

33 **Abstract**

34 The wine industry plays an important role in many national economies, it combines agricultural
35 production with cultural heritage and global trade. In Portugal, it contributes significantly to
36 economic value, regional identity, and rural sustainability. As international wine markets
37 become increasingly complex, the information from financial and production data is essential
38 for regulation, policymaking, and economic analysis. Despite the growing emphasis on
39 viticulture and market dynamics, the wine sector data anomalies have attracted only limited
40 attention so far. This study introduces Benford's Law—a statistical method used to detect
41 irregularities in naturally occurring datasets. Applying first- and second-digit Benford's Law
42 tests to Portuguese wine industry data from 2014 to 2023, which includes company-level
43 financial statements and wine production figures, the analysis reveals data irregularities.
44 Nonconformity between empirical data and theoretical expectations refers to wine must
45 production data and profit before tax. The irregularities must be explained in a further and more
46 detailed survey. The study offers a novel application of Benford's Law in the wine sector.

47

48 **Key Words:** Benford's Law, wine production, reporting

49

50 **1. Introduction**

51 The wine industry plays a central role in many national economies, combining agricultural
52 production with cultural heritage, regional identity, and participation in global trade. In
53 Portugal, where viticulture is deeply embedded in rural structures and economic activity, the
54 availability of reliable financial and production data is essential for informed policymaking,
55 regulatory oversight, and sectoral analysis. As wine markets become increasingly globalised
56 and complex, the accuracy of reported figures—both at the firm level and in official production
57 statistics—assumes strategic importance for tax administration, subsidy allocation, and market
58 monitoring.

59 Although a substantial body of research has examined viticulture, oenology, climate impacts,
60 and market behaviour, relatively little attention has been devoted to the quality and reliability
61 of the numerical data underpinning these analyses. It contrasts with other empirical fields,
62 where data integrity assessment has become a standard component of validation. Benford's
63 Law, which describes the expected logarithmic distribution of digits in many naturally
64 occurring datasets, has been widely applied in areas such as forensic accounting,
65 macroeconomic reporting, environmental monitoring, and financial market analysis to detect
66 irregularities or patterns warranting further investigation [1, 2]. Its usefulness lies in its domain-

67 independence and sensitivity to reporting errors, structural anomalies, and potential
68 manipulation in large numerical datasets.

69 Given the wine sector's reliance on both firm-level accounting data and officially reported
70 production statistics, it represents a particularly relevant yet largely unexplored context for
71 applying Benford's Law. Existing empirical studies have focused mainly on macro-level
72 agricultural aggregates or individual commodity markets, leaving open the question of whether
73 digit distributions remain consistent across different types of wine-sector data and whether
74 deviations may signal issues in reporting practices.

75 This study addresses this gap by examining the conformity of Portuguese wine-sector data with
76 Benford's expected digit distributions. Using first- and second-digit tests evaluated through
77 Mean Absolute Deviation (MAD) and the Kolmogorov–Smirnov (K–S) test and Chi-square
78 test, the analysis covers company-level financial statements and official wine must production
79 statistics for the period from 2014 to 2023. Rather than treating conformity as proof of data
80 accuracy or deviations as evidence of manipulation, the study adopts a conservative perspective,
81 using Benford-based diagnostics as indicators of potential irregularities that may merit closer
82 scrutiny.

83 By integrating two distinct data sources within a single analytical framework, this research
84 contributes to the literature on wine economics and agribusiness by introducing a systematic
85 approach to assessing data quality. It provides empirical evidence on the consistency of reported
86 figures in an economically significant and heavily regulated sector, highlighting the potential
87 of Benford's Law as a complementary tool for enhancing transparency and confidence in
88 sectoral data used for economic, regulatory, and scientific analyses.

89 The paper is structured as follows. Section 2 reviews the relevant literature and outlines the
90 theoretical background and research gap. Section 3 describes the methodology and research
91 sample. Section 4 presents and discusses the results, and Section 5 concludes the study.

92

93 **2. Literature Review**

94 *2.1. Theoretical Background*

95 Research on the quantity of wine grapes and wine production provides fundamental insights
96 into the functioning, efficiency, and resilience of the wine industry. Studies in this area address
97 diverse issues such as yield optimisation, the economic implications of viticulture, tax policy,
98 climate change, and market behaviour. The majority of research focuses on viticultural and
99 oenological determinants of production volumes [3, 4], while relatively few studies investigate
100 the economic and regulatory consequences of these patterns [5–7].

101 Among the studies addressing these interrelated factors, Anderson et al. [8] provide a
102 comprehensive analysis of global wine production dynamics. They showed that climatic
103 variability, technological progress, and economic cycles jointly determine production
104 outcomes. Their findings highlight a general stabilisation of output in traditional European
105 regions and faster growth in emerging markets such as South America and Asia. These
106 observations support the notion that stable production promotes predictable market behaviour
107 and aids in policy design.

108 Earlier work by Jones et al. [4] identified several critical factors affecting production, including
109 vineyard management practices, technological innovation, and, in particular, climate change.
110 Their study demonstrated how rising temperatures have led to shifts in vineyard locations and
111 the adoption of adaptive techniques to maintain both yield and quality. More recently, Del Rey
112 and Loose [9] explored the economic implications of global wine production trends, noting that
113 fluctuations in production volumes have a pronounced effect on price volatility and trade flows.
114 While environmental and technological factors remain crucial determinants of production, a
115 growing body of research highlights the influence of economic and institutional conditions on
116 the performance and stability of the wine sector. Taxation policies also play an essential role in
117 shaping production decisions and market competitiveness. Anderson and Pinilla [8] and
118 Katunar et al. [10] examined how fiscal regulations influence production incentives,
119 emphasising that well-structured and transparent data are indispensable for designing equitable
120 tax systems. Behmiri et al. [5] further linked macroeconomic conditions—GDP growth,
121 exchange rates, and agricultural policy—to production trends across the European Union. In
122 parallel, Rickard et al. [6] investigated the effects of trade liberalisation on the EU and US wine
123 markets, illustrating how the interaction between international trade agreements and domestic
124 regulations affects production stability.

125 Together, fiscal, macroeconomic, and trade factors define the context in which wine producers
126 operate, although environmental conditions continue to play a decisive role in shaping
127 production outcomes. Finally, Ashenfelter and Storchmann [3] discussed the economic
128 implications of climate change, demonstrating how variations in temperature and precipitation
129 influence grape yields, wine quality, and production costs.

130 Overall, these studies highlight the complex interplay of environmental, technological, and
131 economic factors that determine wine production patterns. Understanding these interactions
132 relies heavily on the availability of accurate and transparent production data. Yet, despite its
133 importance, the usefulness of such data has rarely been evaluated systematically. To address

134 this gap, statistical tools capable of identifying irregularities in numerical data can be applied,
135 among which Benford's Law has proven particularly useful.

