An application of Benford's law in the NASDAQGS health care industry

Raymond A. K. Cox
Thompson Rivers University

Natalia Popa Antalovschi
Thompson Rivers University

ABSTRACT

This paper investigates the price distribution of stocks listed in the NASDAQGS Health Care Industry for congruence to the expected price distribution of Benford's Law (BL). BL mirrors a logarithmic function that predicts the leading digits of prices will have a higher probability of lower numbers such as 1, 2, 3 versus higher numbers such as 7, 8, 9. The statistical analysis utilized the chi-square test for the distribution of numbers by digit as well as the z-statistic for the individual numbers by digit. The findings show the first digit, second digit, and fourth digit frequencies of the NASDAQGS conform to BL. Nevertheless, the third digit probabilities did not support the hypothesis of alignment to the BL forecast. These results strongly support the efficacy of BL as typically only the leading two digits comply with BL and trailing digits normally do not fit with BL.

Keywords: Benford's law, equity prices, chi-square, logarithmic law

INTRODUCTION

Originally, in 1881, Simon Newcomb, while exploring logarithm books to carry out logarithmic calculations observed that the earlier pages were far more worn than later pages. This phenomenon indicated that the leading digits were more likely to be the initial numbers (from the set of 1, 2, 3, 4, 5, 6, 7, 8, 9) in mathematical computations and studies. That is, not all numbers are created equal. The probability distribution for the first and second digit was postulated for a series of numbers. Note, digits beyond the first digit (that is, second, third, and fourth) can take on the integers (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Subsequently, an empirical study (Benford, 1938) explored Newcomb's hypothesis by gathering evidence from a total of 20,229 observations from 20 diverse groups such as city populations, molecular weights, addresses and death rates. Afterwards this statistical distribution became known as Benford's Law (BL).

This study focuses on the distribution of stock prices that comprise the NASDAQGS Health Care Industry Index. This category is of particular importance as the US health sector has sustained extraordinary growth in the past several decades and is expected to continue to exceed the US GDP growth rate in the foreseeable future. This research will show that despite expanding health company sizes with increasing stock prices the digit frequency as stated in BL holds.

There has been an explosion of interest in BL across various fields. A relevant application is in the detection of fraud, for example, academic research, voting, and auditing. Given the abundance of BL research the following is a selection of literature, to some extent, concentrated on the finance and accounting research.

LITERATURE REVIEW

Carslaw (1988) investigates if the frequency of occurrence of numbers appearing as the second digit in income numbers of New Zealand companies conforms to the expected random distribution. It was found that there is a much higher than expected frequency of zeros and a less than expected frequency of nines as the second-from-the-left-most digit in reported earnings. This abnormality may provide evidence of goal oriented or goal achieving behaviour.

Thomas (1989), extending Carslaw's paper, offers a preliminary investigation of the rounding of earnings phenomenon. Using data that concerned earnings of U.S firms on COMPUSTAT he confirms Carslaw's findings, but the deviation magnitude observed for U.S. firms are less severe than of those observed for zeros and nines in the New Zealand sample. Additionally, it was found that firms reporting losses exhibit the opposite patterns (fewer zeros and more nines).

Nigrini and Mittermaier (1997) in their empirical study focuses on BL as a mathematical basis of digital analysis tests on: first digits, second digits, first-two digits, number duplication, rounding (the multiple test), and last-two digits. Their main case study is a description of the analysis performed in June 1996 with the Internal Audit Department of a New York Stock Exchange (NYSE) listed oil company. The original data set consisted of 30,084 invoices authorised for payment by the accounts payable system of a business segment which included the dollar amount, payee, expense code, cost centre and payee address fields. After deleting the data elements that were negative numbers (credit vouchers) and numbers less than $10 their data set contained a sample size of 28,736. Using the z-statistics for the first digits it was found that only the differences for the numbers 5, 6, and 9 were not significant at the 0.01 level, However, based on the low mean absolute deviation (MAD) of 0.44 percent it was concluded that the first digits conform to BL. Additionally, it

was concluded that the second digits also conformed to BL based on the low MAD of 0.53 percent. It should be mentioned that the second digit-numbers 0 and 5 have actual frequencies that exceed those of BL.

Das and Zhang (2003) utilizing the concept of BL detected firms rounding-up earnings per share (EPS) to report a one cent EPS increase. Skousen, Guan, and Wetzel (2004) documents pervasive evidence that managers of Japanese firms tend to engage in earnings manipulative activities of rounding earnings numbers to achieve key reference points. Further, similar to other studies ((Carslaw, 1988), (Thomas, 1989), and (Das and Zhang, 2003)) they found that the first digit of earnings numbers is often emphasized by the management. Their analysis includes annual data for net incomes from 1974 to 1997 resulting in a sample of 1,871 companies and 37,900 annual earnings observations. They also found that the second digit, third digit, or even the fourth digit in earnings serves as the the reference points of the rounding earnings behavior. Finally, their results show that the incentives of rounding earnings numbers are negatively associated with the distance of pre-rounded earnings to the next reference point.

Cho and Gaines (2007) explored the data on campaign finance, a field rife with allegations of fraud, cheating, and corruption. For their study they used first digit relative frequencies analysis for all committee-to-committee, and inkind contributions cataloged by the Federal Election Commission (FEC) for each of the six US election cycles during the period from 1994 to 2004. The analysis was done on both aggregate and disaggregate FEC data, according to the size of the contribution. A casual perusal reveals that the fit to the Newcomb-Benford theoretical distribution for the FEC data seems to have gotten worse over time.

Johnson (2009) conducted an exploratory research by using BL to determine if selected company characteristics are associated with a risk of earnings management. Net income and earnings per share quarterly data were collected from the EDGAR data base of the Securities and Exchange Commission (SEC) for twenty-four randomly selected publicly traded companies for fiscal years 1999 through 2004. Companies were classified and analyzed by size (capitalization), age (period of time publicly traded), level of reported insider trading, and Capital Asset Pricing Model (CAPM) stock beta values. Findings show that companies categorized as low capitalization (below USD 45 billion), higher levels of inside trading (3% and higher), and newer to being traded on the public markets (less than 25 years) represent a potential risk of earnings management. Moreover, the mean absolute deviations and correlations were also calculated and their results support the Benford digit breakdown.

Krakar and Žgela (2009) researched to determine if foreign payment message amounts in the Croatian banking system follow BL. The results of the chi-square, z-statistics, and mean absolute deviation tests show that foreign payment messages, when analysed without focusing on special types of messages or certain business entities, do not conform to BL. They conducted additional examinations on specific message types and explained deviations from BL frequencies. It was noticed that as they focused on smaller data subsets of payment messages, some subsets were getting close to BL distribution. On top of that, it was concluded that application of BL is very effective in the auditing of information systems, specifically foreign payment systems.

Hickman and Rice (2010) investigated the extent to which crime statistics conformed to BL for the year 2006 at the national and state level in the US, and the period of 1960-2002 for the local level. By utilizing the first, second, first-two, first-three, and last-two digits their study found that national and state-level summary UCR (Uniform Crime Reporting) data conform to BL.

