



SCHRIFTLICHE AUSAR- BEITUNG

Case Study 1 SE S2020

Till Bieg, David Krug – Fachhochschule Wiener Neustadt



INHALT

Schriftliche Ausarbeitung	1
Tabellenverzeichnis	2
Abbildungsverzeichnis	2
1 Beschreibung des Technologie-Stacks.....	4
2 Beschreibung des ETL-Prozesses.....	4
3 Beschreibung der Datenbasis Und Prüfung der Integrität	5
3.1 Überblick.....	5
3.2 Bemerkungen zur Vollständigkeit der Datenbasis	5
3.3 Grundlegende Überprüfung der Datenintegrität	6
4 Weiterführende Analysen	7
4.1 Prüfungen der buchenden Personen	7
4.2 Prüfungen der Beleg- und Erfassungszeit.....	8
4.3 Prüfung der verwendeten Konten.....	11
4.4 Prüfung der Höhe beziehungsweise Ziffern von Beträgen	12
4.5 Prüfungen auf Unterschiede zum Vorjahr	18
5 Zusammenfassung.....	22
6 Referenzen.....	23

TABELLENVERZEICHNIS

Tabelle 1. Buchungseinträge mit unvollständigen Zeilennummern.....	6
Tabelle 2. Merkmale der Benutzer, die mit lückenhaften Buchungszeilen assoziiert sind.....	8
Tabelle 3. Auffällige Buchungstage (Belegdatum) auf Basis einer Ausreißer-Analyse mittels Isolation Forest.	11
Tabelle 4. Ziffern der positiven Beträge mit der höchsten absoluten Abweichung (First-Two-Digits-Test).	13
Tabelle 5. Ziffern der negativen Beträge mit der höchsten absoluten Abweichung (First-Two-Digits-Test).	14
Tabelle 6. Ziffernpaare mit der höchsten absoluten Abweichung nach positiven Beträgen und Kontotypen.....	15
Tabelle 7. Beträge mit der höchsten Buchungshäufigkeit.....	17
Tabelle 8. Beträge mit der niedrigsten Buchungshäufigkeit.....	18
Tabelle 9. Konten mit den höchsten Abweichungen in der Buchungsanzahl im Vergleich zum Vorjahr.	19
Tabelle 10. Beträge mit den höchsten Abweichungen (Buchungshäufigkeit) im Vergleich zum Vorjahr.....	20
Tabelle 11. Beträge mit den geringsten Abweichungen (Buchungshäufigkeit) Vergleich zum Vorjahr.	21
Tabelle 12. Tage mit der höchsten Anzahl an Buchungen (Jahr 2014).	21

ABBILDUNGSVERZEICHNIS

Abbildung 1. Visualisierung des ETL-Prozesses.	5
Abbildung 2. Verteilung der Buchungshäufigkeit pro Benutzer.	7



Abbildung 3. Anzahl an Buchungen pro Department.....	7
Abbildung 4. Anzahl an Buchungen pro Monat basierend auf dem Belegdatum.	9
Abbildung 5. Verteilung der Buchungshäufigkeit nach Belegdatum.....	9
Abbildung 6. Verteilung der Buchungshäufigkeit nach Erfassungsdatum.....	9
Abbildung 7. Anzahl an Buchungen pro Wochentag (Belegdatum).	10
Abbildung 8. Anzahl der Buchungen pro Wochentag (Erfassungsdatum).	10
Abbildung 9. Auffällige Buchungstage (Belegdatum) auf Basis einer Ausreißer-Analyse mittels Isolation Forest.	11
Abbildung 10. Verteilung der Buchungshäufigkeit nach Konten.	12
Abbildung 11. Verteilung der funktionalen Beträge nach positiven und negativen Beträgen.....	12
Abbildung 12. Ziffernverteilung der ersten zwei Ziffern in Bezug auf alle positiven Beträge im Buchungsjournal.	13
Abbildung 13. Ziffernverteilung der ersten zwei Ziffern in Bezug auf alle negativen Beträge im Buchungsjournal.	14
Abbildung 14. Ziffernverteilung (erste zwei Ziffern) nach positiven und negativen Beträgen und Kontotyp. Kontotyp „Equity“ aufgrund geringer Fallzahl nicht darstellt ($n = 4$).	16
Abbildung 15. Verlauf der summierten Beträge pro Monat.	18
Abbildung 16. Verlauf der Anzahl an Buchungen pro Monat.	19
Abbildung 17. Konten mit den höchsten Abweichungen in der Buchungsanzahl im Vergleich zum Vorjahr.	19
Abbildung 18. Beträge mit den höchsten Abweichungen in der Anzahl im Vergleich zum Vorjahr.	20
Abbildung 19. Anzahl der Buchungen pro Tag.....	21
Abbildung 20. Anzahl der Buchungen pro Tag (2013).....	22



Diese Ausarbeitung wurde im Rahmen der Lehrveranstaltung Case Study 1 SE S2020 an der Fachhochschule Wiener Neustadt erstellt. Sie beschreibt den Technology-Stack, ETL-Prozess und durchgeführte Analysen hinsichtlich eines rechnungslegungsbezogenen Datensatzes, der innerhalb der Lehrveranstaltung zur Verfügung gestellt wurde. Der Schwerpunkt der Analysen liegt dabei auf der Untersuchung möglicher betrügerischer Aktivitäten. Die Ergebnisse der Analysen werden abschließend in Form von Schlussfolgerungen kurz zusammengefasst.

1 BESCHREIBUNG DES TECHNOLOGIE-STACKS

Um die gegebenen Daten zu analysieren, wurden folgende Softwaretools verwendet:

- *Microsoft Excel 2019 (MS Excel)*
- *R*, Version 3.6.2 (R Core Team, 2019), mit *RStudio* und folgenden Packages:
 - *readxl*: Zum Import von *MS Excel*-Dateien in *R*
 - *dplyr*: Zur Datenmanipulation (z.B. Filtern, Selektieren, Aggregieren)
 - *tidyr*: Zur Datenmanipulation (z.B. Pivotieren)
 - *lubridate*: Zur Manipulation von zeitbezogenen Variablen (z.B. Datum)
 - *purrr*: Zum effizienten Arbeiten mit Funktionen und Vektoren
 - *forcats*: Zur Manipulation von nominal- und ordinalskalierten Variablen
 - *stringr*: Zur Manipulation von textbasierten Daten.
 - *ggplot2*: Zur Erstellung statischer Grafiken
 - *ggforce*: Zur Erstellung statischer Grafiken
 - *cowplot*: Zur Erstellung statischer Grafiken
 - *plotly*: Zur Erstellung interaktiver Grafiken
 - *rmarkdown*: Zur Erstellung und Export von Analyseberichten
 - *benford.analysis*: Zur Durchführung von Benford-Analysen
 - *solitude*: Zur Anpassung von Isolation Forests (z.B. Ausreißerdetektion)
 - *naniar*: Zur Analyse fehlender Werte
 - *janitor*: Zur Datenbereinigung
 - *here*: Zum Management von Systempfaden

Zur Erstellung der finalen Ausarbeitung und Präsentation wurden *Microsoft Word 2019 (MS Word)* und *Microsoft Powerpoint 2019 (MS Powerpoint)* verwendet. Außerdem wurde aufgrund des Schwerpunkts der Analyse auf betrügerische Aktivitäten und dolose Handlungen auf die Erstellung eines Dashboards (z.B. mittels *shiny*) verzichtet. Ein Dashboard ist vor allem zur Darstellung von Kennzahlenanalysen sehr gut geeignet – für die detaillierte Darstellung der Betrugsanalysen erscheint es aber eher unflexibel. Für die Präsentation wird stattdessen eine R-Session vorbereitet, innerhalb derer nach Bedarf Detailanalysen durchgeführt werden können.

