MULTIDISCIPLINARY SCIENCE JOURNAL

RESEARCH ARTICLE
Published Online: November 29, 2023
https://doi.org/10.31893/multiscience.2023079

Analyzing crop production statistics of the Philippines using the newcomb-benford law

Samuel John Estenor Parreño[a]

[a]Department of Teacher Education, University of Mindanao Digos College, Digos City, Davao del Sur, Phillipines.

Abstract The agricultural sector plays a crucial role in the Philippines, contributing significantly to its food security and economic growth. However, the sector faces various challenges that hinder its productivity and growth. To address these challenges and promote sustainable agricultural practices, accurate and reliable crop production statistics are essential for informed decision-making and resource allocation. The Newcomb-Benford Law (NBL) provides a statistical distribution pattern for leading digits in numerical datasets, offering insights into data reliability. In this research study, we applied the NBL to analyze crop production statistics for major crops (rice, corn, coconut, sugarcane, banana, cassava, and pineapple) in the Philippines. We assessed the conformity of the datasets to NBL expectations and identified significant deviations, indicating potential data accuracy issues, collection method discrepancies, or reporting irregularities. Although these deviations do not conclusively suggest fraud, they underscore the need for meticulous data validation. The first two-digit test further strengthened the findings, providing a comprehensive understanding of dataset conformity. Transparent data collection and validation processes are crucial for trustworthy agricultural statistics, supporting effective policy-making and resource allocation. Future research should investigate the root causes of the deviations, explore data processing errors, and implement stringent data validation procedures. Additionally, expanding the analysis to include other crops and conducting comparative studies between regions and time periods would enhance data integrity and contribute to sustainable agricultural practices and food security in the Philippines.

Keywords: agriculture, first digit test, first two-digit test, data validation, agricultural policy

1. Introduction

Agriculture plays a vital role in the Philippines, where its abundant agricultural resources and favorable climate contribute to the nation's status as an agricultural country (Yamagishi et al., 2021). Crop production serves as the cornerstone of food security and economic growth in the country. However, the agricultural sector faces various challenges that hinder its growth and productivity. These challenges include land conversions and small farm size, limited access to modern farming technologies, inadequate infrastructure, and inconsistent policy implementation (Briones, 2021). Addressing these problems requires a comprehensive understanding of the agricultural sector, including accurate and reliable crop production statistics. Accurate and dependable crop production statistics play a paramount role in effective policy-making, sustainable agricultural practices, and resource allocation (Hanci, 2022). These statistics furnish critical information on crop yields, production trends, and regional variations, enabling policymakers and stakeholders to make well-informed decisions and allocate resources efficiently. The availability of accurate and timely crop production data is essential for identifying areas requiring improvement, optimizing agricultural practices, and ensuring long-term food security.

The Newcomb-Benford Law (NBL) provides a statistical distribution pattern for the occurrence of leading digits in naturally occurring numerical datasets. It postulates that, in many datasets, smaller leading digits (1 to 9) occur more frequently than larger digits (Newcomb, 1881; Benford, 1938). The extent and direction of deviations between the observed values and the expected frequencies according to the NBL can provide insights into the dataset's reliability. While data with deviations are not conclusive evidence of fraud, the researcher can utilize these deviations to formulate an audit plan that takes into account potential irregularities within the dataset (Gauvrit et al., 2017). The NBL analysis method comprises two stages: general analysis and special analysis tests. The general analysis involves two tests, namely the first digit and second digit tests, which provide an initial understanding of the data. The special analysis tests include the first two digits, first three digits, last two digits, and duplicate recording tests (Yanik & Samanci, 2013). It is important to note that the first- and second-digit tests are unsuitable for control sampling. Nonetheless, the first two-digit test is highly effective in identifying fundamental anomalies within the data (Goh, 2020). The NBL has exhibited successful applications in diverse domains such as finance (Jianu & Jianu, 2021), cryptocurrencies (Vičič & Tošić, 2022), malicious online automated accounts (Madahali & Hall,

Received: August 1, 2023 | Accepted: October 11, 2023

2020), scientific cooperation network (Tošić & Vičič, 2021), epidemiology (Manrique-Hernandez et al., 2017; Parreño, 2023), COVID-19 (Balashov et al., 2021), supply chain management (Kraus & Valverde, 2014), and data analysis (Li et al., 2019), where it has been employed to detect irregularities, identify fraudulent activities, and assess the integrity of numerical datasets (Eckhartt & Ruxton, 2023).

