

Article

From Whence Commeth Data Misreporting? A Survey of Benford's Law and Digit Analysis in the Time of the COVID-19 Pandemic

Călin Vâlsan ¹, Andreea-Ionela Puiu ² and Elena Druică ^{2,*}¹ William School of Business, Bishop's University, Sherbrooke, QC J1M 1Z7, Canada; cvalsan@ubishops.ca² Department of Applied Economics and Quantitative Analysis, Faculty of Business and Administration, University of Bucharest, 030018 Bucharest, Romania; ionela-andreea.puiu@faa.unibuc.ro* Correspondence: elena.druica@faa.unibuc.ro

Abstract: We survey the literature on the use of Benford's distribution digit analysis applied to COVID-19 case data reporting. We combine a bibliometric analysis of 32 articles with a survey of their content and findings. In spite of combined efforts from teams of researchers across multiple countries and universities, using large data samples from a multitude of sources, there is no emerging consensus on data misreporting. We believe we are nevertheless able to discern a faint pattern in the segregation of findings. The evidence suggests that studies using very large, aggregate samples and a methodology based on hypothesis testing are marginally more likely to identify significant deviations from Benford's distribution and to attribute this deviation to data tampering. Our results are far from conclusive and should be taken with a very healthy dose of skepticism. Academics and policymakers alike should remain mindful that the misreporting controversy is still far from being settled.

Keywords: COVID-19 data; Benford's Law; statistical tests; MAD; data conformity; research methodology

MSC: 62-07; 62E99; 62F03; 62P25; 92B10



Citation: Vâlsan, C.; Puiu, A.-I.; Druică, E. From Whence Commeth Data Misreporting? A Survey of Benford's Law and Digit Analysis in the Time of the COVID-19 Pandemic. *Mathematics* **2024**, *12*, 2579. <https://doi.org/10.3390/math12162579>

Academic Editors: Michael Pokojov and Andrews T. Anum

Received: 19 June 2024

Revised: 19 August 2024

Accepted: 20 August 2024

Published: 21 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

We survey the literature on the use of digit analysis based on Benford's distribution applied to COVID-19 case data reporting. We perform a bibliometric analysis, followed by a survey of the content.

Benford's Law represents the methodological cornerstone underpinning fiscal fraud detection based on digit analysis [1–4]. It has been noted that natural processes generate numerical data whose digits follow a specific relative frequency, known as the Benford–Newcombe distribution. Leading digits are more likely to be smaller numbers, such as one or two, while the ending digits are increasingly following a uniform distribution. Naturally generated data are compared to the theoretical Benford distribution, and deviations are deemed red flags that trigger further investigation using other approaches including, but not limited to old-fashioned investigative work.

Since the number of cases during a health crisis is also considered to be naturally generated, digit analysis has been applied to detect misreporting in the wake of the COVID-19 pandemic. It is hardly surprising to see the wide interest received in this topic. Whether COVID-19 data have been misreported is an issue that should be treated with the utmost seriousness. The stakes are very high, given the loss of human lives, the worldwide disruption to the world economy, the abrupt decline in the overall mental health of the population, and, most of all, the polarization of public opinion around data reporting accuracy and vaccine effectiveness.

A multitude of papers published in the last four years investigate data reporting in the healthcare sector [5–8], and a significant number of them examine the incidence of

COVID-19 cases [9–11], attempting to determine the prevalence of over- and/or under-reporting [12–14]. No one, to our knowledge, attempts to centralize and summarize these findings into an integrative survey.

A survey of the literature on Benford’s Law applied to COVID data is important because it would enrich both the academic literature and provide valuable insights for policymakers. Depending on the outcome of the survey, local, national, and international healthcare authorities should be able to ascertain the extent to which their future policy decisions can be informed by the existing evidence on the use of Benford analysis with COVID-19 data. At this point, it is not clear if the aforementioned studies provide any evidence of misreporting or merely apply the Benford methodology to COVID-19 data with the caveat of interpreting the results following further investigation. In the event there is an emerging consensus on (mis)reporting, public authorities should revise statistical data collection and reporting procedures, and policymakers should adapt and fine-tune public health policies. Moreover, a survey that documents a compelling consensus should put some of the controversy to rest and provide much-needed clarity.

Our paper attempts to fill this large gap in the academic literature while providing valuable and possibly actionable information to public administrators.

We set out to survey all the papers published in peer-reviewed academic journals dealing with the application of Benford’s Law to COVID-19 data. Our research will answer a number of important questions. First, it will evince the breadth and depth of academic and scientific collaboration pertaining to this type of research on COVID-19 data. Second, it will reveal the range and diversity of the content and methodology, trying to identify patterns, if any, among the type of data, methodology, results, and interpretation of findings. We hypothesize that we might be able to find that our papers are segregated along methodological lines, similar to the pattern observed in the literature on the Hubble cosmological constant. Since we survey papers that are essentially focused on a very specific methodological ecosystem, it is only natural to expect that their conclusions might be driven by variations in the methodology employed. Third, it will hopefully provide insightful new knowledge for academics and inform the actions of policymakers.

There are several things our research is not doing: This research is not concerned with applying Benford’s Law to COVID-19 data, and hence it is not trying to determine on its own whether there is any misreporting of data; it is merely surveying other studies that apply Benford. Moreover, this study is not evaluating the appropriateness of various statistical tests used as part of the Benford methodological toolkit; this paper is merely reporting what other studies contend about the methodology. Last but not least, this paper is not ranking the importance or significance of other findings; it is merely presenting them in a structured manner and attempting to determine whether there is a consensus or a segregation of findings.

While this study is not a meta-analysis per se, it borrows heavily from the methodology associated with meta-studies. And while it is not trying to directly determine if Benford is suitable for analyzing COVID-19 data, it hopes the findings of the papers that make up the object of this analysis might bring enough arguments for a reasonable clarification of this aspect.

2. Materials and Methods

2.1. Data

In the first stage, we run an exhaustive, blanket search on Google Scholar, across all databases and platforms, with a timeframe ending on 1 January 2024. We look for papers containing the terms “Benford” and “Covid” in the keywords, title, or abstract fields. We found 49 papers. We discard unpublished manuscripts, conference presentations, papers posted on platforms such as ResearchGate or similar, papers including pandemic data other than COVID-19, and reprints or duplicates (including papers indexed in more than one database). We keep published, peer-reviewed papers using Benford’s Law applied to COVID-19 cases only.

We decided to include only peer-reviewed papers to ensure the high academic and scientific quality of our sample. The COVID-19 pandemic was a world-wide crisis that produced intense emotions and a unique counterculture. We tried to avoid poorly informed, politically charged, speculative, and/or controversial works that might have been driven by various agendas other than sound and impartial academic inquiry. Our final sample consists of 32 papers, of which 23 are in Web of Science (WoS henceforth) and 9 in Scopus. Considering that the WoS and Scopus databases use distinct procedures to codify their bibliographic records, we prefer to analyze them separately.

2.2. Method

Our research methodology relies on bibliometric and content analysis. These two approaches are complementary and necessary to handle the complexity entailed by the nature of our research question.

Bibliometric analysis designates a rigorous approach that explores large volumes of scientific data, delivering high-impact research [15]. It has been increasingly used in analyzing the COVID-19 pandemic [16–18] with specific references to the economic environment [16,19,20]. The bibliometric analysis method proposes two principal directions of research: performance analysis and science mapping procedures [15]. A performance analysis quantifies the impact of scientific actors by measuring their contributions to the field [21]. Broad science mapping techniques attempt to emphasize the structural and dynamic components of scientific research by quantifying and visualizing their interrelationships [21].

Our research employs performance analysis and science mapping in a complementary way, using tools from the R software environment version 4.3.1: the bibliometrix package [22] and the biblioshiny web application.

Content analysis aims to systematize information pertaining to the use of methodology, the nature and size of data samples, and results, including both quantitative and qualitative aspects. Benford's Law has been applied in various contexts for many years to many types of "naturally generated" data. Over the years, it has become known that the choice of methodology has some impact on the results. For instance, different levels of data aggregation have the potential to generate inconsistent or conflicting statistical results [23,24]. The choice of individual tests, mainly the use of z-scores and/or Pearson's chi-square statistic applied to large data samples, is more likely to show deviation from Benford. To counter this bias, some authors recommend using the mean average deviation (MAD) instead [25,26].

We generate a fairly small number of quantitative and qualitative variables by taking note of the year the study was published, counting the statistical procedures used, classifying the level of sample aggregation, the extent of the documented deviation from Benford's Law, and the type of interpretation provided.

Most of our content analysis relies on descriptive statistics, but we also attempt a couple of basic quantitative methods to evince the presence of segregation, if any, among our main (mostly binary) qualitative variables. Given the constraints associated with the use of a small data sample, we run Fisher's Exact Test by pairing our qualitative variables, and we attempt the use of classification trees to predict the extent/severity of deviation from Benford's Law, the type of interpretation given, and the presence of a methodological caveat in the discussion of results.