2 BESCHREIBUNG DES ETL-PROZESSES

Für die Datenaufbereitung und -analyse wurde wie bereits erwähnt *R*, Version 3.6.2, verwendet (R Core Team, 2019).

Die Daten aus den bereitgestellten *MS Excel*-Dateien wurden unter Verwendung des *readxl*-Packages unmittelbar in die globale Umgebung von *R* geladen. Hierbei wurde darauf geachtet, dass die Daten innerhalb einer *MS Excel*-Datei gegebenenfalls auf mehrere Arbeitsblätter aufgegliedert sind. Obwohl die Datenmenge insgesamt relativ umfassend ist, war es ohne Weiteres möglich diese Daten in den Arbeitsspeicher zu laden und entsprechende Analysen durchzuführen. Daher war die Verwendung einer Datenbank im Rahmen des ETL-Prozesses nicht unbedingt notwendig. Zur weiteren Analyse wurde auf eine umfangreiche Auswahl an verfügbaren Packages zurückgegriffen (z.B. *dyplr*, *tidyr*, *lubridate*, siehe Abschnitt 1). Bei Bedarf wurden einzelne Datensätze über

Join-Funktionen des *dplyr*-Packages kombiniert. Zur visuellen Exploration der Daten und anschaulichen Darstellung der Analyse wurden Packages zur Datenvisualisierung wie *ggplot2* und *plotly* verwendet. Spezifische Analysen wie Benford- oder Ausreißeranalysen wurden unter Verwendung weiterer, dezidierter Packages durchgeführt (*benford.analysis* und *solitude*).

Zur Dokumentation der Analyse sowie Vorbereitung von Präsentation und schriftlicher Ausarbeitung wurde das Package *rmarkdown* verwendet. Die damit erstellte Analysedokumentation wurde nachfolgend in *docx*- und *pptx*-Formate exportiert und mit *MS Word* und *MS Powerpoint* weiter angepasst. Abbildung 1 stellt den ETL-Prozess visuell dar.

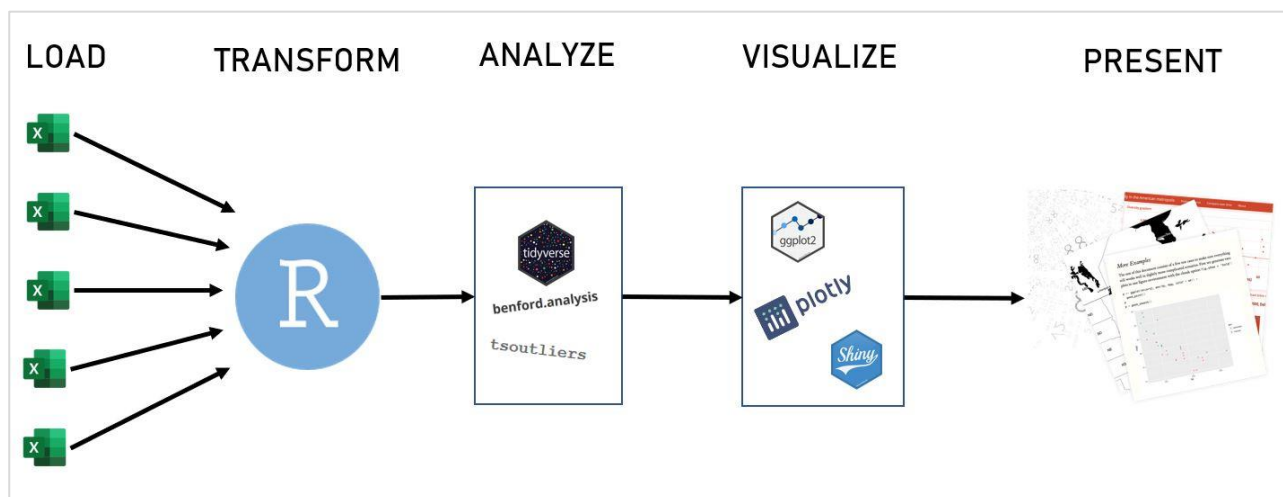


Abbildung 1. Visualisierung des ETL-Prozesses.

3 BESCHREIBUNG DER DATENBASIS UND PRÜFUNG DER INTEGRITÄT

3.1 Überblick

Wie bereits im Konzept beschrieben liegt der Schwerpunkt der Analyse auf der Prüfung der Daten auf dolose Handlungen / betrügerische Aktivitäten. Folgende Datensätze wurden von der ABC-Gesellschaft für die Prüfung zur Verfügung gestellt.

- Buchungsjournal für das Jahr 2014: Diese Datei enthält die Buchungseinträge aus dem Jahr 2014 (entspricht der zu prüfenden Periode).
- Buchungsjournal für das Jahr 2013: Diese Datei enthält die Buchungseinträge aus dem Jahr 2013 (Vorjahr der zu prüfenden Periode).
- Saldenliste („Trial Balances“) für das Jahr 2014: Diese Datei enthält die Anfangs- und Endsalden jedes Kontos für das Jahr 2014.
- Saldenliste für das Jahr 2014 nur mit Daten von Jänner bis September.
- Saldenliste für das Jahr 2013: Beinhaltet die Anfangs- und Endsalden jedes Kontos für das Jahr 2013.
- Kontenplan („Chart of Accounts“): Der Kontenplan beinhaltet die Gliederung von Konten zu Bilanz- oder Gewinn und Verlust-Posten.
- Sourceliste („Source Listing“): Enthält eine Zuordnung von „Sources“ zu Beschreibungen und Gruppen.
- Benutzerliste („User Listing“): Enthält eine Zuordnung verschiedener Systembenutzer zu Namen und Departments.

3.2 Bemerkungen zur Vollständigkeit der Datenbasis

Das Buchungsjournal für die zu prüfende Periode enthält in Bezug auf die meisten Variablen vollständige Angaben. Insbesondere enthalten die Spalten zum „funktionalen Betrag“ („Functional Amount“) und der Benutzer-ID („Preparer ID“) keine fehlenden Werte. Problematisch sind die fehlenden Werte in Bezug auf die Variable „JE Line Description“: Diese Variable enthält anstatt des



Buchungstexts (Beschreibung der Buchung) ausschließlich Platzhalter oder fehlende Werte. Das schränkt die Prüfbarkeit der Daten ein, da der Buchungstext unter Umständen Rückschlüsse auf den Inhalt der jeweiligen Buchung geben kann. Des Weiteren ist anzumerken, dass die Uhrzeit der Buchung beziehungsweise Erfassung nicht zur Verfügung gestellt wurde, was eine weitere Limitation bezüglich der Prüfbarkeit darstellt (die Buchungszeit könnte beispielsweise genutzt werden, um Buchungen¹ außerhalb der Geschäftszeiten zu identifizieren). Darüber hinaus enthält der Kontenplan 15.1% fehlende Werte in Bezug auf den Kontonamen und die Kontoklassen, so dass eine nähere Beschreibung und Eingliederung für diese Konten fehlt. Die Sourceliste blieb im Rahmen der Analyse weitestgehend unberücksichtigt, da diese eine hohe Anzahl von Platzhaltern (bis zu 74.4%) anstelle von aussagekräftigen Beschreibungen enthält. In der Benutzerliste sind Rolle und Titel des jeweiligen Mitarbeiters nicht enthalten. Diese Variablen könnten im Rahmen der Prüfung ebenfalls wertvolle Rückschlüsse erlauben. Die Daten der zur Verfügung gestellten Saldenlisten sind so weit vollständig.