136

137 *2.2. Quality Data Measures – Benford's Law in Real-World Applications*

138 Benford's Law describes the expected frequency of digits in naturally occurring numerical
139 datasets. Rather than following a uniform distribution, the first digit in many datasets is
140 disproportionately likely to be small—most commonly "1"—with probabilities decreasing
141 logarithmically for higher digits [11]. This statistical regularity arises in datasets that span
142 several orders of magnitude and result from multiplicative or exponential processes.

143 Initially discovered in the natural sciences, this method has found widespread application in
144 economics, finance, and environmental monitoring. For example, various geophysical datasets
145 such as measurements of the Earth's geomagnetic field, earthquake depths, and seismic
146 velocities have been shown to conform closely to Benford's distribution [1]. In economics, the
147 method has become a cornerstone of forensic accounting, where it helps detect potential
148 manipulation or fraud in financial statements [2, 12]. Auditors routinely use Benford-based
149 analyses of assets, liabilities, and revenues to identify anomalies that warrant further
150 investigation [13, 14].

151 Benford's Law has also been applied in emerging areas such as cryptocurrency analysis, where
152 it helps identify suspicious transaction patterns [14]. In macroeconomics, it has been used to
153 evaluate the plausibility of reported figures—such as GDP, inflation, and fiscal data—
154 submitted by national governments to Eurostat, revealing significant irregularities in Greece's
155 statistics [12]. Likewise, it has served as a diagnostic tool for validating self-reported water-use
156 data in the United States [15] and detecting anomalies in ecotoxicity datasets [1].

157 The versatility of Benford's Law lies in its simplicity and universality. Because it does not
158 require prior assumptions about the data structure, it can be applied across disciplines to reveal
159 inconsistencies that might otherwise remain unnoticed. Given its proven effectiveness in
160 auditing and environmental monitoring, the method is well-suited to the validation of
161 agricultural and production datasets, where manual reporting and aggregation often introduce
162 potential errors. Despite its wide application, empirical research using Benford's Law in
163 agriculture and the wine industry remains scarce: according to Web of Science database data,
164 more than two hundred studies address its role in fraud detection and only a limited number
165 focus on agricultural data (twelve studies such as Hanci [16], Suzuki et al. [17], Novovic et al.
166 [18]) whereas only two refer to the wine sector [7, 19].

2.3. *Quality Data Measures – Benford's Law in Agricultural Data*

167 Benford's Law has increasingly been recognised as a valuable instrument for assessing the
168 quality of numerical data in agriculture. As a statistical regularity describing the distribution of
169 digits, it provides a straightforward yet powerful means of detecting irregularities and
170 improving the quality of agricultural datasets.

171 Several empirical studies have explored its application in agricultural and related contexts. In
172 the Philippines, Parreño [20] applied Benford's Law to crop production data—covering six
173 major commodities, including bananas—using both Chi-squared and Mean Absolute Deviation
174 (MAD) tests. The results confirmed that Benford's Law is applicable for evaluating data
175 integrity and identifying inconsistencies in national agricultural statistics.

176 A detailed investigation by Hanci [16] extended this approach to official production data in Sri
177 Lanka, encompassing multiple crop categories across several years. Employing first- and
178 second-digit conformity tests, Hanci found that most datasets closely followed Benford's
179 expected pattern, suggesting a generally reliable reporting system. However, moderate
180 deviations in some manually reported crops indicated localised inconsistencies. Importantly,
181 Hanci emphasised the preventive role of Benford's Law, arguing that its use could strengthen
182 national data collection and monitoring practices.

183 Complementing these findings, a large-scale study conducted in China applied Benford's Law
184 to agricultural and precipitation datasets spanning the period from 1951 to 2015. The analysis
185 revealed a high degree of conformity between actual and expected digit distributions,
186 confirming both internal consistency and robustness in long-term agricultural data [21]. Minor
187 deviations were attributed mainly to transcription errors and incomplete time series, again
188 demonstrating the method's diagnostic potential.

189 Beyond crop statistics, Benford's Law has also been successfully applied to fisheries and
190 commodity markets. Noleto-Filho et al. [15] evaluated small-scale fisheries data in Brazil,
191 discovering localised deviations caused by manual reporting, while Domínguez-Bustos et al.
192 [22] used the method to detect structural shifts in tuna catch data following the introduction of
193 total allowable catch (TAC) regulations. Similarly, Martínez-Sánchez [23] analysed Spanish
194 almond prices, noting weaker conformity in first-digit tests but stronger alignment in second-
195 and third-digit analyses. This outcome was attributed to the moderate variability of price data.
196 This observation may also hold for the wine sector, where regional and quality constraints
197 bound prices and production volumes.

198 Collectively, these studies demonstrate the adaptability and robustness of Benford's Law as a
199 tool for assessing data irregularities in agriculture. However, most research has concentrated on
200

201 single data sources or aggregate statistics, leaving open questions regarding the consistency
202 between official datasets and firm-level financial data. This gap is particularly relevant for
203 specialised branches of agriculture such as the wine industry, where both production statistics
204 and firm-level accounting data are systematically collected and reported.

205

206 *2.4. Research Gap and Contribution*

207 While Benford's Law has been widely employed to assess the usefulness of data in various
208 economic and agricultural contexts, its potential application within the wine industry has
209 received little attention. This study, therefore, addresses that gap by examining whether the
210 Law's expected digit patterns hold for both official production data and firm-level financial
211 statements. Existing studies have focused on macro-level agricultural statistics, such as crop
212 yields, production volumes, and fisheries data, confirming that Benford's distribution can serve
213 as a valuable indicator of data integrity. However, no prior research has examined whether the
214 same principles apply to a wine production sector that simultaneously produces official
215 agricultural statistics and firm-level financial data.

216 The wine sector provides a particularly suitable context for such analysis. It combines highly
217 regulated production environments with heterogeneous firm structures, varying from small
218 family-owned vineyards to large-scale cooperatives. Moreover, the coexistence of official
219 production records and independently reported accounting data raises questions about the
220 consistency of information across reporting levels. Understanding whether these datasets
221 conform to Benford's distribution provides valuable insight into data irregularities, which is
222 crucial for both data quality and the effectiveness of institutional reporting frameworks in
223 agribusiness.

224 This study aims to address this research gap by developing a dual-source Benford framework
225 for Portuguese viticulture. The contribution is threefold:

- 226 1. It establishes theoretical and practical criteria for selecting between first- and second-
227 digit tests based on the numerical dispersion and aggregation levels typical of
228 viticultural data.
- 229 2. It applies Benford analysis simultaneously to official production statistics and company-
230 level accounting data, assessing both internal validity and cross-source consistency.
- 231 3. It interprets deviations in light of the institutional and regulatory context of the
232 Portuguese wine industry, where market concentration, climate variability, and
233 reporting obligations may influence numerical regularities.

