

Association for Information Systems AIS Electronic Library (AISeL)

ECIS 2012 Proceedings

European Conference on Information Systems
(ECIS)

5-15-2012

DATA QUALITY IN FINANCIAL PLANNING - AN EMPIRICAL ASSESSMENT BASED ON BENFORD'S LAW

Jochen Martin

Karlsruhe Institute of Technology

Tobias Conte

Karlsruhe Institute of Technology

Follow this and additional works at: <http://aisel.aisnet.org/ecis2012>

Recommended Citation

Martin, Jochen and Conte, Tobias, "DATA QUALITY IN FINANCIAL PLANNING - AN EMPIRICAL ASSESSMENT BASED ON BENFORD'S LAW" (2012). *ECIS 2012 Proceedings*. 159.

<http://aisel.aisnet.org/ecis2012/159>

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2012 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

DATA QUALITY IN FINANCIAL PLANNING – AN EMPIRICAL ASSESSMENT BASED ON BENFORD’S LAW

Martin, Jochen, Karlsruhe Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany, jochen.martin@kit.edu

Conte, Tobias, Karlsruhe Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany, tobias.conte@kit.edu

Abstract

Planning Processes play an important role in almost any business scenario. In particular, induced by the financial crisis, financial planning as a foundation for liquidity management is paid extraordinary attention to. Its quality and reliability is usually ensured by the use of information systems. Besides process efficiency, a key factor in liquidity management is the quality of the delivered planning data. More recently, business intelligence measures to increase data quality, for instance, realized through decision support services, find their way into the planning process. In this paper, we lay the foundation to include digital analyses of reported financial planning numbers into automated decision support services. In this vein, our contribution is twofold: First, based on a large and representative data set from a renowned, multinational enterprise, we empirically prove that financial planning numbers exhibit a certain, characteristic digit distribution, namely, Benford’s Law. Second, we investigate whether decision support services that incorporate intelligence based on Benford’s Law are appropriate to increase perceived financial planning data quality. This question is tackled via analyses that relate detailed properties of the delivered data to Benford’s Law as a prerequisite for the integration of automated decision support services into business intelligence systems.

Keywords: Decision Support Services, Financial Planning, Benford’s Law, Digital Analysis, Data Quality.

1 Introduction

Planning processes play a weighty role in almost any kind of company, not least because it is a highly knowledge-intensive task. Our research is dedicated to the category of planning processes which has gained notoriety in the recent years: The almost ubiquitous financial crisis and, as one of its consequences, the decreasing confidence in the creditworthiness of most enterprises, banks, as well as industrial corporations, has made it painfully obvious that *financial planning* requires particular attention. Of course, it has been recognized long before the crisis that a precise forecast of business figures like sales, production, and investments is essential to accomplish a solid liquidity and exposure planning (Kim et al., 1998; Graham and Harvey, 2001), and thus, enable companies to cope with uncertainties as exemplified above. For globally spread companies, it is even more challenging to compose such a well-founded planning: In the case of central currency-specific liquidity planning, decentralized planning processes have to be coordinated within the local partitions and internal transactions between them have to be monitored to ensure a proper and consistent overall financial planning. Usually, crucial planning tasks, as for instance liquidity planning, are conducted in or at least supported by information systems. In globally-spread companies, corporate financial portals have turned out to be an efficient measure to enhance the process of centralized liquidity risk management (Vo et al., 2007). Today, such information systems are oftentimes included into service-oriented architectures. In this vein, IT-based business intelligence services can be offered to support planning activities, such as market-based prediction services, services to detect complex events as a sequence of defined activities, or *decision support services*. The latter, especially if large amounts of underlying data shall be processed for the planning task, are usually based on data mining methods. However, detecting patterns in huge data sets can be tedious, for instance since it requires a variety of upstream and downstream efforts and may be different for each data set considered (Witten and Frank, 2005).

What if we found properties valid for a variety of data sets, independent of the respective industry, task, or company? Benford's Law, as shown by Benford (1938), provides highly interesting insights into the structure of empirical data: Naively thought it seems obvious that the digits of numbers in the decimal system are equally distributed. Yet, Benford (1938) showed for several kinds of empirically gathered numbers that the leading digits as well as the digits in second position occur with distinct probabilities which clearly differ from an equal distribution.

Going back to the application domain of financial planning data, this property can be exploited for a new kind of automated decision support service: A service that allows for continuous auditing which can be integrated into information systems. If we are able to show that the digit distribution in financial planning data in fact follows Benford's Law, we can combine this result with additional statistical evaluations and expert knowledge to eventually support financial planning managers in their decision which data samples to further investigate. For example, certain rounding behaviors or the creation of duplicate numbers via copy-and-pasting should distort the digit distribution. Hence, the contribution of this work is twofold: We (*i*) introduce Benford's Law to a field of application that has not studied so far by answering the following research question: *Does financial planning data follow Benford's Law?* Based upon this insight, we (*ii*) evaluate relevant patterns in the planning behavior of different planning entities and, based upon that, investigate whether Benford's Law can be applied as a foundation for automated decision support services. This investigation along with expert knowledge about the perceived data quality gained from our industrial partner allows us to address the second research question in this paper: *Are decision support services that incorporate Benford's Law appropriate to increase perceived planning data quality?* The quality assessment beyond data accuracy, i.e. the difference between plan and actual values, is crucial to overcome the lack of missing actual data at the moment of planning data generation and strengthens the important data consumer perspective on planning data quality (Wang and Strong, 1996).