In recent years, researchers have increasingly embraced the application of the Newcomb-Benford Law to analyze agricultural datasets, leveraging its potential to provide valuable insights into the accuracy and reliability of reported agricultural statistics (Tenkcoran, 2022; Hanci, 2022; Qin et al, 2019; Levicar, 2020). By harnessing this statistical tool, researchers have been able to unravel distribution patterns, detect anomalies, and validate the integrity of agricultural data. The utilization of the Newcomb-Benford Law in agricultural studies has demonstrated promise in enhancing data reliability, improving decision-making processes, and identifying areas warranting further investigation.

The Philippines, owing to its favorable climate and fertile soil, cultivates a diverse array of crops. Major crops in the country encompass rice, corn, coconut, sugarcane, banana, cassava, and pineapple (Go & Conag, 2019; Landicho & Balendres, 2022; Pandit et al., 2020). These crops play a pivotal role not only in domestic consumption but also in the country's export market. Analyzing the crop production statistics of these major crops assumes critical significance for comprehending production trends, identifying potential challenges, and formulating effective strategies to maximize agricultural productivity.

The primary objective of this research study is to apply the Newcomb-Benford Law to analyze the crop production statistics of the Philippines, with a specific focus on the major crops, namely rice, corn, coconut, sugarcane, banana, cassava, and pineapple. This analysis aims to unravel the distribution patterns of leading digits in the crop production datasets and identify any deviations from the expected patterns, which may indicate anomalies or potential data manipulation. Through the evaluation of the accuracy and reliability of the reported crop production statistics, this study provide invaluable insights for agricultural planning, policy-making, and resource allocation in the Philippines. Ultimately, the findings of this research endeavor contribute to the promotion of sustainable crop production practices and foster economic growth within the agricultural sector.

2. Materials and Methods

2.1. Research design

This study adopts a quantitative research design to achieve the objectives. Quantitative research design is a systematic and structured method that relies on numerical data and statistical analysis to draw conclusions (Apuke, 2017). In the context of this research, a quantitative approach is well-suited for a rigorous examination of the crop production statistics from the Philippine Statistics Authority (PSA). This approach allows for a comprehensive analysis of the data using statistical tools to uncover patterns, anomalies, and deviations. Historical data provides a valuable foundation for assessing trends and changes in crop production over time. Analyzing historical data allows us to identify long-term patterns and variations that are crucial for informed decision-making in the agricultural sector.

However, it is essential to acknowledge the limitations associated with quantitative research. While quantitative methods offer many advantages, including objectivity and the ability to analyze large datasets, they also come with inherent limitations. One limitation is the potential for oversimplification of complex phenomena, as quantitative research often focuses on measurable variables. Additionally, quantitative analysis relies on predefined metrics, which may not encompass all relevant factors. Furthermore, historical data, such as the crop production statistics provided by the PSA, can be subject to biases and limitations. These biases may arise due to changes in data collection methods, reporting practices, or external factors affecting data accuracy over time. It is crucial to recognize that historical data might not be entirely free from errors or biases. As a result, the findings of this study are based on the assumption that historical data accurately reflect past crop production trends.

To address these potential limitations, this research employs rigorous statistical tests and validation procedures to assess the reliability of the historical crop production data. By doing so, we aim to mitigate the impact of any biases or inaccuracies that may exist in the dataset. This comprehensive approach ensures that the conclusions drawn from the quantitative analysis are robust and provide valuable insights into the crop production statistics in the Philippines.

2.2. Data source

The data for this research study were obtained from the openSTAT open data platform of the Philippine Statistics Authority (PSA), consisting of crop production statistics for major crops including rice, corn, coconut, sugarcane, banana, cassava, and pineapple. The dataset includes quarterly production data for rice and corn from 1987 to the first quarter of 2023, while quarterly production data for coconut, sugarcane, banana, cassava, and pineapple covers the period from 2010 to the first quarter of 2023, all measured in metric tons. With a total of 555 data points, the selected quarters within the specified time frames were used to ensure consistency and comprehensiveness in the analysis. R version 4.2.3 was employed for all statistical analyses and data visualizations, as it offers extensive capabilities for statistical computing and graphics. The use of the openSTAT platform from the PSA ensures the reliability and credibility of the data, providing access to official government statistics, facilitating the derivation of valuable insights into the distribution patterns and potential anomalies in the crop production statistics of the selected major crops in the Philippines.