234 The following section outlines the methodological framework used to test the conformity of
235 Portuguese wine sector data with Benford's Law. It specifies the datasets examined, the
236 structure of the Benford tests applied (first- and second-digit), and the criteria for evaluating
237 conformity through statistical indicators such as the Chi-squared (Chi-2) test, the Kolmogorov-
238 Smirnov (K-S) test, and the Mean Absolute Deviation (MAD) test.

239

240 **3. Methodology, Research Sample**

241 The study draws on two primary data sources: the BvD Orbis database, which provides financial
242 information for companies in the wine sector, and the Statistics Portugal (ine.pt) database,
243 which provides data on wine production.

244 The criteria for selecting financial data from the BvD Orbis database were as follows:

- 245 1. Status – active companies
- 246 2. Country – Portugal
- 247 3. Sector activity NACE Rev. 2 - 1102 - Manufacture of wine from grapes

248 In total, BvD Orbis gives 1366 companies that match our search criteria. In the next step, the
249 dataset was manually refined, and cases with missing or incomplete data were excluded from
250 the analysis. The study period covers 10 years, from 2014 to 2023. The choice of the final year
251 (2023) reflects the availability of the most recently published financial statements.

252 The selected variables from companies' financial statements relate to wine production activities.
253 The first variable is sales revenues (turnover), i.e., the value of economic benefits companies
254 realise when selling wine. The pre-tax financial result is the second variable used to measure
255 the business's effectiveness. The choice of this pre-tax variable is related to the elimination of
256 tax optimisation used by companies and the possibility of tax management. In terms of income,
257 profit before tax and loss before tax are separately calculated. The separate analysis is consistent
258 with the assumption that there might be different irregularities for losses and profits.

259 The variables related to the companies' resources are total assets and inventories. Total assets
260 represent the company's total resources, whereas stock refers to inventories (e.g., materials,
261 work-in-progress, finished products). Wine producers' inventories disclose the value of current
262 assets at all stages of wine production, from the must to bottled wine, which is ready for sale
263 (from the extraction of the must to the final finished product). The variables from financial
264 statements for Benford's Law tests used in the survey follow the literature review
265 considerations. Ausloos et al. [24] analysed pre-taxation data (pre-tax income, separately for
266 profit and loss) and total assets. Nigrini [2] justifies the separate analysis of positive and

267 negative numbers, as there may be different causes for profit and loss manipulation. Revenues
268 are also frequently used for Benford's law irregularity analysis [25–27].

269 Data from the Statistics Portugal database (INE) were subject to the following selection:

- 270 1. Source - Statistics Portugal, Vegetable production statistics
- 271 2. Name of the table - Wine production declared in grape must (hl) by producers by
272 Vinification location (NUTS - 2013) and quality and colour of wine (New regulation) -
273 annual
- 274 3. Measure unit (symbol) - Hectolitre (hl)

275 The data included the locations of wine producers where they produce must through the
276 vinification process. The database lists 308 official geographical areas of Portugal
277 (municipales) where information on the grape must produced was collected. The data collected
278 is the same as the available financial data of companies in the wine production sector (2014 -
279 2023).

280 The irregularities in financial and non-financial data that occurred will be assessed using
281 Benford's Law. This Law is not limited to specific phenomena and units of measurement. It is
282 a universal concept that refers to analysing the distributions of digits occurring in measures of
283 various natural, financial, and non-financial phenomena. The frequency of occurrence of
284 specific digits in a given place in a number can be described by the logarithmic formula
285 proposed by Benford. We calculate the Benford distribution as follows [2]:

$$286 \quad BL(d) = \log \left(1 + \frac{1}{d} \right)$$

287 Where:

288 d – number of digits

289 Based on the calculation formula for the Benford distribution, you can obtain information about
290 the frequency of digits in the number in the first place from the left for the first-digit test. Figure
291 1 shows the Benford distribution for the first digit.

292

293

294

295

296

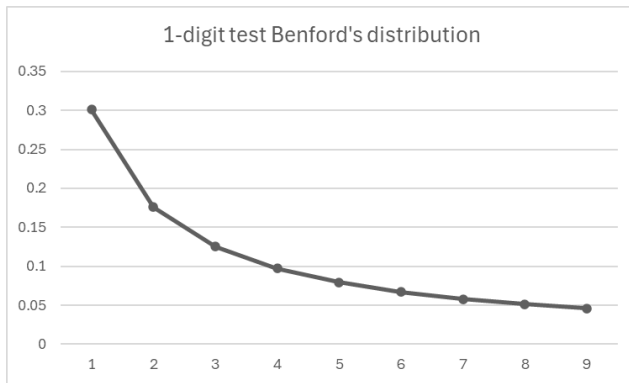
297

298

299

300 **Figure 1. Benford's distribution – 1-digit test**

301



302

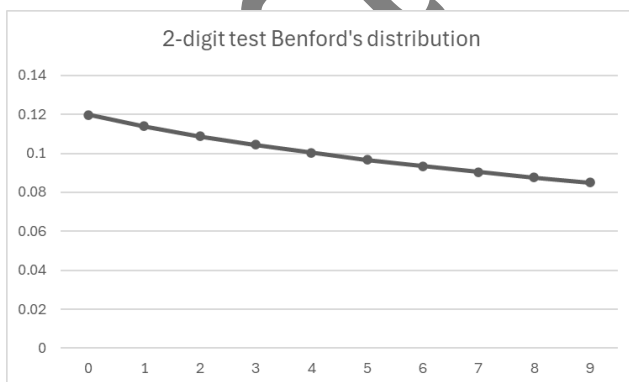
303 Source: Authors' calculation

304

305 Based on Figure 1, it can be observed that the digit 1 appears most frequently (30%) as the
306 leading digit in the data representing the analysed phenomenon, while the digit 9 appears the
307 least often (4%). Discrepancies between the actual digit distribution, such as assets or revenues,
308 and the expected Benford distribution may suggest the presence of anomalies that warrant
309 further investigation. These irregularities might stem from data manipulation, for instance, due
310 to rounding practices employed by companies when reporting financial information or errors in
311 data collection. Additionally, anomalies could result from the construction of the research
312 sample, such as restricting companies based on the size of a specific variable. In such cases,
313 analysing the second-digit distribution may provide further insight.

314

315 **Figure 2. Benford's distribution – 2-digit test**



316

317 Source: Authors' calculation

318

319 The first-digit test examines the frequency distribution of digits ranging from 1 to 9, whereas
320 the second-digit test analyses digits from 0 to 9. The second-digit test is particularly useful for
321 identifying underlying irregularities or inconsistencies within the data. [16].