This paper is structured as follows: In Chapter 2, Benford's Law is introduced along with related fields of applications and domains. Based on the finding that financial planning data is generally suitable to comply with Benford's Law, Chapter 3 includes the data sample, the applied methodology (inferential

analyses) and the hypotheses to be investigated. Chapter 4 includes the evaluation results and shows the practical relevance of our insights. This paper closes with a conclusion and our future work.

2 Scope and Related Work

This chapter summarizes the general idea of Benford's Law (Section 2.1) as the foundation of its application to different domains and problem statements, particularly its suitability for continuous auditing tasks (Section 2.2).

2.1 Benford's Law and its General Applicability to Financial Planning Data

The digital phenomenon today known as *Benford's Law* or the *significant digit law* was initially discovered and described by Newcomb (1881). Benford (1938) found the first empirical evidence for it. The heterogeneous data underlying his studies ranged from numbers on newspaper covers to physical constants. Contrary to intuition, Benford spotted that the digits of these numbers are not uniformly distributed, but rather follow a logarithmic distribution. The probability P_i for each digit j in position i of a number can be calculated as follows (for $i=\{1,2\}$):

$$P_1(j) = \log\left(1 + \frac{1}{j}\right), \quad j \in \{1, 2, \dots, 9\} \quad (1)$$

$$P_2(j) = \sum_{k=1}^9 \log\left(1 + \frac{1}{kj}\right), \quad j \in \{0, 1, \dots, 9\} \quad (2)$$

Besides the striking distribution for $i=\{1,2\}$, Table 1 also depicts the probabilities of the third and fourth position to illustrate the approximation towards a uniform distribution for higher digit positions (Nigrini, 1997). For an intuitive explanation of the above-described phenomenon, please refer to Drake and Nigrini (2000) and Durtschi et al. (2004). As shown by, for instance, Carslaw (1988) and Nigrini (1997), Benford-style analyses can also be made for the combination of two digits. Yet, due to space restrictions, we limit our results to the single digit assessment which is left to be dealt with in future work (cp. Chapter 5).

Digit j	$P_1(j)$	$P_2(j)$	$P_3(j)$	$P_4(j)$
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 %

Table 1. Digit distribution in first to fourth position in "naturally occurring" numbers according to Benford's Law (Benford, 1938; Nigrini, 1997).

Hill (1995) proved that Benford's Law follows a systematical statistical behavior: Since data distributions in nature are usually *random samples* taken from *random distributions* and joined afterwards, they converge to the logarithmic distribution as shown in Equations 1 and 2. Based on that, Nigrini (2000) derived three criteria to decide whether a data sample is likely to comply with Benford's Law:

- (1) The numbers should describe the relative sizes of similar phenomena,
- (2) The numbers should have no fixed upper and lower boundaries, and
- (3) The numbers should not be systematically created and assigned, as, for instance, ID-numbers.

For high quality financial planning data, these criteria are generally fulfilled. For financial planning data to be of great value, its generation must at any point in time include all information available at that time. This information is highly heterogeneous and occurs randomly. Based thereupon, companies can calculate expected invoices and cash flows. Hence, the resulting financial planning numbers are random themselves as they are based on different data sources with different random distributions (Hill, 1995). Summing up, all numbers included in a set of high quality financial planning data are (1) cash-related (same phenomenon), have (2) no pre-fixed boundaries and (3) are not created systematically. Furthermore, Pinkham (1961) showed the scale invariance of Benford's Law, that is, heterogeneous currencies as present in financial planning tasks, should not influence the data's conformity to the expected distribution. Importantly, above-mentioned criteria are necessary, but not sufficient. Therefore, statistical analyses of relevant and representative data are still indispensable to ultimately test whether financial planning numbers satisfy Benford's Law or not.

2.2 Benford's Law in Related Domains and Fields of Application

Benford (1938) initiated an entirely new field of analyses which can be roughly categorized into two groups. Besides papers that provide additional mathematical insights and theoretical evidence for Benford's Law, (cp. Section 2.1), researchers have dealt with its application to different kinds of data sets. Moreover, academia has brought forth a respectable body of empirical work that presents applications of Benford's Law, mostly to detect anomalies and fraud in data. Nigrini and Mittermaier (1997) define such a *digital analysis* as "the analysis of digit and number patterns with the objective of detecting abnormal recurrences of digits, digit combinations and specific numbers".

The conformity of data to Benford's Law has been shown for a couple of domains: For instance, Diekmann (2007) examines the digit distribution in statistical regression coefficients published in scientific literature. However, most work has been done in the field of accounting data. For example, Carslaw (1988) and Thomas (1989) proved that reported earnings satisfy Benford's Law; Nigrini (1996) detected conformity to Benford's Law in tax payments. Due to the development of digital analyses and mathematical investigations, scholar's conclusions on the practical relevance of Benford's Law are quite different: While earlier work, with Raimi (1976) being named as a representative, calls the observations "a curious mathematical phenomenon", more recent publications are aware of the value of Benford's results: Hill (1998) argues that digital analyses have a big impact on daily accounting business. Benford's Law's worth for detecting errors and systematic procedures started with its application to accounting purposes: Based on deviations from the expected Benford distribution, Carslaw (1988) found evidence for systematically rounded up numbers in reported earnings of companies in New Zealand. Thomas (1989) backed up this finding based on data from American companies. Interestingly, in addition, he found a systematic rounding down behavior for reported losses. Checking available tax data against Benford's Law, Nigrini (1996) detected systematic mistakes in tax payments. Based thereupon, Nigrini and Mittermaier (1997) developed and evaluated a standardized set of procedures to analyze huge data sets based upon different kinds of digit distributions. Krakar and Zgela (2009) apply digital analyses to foreign payments in banking transfers. During the last decade, the continuous development of digital analyses and data mining techniques in general has been accompanied by a continuous growth of available data in all business areas. As stated by Rezaee et al. (2002), this development requires new auditing structures and systems providing continuous auditing procedures. Suggesting continuous auditing based on digital analyses, Nigrini (2000) points to the same direction.