2.3. Data analysis

The data analysis for this research study involves the application of the Newcomb-Benford Law to analyze the leading digit frequencies within the dataset of crop production statistics for major crops in the Philippines. The following statistical techniques was utilized to assess the reliability and distribution patterns of the reported data:

Newcomb-Benford Law Analysis. The Newcomb-Benford Law was applied to the dataset to examine the occurrence of leading digits (1 to 9) in the reported crop production values. This analysis test whether the data adhered to the expected distribution pattern according to the law, where smaller leading digits occur more frequently than larger digits (Newcomb, 1881; Benford, 1938). The first digit and first two-digit tests were conducted as part of the general analysis to provide initial insights into the distribution patterns (Goh, 2020). The probability (P) of the first significant digit (d) occurring is given by:

$$P(d) = \log_{10}\left(\frac{1+d}{d}\right) \text{ for } d \in \{1,2,\ldots,9\} \tag{1}$$

Chi-square Test. To assess the significance of the differences between the observed leading digit frequencies and the expected frequencies based on the NBL, the chi-square test was employed. This test determines if the deviations from the expected distribution were statistically significant, indicating potential anomalies or irregularities in the dataset (Fewster, 2009). The chi-square statistic (χ^2) is calculated as follows:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{2}$$

where O_i is the observed leading digit frequencies and E_i is the expected frequencies based on the NBL.

Mantissa Arc Test. In addition to the chi-square test, the mantissa arc test was utilized to further evaluate the distribution patterns of the leading digits. The test involves calculating the arcsine transformation of the observed mantissas (a) and comparing them with the expected arcsine values (b) derived from the Newcomb-Benford distribution (Madahali & Hall, 2020). The test statistic (M) is computed as:

$$M = \sum(\sin^{-1}(a) - \sin^{-1}(b))^2 \tag{3}$$

Mean Absolute Deviation. The Mean Absolute Deviation (MAD) measures the average absolute difference between the observed leading digit frequencies (O_i) and the expected frequencies (E_i) according to the NBL. A lower MAD value indicates a closer alignment between the observed and expected leading digit frequencies, suggesting a higher level of conformity to the Newcomb-Benford distribution. Conversely, a higher MAD value implies greater deviation from the expected distribution, indicating potential anomalies or irregularities in the dataset. A significant MAD value may warrant further investigation to identify the underlying reasons for the deviations from the expected distribution (Nigrini, 2012). The MAD is calculated as:

$$MAD = \frac{\sum|O_i - E_i|}{n} \tag{4}$$

where n is the total number of leading digit frequencies.

Distortion Factor. The distortion factor was calculated as a measure of the overall deviation between the observed and expected leading digit frequencies. The distortion factor quantifies the extent of distortion in the dataset and can reveal the presence of potential irregularities or manipulations (Geyer & Williamson, 2004). The distortion factor is calculated as the ratio of the sum of observed leading digit frequencies (O_i) to the sum of expected frequencies (E_i) based on the NBL. The formula is:

$$DF = \frac{\sum O_i}{\sum E_i} \tag{5}$$

3. Results and Discussion

Before implementing the NBL analysis, we conducted a preliminary check to ensure that the numeric data of the major crops satisfied the constraints outlined by Tošić and Vičič (2021). This step was essential to avoid obtaining misleading results from the NBL analysis. The examination revealed that the datasets for coconut, banana, cassava, and pineapple violated the property that requires data to be right-skewed, rendering the NBL inapplicable to these crops. These deviations in the data source can potentially be attributed to variations in reporting practices, data collection methods, or unique characteristics of these crops. For example, the absence of the leading digit "1" in these datasets may point towards specific data entry practices or data processing algorithms that differ from those applied to other crops. It is crucial to acknowledge that such discrepancies can result from both human errors and systematic factors within the data reporting system. Further investigation into these specific factors could provide valuable insights into the observed deviations.