322 The study of deviations between the empirical (for each variable) and theoretical (Benford)
323 distributions can be measured using the Kolmogorov-Smirnov, Chi-squared, and Mean
324 Absolute Deviation (MAD) test [1, 2, 20, 28]. The conformity with Benford's Law is tested by
325 Isaković-Kaplan, Demirović, and Proho [25] or Adahali and Hall [29], who used all three
326 approaches. MAD is also used by Půček, Plaček, Ochrana [30] or Van Caneghem [31]. The
327 previous application of the Chi-squared test for this purpose can be found in the research of
328 González [32], Cella and Zanolla [33], or Geyer and Drechsler [34]. Aggarwal, Dharni [35] and
329 Badal-Valero, Alvarez-Jareño, Pavía [36] have already used the Kolmogorov-Smirnov test for
330 this kind of comparison. The K-S test applies to the study of the fit of distributions for smaller
331 research samples [37]. Carno [38] analysed the conformity of empirical data (pandemic-related
332 data) with Benford's distribution through a literature review and found that the most frequently
333 used conformity tests were the Chi-square test (21 out of 26 studies), the Mean Absolute
334 Deviation (MAD) test (11 out of 26), and the Kolmogorov-Smirnov test (9 out of 26).
335 Moreover, seven studies used both Chi-square and K-S tests, eight used Chi-square and MAD,
336 and four employed all three methods simultaneously.

337 Figueiredo and Silva [39] utilised multiple conformity tests, including the Chi-squared, K-S,
338 and MAD tests. They observed that as the sample size increases, statistical tests tend to become
339 more sensitive, raising the likelihood of detecting smaller deviations. It has been noted that the
340 Chi-squared test is particularly sensitive to sample size [40]. However, the excessive statistical
341 power typically becomes noticeable only for datasets exceeding 5,000 records [2]. In our
342 research, the variables: profit before tax, loss before tax and wine production, all consist of
343 fewer than 5,000 observations, thereby mitigating this concern (Tables 2, 3 and 4). The other
344 variables, revenues, assets, and stock, exceeded 5,000 observations. In the study, the conformity
345 of distributions (both empirical and theoretical) is measured using the Chi-2 test, K-S test and
346 MAD [41].

347 MAD provides a more robust measure of conformity with Benford's Law than other statistical
348 measures. It can handle outliers, detect minor deviations, and is based on a strong statistical
349 foundation [2, 42, 43]. In this study, MAD:

350

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_i - m(X)|$$

351 Where:

352 $m(X)$ – average data value

353 n – number of data values

354 x_i – data values in the set

355 Depending on the level of MAD, it can be considered a close conformity, an acceptable
356 conformity, a marginally acceptable conformity, or a nonconformity. The ranges for MAD are
357 presented in Table 1. MAD does not test the hypothesis of whether two samples originate from
358 the same or different distributions. MAD provides the level of conformity.

359

360 **Table 1. Ranges for Mean Absolute Deviation**

Digit	Range	Conclusion
First-digit	0,000 - 0,006	close conformity
	0,006 - 0,012	acceptable conformity
	0,012 - 0,015	marginally acceptable conformity
	above 0,015	Nonconformity
Second-digit	0,000 - 0,008	close conformity
	0,008 - 0,010	acceptable conformity
	0,010 - 0,012	marginally acceptable conformity
	above 0,012	Nonconformity

361 Source: based on Nigrini [2]

362

363 Measuring the fit of MAD distributions is commonly used to determine the conformity of an
364 empirical distribution with the Benford distribution [44, 45]. The rejection threshold for
365 distribution conformity is 0.015 for the first-digit test. For the second-digit test, the
366 nonconformity threshold is 0.012.

367 The Kolmogorov–Smirnov test, in contrast, is a nonparametric statistic based on the empirical
368 distribution function [2, 25]. It is used to determine whether a random sample originates from
369 a specified continuous distribution.

370 Kolmogorov-Smirnov test (K-S)

$$371 \quad D_{n_1, n_2} = \sup_{-\infty < x < \infty} |F_{1, n_1}(x) - F_{2, n_2}(x)|$$

372 Where:

373 $F_{1, n_1}(x)$ – empirical distribution function of the first sample

374 $F_{2, n_2}(x)$ – empirical distribution function of the second sample

375

376 Two distribution functions are compared using the following test criterion:

377
$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_{n_1 n_2} > K_\alpha$$

378 Where:

379 n_1 - number of observations of the first sample

380 n_2 - number of observations of the second sample

381 K_α - test criterion at level α

382 The tested hypotheses are:

383 H_0 : Two univariate random variables come from the same probability distribution.

384 H_1 : Two univariate random variables do not come from the same probability distribution.

385

386 The Chi-square test (Chi-2) is a nonparametric statistical method used to determine whether an
387 observed frequency distribution differs significantly from an expected theoretical distribution
388 [2]. It is commonly applied to categorical data to assess how well the observed data fit a
389 specified model.

390 Chi-square test (χ^2)

391
$$\chi^2 = \sum_{i=1}^n \frac{(e - t)^2}{t}$$

392 Where:

393 n – total number of phenotypic classes

394 e – experimental frequency of the i -th class

395 t – theoretical frequency of the i -th class

396 The tested hypotheses are:

397 H_0 : The observed frequency distribution is consistent with the theoretical distribution.

398 H_1 : The observed frequency distribution is not consistent with the theoretical distribution.

399 Carno [38] reported that at least one of the conformity tests indicated nonconformity even when
400 other tests did not. Consequently, the assumption applied in our research is that if at least one
401 of the applied tests (Chi-2, K-S or MAD) indicates nonconformity between empirical and
402 theoretical distributions, this may serve as evidence of potential anomalies.

403 The analysis and statistical evaluation were processed using MS Excel for extracting digits and
404 the MAD test, and Statistica 13.3 for the Chi-2 test and the KS test.

405 4. Research results and discussion

406 Upon examining the financial data of companies from the wine production sector, a subset of
407 companies met the criteria described in the methodological section. The mere indication of a
408 company belonging to the NACE 1102 sector in Portugal does not mean that the BvD Orbis
409 database will include figures on financial results and company resources. The first-digit test
410 requires the first digit of the values of the relevant variables in the study to be analysed between
411 1 and 9. Table 2 presents the final research sample, excluding companies with missing financial
412 data from the database and those with variables of 0.