Archambault and Archambault (2011) explored the existence of financial statement manipulation on U.S. firms by using the BL analysis. Based on the 1915 Moody's Analyses of Investments (before the creation of the US Securities Exchange Commission) the sample of firms were broken into industrial companies, which faced no U.S. federal government regulation, and railroads and utilities, which did have government imposed rate regulations. This study showed results that are consistent with the assertion that regulated and unregulated companies managed reported results and did so in different ways. The only common area of manipulation was current liabilities. While unregulated companies manipulated revenue and income, regulated companies were more likely to manipulate amounts in non-value maximizing ways and did so in accounts that were more difficult to verify. This suggests that the regulatory process provided some scrutiny of results and did influence the types of manipulations made.

Jordan and Clark (2011) extended previous work ((Thomas, 1989) and (Das and Zhang, 2003)) by testing for the presence of cosmetic earnings management (CEM) by investigating Compustat data for two time periods, 1997 to 2000 and 2003 to 2006, i.e., before and after the implementation of the Sarbanes Oxley Act (SOX) in 2002. Based on the first two digits analysis versus Benford's expected distributions it was found that the period 1997 to 2000 shows a continuation of the CEM that had been documented in earlier studies while the period 2003 to 2006 indicates a noticeable decline in the level of CEM. Further empirical evidence of earnings management by rounding up EPS was provided (Van Caneghem, 2002) with a British sample.

Henselmann, Scherr, and Ditter (2012) analyzed digit frequencies for a sample that is based on all available SEC XBRL 10-K reports (XBRL – eXtensible Business Reporting Language) filed with the SEC EDGAR system for the fiscal year 2012. Extracting all monetary line items that are contained in single annual reports, on average, they found that the distribution of first digits conforms very well to the expected distribution according to BL. Furthermore, their results indicate several line items with an abnormal digit frequency potentially indicating human interaction.

Tilden and Janes (2012) used BL to investigate the occurrence of the intentional manipulation of reported financial statement numbers (net sales, net income, inventory, and allowance for doubtful accounts) during economic recessionary times. In an examination of financial reporting data surrounding recessions occurring from 1950 to 2006, this study provides evidence of increased financial statement manipulation during economic recessions. Results strongly indicate the presence of manipulated or falsified data in allowance for doubtful accounts and net income, whilst weaker evidence of manipulations in inventories and net sales. Moreover, the results of this study indicate that during recessionary times, there is a certain level of financial statement manipulation that goes undetected; most likely because the manipulations are corrected when the economy improves and are not exposed by events such as bankruptcy. It is also important to note that the tests in their study cannot distinguish between manipulations that may be within the parameters of generally accepted accounting procedures and those that may cross the line into fraud.

Hsieh and Lin (2013) examined the extent of window dressing behaviour (specifically reporting rounded earnings) among firms in the US marine industry. Their findings suggest that window dressing is a significant practice among the marine firms. However, the extent of the pervasiveness of such behaviour is less severe among marine firms than among all publicly listed firms in the entire US economy, suggesting that the quality of financial statements of marine companies is higher than the overall population of public companies.

Amiram, Bozanic, and Rouen (2015) analysed 43,332 financial statement numbers from numerous accounts employing the match up to the expected BL digit to check for fraud in US public company reports . Significant misrepresentation was found.

Horton, Kumar, and Wood (2020) tested whether BL can be used to differentiate retracted academic papers that have employed fraudulent/manipulated data from other academic papers that have not been retracted. By using the case of Professor James Hunton who had 37 of his articles retracted because there were grave concerns that they contained mis-stated or fabricated datasets, their results clearly indicate that Hunton's retracted papers significantly deviate from BL, relative to the control group of papers. In additional analysis it was also found that these results are generalisable to other authors with retracted papers.

Hypothesis

We state the hypotheses as follows:

H_{0i}: The occurrence of numbers in the ith place (digit) of a stock price will conform to the expected Benford's Law distribution,

where i can take the value of 1, 2, 3, or 4.

The alternative hypothesis posits that the occurrence of numbers in the 1st, 2nd, 3rd, or 4th place (digit) of stock price does not conform to the expected distribution.

SAMPLE AND METHODOLOGY

All US Health Care Industry public firms listed on the NASDAQ Global Select Market (NASDAQGS) are selected in the initial sample. The last business day of the month Daily Close Prices are collected from the Capital IQ (of Standard and Poor's) database for the years 2020 and 2021. After observations with missing data are deleted, the final samples contain 238 and 296 firms corresponding to 2,856 and 3,552 firm-months in 2020 and 2021 respectively. Trades are sometimes reported with prices with three decimal places (even though the prices are above $1). There is SEC Rule 612 (Minimum Pricing Increment) of Regulation NMS requiring prices for securities below USD 1.00 to be quoted with four decimal places and above USD 1.00 with two decimal places.

The data from S & P Capital IQ is downloaded onto an Excel file and then uploaded into SPSS Statistics (Statistical Package for the Social Sciences) software for statistical analysis. The chi-square goodness of fit test and z-statistics are the two statistical tests used in this paper to measure deviations from expected Benford's frequencies on the first digit, the second digit, the third digit, and the fourth digit of stock prices. The chi-square goodness of fit test is used as an all-digit-at-once test, while the z-statistic is the correct test to assess whether the actual proportion for a specific number by each digit differs significantly from the expectation of BL.

Benford postulated that the expected occurrence of a number as the first significant digit d is:

$$P(d\ is\ the\ first\ digit) = log_{10}\left(\frac{1}{d}+1\right), d = 1,2,\dots,9$$

The general form of BL that specifies the probabilities of occurrence of the first and higher significant digits, and more generally, the joint distribution of all the significant digits is:

$$Prob(n,d) = \left(\sum_{j=10^{n-2}}^{10^{n-1}-1} log_{10}\left(1+\frac{1}{10j+d}\right)\right),$$

where n is the nth leading digit $n > 1$, d is an integer in $\{0,1,\dots,9\}$.

As indicated in Table 1 (Appendix), this contains the BL distribution for the first four digits, i.e. the expected occurrences of each number zero, one, two, three, four, five, six, seven, eight, and nine in the first, second, third, and fourth (digit) places.

DESCRIPTIVE ANALYSIS OF STOCK DATA

As indicated in Tables 2 and 3 (Appendix), the content contains the descriptive analysis of the NASDAQGS Health Industry stock prices for the last trading day of each month (January through December) for the years 2020 and 2021 respectively.

The monthly mean stock price (end of the month, last business (trading) day) for 2020 and 2021 ranged from $41.29 to $63.43. The minimum and maximum monthly stock prices for individual stocks comprising the NASDAQGS health sector index were $0.35 and $719.19 respectively. The stock price is highly positive skewed with the skewness statistic ranging from a low of 3.057 in December to 4.606 in January in 2020, and in 2021 the low of 3.376 occurred in January and the high of 3.928 happened in December. Kurtosis is greatly peaked with the trough in December of 11.321 and peak in January of 32.397 in 2020, whereas for 2021 the low of 13.996 is in January and the high of 17.700 is in December. Clearly there is an end-of-the-year and beginning-of-the-year effects in the stock price distributions. Other statistics provided for each month across the 2 years include the lower quartile (25 percentile), median (50 percentile), upper quartile (75 percentile), and standard deviation.

