Logic of Historical Thought. New York: Harper & Row, 1979, 337–338

[22] Kahneman, Daniel. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux, 2013, 102

[23] Spirer, Herbert F., Louise Spirer, and Jon Jaffe. Misuse of Statistics: Concepts, Tools, and a Research Agenda. New York: Marcel Dekker, 1998, chapters 3 and 4

[24] Weatherburn, Don. "Uses and Abuses of Crime Statistics." Crime and Justice Bulletin no. 153 (2011)

objectivity" and typically avoids critical scrutiny. As *Darrell Huff* noted, deceptive graphs “have no qualifying words or adverbs to spoil the show.”[25]

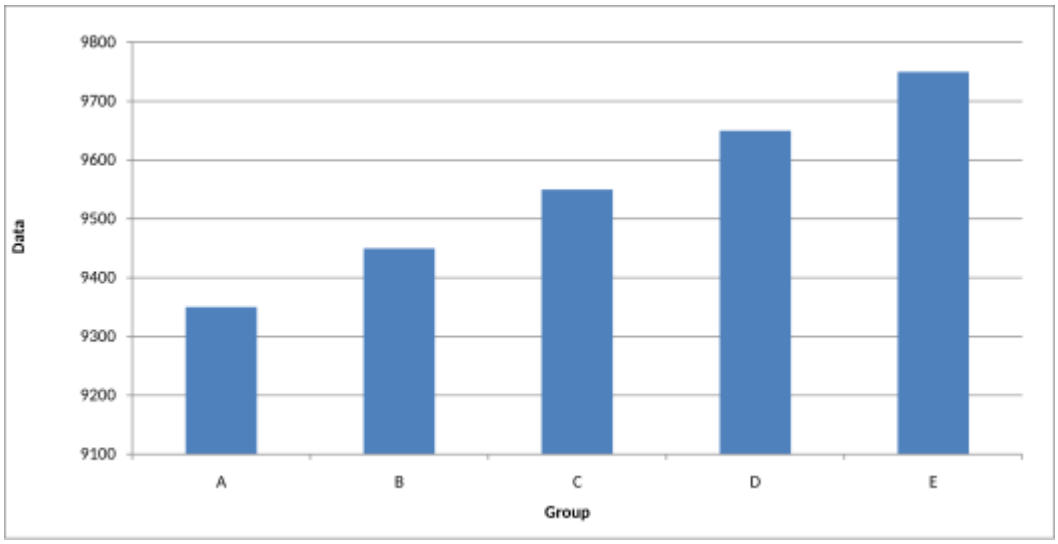
Marketers often exploit scale manipulation by using an overly broad range to downplay declines or narrowing the scale to exaggerate minor gains.[26] Politicians, too, have been known to manipulate graphs, such as starting a y-axis at 90 instead of 0 to amplify small fluctuations in unemployment or budget statistics.[27] Three-dimensional graphics and skewed perspectives, particularly in advertising pie charts, can further distort relative proportions, leading viewers to misinterpret which category is most significant.[28] A study from the *Texas State Auditor’s Office* found that even when viewers are informed about truncated axes, they still consistently overestimate differences, illustrating the deep cognitive impact of visual distortions.[29]

Other deceptive practices include omitting axis labels or scales, using inconsistent interval markings, and selectively framing time periods in line graphs to favor trends. Misleading visualizations have been criticized for decades, with scholars like Edward Tufte emphasizing the ethical obligation of graph designers to prioritize clarity and honesty over persuasive distortion.[30] Misleading graphics are so prevalent that they have been a recurring feature in false advertising cases, such as in promotions for weight-loss supplements where dramatic slopes are presented as factual results, despite being fictional illustrations.

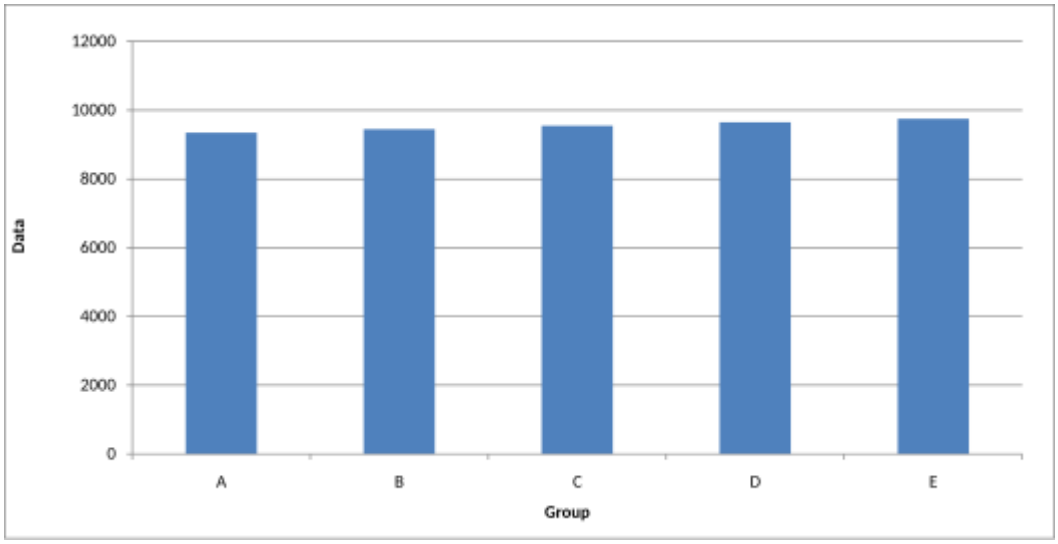
Ultimately, it is the responsibility of the viewer to scrutinize axes, scales, and context, recognizing when a graph’s design serves to manipulate perception rather than accurately represent reality.

---

[25] Darrell Huff, *How to Lie with Statistics* (New York: W.W. Norton & Company, 1954), 63
[26] John Burn-Murdoch, "Why You Should Never Trust a Data Visualisation," The Guardian, July 24, 2013, https://www.theguardian.com/news/datablog/2013/jul/24/why-you-should-never-trust-a-data-visualisation
[27] Kirk, Andy. Data Visualisation: A Handbook for Data Driven Design (London: SAGE, 2016), 52
[28] Michael Siegrist, "The Use or Misuse of Three-Dimensional Graphs to Represent Lower-Dimensional Data," Behaviour & Information Technology 15, no. 2 (1996): 96–100
[29] Texas State Auditor's Office, "Methodology Manual: Data Analysis: Displaying Data - Deception with Graphs," January 4, 1996, https://web.archive.org/web/20030402194810/http://www.sao.state.tx.us/methmanual/da9.html
[30] Edward R. Tufte, The Visual Display of Quantitative Information (Cheshire, CT: Graphics Press, 2006), 178



*Example of a **truncated bar chart** (y-axis does not start at zero). The values for groups A–E range roughly 9100 to 9800, but by starting the vertical axis at 9100, the differences are magnified visually. Such a graph creates the impression of significant change where the actual differences are minor.*[31]



*Same data plotted with a **full y-axis (starting at zero)**. Here the bars for groups A–E appear nearly equal, correctly showing that the differences are very small. The comparison between the two figures highlights how truncating the axis inflates the perceived variation.*[32]

These graphs display identical data; however, in the first truncated bar graph, the data appear to show significant differences, whereas, in the regular bar graph, these differences are hardly visible.[33]

---

[31] "Misleading Graph," Wikipedia, last modified April 26, 2025, https://en.wikipedia.org/wiki/Misleading_graph.
[32] Ibid.
[33] Ibid.