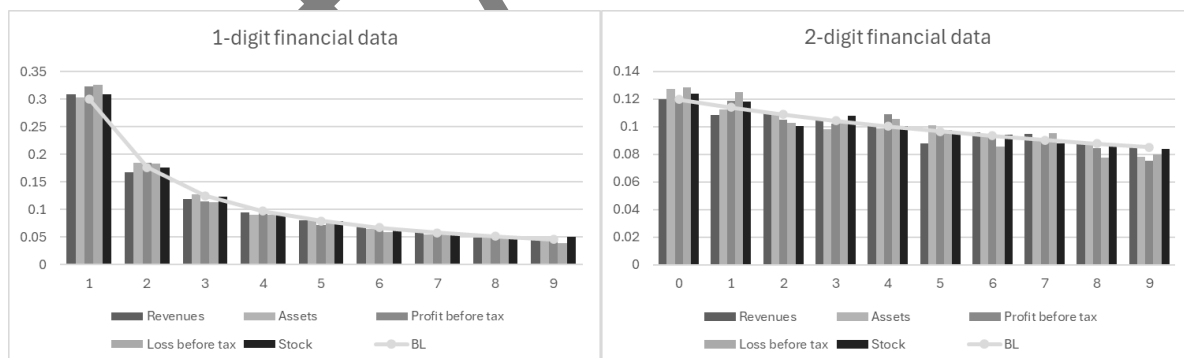
413 The second-digit test requires seeing which variables have values greater than 10. It is a
414 condition for the second digit in the number. Therefore, only those companies that reported data
415 for individual variables were selected from the database: stock, total assets, operating revenue,
416 or profit before tax with a value of at least 10. Table 3 shows the number of observations for
417 each variable.

418 Figure 3 illustrates the first- and second-digit tests for data from the financial statements of
419 companies in the Portuguese wine industry. It can be observed that the digits in the first and
420 second positions of operating revenue, total assets, profit before tax, loss before tax, and
421 inventories conform to the theoretical values implied by Benford's Law.

422

423 **Figure 3. Financial data: assets, revenues, profit before tax, loss before tax, inventories**
424 **(2014 - 2023)**

425



426

427 Source: Authors' calculation.

428

429 Table 2 presents the results of three goodness-of-fit tests for financial variables: revenues,
430 assets, profit before tax, loss before tax, and stock. It can be noted that for the profit before tax
431 variable, two tests indicated acceptance of the alternative hypothesis: chi-square (the observed
432 frequencies differ from the expected frequencies) and K-S (the sample does not follow the

433 specified distribution). The lack of consistency between the distributions demonstrated by the
 434 chi-square and K-S tests may suggest anomalies in determining profit before tax. A closer
 435 examination of profit formation in Portuguese wine sector companies would enable an
 436 assessment of the potential scale of earnings management practices in these companies.

437 **Table 2. 1-digit test Chi2, K-S and MAD results**

Variable	No. Obs.	Chi-2	decision	K-S test	decision	MAD	decision
Revenues	6766	12.061	conformity	0.500	conformity	0.004	close
p-value		p = 0.149		p > 0.05			conformity
Assets	7503	9.924	conformity	0.797	conformity	0.003	close
p-value		p = 0.270		p > 0.05			conformity
Profit before tax	4345	18.967	nonconformity	1.469	nonconformity	0.007	acceptable
p-value		p = 0.015		p < 0.05			conformity
Loss before tax	2506	14.945	conformity	1.135	conformity	0.008	acceptable
p-value		p = 0.060		p > 0.05			conformity
Stock	6098	7.128	conformity	0.485	conformity	0.003	close
p-value		p = 0.523		p > 0.05			conformity

438 Df = 8, at 5%, critical value Chi2 (8) = 15.507; critical value K-S test = 1.36

439 Source: Authors' calculation.

440
 441 For the remaining variables: revenues, assets, loss before tax, and stock, the results of the first-
 442 digit test using the chi-square and K-S goodness-of-fit measures indicate no basis for rejecting
 443 the null hypothesis that the observed counts for digits 1 to 9 do not differ from the expected
 444 counts based on the Benford distribution. The MAD measure additionally introduces a range of
 445 fit, and for all these variables, the range of fit is either close conformity or acceptable
 446 conformity.

447 **Table 3. 2-digit test Chi2, K-S and MAD results**

Variable	No. Obs	Chi2	decision	K-S test	decision	MAD	decision
Revenues	6321	9.746	conformit y	0.532	conformit y	0.003	close
p-value		p = 0.371		p > 0.05			conformit y
Assets	7190	12.310	conformit y	0.471	conformit y	0.003	close

		p = 0.196		p > 0.05			conformity
Profit before tax	3114	7.799	conformity	0.512	conformity	0.004	acceptable conformity
		p = 0.554		p > 0.05			
Loss before tax	1768	8.418	conformity	0.589	conformity	0.007	acceptable conformity
		p = 0.493		p > 0.05			
Stock	5561	6.433	conformity	0.447	conformity	0.003	close conformity
		p = 0.696		p > 0.05			

448 Df = 9, at 5%, critical value Chi2 (9) = 16.919, critical value K-S test = 1.36

449 Source: Authors' calculation.

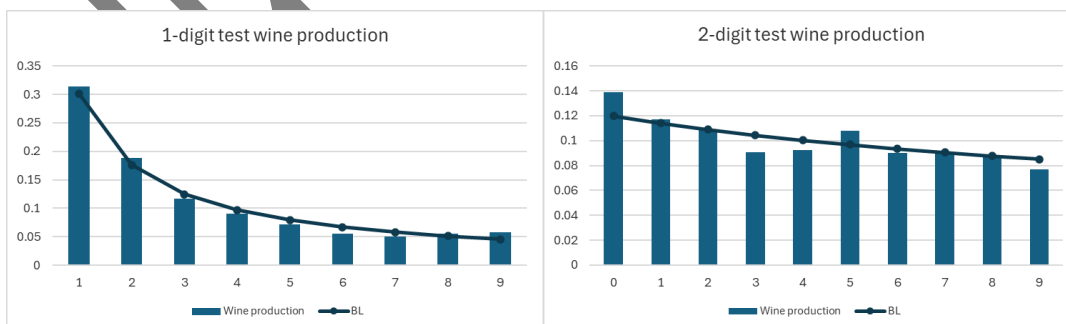
450

451 Table 3 presents the results for the second-digit test. All goodness-of-fit tests indicate goodness-
452 of-fit distributions or no basis for rejecting the hypothesis that there is no difference in observed
453 and expected numbers. For the second-digit test, the analysis ranged from digits 0 to 9.

454 The next step in the study is to verify the statistical data on wine must production in Portugal.
455 The study examined anomalies in the distribution of the amount of wine must produced in 308
456 regions of Portugal between 2014 and 2023 using the first- and second-digit tests. The data
457 collectively refer to the production of must from white and red grapes and various appellations:
458 Generous wine by protected designation of origin, wine by protected designation of origin, wine
459 by protected geographical indication, wine with grape variety indication, and wine without
460 certification.

