STATISTICS IN TRANSITION new series, September 2025
Vol. 26, No. 3, pp. xx–xx, https://doi.org/10.59139/stattrans-2025-xxx
Received – xx.xx.xxxx; accepted – xx.xx.xxxxx

The perils of premature evaluation: reassessing the application of Benford's Law to the USA's COVID-19 data

Ravi Dutta-Powell[1]

Abstract

This paper reviews earlier applications of Benford's Law to the COVID-19 data in the United States that claimed these data's non-conformity with Benford's Law, and uses later and more granular data to demonstrate that this was likely due to the earlier data being unsuitable for such applications. It also demonstrates that the same dataset, analyzed in different ways, can show vastly different levels of conformity with Benford's Law. Specifically, most US states show high degrees of conformity for the COVID-19 cases and cumulative deaths when the Robust Order of Magnitude (ROM) is over 3 and data at the county level is used to analyze state outcomes. Conversely, when the county data is aggregated to the state level and analyzed (i.e. case totals for all counties are summed to create a single state figure for each day of the pandemic), every state shows non-conformity. Only new deaths showed the reverse pattern - this is likely because new deaths at the county level do not span sufficient orders of magnitude, and aggregation to the state level overcomes this. This suggests that some instances of non-conformity with Benford's Law in the literature may be caused by its applications to inappropriate datasets or methodological issues

Key words: Benford's Law, COVID-19 data.

1. Introduction

Since 2020, the COVID-19 pandemic has impacted nearly every aspect of human existence, with over three quarters of a billion cases and over 7 million deaths reported to the World Health Organisation (World Health Organisation, 2024). COVID-19 data is key for understanding the pandemic, both in terms of measuring its impact but also in terms of developing strategies to mitigate its spread. However, since the early days of the pandemic, questions have been raised with respect to the quality of the data that is being reported (Campolieti, 2022; Sambridge & Jackson, 2020). There have been suggestions of undercounting and misreporting of data, which has persisted through to the present day (Neumayer & Plümper, 2022). One common tool used to identify potential anomalies is Benford's Law (Benford, 1938; Nigrini, 2012), which posits that

---

[1] Independent researcher. L12 309 Kent St Sydney NSW 2000, Australia. E-mail: ravi.dutta.powell@gmail.com. ORCID: https://orcid.org/0000-0003-4197-3477.

the first digits of a sufficiently large and naturally generated dataset will conform to a particular frequency.

Benford's Law has previously been applied to the H1N1 influenza epidemic (Idrovo et al., 2011), and has seen renewed interest during the COVID-19 pandemic. Early analysis using Benford's Law yielded mixed results – for example, some analysis suggested broad conformity (Sambridge & Jackson, 2020), whilst others highlighted nonconformity, but noted this was consistent across a range of countries, and did not necessarily suggest data manipulation (Koch & Okamura, 2020). Farhadi (2021) conducted an analysis covering a period up to September 2020, and found that only a minority of countries adhered well to Benford's Law, whilst Isea (2020) found that Italy, Portugal, Netherlands, United Kingdom, Denmark, Belgium and Chile all did not conform to Benford's Law based on data up until April 2020. Campolieti (2022) found that nearly every US state up until June 2020, as well as New York City and the District of Columbia, deviated substantially from Benford's Law with respect to new daily cases, although as will be demonstrated this is likely due to the data not providing a sufficiently large sample size and not covering sufficient orders of magnitude – two key prerequisites for conformity with Benford's Law. Similarly, early data for countries in the European Union found a wide variance in conformity, with the perhaps counterintuitive result that countries with higher vaccination rates saw deviations from Benford's law (Kolias, 2022). Neumayer and Plumper (2022) found that autocratic regimes appear to have lower COVID-19 mortality rates on the surface, but in fact an analysis of excess mortality suggests that this is largely due to data manipulation.

Later work has shown more widespread conformity, as more data has emerged. Farhadi and Lahooti (2021) showed that as time had progressed, more countries showed greater conformity to Benford's Law, and two years in most countries showed conformity with Benford's Law (Farhadi & Lahooti, 2022a). Similarly, Campanelli (2023) found that a majority of countries showed conformity with Benford's Law when using the Euclidean Distance statistic, with only a minority showing significant non-conformity. Others have shown that conformity with Benford's Law is often positively correlated with indices of development (Balashov et al., 2021), or with indices of democratic freedom (Kilani, 2021). In the US, research has suggested that partisan bias may have led to data manipulation and misreporting of data at the county level (Eutsler et al., 2023).