CHI-SQUARE TEST

The chi-square goodness of fit test was performed to compare if collected stock price data from the NASDAQGS Health Care Industry conforms to BL for each of the first 4 digits of the stock price. The null hypothesis is that the digits conform to BL. The chi-square statistic is calculated as is shown in Equation 1:

$$\chi^2 = \sum_{i=1}^{K} \frac{(AC-EC)^2}{EC}, \quad (1)$$

where AC and EC represent the actual count and expected count respectively, and K represents the number of bins (in our case this equals to 9 or 10 depending on if it is the first digit or instead the second, third, or fourth digit respectively). The number of degrees of freedom is $(K-1)$, which for the first digit, the test is evaluated using 8 degrees of freedom. For the second, third, and fourth digits the degrees of freedom are 9 as there is an additional possible digit, namely zero.

Before proceeding with the chi-square goodness of fit test, it is necessary to check the two important assumptions of this test:

1. The independence of observations – this assumption is met, since each observation recorded in the contingency table have one entry in the chi-square table.
2. The expected frequencies should be greater than 5. Although it is acceptable in larger contingency tables to have up to 20% of expected frequencies below 5, the result is a loss of statistical power. Even in larger contingency tables no expected frequencies should be below 1.

Since the above assumptions are met, the chi-square statistical test was suitable.

The criterion values to test for statistical significance using the chi-square test at an alpha level of 5%, one tail as all chi-square goodness of fit tests are one tail, for degrees of freedom 8 and 9 are 15.5073 and 17.5346 respectively.

Z-STATISTIC

The second statistical test used in our research to evaluate whether the actual proportion for a digit differs significantly from the expectation of BL is the z-statistic:

$$Z = \frac{|p_a - p_e| - \left(\frac{1}{2n}\right)}{\sqrt{\frac{p_e(1-p_e)}{n}}} \quad (2)$$

where p_a denotes the actual proportion, p_e denotes the expected proportion, and n denotes the number of stock prices (sample size). The second term in the numerator $(1/2n)$ is a continuity correction term to bring normal and binomial probability curves into closer agreement and is used only when it is smaller than the first term in the numerator. For an alpha significance level of 5%, two-tail test, the cut-off level is 1.96.

Note, the chi-square goodness of fit test is used as an all-numbers-at-once test (each of the numbers zero, one, two, three, four, five, six, seven, eight and nine is evaluated for each digit (first, second, third, and fourth) separately. Whereas the z-statistic is a test to appraise whether the actual proportion for a specific number (done for each digit individually) differs significantly compared to the BL predicted probability of occurrence.

RESULTS

As indicated in Tables 4 (for 2020) and 5 (for 2021) (Appendix) this reports the observed actual proportions of the numbers for each month in panels A, B, C, and D for the first, second, third, and fourth digits respectively. The first entry in each cell reports the observed deviation percent between actual and predicted (BL) proportions, while the second entry reports the z-statistic values. The bottom four rows of each of the panels of the tables reports the sample size, the chi-square values, the degree of freedom, and the p-values for the chi-square test for each of the 24 months across the 2-year period.

Observing Table 4, Panel A, (Appendix) for the first digit (across the 12 months of 2020) we can see that there are four out of 108 z-statistic values that exceed the cut-off level of 1.96, leading us to conclude that the actual proportion differs significantly from the expected proportion for these four cases. These four cases, with z-statistic values of 2.72, 2.29, 2.58, and 2.29, are associated with the number 4 for June 30 and July 31, with number 7 for October 30, and with number 9 for June 30 respectively. Based on the chi-square test statistic of 16.86, the p-value of 0.032 for June 30 is less than the 0.05 level of statistical significance, a difference from BL mostly due to digits 4 and 9. The results in Panel B in Table 4 (Appendix) for the second digit reveal that there is only one significant z-statistic value of 4.019, associated with digit 4 for March 31, 2020, that exceeds the cut-off level of 1.96. The result of this test is consistent with the chi-square test statistic of 21.54 for the same month-year, as the p-value of 0.010 is less than 0.05. Looking at Panel C in Table 4 (Appendix) for the third digit indicates nonconformity to BL for the January 31 and April 30 dates in 2020 with chi-square test statistics of 18.71 and 19.86 associated with p-values of 0.028 and 0.019 respectively. Lastly for 2020 seeing Panel D of Table 4 (Appendix) for the fourth digit each month the stock price numbers distribution fits with the forecast of BL. This is strong evidence of the efficacy of BL as many researchers find poor results at the fourth digit level. That said, the z-statistics display nonconformity for the number 0 on January 31 and April 30, number 4 on March 31, and number 8 on February 28 and September 30, in 2020.

Viewing Table 5 for 2021, Panel A (Appendix) for the first digit only June 30 is not statistically significant fitting with BL with a chi-square test statistic of 15.73 (p-value of 0.046). Otherwise, there are 5 individual months (April 30, May 28, July 30, August 31, and September 30) where the number 0 frequency does not match the BL prediction. Likewise, the number 2 has 2 months (May 28 and June 30) along with 1 month (November 30) for number 5 that is statistically significant from the forecast. Eyeing Table 5, Panel B, (Appendix) the chi-square tests in all month's dovetail with BL. However, using the z-statistics at the numbers level, the number 3 on October 29, and the number 4 on November 30, as well as the number

6 on May 28, plus the number 8 on September 30 and finally the number 9 on August 31 are statistically different to what is predicted. Reviewing Table 5, Panel C (Appendix) for the third digit all the stock price distributions fit BL. The instances of individual numbers not fitting the predicted values occurs for number 0 (July 30), number 1 (October 29), number 2 (April 30), number 3 (September 30), and number 8 (July 30). Finally, Panel D (fourth digit) of Table 5 (Appendix) shows every month following BL. Nevertheless, individual numbers do significantly deviate including number 0 on February 26 and March 31, number 1 on June 30, number 5 on January 29 and May 28, and number 9 on April 30 and August 31. In general, there appears to be no systematic pattern of nonconformity of the numbers (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9) by month for each of the 4 digits.

The 4 panels of Figure 1 (Appendix) show in picture form the actual proportion distributions of each of the 4 digits by year (2020 and 2021) and that predicted by BL. While statistical significance cannot be determined by looking at the 4 figures it is readily apparent that the stock price distributions across the 24-month period closely follow the distribution of BL.

CONCLUSION

This study examines the conformity of stock prices in the NASDAQGS Health Care Industry to that predicted by Benford's Law. It was found for the period 2020 to 2021 using end-of-the-month stock prices compliance with the expected price by digit distribution. These results are not astonishing as they are in congruence with much of the literature. Nevertheless, these findings do not purport future harmony between actual and forecasted price distributions. Further, as NASDAQGS Health Care Industry stock prices increase there may be deviations from the hypothesized BL distribution. Moreover, why distributions so often follow the frequencies stated by Benford remains a mystery.

REFERENCES

Amiram, D., Bozanic, Z., & Rouen, E. (2015). Financial statement errors: evidence from the distributional properties of financial statement numbers. *Review of Accounting Studies*, 20(4), 1540-1593. doi:10.1007/s11142-015-9333-z

Archambault, J. J., & Archambault, M. E. (2011). Earnings management among firms during the pre-SEC-era: A Benford's law analysis. *Accounting Historians Journal*, 38(2), 145-170. doi:10.2308/0148-4184.38.2.145

Carslaw, C. A. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review*, 63(2), 321-327.