We pick up these arguments by laying the foundation for efficient auditing systems to continuously validate financial planning, that is, forecast data. On the one hand, this foundation is reflected in the very assessment of financial planning data: As mentioned above, to date, it has not been proven whether numbers from this domain satisfy Benford's Law or not, though the necessary criteria are met. We prove, based upon a large set of representative empirical data (cp. Section 3.1) that the digit distribution in financial planning data is conform to Benford's Law. In addition, we conduct more detailed analyses of clustered data in order to evidence the suitability of digital analyses for quality

improvement measures integrated into automated decision support services as a part of business intelligence systems.

3 Methodology and Hypotheses

Chapter 3 starts with the basic data sample and the preparation of all sub-samples (Section 3.1). Based upon that, Section 3.2 describes the conducted inferential analysis combined with the precision measure we introduce – the average fulfillment rate. Based on these preparatory explanations and the contents presented in Chapter 2, Section 3.4 formulates the hypotheses to be evaluated in this work.

3.1 Data Sample and Preparation

The data set to be evaluated is the cashflow-oriented financial planning data we have access to at our *industry partner*, a globally acting large enterprise in the pharmaceutical and chemical industry. It can be seen as an archetypical *multinational enterprise* with subsidiaries spread all over the world which implies a decentralized data generation. In more detail, we have available 25 data pools from May 2005 to June 2011 each of which reflects a data delivery. Each data pool contains the planned values of all entities that have handed in planning data in this delivery. The planning data per entity reports expected issued and received invoices, cashflows, tax payments and so on for 15 months into the future. Initially, the financial planning data was delivered quarterly in the months of February, May, August and November. Due to internal re-structuring reasons, the delivery months have changed to March, June, September and November since 2008. To increase the robustness of our results, we conduct the evaluations for 9 data samples: these are the complete data sample and 8 sub-samples that provide different perspectives on the available data. More detailed, we first divide the basic set ([1] *complete*) into data delivered by [2] *large* and [3] *small entities*. Thereby, the size of the entity is determined by the number of planned values. This distinction results from the expectation that large entities with a high number of planned values due to economies of scale (cp. Williamson, 1991) can put more effort in their planning data generation and, hence, achieve a higher quality. Since planning data has not been examined in literature before, this distinction is based on expert knowledge gained from interviews within the enterprise. The second perspective distinguishes data with [4] *positive* and [5] *negative* prefix. This separation is based on observations made, for instance, by Carslaw (1988) and Thomas (1989), who observed different digit distributions for reported positive and negative numbers. The remaining four sub-samples are combinations from the distinctions listed above, that is [6] *positive large*, [7] *positive small*, [8] *negative large*, and [9] *negative small*.

Generally, the number of delivered planned values increased from 27,511 in May 2005 to 72,141 in June 2011. Additionally, the creation of the sub-samples as described above brings about highly different sample sizes. To tackle this issue and to create a solid and comparable basis for the inferential analyses (cp. Section 3.2), we conduct a *normalization* as follows: We set the minimum sample size N to the number of items included in the minimum of all delivery-specific sample sizes. Afterwards, we reduce all other sample sizes to N . To do so, we uniformly draw N items from the respective sample. If the reduced set is not significantly smaller than the original one, this reduction does not change the digit distribution, yet to avoid biases in case of a stronger reduction, the reduction is carried out multiple times and the average of the resulting distributions is calculated. To obtain a reasonable trade-off between evaluation performance and accuracy of the result, we performed a simulation study and find out the required number of reductions that have to be carried out considering the distribution of the original sample size and the distribution of the reduced sample size. For instance, we found out that in case of reduction higher than 90% (e.g., the *complete* data sample size in June 2011 was 72,141 which had to be reduced to the minimum number 5,225), we have to take the average of 50 reductions. To summarize, in order to secure the comparability of our results, the normalization of the sample size facilitates the creation of a solid data pool and, thus, robust results. The above-described data preparation (sub-samples and normalization) also includes the deletion of all planned values with an absolute nominal less than 10 (depends on the currency and includes zero values) to avoid procedural problems with Benford's Law according to Nigrini (1997).

3.2 Interferential Analyses and Average Fulfillment Rate

To date, related literature (cp. Section 2.2) consults two kinds of interferential analyses to investigate if data is conform to Benford's Law: interferential analyses for (i) single digits (e.g. *z-statistic*), and (ii) the complete distribution (e.g. *chi-square statistic*, *mean absolute deviation*). Yet, all of these approaches aim to detect significant deviations from the expected distribution yielded by Benford's Law. None of the approaches offers the possibility to make statements on the degree a distribution satisfies Benford's Law. The mean absolute deviation (MAD) only allows for a descriptive indication; the chi-squared test is suitable for a pure true-false view of the distribution without any differentiation in between. Moreover, the results of these analyses strongly depend on the size of the data sample. For instance, the width of a confidence interval is solely calculated based on the p-value and the number of observations.