In this paper, we review data for the United States of America (henceforth the US), covering the full period of the pandemic (January 2020 – March 2024), and revisit previous analyses. Section 2 below outlines the data sources used and the methodology, where we identify the key requirements of Benford's Law and appropriate measures. Section 3 covers how this approach is then applied to the various sources of US data, where we identify potential shortcomings of previous applications. We also demonstrate multiple ways of analyzing the same dataset using Benford's Law, and demonstrate how different approaches can lead to radically different results. Section 4 provides a discussion of the results and our conclusions.

2. Data and methods

This paper uses two datasets to assess cases and deaths due to COVID-19 – the Centre for Disease Control's (CDC) dataset on daily cases and deaths at a state level, and the New York Times (NYT) COVID-19 cases and deaths tracker, available at https://raw.githubusercontent.com/nytimes/covid-19-data/master/us-counties.csv. CDC data for the period 6 March 2020 to 5 August 2020 was originally used by Campolieti (2022), and a more recent update of the data has also been included. The New York Times data covers slightly over three years, and has been used by several authors more recently (Eutsler et al., 2023; Rocha Filho et al., 2023). The datasets differ in terms of their coverage and figures – the CDC has its own collection methods, provides only state level data, and provides data on new cases and deaths, whilst the New York Times sources data directly from states and counties, provides county level data, and provides the cumulative tally of cases and deaths. All data and analysis code is available at https://osf.io/j2s9d/.

One potential shortcoming of cumulative data, particularly with respect to the spread of an infectious disease, is that the data can have "plateaus" where there are few new deaths or cases. The cumulative total can then be "stuck" in a particular numerical range, which may lead to certain leading digits being overrepresented and thus lead to nonconformity with Benford's Law. An alternative to cumulative data is thus to generate daily changes, which may show greater conformity (daily change data can easily be calculated from cumulative data, and vice-versa). Copies of the datasets used, and all analysis code, are available at https://osf.io/592bu/.

To assess whether data complies with Benford's Law, there are several criteria that must be met: the data should have a large sample size, cover several orders of magnitude, have a positively skewed distribution, and be the result of some natural or semi-random process (Goodman, 2016). The first two criteria are particularly notable, as whilst there is agreement on the overall principles, there are a range of definitions for acceptable thresholds. A common threshold for minimum sample size of 1,000 is often cited (Farhadi, 2021; Nigrini, 2012), but others have suggested lower thresholds of 500 (Cerqueti & Provenzano, 2023). Regardless, it is likely that there is no hard and fast rule, and that some smaller samples will show conformity whilst some other, larger samples will not. Similarly, there is no definitive answer for how many orders of magnitude (OOM) a dataset must span to ensure compliance, although it is generally agreed that there should be at least 3 (Farhadi & Lahooti, 2022b; Fewster, 2009; Kossovsky, 2021). Notably, Kossovsky (2021) suggests using the "Robust Order of Magnitude" (ROM), which has the advantage of excluding any outliers, and requiring the ROM to exceed at least 3 to deem a dataset suitable for Benford analysis (though it may be possible to observe conformity at values below 3). The ROM is defined as:

$$\text{ROM} = Log_{10}(P_{99\%} \,/\, P_{1\%})$$

where $P_{99\%}$ and $P_{1\%}$ represent the 99th percentile and 1st percentile of the data, respectively.

There are a range of methods for assessing compliance with Benford's Law, but research has identified issues with some. For example, some research has used Pearson's x^2 test as a prominent method for evaluating the conformity of data with Benford's Law (Campolieti, 2022). However, the x^2 test is limited, in ways that are particularly problematic for use with data expected to conform with Benford's Law – larger datasets are more likely to show conformity with Benford's Law, whereas the χ^2 test is more likely to show non-conformity as the sample size increases, making it almost paradoxically unsuitable for evaluating conformity with Benford's Law (Cerqueti & Lupi, 2023; Koch & Okamura, 2020; Kossovsky, 2021; Nigrini, 2012). Notably, the χ^2 test also typically has a minimum sample size requirement, with each potential "cell" (in this case, the 9 potential values for the frequency of each leading digit) required to have a value of 5 or more (McHugh, 2013). For a test of Benford's Law, where there are 9 potential cells, this implies a minimum sample size of 45.

