Vol.2 | Issue 1 | 2019

MAKTABAH JAFARIYAH
Knowledge & Research Academy

The International Journal

ISSN 2581-4354

Analysis of Misleading Statistics

Mr. Faiz Mohammed Masi
M. Tech, Power Electronics.
Charotar University of Science and Technology, Changa.

Abstract:
With a sense of certainty associated with it, Statistics are popularly used in all sorts of things ranging from research to advertisements and even political campaigns. Now, decision making is also dependent on statistics. This paper delineates concerns regarding misuse of statistics in various domains with special attention to the scientific domain. Specific ways by which this malpractice is done are identified and listed along with real-life examples. The key terms pertaining to this domain are also included to give the context of the topic and to provide a better understanding. There are various methods employed to manipulate statistics and persuade people. And there is a real urgency to figure out their demerits and consequences because of the sheer number of people affected and let astray because of it. In the later part of the paper, different techniques are put forward to counter this unethical practice.

Key Words: Statistics, Manipulation of data, Statistical literacy.

Rationale:
The researcher has tried to find and analyze different ways by which statistics are manipulated, in the hope that it will create awareness about a vital yet often overlooked skill i.e. statistical literacy.

Introduction:
There are plenty of ways and methods by which opinions of people on a large scale can be influenced. Such as by way of manipulative and loaded words, by using faulty statistics, by way of graphical presentation to name a few. For the course of this research, only statistical and graphical manipulation are considered for the analysis.

Statistics are cardinal in today's market as well as research. They are very much part of decision making and are used to justify certain decisions. They are essential for research also because we need data to test the hypothesis ¬— whether for or against it. Manipulation of data and manufacturing of deceptive and frivolous data is not new at this point. Misleading statistics are widely used because people see certitude when they are presented with graphs, stats, and numbers, which they don't necessarily associate with words. It is believed that graphs and statistics are tough to twist and manipulate, but, we will extensively analyse in this research article that it is not the case.

Literature Review:

Looking at the rampant misuse of statistics, one might wonder why statistics are given so much importance — given their wide misuse.

The American Statistical Association emphasis on the importance of statistics as follows:

"The professional performance of statistical analyses is essential to many aspects of society. The use of statistics in medical diagnoses and biomedical research may affect whether individuals live or die, whether their health is protected or jeopardized, and whether medical science advances or gets sidetracked. Life, death, and health, as well as efficiency, may be at stake in statistical analyses of occupational, environmental, or transportation safety. Early detection and control of new or recurrent infectious diseases depend on sound epidemiological statistics. Mental and social health may be at stake in psychological and sociological applications of statistical analysis."[1] (Ethical Guidelines for Statistical Practice, 1999)

(John S. Gardenier, 2002) talks in great detail about ethics in research as well as statistics and analyses what does it mean to misuse statistics, why statistics are misused, what are the factors behind it, what are its implications, what can be done to overcome this huge problem, and what sorts of questions we should be asking while reading a research paper and why it is essential to learn about Statistical literacy.[2]

In her review article titled, "Misleading Statistical Studies" (Wolf, 2007) thoroughly investigates about misleading statistics, their effect on policymaking. She suggests some steps to minimize misleading statistics. She then presents her case about why policymaking should not be over-reliant on statistics with many studies and concludes that,

"In the policy debate, a badly interpreted or misused number may be better than no number at all, because it can stimulate correction. But a misleading number may become embedded in the policy milieu with no further scrutiny. The less misleading a number is, the less will be the collateral damage from misinterpretation and misuse." (Wolf, 2007)[3]

Some relevant terms:

Statistics: Statistics is the process of collecting, organizing, presenting, analysing and interpreting the data. These steps are done in chronological order i.e. First Data Collection is done, followed by

1 (Ethical Guidelines for Statistical Practice, 1999)

2 (John S. Gardenier, 2002)

3 (Wolf, 2007)

other steps respectively. It is the science of drawing informed conclusions from large amounts of data.

Data Analysis: It is a process of collecting, modelling, and transforming data.

Methods of persuasion:

There are many techniques employed to deceive people and achieve goals whether by presenting or fabricating frivolous data or by other means. Some of the widely utilized ones are identified here. Listing factors which lead people to forge statistics is beyond the scope of this study.

- **False correlation:**

One of the widely used technique to create faulty statistics is a false correlation between unrelated data sets.

As an example, look at this survey. A group of people was surveyed about what will they prefer in case of emergency—an iPhone or a Book? The result of this is obvious. Now, isn't it possible that we create a chart showing 100% of people prefer an iPhone over a Book? Does this represent the whole picture and nuances of it? Although we have taken an extreme example here, the same can be true for other types of statistics as well.

Then the way in which a survey was conducted also plays a crucial role. Because here many different biases come into play. If people are asked a biased question they are likely to answer in a way that would be socially acceptable even if they don't agree with it.

- **Correlation doesn't necessarily mean causation:**

Correlation refers to the relationship between two variables. The Correlation coefficient measures the extent to which they are related. This fallacy is referred to as "post hoc, ergo propter hoc", which means after this, therefore because of this.