As an alternative, Nigrini (2000) states that a desirable approach to investigate conformity to Benford's Law should fulfill the following requirements:

- (1) The test shall measure the conformity to the expected distribution, not only with single digits or digit combinations,
- (2) the result shall be independent of the sample size,
- (3) the test shall be implementable and understandable for users in practice, and
- (4) the conclusion of the test shall be objectively determinable.

Existing approaches fail to fulfill the independence requirement (2). To fulfill (2), the normalization approach to eliminate the dependence on the basic sample size was presented in Section 3.1.

Independent of the application field, the main challenge for all inferential analyses is the interpretation of deviations from the expected Benford distribution. Nigrini (1997) and Durtschi et al. (2004) apply single digit analyses based upon z-statistics and distribution analyses (chi-squared test). However, other scholars, e.g. Busta and Weinberg (1998), use neural networks to evaluate deviations from the expected distribution. The latter approach performs better than the above-mentioned digit analyses, however, to the disadvantage of type I errors which are an indication for fraud in case of correct data (so-called "over auditing"). Since we want to keep the rate of type I errors low and to incorporate a more differentiated view of the fulfillment degree (cp. requirement 1), we introduce the average fulfillment rate (*AFR*) as a heuristic to measure the degree of confirmation between two distributions.

The *AFR* of a data sample reflects the percentage of digits not deviating significantly from the expectation with respect to the digit position i . Accordingly, the calculation of the average fulfillment rate per distribution is based on the z-statistic per digit. For each digit, we calculate the 95% confidence interval in dependence of the basic p-value .05. Based upon this interval, we can decide if the digit probability significantly deviates from the expectation. In case of a deviation, we assign the digit with 0 and in case of conformity to 1. The *AFR* is then the mean of all decisions, for instance, in case of one digit deviating significantly in i , the AFR_i would be 88.9%. We can calculate AFR_i of the digit position i in the following way:

$$AFR_1(P_1(d)) = \frac{1}{9} \sum_{j=1}^9 1_{[l_j, u_j]}(P_1(j)), \quad (3)$$

$$AFR_2(P_2(d)) = \frac{1}{10} \sum_{j=0}^9 1_{[l_j, u_j]}(P_2(j)), \quad (4)$$

where $P_i(d) \in [0,1]^n$ denotes the vector of digit probabilities per position $P_i(j)$ with dimension n , $n = 9$ if $i = 1$ and $n = 10$ if $i = 2$. $1_{[l_j, u_j]}$ is the indicator function which is 1 in the denoted interval $[l_j, u_j]$ and 0 else. For the confidence interval, the lower bound l_j and upper bound u_j are calculated separately for each digit j . To demonstrate the indication of the *AFR*, we compared the chi-square value and the *AFR* for the *complete* data sample. This examination reveals a highly significant dependence with $\tau = -.76$, $p < .001$. Indeed, particularly a chi-square value less than 17.53, i.e., there is no significant deviation between the distributions based upon $p = .05$, is significantly correlated ($\tau = .52$, $p < .01$) to an AFR_1 of 77.8% (i.e. 7 of 9 digits do not deviate significantly). Hence, the *AFR* yields the same significance as the chi-square test, yet, firstly, it is much easier to understand and to interpret than the chi-square value. Secondly, it offers a differentiated indication of the analyzed data sample's conformity to the

expected distribution. According to Nigrini (2000), both properties are weighty advantages for the implementation into a decision support service (cp. (3) and (4) in the requirements list).

For the statistical evaluation of the results generated based upon the AFR, we conduct two kinds of non-parametric inferential analyses: (i) Kendall's correlation, and (ii) the Wilcoxon signed-rank test. We decided to utilize non-parametric approaches since we found, as a result of a Shapiro-Wilk distribution test, a significant deviation from the normal distribution for approximately 72% of the tested treatments (for a more detailed explanation of the treatments, please refer to Section 3.3). We opt for the Kendall correlation coefficient τ since we have a relatively small sample size ($N=25$) along with many tied ranks (Field, 2009). The Wilcoxon signed-rank test is chosen as it is the most common non-parametric test. According to Field (2009), we always add the 1-tailed level of significance p and the test statistic T (denoting the smaller value of the two rank sums) to the reported absolute value. In order to back up the robustness of our analyses, we also calculated Pearson correlations and t-tests, yet the parametric results of both analyses were identical to the non-parametric tests as to the level of significance.

3.3 Hypotheses

As we have argued in Section 2.1, financial planning data meets the necessary criteria to follow Benford's Law due to its structure. Additionally, a major foundation of our evaluation is expert knowledge gained from our industry partner about the quality assurance measures carried out during the 6 year time period which spans our data sample. We can demonstrate that, due to process optimization and compliance enhancement, the output data of the financial planning process was continuously increased. We can also state that none of the improvements were directly and explicitly related to pushing the data towards the Benford distribution. Based upon this determining factor, the answer to our first research question (*Does financial planning data follow Benford's Law?*) is in the verification of the conformity between the digit distribution in the underlying financial planning data sample and the logarithmic distribution present in Benford's Law. To do so, we set up the following hypothesis:

H1: *The average fulfillment rate increases over the considered time period.*

To extract robust evidence for H1, formally, we investigate *18 sub-hypotheses*: For each of the investigated 9 data samples [1]-[9] (cp. Section 3.1), the conformity of the first *and* second digit of the underlying numbers to Benford's Law is tested. This consideration leads to 18 *treatments* to be tested. In more detail, these are: [1.1] complete first digit, [1.2] complete second digit, [2.1] large first digit, [2.2] large second digit, and so forth, ending with [9.1] negative small first digit and [9.2] negative small second digit.