We expect to find either some emerging consensus among scholars with respect to the under- or over-reporting of COVID cases, if any, or a discernible segregation of findings, most probably along methodological lines. Why these expectations? If one were to survey the research on the effects of smoking on human health, one would find a broad consensus among researchers that smoking is clearly detrimental to one's health [27]. If, on the other hand, we were to survey the state of the art on the value of the Hubble cosmological constant, we would find the scientific community segregated among methodological lines. Those who measure the Hubble constant using the Cosmic Microwave Background end up

with values that are significantly different from those who measure it using the Cepheid-type Ia supernova cosmic ladder. This issue is known as the Hubble tension problem [28].

We contend we might be able to observe a similar differentiation of results driven by the choice of statistical tests.

3. Results of Bibliometric Analysis

3.1. Performance Analysis

Performance analysis in bibliometric research depends on the contributions of research metrics [29]. In other words, performance analysis permits the summarization of authors, journals, institutional details, publication indicators, and journal citation metrics [15].

3.1.1. Publications Related Metrics

Our final WoS database comprises 23 articles published between 2020 and 2023 in 20 distinct peer-reviewed sources. The most relevant publishing source is the *Journal of Public Health*, with three scientific papers on the “Benford Law and Covid Data” topic, followed by *Biomedica*, with two publications. The most cited article was published in 2020 in the scientific journal *Physica A: Statistical Mechanics and its Applications*. There are 57 authors and five single-authored documents. The mean number of co-authors per document is 2.91, and 17.39% of papers represent international co-authorships.

The Scopus database includes nine articles published between 2021 and 2023. All articles are published in different journals. There are 18 contributors in various co-authorships and three single-authored documents. The average number of authors per document is 2.33, and only one article is the result of international co-authorships. In both databases, most scientific papers were published in 2021.

There appears to be quite a diversity among journals. Most publications appear scattered over a variety of journals, with one notable exception, namely, *The Journal of Public Health*, which is not surprising given the main research theme pertains to COVID-19. Other than that, there are no preferred publications. Table 1 presents descriptive statistics of both databases.

Table 1. Descriptive Statistics.

Database	No. of Articles	No of Publishing Sources	Average Citation per Article	Highest Number of Citations per Article
WoS	23	20	5.783	24
Scopus	9	9	1.33	7

3.1.2. Citation Analysis

The citation analysis method relies on the observation that citations delineate academic networks generated among publications in the wake of cross-citations, that is, when one paper cites another [30]. The first positions are claimed by the following academic journals: *Physica A: Statistical Mechanics and its Applications*, *Journal of Public Health*, and *Economics Letters*. In the Scopus database, the most globally cited scientific journals are *Heritage and Sustainable Development* and *Model Assisted Statistics and Applications*. The most cited local references are Benford’s own work [31], followed by Diekmann in 2007 [32], and Koch in 2021 [33]. In the Scopus database, the most cited references are Benford’s own paper [31], followed by Fewster in 2009 [34]. It appears that most citations refer chiefly to the mechanics of testing deviations from Benford’s Law. Paradoxically, most papers citing Diekmann limit themselves to the testing of the first digit only.

3.1.3. Collaboration Analysis

Collaboration investigation enables an overview of the scientific cooperation and research communities at various aggregation levels [35]. We consider authors and countries as units of investigation.

The country’s scientific production indicator quantifies the frequency of authors’ appearances associated with a given country affiliation. If an article has multiple authors, it is attributed to all the countries of co-authorship [36]. For the WoS database, the results indicate that Brazil, the US, and Colombia contributed the most to the production of Benford Law and COVID-19 data studies (Figure 1). The results are interesting, considering that, in terms of the confirmed number of daily cases, all three countries have a similar profile [37]. In terms of the cumulative number of cases, the US recorded twice as many numbers compared to Brazil and Colombia (the same link as above). Similar values appear to describe the number of deaths in these three countries, both daily and cumulatively. There seems to be no direct collaboration between these three countries. In Europe, Sweden leads the ranking, followed by Portugal (Figure 1). In the case of the Scopus database, Brazil and Malaysia are the most visible countries, followed by the United Kingdom (Figure 1).

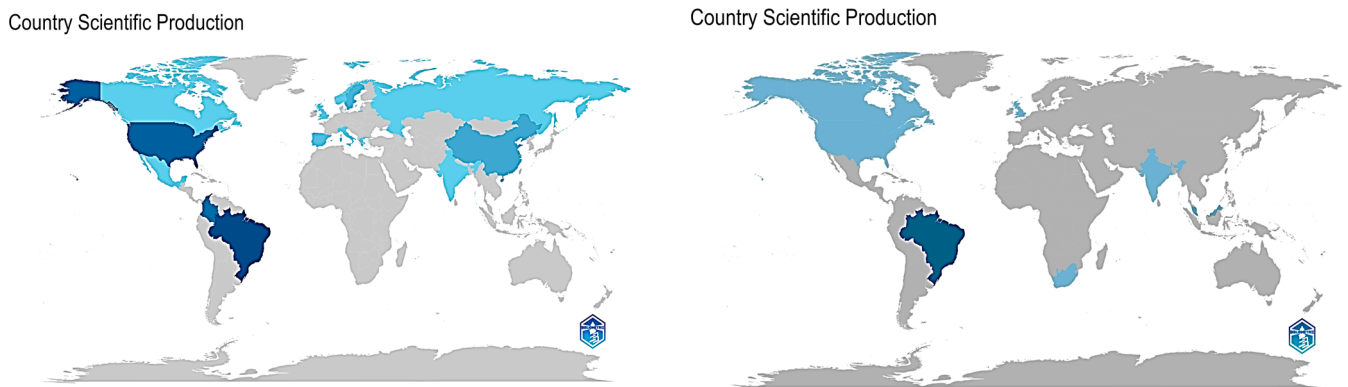


Figure 1. Scientific Production by Country (left—Scopus database; right—WoS database).

In terms of citation performance, the US stands out as the most cited country in the WoS database, followed by Colombia and Brazil. In the Scopus database, the most cited country is Malaysia, followed by Brazil.

Cross-country collaboration analysis reveals that the most salient research partnerships are between Brazil and Portugal, China and Norway, and Sweden and Singapore (WoS). Collaboration between the US and UK was identifiable in both databases (Figure 2).

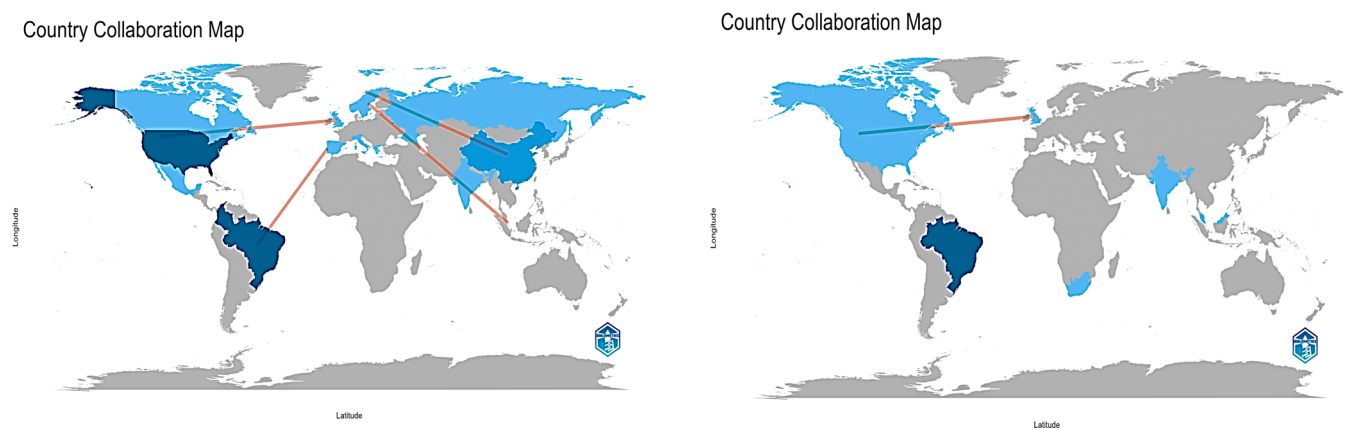


Figure 2. Collaboration World Map by Country (left—Scopus database; right—WoS database).

3.2. Science Mapping

The science mapping technique designates an instrument that develops and employs computational procedures to visualize, measure, and study diverse technical and scientific information [38]. Science mapping procedures incorporate co-citation analysis, bibliographic coupling, co-word investigation, and co-authorship analysis [15], facilitating the evaluation of authors’ and research institutions’ productivity [39]. In our research, science mapping was performed through co-authorship, co-citation, and co-word analysis.