Zusammenfassend sind die zu Verfügung gestellt Daten weitgehend vollständig – allerdings schränken einige fehlende Werte in den verschiedenen Datensätzen die Prüfbarkeit ein.

3.3 Grundlegende Überprüfung der Datenintegrität

Zunächst wurde überprüft, ob die Summe aller funktionalen Beträge für das Jahr 2014 0 entspricht. Dieses Integritätskriterium konnte bestätigt werden. Auch die Summen innerhalb jedes Journaleintrags (Summe der Buchungszeilen pro Journaleintrag) entsprechen jeweils 0, so dass auch dieses Kriterium erfüllt ist. Darüber hinaus wurde geprüft, ob die jeweiligen Anfangs- und Endsalden pro Konto zwischen dem Buchungsjournal und der Saldenliste übereinstimmen. Auch das konnte für alle Konten bestätigt werden. Allerdings zeigte eine weitere Prüfung, dass die **Zeilennummern in Bezug auf die jeweiligen Journaleinträge nicht immer lückenlos** sind. Es gibt zumindest 9 Journaleinträge, die nicht lückenlos sind. Diese Journaleinträge sind in Tabelle 1 näher beschrieben.

Tabelle 1. Buchungseinträge mit unvollständigen Zeilennummern.

JE-Nummer	Tatsächliche Anzahl an Zeilen	Erwartete Anzahl an Zeilen
2014-0001-2400008088	4	8
2014-0001-2400013762	4	8
2014-0001-4200036317	2	3
2014-0001-4200036320	2	3
2014-0001-4200036323	2	3
2014-0001-4200036332	2	3
2014-0001-4200039194	4	7
2014-0001-4200046717	2	3
2014-0001-4200050312	2	3

Wie in der Tabelle zu sehen fehlen je nach Eintrag bis zu vier Buchungszeilen. Das kann ein erster Hinweis auf dolose Handlungen oder ähnliches sein. Die Buchungen sind assoziiert mit insgesamt zwei individuellen Benutzern („SYS“ und „SheA101“), zwei Sourcegruppen („Accruals“ und „Goods receipts“) und vier verschiedenen Konten („0060000500“, „0140520850“, „0140203005“, „0140203004“).

¹ Im Rahmen dieser Ausarbeitung wird der Begriff „Buchung“ synonym mit „Buchungszeile“ verwendet.

Zusammenfassend erfüllen die zur Verfügung gestellten Daten einen überwiegenden Teil der grundlegenden Integritätskriterien. Allerdings gibt es Journaleinträge mit unvollständigen Zeilennummern, was ein erster Hinweis auf dolose Handlungen ist und bei den weiterführenden Analysen genauer untersucht wird.

4 WEITERFÜHRENDE ANALYSEN

4.1 Prüfungen der buchenden Personen

Die folgenden Analysen beziehen sich schwerpunktmäßig auf die buchenden Personen, um mögliche Auffälligkeiten in diesem Kontext zu identifizieren und genauer zu untersuchen. Im Buchungsjournal haben insgesamt 48 Benutzer Buchungen durchgeführt. Abbildung 2 stellt die Verteilung der Buchungshäufigkeit nach Benutzern dar (die Buchungshäufigkeit ist dabei auf einer logarithmischen Skala visualisiert). Wie in der Abbildung zu erkennen gibt es erhebliche Unterschiede in der Buchungshäufigkeit zwischen Nutzern: So gibt es einen Nutzer mit fast 100.000 Buchungen, andere Benutzer haben weniger als 10 Buchungen. Der Median der Buchungshäufigkeit liegt bei ca. 1000 Buchungen, die Spannweite ist allerdings erheblich. Die Buchungshäufigkeit kann als Indikator für die Expertise von Benutzern gesehen werden – dahinter steht die Annahme, dass Mitarbeiter, die häufiger Buchungen durchführen, insgesamt erfahrener sind und daher weniger Fehler machen.

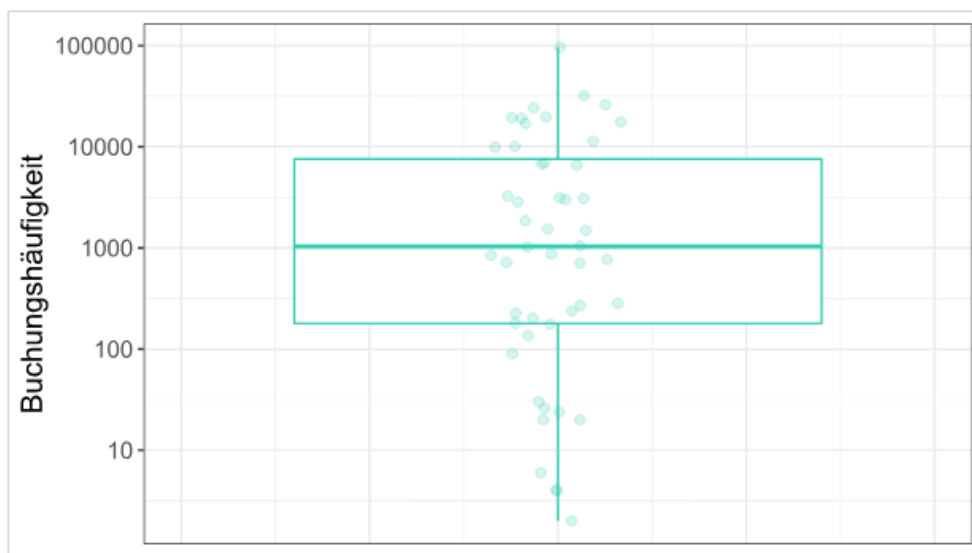


Abbildung 2. Verteilung der Buchungshäufigkeit pro Benutzer.

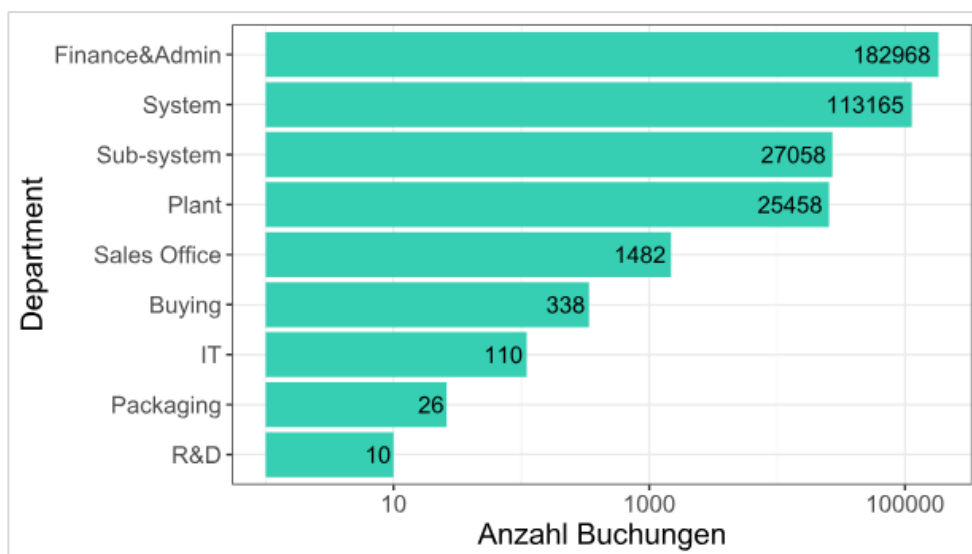


Abbildung 3. Anzahl an Buchungen pro Department.