In order to find evidence for the second research question (*Are decision support services that incorporate Benford's Law appropriate to increase perceived planning data quality?*), we have to delve deeper into details of the data structure. In this work, we concentrate on (i) differences between reported negative numbers and positive numbers (cp. also Carslaw, 1988; Thomas, 1989) and (ii) large and small entities (as large companies, due to their size and their planning volume, may be able to put a greater effort as well as more expertise into the financial planning). Thus, we investigate the following hypotheses H2 and H3:

H2: *The average fulfillment rate in the underlying financial planning data delivered by large entities is higher than in the financial planning data delivered by small entities.*

H3: *The average fulfillment rate in the underlying financial planning data with a positive prefix is higher than in financial planning data with negative prefix.*

To validate these hypotheses, we again conducted analyses for the first and second digits in multiple data sub-samples: For H2, these data samples are [1.1] *complete first digit*, [1.2] *complete second digit*, [4.1] *positive first digit*, [4.2] *positive second digit* and [5.1] *negative first digit*, [5.2] *negative second digit*. For H3, we consult [1.1] *complete first digit*, [1.2] *complete second digit*, [2.1] *small first*

digit, [2.2] *small second digit* and [3.1] *large entities first digit*, [3.2] *large entities first digit*. H2 and H3 potentially provide us with insights that we can include into a decision support service for financial planning managers.

4 Evaluation

This chapter presents a detailed examination of the hypotheses set up in Section 3.3. To address H1, we investigate the dependency between progressing time and the data's conformity to Benford's Law for the first and second digits in 9 data samples (cp. Section 4.1). In Section 4.2, we generate detailed knowledge about the data characteristics in order to address H2 and H3. Finally, Section 4.3 merges the insights gained through the statistical analyses and translates them into the practical value of our results.

4.1 Trend Analyses and Robustness

In order to address H1, we perform the trend analyses for all 9 data samples described in Section 3.1, each of them for the first and second digit. Altogether, 18 treatments as noted in Section 3.3, can thus be tested. The results of this evaluation for AFR_1 and AFR_2 are listed in Table 2. We calculated the mean and the standard deviation over 25 deliveries along with the correlation τ and the significance level p . Although we chose the rather conservative Kendall's τ (the average absolute value is only around 2/3 of a Spearman correlation according to Field 2009), we found a medium or strong correlation for 15 of 18 treatments. The only non-significant trends were discovered for *positive* data, yet even for [4.1] ($\tau = .19$), [6.1] ($\tau = .09$) and [7.2] ($\tau = .19$) we found at least a small positive correlation.

Data sample	AFR_1				AFR_2			
	Mean	SD	τ	p	Mean	SD	τ	p
[1] Complete	87.2%	12.2%	.59	.000***	29.6%	15.7%	.51	.000***
[2] Large	81.4%	18.4%	.44	.002**	57.6%	15.4%	.36	.010**
[3] Small	70.7%	23.5%	.65	.000***	10.4%	10.2%	.27	.047*
[4] Positive	79.2%	19.1%	.19	.102	48.4%	19.3%	.44	.002**
[5] Negative	79.3%	14.0%	.54	.000***	20.0%	9.6%	.50	.001***
[6] Positive Large	70.0%	15.5%	.09	.289	67.6%	13.3%	.32	.020*
[7] Positive Small	60.6%	19.7%	.53	.000***	26.8%	11.8%	.19	.114
[8] Negative Large	64.9%	24.5%	.46	.001***	47.2%	15.7%	.28	.033*
[9] Negative Small	60.6%	18.3%	.45	.001***	10.8%	10.0%	.42	.004**

Table 2. Results (Mean/standard deviation SD/correlation τ /p-value) of the AFR trend analyses for first and second digit position (AFR_1 and AFR_2) in nine data samples (* $p < .05$, ** $p < .01$, *** $p < .001$; $N=25$).

For AFR_1 , the effect size ranges from $\tau = .44$, $p < .01$ in [2.1] and $\tau = .45$, $p < .001$ in [9.1] to $\tau = .59$, $p < .001$ in [1.1] and even $\tau = .65$, $p < .001$ in [3.1]. For AFR_2 , in analogy with the general AFR level, the trends are alleviated compared to AFR_1 . Nevertheless, they are all at least medium strong ($\tau = .27$, $p < .05$ in [3.2] or $\tau = .28$, $p < .05$ [2.2]) and even range to strong dependencies in [5.2] ($\tau = .50$, $p < .001$) and in [1.2] ($\tau = .51$, $p < .001$).