Another frequent manipulation is selectively using a statistical indicator that does not represent the typical situation. "Average" can denote disparate measurements, most typically the mean (arithmetic average) or the median (middle value of a distribution). Under a skewed distribution, with a minority of outlier cases significantly deviating from the majority, the mean is pulled towards upwardly or downwardly by the outliers, while the median gives a better sense of the "typical" instance. Manipulators will purposely select the indicator that more favorably supports their case.

For example, a government may pride itself, "The average (mean) yearly income of our nation is $15,000," suggesting overall prosperity, while the opposition responds with, "But half of our people have a yearly income of under $10,000," putting the emphasis instead on the median.[34] Both can be factually true: a handful of millionaires can set the mean up significantly even if the median standard of living is modest. This old standby, serves to illustrate how selectively emphasizing various measures of "average" can mislead without being false.[35]

It is the same with marketing. "On average, users lost 10 pounds" might be advertised by a company, but whether we are measuring from the mean or the median, often we are left guessing. If a few people lost a huge amount of weight while a majority lost a little, using the mean would overemphasize the average outcome. If a minority lost no weight, a marketer might want to use the median to hide a smaller mean. Without a direct statement, "average" is a vague word and can be a manipulative tool for presenting the message the communicator wants.[36]

Finally, leaving context involves presenting a statistic or fact without the background or scope needed to provide a reasonable interpretation. Even facts are misleading if a the broader scope is left out. Leaving context involves presenting a statistic or fact without the background or scope needed to provide a reasonable interpretation.

Omitting context is common for advertising as well as journalism. Companies might say, “Our product reduced defects by 30%,” without noting that it was from a very high starting baseline of defects (so things are still broken), or that it was only a trial for a short time. Political campaigns will quote grand-sounding statistics without context – reporting the number of new jobs added for a month without adjusting for the season, or for the size of the populace. Statistics, without context such as timeframes, denominators, or comparative baselines, are very misleading. Honesty presents the full picture (historical trends, per-capita rates, for instance), but those who are controlling the message will withhold such context as a means of shaping perceptions.

---

[34] "Virtual Mentor: 'Mean, Median, and Mode: Measures of Central Tendency'," *AMA Journal of Ethics* 15, no. 1 (2013): 77–81, https://doi.org/10.1001/virtualmentor.2013.15.1.mnar1-1301

[35] Ibid.

[36] Ibid.

Part II – Practical Examples

I. DATA MANIPULATION IN MARKETING: PERSUASIVE OR MISLEADING?

In marketing and advertising, statistics are powerful persuasive tools, but also frequent instruments of deception. Numbers give credibility and a sense of objectivity to promotional content, which marketers may exploit to influence consumers. As a study from *Leung and Chan* observe, misleading statistical information in advertising not only distorts consumer understanding but can also go unnoticed, even among educated audiences.[37] The study found that many university-level participants, despite being mathematically literate, struggled to detect deceptive claims in marketing messages, underscoring the subtlety and effectiveness of these tactics.

Common techniques were already mentioned above include cherry-picking favorable data, improper or misleading comparisons, and visual distortion through graphs. A vivid example comes from a blog by Emeritus Professor *John Mangan*: a newspaper claimed men visit shopping centers "to meet women" based on a survey. But closer inspection revealed that these men were simply shopping with their partners.[38] The data wasn't false, but its framing created a sensational, misleading narrative, an example of selective sampling used to distort meaning.

Academic and regulatory case studies further illustrate how pervasive and problematic this issue is. In fact skewed marketing statistics can “distort truth and influence decision-making,” often leading to consumer harm.[39] A widely cited case involves *Colgate’s “80% of dentists recommend”* campaign. While technically accurate, dentists did name Colgate in a multi-brand survey, the ad implied exclusive preference.[40] The *UK’s Advertising Standards Authority* banned the ad, recognizing that the omission of context misled viewers. This case is now a classic example of truthful statistics used deceptively.

Fortunately, regulatory bodies are catching up and some have already responded. For instance, between 2015 and 2020, the *U.S. Federal Trade Commission* filed 172 cases against misleading advertising, resulting in over $190 million in settlements compensations.[41] These often involved

---

[37] Leung, Wing Sze, and Andrew Chi Kit Chan. “Consumer Attitudes and Statistical Deception in Marketing: A Survey Study.” Journal of Consumer Psychology, vol. 33, no. 2, 2023, pp. 198–210.
[38] John Mangan, “Use and Abuse of Statistics,” John Mangan Blog, 2022, https://mangansblog.com/use-and-abuse-of-statistics/
[39] Hidalgo Torres, María. La estadística manipulada en marketing: distorsión de la verdad y consecuencias sociales. Madrid: Editorial UCM, 2025
[40] Advertising Standards Authority, “ASA Adjudication on Colgate-Palmolive (UK) Ltd,” July 15, 2009, https://www.asa.org.uk/rulings/colgate-palmolive-uk-ltd-a08-84484.html
[41] Federal Trade Commission (FTC), “FTC Cases Resulting in Refunds to Consumers,” https://www.ftc.gov/enforcement/cases-proceedings/refunds

numeric claims in sectors such as health, finance, and e-commerce, confirming that statistical distortion is not only unethical, but actionable.

From a research standpoint, scholars are exploring both the motivations behind data manipulation and the consumer response to it. Studies suggest that in the digital environment, unverified or misleading statistics spread quickly, sometimes even unintentionally, as companies echo trending claims without due diligence. On the consumer side, numeracy and data literacy remain critical gaps. It has been found that even well-educated consumers showed positive attitudes toward data but lacked the skills to critically evaluate statistical claims in ads.[42] The literature reveals a clear picture: marketers frequently deploy statistical techniques not just to inform, but to influence, exaggerate, or mislead. Whether through biased surveys, misleading visuals, or selective framing, the goal is often persuasion over accuracy. While these strategies may offer short-term gains in attention or sales, they ultimately risk damaging consumer trust and undermining credibility. Academic studies and regulatory actions alike stress the need for greater transparency, critical education, and accountability in how data is used in marketing.

Aside from misunderstandings coming from the biased use of data that would inherently be correct and reliable, it is more interesting and timelier to examine cases of manipulation stemming from the use of false statistical sources and difficulties in accessing reliable sources: these cases involve advertisers referencing statistics or "studies" that were false, biased, or did not actually support the advertising claims. In each, the supposedly authoritative source was misleading or fabricated:

GlaxoSmithKline's Ribena Vitamin C Claims (New Zealand, 2007): GSK advertised Ribena juice with the claim that "the blackcurrants in Ribena have four times the vitamin C of oranges." This statistic was misleadingly sourced: funnily enough, two schoolgirls' science project revealed the ready-to-drink Ribena had almost no vitamin C, while a competitor's orange juice had nearly four times more. New Zealand's Commerce Commission found the claim false, and Ribena's ads "left consumers with a wrong impression" and GSK pleaded guilty to misleading advertising. The company was in the end fined $130,000 for a "massive breach of trust" in citing a nutritional fact that did not hold true for the actual product.[43]

Skechers Shape-Ups "Independent Study" (USA, 2012): Footwear maker Skechers marketed its Shape-ups toning shoes with purported scientific backing, an "independent clinical study" by a chiropractor to claim users would lose weight and tone muscle just by walking in the shoes. The FTC found this study was misrepresented and biased: the results did not actually show the advertised

---

[42] Leung and Chan, "Consumer Attitudes," 2023

[43] Rebecca Smithers, "Students Expose Vitamin C Claims of Ribena," The Guardian, March 27, 2007, https://www.theguardian.com/world/2007/mar/27/schoolsworldwide.foodanddrink

benefits, and Skechers failed to disclose the researcher was hired by the company and married to a Skechers marketing executive. In short, the cited "research" was not credible and under heavy conflict of interest. Skechers' unfounded health claims, amplified by celebrity ads, led to a $40 million settlement and an injunction barring future claims without reliable scientific evidence.[44]