In diesem Zusammenhang erscheint auch ein Blick auf die Buchungshäufigkeiten der verschiedenen Departments sinnvoll, denen die Benutzer zugeordnet sind. Das gibt Aufschluss darüber, welche Departments typischerweise Buchungen durchführen. Abbildung 3 stellt die Anzahl an Buchungen für die verschiedenen Departments dar. Auch zwischen den Departments zeigen sich erhebliche Unterschiede bei den Buchungshäufigkeiten: Das Department „Finance&Admin“ hat über 180.000 Buchungszeilen erfasst, das Department „R&D“ nur insgesamt 10 Buchungen.

Im Folgenden sollen darüber hinaus jene zwei Benutzer genauer betrachtet werden, die mit den lückenhaften Journaleinträgen assoziiert sind. Tabelle 2 stellt Merkmale dieser Benutzer da.

Tabelle 2. Merkmale der Benutzer, die mit lückenhaften Buchungszeilen assoziiert sind.

Benutzername	Ganzer Name	Department	Rolle	Titel	Anzahl Buchungen
SheAI01	Alan Shepard	Finance&Admin	<i>Fehlend</i>	<i>Fehlend</i>	17561
SYS	SYS SYS	System	<i>Fehlend</i>	<i>Fehlend</i>	16952

Beide Benutzer gehören Departments an, die viele Buchungen durchgeführt haben. Die Benutzer selbst haben auch überdurchschnittlich viele Buchungen durchgeführt. Das lässt darauf schließen, dass es sich um „erfahrene“ Benutzer, die vergleichsweise wenige Fehler bei der Systembenutzung machen. Es erscheint also tendenziell unwahrscheinlich, dass die lückenhaften Journaleinträge durch Irrtum zustande gekommen sind. Eine alternative Erklärung wären allerdings geteilte Benutzerkonten, bei denen mehrerer natürliche Personen Zugriff auf dasselbe Benutzerkonto haben. Bei dem Benutzer „SYS“ lässt sich darüber hinaus vermuten, dass es sich hierbei um ein automatisiertes Buchungsprogramm handelt. Dies sollte genauer betrachtet werden, da ein fehlerhaftes automatisiertes Buchungssystem sich gegebenenfalls systematisch auf die Datenintegrität auswirken kann. Es empfiehlt sich jedenfalls weitere Details in Bezug auf diese Benutzer mit dem Management der ABC-Gesellschaft zu besprechen.

4.2 Prüfungen der Beleg- und Erfassungszeit

Uhrzeit und Datum von Buchungen können dazu dienen weitere Muster zu identifizieren, die auf Unstimmigkeiten beziehungsweise betrügerische Aktivitäten hindeuten. Wie bereits erwähnt fehlt die Uhrzeit von Buchungen in dem zur Verfügung gestellten Datensatz, so dass nur das Datum der Buchungen betrachtet werden kann.

Alle Belegdaten liegen im Jahr 2014, in Bezug auf die Erfassungsdaten ist das nicht der Fall (ca. 1.6% aller Buchungszeilen): Das Erfassungsdatum dieser Buchungszeilen liegt entweder im Jahr 2015 (4.847 Zeilen) oder im Jahr 2013 (842 Zeilen). Die Erfassungsdaten im Jahr 2013 erscheinen problematisch, da sie jedenfalls vor dem jeweiligen Belegdatum liegen. Das ist allerdings nicht nur für die Buchungszeilen aus dem Jahr 2013 der Fall: Insgesamt wurden 19.638 Buchungszeilen identifiziert, bei denen das Erfassungsdatum vor dem Belegdatum sind. Dies weist auf Verletzungen der Datenintegrität hin.

Um einen Überblick der Buchungen pro Monat zu erhalten, stellt Abbildung 4 diese visuell dar (basierend auf dem Belegdatum). Anhand dieser Darstellung zeigen sich keine auffälligen Muster in der Buchungsverteilung, insbesondere ist keine starke Häufung von Buchungen am Jahresende auffällig. Ähnlich verhält es sich, wenn die Buchungshäufigkeiten pro Monat basierend auf dem Erfassungsdatum betrachtet werden (hier nicht dargestellt).

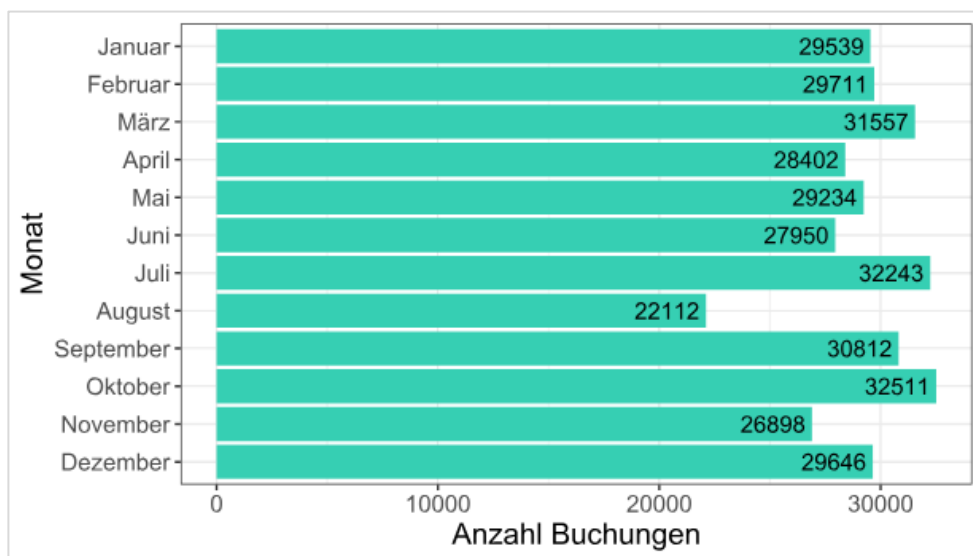


Abbildung 4. Anzahl an Buchungen pro Monat basierend auf dem Belegdatum.