461

462 **Figure 4. Wine production stat data (2014 – 2023)**



463

464 Source: Authors' calculation, based on INE database.

465

466 Table 4 presents the results of the first and second digit tests for the grape must production
 467 variable. The chi-square goodness-of-fit test supported the alternative hypothesis that the
 468 empirical and theoretical observation counts were significantly different for both the first and
 469 second digit tests. Although the remaining K-S and MAD goodness-of-fit tests did not reveal
 470 any discrepancies in the distributions, the chi-square test indicates irregularities that require
 471 further investigation.

472

473 **Table 4. 1-digit and 2-digit test Chi-2, K-S and MAD wine production results**

Variable	No.	Chi2	decision	K-S test	decision	MAD	decision
1-digit test wine production	2349	22.859 p = 0.004	nonconformity	0.835 p > 0.05	conformity	0.009	acceptable conformity
2-digit test wine production	2281	17.815 p = 0.037	nonconformity	0.758 p > 0.05	conformity	0.007	close conformity

474 Df = 8, at 5%, critical value Chi2 (8) = 15.507; critical value K-S test = 1.36

475 Df = 9, at 5%, critical value Chi2 (9) = 16.919, critical value K-S test = 1.36

476 Source: Authors' calculation.

477

478 Based on the results of the first-digit and second-digit tests (Table 4) for grape must production,
 479 it can be seen that the chi-square test indicates a lack of basis for accepting the null hypothesis
 480 of equality between the empirical and theoretical digit distributions. The lack of conformity
 481 between the distributions, as indicated by the chi-square test, may result from difficulties in
 482 obtaining this data by statistical offices [16]. Possible causes of nonconformity of distributions
 483 may include data entry errors or misinterpretation during data collection, as well as
 484 inconsistencies in data processing methods [20].

485

486 5. Conclusions

487 This study examined the application of Benford's Law to identify irregularities in financial and
 488 production data within the Portuguese wine industry. By examining company-level financial
 489 variables — such as operating revenues, total assets, profit before tax, loss before tax, and
 490 inventories — alongside official statistics on wine must production across 308 regions, the
 491 research provides a robust evaluation of data integrity over ten years (2014–2023). The first-
 492 and second-digit test results demonstrate strong conformity with Benford's expected digit

493 distributions, with Mean Absolute Deviation (MAD) values falling within the ranges of close
494 or acceptable conformity. To confirm the consistency of the distributions, the study employed
495 the Chi-squared and the K-S tests in addition to the MAD. The Chi-2 and K-S test results
496 indicate a lack of consistency in one case of the financial variable (profit before tax) and in the
497 wine production variable (wine must production).

498 These findings hold important implications. First, they validate the use of financial and
499 statistical data in subsequent economic modelling, sectoral studies, and policy assessments in
500 viticulture. Second, they demonstrate the applicability of Benford's Law in the wine production
501 sector, providing a replicable methodology for auditing financial and non-financial data.
502 Parreño [20] revealed significant deviations from the expected Benford distribution, indicating
503 potential issues with data accuracy, collection methods, or reporting irregularities. These
504 deviations highlight the importance of data validation in wine production statistics. The
505 problem of obtaining precise data in the agriculture sector was highlighted by Hanci [16]. Our
506 findings revealed that some irregularities were detected in the Portuguese wine sector, as
507 indicated in data collected by the statistical office.

508 The causes of irregularities may differ between companies that report profits and those that
509 report losses. A discrepancy in the distribution of pre-tax profits may suggest actions related to
510 corporate earnings management. The literature offers explanations for the detected irregularities
511 in financial figures, including rounding errors [34, 46, 47], fraud detection [2, 11], and
512 adjustments to comply with legal regulations [48, 49]. Revealed in the survey, nonconformity
513 in reported profit before tax indicates the threat of earnings management and financial figure
514 manipulation [2, 50]. The findings revealed the need to assess earnings management in the wine
515 production sector and identify the reasons for nonconformity. These actions prompt further
516 research to conduct a first two-digit test to select companies whose first-digit pre-tax profits
517 had the most significant deviations.

518 Moreover, the results of this study may inform the work of public institutions and regulatory
519 bodies, particularly in areas such as tax compliance, subsidy allocation, and agricultural
520 monitoring. By ensuring that datasets are reliable, authorities can base their decisions on solid
521 evidence. In the private sector, wineries and investors may employ similar analytical
522 approaches to audit internal data, enhance transparency, and mitigate the risk of reporting errors
523 or fraud.

524 However, the study has certain limitations. Benford's Law can identify anomalies, but does not
525 reveal the specific causes of irregularities. Additionally, the data may be affected by limitations

526 in availability or representativeness, particularly for smaller producers. Future research could
527 expand its scope to other wine-producing countries or examine trends in data quality over time.

528

529 References

- 530 [1] P. de Vries and A.J. Murk, “Compliance of LC50 and NOEC data with Benford’s Law:
531 An indication of reliability?,” *Ecotoxicology and Environmental Safety*, vol. 98, pp.
532 171–178, 2013, <https://doi.org/10.1016/j.ecoenv.2013.09.002>.
- 533 [2] M.J. Nigrini, *Benford’s law: Applications for forensic accounting, auditing, and fraud*
534 *detection*. 2012. <https://doi.org/10.1002/9781119203094>.
- 535 [3] O. Ashenfelter and K. Storchmann, “Climate Change and Wine: A Review of the
536 Economic Implications,” *Journal of Wine Economics*, vol. 11, no. 1, pp. 105–138,
537 2016, <https://doi.org/10.1017/jwe.2016.5>.
- 538 [4] G.V. Jones, E.J. Edwards, M. Bonada, et al., “17 - Climate change and its
539 consequences for viticulture,” in A.G. Reynolds, ed. *Managing Wine Quality (Second*
540 *Edition)*, Woodhead Publishing, 2022, pp. 727–778. [https://doi.org/10.1016/B978-0-](https://doi.org/10.1016/B978-0-08-102067-8.00015-4)
541 [08-102067-8.00015-4](https://doi.org/10.1016/B978-0-08-102067-8.00015-4).
- 542 [5] N.B. Behmiri, L. Correia, and S. Gouveia, “Drivers of wine production in the European
543 Union: a macroeconomic perspective,” *New Medit*, vol. 18, no. 3, pp. 85–96, 2019,
544 <https://doi.org/10.30682/nm1903g>.
- 545 [6] B.J. Rickard, Gergaud ,Olivier, Ho ,Shuay-Tsyr, et al., “Trade liberalization in the
546 presence of domestic regulations: public policies applied to EU and U.S. wine
547 markets,” *Applied Economics*, vol. 50, no. 18, pp. 2028–2047, Apr. 2018,
548 <https://doi.org/10.1080/00036846.2017.1386278>.
- 549 [7] N. Sequeira, M. Mota, R. Costa, et al., “The Influence of the COVID-19 Crisis on
550 Financial Statements Manipulations in the Portuguese Wine and Tourism Sector,” in A.
551 Abreu, J.V. Carvalho, P. Liberato, et al., eds. *Advances in Tourism, Technology and*
552 *Systems*, Singapore: Springer Nature, 2024, pp. 483–493. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-981-99-9765-7_42)
553 [981-99-9765-7_42](https://doi.org/10.1007/978-981-99-9765-7_42).
- 554 [8] , *Wine Globalization: A New Comparative History*. Cambridge: Cambridge University
555 Press, 2018. <https://doi.org/10.1017/9781108131766>.
- 556 [9] S.M. Loose and R. del Rey, “State of the International Wine Markets in 2023. The wine
557 market at a crossroads: Temporary or structural challenges?,” *Wine Economics and*
558 *Policy*, vol. 13, no. 2, pp. 3–14, Sept. 2024, <https://doi.org/10.36253/wep-16395>.
- 559 [10] J. Katunar, M. Grdinić, and D. Maradin, “EU Tax and Agricultural Policy in the Wine
560 Sector,” in B. Olgjić Draženović, V. Buterin, and S. Suljić Nikolaj, eds. *Real and*
561 *Financial Sectors in Post-Pandemic Central and Eastern Europe: The Impact of*
562 *Economic, Monetary, and Fiscal Policy*, Cham: Springer International Publishing,
563 2022, pp. 109–120. https://doi.org/10.1007/978-3-030-99850-9_7.
- 564 [11] T.P. Hill, “A statistical derivation of the significant-digit law,” *Statistical Science*, vol.
565 10, no. 4, pp. 354–363, 1995, <https://doi.org/10.1214/ss/1177009869>.