When two variables are found to be correlated, it is implied that one variable causes the other and one thing was responsible for the other. But, even if correlations weren't false, they can't always imply causation. Statistics create false dichotomy to suit an agenda.

Like $A \sim B$ then $A \rightarrow B$

This may be true, but the thing is by simply knowing correlation we can't predict which way the causality goes, we need more than one sample or relation. We can infer the causality by accessing additional information. Like knowing which thing occurred first. Is there any solid evidence to suggest that one thing causes the other, etc.

- **Confirmation bias:**

Even if the survey is 100% accurate and tries to show the true picture. Still, there are chances that the input can't be relied upon. Data gathering is often limited by ethical, practical and sometimes even financial constraints. Stats don't take context into consideration. Neither do they consider humane side and ground reality. The expression "garbage in, garbage out." is applicable here.

- **Manipulation of Graphs:**

Statistics are often represented graphically.

The reason behind this is they make analysis easy and help us to understand complex concepts intuitively. However, they can be misleading too.

Whereas manipulation of statistics is used to get the desired result. Alteration of scale is done to emphasize something. The scale of the graph is altered to emphasize the difference, often to amplify it.

We all are well familiar with the story of a boy who was asked by his teacher to decrease the size of the line without rubbing it. So the boy draws a bigger line beside it. That way the first line looks smaller. Well, something similar to this story is also being done with statistics.

Some common characteristics of misleading graphs are as follows:

Graphs with no title, no labeled axis, axis don't start at 0, unequal intervals, different sized bars, and amplified axis and by representing one-dimensional quantities as a two- or three-dimensional objects to compare their sizes.

- **Testing many hypotheses:**

Scientific works are rated based on their reproducibility. One way to overcome this is by testing many hypotheses and reporting only those that match with the study. A slightly different version of this tactic is purposefully making hypothesis whose outcome is most likely to be inconclusive. Moreover, assumptions underlying that research or hypothesis are also hidden from the public domain.

- **Cherry-picking data:**

Purposefully choosing the sample which is likely to give desired answer or response. The audience chosen is not representative of the whole population upon which the survey is conducted.

Constraints in analyzing the data:

Statistics usually produce probabilities; conclusions are provisional, which gives rise to certain paradoxes. Most famous of the bunch is Simpson's paradox.

Simpson's paradox:

"Simpson's paradox, also called Yule-Simpson effect, in statistics, is an effect that occurs when the marginal association between two categorical variables is qualitatively different from the partial association between the same two variables after controlling for one or more other variables."[4] (Carlson, 2019)

In simpler terms, it means that aggregated datasets show a reverse trend than the original one when they were separately presented.

It can be understood as follows:

$$A < B$$

$$C < D$$

$$A+C > B+D$$

Due to Simpson's paradox, it is possible to draw two completely opposite conclusions from the same data depending upon how data sets are divided. In other words, the grouping of data is done in such a way that it enables people to draw a completely different conclusion than the original conclusion which was drawn from

4 (Carlson, 2019)

ungrouped data.

Real life example of Simpson's paradox is the famous case of UC Berkley in 1973 when the university was sued for potential gender bias in admissions. It was alleged that male applicants were 1.8 times more likely to get admission compared to female applicants. However, after an extensive examination of admission process of all the departments individually it was established that admissions were slightly better for women applicants!

	Applicants	Admitted
Men	2691	45%
Women	1835	30%

Table 1: Aggregate results

If we look at the data presented in Table 1, it appears that the university might have discriminated on the basis of gender. But, when confronted with detailed data, the trend swings in the exact opposite direction.

Enrollment in UC Berkley Department wise, 1973

Department	**Men**		**Women**	
	Applicants	Admitted	Applicants	Admitted
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Source: (Berkeley admissions data)[1]

Simpson's paradox happens generally due to the presence of hidden Confounding variable or Lurking variable. In our example, the rejection rate is a confounding variable.

The analysis in scientific domain:

The misuse and manipulation of statistics to reach a specified goal are not just limited to advertisements. It has also reached the scientific research field. According to research (Wang MQ, 2018), 1 in 4 Statisticians admit they were asked to commit scientific fraud by manipulating statistics.[2] This number is based on statisticians which admitted to questionable research practices, Actual number could be higher than this.

1 (Berkeley admissions data)
2 (Wang MQ, 2018)