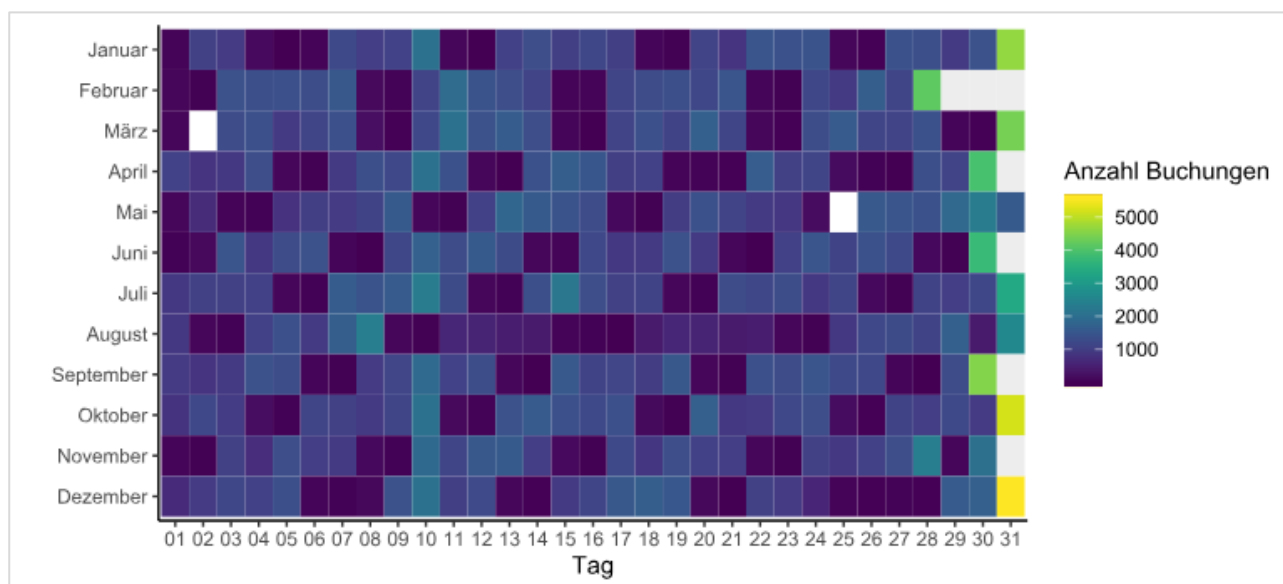


Abbildung 5. Verteilung der Buchungshäufigkeit nach Belegdatum.

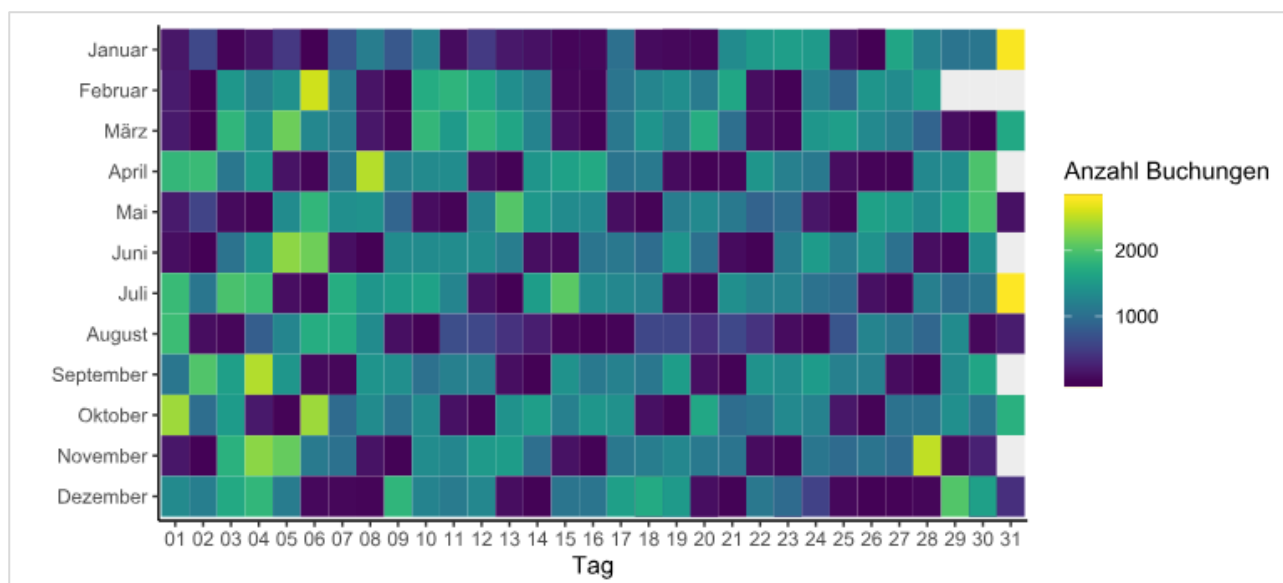


Abbildung 6. Verteilung der Buchungshäufigkeit nach Erfassungsdatum.

Interessante Muster zeigen sich allerdings bei der Betrachtung von Buchungshäufigkeiten pro Tag. In Bezug auf das Belegdatum ist eine starke Häufung von Buchungen zum Monatsende auffällig (siehe Abbildung 5). An sich ist der Umstand nicht verdächtig und hat vermutlich seinen Ursprung in bestimmten betrieblichen Abläufen (beispielsweise Projektabschlüsse), sollte aber ebenfalls mit dem Management der ABC-Gesellschaft besprochen werden.

In Bezug auf das Erfassungsdatum ist diese Häufung weniger stark ausgeprägt, allerdings gibt es auch hier Tage, die durch eine überdurchschnittliche hohe Anzahl von Buchungen auffällig sind (z.B. der 31. Januar, siehe Abbildung 6). Des Weiteren zeigt sich, dass sowohl das Datum der Erfassung als auch das Datum des Belegs von einigen Buchungen auf Samstage und Sonntage fallen (für eine detaillierte Darstellung siehe Abbildung 7 und Abbildung 8). Auch in Bezug auf diese Ergebnisse empfiehlt sich eine Betrachtung gemeinsam mit dem Management der ABC-Gesellschaft, um mögliche Erklärungen zu finden.

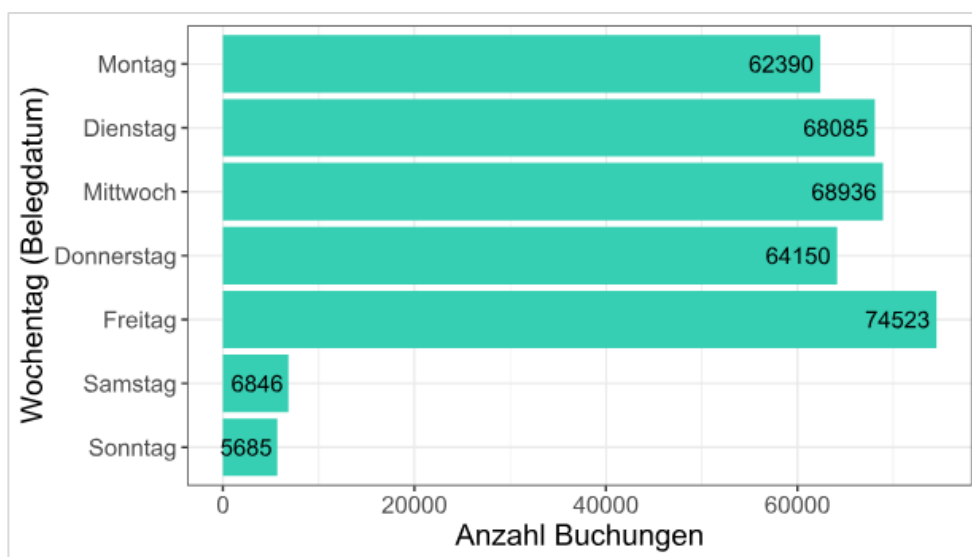


Abbildung 7. Anzahl an Buchungen pro Wochentag (Belegdatum).