- 566 [12] C. Durtschi, W. Hillison, and C. Pacini, "The Effective Use of Benford's Law to Assist
567 in Detecting Fraud in Accounting Data," *J. Forensic Account*, vol. 5, Jan. 2004,
- 568 [13] S. Günnel and K.-H. Tödter, "Does Benford's Law hold in economic research and
569 forecasting?," *Empirica*, vol. 36, no. 3, pp. 273–292, 2009,
570 <https://doi.org/10.1007/s10663-008-9084-1>.
- 571 [14] J. Vičić and A. Tošić, "Application of Benford's Law on Cryptocurrencies," *Journal of*
572 *Theoretical and Applied Electronic Commerce Research*, vol. 17, no. 1, pp. 313–326,
573 2022, <https://doi.org/10.3390/jtaer17010016>.
- 574 [15] E.M. Noleto-Filho, A.R. Carvalho, M.J.F. Thomé-Souza, et al., "Reporting the
575 accuracy of small-scale fishing data by simply applying Benford's law," *Frontiers in*
576 *Marine Science*, vol. 9, 2022, <https://doi.org/10.3389/fmars.2022.947503>.
- 577 [16] F. Hanci, "Application of Benford's law in agricultural production statistics," *Journal*
578 *of the National Science Foundation of Sri Lanka*, vol. 50, no. 2, pp. 387–393, 2022,
- 579 [17] T. Suzuki, T. Kamimasu, T. Nakatoh, et al., "Identification of Unnatural Subsets in
580 Statistical Data," in *2018 7th International Congress on Advanced Applied Informatics*
581 *(IIAI-AAI)*, 2018, pp. 74–80. <https://doi.org/10.1109/IIAI-AAI.2018.00024>.
- 582 [18] D. Petrović, M. Novović, and M. Šoškić, "Comparative analysis of milling-bakery and
583 confectionery industry in Serbia based on Benford's Law," *Economic of Agriculture*,
584 vol. 72, no. 1, pp. 13–32, Mar. 2025, <https://doi.org/10.59267/ekoPolj250113P>.
- 585 [19] M. Parker and C. Jeynes, *A Maximum Entropy Resolution to the Wine/Water Paradox*.
586 2023. <https://doi.org/10.20944/preprints202306.1551.v1>.
- 587 [20] S.J.E. Parreño, "Analyzing crop production statistics of the Philippines using the
588 newcomb-benford law," *Multidisciplinary Science Journal*, vol. 6, no. 6, 2024,
589 <https://doi.org/10.31893/multiscience.2024079>.
- 590 [21] L. Qin, S. Han, L. Xing, et al., "Application Research of Benford's Law in Testing
591 Agrometeorological Data," in 2019, . <https://doi.org/10.1088/1755-1315/310/5/052030>.
- 592 [22] Á.R. Domínguez-Bustos, R. Cabrera-Castro, M.L. Ramos, et al., "Using Benford's
593 Law to Detect Possible Biases in Reported Catches of Tropical Tuna From the Indian
594 Ocean," *Fisheries Management and Ecology*, vol. 32, no. 1, pp. e12749, 2025,
595 <https://doi.org/10.1111/fme.12749>.
- 596 [23] F. Martínez-Sánchez, "Tracking the Price of Almonds in Spain," *Journal of*
597 *Competition Law and Economics*, vol. 17, no. 3, pp. 750–763, 2021,
598 <https://doi.org/10.1093/joclec/nhab002>.
- 599 [24] M. Ausloos, P.E. Sastroredjo, and P. Khrennikova, "Note on Pre-Taxation Data
600 Reported by UK FTSE-Listed Companies: Search for Compatibility with Benford's
601 Laws," *Stats*, vol. 8, no. 1, pp. 15, Mar. 2025, <https://doi.org/10.3390/stats8010015>.
- 602 [25] Š. Isaković-Kaplan, L. Demirović, and M. Proho, "Benford's law in forensic analysis
603 of income statements of economic entities in Bosnia and Herzegovina," *Croatian*
604 *Economic Survey*, vol. 23, no. 1, pp. 31–61, 2021, <https://doi.org/10.15179/ces.23.1.2>.