A more common alternative is the Mean Absolute Deviation or MAD (Kolias, 2022; Kossovsky, 2021; Nigrini, 2012), which is not sensitive to sample sizes, although recent research suggests that at very large sizes it can also have limitations (Cerqueti & Lupi, 2023). Instead of significance testing, a series of thresholds have been to classify the levels of conformity: close conformity ($\text{MAD} < 0.006$), acceptable conformity ($0.006 < \text{MAD} < 0.012$), marginally acceptable conformity ($0.012 < \text{MAD} < 0.015$), and nonconformity ($\text{MAD} > 0.015$) (Nigrini, 2012). Note these are not the only thresholds for MAD – earlier research has used lower thresholds (Drake & Nigrini, 2000), however more recently the consensus in the literature has coalesced around the thresholds listed (see, for example, Campolieti, 2022; Cerqueti & Provenzano, 2023; Kolias, 2022, Kossovsky, 2021). As such, to ensure consistency with the broader literature and specific works cited, we used these thresholds. MAD is defined as:

$$\text{MAD} = \frac{1}{9}\sum_{i=1}^{9} |Ex - Obs|$$

where *Ex* is the expected proportion for any given leading digit, and *Obs* is the actual observed proportion for that leading digit.

More recently, the Sum of Squared Deviations (SSD) has emerged as a more scale-agnostic alternative to MAD (Kossovsky, 2021). Compared to MAD, it also treats fewer, larger deviations from Benford's Law more severely than several smaller deviations - the latter of which is much more likely to occur naturally due to random chance (Kossovsky, 2021). SSD is defined as

$$\text{SSD} = \sum_{i=1}^{9} (Ex - Obs)^2 \times 10^4$$

Hence this paper will apply both SSD and MAD, but with preference given to the findings of SSD, given its simplicity of not dividing deviations by the number of digits involved, especially in cases of a large base in a number system. Similarly, datasets will note whether they have sufficient sample size ($N > 1{,}000$), and have sufficient robust orders of magnitude (i.e. ROM > 3).

3. Results

3.1. Early analysis

We first review the data on daily deaths at the state level in the US, initially analyzed in the early period of the COVID-19 pandemic (Campolieti, 2022) to evaluate whether it meets the minimum requirements for potential conformity with Benford's Law. Table S1 in the supplementary materials provides a summary of each state, as well as Washington DC and New York City, along with brief summary statistics of the total sample, range, and OOM/ROM measures.

Concerningly, none of the states even comes close to the minimum N of 1,000 – every state has an $N < 150$ (full details supplementary materials). In fact, 5 states have $N < 45$, making the use of x^2 inappropriate. Similarly, no state has a ROM over 3 – only 3 states exceed even a ROM of 2.5. In fact, several states have data that *does not even cover a single order of magnitude.* Thus, it is perhaps unsurprising that Campolieti (2022) previously found widespread nonconformity with Benford's Law – in many cases, the data was such that it was not even mathematically possible for conformity to exist. It is noteworthy that other research around this time also found deviations from Benford's Law in the US for some time periods (Koch & Okamura, 2020).

3.2. Later data

Subsequently, as the pandemic progressed, more data has become available. We now turn to an analysis of CDC data covering a longer period (until October 2022). For deaths, no states had a $N > 1{,}000$ (although some came close, with over $N > 900$), and no states had a ROM of over 3 - just three states and New York City had a ROM above even 2.5, making deaths likely unsuitable for Benford's Law (see supplementary materials for full details).

Similarly, no state had cases where $N > 1{,}000$, but 22 had a ROM of over 3. Table 1 shows a summary of the results, with full results for all states in the supplementary materials. For states with a ROM over 3, 12 showed nonconformity on MAD, but just two of them showed nonconformity on SSD. For the 30 states with a ROM below 3, the results were different – 5 states showed nonconformity on SSD and 18 states showed nonconformity on MAD. Of note, states that showed nonconformity on SSD were a subset of those who showed nonconformity on MAD.