Airborne Supplement's Bogus Clinical Trial (USA, 2008): Airborne, a popular herbal supplement, was sold as a "miracle cold buster" with the help of a claimed clinical study by a schoolteacher co-founder. In truth, no legitimate scientific trial backed its cold-prevention claims, an investigation revealed the so-called study was conducted by a two-person outfit with no medical experts. "There was no credible scientific evidence" for Airborne's efficacy. Following pressure from consumer groups and regulators, Airborne agreed to pay $23.3 million in a class settlement for false advertising. The FTC also charged the ads as deceptive, noting the company's marketing blurred the line between a mere vitamin pill and a real medicine.[45]

Volkswagen's "Clean Diesel" Emissions Claims (Global, 2008–2015): This case was very discussed on the news at the time. Volkswagen ran an international *"Clean Diesel"* ad campaign portraying its diesel cars as low-emission and eco-friendly. The company touted that its diesel technology cut harmful emissions by "90%" compared to conventional engines. This statistic was egregiously false: in 2015 regulators discovered VW had installed software to cheat emissions tests. In reality the cars were emitting up to 40 times the legal NOx limit (4,000% more than allowed) on the road. The Federal Trade Commission sued Volkswagen, stating the automaker's ads deceived over 550,000 customers with false environmental claims. The ensuing *"Dieselgate"* scandal led to billions in fines and a global recall, underscoring how a fabricated statistic (the *"90% cleaner"* claim) can undermine public trust on a massive scale.[46]

Finally, *Lumosity "Brain Training" Efficacy (USA, 2015)*: Lumos Labs advertised its Lumosity brain-training games as scientifically proven to boost memory and even prevent cognitive decline (including Alzheimer's). The FTC found Lumosity's cited "science" to be wildly misleading: the company "simply did not have the science to back up its ads," despite invoking studies and expert-

---

[44] Roni Caryn Rabin, "Skechers Toning Shoe Customers to Get Refund," The New York Times, May 16, 2012, https://archive.nytimes.com/well.blogs.nytimes.com/2012/05/16/skechers-toning-shoe-customers-to-get-refund/

[45] Center for Science in the Public Interest, "Airborne Agrees to Pay $23.3 Million to Settle Lawsuit over False Advertising of Its 'Miracle Cold Buster,'" August 14, 2008, https://www.cspinet.org/news/airborne-agrees-pay-233-million-settle-lawsuit-over-false-advertising-its-miracle-cold-buster

[46] European Parliamentary Research Service, "The Volkswagen Emissions Scandal," European Parliament Briefing, January 2016, https://www.europarl.europa.eu/RegData/etudes/BRIE/2016/583793/EPRS_BRI(2016)583793_EN.pdf

sounding claims. In 2016, Lumosity's makers paid $2 million in fines and agreed to stop using such false scientific claims unless backed by competent evidence.[47]

These are just a few examples of how numbers can be bent and trimmed around the edges, only to eventually add a few extra zeros to the bottom line. The issue becomes far more concerning when those "edges" are national borders, and the manipulation isn't done by some S&P500 marketing executive, but by political candidates from major global players or autocrats who keep the press and national political discourse in a chokehold.

II. DATA MANIPULATION IN POLITICS: THE MARGIN OF (MIS)INTERPRETATION

Statistics are deeply entertwined with political discourse. From shaping legislation to defending policy positions, numerical data serves as both foundation and ammunition. Yet, the persuasive power of statistics often relies less on their accuracy and more on the public's limited statistical literacy. Politicians and the media that amplify them, frequently exploit this gap, presenting figures in misleading or decontextualized ways. This is especially apparent in debates over contentious issues such as abortion or migration issues, where both sides selectively frame statistics to reinforce ideological narratives rather than to inform.

A striking example of how objective truth is often overshadowed by emotionally charged narratives is the recent trend of omitting actual fact-checking during televised head-to-head political debates. This shift favors sensational storytelling that appeals to the public's gut reactions rather than sparking informed discussions. These programs thrive on heated debates, rapid exchanges, and simplified talking points, conditions that rarely allow for critical scrutiny of the data being cited. As a result, statistical claims are frequently presented without context, qualification, or verification. Hosts and guests alike often cite numbers that sound authoritative but may be outdated, misinterpreted, or entirely inaccurate. The lack of real-time correction enables these figures to circulate unchallenged, reinforcing viewers' preexisting beliefs rather than encouraging informed judgment. In many cases, the spectacle of argument matters more than the substance of evidence.

Clearly, while in democratic societies the misuse of data tends to remain confined to the television screen, primarily as a tool to sway public opinion during elections rather than influencing the actual legislative process (hopefully), in autocratic contexts, a government's manipulation of information and the lack of access to accurate, independently verified data can become a powerful

---

[47] Federal Trade Commission, "Lumosity to Pay $2 Million to Settle FTC Deceptive Advertising Charges for Its 'Brain Training' Program," January 5, 2016, https://www.ftc.gov/news-events/news/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges-its-brain-training-program

instrument for political repression. It allows regimes to distort the country's economic reality, suppress dissent, and project an exaggerated image of strength on the international stage.

Let's begin, however, with democratic contexts: what are some examples of statistical manipulation primarily stemming from the use of false data sourcest rather than the difficulty of accessing reliable ones?

Brexit's "£350 Million per Week" Claim (UK): During the 2016 EU referendum, the *Vote Leave* campaign emblazoned a bus with the claim that Britain sent £350 million a week to the EU, implying this money could fund the National Health Service. This figure was misleading: it was a gross amount not accounting for the UK's rebate or funds returned from the EU. The *UK Statistics Authority* publicly rebuked the campaign, calling the claim "misleading and [one that] undermines trust in official statistics".[48] Despite the figure being debunked, its repetition influenced public opinion ahead of the Brexit vote, illustrating how a false statistical source can sway voters.[49]

Fabricated Crime Statistics Graphic (USA): This graphic, tweeted by *Donald Trump* in 2015, purported to show "USA Crime Statistics ~2015" by race, claiming, for example, that 81% of white homicide victims were killed by Black perpetrators. In reality, the image's numbers were completely fabricated and cited a fictitious "Crime Statistics Bureau – San Francisco". Again in reality, no such bureau exists and none of the percentages were supported by official data. By retweeting this chart to millions of followers, Trump injected a false statistical narrative into public discourse, stoking racial fears with manipulated data in order to influence public opinion.[50]

Argentina's Inflated Inflation Rates (2007–2015): For nine years the Argentine government systematically understated its inflation statistics to downplay economic troubles. Officials even fired veteran statisticians and essentially "just made things up" when reporting the Consumer Price Index. During this period, independent economists' price indexes ran roughly 10 percentage points higher than the official rate, revealing a huge gap between true inflation and the rosy official figures. This falsification of data was aimed at fooling the public and investors into believing inflation was under control, thereby influencing wage negotiations and economic policy based on a false premise.[51]

At least in principle, independent verification and access to reliable data remain possible. The same cannot be said for authoritarian and semi-authoritarian regimes. In these contexts, the issue is

---

[48] Leo Barasi, "EU Referendum: Leave Voters Twice as Likely to Disbelieve Facts and Science," Eco-Business, June 20, 2016, https://www.eco-business.com/opinion/eu-referendum-leave-voters-twice-as-likely-to-disbelieve-facts-and-science/.