Cho, W. K., & Gaines, B. J. (2007). Breaking the (Benford) law: Statistical fraud detection in campaign finance. *The American Statistician*, 61(3), 218-223.

Das, S., & Zhang, H. (2003). Rounding-up in reported EPS, behavioral thresholds, and earnings management. *Journal of Accounting and Economics*, 35(1), 31-50. doi:10.1016/S0165-4101(02)00096-4

Henselmann, K., Scherr, E., & Ditter, D. (2012). Applying Benford's law to individual financial reports. An empirical investigation on the basis of SEC XBRL filings. *Working Papers in Accounting Valuation Auditing* Nr. 2012-1, 1-21.

Hickman, M. J., & Rice, S. K. (2010). Digital analysis of crime statistics: Does crime conform to Benford's law? *Journal of Quantitative Criminology*, 26(3), 333-350. doi:10.1007/s10940-010-9094-6

Horton, J., Kumar, D. K., & Wood, A. (2020). Detecting academic fraud using Benford law: The case of Professor James Hunton. *Research Policy*, 49(8), 1-19. doi:10.1016/j.respol.2020.104084

Hsieh, C. H., & Lin, F. (2013). Applying digital analysis to detect fraud: an empirical analysis of US marine industry. *Applied Economics*, 45(1), 135-140. doi:10.1080/00036846.2011.605759

Johnson, G. C. (2009). Using Benford's law to determine if selected company characteristics are red flags for earnings management. *Journal of Forensic Studies in Accounting & Business,* 1(2), 39-65.

Jordan, C. E., & Clark, S. J. (2011). Detecting cosmetic earnings management using Benford's law. *CPA Journal*, 81(2), 32-37.

Krakar, Z., & Žgela, M. (2009). Application of Benford's law in payment systems auditing. *Journal of Information and Organizational Sciences*, 33(1), 39-51.

Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing: A Journal of Practice & Theory*, 16(2), 52-67.

Skousen, C. J., Guan, L., & Wetzel, T. S. (2004). Anomalies and unusual patterns in reported earnings: Japanese managers round earnings. *Journal of International Financial Management and Accounting,* 15(3), 212-234. doi:10.1111/j.1467-646X.2004.00108.x

Thomas, J. K. (1989). Unusual patterns in reported earnings. *The Accounting Review*, 64(4), 773-787.

Tilden, C., & Janes, T. (2012). Empirical evidence of financial statement manipulation during economic recessions. *Journal of Finance and Accountancy*, 10, 1-15.

Van Caneghem, T. (2002). Earnings management induced by cognitive reference points, *The British Accounting Review*, 34, 167-178.

APPENDIX

Table 1. Expected frequency occurrences (in %) for each digit according to Benford's Law

	Position in number			
Digit	**1st**	**2nd**	**3rd**	**4th**
0	N/A*	11.968	10.178	10.018
1	30.103	11.389	10.138	10.014
2	17.609	10.882	10.097	10.010
3	12.494	10.433	10.057	10.006
4	9.691	10.031	10.018	10.002
5	7.918	9.668	9.979	9.998
6	6.695	9.337	9.940	9.994
7	5.799	9.035	9.902	9.990
8	5.115	8.757	9.864	9.986
9	4.576	8.500	9.827	9.982

* Not applicable.

Table 2. Descriptive analysis (for year 2020)

Year, 2020		Day Close Price, Last business day of Month of					
		Jan	Feb	Mar	Apr	May	Jun
Sample size,	n	238	238	238	238	238	238
Mean		46.452	44.954	41.291	46.995	51.791	52.97
Percentiles	25	8.58	8.537	6.688	8.8	10.05	10.46
	50	22.04	22.7	17.83	21.445	23.525	24.22
	75	51.51	49.74	42.795	50.953	53.803	57.135
Minimum		1.24	0.73	0.838	0.96	0.9	0.822
Maximum		719.19	616.31	650	631.83	641.99	637.01
Std. Deviation		72.956	70.229	72.239	77.228	84.446	85.388
Skewness*		4.606	4.136	4.563	4.001	3.876	3.841
Kurtosis**		32.297	23.763	28.218	21.143	19.461	19.038

Year, 2020		Day Close Price, Last business day of Month of					
		Jul	Aug	Sept	Oct	Nov	Dec
Sample size,	n	238	238	238	238	238	238
Mean		54.065	54.894	55.055	54.862	59.634	63.431
Percentiles	25	9.415	10.273	10.035	9.355	10.375	10.785
	50	22.49	24.98	24.255	24.52	27.185	27.31
	75	52.635	58.73	54.885	54.135	61.983	70.36
Minimum		0.885	1.11	1.09	1.04	1.34	1.15
Maximum		632.07	631.72	626	601.15	600	642.24
Std. Deviation		89.801	89.378	87.669	86.34	89.465	95.228
Skewness*		3.59	3.549	3.335	3.27	3.08	3.057
Kurtosis**		15.938	15.756	13.849	13.243	11.536	11.321

* Standard Error of Skewness is of 0.158

** Standard Error of Kurtosis is of 0.314

Table 3. Descriptive analysis (for year 2021)

Year, 2021		Day Close Price, Last business day of Month of					
		Jan	Feb	Mar	Apr	May	Jun
Sample size,	n	296	296	296	296	296	296
Mean		59.684	58.978	55.861	57.399	54.759	58.033
Percentiles	25	13.023	11.66	11.12	11.138	10.27	10.313
	50	30.28	31.08	28.97	27.175	24.515	25.84
	75	61.283	61.295	59.195	61.085	59.033	62.303
Minimum		1.53	1.52	1.34	1.21	1.21	1.19
Maximum		651.47	624.9	641.31	638.6	621	631.55
Std. Deviation		88.279	87.506	85.421	90.468	88.555	96.322
Skewness*		3.376	3.475	3.587	3.574	3.724	3.666
Kurtosis**		13.996	14.732	15.806	15.352	16.633	15.672

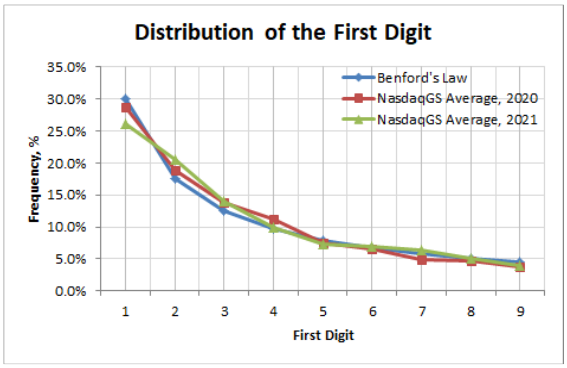
Year, 2021		Day Close Price, Last business day of Month of					
		Jul	Aug	Sept	Oct	Nov	Dec
Sample size,	n	296	296	296	296	296	296
Mean		57.826	59.9	56.957	57.179	53.338	54.036
Percentiles	25	8.458	9.065	8.423	8.083	7.253	6.698
	50	23.155	23.795	22.92	21.68	18.255	18.73
	75	57.673	59.288	55.9	53.113	46.793	46.983
Minimum		1	0.795	0.741	0.728	0.495	0.356
Maximum		695.8	709	697.5	728	709.99	704.9
Std. Deviation		103.474	107.893	103.096	106.793	101.689	104.262
Skewness*		3.787	3.84	3.82	3.906	3.915	3.928
Kurtosis**		16.422	16.932	16.738	17.392	17.621	17.7