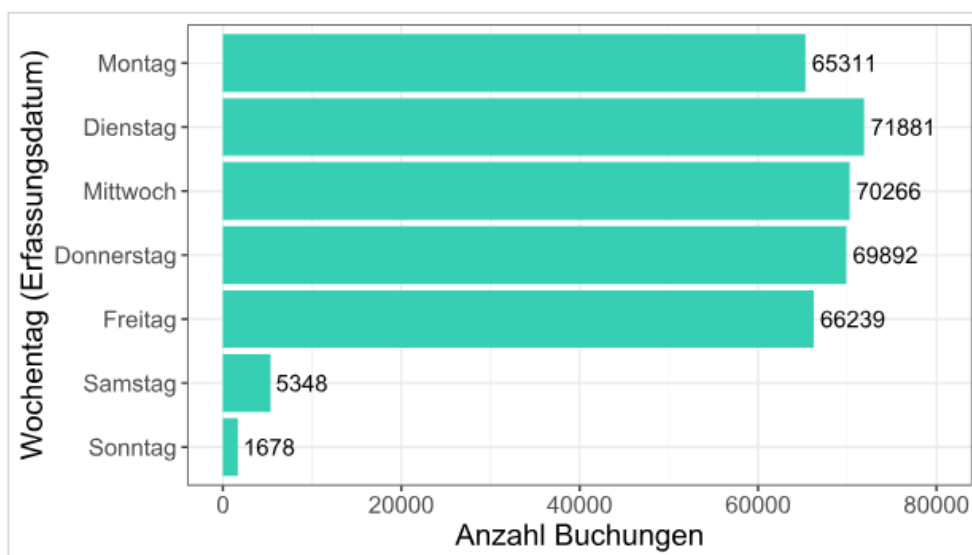


Abbildung 8. Anzahl der Buchungen pro Wochentag (Erfassungsdatum).

Außerdem wurden die Buchungszeiten von jenen Journaleinträgen untersucht, die mit den lückenhaften Buchungszeilen assoziiert sind. Bei dieser Betrachtung erscheinen die Buchungsdaten des Journaleintrags „2014-0001-4200050312“ besonders auffällig. Das Belegdatum dieser Buchungen ist der 31. Dezember 2014, erfasst wurden die Buchungen am 13. Januar 2015. Das Datum der

Buchungen ist verdächtig, da eine betrügerisch handelnde Person wissen könnte, ob bestimmte Jahresziele schon erreicht wurden und so leichter Buchungen einschleichen könnte. Der Buchungsbetrag beläuft sich nur auf 100,00 Euro, dennoch ist diese Auffälligkeit erwähnenswert (aufgrund der Kombinationen einer fehlenden Buchungszeile mit einem verdächtigen Buchungsdatum). Die Buchungsdaten anderer Journaleinträge mit lückenhaften Zeilen erscheinen diesbezüglich eher unverdächtig. Auch wurden keine lückenhaften Journaleinträge an Wochenende gebucht.

Schließlich wurde das Belegdatum mittels Isolation Forest auf Anomalien / Ausreißer untersucht. Isolation Forests sind eine Methode aus dem Bereich des maschinellen Lernens zur Detektion von Ausreißern. Bei der Anwendung des Isolation Forests wurde folgende Merkmale berücksichtigt: Die Summe der Beträge pro Tag und die Anzahl der Buchungen pro Tag.

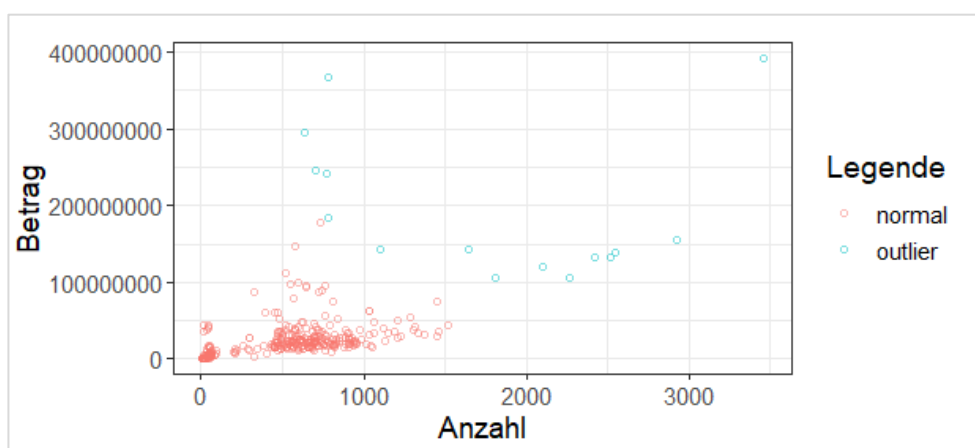


Abbildung 9. Auffällige Buchungstage (Belegdatum) auf Basis einer Ausreißer-Analyse mittels Isolation Forest.

In Abbildung 9 sind zwei Cluster erkennbar. Eines kennzeichnet die Daten, die aufgrund der Höhe der Summe der täglichen Beträge beziehungsweise der Häufigkeit der täglichen Buchungen nicht auffällig sind (rote Punkte). Das zweite Cluster enthält Ausreißer (blaue Punkte). Weitere Details zu diesen Ausreißern sind in Tabelle 3 dargestellt. Ob diese Ausreißer inhaltliche Bedeutung haben ist nach Rücksprache mit der ABC-Gesellschaft zu beurteilen.

Tabelle 3. Auffällige Buchungstage (Belegdatum) auf Basis einer Ausreißer-Analyse mittels Isolation Forest (zu zusätzlich zu Tagen mit einer sehr hohen Anzahl an Buchungen).

Datum	Betrag	Anzahl der Buchungen
2014-02-03	366 366 763	782
2014-02-04	240 540 541	771
2014-02-05	295 927 354	636
2014-07-31	142 443 704	1 649
2014-12-03	245 135 053	706
2014-12-05	183 736 906	784
2014-12-30	143 615 177	1 097

4.3 Prüfung der verwendeten Konten

Die Betrachtung von ungewöhnlichen beziehungsweise selten verwendeten Konten kann weitere Hinweise auf betrügerische Aktivitäten geben. Für einen Überblick über die Buchungshäufigkeit nach Konten siehe Abbildung 10.

Insgesamt enthält der Datensatz Buchungszeilen von 476 Konten. Die Buchungshäufigkeit nach Konten erstreckt sich über mehrere Größenordnungen, ungefähr die Hälfte aller Konten ist mit 30 oder weniger Buchungen assoziiert. Es wurden 121 Konten mit 10 Buchungen oder weniger identifiziert. Journaleinträge mit lückenhaften Buchungszeilen sind diesbezüglich kaum auffällig, da die

jeweiligen Konten mindestens 45 Buchungen verzeichnet haben (siehe Tabelle 4). Diese Untersuchung wurde als Vorbereitung auf weitere Detailbetrachtungen im Rahmen der Benford-Analysen durchgeführt, da Konten mit einer geringen Buchungshäufigkeit in Kombination mit anderen Auffälligkeiten auf betrügerische Aktivitäten hinweisen können.

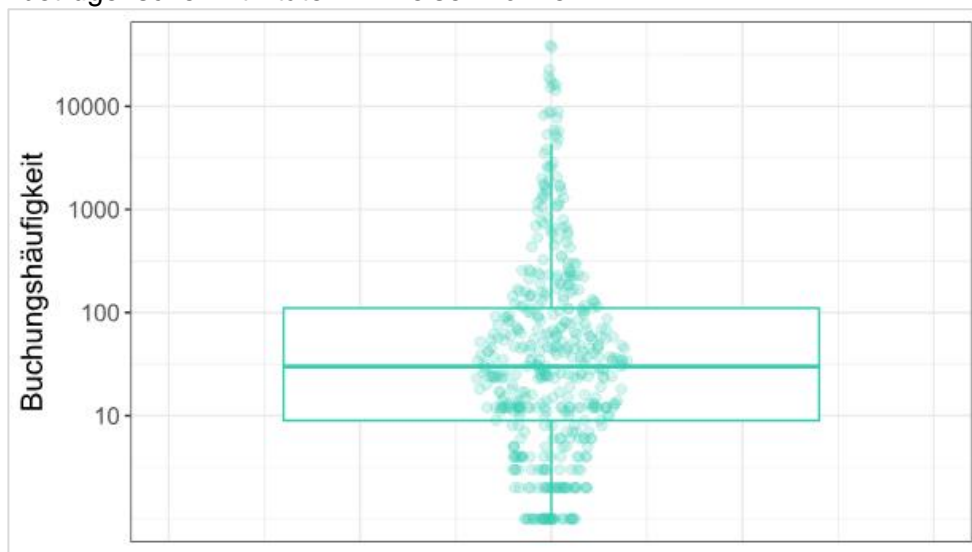


Abbildung 10. Verteilung der Buchungshäufigkeit nach Konten.