[49] Full Fact, "£350 Million a Week: The Statistic the UK Statistics Authority Said Was Misleading," Full Fact, last modified July 16, 2021, https://fullfact.org/europe/350-million-week-boris-johnson-statistics-authority-misuse/

[50] Robert Farley, "Trump Retweets Bogus Crime Graphic," FactCheck.org, November 23, 2015, https://www.factcheck.org/2015/11/trump-retweets-bogus-crime-graphic/

[51] Charles Nathanson and Christopher Tang, "The Inflation Lie," Anderson Review (UCLA Anderson School of Management), January 18, 2023, https://anderson-review.ucla.edu/inflation-lies/

not merely the misinterpretation or selective use of data, but the possible complete inaccessibility of verified information. With control over national statistics offices, suppression of independent journalism, and the absence of third-party oversight, governments are free to construct internal narratives with virtually no accountability. This monopolization of data becomes a tool not just for misinformation, but for shaping public perception, obscuring economic realities, and silencing dissent. Against this premises, the misuse of statistics in democratic contexts, while problematic, pales in comparison to the systemic opacity found in authoritarian states.

Underreported COVID-19 Data in Tanzania: In 2020, Tanzania's government, under President *John Magufuli*, stopped releasing COVID-19 case and death data entirely, creating the illusion that the country had no pandemic. Magufuli declared Tanzania virus-free and stopped testing and reporting after April 2020. By cutting off reliable data, the regime misrepresented the outbreak's reality. For over a year no official updates were given, even as anecdotal evidence (overfilled hospitals, surges in funerals) suggested a hidden crisis. This manipulation through data suppression was intended to bolster public confidence and avoid policy action, at the cost of denying the public accurate information about a health emergency.[52]

Chinese GDP Growth Figures: Analysts have long questioned the accuracy of China's economic statistics due to political pressures to meet growth targets. Research by *Yale* scholars finds clear evidence that some local officials inflate GDP numbers, in some cases by simply fabricating data to hit quotas. Because independent verification is limited, these dubious figures enter official records. Even China's overall national data have been met with skepticism: journalists and economists debate whether the meteoric growth rates reported in past years are trustworthy. Inflated statistics give a false picture of economic health, potentially skewing policy decisions (both within China and by foreign governments or investors) based on over-optimistic data.[53]

North Korea's "99.99%" Election Turnout: authoritarian regimes often exploit limited data access to push implausible statistics. North Korea routinely announces near-100% voter turnout and unanimous support its single-party "elections." For instance, state media claimed a 99.99% turnout in the 2019 parliamentary elections. In 2023, even after allowing a token mention of a few dissenting votes, the official results still showed over 99.9% of voters approving the regime's candidates. With no independent observers, these figures cannot be verified and serve as propaganda to legitimize Kim Jong-un's rule. The manipulation of election data, made possible by strict control over information,

---

[52] Dickens Olewe, "Coronavirus in Tanzania: The Country That's Rejecting the Vaccine," BBC News, June 10, 2020, https://www.bbc.com/news/world-africa-52966016

[53] Michael Blanding, "Study Suggests That Local Chinese Officials Manipulate GDP," Yale Insights, November 3, 2021, https://insights.som.yale.edu/insights/study-suggests-that-local-chinese-officials-manipulate-gdp

aims to influence domestic public opinion and project an image of absolute national unity to the world.[54]

Figures on Turkey's Inflation: In 2022, Turkey's official statistics agency reported annual inflation around 79%, a multi-decade high. However, independent economists (hampered by limited access to official raw data) calculated that real inflation was roughly double the official rate. An unofficial research group *(ENAG)* using online price data found prices had actually surged ~175% year-on-year, far above the government's number. This big discrepancy, widely covered in media and evident to Turkish households, suggested deliberate underreporting. With strict control over data publication, authorities could present a better economic picture to influence public sentiment and avoid admitting a severe crisis. The Turkey case shows how restricting data transparency enables statistical misrepresentation that affects economic policy debates and public trust.[55]

[54] "North Korea Cites Rare Dissent in Elections, Even with 99% Backing for Candidates," Reuters, November 28, 2023, https://www.reuters.com/world/asia-pacific/north-korea-cites-rare-dissent-elections-even-99-back-candidates-2023-11-28/

[55] Phillip Inman, "Turkey Hit with Soaring Prices as Inflation Reaches 24-Year High," The Guardian, July 4, 2022, https://www.theguardian.com/business/2022/jul/04/turkey-hit-with-soaring-prices-as-inflation-reaches-24-year-high-erdogan.

Part III – How to mitigate misinformation: current landscape and future challenges

I. ETHICAL DATA: THE ROLE OF SCIENTISTS AND PUBLIC INSTITUTIONS

Debates over the social and ethical responsibilities of scientists have a long history. In the early 20th century, many researchers embraced the idea of value-free science, believing their sole duty was to conduct objective research, not address social impacts, however, much has changed since then. A growing consensus holds that scientists do bear responsibility for the consequences of their work. As members of society, scientists share general moral duties: to avoid causing harm and to help others and they have in fact special obligations, also because their training and research often rely on public support. Public institutions likewise play a crucial role in following ethical norms and ensuring that science serves the public good.

The ethical consciousness of the scientific community was indeed awakened by events of the mid-20th century. Most notably, the development of the atomic bomb in the Manhattan Project forced scientists to confront the moral implications of their work. For example, physicist Joseph Rotblat chose to leave the Manhattan Project in late 1944 upon learning the weapon's aim had shifted beyond deterring Nazi Germany; he then devoted his life to nuclear disarmament and ultimately shared the 1995 Nobel Peace Prize for his efforts.[56] Many Manhattan Project researchers wrestled with such dilemmas: they felt an obligation to help end World War II, yet feared how their inventions would be used. Following these footsteps, by the late 20th century, formal statements of scientists' social responsibilities began to emerge.[57]

Contemporary ethical frameworks build on these lessons. Today it is widely accepted that scientists must consider the societal impacts of their research. Whereas a century ago many scientists disavowed responsibility for how others used their analysis, now there is recognition that researchers and statisticians should at least minimize potential harms and actively promote good outcomes.

Scientists and statisticians inevitably interact with politics and policymaking, whether as advisers, advocates, or experts informing public debate. This raises critical questions: Should experts remain neutral technical advisors, or actively champion causes informed by their expertise? What ethical lines must they draw when data meets politics?

---

[56] Naomi Oreskes, "The Manhattan Project Shows Scientists' Moral and Ethical Responsibilities," Scientific American, August 4, 2020, https://www.scientificamerican.com/article/the-manhattan-project-shows-scientists-moral-and-ethical-responsibilities/.

[57] Resnik DB, Elliott KC. The Ethical Challenges of Socially Responsible Science. Account Res. 2016;23(1):31-46. doi: 10.1080/08989621.2014.1002608. PMID: 26193168; PMCID: PMC4631672.

On one hand, the ideal of impartiality is valued: scientists are expected to provide objective evidence without political spin. On the other hand, when scientific evidence is ignored or misrepresented by policymakers, many argue that experts in fact do have a duty to speak up, for example during the COVID-19 pandemic. In such circumstances, staying silent can be ethically problematic. One could argue that embracing political engagement allows scientists to fulfill their responsibility to protect the integrity of their work and contribute to society. This view echoes the stance of climate scientists-turned-activists who strongly states that with knowledge comes an obligation to warn society of impending dangers, even if it means taking on an advocacy role.