Table 1 presents the NBL analysis and statistical tests on the coconut, banana, cassava, and pineapple datasets. The datasets exhibit an absence of the leading digit "1," indicating that none of the reported values start with this digit. Moreover, the Mean Absolute Deviation (MAD) of each crop were classified as “nonconformity” (Nigrini, 2012) while all statistical tests have p-values less than 0.05 since the crops do not conform to NBL.

Table 1 Misleading results of the NBL analysis.

Digit	NBL	Coconut	Banana	Cassava	Pineapple
1	30.1	0.0	0.0	0.0	0.0
2	17.6	0.0	100.0	0.0	0.0
3	12.5	66.0	0.0	0.0	0.0
4	9.7	34.0	0.0	20.8	3.8
5	7.9	0.0	0.0	24.5	18.9
6	6.7	0.0	0.0	22.6	52.8
7	5.8	0.0	0.0	22.6	24.5
8	5.1	0.0	0.0	9.5	0.0
9	4.6	0.0	0.0	0.0	0.0
n		53	53	53	53
χ^2		195.08*	247.98*	107.49*	247.55*
Mantissa		0.92413*	0.97343*	0.7399*	0.92191*
MAD		0.1729225	0.1830908	0.1439594	0.1684765
MAD Conformity		Nonconformity	Nonconformity	Nonconformity	Nonconformity
DF		-5.308788	-42.15151	60.69312	65.3418

*p-value < 0.05.

The rice, corn, and sugarcane datasets have been found to satisfy the properties outlined by Tošić and Vičič (2021), including mathematical combination, variety in the number figures, minimum number of observations, right-skewedness, and absence of predefined maximum and minimum values. As a result, we can confidently perform the NBL analysis on these datasets. The comprehensive analysis of the crop production datasets reveals significant variations in the leading digit distributions for rice, corn, and sugarcane concerning NBL expectations (Table 2 and Figure 1). For the rice dataset, the observed frequency of the digit "1" at 11.0% is lower than the expected 30.1%, while the corn dataset exhibits a leading digit "1" frequency of 51.0%, surpassing the expected proportion. Similarly, the observed frequency of the digit "1" for sugarcane, at 37.7%, was above the expected proportion. The substantial discrepancies in the leading digit distributions for rice, corn, and sugarcane raise concerns about data accuracy, data collection methodologies, or potential reporting irregularities. To delve further into potential explanations for these observed discrepancies, it is worth considering factors such as data entry errors and variations in data processing methods. Data entry errors, including typographical mistakes or misinterpretations during data collection, could contribute to deviations from the expected patterns. Additionally, differences in data processing methods over time or between crops may introduce variations that impact the reported digit frequencies. Examining these factors in future research could provide valuable insights into the root causes of the observed deviations.

Table 2 Results of the first digit analysis.

Digit	NBL	Rice	Corn	Sugarcane
1	30.1	11.0	51.0	37.7
2	17.6	26.2	24.8	3.8
3	12.5	24.8	2.8	9.4
4	9.7	21.4	0.7	5.7
5	7.9	4.8	1.4	5.7
6	6.7	3.5	0.7	13.2
7	5.8	6.9	4.8	9.4
8	5.1	0.7	8.3	11.3
9	4.6	0.7	5.5	3.8
n		145	145	53
χ^2		76.367*	67.505*	17.044*
Mantissa		0.19814*	0.23004*	0.1911*
MAD		0.074927	0.07167121	0.05330219
MAD Conformity		Nonconformity	Nonconformity	Nonconformity
DF		-4.521611	-18.64909	11.88005

*p-value < 0.05.

The statistically significant outcomes of the Chi-square test and Mantissa test for the crops underscore the presence of deviations from NBL expectations. Moreover, based on the MAD values, it is evident that all of the crops analyzed exhibit nonconformity to NBL expectations (Nigrini, 2012). The MAD values provide a measure of the average magnitude of deviations between the observed and expected leading digit frequencies, and their nonconformity further reinforces the need for thorough investigation and validation of the reported crop production data for rice, corn, and sugarcane. These tests provide valuable insights into the reliability of the reported crop production statistics and warrant further scrutiny of the data sources, collection methodologies, and reporting practices for each crop.