Due to page restriction, we are only able to show graphs for two exemplary data samples: Figure 1 shows the AFR_1 (for all 25 deliveries); for [1.1], i.e. *complete* on the left hand side, and for [3.1], i.e. *small* on the right hand side. As easily can be seen, for [1.1], the AFR_1 clearly increases over time from a minimum value of 66.7% to 100% in the last four deliveries. A highly significant correlation between time and AFR_1 ($\tau = .59$, $p < .001$) is present. In [3.1], the trend itself is even stronger with $\tau = .65$, $p < .001$, yet, the mean AFR_1 over all 25 deliveries is higher in [1.1] (87.2%) than in [3.1] (70.7%). These differences provide first indications for the evaluation of H2 in Section 4.2. The difference described above can also be observed in the mean AFR_2 (29.6% in [1.2] and 10.4% in [3.2])

data). Yet, as clearly demonstrated in Figure 2, the degree of conformity to Benford's Law in financial planning data is on a lower level in the second digit than in the first digit. Still, we observe a strongly positive trend in [1.2] ($\tau = .51, p < .001$) and at least a medium positive trend [3.1] ($\tau = .27, p < .05$).

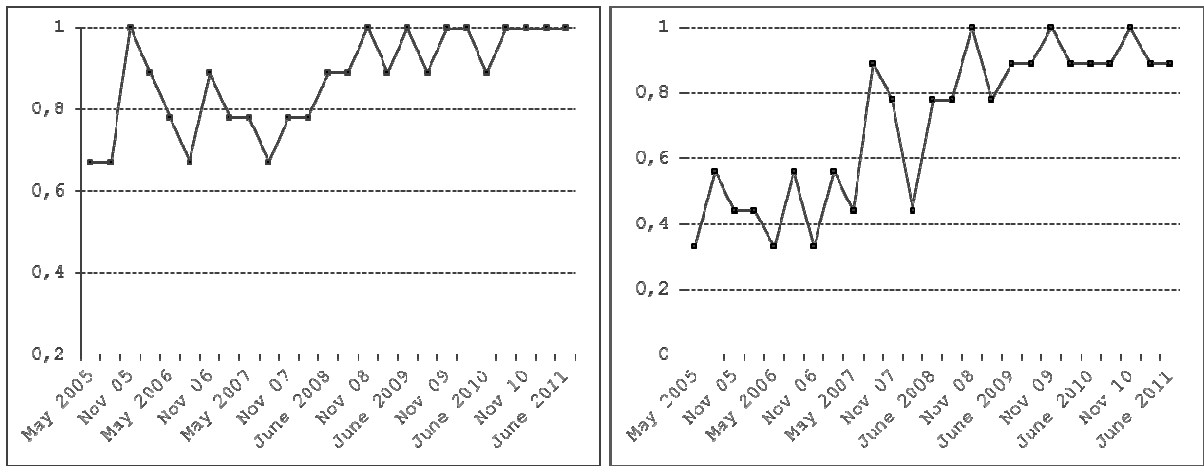


Figure 1. Development of AFR_1 in the complete data sample (on the left) and in the small data sample (on the right) over all deliveries (from May 2005 to June 2011).

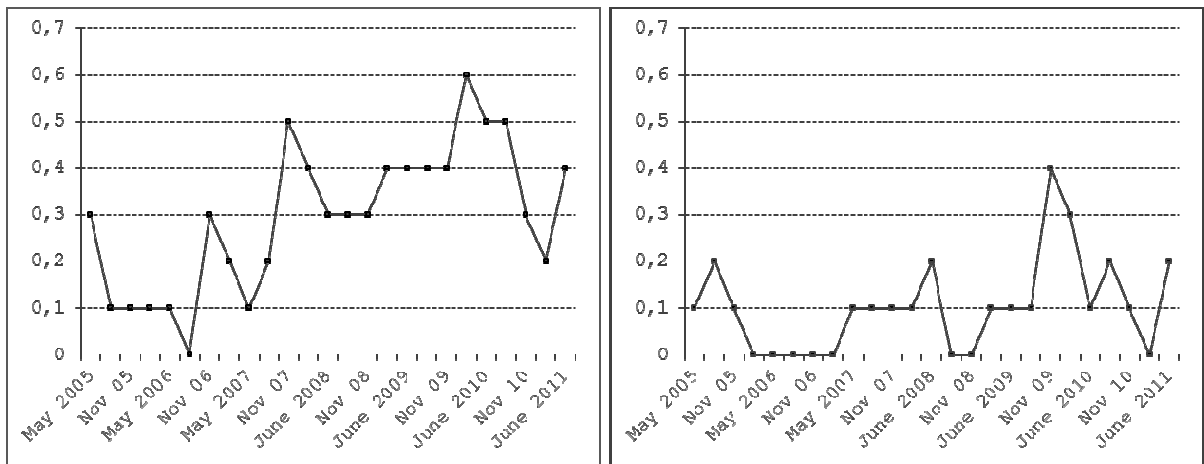


Figure 2. Development of AFR_2 in the complete data sample (on the left) and in the small data sample (on the right) over all deliveries (from May 2005 to June 2011).

Altogether, we are able to prove a significant and robust positive trend in AFR_1 and AFR_2 in 15 of 18 data samples. For the data samples [4.1], [6.1] and [7.2] we found at least a small correlation (cp. Field 2009). Based upon these findings, we can confirm H1. Furthermore, the results clearly point to a difference in the quality indication of AFR_1 and AFR_2 : An increased data quality is likely to first lead to an increased AFR_1 and, in a second step, leads to an increased AFR_2 . That is why we (i) observe a generally smaller AFR_2 throughout all samples, and (ii) AFR_2 is higher in data samples with a higher AFR_1 . Yet, generally, the robust positive trend is present both for AFR_1 and AFR_2 , which verifies H1. Although the average AFR_2 in our data set is not (yet) conform to the expected distribution, there is a strong indication for the conformity of high quality financial planning data to Benford's Law. Thus, RQ1 can be confirmed.