At the same time, ethical engagement in politics does not mean abandoning objectivity or truth. Both statisticians and public officials carry responsibilities for honest communication. The *International Science Council* (ISC), in a 2020 statement, articulated several guiding principles for scientists in the policy realm. Some of these principles include:

- Use the best available evidence: "Health and social policies should be guided by the best possible scientific evidence… scientists therefore have the responsibility to use the best evidence that they can". In other words, this means rigorously choosing data and not cherry-picking or distorting findings when advising leaders or the public.
- Counter misinformation: Misinformation can have dire consequences. The scientific community "has the responsibility to be vigilant in the face of such anti-scientific acts, to make publicly known their lack of validity, and to advocate strongly for scientific values and the scientific method". For example, during the COVID-19 crisis, scientists worked to debunk false claims about cures and origins of the virus, recognizing that lives were at stake.
- Communicate uncertainties: Especially on complex issues (like climate change or the pandemic), scientists must be transparent about what is known and unknown. As the ISC emphasized, "scientists have the responsibility to communicate uncertainties, where they exist, to policy-makers and the public"
- Clearly conveying the margins of error or confidence intervals in research can prevent public misperceptions and build trust.
- Inform policy, don't dictate: Science alone cannot decide policy, which also involves social values, economics, and public priorities. Thus, "science should aim to best inform policy, not dictate it".[58]

[58] International Science Council, "Scientific Freedom and Responsibility: CFRS Statement – 15 June 2020," International Science Council, June 15, 2020, https://council.science/news/cfrs-statement-15-june-2020/.

An important aspect of scientists' political engagement is scientific integrity within government institutions. In the United States and elsewhere, there have been controversies over political interference in research: from censorship of climate reports to pressuring experts at public health agencies.

Public institutions must protect researchers who report inconvenient truths rather than allowing science to be twisted for political ends. Independent science advisory boards, whistleblower protections, and strong scientific integrity policies are mechanisms that uphold this ethical standard. For example, the *Union of Concerned Scientists* (a science advocacy NGO) has documented numerous cases of suppression or distortion of scientific and statistical findings by political appointees, and it urges reforms so that government scientists can communicate openly without fear of retribution.[59] The ethical bottom line is that policy-making should be informed by the best science available and presented honestly, a principle that serves not only democracy but also human welfare, as poor decisions on issues like public health, crime statistics or environmental safety can cost lives.

Science does not only meet society in politics; it is also entangled with marketing and media, which raises another set of ethical considerations. Scientific claims are used to sell products, to gain public attention, and to persuade, sometimes in ways that skirt honesty.

Historically, abuses of scientific authority in advertising provide cautionary tales. A notorious example is tobacco advertising in the 1940s–50s. Cigarette companies famously ran ads featuring doctors to bolster the health image of their products. "More Doctors Smoke Camels than Any Other Cigarette!", one 1946 magazine ad declares, showing a physician figure to imply medical endorsement.[60] In reality, this was pseudoscience and marketing spin: the tobacco industry conducting flimsy "surveys" of doctors to support a misleading claim. The campaign leveraged the public's trust in doctors to sell cigarettes, despite growing evidence of harm. By today's standards, this is a gross violation of ethical marketing; it conflicts with the duty to do no harm and the responsibility to accurately represent scientific knowledge.

In response to past abuses, modern societies have established rules and norms for ethical marketing, especially when scientific or health claims are involved. Advertising regulations in many countries require that marketing claims be truthful, not misleading, and backed by evidence. For instance, the *U.S. Federal Trade Commission* stipulates: "Under the law, claims in advertisements must be truthful, cannot be deceptive or unfair, and must be evidence-based".[61] This principle applies

---

[59] Union of Concerned Scientists, "Scientific Integrity in Policy Making," Union of Concerned Scientists, accessed May 22, 2025, https://www.ucs.org/resources/scientific-integrity-policy-making.
[60] Stanford Research into the Impact of Tobacco Advertising, "Cigarette Advertising: Image 0027," Stanford University, accessed May 22, 2025, https://tobacco.stanford.edu/cigarette/img0027/.
[61] Federal Trade Commission, "Advertising and Marketing Basics," Consumer Advice, accessed May 22, 2025, https://consumer.ftc.gov/business-guidance/advertising-marketing.

whether a company is marketing a new drug, a dietary supplement, or even a household cleaner claiming to be "scientifically proven" or backed by statistical data to kill germs. Regulatory bodies can penalize advertisers for false scientific claims. Ethical marketing thus demands to be truthful to science itself: claims should be substantiated by reliable data, and qualifications or uncertainties should be clearly stated to avoid false impressions.[62]

Scientists working in industry or acting as product endorsers carry particular ethical responsibility. A researcher should not lend credibility to a product that they know lacks solid evidence. Professional associations sometimes provide guidelines here; for example, medical ethics would forbid a physician promote a treatment not supported by clinical evidence. The appearance of scientific approval is powerful in marketing, so using it deceitfully is considered a serious breach of public trust. This extends to more subtle forms of marketing, such as corporate-funded research that is later used in advertising. Scientists and institutions must ensure their research sponsorships and potential conflicts of interest are transparent, to avoid eroding credibility.

Beyond commercial advertising, marketing in the context of science communication also demands attention. For example, universities and research institutions "market" science when they issue press releases about discoveries, hoping to attract media coverage. There is an inherent tension: they want to celebrate and promote their scientists' work, but overselling results can mislead the public. A revealing 2014 study in *BMJ* examined health-related news and press releases, and found that exaggeration in news stories was strongly associated with exaggeration in the academic press releases that fed those stories.[63] In other words, hype often starts at the source. If a university press office claims a "breakthrough cure" where the actual research is a preliminary mouse study, journalists and the public may be misled about the certainty or importance of the finding. Ethically, institutions should exercise restraint and accuracy in science communication.

Similarly, individual scientists who communicate directly through social media, blogs and interviews should treat those interactions as extensions of their professional ethics. The rise of Twitter, YouTube, and podcasts means researchers often act as their own marketers to the public. This can humanize science and rapidly disseminate knowledge, but it also carries risks. Scientists must be careful that in trying to popularize their work, they don't sacrifice nuance or feed into sensationalism. This means checking facts, acknowledging uncertainty, and correcting mistakes in public forums. It also means understanding the audience's perspective: effective communication should inspire trust, not confusion or fear.

---

[62] Sandy Walsh, "How Advertising Self-Regulation Can Provide Guidance on Health-Related Product Claims," SupplySide SJ, accessed May 22, 2025, https://www.supplysidesj.com/supplement-regulations/how-advertising-self-regulation-can-provide-guidance-on-health-related-product-claims.

[63] Peter Doshi, "Pfizer and Moderna's '95% Effective' Vaccines—Let's Be Cautious and First See the Full Data," BMJ 349 (November 2020): g7015, https://www.bmj.com/content/349/bmj.g7015.

II. EDUCATING THE PUBLIC: STATISTICAL LITERACY

In an age awash with data and misleading claims, data literacy, the ability to understand and critically evaluate data and statistics has become an essential skill. Misinformation often exploits statistics or "authoritative" looking figures to lend false credibility to incorrect claims. Without the skills to interpret graphs and numbers critically, people are at a disadvantage in recognizing truth: "if you torture numbers, they'll confess to anything",[64] highlighting how easily statistics can be manipulated.

Governments and experts worldwide emphasize that improving public data literacy is crucial for societal resilience against "fake news" and statistical manipulation. Widespread statistical literacy as the capacity to interpret and question statistical information is seen as vital for a healthy democracy. Misleading use of data ranges from manipulated charts to contextless figures. Understanding such tricks, whether it's skewed axes, cherry-picked statistics, or "base rate" fallacies requires numerate critical thinking. A fact-checking analysis by *Full Fact (UK)* notes how poor numeracy can distort how we see the truth, citing examples from COVID-19 vaccine stats (where a higher absolute number of vaccinated people died without implying the vaccine was ineffective) to absurd misreadings of drug data.[65] These cases underscore that data literacy is a new survival skill in the information age– empowering individuals to see through numerical deception and make informed decisions.