Table 1: Summary of results for states from CDC data (cases)

Specification	MAD		SSD	
Subset	Conforming	Nonconforming	Conforming	Nonconforming
ROM > 3	10	12	20	2
ROM < 3	12	18	25	5
Total	22	30	45	7

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data are contained in supplementary materials.

However, when combining all states and analyzing the US as a whole, leading to N = 42,889 and ROM = 3.879, the results show very close conformity with Benford's Law (MAD = 0.002, SSD = 0.608).

3.3. County level data

Finally, we consider the NYT county-level data, covering a period up to March 2023. This allows for data from each county in a state to be pooled together, resulting in substantially larger sample sizes for each state of well over N > 1,000, with N > 10,000 for the vast majority of states. All states also have full coverage of first digits in the lowest order of magnitude, meaning that any non-conformity is not due to a lack of coverage of digits in the lowest order of magnitude (Goodman, 2023). The data is provided as a cumulative daily tally of all cases and deaths, by county.

3.3.1. Couty level data – deaths

Table 2: Summary of results for states from NYT data (cumulative daily deaths, by county)

Specification	MAD		SSD	
Subset	Conforming	Nonconforming	Conforming	Nonconforming
ROM > 3	23	8	28	3
ROM < 3	7	12	13	6
Total	30	20	41	9

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data is contained in supplementary materials.

Table 2 summarizes the results for cumulative daily deaths by county, analyzed at the state level (i.e. each county's daily figure is treated as a separate observation, and combined with other county-day observations within the state to create the total dataset). Thirty-one states show a ROM of over 3. Of these, just 3 states show nonconformity on SSD and MAD (Connecticut, Massachusetts and Hawaii), with 5 further states (Nevada, Oregon, Louisiana, Ohio and Kansas) showing nonconformity on MAD only. Of note, Connecticut, Massachusetts, and Hawaii are the three smallest states by *N* in this group.

For completeness, we also review the 19 states with a ROM below 3. Of these, 6 show non-conformity on SSD and MAD, whilst a further 6 show nonconformity on MAD only. Notably, just one state (Kentucky) shows close conformity on MAD (none do so on SSD) - interestingly, it has the second highest *N* of this group.

Notably, the US as a whole (N = 3,162,163, ROM = 3.436) shows strong conformity with Benford's Law (MAD = 0.003, close conformity; SSD = 2.35, acceptable).

Another option for analyzing this information is to aggregate the daily county-level observations into a single, state-level observation. That is, for each given date, we can sum all observations across every county in each state, to get a single figure for cumulative deaths by state. This does reduce the overall sample size for each state, but each state still has $N > 1{,}000$. Strikingly, every single state shows nonconformity on MAD and SSD using this analysis, suggesting that this approach may not be suitable. However, once again, the US as a whole (N = 54,976, ROM = 4.052) shows acceptable conformity (MAD = 0.0197, SSD = 16.97).

Alternatively, instead of cumulative deaths, we can analyze new daily deaths, again using individual county-level and then aggregated state-level data. Perhaps unsurprisingly, county-level new death data does not show conformity for any states - this is to be expected, however, as despite most states have sizeable N, they all have limited ROMs. Not a single state had a ROM over 3 (few states even had a raw OOM of over 3), and several states barely covered even a single order of magnitude. Hence in this situation, it may be appropriate to aggregate county-level data to state level, as it will increase the ROM for most states, even though it will reduce the *N* substantially.

Table 3: Summary of results for states from New York Times data (new daily deaths, aggregated)

Specification	MAD		SSD	
	Conforming	Nonconforming	Conforming	Nonconforming
All states	28	22	44	6

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data is contained in supplementary materials.

Surprisingly, despite this leading to $N < 1{,}000$ for nearly every state, and no state having a ROM over 3, there was a moderate degree of conformity (Table 3). It should be noted that the ROMs for states were generally higher in aggregated new deaths scenario, as compared to county-level new deaths, which may have contributed to this result.