* Standard Error of Skewness is of 0.142

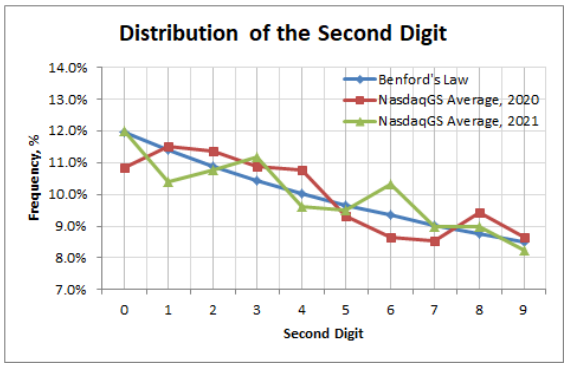
** Standard Error of Kurtosis is of 0.282

Figure 1. Distribution of the digits (Health Care Industry)

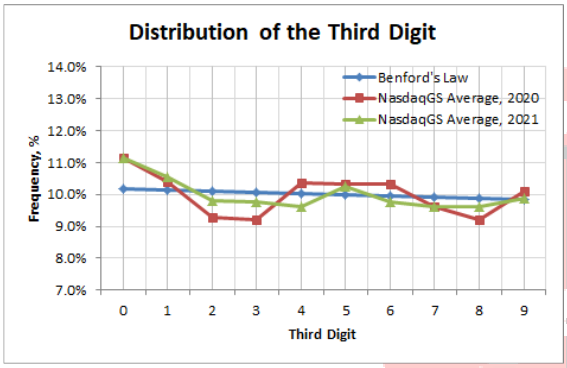
Panel A.



Panel B.



Panel C.



Panel D.

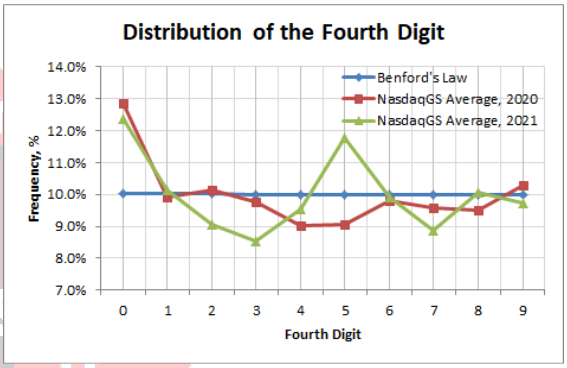


Table 4. NASDAQGS Health Care Industry (2020)

Panel A. First Digit

Digit value	Jan 31	Feb 28	Mar 31	Apr 30	May 29	Jun 30	Jul 31	Aug 31	Sept 30	Oct 30	Nov 30	Dec 31
0	Not Applicable											
1	-4.89[a]	-3.63	-2.37	-0.69	0.15	-2.79	-1.53	-0.27	-0.27	0.15	0.57	-0.27
	1.57[b]	1.15	0.73	0.16	0.05	0.87	0.44	0.02	0.02	0.05	0.12	0.02
2	2.14	0.46	0.04	0.46	0.88	3.82	3.40	1.72	1.30	0.46	-0.38	0.88
	0.78	0.10	0.02	0.10	0.27	1.46	1.29	0.61	0.44	0.10	0.07	0.27
3	0.95	1.37	2.63	3.47	3.05	-1.15	0.11	1.79	4.31	1.37	0.11	-1.15
	0.35	0.54	1.13	1.52	1.33	0.44	0.05	0.74	1.91	0.54	0.05	0.44
4	1.65	2.49	-0.45	-1.29	0.81	5.44	4.59	2.91	-0.45	0.39	0.39	2.07
	0.75	1.19	0.12	0.56	0.31	2.72*	2.29*	1.41	0.12	0.10	0.10	0.97
5	-2.46	-1.20	-0.36	1.33	-2.04	0.49	-2.04	-1.62	-1.20	1.75	1.75	0.49
	1.28	0.56	0.08	0.64	1.04	0.16	1.04	0.80	0.56	0.88	0.88	0.16
6	0.03	0.03	0.03	-1.65	2.55	-1.23	-0.81	1.71	-2.49	-0.81	1.29	0.45
	0.02	0.02	0.02	0.89	1.44	0.63	0.37	0.93	1.41	0.37	0.67	0.15
7	1.34	-0.34	-1.18	-1.18	-2.86	-0.34	-1.60	-2.86	0.50	-4.12	-0.76	1.34
	0.75	0.08	0.64	0.64	1.75	0.08	0.92	1.75	0.19	2.58*	0.36	0.75
8	0.77	2.45	-0.07	0.35	-2.59	-0.91	-1.75	-0.91	-1.33	0.35	-0.91	-1.33
	0.39	1.57	0.05	0.10	1.67	0.49	1.08	0.49	0.79	0.10	0.49	0.79
9	0.47	-1.63	1.73	-0.79	0.05	-3.32	-0.37	-2.47	-0.37	0.47	-2.05	-2.47
	0.19	1.05	1.12	0.43	0.03	2.29*	0.12	1.67	0.12	0.19	1.36	1.67
Sample size, n	238	238	238	238	238	238	238	238	238	238	238	238
χ^2	6.30	7.61	3.97	5.22	12.08	16.86	10.97	11.85	7.47	8.71	4.41	6.31
df	8	8	8	8	8	8	8	8	8	8	8	8
p-value	0.614	0.472	0.859	0.734	0.148	0.032*	0.204	0.158	0.487	0.368	0.818	0.612

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 4. NASDAQGS Health Care Industry (2020) (continuation)