Tabelle 4. Konten assoziiert mit lückenhaften Journaleinträgen.

Nummer	Name	Gruppe	Anzahl Buchungen
0060000500	Bills to receive goods	Liabilities	37390
0140520850	Goodwill charges	Expenses	1603
014020300	Finished product purchases	Expenses	45
014020004	Packaging purchases	Expenses	15189

4.4 Prüfung der Höhe beziehungsweise Ziffern von Beträgen

Die Höhe und Ziffernverteilung von Beträgen kann ein wichtiger Anhaltspunkt für betrügerische Aktivitäten sein. Zur Prüfung von Auffälligkeiten in diesem Zusammenhang wurden Benford-Analysen durchgeführt und verdächtige Beträge begutachtet. Da die Verteilung von Ziffern in Bezug auf viele finanzielle Daten in der Regel der Benford-Verteilung folgt, sind Abweichungen von dieser Verteilung (das heißt, wenn die Ziffernverteilung von Buchungsbeträgen nicht der Benford-Verteilung folgt) ein Hinweis auf eine Manipulation der Daten.

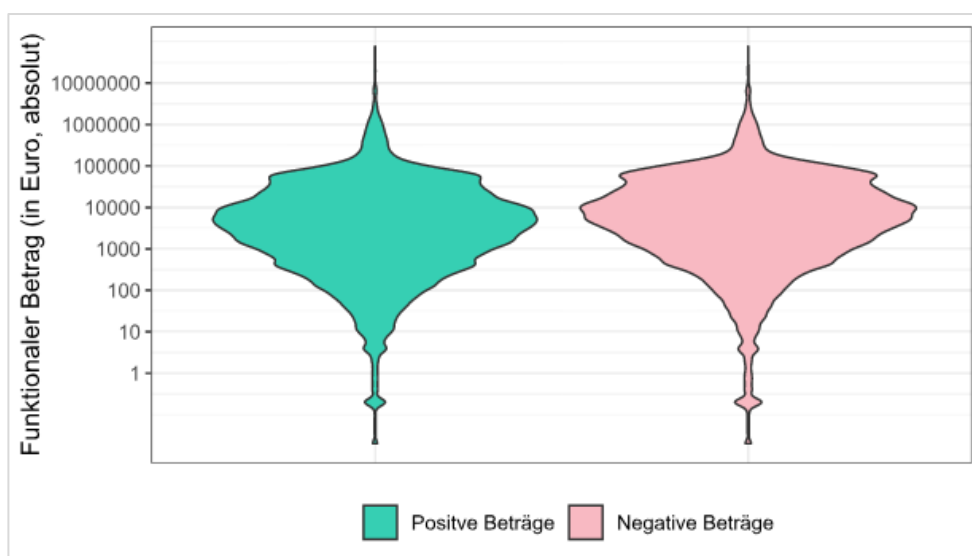


Abbildung 11. Verteilung der funktionalen Beträge nach positiven und negativen Beträgen.

Abbildung 11 stellt die Verteilung von Beträgen nach negativen und positiven Beträgen dar. Zu erkennen ist, dass sich erwartungsgemäß positive und negative Beträge annähernd gleich verteilen. Außerdem lässt sich erkennen, wo die „Schwerpunkte“ der Verteilung liegen. Beispielsweise gibt es relativ viele Beträge (positiv und negativ), die knapp unter 10.000 Euro liegen.

Zur Durchführung der Benford-Analysen wurden in jedem Fall separate Analyse für positive und negative Beträge durchgeführt. Dieses Vorgehen begründet sich darin, dass die Motivation zur Manipulation von positiven oder negativen Beträgen in der Regel unterschiedlich ist (Nigrini & Miller, 2009). Bei den durchgeführten Analysen wurden schwerpunktmäßig die ersten zwei Ziffern des jeweiligen Betrags gemeinsam betrachtet (der sogenannte „First-Two-Digits-Test“). Diese Betrachtungsweise ist auch dazu geeignet, Stichproben für weitere Prüfungshandlungen zu identifizieren, wohingegen die alleinige Betrachtung der ersten Ziffer der Beträge („First-Digit-Test“) nur sehr grobe Abweichungen in der Verteilung identifizieren kann (Association of Certified Fraud Examiners, n.d.).

Zunächst wurde eine Benford-Analyse für die Gesamtdatensatz untergliedert nach positiven und negativen Beträgen durchgeführt (ohne weitere Untergliederung beispielsweise nach Posten). In Bezug auf die positiven Beträge des Buchungsjournals zeigt sich eine insgesamt hohe Konformität der tatsächlichen Ziffernverteilung zur Benford-Verteilung: Die Mean Absolute Deviation (MAD) ist kleiner als 0.0001, was nach Nigrini (2012) einer hohen Konformität entspricht. Die Ergebnisse des χ^2 -Tests sind aufgrund des großen Stichprobenumfangs eher zu vernachlässigen ($\chi^2 = 3414.8$, $df = 89$, $p < 0.0001$). Abbildung 12 stellt die Verteilung der ersten beiden Ziffern der positiven Beträge visuell dar.

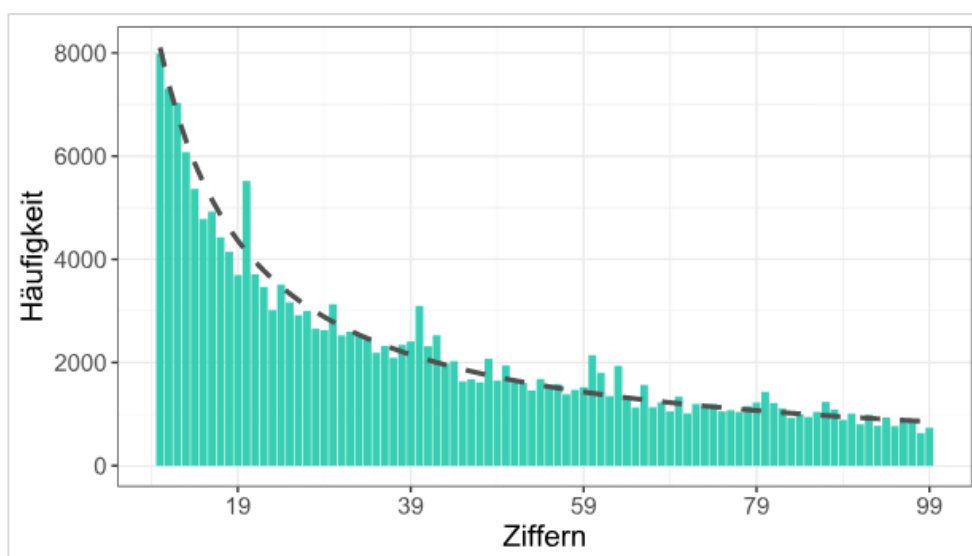


Abbildung 12. Ziffernverteilung der ersten zwei Ziffern in Bezug auf alle positiven Beträge im Buchungsjournal.