3.3.2. Cases

For cumulative cases, all states have very high *Ns* and high ROMs. This analysis shows just 2 states with a non-conforming SSD for cases – Connecticut and Hawaii, although Delaware is also very close to the threshold (Table 4). Notably, these three states represent the three of the four smallest by *N*, suggesting that again a very large sample size can be required to find a high degree of compliance with Benford's Law. Under MAD, these three states are also non-conforming, along with a further 8 - Rhode Island, New Hampshire, Vermont, Arizona, Massachusetts, Nevada, New Jersey, and Iowa. However, 10 of these states are in the bottom 11 by *N*, again suggesting that even large *Ns* may not always show conformity. The outlier is Iowa. One potential explanation for Iowa's nonconformity is the fact that they are a particularly rural state, with a large number of sparsely populated counties. Hence, these individual counties are likely to have a large number of relatively small numbers of cases, leading to fairly stable cumulative case counts and potentially causing non-compliance with Benford's Law.

Table 4: Summary of results for states from New York Times data (cumulative daily cases, county-level)

Specification	MAD		SSD	
	Conforming	Nonconforming	Conforming	Nonconforming
All states	39	11	48	2

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data is contained in supplementary materials.

As with cumulative deaths above, we also try aggregating cumulative case data at the county-level to generate daily state-level observations. Strikingly, once again every single state shows nonconformity, despite all states having $N > 1{,}000$ and ROM of well over 3 (most are well over 3.5). However, despite this, the US as a whole ($N = 55{,}876$, ROM = 5.82) shows acceptable conformity with Benford's Law (MAD = 0.0063, SSD = 6.244).

Table 5: Summary of results for states from NYT data (new daily cases, by county)

Specification	MAD		SSD	
Subset	Conforming	Nonconforming	Conforming	Nonconforming
ROM > 3	14	1	15	0
ROM < 3	27	8	33	2
Total	41	9	48	2

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data is contained in supplementary materials.

When considering new cases, 15 states have a ROM above 3, with all showing conformity on SSD and just one (Texas) showing nonconformity on MAD (Table 5). Even states with a ROM below 3 show fairly high levels of conformity - just two states show nonconformity on SSD, and 8 on MAD. And, as with other results, the US as a whole (N = 1,764,406, ROM = 2.914) shows acceptable conformity (MAD = 0.0105. SSD = 19.39).

Finally, new cases can also be aggregated at the state level. In line with most of the other results, states generally have greater nonconformity when considering aggregated data, as compared to when county-level data remains disaggregated (Table 6).

Table 6: Summary of results for states from NYT data (new cases, aggregated)

Specification	MAD		SSD	
Subset	Conforming	Nonconforming	Conforming	Nonconforming
ROM > 3	15	17	32	2
ROM < 3	11	7	16	2
Total	26	24	48	2

MAD nonconformity: MAD > 0.015. SSD nonconformity: SSD > 100. Full data is contained in supplementary materials.

4. Discussion and Conclusion

The results of this work show how different views of the same data can lead to very different results, and the importance of selecting an appropriate sample for analysis of conformity with Benford's Law. Based on the results, it seems that if data is available at different levels (e.g., state and national, or state and county), then the level of data below the level of interest should be used if it provides sufficient orders of magnitude (ideally 3, but it is possible to show conformity below this). That is, state level data should be used to calculate national level results, and county level data should be used to calculate state level results. Data should not be aggregated and summed together to a higher level (for example, summing daily county-level observations to get a single state level observation for each day), unless absolutely necessary. In general, when each county-day observation was treated as its own datapoint, most states showed more conformity with Benford's Law. In contrast, when observations for counties were summed each day to produce a single, aggregated daily figure for states, conformity was much lower. Perhaps most strikingly, cumulative deaths and cumulative cases saw the majority of states showing conformity when using county-level data, but literally every state showed nonconformity on both cases and deaths when analyzing aggregate data (despite the US as a whole also showing conformity).

However, there was one exception to this, namely the NYT data on new deaths. In this case, county level data does not cover sufficient orders of magnitude – as such, aggregating several observations together can make sense and may in fact be necessary. When county-day observations for new deaths (which spanned fewer orders of magnitude than new cases, or cumulative cases/deaths) were treated as individual data points, no state showed conformity. This was expected, as no state had a ROM of over 3 and in fact several had a ROM below 1. However, when new deaths per county were summed to provide a single, daily state figure, many states actually showed some degree of conformity, despite nearly every state having $N < 1{,}000$ and ROM < 3.