- 605 [26] X. Garza-Gomez, X. Dong, and Z. Yang, "Unusual patterns in reported segment
606 earnings of US firms," *Journal of Applied Accounting Research*, vol. 16, no. 2, pp.
607 287–304, 2015, <https://doi.org/10.1108/JAAR-04-2013-0031>.
- 608 [27] M.J. Nigrini, "An Assessment of the Change in the Incidence of Earnings Management
609 Around the Enron-Andersen Episode," *Review of Accounting and Finance*, vol. 4, no.
610 1, pp. 92–110, 2005, <https://doi.org/10.1108/eb043420>.
- 611 [28] P. Fernandes, Sé. Ciardhuáin, and M. Antunes, "Unveiling Malicious Network Flows
612 Using Benford's Law," *Mathematics*, vol. 12, no. 15, 2024,
613 <https://doi.org/10.3390/math12152299>.
- 614 [29] L.M. Adahali and M. Hall, "Application of the Benford's law to Social bots and
615 Information Operations activities," in 2020, .
616 <https://doi.org/10.1109/CyberSA49311.2020.9139709>.
- 617 [30] M. P\uuček, M. Plaček, and F. Ochrana, "Do the data on municipal expenditures in the
618 Czech Republic imply incorrectness in their management?," *Economics and
619 Management*, 2016,
- 620 [31] T. Van Caneghem, "NPO Financial Statement Quality: An Empirical Analysis Based
621 on Benford's Law," *Voluntas*, vol. 27, no. 6, pp. 2685–2708, 2016,
622 <https://doi.org/10.1007/s11266-015-9629-4>.
- 623 [32] F.A.I. González, "Self-reported income data: are people telling the truth?," *Journal of
624 Financial Crime*, vol. 27, no. 4, pp. 1349–1359, 2020, [https://doi.org/10.1108/JFC-08-
2019-0113](https://doi.org/10.1108/JFC-08-
625 2019-0113).
- 626 [33] R.S. Cella and E. Zanolla, "Benford's Law and transparency: An analysis of municipal
627 expenditure," *Brazilian Business Review*, vol. 15, no. 4, pp. 331–347, 2018,
628 <https://doi.org/10.15728/bbr.2018.15.4.2>.
- 629 [34] D. Geyer and C. Drechsler, "Detecting cosmetic debt management using Benford's
630 law," *Journal of Applied Business Research*, vol. 30, no. 5, pp. 1485–1492, 2014,
631 <https://doi.org/10.19030/jabr.v30i5.8801>.
- 632 [35] V. Aggarwal and K. Dharni, "Deshelling the Shell Companies Using Benford's Law:
633 An Emerging Market Study," *Vikalpa*, vol. 45, no. 3, pp. 160–169, 2020,
634 <https://doi.org/10.1177/0256090920979695>.
- 635 [36] E. Badal-Valero, J.A. Alvarez-Jareño, and J.M. Pavía, "Combining Benford's Law and
636 machine learning to detect money laundering. An actual Spanish court case," *Forensic
637 Science International*, vol. 282, pp. 24–34, 2018,
638 <https://doi.org/10.1016/j.forsciint.2017.11.008>.
- 639 [37] R.B. Sowby, "Conformance of Public Water Use Data to Benford's Law," *Journal -
640 American Water Works Association*, vol. 110, no. 12, pp. E52–E59, 2018,
641 <https://doi.org/10.1002/awwa.1161>.
- 642 [38] C.R.S. Carmo, F.C. Nunes, and F. de L. Caneppele, "The limits of conformity analysis
643 under the Newcomb-Benford law and the COVID-19 pandemic in Brazil," *Brazilian*

- 644 *Journal of Biometrics*, vol. 41, no. 3, pp. 234–248, Sept. 2023,
645 <https://doi.org/10.28951/bjb.v41i3.626>.
- 646 [39] D. Figueiredo and L. Silva, “Multiple conformity tests to assess deviations from the
647 Newcomb-Benford Law (NBL): A replication of Koch and Okamura (2020),” *Model*
648 *Assisted Statistics and Applications*, vol. 19, no. 1, pp. 117–122, 2024,
649 <https://doi.org/10.3233/MAS-231459>.
- 650 [40] A.W. Meade, E.C. Johnson, and P.W. Braddy, “Power and Sensitivity of Alternative
651 Fit Indices in Tests of Measurement Invariance,” *Journal of Applied Psychology*, vol.
652 93, no. 3, pp. 568–592, 2008, <https://doi.org/10.1037/0021-9010.93.3.568>.
- 653 [41] P. Luty, M. Petković, and R. Vavrek, “Applying Benford’s Law on assessing the
654 reliability of financial information in European companies from the rental and leasing
655 sector before and after the adoption of IFRS 16,” *Zeszyty Teoretyczne Rachunkowosci*,
656 vol. 46, no. 4, pp. 51–68, 2022, <https://doi.org/10.5604/01.3001.0016.1302>.
- 657 [42] R. Cerqueti and M. Maggi, “Data validity and statistical conformity with Benford’s
658 Law,” *Chaos, Solitons & Fractals*, vol. 144, pp. 110740, Mar. 2021,
659 <https://doi.org/10.1016/j.chaos.2021.110740>.
- 660 [43] C. Goh, “Applying visual analytics to fraud detection using Benford’s law,” *Journal of*
661 *Corporate Accounting and Finance*, vol. 31, no. 4, pp. 202–208, 2020,
662 <https://doi.org/10.1002/jcaf.22440>.
- 663 [44] D. Figueiredo Filho, L. Silva, and H. Medeiros, ““Won’t get fooled again”: statistical
664 fault detection in COVID-19 Latin American data,” *Globalization and Health*, vol. 18,
665 no. 1, 2022, <https://doi.org/10.1186/s12992-022-00899-1>.
- 666 [45] H.R. Morales, M. Porporato, and N. Epelbaum, “Benford’s law for integrity tests of
667 high-volume databases: a case study of internal audit in a state-owned enterprise,”
668 *Journal of Economics, Finance and Administrative Science*, vol. 27, no. 53, pp. 154–
669 174, 2022, <https://doi.org/10.1108/JEFAS-07-2021-0113>.
- 670 [46] M. Miranda-Zanetti, F. Delbianco, and F. Tohmé, “Tampering with inflation data: A
671 Benford law-based analysis of national statistics in Argentina,” *Physica A: Statistical*
672 *Mechanics and its Applications*, vol. 525, pp. 761–770, July 2019,
673 <https://doi.org/10.1016/j.physa.2019.04.042>.
- 674 [47] C.J. Skousen, L. Guan, and T.S. Wetzel, “Anomalies and unusual patterns in reported
675 earnings: Japanese managers round earnings,” *Journal of International Financial*
676 *Management and Accounting*, vol. 15, no. 3, pp. 212–234, 2004,
677 <https://doi.org/10.1111/j.1467-646X.2004.00108.x>.
- 678 [48] B. Demir and B. Javorcik, “Trade policy changes, tax evasion and Benford’s law,”
679 *Journal of Development Economics*, vol. 144, pp. 102456, May 2020,
680 <https://doi.org/10.1016/j.jdeveco.2020.102456>.
- 681 [49] P. Luty and Z. Zawolska, “Detecting Anomalies in Tax Revenues Using Benford’s
682 Law. The Case of Polish Adjustment,” *Central European Business Review*, vol. 14,
683 June 2025, <https://doi.org/10.18267/j.cebr.401>.

684 [50] D. Amiram, Z. Bozanic, and E. Rouen, "Financial statement errors: evidence from the
685 distributional properties of financial statement numbers," *Review of Accounting*
686 *Studies*, vol. 20, no. 4, pp. 1540–1593, 2015, [https://doi.org/10.1007/s11142-015-9333-](https://doi.org/10.1007/s11142-015-9333-z)
687 z.

688

Just Accepted