Panel B. Second Digit

Digit value	Jan 31	Feb 28	Mar 31	Apr 30	May 29	Jun 30	Jul 31	Aug 31	Sept 30	Oct 30	Nov 30	Dec 31
0	-3.98[a]	0.22	-1.46	0.22	0.22	-2.72	-2.72	-1.46	1.48	-1.04	0.22	-2.30
	1.79[b]	0.00	0.60	0.00	0.00	1.19	1.19	0.60	0.60	0.40	0.00	1.00
1	0.38	2.90	-0.88	-0.04	-2.99	0.80	-0.04	-0.88	0.38	2.06	-1.73	1.64
	0.08	1.30	0.33	0.02	1.35	0.28	0.02	0.33	0.08	0.90	0.74	0.69
2	3.40	-2.48	-3.74	1.72	0.04	2.14	0.46	0.88	-1.22	3.40	-1.64	2.98
	1.58	1.12	1.75	0.75	0.02	0.96	0.12	0.33	0.50	1.58	0.71	1.37
3	2.59	-1.19	-1.61	0.49	0.91	3.85	-1.19	3.85	-1.19	-0.77	0.91	-1.19
	1.20	0.49	0.71	0.14	0.35	1.84	0.49	1.84	0.49	0.28	0.35	0.49
4	-2.89	1.73	8.04	-1.21	0.47	0.47	-0.79	-0.79	3.83	-0.79	0.05	0.89
	1.38	0.78	4.02*	0.51	0.14	0.14	0.30	0.30	1.86	0.30	0.03	0.35
5	0.42	-0.84	1.26	-0.42	-2.52	-3.37	0.42	0.00	0.42	-2.95	2.94	0.42
	0.11	0.33	0.55	0.11	1.21	1.65	0.11	0.00	0.11	1.43	1.42	0.11
6	1.17	0.75	-0.51	-0.51	-0.51	-0.51	0.33	-3.03	-1.35	-0.51	-1.77	-1.77
	0.51	0.28	0.16	0.16	0.16	0.16	0.06	1.50	0.61	0.16	0.83	0.83
7	-1.47	0.21	1.47	-2.31	0.63	-0.63	1.05	-2.31	0.21	-0.21	-0.21	-2.31
	0.68	0.11	0.68	1.13	0.23	0.23	0.45	1.13	0.11	0.00	0.00	1.13
8	2.17	-2.45	-1.61	1.75	2.17	0.49	-0.35	0.07	0.07	2.17	2.59	0.91
	1.07	1.22	0.77	0.84	1.07	0.15	0.08	0.04	0.04	1.07	1.30	0.38
9	-1.78	1.16	-0.94	0.32	1.58	-0.52	2.84	3.69	-2.62	-1.36	-1.36	0.74
	0.87	0.53	0.40	0.06	0.76	0.17	1.46	1.92	1.33	0.63	0.63	0.30
Sample size, n	238	238	238	238	238	238	238	238	238	238	238	238
χ^2	12.35	6.49	21.54	3.44	5.84	9.15	4.65	11.85	7.04	7.92	6.68	6.71
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.194	0.690	0.010*	0.944	0.756	0.423	0.864	0.222	0.633	0.542	0.670	0.668

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 4. NASDAQGS Health Care Industry (2020) (continuation)

Panel C. Third Digit

Digit value	Jan 31	Feb 28	Mar 31	Apr 30	May 29	Jun 30	Jul 31	Aug 31	Sept 30	Oct 30	Nov 30	Dec 31
0	2.43[a]	-1.74	2.01	7.05	-1.28	-0.09	-0.09	2.43	1.59	-1.78	-0.51	1.59
	1.13[b]	0.78	0.92	3.49*	0.54	0.05	0.05	1.13	0.70	0.80	0.16	0.70
1	-4.68	0.41	-1.31	-0.05	-0.39	-0.05	3.31	2.47	2.05	-0.47	-0.79	1.21
	2.28*	0.10	0.56	0.03	0.09	0.03	1.58	1.15	0.94	0.13	0.29	0.51
2	-0.43	-3.77	-2.11	-2.95	3.46	-0.85	-2.11	-0.85	-3.79	1.67	0.83	1.25
	0.11	1.82	0.97	1.41	1.66	0.33	0.97	0.33	1.84	0.75	0.32	0.53
3	-1.65	-0.35	-1.23	-3.33	-4.55	2.55	2.13	0.03	-0.81	2.13	-2.07	-2.91
	0.74	0.07	0.53	1.60	2.22*	1.20	0.98	0.01	0.31	0.98	0.96	1.39
4	0.07	1.37	0.49	1.33	0.58	-1.61	0.91	-0.35	1.33	1.33	0.49	-1.61
	0.03	0.60	0.14	0.57	0.19	0.72	0.36	0.07	0.57	0.57	0.14	0.72
5	4.31	-0.27	4.31	-0.74	-1.08	0.53	-1.16	0.95	-1.16	-2.42	0.53	0.53
	2.11*	0.03	2.11*	0.27	0.45	0.16	0.49	0.38	0.49	1.14	0.16	0.16
6	3.93	-1.50	-1.54	0.14	1.92	-0.28	0.56	-1.12	2.24	0.56	0.98	-1.54
	1.92	0.66	0.68	0.07	0.88	0.03	0.18	0.47	1.05	0.18	0.40	0.68
7	-2.76	2.76	-0.24	-3.18	1.54	-2.34	1.02	-0.66	1.44	-0.66	0.60	-1.08
	1.32	1.31	0.01	1.53	0.68	1.10	0.42	0.23	0.64	0.23	0.20	0.45
8	-2.30	-0.58	-1.04	0.22	-0.54	3.16	-4.40	-3.56	-1.88	1.90	1.48	-0.20
	1.08	0.19	0.43	0.01	0.17	1.53	2.17*	1.73	0.86	0.87	0.66	0.10
9	1.10	3.68	0.68	1.52	0.34	-1.00	-0.16	0.68	-1.00	-2.26	-3.10	2.78
	0.46	1.79	0.24	0.68	0.07	0.41	0.08	0.24	0.41	1.06	1.50	1.33
Sample size, n	238	237	238	238	236	238	238	238	238	238	238	238
χ^2	18.71	10.27	8.19	19.86	9.97	6.38	10.22	6.80	8.66	6.62	4.69	6.72
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.028*	0.329	0.515	0.019*	0.353	0.701	0.333	0.658	0.469	0.677	0.860	0.667

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 4. NASDAQGS Health Care Industry (2020) (continuation)

Panel D. Fourth Digit

Digit value	Jan 31	Feb 28	Mar 31	Apr 30	May 29	Jun 30	Jul 31	Aug 31	Sept 30	Oct 30	Nov 30	Dec 31
0	5.77[a]	4.69	2.89	5.77	2.83	2.00	-0.36	2.69	2.83	-0.25	4.43	0.97
	2.39[b]*	1.91	1.06	2.39*	1.14	0.78	0.03	1.08	1.14	0.11	1.85	0.31
1	-2.41	-4.13	3.53	-0.07	1.16	0.37	2.49	-1.73	1.16	0.33	-1.68	-0.12
	0.92	1.67	1.33	0.03	0.39	0.04	0.97	0.65	0.39	0.02	0.63	0.06
2	-1.24	-1.19	-0.98	4.03	0.05	2.56	-1.49	0.49	-1.07	0.34	0.55	-0.67
	0.41	0.39	0.27	1.63	0.02	1.03	0.53	0.09	0.35	0.02	0.12	0.18
3	-0.06	-1.18	1.61	1.69	-4.42	-0.72	-0.35	-0.06	-1.07	-1.96	-0.56	4.28
	0.03	0.39	0.53	0.61	1.85	0.20	0.03	0.03	0.35	0.74	0.13	1.80
4	-0.06	2.94	-6.78	-4.15	-2.74	1.47	-3.18	-0.61	-1.62	0.92	0.00	2.09
	0.03	1.15	2.68*	1.68	1.10	0.54	1.28	0.15	0.60	0.28	0.00	0.81
5	-0.06	-1.76	1.61	-2.98	-1.62	-0.16	-1.48	-1.71	1.18	1.50	-2.78	-2.86
	0.02	0.64	0.54	1.17	0.60	0.07	0.53	0.64	0.40	0.53	1.12	1.16
6	-2.39	-2.35	2.26	-0.64	-0.50	-0.70	1.37	1.06	2.86	-3.10	-2.77	2.64
	0.92	0.89	0.81	0.15	0.10	0.19	0.48	0.35	1.15	1.24	1.12	1.07
7	-0.05	-0.58	-2.89	-1.80	0.07	-2.34	1.94	-1.15	3.42	-0.22	4.45	-5.59
	0.02	0.12	1.07	0.66	0.03	0.93	0.73	0.39	1.40	0.10	1.87	2.39*
8	-1.80	5.31	0.34	-0.04	1.75	-0.70	-0.90	-1.15	-6.08	0.93	-2.21	-1.20
	0.66	2.18*	0.01	0.02	0.65	0.19	0.27	0.39	2.59*	0.28	0.86	0.41
9	2.30	-1.75	-1.60	-1.80	3.43	-1.79	1.95	2.17	-1.60	1.51	0.57	0.46
	0.88	0.63	0.53	0.66	1.40	0.68	0.74	0.85	0.59	0.54	0.13	0.08
Sample size, n	171	170	155	171	179	183	176	181	179	183	176	176
χ^2	9.38	15.41	13.80	14.592	9.67	4.22	5.492	4.02	13.43	3.48	11.42	13.14
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.403	0.080	0.130	0.103	0.378	0.896	0.790	0.910	0.144	0.942	0.248	0.157