Overall, this highlights the importance of large sample sizes and orders of magnitude, particularly the ROM measure. Some datasets with very large *Ns* (e.g., new daily deaths at the county level) but low orders of magnitude show high nonconformity, whilst others with high ROMs but relatively lower *Ns* (e.g., cumulative cases, aggregated to the state level) also showed nonconformity. Whilst there were some cases where despite a low ROM and a low N, some states did show conformity with Benford's Law, these were in the minority. Indeed, just because some data with low *N* and limited ROM complies with Benford's Law does not mean that we should expect all data to comply (Goodman, 2023). Of note, whenever states were divided into groups based on a ROM above or below 3, there was typically greater conformity in the group with ROMs above 3, suggesting that the thresholds for *N* and ROM, whilst important, are more in the nature of guidelines than hard and fast rules.

This highlights the importance of a comparative approach, using multiple, similar datasets to test conformity, and using the most appropriate level of analysis where multiple levels are available. The results of this work confirm more recent work that has found broad conformity with Benford's Law for the US as a whole and for most states (Rocha Filho et al., 2023). It fits with the broader trend of early indications of non-conformity, but later data showing clear conformity (Farhadi & Lahooti, 2021, 2022a). It reinforces the need to ensure that data has the potential to actually be compliant with Benford's Law, before testing for compliance (Ausloos et al., 2021), and indeed helps to identify why earlier analysis was more prone to finding deviations from Benford's Law. It can be tempting to try and find sources of non-compliance, and forget the underlying prerequisites of Benford's Law. It is likely that a thorough review of previous work of Benford's Law, particularly from early in the COVID-19 pandemic but potentially in a range of fields, would likely find many such examples of its application to datasets that were not at all suitable, resulting in spurious claims of nonconformity.

References

Ausloos, M., Ficcadenti, V., Dhesi, G. and Shakeel, M., (2021). Benford's laws tests on S&P500 daily closing values and the corresponding daily log-returns both point to huge non-conformity. *Physica A: Statistical Mechanics and Its Applications*, 574, 125969. https://doi.org/10.1016/j.physa.2021.125969.

Balashov, V. S., Yan, Y. and Zhu, X., (2021). Using the Newcomb–Benford law to study the association between a country's COVID-19 reporting accuracy and its development. *Scientific Reports*, 11(1), 22914. https://doi.org/10.1038/s41598-021-02367-z.

Benford, F. R., (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*, *78*(4), pp. 551–572.

Campanelli, L., (2023). Breaking Benford's law: a statistical analysis of COVID-19 data using the Euclidean distance statistic. *Statistics in Transition new series*, 24(2), pp. 201–215. https://doi.org/10.59170/stattrans-2023-028.

Campolieti, M., (2022). COVID-19 deaths in the USA: Benford's law and under-reporting. *Journal of Public Health*, 44(2), e268–e271. https://doi.org/10.1093/pubmed/fdab161.

Cerqueti, R., Lupi, C., (2023). Severe testing of Benford's law. *Test*, 32(2), pp. 677–694. https://doi.org/10.1007/.

Cerqueti, R., Provenzano, D., (2023). Benford's Law for economic data reliability: The case of tourism flows in Sicily. *Chaos, Solitons & Fractals*, 173, 113635. https://doi.org/10.1016/j.chaos.2023.113635.

Drake, P. D., Nigrini, M. J., (2000). Computer assisted analytical procedures using Benford's Law. *Journal of Accounting Education*, 18(2), pp. 127–146. https://doi.org/doi:10.1016/s0748-5751(00)00008-7.

Eutsler, J., Kathleen Harris, M., Tyler Williams, L. and Cornejo, O. E., (2023). Accounting for partisanship and politicization: Employing Benford's Law to examine misreporting of COVID-19 infection cases and deaths in the United States. *Accounting, Organizations and Society*, 108, 101455. https://doi.org/10.1016/j.aos.2023.101455.

Farhadi, N., (2021). Can we rely on COVID-19 data? An assessment of data from over 200 countries worldwide. *Science Progress*, 104(2), 00368504211021232. https://doi.org/10.1177/00368504211021232.