3.2.1. Social Structure—Co-Authorship Analysis

Co-authorship analysis facilitates the identification of collaborative networks by looking at the author’s professional background, research interests, or geographical residence. In the WoS database, we identify twelve collaborative clusters, as presented in Table 2.

Table 2. Co-authorship clusters—WoS database.

Cluster	Country/Countries	Authors’ Research Interest
Cluster 1	Colombia	Public Health
Cluster 2	Brazil	Political Science and Quantitative Methods
Cluster 3	Mexico	Statistics
Cluster 4	Italy	Data Analytics and Engineering
Cluster 5	Brazil	Business Admin and Pharmaceutical Sciences
Cluster 6	Spain	Applied Economics and Corporate Finance
Cluster 7	US	Audit, Accounting, Corporate Governance, and Population Genetics
Cluster 8	US	Causal Inference and Business Analytics
Cluster 9	China and Norway	Data mining, Machine Learning, and Big Data Analysis
Cluster 10	US and UK	Monetary Policy, Macroeconomics, and Financial Stability
Cluster 11	Singapore and Sweden	Public Health and Medical Sciences
Cluster 12	Brazil and Portugal	Complex Systems and Statistical Methods

Likewise, in the Scopus database, we identify five clusters, as presented in Table 3.

Table 3. Co-authorship clusters—Scopus database.

Cluster	Country/Countries	Authors’ Research Interest
Cluster 1	Brazil	Data Analysis and Artificial Intelligence
Cluster 2	US and UK	Healthcare
Cluster 3	Malaysia	No particular pattern emerges
Cluster 4	Brazil	No particular pattern emerges
Cluster 5	India	No particular pattern emerges

The fact that there are very few interdisciplinary clusters appears to reinforce the notion that the interest is mainly academic, focused on applying the Benford methodology to new data. Three appears to be little connection to healthcare workers or policymakers.

3.2.2. References Co-Citation Analysis

The co-citation technique captures the intellectual structure of the field under analysis [40–42]. It describes a procedure that considers articles with similar themes. In a

co-citation network, two papers are connected when they appear in the bibliographical list of a source publication [43]. This type of analysis facilitates the identification of thematic clusters.

The network nodes and the edges are components related to the thematic clusters [44]. The network nodes designate cited papers, and their size is proportional to the number of citations [44,45]. The edges describe co-citation networks. Their weights depend on how frequently two papers have been jointly cited [44,45].

The co-citation analysis performed on the WoS database was carried out with a threshold of 49 nodes. The results indicate a network with two significant and central nodes, “benford f. 1938” [31] (betweenness centrality = 494.83) and “nigrini m.j. 1996” [46] (betweenness centrality = 176.56). The 1938 paper authored by Benford is the first to introduce the expected distributions of significant digits in naturally generated numbers [31]. Nigrini [46] is the first to apply digit frequency analysis to the issue of tax evasion.

The co-citation investigation of the WoS database reveals four clusters (Figure 3). The first cluster (in red) is the largest and densest, with 16 publications. This cluster includes articles from different areas that use Benford’s Law to detect fraudulent data reports. The second cluster (in blue) considers two papers that use Benford’s Law to detect suspicious or fraudulent activities. The third cluster (in green) consists of 17 papers related to applications of Benford’s Law to assess the quality of COVID-19 data reports. The fourth cluster (in purple) comprises 14 papers related to data manipulation procedures. Again, it appears as though the main focus is on the mechanics of statistical tests applied to COVID-19 data.

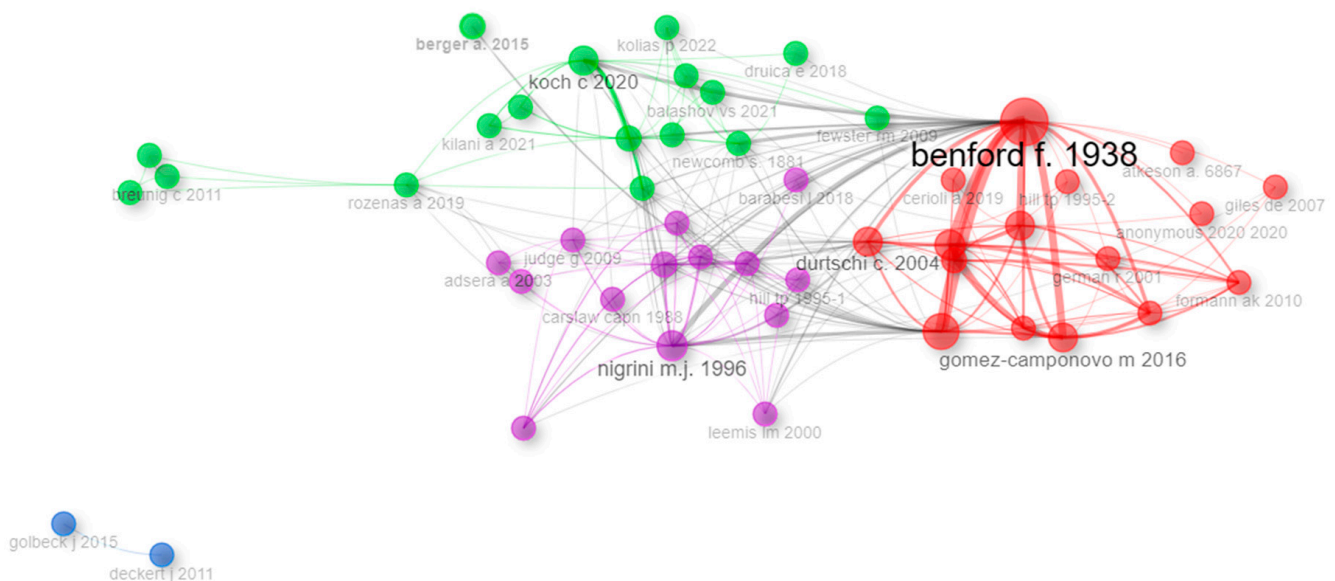


Figure 3. Co-citation analysis of references—WoS database.

The Scopus co-citation analysis reveals the emergence of only one cluster (Figure 4). Similar to the WoS database, “benford f. 1938” [31] (betweenness centrality = 1.892) and “fewster r.m. 2009” (betweenness centrality = 1.108) are the most cited documents from the Scopus database. The thematic cluster identified in the Scopus database encompasses different conceptualizations of contexts, motives, and manners related to the application of Benford’s Law.

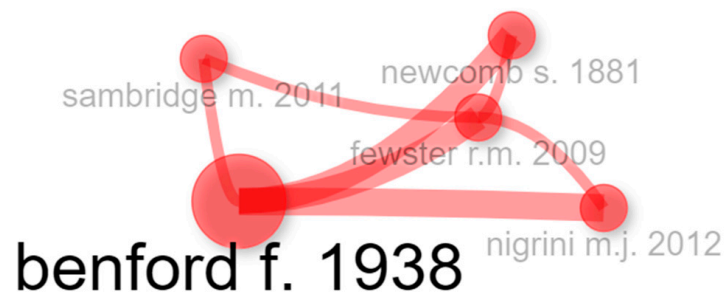


Figure 4. Co-citation analysis of references—Scopus database.

3.2.3. Conceptual Structure—Co-Word Analysis

Co-word analysis represents a method of investigating the co-occurrences of key terms in the keywords, abstracts, or titles of documents, assuming that notions that occur together more frequently represent thematic associations [15]. The co-word analysis of the author’s keywords, paper titles, and abstracts enables mapping the main themes in the field by recognizing and visualizing networks that represent conceptual clusters of various matters discussed in a specific research field [44].

We employ a thematic map and plot the bigram terms from paper abstracts into four clusters based on centrality and density rank scores [47,48]. Centrality quantifies the level of interaction between networks and assesses the level of significance of a concept in the respective field. Density estimates the inner strength of the group and determines the level of development of a subject [47,48]. The cluster size is determined by the frequency notions that it contains and the number of related documents [44]. The cluster name is considered the predominant bigram term.

The thematic map considers four thematic clusters: motor themes, niche themes, emerging/declining themes, and basic themes [47].

The motor themes group comprises well-developed notions in the structure of the domain under analysis [47]. In the WoS database abstracts, we find two motor themes—“Benford’s law” and “Newcomb Benford law”—used interchangeably to express the same concept—the probability distribution associated with the digits of naturally generated numbers.

The niche themes quadrant considers specialized and marginal notions in the overall field [47]. COVID-19 data are identified at the border between niche and motor themes. The emerging or declining themes quadrant considers subjects characterized by low density and centrality [47]. In this quadrant, we display “epidemic growth” and “countries tendency”.

The last quadrant considers the basic themes that are important to the research domain but appear undeveloped. Those themes display high centrality and low density [47]. “Coronavirus disease” is identified as a basic theme.

Figure 5 presents graphically the results of the co-word analysis of the WoS database.