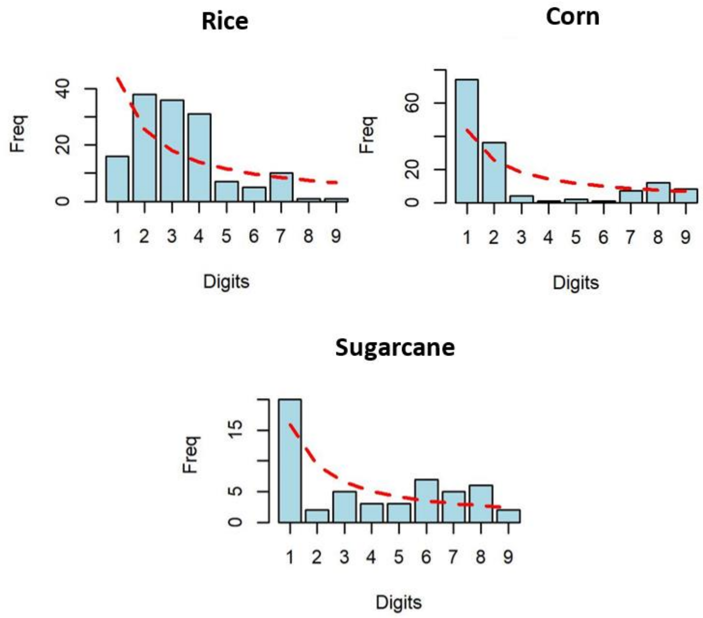

Figure 1 Relative frequencies of the first digits.

After conducting the initial analysis using the first digit test to assess the conformity of crop production datasets to NBL, first two-digit test was performed to further investigate the distribution patterns and potential anomalies within the data. The first two-digit test is a special statistical approach that provides additional insights into the reliability of the reported crop production statistics. This comprehensive assessment ensures a thorough evaluation of the datasets and enhances the validity of the findings. The application of the first two-digit test reinforces the significance of the results, highlighting potential areas of concern and offering valuable guidance for policymakers and stakeholders in the agricultural sector (Yanik & Samanci, 2013). The combination of both the first- and second-digit tests into first two-digit test contributes to a more comprehensive understanding of the conformity of crop production statistics to NBL and enhances the credibility of the research findings (Goh, 2020).

The first two-digit test results for rice, corn, and sugarcane crop production datasets reinforced the findings obtained from the first digit test. To provide potential explanations for the observed discrepancies in leading digit frequencies, we considered various factors. Data entry errors, data processing methods, and even changes in reporting practices over time could contribute to these discrepancies. For instance, variations in data collection procedures among different regions or time periods might influence the leading digit frequencies. It is important to acknowledge that historical data, while valuable, can be subject to biases and inaccuracies due to such factors. Therefore, these findings should prompt further investigation into the specific data collection and reporting processes for rice, corn, and sugarcane to ensure data accuracy and integrity.

Only the important results were included in Table 3 since there were 90 (values between 10 and 99) categories. The results of the first two-digit test indicate statistically significant deviations from NBL expectations (χ^2 values: 139.03, 141.89, and 112.08, respectively). Similarly, the Mantissa values for these crops (0.19814, 0.23004, and 0.1911, respectively) highlight the substantial differences between the observed and expected leading digit frequencies. The MAD values of rice, corn, and sugarcane indicate nonconformity to the NBL (Nigrini, 2012). Additionally, the Distortion Factor (DF) values provide insights into the overall deviation for each crop, with rice having -4.5216, corn with -18.6491, and sugarcane with 11.8801. Upon closer examination of the specific two-digit combinations, for rice, the five largest deviations occur with digits 10, 14, 30, 16, and 42, indicating significant deviations from the expected frequencies for these specific two-digit combinations. Similarly, corn shows the five largest deviations with digits 14, 22, 24, 16, and 10, reflecting significant differences between the observed and expected frequencies for these particular two-digit combinations. For sugarcane, the five largest deviations are with digits 11, 12, 10, 87, and 13, further emphasizing substantial deviations from NBL expectations for these specific two-digit combinations. These results highlight the presence of significant deviations in the first two-digit distributions for rice, corn, and sugarcane crop production statistics.

Table 3 Results of the first two-digit analysis.

Crop	χ^2	Mantissa	MAD	MAD Conformity	DF
Rice	139.03*	0.19814*	0.009325038	Nonconformity	-4.521611
Corn	141.89*	0.23004*	0.009223114	Nonconformity	-18.64909
Sugarcane	112.08*	0.1911*	0.01363337	Nonconformity	11.88005

*p-value < 0.05.