4.2 Group Analyses

To validate H2 and H3, we again conduct analyses for multiple data samples to assure either the robustness of the results if they are unique or to detect characteristics of the data clusters if the results

are controversial. In case of H2 the investigated data samples are *complete*, *positive* and *negative*, both for the first and the second digit. Furthermore, we investigate the *complete*, *small* and *large* data sample to validate H3, again for both digit positions.

Data sample	AFR ₁			AFR ₂		
	[Sample] Mean		p	[Sample] Mean		p
	Large	Small		Large	Small	
Complete	[2.1] 81.4%	[3.1] 70.7%	.024*	[2.2] 57.6%	[3.2] 10.4%	.000***
Positive	[6.1] 70.0%	[7.1] 60.6%	.023*	[6.2] 67.6%	[7.2] 26.8%	.000***
Negative	[8.1] 64.9%	[9.1] 60.6%	.251	[8.2] 47.2%	[9.2] 10.8%	.000***

Table 3. Results (Mean/p-value) of the AFR comparison between large and small data for first and second digit in the complete, positive, and negative data sample (*p<.05, **p<.01, ***p<.001; N=25).

The results for the investigation of H2 are listed in Table 3. We were able to show significant differences between small and large entities for 5 of 6 investigations. The differences in AFR₁ are significant for the *complete* data with 10.7 percentage points ($T = 2.32, p < .05$) and *positive* data with 9.4 percentage points ($T = 2.16, p < .05$), yet not in *negative* data (3.3 percentage points). The largest differences can be observed in AFR₂. Here, we observe significant differences in all samples: 47.2 percentage points in *complete* ($T = 12.12, p < .001$), 40.8 percentage points in *positive* ($T = 10.54, p < .001$), and 36.4 percentage points in *negative* ($T = 10.11, p < .001$).

Data sample	AFR ₁			AFR ₂		
	[Sample] Mean		p	[Sample] Mean		p
	Positive	Negative		Positive	Negative	
Complete	[4.1] 79.2%	[5.1] 79.3%	.895	[4.2] 48.4%	[5.2] 20.0%	.000***
Large	[6.1] 70.0%	[8.1] 64.9%	.313	[6.2] 67.6%	[8.2] 47.2%	.000***
Small	[7.1] 60.6%	[9.1] 60.6%	.867	[7.2] 26.8%	[9.2] 10.8%	.000***

Table 4. Results (Mean/p-value) of the AFR comparison between positive and negative data for first and second digit in the complete, large, and small data sample (*p<.05, **p<.01, ***p<.001; N=25).

Table 4 contains the evaluation results for AFR₁ and AFR₂ with respect to H3. Interestingly, in none of the data samples *complete*, *large*, and *small*, a significant difference between *positive* and *negative* numbers can be shown for AFR₁. In contrast, in AFR₂ we find a highly significant deviation for all data samples. The difference varies from 16.0 percentage points to 28.4 percentage points, $T = 7.53, T = 6.51$ and $T = 5.06, p < .001$ in the *complete*, *positive* and *negative* data sample.

To summarize, we can confirm H2 for AFR₁ and AFR₂. The only non-significant difference (*negative*) has the correct direction, too. However, H3 can only be confirmed for AFR₂. Nevertheless, the results of this section are highly important for the design of a business intelligence service based on digital analyses: Whereas for an application of the AFR₁ as an error indicator only the company size is crucial and the sign can be ignored, AFR₂ is sensitive to company size and sign. Furthermore, according to the findings of Section 4.1, AFR₂ should only be applied to data samples of *large* entities.

4.3 Practical Relevance

The analyses presented in Sections 4.1 and 4.2 are the foundations for the design of our decision support service. The basic findings about the increasing conformity of financial planning data to Benford's Law over the time (cp. H1) form the very justification for the application of digital analyses in business intelligence systems. In accordance with the discovery of characteristics for the different data sub-samples (cp. H2 and H3), the knowledge about the conformity of financial planning data to Benford's Law enables us to provide specified recommendations. Transferring these results into IT-based, automated decision support services, we can enrich the digit analyses with both additional

statistical evaluations and expert knowledge. For example, different rounding behaviors or the consciously intended creation of duplicate numbers (e.g. through copy-and-pasting) affect the expected distribution of the delivered data. Adding expert knowledge about *compliance requirements of the respective company*, acceptable and non-acceptable adaptations can be classified.

An exemplary, real-world decision support service taken from our industry partner, could be the following: Since numbers above 100,000 require further planning details due to compliance rules, numbers slightly below 100,000 may be overrepresented to save the knowledge workers time and effort. Such knowledge can be transferred and automated into the decision support service along with the knowledge about the Benford distribution in financial planning data as shown in this paper. The manager, who accesses such a service through a business intelligence system, e.g. a corporate financial portal, may then be pointed to a following pattern: Numbers beginning with a “9” (e.g. 99,000) are highly overrepresented in the first digit with a probability of 11.5% (instead of the expected 4.9%). Based upon the results for H2 and H3, the recommendation can be further enhanced, for instance, by an investigation of the second digit for *positive* data of a *large* entity. As a result of this “alert”, the manager is able to further investigate this issue.