Among Scopus abstracts, “newcomb-benford law” is identified at the border between niche and motor themes. “Daily Covid” is identified as an emerging theme. This is to be expected, considering that daily data lies at the core of the analysis, while “benford law” is at the limit between motor themes and basic themes. The results of the co-word analysis from the Scopus database are presented in Figure 6.

What do we learn from the bibliometric analysis, above and beyond the descriptive aspect? We notice there is very little diversity among concepts. Conceptual themes occupy a relatively narrow range. Key words appear to describe generic notions, lacking conceptual depth and specialization. In fact, it appears there are no niche themes. The general impression is that studies focused on the mechanics of statistical testing the deviation from Benford’s Law on datasets that happen to describe COVID-19 cases. It is as though COVID-19 is merely incidental to the statistical method, driven by a purely academic interest. Bibliometric analysis cannot identify themes focused on practical implications relating to healthcare workers and policymakers. Moreover, there is very little interdisciplinary research.

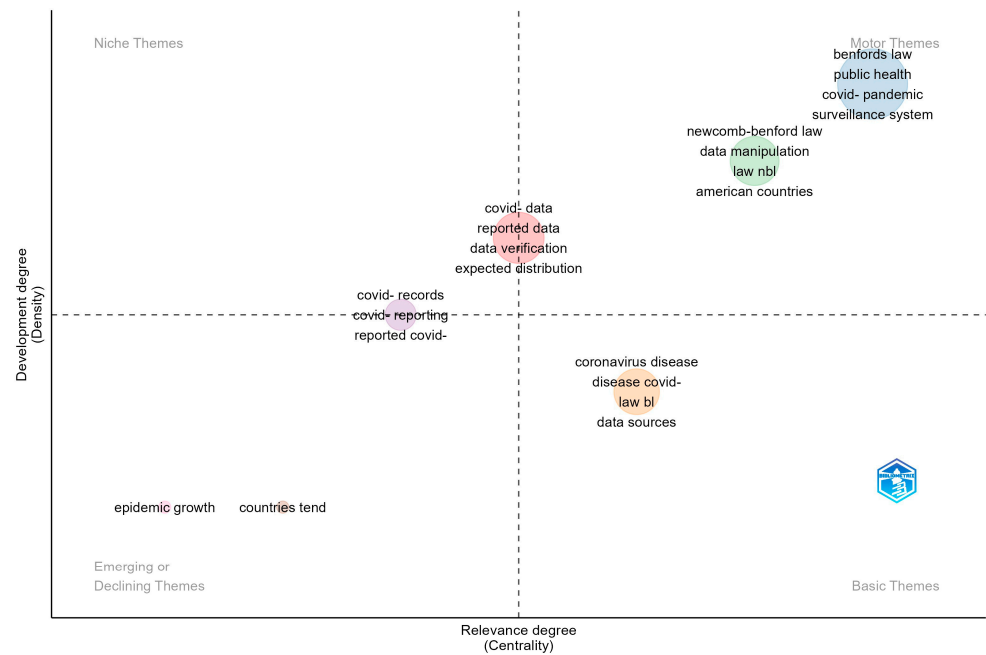


Figure 5. Social structure—co-word analysis WoS database.

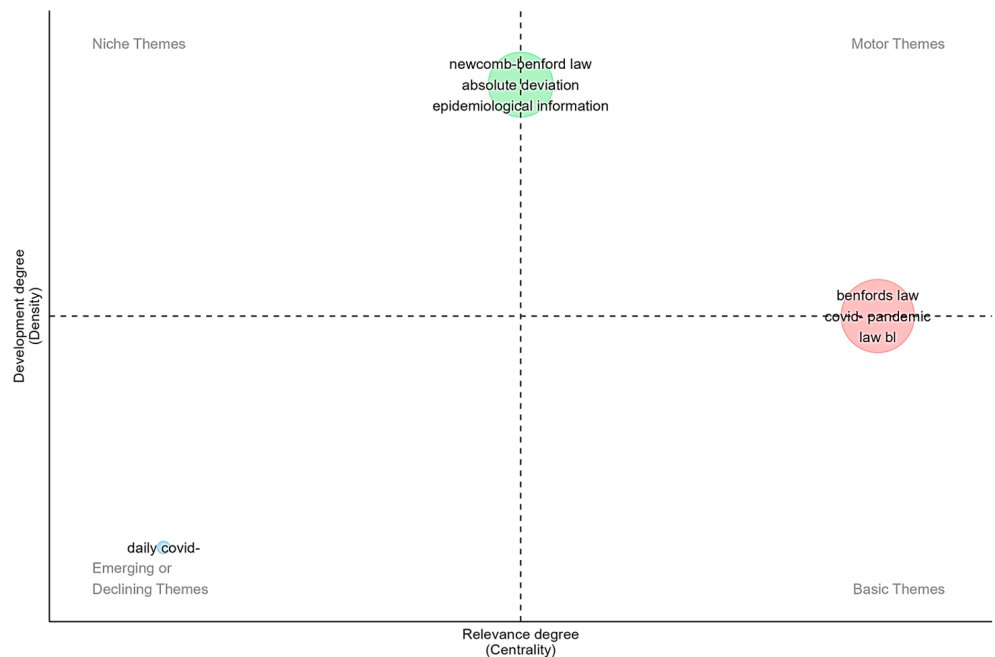


Figure 6. Social structure—co-word analysis Scopus database.

As will be seen later, the situation is more complex. Quite a few papers go beyond the mere technical aspect of statistical testing and attempt an interpretation of results with real-life, practical implications. Some of them even conclude that data tampering is very likely. For a more nuanced understanding of the overall picture, we need to resort to a structured content analysis. To this task, we turn to the next sections.

4. Results of Content Survey and Analysis

In addition to the bibliometric analysis, we conducted a survey of the content, methodology, and findings. Of the initial sample of 32 papers, we kept 25 that have a clearly identifiable research question, methodology, data sample, and results. Twelve were pub-

lished in 2021. Only five papers were published in 2020, four were published in 2022, and another four in 2023.

The year 2021 appears pivotal to both the evolution of the pandemic and our understanding of it. It is somewhat expected to see a relatively small research output on COVID-19 in 2020. The pandemic is still relatively novel, having taken everyone by storm. There is little data available, and the world is still reeling from the lockdown. One should also factor in the impact of the publication cycle. It takes at least a few months to gather data and conduct research; another few months between the time a paper is submitted and the completion of the peer review process; and yet another period between the moment the paper is accepted for publication and the actual publication. Given the traumatic impact of COVID-19 across the world, a lot of resources have been made available to investigate the pandemic, and top priority has been given to medical research related to COVID-19. Still, the publication cycle could not be compressed or shortened too much without compromising quality standards. Hence the relatively modest output of 2020.

The year 2021 represents the peak of COVID-related research. The pandemic is in full swing, vaccines are rolled out, and more data becomes available. This is also when scrutiny is at its maximum, fuelled by public health concerns, growing economic worries, and increasing controversies—often originating on social media and in selected political circles. The publication cycle finally caught up with the pandemic, and we see a significant jump in the research output. The years 2022 and 2023 are turning points. They witnessed the winding down of the pandemic and the relative waning of interest in COVID-19. The beginning of 2022 is marred by the Russian invasion of Ukraine. As a matter of fact, the war in Ukraine has totally upstaged the pandemic, and COVID-19 is relegated to the status of second-rate news. It is also the case that the virus appears to have mutated into a more infectious yet less deadly strain. Restrictions are gradually lifted across the world, and new infections become less of a concern in comparison to the initial outbreak. As it will be explained later, we expect to find a relationship between the year in which the papers are published and their findings. Common sense dictates that the passage of time has allowed more information to emerge, has led to the containment of the virus, and has witnessed a change in our understanding of the phenomenon.

The breakdown of the number of papers by the type of statistical tests used with Benford analysis is shown in Table 4. These tests offer a glimpse into the sophistication and depth of the analysis. It does not necessarily mean, however, that papers using fewer tests are less rigorous. The z-statistic, chi-square, and Mean Average Distance (MAD) are the most commonly used measures and test statistics in the Benford’s Law literature. They are used in 24 instances, 18 instances, and 10 instances, respectively. The Kuiper and KSD tests are used in five different instances each. The log likelihood ratio test, goodness of fit, and Euclidian distance are used three times each. The Morena-Montoya test, the Chebyshev distance M-statistic, Leemis M-statistic, Cho and Gaines D-statistic, SSD, RSMD, and RMSE are used once each.

Table 4. The breakdown of the number of papers by the type of statistical tests used with Benford analysis.

Type of Statistical Test	Frequency
Z-statistic	24
CHI-SQUARE	18
MAD	10
Goodness of Fit	3
KUIPER	5
Log likelihood ratio test	3
Distortion factor (DF)	3

Table 4. *Cont.*

Type of Statistical Test	Frequency
Euclidean distance	3
KSD statistic	5
Moreno-Montoya test	1
Chebyshev distance M statistic	1
Leemis M statistic	1
Cho and Gaines D-Statistic	1
SSD	1
RMSD/RMSE	2

As shown in Table 5, most papers use a combination of two or more statistical tests. The typical paper uses, on average, four different tests. Three papers use five or more different tests.