The application of NBL analysis to finance and accounting typically involves larger sample sizes compared to the crop production statistics considered in this research study. The disparity in sample sizes presented a limitation of the research study, which might lead to conflicting conclusions when applying different testing approaches (Nigrini, 2012; Koch and Okamura, 2020). Therefore, deviations from the NBL do not always indicate evidence of suspicious activities (Parreño, 2023). However, the results of various statistical tests consistently showed significant deviations in crop production, reinforcing our conclusion that the current agricultural reporting system lacks reliability and trustworthiness in providing data on crop production in the Philippines. While the sample size limitation is acknowledged, it is essential to recognize that the observed deviations in the crop production data are statistically significant, indicating potential issues with data accuracy and reliability. These deviations may not necessarily imply fraudulent activities but underscore the need for a comprehensive examination of data collection, reporting, and validation processes within the agricultural sector.

The results of the first digit and first two-digit NBL analyses highlight the importance of meticulous data validation and verification in agricultural statistics. The observed discrepancies in the leading digit distributions across the crop production datasets emphasize the need for transparent data collection, reporting, and validation processes to ensure the accuracy, reliability, and integrity of the reported crop production figures. Addressing these irregularities is fundamental to providing trustworthy and informed crop production statistics that can effectively guide decision-making, resource allocation, and sustainable agricultural practices in the Philippines.

4. Conclusions

The results of this research study underscore the significance of accurate and reliable crop production statistics for effective agricultural planning and policy-making in the Philippines. By applying the Newcomb-Benford Law (NBL) to the major crop production datasets, we have gained valuable insights into the distribution patterns and potential anomalies within the reported data. However, before conducting the NBL analysis, we rigorously assessed the datasets to ensure their suitability for the statistical tests, excluding coconut, banana, cassava, and pineapple due to violations of the NBL constraints. This preliminary check highlights the importance of adhering to the appropriate data requirements to avoid misleading results. Upon analyzing the rice, corn, and sugarcane datasets, the NBL analysis revealed statistically significant deviations from the expected distribution patterns. These discrepancies in the leading digit frequencies indicate potential data accuracy issues, discrepancies in collection methods, or reporting irregularities that require thorough validation. While the deviations do not conclusively suggest fraud, they do emphasize the need for meticulous examination and verification of the reported crop production figures. It is imperative to acknowledge the limitations of this study, including the absence of a detailed exploration of the root causes of these deviations. Further research is needed to delve into potential data processing errors, data collection methodologies, and other factors contributing to these discrepancies. A more comprehensive understanding of these issues can provide valuable insights into improving data accuracy and reliability in agricultural statistics.

The application of the first two-digit test further strengthens these findings, as it provides a more comprehensive understanding of the conformity of the datasets to the NBL. The findings emphasize the importance of transparent data collection and validation processes to ensure trustworthy and informed decision-making in the agricultural sector. The identification of deviations in leading digit frequencies within the datasets calls for further scrutiny and validation to enhance the integrity of the reported crop production figures, ultimately contributing to sustainable agricultural practices and fostering economic growth in the Philippines.

Future research should focus on investigating the root causes of these deviations, exploring potential data processing errors, and implementing more stringent data validation procedures. In addition, expanding the analysis to include other crops in the Philippines and conducting comparative studies between different regions, provinces, or time periods would be beneficial to assess the consistency and reliability of crop production statistics. By ensuring the reliability of agricultural statistics, we can effectively facilitate informed policy-making and strengthen food security in the Philippines, contributing to the overall growth and stability of the agricultural sector. These efforts will help address the limitations identified in this study and further enhance the quality of agricultural data for better decision-making and planning.

Ethical considerations

Not applicable.

Conflict of Interest

The author declares no conflicts of interest.

Funding

This research did not receive any financial support.

References

Apuke, O. D. (2017). Quantitative research methods: A synopsis approach. *Kuwait Chapter of Arabian Journal of Business and Management Review, 33*(5471), 1-8.

Balashov, V. S., Yan, Y., & 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.

Benford, F. (1938). The law of anomalous numbers. *Proc Am Philos Soc,* (78), 551-572.

Briones, R. M. (2021). Philippine agriculture: Current state, challenges, and ways forward.