Educators and researchers stress that data literacy isn't just about mathematical ability, but about critical reasoning and healthy skepticism. Psychological studies and meta-analyses find that media and information literacy training can significantly improve people's resilience to misinformation (with a notable medium-to-large effect size) by instilling habits of questioning and verification.[66] In other words, teaching citizens how to think critically about statistics and news reports can reduce their likelihood of believing or spreading false claims. This is especially urgent as digital platforms amplify statistical misinformation, from viral social media graphs to dubious "studies" that propagate unvetted data.

Scholars have developed theoretical frameworks to guide data literacy education for the public. Statistical literacy is commonly defined as the ability to interpret, critically evaluate, and communicate messages involving data.

In essence, a statistically literate citizen should ask probing questions about how data were collected, what is being claimed, and whether conclusions are justified. This notion of "statistical

---

[64] Good, I. J. "Statistics and Today's Problems." The American Statistician 26, no. 3 (1972): 11–19. https://doi.org/10.2307/2682859.
[65] Claire Milne, "From COVID Claims to Cocaine Stats: Why Misleading Numbers Matter," Full Fact, February 16, 2023, https://fullfact.org/education/from-covid-claims-to-cocaine-stats-why-misleading-numbers-matter/.
[66] Fake News and Statistical Literacy." Statistics Teacher, November 13, 2023. https://www.statisticsteacher.org/2023/11/13/fake-news-statistical-literacy/

citizenship" highlights that citizens need skepticism tempered with openness: not accepting every statistic at face value, but also not dismissing evidence solely because it's counterintuitive.[67]

A growing perspective is critical data literacy, which situates data skills within a broader understanding of power and context. As *Spiranec et al. (2019)* argue, data literacy today means being able to "challenge, object and protest" the power asymmetries in a datafied society.[68] This framework goes beyond reading charts as it includes recognizing biases in data collection, questioning who is presenting the numbers and why, and understanding the societal impact of data-driven claims. Similarly, media literacy scholars call for integrating data literacy with media literacy, coining terms like "media and data literacy education" to address phenomena like conspiracy theories and propaganda. The underlying principle is that to counter complex misinformation, people must grasp both the content of factual data and the context of how information can be spun or weaponized.

Educating the public in data literacy is a long-term endeavor, and measuring its success is complex. Studies so far are encouraging: research syntheses indicate that well-designed literacy interventions can improve people's ability to discern misinformation and reduce their susceptibility to misleading claims. For example, a meta-analysis of media literacy programs (including those teaching how to analyze statistics in news) found significant improvements in critical thinking and mistaken-belief correction among participants.[69] Furthermore, surveys have shown correlations between higher numeracy and lower belief in COVID-19 hoaxes or conspiracy theories, suggesting that boosting numerical understanding might directly contribute to a more informed public.[70]

That said, challenges remain. One is scale and accessibility: reaching all age groups from schoolchildren to seniors requires diverse channels and continuous effort. Many adults missed out on data education in school, so initiatives must meet them through workplaces, community programs, or popular media. Another challenge is keeping up with the evolving landscape of misinformation. As manipulation tactics become more sophisticated (deepfake data visualizations or AI-generated "studies" for example), educational content must also evolve. Experts emphasize the importance of teaching core principles that transfer to new scenarios: understanding uncertainty, demanding transparency of sources, and maintaining a healthy skepticism.

---

[67] Fake News and Statistical Literacy," Statistics Teacher, November 13, 2023, https://www.statisticsteacher.org/2023/11/13/fake-news-statistical-literacy/

[68] , Lockley, Eleanor and Pawluczuk, Alicja. 2020. "Data citizenship: rethinking data literacy in the age of disinformation, misinformation, and malinformation". Internet Policy Review 9 (2). DOI: 10.14763/2020.2.1481. https://policyreview.info/articles/analysis/data-citizenship-rethinking-data-literacy-age-disinformation-misinformation-and.

[69] Seoyoon Woo and Zhan Zhang, "Seeing Is Believing? How Data Visualization Formats Influence Public Trust and Misinformation," Communication Research, published online April 8, 2024, https://doi.org/10.1177/00936502241288103

[70] Sarah Boseley, "Poor Numerical Literacy Linked to Greater Susceptibility to Covid-19 Fake News," The Guardian, October 14, 2020, https://www.theguardian.com/world/2020/oct/14/poor-numerical-literacy-linked-to-greater-susceptibility-to-covid-19-fake-news

Lastly, there's the question of motivation. People often don't engage with numeracy for its own sake; thus, tying data literacy to personally relevant goals is key. Programs that frame statistical skills as empowering: protecting oneself from scams, making informed health choices, or holding officials accountable with facts tend to resonate better than abstract math lessons. This aligns with the concept of "data citizenship," where the aim is not just to comprehend data but to use that understanding to participate in society and advocate for truth.

By cultivating this sense of agency, data literacy initiatives hope to create a public that not only spots when "the numbers don't add up," but also demands honesty and clarity in the use of statistics.

III. AI AND PROCEDURALLY GENERATED CONTENT: CURRENT AND FUTURE CHALLENGES

Advances in artificial intelligence have enabled procedurally generated content, namely the algorithmic creation of text, images, video, and even data at an unprecedented scale. From AI-written news articles to auto-generated game worlds, these technologies promise greater creativity and efficiency. However, they also raise serious concerns about the manipulation of information and statistical data. When anyone can easily generate realistic but fabricated content, it becomes challenging to trust what we see, read, or even measure.

AI-generated media can be weaponized to spread false or misleading information on a large scale. Generative models can produce highly convincing fake images, videos, audio, or text that mimic authoritative sources. This capability has already been misused in political propaganda: for example, deepfake videos and images have been deployed to influence elections and fuel geopolitical falsehoods.[71] Such AI-driven disinformation campaigns are cheap and scalable, allowing bad actors to automate the creation of fake news stories, synthetic statistics, or bogus "expert" commentary. One European report notes that AI-generated content has been used in deceptive ads and social media posts (a deepfake of a president for example) and even to impersonate voices of politicians, all with the intent to mislead the public.[72]

Beyond creating fake content, the mere existence of generative AI gives rise to a "liar's dividend": opportunists can falsely claim that authentic images or videos are "AI fakes," undermining the truth. A recent example saw corporate lawyers dismissing a legitimate video as a deepfake to avoid

[71] European Digital Media Observatory. Generative AI and Disinformation: White Paper. December 2023. https://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation_-White-Paper-v8.pdf
[72] Ibid.

accountability. Overall, the spread of AI-facilitated misinformation erodes trust in media and makes it harder for citizens to discern reality from fabrication.

It is also interesting to focus on the ethical part of the issue, not just the practical side. For example, AI researchers recognize that algorithms improving image recognition or text generation could be applied to positive ends (medical diagnostics, educational tools) or negative ones (mass surveillance, deepfake disinformation). Thus, many AI scientists are now engaging in ethics by design creating frameworks to ensure transparency, fairness, and accountability in AI systems. In 2023, hundreds of tech leaders and researchers signed an open letter calling for a pause in training the most powerful AI models until safety protocols are in place, citing the ethical imperative to assess societal risks before forging ahead.[73]

AI-driven procedural content generation is transforming how content and data are created across media, games, journalism, and science. Its promise is immense from automating creative tasks to simulating data for research, but so are the perils of misuse. Key ethical concerns revolve around misinformation, embedded bias, and the erosion of trust when we can no longer be sure if data or media are genuine. Technically, society faces an uphill battle in detecting synthetic content, ensuring AI outputs are reliable, and maintaining quality control. The policy landscape is rapidly evolving to catch up with the technology, through laws mandating transparency, initiatives to develop standards (like content watermarking), and guidelines for responsible use.