Table 5. The breakdown of the number of papers by the number of different statistical tests used in the Benford analysis.

Number of Different Tests Used	Frequency
1	2
2	6
3	4
4	10
5	2
6	1

The source of the data is shown in Table 6. There are ten papers that use a very large and diverse panel of international data, that is, data pertaining to more than fifteen countries at once, from virtually all continents (except Antarctica). North American data are identified in five studies, Latin-American data are present in another five studies, Western European data are used in three studies, two studies focus on China, and two more focus on other Asian countries. It is worth noting that there is a small degree of overlapping in the use of country samples (i.e., papers that use data from Canada and China, or Italy and China). There appears to be a good reason as to why, in some cases, the data are very country-specific and why a handful of countries represent the focus of several different papers. China offers a very substantial sample size by virtue of its huge population and represents the place from which the pandemic began spreading. Italy represents the first major European nation to be hit by COVID-19. The United States is a country with a relatively large population, a nation that has been significantly impacted by COVID-19, a country with extensive data collection capabilities, and a federal political system in which different states implemented different lockdown and vaccination policies. Brazil is another relatively large federal country, somewhat similar to the United States, in that it offers a large number of cases to be analyzed. The data have various levels of local, national, and regional aggregation, and most samples can be broken down by administrative structures and levels.

Table 6. The breakdown of the number of papers by the geographical provenience of the data used in the Benford analysis.

Source of the Data	Number of Papers
North America	5
Latin-America	5
Western Europe	3
Eastern Europe	0
China	2
India	0
Other Asia	2
World (multiple countries and continents)	10

The presentation of findings from 25 articles in a highly structured and systemized manner is nothing short of challenging. We settle on a breakdown focused on the severity of the data misalignment with Benford's distribution and the interpretation given to those findings.

Not surprisingly, 20 of 25 papers find at least some measure of deviation from Benford's distribution. Only two authors conclude (rather vaguely) that the data by-and-large fits the distribution. At least one of these two papers uses only visual inspection to conclude all seems to be conforming to the expected distribution. In three other papers, it is not at all clear what the nature of deviations is, if any.

Of the 20 papers that find statistically significant deviations, 13 claim these deviations appear more or less minor and/or mild. Seven papers acknowledge severe and/or persistent deviations from the expected Benford's distribution. The more delicate aspect, however, is the interpretation of the results. There are 16 papers suggesting innocuous reasons for the observed deviations. These include, but are not limited to, clerical errors, omissions, involuntary misreporting; poorly trained and/or overworked personnel; poorly designed information systems, or lack thereof; poor coordination among various levels of government agencies, voluntary self-censorship, and unconscious biases; and a lack of overall resources, lack of consistent reporting standards, plain incompetence, and other reasons along similar lines.

In addition to and/or in lieu of these rationalizations, at least six other papers conjecture a less innocuous interpretation, suggesting the results are driven by a malevolent manipulation of the data reporting process. Although these researchers claim that we are most likely dealing with data tampering, they keep away from embracing the highly controversial narratives peddled by various social media and political circles. They argue, and rightly so, that data tampering can occur for a variety of reasons, including but not limited to the pursuit of an electoral political agenda, the suppression of information that might prove politically and economically destabilizing, the cover-up of corruption and embezzlement, and control over the pandemic narrative.

One paper evaluates COVID-19 data from the United States and finds that deviations from Benford's Law are related to the political color of each state. Blue states, that is, states dominated by Democrats, show a distinctly different pattern of data reporting compared to red states, dominated by the GOP. The authors believe they identify the impact that different policies and measures related to mask mandates, lockdowns, and vaccination requirements had on the progression of COVID-19 cases across various American states.

In addition to, instead of, or unrelated to the above interpretations of results, nine papers offer a vigorous and rigorous discussion of the methodology. It is argued that it is difficult to disentangle the impact of the methodology from that of misreporting. One cannot exclude the possibility that some of the observed deviations might be partially or even totally due to the choice of statistical tests and the manner in which they have

been implemented. The application of Benford's Law has a long history, resulting in many studies dealing with various types of data, long before the COVID-19 pandemic. In almost every instance, the interpretation of results hinged on the choice of methodology. For example, it has been known for a long time that the analysis of data at different levels of aggregation is bound to produce inconsistent or conflicting statistical results [23,24]. Some authors even question the appropriateness of the overall Benford's approach to evaluating the data generated by the COVID-19 pandemic. The summary of findings is presented in Table 7.

Table 7. The breakdown of the number of papers by type of findings and interpretation.

Summary of Findings	Frequency
Some deviation from Benford's Law	13
Persistent deviation from Benford's Law	7
Plausible misreporting	16
Tampering with data	6
Methodological rationalization	9

Next, we attempt to determine if there is any segregation or pattern among the findings in our sample. It has been known for some time that hypothesis testing using z-scores and/or Pearson's chi-square statistic applied to large data samples can yield misleading results, which is why some authors recommend using the mean average deviation (MAD) instead [25,26]. One might expect that a more extensive and sophisticated methodology going beyond the minimum z-score and chi-square test would result in a more in-depth discussion and interpretation of results. Such a discussion might reduce some of the uncertainty around the extent and nature of deviations from Benford and might bring nuance with respect to the limitations of the various methodologies used. We therefore expect that studies using only hypothesis testing will find more deviations from Benford. The same goes for studies using large, aggregate samples. One would expect to find more deviations, as noted by [25,26].

Moreover, we expect to find some relationship between the period in which the paper was published and their findings. Although there is no theoretical background for this expectation, common sense dictates that as time passes, more and better information becomes available. At the same time, widespread lockdowns and vaccination campaigns would have already constrained the natural progression of the pandemic. While we have no clear direction for this relationship, we suspect we might be able to find a link between the period on the one hand and the extent of deviation from Benford and/or the type of rationalization on the other hand.

Last but not least, we would also expect that studies using a more elaborated and sophisticated methodology, using more numerous and novel procedures at the same time, and finding significant deviation from the expected distribution would be more inclined to question the overall appropriateness of applying Benford analysis to COVID-19 cases. At the very least, finding a significant deviation from Benford in the absence of other extra-statistical investigations should yield competing explanations. It seems unlikely the default interpretation would be that of data tampering when the most obvious question on the minds of data scientists should be: Does it make sense to apply Benford to COVID-19 data?

Next, we rearrange the data into (mostly binary) categorical variables and attempt to resolve the degree of association among them by pairing them in contingency tables. Given the small size of our sample, we cannot use the chi-square test because, in most cases, the requirement that at least 80% of the expected frequency be equal to or larger than 5 is not met. We resort to Fisher's Exact Test instead, which relies on a hypergeometric distribution and is robust to small-size samples [49]. For larger than 2×2 tables, we compute the corresponding *p*-values of the test using Monte Carlo simulation with 2000 replicates. The results are presented in Table 8. The test is implemented in R.

Table 8. Fisher’s Exact Test results for pairs of categorical variables.

Pairs of Categorical Variables (H0: Independence)	Fisher’s Exact Test Two-Tailed p -Value	ODDS RATIO
Period vs. severity of deviation from Benford	0.645	2.153
Period vs. rationalization	0.344	0.374
No. of different tests vs. severity of deviation from Benford	0.076 *	n.a.
No. of different tests vs. rationalization	0.095	n.a.
Severity of deviation from Benford vs. rationalization	0.032 **	9.347
No. of countries vs. severity of deviation from Benford	0.073 *	8.63
No. of countries vs. rationalization	0.644	2.153
Deviation from Benford vs. methodological caveat	0.205	3.281
No. of different tests vs. methodological caveat	0.616	n.a.

* $p < 0.1$, ** $p < 0.05$.

The variable “Period” has the levels “peak”—for the years 2020 and 2021—and “ebb”—for the years 2022 and 2023. The variable “Severity of deviation from Benford” has the levels “none or some mild deviation,” and “severe deviation.” The variable “Rationalization” has the levels “plausible misreporting, if any” and “tampering with data.” The variable “No of countries” has the levels “single” and “multiple,” depending on the number of countries covered in the respective data sample. The variable “Methodology” is “0” when only z-score and chi-square tests are used and “1” when MAD and/or additional, more sophisticated methods are used. The variable “Methodological caveat” has the levels “0” when no methodological caveat is offered as an alternative or in addition to discussions and conclusions and “1” when a methodological caveat is considered as an alternative or in addition to the main discussions and/or conclusions.

We note that there are only three instances in which the pairs of variables do not appear independent, although, in two cases, the test is only marginally significant. In the first case, it appears that the severity of the deviation from Benford might be related to the number of different statistical tests used ($p = 0.076$), although this result is marginally significant. In the second case, it appears that the severity of the deviation from Benford might be related to the interpretation given to this finding ($p = 0.032$ and odds ratio = 9.347); data tampering appears to be inferred more readily whenever large and significant deviations from the expected distribution are encountered.