Eckhartt, G. M., & Ruxton, G. D. (2023). Investigating and preventing scientific misconduct using Benford's Law. *Research Integrity and Peer Review, 8*(1), 1-10.

Fewster, R. M. (2009). A simple explanation of Benford's Law. *The American Statistician, 63*(1), 26-32.

Gauvrit, N. G., Houillon, J. C., & Delahaye, J. P. (2017). Generalized Benford's Law as a lie detector. *Advances in cognitive psychology, 13*(2), 121.

Geyer, C. L., & Williamson, P. P. (2004). Detecting fraud in data sets using Benford's law. *Communications in Statistics-Simulation and Computation, 33*(1), 229-246.

Go, A. W., & Conag, A. T. (2019). Utilizing sugarcane leaves/straws as source of bioenergy in the Philippines: A case in the Visayas Region. *Renewable energy, 132*, 1230-1237.

Goh, C. (2020). Applying visual analytics to fraud detection using Benford's law. *Journal of Corporate Accounting & Finance, 31*(4), 202-208.

Hanci, F. (2022). Application of Benford's law in agricultural production statistics. *Journal of the National Science Foundation of Sri Lanka, 50*(2), 387-393.

Jianu, I., & Jianu, I. (2021). Reliability of financial information from the perspective of Benford's law. *Entropy, 23*(5), 557.

Koch, C., & Okamura, K. (2020). Benford's law and COVID-19 reporting. Economics letters, 196, 109573.

Kraus, C., & Valverde, R. (2014). A data warehouse design for the detection of fraud in the supply chain by using the benford's law. *American Journal of Applied Sciences, 11*(9), 1507-1518.

Landicho, D., & Balendres, M. A. (2022). Possible incursion of cassava virus diseases: risks and potential threats to the Philippine cassava industry. Archives of Phytopathology and Plant Protection, 55(15), 1725-1749.

Levičar, S. (2020). Potential of Benford's law and machine learning based verification in agricultural logistics. In XIV. International Conference on Logistics in Agriculture, (p. 4). 63 (082)(0.034. 2)

Li, F., Han, S., Zhang, H., Ding, J., Zhang, J., & Wu, J. (2019, February). Application of Benford's law in Data Analysis. In Journal of Physics: Conference Series 1168(3), 032133. IOP Publishing.

Madahali, L., & Hall, M. (2020, June). Application of the Benford's law to Social bots and Information Operations activities. In 2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA) (pp. 1-8). IEEE.

Manrique-Hernández, E. F., Fernández-Niño, J. A., & Idrovo, A. J. (2017). Global performance of epidemiologic surveillance of Zika virus: rapid assessment of an ongoing epidemic. *Public Health, 143*, 14-16.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *Am J Math, 4*: 39.

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

Pandit, P., Pandey, R., Singha, K., Shrivastava, S., Gupta, V., & Jose, S. (2020). Pineapple leaf fibre: cultivation and production. Pineapple Leaf Fibers: Processing, Properties and Applications, 1-20.

Parreño , S. . (2023). Assessing the quality of dengue data in the Philippines using Newcomb-Benford law. *Sapienza: International Journal of Interdisciplinary Studies, 4*(3), e23039. https://doi.org/10.51798/sijis.v4i3.662

Qin, L., Han, S., Xing, L., Zhang, J., Zhang, J., Yang, R., & Wu, J. (2019, August). Application Research of Benford's Law in Testing Agrometeorological Data. In IOP Conference Series: Earth and Environmental Science, 310(5), 052030. IOP Publishing.

Tenkorang, F. (2022). Manipulation in Agricultural Commodities Futures Market: Application of Benford's Law.

Tošić, A., & Vičič, J. (2021). Use of Benford's law on academic publishing networks. *Journal of Informetrics, 15*(3), 101163.

Vičič, J., & Tošić, A. (2022). Application of Benford's law on cryptocurrencies. *Journal of Theoretical and Applied Electronic Commerce Research, 17*(1), 313-326.

Yamagishi, K., Gantalao, C., & Ocampo, L. (2021). The future of farm tourism in the Philippines: challenges, strategies and insights. Journal of Tourism futures.

Yanik, R., & Samanci, T. H. (2013). Benford's Law and a Practical Implementation in Public Sector About its Application to Accounting Data. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 17*(1), 335-348.