The manipulation of statistical data is a thread that ties these issues together: whether it's fake numbers in a news story or fabricated results in a research paper, the integrity of empirical truth is at stake. Going forward, meeting these challenges will require a multipronged strategy. Stakeholders must invest in better tools to verify authenticity, educate the public to be critical of what they consume, design AI systems with ethical guardrails, and implement smart regulations that deter malicious behavior without stifling innovation. If successful, we can reap the benefits of generative AI and procedural content in richer games, personalized media, and accelerated research while preserving truth and trust in our digital society. The road ahead demands vigilance and collaboration, ensuring that these powerful new content technologies become tools for progress, not weapons of deception.

---

[73] Future of Life Institute. "Pause Giant AI Experiments: An Open Letter." March 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Conclusion

This thesis has demonstrated, through the analysis of real-world examples and reflective inquiry, that statistical data occupies a central though frequently misunderstoodposition in shaping how societies function, communicate, and decide. Across domains like marketing, politics, and digital media, data has evolved into both an essential instrument and a potential tool for manipulation. In an era overwhelmed by information, the ability to access trustworthy data has become increasingly difficult, while the potential to distort numbers, deliberately or inadvertently, has only grown.

What has emerged clearly is the pressing need for greater statistical literacy among the general public. The public's persistent inability to decipher numbers, especially when presented in isolation or with an intent to persuade leaves the field vulnerable to mass disinformation. When figures are stripped of context or deployed with persuasive intent, the risk of widespread misinformation becomes not just likely but inevitable. The chapters focusing on political messaging and advertising campaigns reveal just how blurred the boundary can be between informing and misleading when statistics are involved.

Yet, the road ahead need not be bleak. The key lies in restoring a commitment to ethical standards in the communication of data, reinforcing educational efforts, and holding institutions to a higher level of transparency and accountability. Scientists and policymakers must act as stewards of clarity and integrity, while citizens must be equipped with the tools to question and interpret the data that surrounds them. As new challenges arise on the horizon and most particularly, those brought about by AI-generated content and algorithmic censorship our collective ability to identify truth amidst distortion will no longer be an asset, but a societal requirement

.

Bibliography

.

- *Advertising Standards Authority. "ASA Adjudication on Colgate-Palmolive (UK) Ltd." July 15, 2009. https://www.asa.org.uk/rulings/colgate-palmolive-uk-ltd-a08-84484.html.*
- *Barasi, Leo. "EU Referendum: Leave Voters Twice as Likely to Disbelieve Facts and Science." Eco-Business, June 20, 2016. https://www.eco-business.com/opinion/eu-referendum-leave-voters-twice-as-likely-to-disbelieve-facts-and-science/.*
- *Burn-Murdoch, John. "Why You Should Never Trust a Data Visualisation." The Guardian, July 24, 2013. https://www.theguardian.com/news/datablog/2013/jul/24/why-you-should-never-trust-a-data-visualisation.*
- *Butt, S., et al. Public Confidence in Official Statistics 2021. NatCen Social Research, 2022. https://uksa.statisticsauthority.gov.uk/publication/public-confidence-in-official-statistics-2021/.*
- *Center for Science in the Public Interest. "Airborne Agrees to Pay $23.3 Million to Settle Lawsuit over False Advertising of Its 'Miracle Cold Buster.'" August 14, 2008. https://www.cspinet.org/news/airborne-agrees-pay-233-million-settle-lawsuit-over-false-advertising-its-miracle-cold-buster.*
- *Charles Nathanson and Christopher Tang, "The Inflation Lie," Anderson Review (UCLA Anderson School of Management), January 18, 2023, https://anderson-review.ucla.edu/inflation-lies/.*
- *Claire Milne, "From COVID Claims to Cocaine Stats: Why Misleading Numbers Matter," Full Fact, February 16, 2023, https://fullfact.org/education/from-covid-claims-to-cocaine-stats-why-misleading-numbers-matter/.*
- *Dani, A., and E. Al Quraan. "Investigating Research Students' Perceptions About Statistics and Its Impact on Their Choice of Research Approach." Heliyon 9, no. 10 (September 29, 2023): e20423. https://doi.org/10.1016/j.heliyon.2023.e20423.*
- *Dickens Olewe, "Coronavirus in Tanzania: The Country That's Rejecting the Vaccine," BBC News, June 10, 2020, https://www.bbc.com/news/world-africa-52966016.*
- *European Digital Media Observatory. Generative AI and Disinformation: White Paper. December 2023. https://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation_-White-Paper-v8.pdf.*
- *European Parliamentary Research Service. "The Volkswagen Emissions Scandal." European Parliament Briefing, January 2016. https://www.europarl.europa.eu/RegData/etudes/BRIE/2016/583793/EPRS_BRI(2016)583793_EN.pdf.*
- *Farley, Robert. "Trump Retweets Bogus Crime Graphic." FactCheck.org, November 23, 2015. https://www.factcheck.org/2015/11/trump-retweets-bogus-crime-graphic/.*
- *Federal Trade Commission. "Advertising and Marketing Basics." Consumer Advice. Accessed May 22, 2025. https://consumer.ftc.gov/business-guidance/advertising-marketing.*
- *Federal Trade Commission. "FTC Cases Resulting in Refunds to Consumers." https://www.ftc.gov/enforcement/cases-proceedings/refunds.*
- *Federal Trade Commission. "Lumosity to Pay $2 Million to Settle FTC Deceptive Advertising Charges for Its 'Brain Training' Program." January 5, 2016.*

https://www.ftc.gov/news-events/news/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges-its-brain-training-program.

- *Fischer, David. Historians' Fallacies: Toward a Logic of Historical Thought. New York: Harper & Row, 1979.*
- *Full Fact. "£350 Million a Week: The Statistic the UK Statistics Authority Said Was Misleading." Last modified July 16, 2021. https://fullfact.org/europe/350-million-week-boris-johnson-statistics-authority-misuse/.*
- *Future of Life Institute. "Pause Giant AI Experiments: An Open Letter." March 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/.*
- *Gardenier, John, and David Resnik. "The Misuse of Statistics: Concepts, Tools, and a Research Agenda." Accountability in Research 9, no. 2 (2002): 65–74.*
- *Gigerenzer, Gerd. Risk Savvy: How to Make Good Decisions. New York: Viking, 2014.*
- *Good, I. J. "Statistics and Today's Problems." The American Statistician 26, no. 3 (1972): 11–19. https://doi.org/10.2307/2682859.*
- *Grant, Gareth. "Relative Measures of Effects Can Be Misleading." Students 4 Best Evidence (Cochrane), February 16, 2018. https://s4be.cochrane.org/blog/2018/02/16/relative-measures-effects-can-misleading/.*
- *Hidalgo Torres, María. La estadística manipulada en marketing: distorsión de la verdad y consecuencias sociales. Madrid: Editorial UCM, 2025.*
- *Huff, Darrell. How to Lie with Statistics. New York: W.W. Norton & Company, 1954.*
- *Institute for Propaganda Analysis. The Fine Art of Propaganda: A Study of Father Coughlin's Speeches. New York: Harcourt Brace and Company, 1939.*
- *Institute of Data. "Cherry Picking in Data Analytics: How to Avoid Common Pitfalls." Accessed April 2025. https://www.institutedata.com/blog/cherry-picking-in-data-analytics/.*
- *International Science Council. "Scientific Freedom and Responsibility: CFRS Statement – 15 June 2020." International Science Council, June 15, 2020. https://council.science/news/cfrs-statement-15-june-2020/.*
- *Kahneman, Daniel. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux, 2013.*
- *Kim, John C. S. The Art of Creative Critical Thinking. Lanham, MD: University Press of America, 1993.*
- *Kirk, Andy. Data Visualisation: A Handbook for Data Driven Design. London: SAGE, 2016.*
- *Leung, Wing Sze, and Andrew Chi Kit Chan. "Consumer Attitudes and Statistical Deception in Marketing: A Survey Study." Journal of Consumer Psychology 33, no. 2 (2023): 198–210.*
- *Lockley, Eleanor, and Alicja Pawluczuk. "Data Citizenship: Rethinking Data Literacy in the Age of Disinformation, Misinformation, and Malinformation." Internet Policy Review 9, no. 2 (2020). https://doi.org/10.14763/2020.2.1481.*
- *Mangan, John. "Use and Abuse of Statistics." John Mangan Blog, 2022. https://mangansblog.com/use-and-abuse-of-statistics/.*
- *Michael Blanding, "Study Suggests That Local Chinese Officials Manipulate GDP," Yale Insights, November 3, 2021, https://insights.som.yale.edu/insights/study-suggests-that-local-chinese-officials-manipulate-gdp.*
- *Motta, Matt, Dominik Stecula, and Christina Farhart. How Visualizing Data Legitimizes Unorthodox Science. arXiv, 2021. https://arxiv.org/.*