In the third case, results suggest that there might be a marginally significant relationship between the geographical coverage of the data (single country vs. multiple countries) and the severity of the deviation from Benford ($p = 0.073$ and odds ratio = 8.63).

As already mentioned above, the use of hypothesis testing (z-scores and/or Pearson’s chi-square statistic applied to large data samples) can bias the results. At the very least, one should use the mean average deviation (MAD), possibly in combination with other statistical procedures instead [25,26].

We take note of the relative lack of strong statistical significance in our results and conclude that the issue lies either with the very modest power of Fisher’s Exact Test or, less likely, the types of methodologies used did not systematically influence the nature of findings and/or interpretation.

Interestingly, there is also no apparent relationship between the period in which the studies were conducted and the type of interpretation given to the findings. When the pandemic moved beyond its peak and started to subside in 2022, the extent of measured deviations, and the propensity to attribute these deviations (if any) to either misreporting or tampering did not appear to shift systematically. Again, either Fisher’s Exact Test is too weak or the passage of time did not bring about more relevant information and insight.

Given the categorical nature of our data, we also attempt to predict the extent of deviation from Benford, the type of rationalization given (tampering with data or not), and whether there is a methodological caveat for the results. We use classification trees, in which 85% of our small sample is used for training and 15% for prediction. We are fully aware of the limitations of this method in the presence of a small sample, but we believe it is worth reporting the results.

Figure 7 presents the classification tree for the extent/severity of deviation from Benford. Here, we allow the variable to take three values, “no deviation”, “some deviation”, and “persistent deviation”. Accuracy is 75%, but the True Positive Rate appears to vary in the same proportion as the False Positive Rate, making for arguably weak results. The model, however, seems to suggest that when research uses a large, aggregate data sample, extracted from multiple countries and regions, it is more likely than not to find a persistent deviation from Benford. When analyzing only one country using MAD and other advanced methods, research is more likely than not to find only mild deviations from Benford. Notwithstanding the modest accuracy of the model, these predictions are largely consistent with the issues of sample aggregation and hypothesis testing already signaled by [25,26].

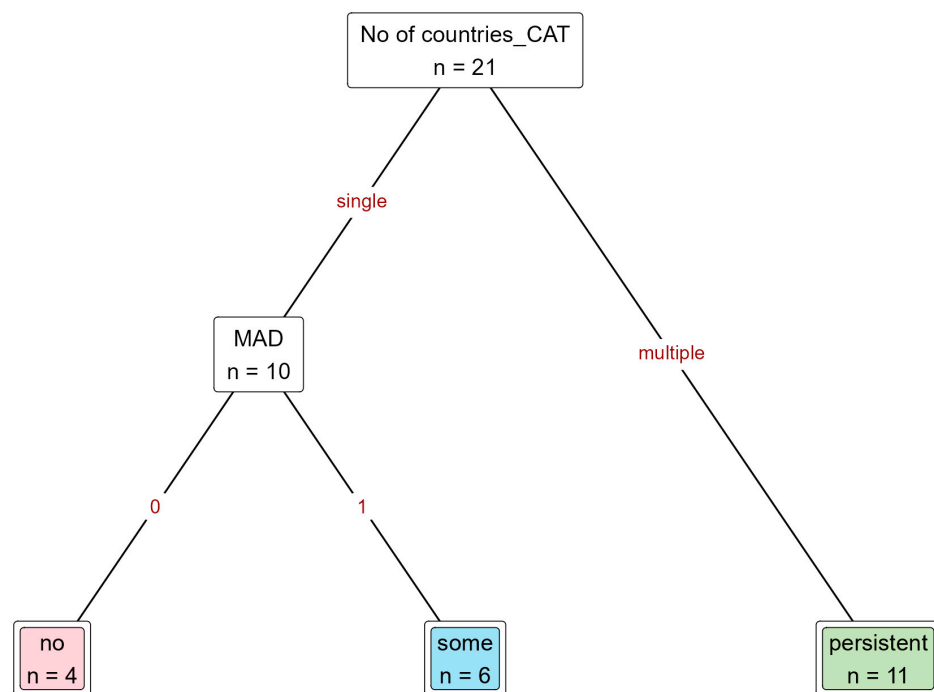


Figure 7. Prediction tree: extent/severity of deviation from Benford.

Predicting whether research concludes data have been tampered with yields a far lower prediction accuracy. When persistent deviations from Benford are present and the research methodology is limited to the use of z-score and chi-square tests or similar, it is marginally more likely than not to conclude data tampering. The visual presentation of the results is available in Figure 8.

The last prediction is obtained for the methodological caveat and is displayed in Figure 9. Whenever research finds a persistent deviation from Benford and MAD and/or more sophisticated methods are used, it is more likely than not that a methodological caveat is included with the discussion and/or conclusions. Here, we have encoded the variable “Persistent deviation from Benford” as binary.

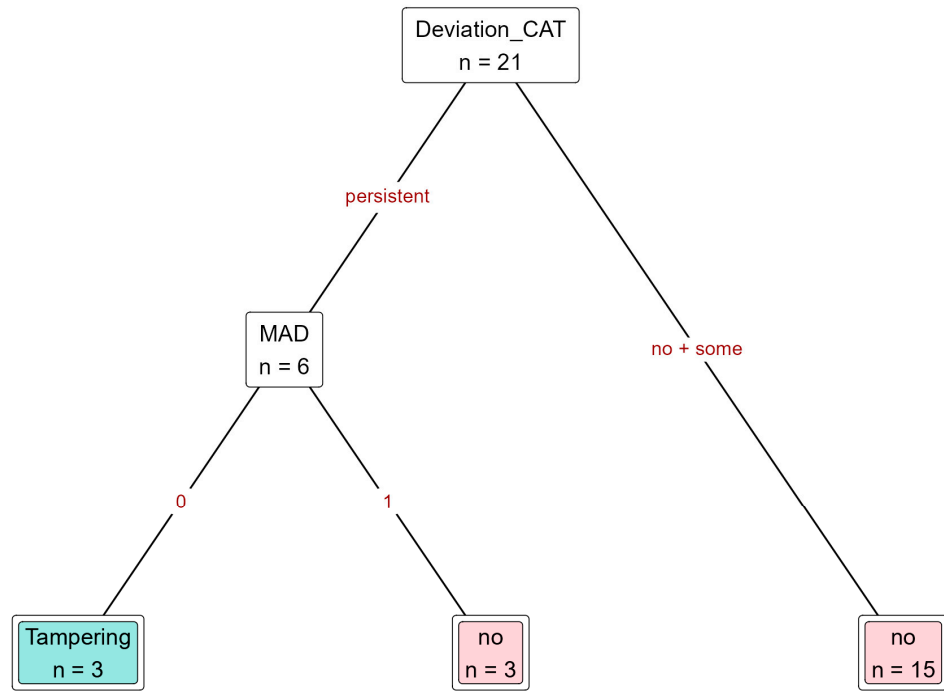


Figure 8. Rationalization: tampering with data.

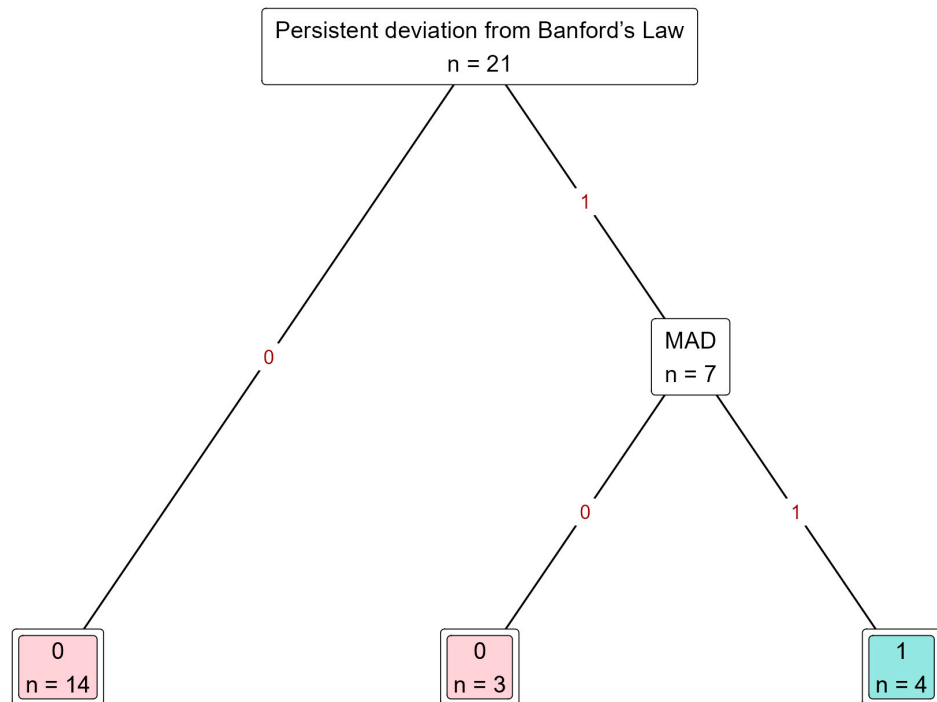


Figure 9. Methodological caveat: decision tree plot.