Farhadi, N., Lahooti, H., (2021). Are COVID-19 Data Reliable? A Quantitative Analysis of Pandemic Data from 182 Countries. *COVID*, 1(1), Article 1. https://doi.org/10.3390/covid1010013.

Farhadi, N., Lahooti, H., (2022a). Forensic Analysis of COVID-19 Data from 198 Countries Two Years after the Pandemic Outbreak. *COVID*, 2(4), Article 4. https://doi.org/10.3390/covid2040034.

Farhadi, N., Lahooti, H., (2022b). Reply to Morillas-Jurado et al. Benford Law to Monitor COVID-19 Registration Data. Comment on "Farhadi, N.; Lahooti,

H. Forensic Analysis of COVID-19 Data from 198 Countries Two Years after the Pandemic Outbreak. COVID 2022, 2, pp. 472–484”. *COVID*, 2(7), Article 7. https://doi.org/10.3390/covid2070070.

Fewster, R. M., (2009). A Simple Explanation of Benford’s Law. *The American Statistician*, *63*(1), pp. 26–32. https://doi.org/10.1198/tast.2009.0005.

Goodman, W., (2016). The promises and pitfalls of Benford’s law. *Significance*, 13(3), pp. 38–41. https://doi.org/10.1111/j.1740-9713.2016.00919.x.

Goodman, W., (2023). Applying and Testing Benford’s Law Are Not the Same. *Spanish Journal of Statistics*, 5(1), pp. 43–53. https://doi.org/10.37830/SJS.2023.1.03.

Idrovo, A. J., Fernández-Niño, J. A., Bojórquez-Chapela, I. and Moreno-Montoya, J., (2011). Performance of public health surveillance systems during the influenza A(H1N1) pandemic in the Americas: Testing a new method based on Benford’s Law. *Epidemiology and Infection*, 139(12), pp. 1827–1834. https://doi.org/10.1017/S095026881100015X.

Isea, R., (2020). How Valid are the Reported Cases of People Infected with Covid-19 in the World? *International Journal of Coronaviruses*, 1, pp. 53–56. https://doi.org/10.14302/issn.2692-1537.ijcv-20-3376.

Kilani, A., (2021). Authoritarian regimes’ propensity to manipulate Covid-19 data: A statistical analysis using Benford’s Law. *Commonwealth & Comparative Politics*, 59(3), pp. 319–333. https://doi.org/10.1080/14662043.2021.1916207.

Koch, C., Okamura, K., (2020). Benford’s Law and COVID-19 reporting. *Economics Letters*, 196, 109573. https://doi.org/10.1016/j.econlet.2020.109573.

Kolias, P., (2022). Applying Benford’s law to COVID-19 data: The case of the European Union. *Journal of Public Health*, 44(2), e221–e226. https://doi.org/10.1093/pubmed/fdac005.

Kossovsky, A. E., (2021). On the Mistaken Use of the Chi-Square Test in Benford’s Law. *Stats*, *4*(2), Article 2. https://doi.org/10.3390/stats4020027.

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, pp. 143–149. https://doi.org/10.11613/BM.2013.018.

Neumayer, E., Plümper, T. (2022). Does ‘Data fudging’ explain the autocratic advantage? Evidence from the gap between Official Covid-19 mortality and excess mortality. *SSM – Population Health*, 19, 101247. https://doi.org/10.1016/j.ssmph.2022.101247.

Nigrini, M. J., (2012). *Benford’s Law: Applications for forensic accounting, auditing, and fraud detection* (Vol. 586). John Wiley & Sons.

Rocha Filho, T. M., Mendes, J. F. F., Lucio, M. L. and Moret, M. A., (2023). COVID-19 data, mitigation policies and Newcomb–Benford law. *Chaos, Solitons & Fractals*, 174. http://arxiv.org/abs/2208.11226.

Sambridge, M., Jackson, A., (2020). National COVID numbers—Benford's law looks for errors. *Nature*, 581(7809), pp. 384–385.

World Health Organisation, (2024). *COVID-19 deaths reported (2024 global)*. WHO COVID-19 Dashboard. https://data.who.int/dashboards/covid19/circulation.