5 Conclusion and Future Work

This work transfers digital analyses (that is, analyses based on Benford’s Law) into a new domain: financial planning data. Our results are based on a substantial set of empirical data we were provided with by a globally acting, renowned large enterprise from the pharmaceutical and chemical industry. Via statistical analyses of this data, we were able to generally show that financial planning data in fact follows Benford’s Law as a contribution to the state of the science. To this end, we introduced the average fulfillment rate (AFR) as a new quality measure to enhance the interpretation of deviations from the Benford distribution. In more detail, we could show that the conformity of our data sample to Benford’s Law increases over the considered time period: a significant dependency between progressing time and increasing AFR was demonstrated for 15 of a total of 18 data treatments. In order to transfer these findings into business relevant decision support services and to enhance the assessment of financial planning data quality, the data structure was investigated in detail in two group analyses. That way, we conducted valuable analyses to validate if decision support services that incorporate Benford’s Law are appropriate to increase perceived planning data quality. In more detail, we significantly showed that the AFR in the underlying financial planning data depends on the entity size and the AFR in data delivered by large entities is higher than in the financial planning data delivered by small entities. For the second group analysis that was designed to validate if the AFR in the underlying financial planning data with a positive prefix is higher than in financial planning data with negative prefix, we found controversial results. Thus, the suggested separation of positive and negative numbers related to Benford’s Law (cp., for instance, Carslaw, 1988; Thomas, 1989) does not play a significant role in the underlying domain.

Equipped with these results, we are able to address the integration of digital analyses into information systems in our future work as required by recent papers that deal with the application of Benford’s Law (cp. e.g. Nigrini, 2000; Rezaee et al., 2002). Yet, so far, digital analyses have mostly been applied to static data. With the results achieved in this work, we will be able to realize a concrete implementation of digital analyses in of such a service within a business intelligence system. Beyond that, our results will enable us to present a service able to cope with dynamically growing data sets in the planning domain. For additional evaluation of our findings and to further assess the financial planning data quality, we will benchmark digital analyses against the concept of *weak planning efficiency* (Nordhaus, 1987) and *planning accuracy*, i.e. the difference between plan and actual values. Altogether, an automated decision support service will tremendously decrease the complexity of the planning data review and, at the same time, improve the quality of the forecast data.

References

- Benford, F. (1938). The law of anomalous numbers. In *Proceedings of the American Philosophical Society*, 78(4), 551-572, Lancaster Press, Lancaster.
- Busta, B. and R. Weinberg (1998). Using Benford's law and neural networks as a review procedure. *Managerial Auditing Journal*, 13(6), 356-366.
- Carlsaw, C.A. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *Accounting Review*, 63(2), 321-327.
- Diekmann, A. (2007). Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. *Journal of Applied Statistics*, 34(3), 321-329.
- Drake, P. D. and M. J. Nigrini (2000). Computer assisted analytical procedures using Benford's Law. *Journal of Accounting Education*, 18(2), 127-146.
- Durtschi, C., W. Hillison, and C. Pacini (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting*, 5(1), 17-43.
- Field, A. (2009). *Discovering Statistics Using SPSS*. 3rd Edition. SAGE, London.
- Graham, J. R. and C. R. Harvey (2001). The theory and practice of corporate finance: evidence from the field. *Journal of financial economics*, 60(2-3), 187-243.
- Hill, T. (1995). A statistical derivation of the significant-digit law. *Statistical Science*, 10(4), 354-363.
- Hill, T. (1998). The First Digit Phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist*, 86(4), 358-363.
- Kim, C.-S., Mauer, D. C., & Sherman, A. E. (1998). The Determinants of Corporate Liquidity: Theory and Evidence. *The Journal of Financial and Quantitative Analysis*, 33(3), 335-359.
- Krakar, Z. and M. Zgela (2009). Application of Benford's law in payment systems auditing. *Journal of Information and Organizational Sciences*, 33(1), 39-51.
- Newcomb, S., (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, 4(1), 39-40.
- Nigrini, M.J. (1996). A taxpayer compliance application of Benford's law. *The Journal of the American Taxation Association*, 18(1), 72-91.
- Nigrini, M.J. and L.J. Mittermaier (1997). The use of Benford's Law as an aid in analytical procedures. *Auditing*, 16(2), 52-67.
- Nigrini, M. J. (2000). *Continuous Auditing*. Ernst & Young Center for Auditing Research and Advanced Technology University of Kansas.
- Nordhaus, W. D. (1987). Forecasting Efficiency: Concepts and Applications. *The Review of Economics and Statistics*, 69(4), 667-674.
- Pinkham, R.S. (1961). On the distribution of first significant digits. *The annals of mathematical statistics*, 32(4), 1223-1230.
- Raimi, R.A. (1976). The first digit problem. *American Mathematical Monthly*, 83(7), 521-538.
- Rezaee, Z., A. Sharbatoghlie, R. Elam and P.L. McMickle (2002). Continuous auditing: Building automated auditing capability. *Auditing*, 21(1), 147-164.
- Thomas, J.K. (1989). Unusual patterns in reported earnings. *The Accounting Review*, 64(4), 773-787.
- Vo, H. T. K., Glaum, M., & Wojciechowski, R. (2007). Streamlining Foreign Exchange Management Operations with Corporate Financial Portals. *Handbook on Information Technology in Finance*, 123-139.
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5-33.
- Williamson, O. E. (1991). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly*, 36(2), 269-296.
- Witten, I., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.): Elsevier Inc.