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 5. NASDAQGS Health Care Industry (2021)

Panel A. First Digit

Digit value	Jan 29	Feb 26	Mar 31	Apr 30	May 28	Jun 30	Jul 30	Aug 31	Sept 30	Oct 29	Nov 30	Dec 31
0	Not Applicable											
1	-1.05[a]	-1.72	-4.77	-5.44	-5.44	-5.10	-7.81	-5.78	-5.78	-4.09	0.98	-2.06
	0.33[b]	0.58	1.72	1.98*	1.98*	1.85	2.86*	2.10*	2.10*	1.47	0.30	0.71
2	0.63	0.30	2.32	1.99	4.69	5.03	4.35	2.66	4.01	2.32	0.97	6.04
	0.21	0.06	0.97	0.82	2.04*	2.19*	1.89	1.13	1.74	0.97	0.36	2.65*
3	1.36	1.36	2.71	2.37	2.37	-0.67	2.37	2.03	1.02	1.02	2.37	0.34
	0.62	0.62	1.32	1.15	1.15	0.26	1.15	0.97	0.44	0.44	1.15	0.09
4	0.11	2.81	2.47	1.46	-1.25	1.46	-1.58	-1.25	-0.91	-0.23	-0.23	-1.25
	0.06	1.54	1.34	0.75	0.63	0.75	0.82	0.63	0.43	0.04	0.04	0.63
5	2.22	0.87	-1.16	-0.49	-1.16	-2.17	0.53	-0.49	0.19	0.87	-3.86	-2.85
	1.30	0.44	0.63	0.20	0.63	1.28	0.23	0.20	0.01	0.44	2.35*	1.71
6	0.06	-1.63	1.08	0.40	0.40	1.75	0.40	0.74	1.41	0.06	-0.28	-0.61
	0.04	1.00	0.62	0.16	0.16	1.09	0.16	0.39	0.86	0.04	0.07	0.31
7	-0.06	0.96	-0.73	1.30	0.62	-1.41	0.96	2.31	1.30	-0.39	1.30	-0.06
	0.04	0.58	0.41	0.83	0.33	0.91	0.58	1.58	0.83	0.17	0.83	0.04
8	-1.06	-0.72	-2.41	-2.07	-0.05	2.32	1.64	-0.05	-1.06	0.29	1.30	0.29
	0.70	0.43	1.75	1.49	0.04	1.68	1.15	0.04	0.70	0.09	0.89	0.09
9	-2.21	-2.21	0.49	0.49	-0.18	-1.20	-0.86	-0.18	-0.18	0.15	-2.55	0.15
	1.68	1.68	0.27	0.27	0.01	0.85	0.57	0.01	0.01	0.13	1.96	0.13
Sample size, n	296	296	296	296	296	296	296	296	296	296	296	296
χ^2	6.27	8.54	11.56	9.22	9.21	15.73	13.95	9.00	8.91	3.24	13.26	10.32
df	8	8	8	8	8	8	8	8	8	8	8	8
p-value	0.617	0.383	0.172	0.324	0.325	0.046*	0.083	0.342	0.350	0.918	0.103	0.243

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 5. NASDAQGS Health Care Industry (2021) (continuation)

Panel B. Second Digit

Digit value	Jan 29	Feb 26	Mar 31	Apr 30	May 28	Jun 30	Jul 30	Aug 31	Sept 30	Oct 29	Nov 30	Dec 31
0	-2.17[a]	1.88	2.90	2.90	-2.17	-1.16	0.87	-1.49	0.19	-2.85	3.23	-1.83
	1.06[b]	0.91	1.45	1.45	1.06	0.52	0.37	0.70	0.01	1.42	1.63	0.88
1	-3.62	1.11	-2.27	-0.92	-1.25	-2.27	1.11	-1.25	-1.93	1.79	-2.27	-0.24
	1.87	0.51	1.14	0.40	0.59	1.14	0.51	0.59	0.95	0.88	1.14	0.04
2	0.94	1.28	0.94	-2.10	-0.07	-0.41	-0.07	-0.75	-1.42	-1.76	0.94	1.28
	0.43	0.61	0.43	1.07	0.04	0.13	0.04	0.32	0.69	0.88	0.43	0.61
3	-0.97	-2.66	0.04	3.08	1.05	3.08	-2.32	-0.97	0.72	3.76	0.38	3.76
	0.45	1.40	0.02	1.64	0.50	1.64	1.21	0.45	0.31	2.02*	0.12	2.02*
4	2.13	0.10	2.81	-1.25	-0.57	0.44	1.46	-2.94	-1.25	0.10	-4.63	-1.25
	1.12	0.06	1.51	0.62	0.23	0.16	0.74	1.58	0.62	0.06	2.55*	0.62
5	2.16	-0.88	1.14	-1.56	-2.57	0.13	-1.22	2.49	0.13	-0.21	-0.88	-0.88
	1.16	0.42	0.57	0.81	1.40	0.08	0.61	1.35	0.08	0.02	0.42	0.42
6	0.80	-1.23	-0.89	-1.23	5.19	-0.22	-0.89	2.15	2.49	2.15	1.14	2.49
	0.37	0.63	0.43	0.63	2.97*	0.03	0.43	1.17	1.37	1.17	0.57	1.37
7	-0.59	-2.28	-1.94	0.09	-0.93	1.44	2.79	-1.26	-1.60	-0.25	2.79	1.10
	0.25	1.27	1.06	0.05	0.45	0.76	1.57	0.66	0.86	0.05	1.57	0.56
8	0.03	2.39	-2.34	1.72	0.70	0.70	-0.65	0.03	3.41	-0.65	-0.65	-2.00
	0.02	1.35	1.32	0.94	0.32	0.32	0.29	0.02	1.97*	0.29	0.29	1.11
9	1.30	0.28	-0.39	-0.73	0.62	-1.74	-1.07	4.00	-0.73	-2.08	-0.05	-2.42
	0.70	0.07	0.14	0.35	0.28	0.97	0.55	2.36*	0.35	1.18	0.03	1.39
Sample size, n	296	296	296	296	296	296	296	296	296	296	296	296
χ^2	8.75	8.04	9.77	9.05	13.14	6.39	6.47	13.39	9.04	10.83	13.86	11.74
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.461	0.530	0.370	0.433	0.157	0.701	0.693	0.146	0.433	0.287	0.127	0.228