- *Naomi Oreskes, "The Manhattan Project Shows Scientists' Moral and Ethical Responsibilities," Scientific American, August 4, 2020, https://www.scientificamerican.com/article/the-manhattan-project-shows-scientists-moral-and-ethical-responsibilities/.*
- *Nikiforova, A., and K. McBride. "Open Government Data Portal Usability: A User-Centred Usability Analysis of 41 Open Government Data Portals." Telematics and Informatics 58 (2021): 101539. https://www.sciencedirect.com/science/article/pii/S0736585321000542.*
- *Norman, Sarah E. How Data Literacy Can Keep America Safe. TIME, June 29, 2023. https://time.com/6290961/data-literacy-statistics-.*
- *Open Data Watch. Overcoming Data Graveyards in Official Statistics. 2022. https://opendatawatch.com/publications/overcoming-data-graveyards-in-official-statistics/.*
- *Pennycook, Gordon, and David G. Rand. How Online Search Can Increase Belief in Misinformation. Nature 621 (2023): 822–827. https://www.nature.com/articles/s41586-023-06883-y.*
- *Peter Doshi, "Pfizer and Moderna's '95% Effective' Vaccines—Let's Be Cautious and First See the Full Data," BMJ 349 (November 2020): g7015, https://www.bmj.com/content/349/bmj.g7015.*
- *Phillip Inman, "Turkey Hit with Soaring Prices as Inflation Reaches 24-Year High," The Guardian, July 4, 2022, https://www.theguardian.com/business/2022/jul/04/turkey-hit-with-soaring-prices-as-inflation-reaches-24-year-high-erdogan.*
- *Rabin, Roni Caryn. "Skechers Toning Shoe Customers to Get Refund." The New York Times, May 16, 2012. https://archive.nytimes.com/well.blogs.nytimes.com/2012/05/16/skechers-toning-shoe-customers-to-get-refund/.*
- *Resnik, David B., and Kevin C. Elliott. "The Ethical Challenges of Socially Responsible Science." Accountability in Research 23, no. 1 (2016): 31–46. https://doi.org/10.1080/08989621.2014.1002608.*
- *Ruchlis, Hyman, and Sandra Oddo. Clear Thinking: A Practical Introduction. Buffalo, NY: Prometheus Books, 1990.*
- *Sarah Boseley, "Poor Numerical Literacy Linked to Greater Susceptibility to Covid-19 Fake News," The Guardian, October 14, 2020, https://www.theguardian.com/world/2020/oct/14/poor-numerical-literacy-linked-to-greater-susceptibility-to-covid-19-fake-news.*
- *Scharrer, Lisa, et al. "When Science Becomes Too Easy: Science Popularization Inclines Laypeople to Underrate Their Dependence on Experts." Public Understanding of Science 26, no. 8 (2016): 1003–1018. https://journals.sagepub.com/doi/abs/10.1177/0963662516646047.*
- *Seoyoon Woo and Zhan Zhang, "Seeing Is Believing? How Data Visualization Formats Influence Public Trust and Misinformation," Communication Research, published online April 8, 2024. https://doi.org/10.1177/00936502241288103.*
- *Siegrist, Michael. "The Use or Misuse of Three-Dimensional Graphs to Represent Lower-Dimensional Data." Behaviour & Information Technology 15, no. 2 (1996): 96–100.*
- *Sostek, K., D. M. Russell, et al. "Discovering Datasets on the Web Scale: Challenges and Recommendations for Google Dataset Search." Harvard Data Science Review, Special Issue, 2024. https://hdsr.mitpress.mit.edu/pub/dataset-search/release/1.*

- *Spirer, Herbert F., Louise Spirer, and Jon Jaffe. Misuse of Statistics: Concepts, Tools, and a Research Agenda. New York: Marcel Dekker, 1998.*
- *Stanford Research into the Impact of Tobacco Advertising. "Cigarette Advertising: Image 0027." Stanford University. Accessed May 22, 2025. https://tobacco.stanford.edu/cigarette/img0027/.*
- *Texas State Auditor's Office. "Methodology Manual: Data Analysis: Displaying Data - Deception with Graphs." January 4, 1996. https://web.archive.org/web/20030402194810/http://www.sao.state.tx.us/methmanual/da9.html.*
- *The New Atlantis. "Correlation, Causation, and Confusion." n.d. https://www.thenewatlantis.com/publications/correlation-causation-and-confusion.*
- *Tufte, Edward R. The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press, 2006.*
- *U.S. Census Bureau. Understanding and Using American Community Survey Data: What All Data Users Need to Know. U.S. Department of Commerce, 2018. https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch07.pdf.*
- *Union of Concerned Scientists. "Scientific Integrity in Policy Making." Accessed May 22, 2025. https://www.ucs.org/resources/scientific-integrity-policy-making.*
- *University of Oxford. Major Research on Hallucinating Generative Models Advances Reliability of Artificial Intelligence. June 20, 2024. https://www.ox.ac.uk/news/2024-06-20-major-research-hallucinating-generative-models-advances-reliability-artificial.*
- *Walsh, Sandy. "How Advertising Self-Regulation Can Provide Guidance on Health-Related Product Claims." SupplySide SJ. Accessed May 22, 2025. https://www.supplysidesj.com/supplement-regulations/how-advertising-self-regulation-can-provide-guidance-on-health-related-product-claims.*
- *Weatherburn, Don. "Uses and Abuses of Crime Statistics." Crime and Justice Bulletin, no. 153 (2011).*
- *Wikipedia. "Cherry Picking." Last modified July 2023. https://en.wikipedia.org/wiki/Cherry_picking.*
- *"Misleading Graph." Wikipedia. Last modified April 26, 2025. https://en.wikipedia.org/wiki/Misleading_graph.*
- *"Virtual Mentor: 'Mean, Median, and Mode: Measures of Central Tendency'." AMA Journal of Ethics 15, no. 1 (2013): 77–81. https://doi.org/10.1001/virtualmentor.2013.15.1.mnar1-1301.*
- *"Fake News and Statistical Literacy." Statistics Teacher, November 13, 2023. https://www.statisticsteacher.org/2023/11/13/fake-news-statistical-literacy/.*
- *"North Korea Cites Rare Dissent in Elections, Even with 99% Backing for Candidates," Reuters, November 28, 2023, https://www.reuters.com/world/asia-pacific/north-korea-cites-rare-dissent-elections-even-99-back-candidates-2023-11-28/*