Table 1. Biostatician-Reported Frequency and Severity Rating of Requests for Inappropriate Analysis and Reporting (n = 390)*				
Violation Request	**Respondents Rating the Item as "Most Severe," %†**	**Reported Requests During the Past 5 Years, %**		
		0	**1-9**	**≥10**
Falsify the statistical significance (such as the *P* value) to support a desired result	84	97	2	1
Change data to achieve the desired outcome (such as the prevalence rate of cancer or another disease)	84	93	7	-
Remove or alter some data records (observations) to better support the research hypothesis	80	76	22	2
Interpret the statistical findings on the basis of expectations, not the actual results	68	70	28	2
Do not fully describe the treatment under study because protocol was not exactly followed	62	85	15	-
Do not report the presence of key missing data that could bias the results	68	76	23	1
Ignore violations of assumptions because results may change to negative	64	71	28	1
Modify a measurement scale to achieve some desired results rather than adhering to the original scale as validated	55	79	20	1
Report power on the basis of a post hoc calculation, but make it seem like an a priori statement	54	76	23	2
Request to not properly adjust for multiple testing when "a priori, originally planned secondary outcomes" are shifted to an "a posteriori primary outcome status"	56	80	18	2
Conduct too many post hoc tests, but purposefully do not adjust α levels to make results look more impressive than they really are	54	60	36	4
Remove categories of a variable to report more favorable results	48	68	31	1
Do not mention interim analyses to avoid "too much testing"	50	81	18	1
Report results before data have been cleaned and validated	48	56	39	5
Do not discuss the duration of follow-up because it was inconsistent	45	84	15	1
Stress only the significant findings, but underreport nonsignificant ones	42	45	48	7
Do not report the model statistics (including effect size in ANOVA or R^2 in linear regression) because they seemed too small to indicate any meaningful changes	42	76	23	1
Do not show plot because it did not show as strong an effect as you had hoped	33	58	39	3

ANOVA = analysis of variance.
* Based on findings from questions 1-18 of the Bioethical Issues in Biostatistical Consulting Questionnaire, which asked biostatisticians "to estimate the number of times—during the past 5 years—that you, personally, have been DIRECTLY asked to do this." Data are presented in decreasing order by the percentage of respondents with a perceived severity score of 4 or 5.
† Items were defined as "most severe" if respondents ranked the severity as 4 or 5 on a scale of 0-5.

Table 2: List of Questions about scientific data manipulation

If we look at the questions from top to bottom, the severity of questions decreases. Some of the questions included "falsifying the statistical significance to support a desired result" which means producing bogus statistics, and "remove or alter some data records to better support the research hypothesis" which basically means cherry picking the data to suit the hypothesis.

"Other malpractices include modifying results to improve the outcome, questionable interpretation of data, withholding methodological or analytical details, dropping observations or data points from analyses and deceptive or misleading report of design, data or results."[1] (Fanelli, 2009)

1 (Fanelli, 2009)

Overcoming faulty statistics:

Computer software and the Internet have aggravated the complexity with which data can be analyzed, but it also enables us to cross-check and verify them easily.

At first glance a graph matches our hypothesis, however, if we look deeply we begin to see faults in it. To figure out if a graph is misleading we need to check the scale of the X and Y axis and look for any inconsistency and at the context of the graph.

Another important factor to take into consideration is the sample size of statistics or the survey upon which statistics are based. The small sample size also makes it easy to create a false equivalence. Statistics with small sample size are considered less accurate and statistics with large sample size are considered more accurate. If we interview 10-15 people out of 1000 and then frame statistics based on that it won't give us the whole picture. We must consider how data was collected, under which conditions and the audience based on whom the data were created. See Confirmation Bias and Sampling Bias.

Recently, More than 800 researchers at the Nature journal urged scientists to ditch the notion of statistical significance. They argued that statistical significance can be misleading because it sets an arbitrary threshold on the level of uncertainty science should be willing to accept.[2] (Valentin Amrhein, 2019)

Uncertainty is expressed as the likelihood of observing an experimental result by chance, assuming the effect being tested doesn't actually exist. In statistical language, this likelihood is known as the p-value. Statistical significance typically requires the p-value of less than 5 percent, or 0.05.

Conclusion:

After studying the different methods of representing statistics and different ways to manipulate, it is of little wonder that statistics are now not taken at its face value and people have correctly started to scrutinize it. We shouldn't take all the statistics at their face value and properly analyze and scrutinize them. We need to remember that Statistics aren't created themselves, they are created by people. With a purpose unknown to us. So to rely on them 100% and believe them to be infallible isn't a very good idea.

2 (Valentin Amrhein, 2019)

Bibliography

(n.d.). Berkeley admissions data. Retrieved April 10, 2019, from http://www.randomservices.org/random/data/Berkeley.html

Carlson, B. W. (2019, January 07). Simpson's paradox. Encyclopædia Britannica. Retrieved from https://www.britannica.com/topic/Simpsons-paradox

(1999). Ethical Guidelines for Statistical Practice. American Statistical Association.

Fanelli, D. (2009, May 29). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. (T. Tregenza, Ed.) doi:10.1371/journal.pone.0005738

John S. Gardenier, D. B. (2002). The misuse of statistics: concepts, tools, and a research agenda. Accountability in Research. doi:10.1080/08989620212968

Valentin Amrhein, S. G. (2019, March 20). Scientists rise up against statistical significance. Retrieved from https://www.nature.com/articles/d41586-019-00857-9

Wang MQ, Y. A. (2018, October 16). Researcher Requests for Inappropriate Analysis and Reporting: A U.S. Survey of Consulting Biostatisticians. Ann Intern Med. doi:10.7326/M18-1230

Wolf, A. (2007, November 1). Misleading Statistical Studies. Policy Quarterly, 3(4). doi:10.26686/pq.v3i4.4231