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 5. NASDAQGS Health Care Industry (2021) (continuation)

Panel C. Third Digit

Digit value	Jan 29	Feb 26	Mar 31	Apr 30	May 28	Jun 30	Jul 30	Aug 31	Sept 30	Oct 29	Nov 30	Dec 31
0	0.97[a]	3.34	1.65	0.29	-1.73	0.29	5.02	-0.04	0.29	2.36	1.35	-2.07
	0.46[b]	1.80	0.84	0.07	0.89	0.07	2.76*	0.02	0.07	1.25	0.67	1.08
1	0.67	2.02	0.67	0.34	3.04	2.02	0.34	2.36	3.38	-5.05	-2.34	-2.37
	0.29	1.06	0.29	0.09	1.64	1.06	0.09	1.25	1.83	2.78*	1.24	1.25
2	-0.30	-3.00	1.73	-4.35	0.38	1.73	-1.65	3.08	-0.98	-0.94	0.75	0.04
	0.07	1.62	0.89	2.39*	0.12	0.89	0.85	1.66	0.46	0.44	0.33	0.02
3	-0.60	-0.60	1.77	-0.26	-0.26	-0.60	0.42	-1.61	-3.64	1.81	-0.57	0.75
	0.25	0.25	0.91	0.05	0.05	0.25	0.14	0.83	1.98*	0.94	0.23	0.33
4	-0.22	2.82	-2.25	-1.23	-0.56	-1.91	-0.90	-2.92	-0.22	2.86	-0.19	-0.22
	0.03	1.52	1.19	0.61	0.22	1.00	0.42	1.58	0.03	1.54	0.01	0.03
5	0.83	-2.55	-0.86	0.83	3.53	-2.88	-0.52	2.52	2.52	0.19	0.19	-0.52
	0.38	1.36	0.40	0.38	1.93	1.56	0.20	1.35	1.35	0.01	0.01	0.20
6	-1.16	-1.16	-1.83	2.90	-1.16	-0.48	0.20	-0.48	-0.48	1.25	-1.47	1.55
	0.57	0.57	0.96	1.57	0.57	0.18	0.01	0.18	0.18	0.62	0.74	0.79
7	1.58	1.92	-2.47	0.91	-1.46	-3.15	0.57	0.91	-2.47	-0.41	-2.11	2.94
	0.82	1.01	1.33	0.43	0.74	1.71	0.23	0.43	1.33	0.14	1.11	1.59
8	0.27	-1.42	1.96	-0.40	-0.40	1.62	-3.78	-2.43	0.95	-2.07	2.68	-0.07
	0.06	0.72	1.03	0.14	0.14	0.84	2.09*	1.31	0.45	1.09	1.45	0.04
9	-2.06	-1.38	-0.37	0.98	-1.38	3.35	0.31	-1.38	0.65	0.00	1.70	-0.03
	1.09	0.70	0.11	0.47	0.70	1.84	0.08	0.70	0.28	0.00	0.88	0.02
Sample size, n	296	296	296	296	296	296	296	296	296	295	295	296
χ^2	3.202	14.134	8.440	9.379	9.083	12.941	12.974	11.715	11.715	14.482	7.369	6.435
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.956	0.118	0.490	0.403	0.430	0.165	0.164	0.230	0.230	0.106	0.599	0.696

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.

Table 5. NASDAQGS Health Care Industry (2021) (continuation)

Panel D. Fourth Digit

Digit value	Jan 29	Feb 26	Mar 31	Apr 30	May 28	Jun 30	Jul 30	Aug 31	Sept 30	Oct 29	Nov 30	Dec 31
0	1.41[a]	4.80	4.57	2.26	3.38	-1.68	0.88	1.61	-0.93	2.54	4.20	5.29
	0.63[b]	2.38*	2.22*	1.03	1.57	0.74	0.31	0.67	0.33	1.10	1.88	2.35*
1	-1.03	0.27	-2.29	0.51	-1.09	4.90	-2.90	-0.25	1.95	0.61	0.28	0.70
	0.43	0.04	1.05	0.15	0.43	2.35*	1.29	0.01	0.82	0.18	0.02	0.21
2	-1.44	-0.54	-3.57	2.27	-1.53	1.39	-3.85	2.55	-2.35	-0.35	-0.70	-3.38
	0.64	0.18	1.71	1.03	0.65	0.59	1.75	1.13	1.02	0.05	0.21	1.46
3	-0.62	-2.19	-0.99	-1.23	-2.86	-3.43	0.89	-2.10	-3.79	1.59	-1.67	-1.33
	0.22	1.03	0.40	0.51	1.32	1.61	0.32	0.91	1.71	0.65	0.68	0.50
4	1.02	-3.01	0.73	0.09	1.61	1.40	0.42	-3.49	1.96	-1.79	-4.12	-0.31
	0.42	1.45	0.26	0.04	0.69	0.60	0.09	1.59	0.83	0.74	1.84	0.02
5	4.29	0.70	2.02	2.28	4.73	-0.79	2.80	-0.70	2.44	2.08	1.28	0.21
	2.13*	0.26	0.92	1.04	2.25*	0.29	1.24	0.23	1.06	0.88	0.49	0.10
6	-2.24	-0.53	-1.84	1.41	2.06	0.09	-0.04	-3.48	1.97	-1.30	4.22	-1.32
	1.06	0.17	0.83	0.60	0.92	0.05	0.02	1.59	0.83	0.51	1.89	0.50
7	-0.19	1.53	-0.12	-3.41	-1.06	-3.41	-0.99	0.71	-3.77	0.15	-1.17	-1.83
	0.10	0.69	0.06	1.61	0.42	1.61	0.36	0.23	1.70	0.07	0.44	0.73
8	0.22	-2.99	0.74	0.98	-2.40	0.98	0.91	0.71	2.93	-0.32	0.31	-1.31
	0.01	1.45	0.27	0.38	1.09	0.38	0.33	0.23	1.30	0.04	0.03	0.49
9	-1.41	1.95	0.75	-5.16	-2.84	0.54	1.87	4.44	-0.41	-3.22	-2.63	3.28
	0.63	0.91	0.27	2.49*	1.31	0.16	0.79	2.06*	0.08	1.43	1.14	1.41
Sample size, n	245	243	233	228	224	228	211	215	209	207	204	196
χ^2	7.846	12.892	11.411	13.341	15.068	12.761	8.042	12.698	12.790	6.044	13.416	11.626
df	9	9	9	9	9	9	9	9	9	9	9	9
p-value	0.550	0.168	0.249	0.148	0.089	0.174	0.530	0.177	0.172	0.735	0.145	0.235

a. Observed Deviation (in %), computed as a difference between Observed and Expected Benford's frequency.
b. Z-value computed as per formula (2).
* Significant at the .05 level.