5. Discussion of Results

Our bibliometric analysis emphasizes the breadth and depth of the collaborative effort put into researching the accuracy of COVID-19 data reporting using Benford’s Law. Most studies represent the result of teamwork among researchers from almost all continents and from a variety of academic backgrounds and institutions. The most active collaborative efforts appear between authors from the US and UK, Brazil and Portugal, China and Norway, and Sweden and Singapore. Yet, most of this work is focused mainly on the

narrow application of the Benford methodology to COVID-19 data. While interdisciplinary work is present, its occurrence is surprisingly rare, in our opinion.

Our survey of the content is largely descriptive and mostly qualitative. We document 25 papers analyzing data obtained from one or multiple countries using a wide array of methods. Most findings show at least some deviation of COVID-19 data from the expected Benford distribution, and an important number of studies document persistent deviations from Benford. These findings are explained as plausible misreporting, data tampering, and/or artifacts of the methodology employed. However, the most important result so far is the absence of an emerging consensus among scholars with respect to whether data reporting has been systematically manipulated.

We are nevertheless able to combine some of the qualitative data with quantitative methods to arguably identify the presence of very weak segregation among results, the use of methodology, and the explanations advanced by the authors of these studies. Fisher's Exact Test results suggest there might be a link between the number of countries from which the data were obtained and the presence of persistent deviations from Benford, and between the aforementioned deviation and the interpretation given to those findings.

Fisher's Exact Test results fit relatively well with the results of classification tree predictive models. In spite of the shortcomings associated with the small sample and low accuracy of prediction, we discern a narrative that is coherent and consistent. Research relying on hypothesis testing (z-score and chi-square) and focusing on large, aggregate samples from multiple countries is marginally more likely than not to identify persistent deviations from Benford and conclude data tampering is present. Among the studies finding persistent deviations from Benford, those using MAD and/or other advanced methods are marginally more likely than not to include a methodological caveat in their discussion of findings.

Our results, however, should be approached with a very healthy dose of skepticism. The overall picture is far from compelling. Apart from descriptive statistics, we use a test known for its low power and obtain results that are marginally significant. Our prediction models are notoriously fickle in the presence of small samples. Their prediction accuracy is low, barely above that of a coin toss.

6. Implications and Conclusions

As mentioned in the introduction, this paper is not quite a meta-analysis but borrows from the meta-analysis methodology and sets out to answer a couple of important questions pertaining to the nature of misreporting and the appropriateness of the entire digit analysis approach in the context of the pandemic.

In spite of wide-ranging academic collaborations among researchers from various countries using large samples of COVID-19 cases and generating over 25 recent studies, there is no emerging consensus on whether COVID-19 data have been manipulated and tampered with. That many studies found at least some measure of mild deviation from Benford is a relatively trivial result and does not provide the much-needed clarity.

It is debatable whether the evidence of segregation among the findings of these papers that we document here is strong enough to warrant definitive conclusions. We would hope for an easily discernible pattern among sample aggregation levels, statistical tools, the extent/severity of deviations from Benford, and their interpretation. Instead, we conduct tests and predictive models, producing results that are rather weak. There is no conclusive picture emerging, only blurred traces. If Benford were indeed the ultimate tool for the detection of misreporting, we should see the emergence of a vigorous consensus. This is obviously not the case. We would have to conclude that either there is a fundamental problem with applying digit analysis to pandemic data or, less likely, there must be some unconscious local bias present among researchers. A handful of papers cited here echo this concern and raise questions, not only about the validity of individual procedures but also about the entire concept of COVID data manipulation detection based on digit analysis.

The main takeaway for academics is that future research needs to provide definitive clarification regarding the appropriateness of applying Benford's Law to COVID-19 and other pandemic data.

The use of digit analysis applied to tax fraud detection has shown that deviations from the expected Benford distribution represent at most mere starting points for a more in-depth investigation using traditional law enforcement protocols. Deviations in and by themselves represent by no means conclusive proof of data manipulation and/or fraud.

Notwithstanding these caveats, policymakers, public administrators, and healthcare officials should always be mindful of biased statistics. Some type of misreporting is without a doubt the default state, especially when dealing with complex bureaucracies in a health crisis context. This is by no means breaking news. The only question is about its severity and causes. And this is precisely why a lot of resources should be dedicated to personnel training, improved and streamlined reporting procedures, and crisis management.

We know distortions might be present when data are aggregated into large, stratified samples. Another source of distortions is represented by policy measures aimed at containing the crisis. They end up interfering with the natural course of the pandemic and skewing the numbers. Since it would be impractical and unethical to conduct controlled experiments during healthcare emergencies, it is nearly impossible to disentangle the interference effect generated by remedial policies from incompetence and/or breakdowns in healthcare and administrative procedures. Moreover, one could never categorically exclude the impact of various political agendas and interest groups.

In the end, the overriding consideration for policymakers and academics alike is that existing research has not been able to settle the question of data manipulation associated with COVID data. And it is unlikely it will be able to do so in the near future.

Author Contributions: Conceptualization, E.D. and C.V.; methodology, A.-I.P., C.V. and E.D.; software, A.-I.P.; formal analysis, A.-I.P., C.V. and E.D.; writing—original draft preparation, A.-I.P. and C.V.; writing—review and editing, C.V. and E.D.; supervision, E.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been made possible with a grant from the Senate Research Committee of Bishop's University, Canada, grant number RACG-102703.

Data Availability Statement: Data consists of already published papers.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nigrini, M.J. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*; Wiley Corporate F&A Series; Wiley: Hoboken, NJ, USA, 2012; ISBN 978-1-118-15285-0.
2. Azevedo, C.D.S.; Gonçalves, R.F.; Gava, V.L.; Spinola, M.D.M. A Benford's Law Based Methodology for Fraud Detection in Social Welfare Programs: Bolsa Familia Analysis. *Phys. A Stat. Mech. Its Appl.* **2021**, *567*, 125626. [[CrossRef](#)]
3. Noorullah, A.S.; Jari, A.S.; Hasan, A.M.; Flayyih, H.H. Benford Law: A Fraud Detection Tool Under Financial Numbers Game: A Literature Review. *Soc. Sci. Humanit. J.* **2020**, *4*, 1909–1914.
4. Durtschi, C.; Hillison, W.; Pacini, C. The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data. *J. Forensic Account.* **2004**, *5*, 17–34.
5. Idrovo, A.J.; Fernández-Niño, J.A.; Bojórquez-Chapela, I.; Moreno-Montoya, J. Performance of Public Health Surveillance Systems during the Influenza A(H1N1) Pandemic in the Americas: Testing a New Method Based on Benford's Law. *Epidemiol. Infect.* **2011**, *139*, 1827–1834. [[CrossRef](#)] [[PubMed](#)]
6. Lu, F.; Boritz, J.E. Detecting Fraud in Health Insurance Data: Learning to Model Incomplete Benford's Law Distributions. In *Machine Learning: ECML 2005*; Gama, J., Camacho, R., Brazdil, P.B., Jorge, A.M., Torgo, L., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3720, pp. 633–640. ISBN 978-3-540-29243-2.
7. Crocetti, E.; Randi, G. Using the Benford's Law as a First Step to Assess the Quality of the Cancer Registry Data. *Front. Public Health* **2016**, *4*, 225. [[CrossRef](#)]
8. Daniels, J.; Caetano, S.-J.; Huyer, D.; Stephen, A.; Fernandes, J.; Lytwyn, A.; Hoppe, F.M. Benford's Law for Quality Assurance of Manner of Death Counts in Small and Large Databases. *J. Forensic Sci.* **2017**, *62*, 1326–1331. [[CrossRef](#)]
9. Morillas-Jurado, F.G.; Caballer-Tarazona, M.; Caballer-Tarazona, V. Applying Benford's Law to Monitor Death Registration Data: A Management Tool for the Covid-19 Pandemic. *Mathematics* **2022**, *10*, 46. [[CrossRef](#)]

10. Natashekara, K. COVID-19 Cases in India and Kerala: A Benford's Law Analysis. *J. Public Health* **2022**, *44*, E287–E288. [[CrossRef](#)]
11. Wong, W.K.; Juwono, F.H.; Loh, W.N.; Ngu, I.Y. Newcomb-Benford Law Analysis on COVID-19 Daily Infection Cases and Deaths in Indonesia and Malaysia. *Herit. Sustain. Dev.* **2021**, *3*, 102–110. [[CrossRef](#)]
12. Kilani, A.; Georgiou, G.P. Countries with Potential Data Misreport Based on Benford's Law. *J. Public Health* **2021**, *43*, E295–E296. [[CrossRef](#)]
13. Campolieti, M. COVID-19 Deaths in the USA: Benford's Law and under-Reporting. *J. Public Health* **2022**, *44*, E268–E271. [[CrossRef](#)]
14. Balashov, V.S.; Yan, Y.; Zhu, X. Using the Newcomb–Benford Law to Study the Association between a Country's COVID-19 Reporting Accuracy and Its Development. *Sci. Rep.* **2021**, *11*, 22914. [[CrossRef](#)]
15. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to Conduct a Bibliometric Analysis: An Overview and Guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [[CrossRef](#)]
16. Zhong, M.; Lin, M. Bibliometric Analysis for Economy in COVID-19 Pandemic. *Heliyon* **2022**, *8*, e10757. [[CrossRef](#)] [[PubMed](#)]
17. Chen, Y.; Zhang, X.; Chen, S.; Zhang, Y.; Wang, Y.; Lu, Q.; Zhao, Y. Bibliometric Analysis of Mental Health during the COVID-19 Pandemic. *Asian J. Psychiatry* **2021**, *65*, 102846. [[CrossRef](#)]
18. Farooq, R.; Rehman, S.; Ashiq, M.; Siddique, N.; Ahmad, S. Bibliometric Analysis of Coronavirus Disease (COVID-19) Literature Published in Web of Science 2019–2020. *J. Fam. Community Med.* **2021**, *28*, 1. [[CrossRef](#)]
19. Mahi, M.; Mobin, M.A.; Habib, M.; Akter, S. A Bibliometric Analysis of Pandemic and Epidemic Studies in Economics: Future Agenda for COVID-19 Research. *Soc. Sci. Humanit. Open* **2021**, *4*, 100165. [[CrossRef](#)] [[PubMed](#)]
20. Viana-Lora, A.; Nel-lo-Andreu, M.G. Bibliometric Analysis of Trends in COVID-19 and Tourism. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 173. [[CrossRef](#)]
21. Heradio, R.; Perez-Morago, H.; Fernandez-Amoros, D.; Javier Cabrerizo, F.; Herrera-Viedma, E. A Bibliometric Analysis of 20 Years of Research on Software Product Lines. *Inf. Softw. Technol.* **2016**, *72*, 1–15. [[CrossRef](#)]
22. Aria, M.; Cuccurullo, C. Bibliometrix: An R-Tool for Comprehensive Science Mapping Analysis. *J. Informetr.* **2017**, *11*, 959–975. [[CrossRef](#)]
23. Kaiser, M. Benford's law as an indicator of survey reliability—Can we trust our data? *J. Econ. Surv.* **2019**, *33*, 1602–1618. [[CrossRef](#)]
24. Druică, E.; Oancea, B.; Vălsan, C. Benford's Law and the Limits of Digit Analysis. *Int. J. Account. Inf. Syst.* **2018**, *31*, 75–82. [[CrossRef](#)]
25. Nigrini, M.J. Audit Sampling Using Benford's Law: A Review of the Literature with Some New Perspectives. *J. Emerg. Technol. Account.* **2017**, *14*, 29–46. [[CrossRef](#)]
26. Barney, B.J.; Schulzke, K.S. Moderating “Cry Wolf” Events with Excess MAD in Benford's Law Research and Practice. *J. Forensic Account. Res.* **2016**, *1*, A66–A90. [[CrossRef](#)]
27. Dai, X.; Gil, G.F.; Reitsma, M.B.; Ahmad, N.S.; Anderson, J.A.; Bisignano, C.; Carr, S.; Feldman, R.; Hay, S.I.; He, J.; et al. Health Effects Associated with Smoking: A Burden of Proof Study. *Nat. Med.* **2022**, *28*, 2045–2055. [[CrossRef](#)]
28. Wang, B.; López-Corredoira, M.; Wei, J.-J. The Hubble Tension Survey: A Statistical Analysis of the 2012–2022 Measurements. *Mon. Not. R. Astron. Soc.* **2023**, *527*, 7692–7700. [[CrossRef](#)]
29. Cucari, N.; Tutore, I.; Montera, R.; Profita, S. A Bibliometric Performance Analysis of Publication Productivity in the Corporate Social Responsibility Field: Outcomes of SCIVAL Analytics. *Corp. Soc. Responsib. Environ. Manag.* **2023**, *30*, 1–16. [[CrossRef](#)]
30. Andrikopoulos, A.; Economou, L. Coauthorship and Subauthorship Patterns in Financial Economics. *Int. Rev. Financ. Anal.* **2016**, *46*, 12–19. [[CrossRef](#)]
31. Benford, F. The Law of Anomalous Numbers. *Proc. Am. Philos. Soc.* **1938**, *78*, 551–572.
32. Diekmann, A. Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. *J. Appl. Stat.* **2007**, *34*, 321–329. [[CrossRef](#)]
33. Koch, P. Economic Complexity and Growth: Can Value-Added Exports Better Explain the Link? *Econ. Lett.* **2021**, *198*, 109682. [[CrossRef](#)]
34. Fewster, R.M. A Simple Explanation of Benford's Law. *Am. Stat.* **2009**, *63*, 26–32. [[CrossRef](#)]
35. Yan, E.; Ding, Y. Scholarly Network Similarities: How Bibliographic Coupling Networks, Citation Networks, Cocitation Networks, Topical Networks, Coauthorship Networks, and Coword Networks Relate to Each Other. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 1313–1326. [[CrossRef](#)]
36. K-Synth Team. Frequently Asked Questions. 2023. Available online: <https://www.bibliometrix.org/home/index.php/about-us-2/k-synth-team> (accessed on 14 May 2024).
37. Our World in Data. Brazil: Coronavirus Pandemic Country Profile. 2024. Available online: <https://ourworldindata.org/coronavirus/country/brazil> (accessed on 31 July 2024).
38. Chen, C.; Dubin, R.; Schultz, T. Science Mapping. In *Advances in Information Quality and Management*; Mehdi Khosrow-Pour, D.B.A., Ed.; IGI Global: Hershey, PA, USA, 2014; pp. 4171–4184. ISBN 978-1-4666-5888-2.
39. Durieux, V.; Gevenois, P.A. Bibliometric Indicators: Quality Measurements of Scientific Publication. *Radiology* **2010**, *255*, 342–351. [[CrossRef](#)] [[PubMed](#)]
40. Osareh, F. Bibliometrics, Citation Analysis and Co-Citation Analysis: A Review of Literature I. *Libri* **1996**, *46*, 149–158. [[CrossRef](#)]
41. Small, H. Co-Citation in the Scientific Literature: A New Measure of the Relationship between Two Documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [[CrossRef](#)]

42. Navarro-Ballester, A.; Merino-Bonilla, J.A.; Ros-Mendoza, L.H.; Marco-Doménech, S.F. Publications on COVID-19 in Radiology Journals in 2020 and 2021: Bibliometric Citation and Co-Citation Network Analysis. *Eur. Radiol.* **2022**, *33*, 3103–3114. [[CrossRef](#)]
43. Mas-Tur, A.; Roig-Tierno, N.; Sarin, S.; Haon, C.; Segó, T.; Belkhouja, M.; Porter, A.; Merigó, J.M. Co-Citation, Bibliographic Coupling and Leading Authors, Institutions and Countries in the 50 Years of Technological Forecasting and Social Change. *Technol. Forecast. Soc. Change* **2021**, *165*, 120487. [[CrossRef](#)]
44. Fusco, F.; Marsilio, M.; Guglielmetti, C. Co-Production in Health Policy and Management: A Comprehensive Bibliometric Review. *BMC Health Serv. Res.* **2020**, *20*, 504. [[CrossRef](#)]
45. Trujillo, C.M.; Long, T.M. Document Co-Citation Analysis to Enhance Transdisciplinary Research. *Sci. Adv.* **2018**, *4*, e1701130. [[CrossRef](#)]
46. Nigrini, M.J. Taxpayers Compliance Application of Benford's Law. *J. Am. Tax. Assoc.* **1996**, *18*, 72–92.
47. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. An Approach for Detecting, Quantifying, and Visualizing the Evolution of a Research Field: A Practical Application to the Fuzzy Sets Theory Field. *J. Informetr.* **2011**, *5*, 146–166. [[CrossRef](#)]
48. Cobo, M.J.; Martínez, M.A.; Gutiérrez-Salcedo, M.; Fujita, H.; Herrera-Viedma, E. 25years at Knowledge-Based Systems: A Bibliometric Analysis. *Knowl.-Based Syst.* **2015**, *80*, 3–13. [[CrossRef](#)]
49. Kim, H.-Y. Statistical Notes for Clinical Researchers: Chi-Squared Test and Fisher's Exact Test. *Restor. Dent. Endod.* **2017**, *42*, 152. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.