Tabelle 5. Ziffern der positiven Beträge mit der höchsten absoluten Abweichung (First-Two-Digits-Test).

Ziffern	Absolute Abweichung
20	1387.82
40	998.1
60	737.33
15	711.44
19	666.05

Bei visueller Inspektion des Plots zeigen sich trotz insgesamt hoher Konformität auffällige Abweichungen. Zum Beispiel gibt es viele Beträge die mit den Ziffern „20“, „40“, „60“ beginnen. Tabelle 5 stellt die fünf Ziffernpaare mit der höchsten absoluten Abweichung dar. Eine weitere Inspektion der Daten ergab, dass Beträge mit diesen Anfangsziffern über viele Größenordnungen verteilt sind (zum Beispiel gibt es viele Beträge in Höhe von 0.02 Euro, aber auch relativ viele Beträge in Höhe von ca. 20.000 Euro). Da die Ursachen für die beobachteten Abweichungen nicht notwendigerweise in der Manipulation von Beträgen liegen und Kontextwissen über Eigenschaften des geprüften Unternehmens notwendig ist, um diese Abweichungen sinnvoll interpretieren zu können, erscheint eine Abklärung dieser Ergebnisse mit dem Management der ABC-Gesellschaft notwendig (Tabelle 5. Ziffern der positiven Beträge mit der höchsten absoluten Abweichung (First-Two-Digits-Test)).

Die Ergebnisse in Bezug auf die negativen Beträge sind sehr ähnlich zu denen der positiven Beträge. Auch hier zeigt sich eine hohe Konformität der tatsächlichen Ziffernverteilung zur Benford-Verteilung ($MAD < 0.00001$). Die Ergebnisse des χ^2 -Tests sind wieder aufgrund des großen Stichprobenumfangs eher zu vernachlässigen ($\chi^2 = 2425.5$, $df = 89$, $p < 0.0001$).

Abbildung 13 stellt die Ziffernverteilung der ersten zwei Ziffern in Bezug auf alle negativen Beträge dar. Hierbei sind ähnliche Abweichungen wie bei den positiven Beträgen zu erkennen. Tabelle 6 stellt darüber hinaus die Ziffernpaare mit der höchsten absoluten Abweichung in Bezug auf die negativen Beträge dar.

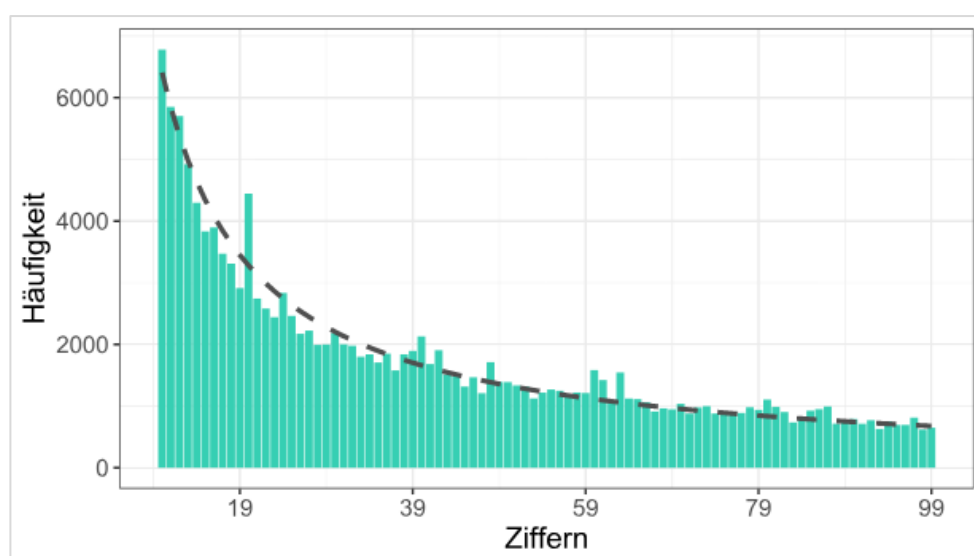


Abbildung 13. Ziffernverteilung der ersten zwei Ziffern in Bezug auf alle negativen Beträge im Buchungsjournal.

Tabelle 6. Ziffern der negativen Beträge mit der höchsten absoluten Abweichung (First-Two-Digits-Test).

Ziffern	Absolute Abweichung
20	1164.90
19	536.39
15	505.85
63	485.26
60	471.76

Auffällig ist vor allem wieder eine erhöhte Anzahl von Beträgen mit den Anfangsziffern „20“ und eine (entsprechend) niedrige Anzahl von Beträgen mit den Anfangsziffern „19“. Dies könnte auf die Aufrundung von Beträgen hinweisen.



Auf Basis dieser beiden Benford-Analysen wurden die verdächtigen Beträge (das heißt mit Anfangsziffern mit der höchsten absoluten Abweichung) genauer untersucht. Dazu wurden diese Beträge hinsichtlich weiterer Merkmale betrachtet, die im Vorhinein als auffällig identifiziert wurden:

- Von welchen Benutzern wurden die Buchungen durchgeführt?
- Wurden diejenigen Buchungen am Wochenende durchgeführt?
- Handelt es sich um Buchungen, deren Belegdatum nach dem Erfassungsdatum liegt?
- Beziehen sich die jeweiligen Buchungen auf Konten mit einer geringen Buchungshäufigkeit?

Bei dieser Betrachtung zeigte sich, dass viele individuelle Benutzer (bis zu 41 BenutzerInnen) mit den „verdächtigen“ Beträgen assoziiert sind. Dieser Umstand lässt die Manipulation der Daten durch einzelne Nutzer eher unwahrscheinlich wirken. Es wurden des Weiteren auch „verdächtige“ Buchungen identifiziert, die 1.) am Wochenende durchgeführt wurden, 2.) deren Belegdatum nach dem Erfassungsdatum liegt, 3.) mit Konten mit geringer Buchungshäufigkeit assoziiert sind. Aufgrund des Umfangs der Ausarbeitung werden diese nicht im Detail berichtet. Auf den ersten Blick ließen sich keine eindeutigen Muster in Bezug auf diese Buchungen feststellen, die auf Betrug oder ähnliches hinweisen. Allerdings liegt hier ein Ansatzpunkt für zukünftige Untersuchungen. Zur genaueren Betrachtung der Ziffernverteilung wurden auch Benford-Analysen für die jeweiligen Kontotypen (*Assets*, *Liabilities*, *Expenses*, *Revenue*) untersucht. Diese Aufgliederung bietet sich gut an, da 1.) anzunehmen ist, dass Manipulationen in Bezug auf die verschiedenen Kontotypen jeweils mit unterschiedlichen Motiven verknüpft sein können und 2.) die Fallzahl in Bezug auf die einzelnen Kontotypen ausreichend groß ist, um eine Benford-Analyse durchzuführen (ausgenommen Buchungen des Kontotyps *Equity*). Tabelle 7 stellt die entsprechenden Ziffernpaare mit der höchsten absoluten Abweichung dar. Abbildung 14 stellt darüber hinaus die Ziffernverteilung nach Kontotyp sowie positiven und negativen Beträgen dar. Hier zeigen sich einige interessante Auffälligkeiten, zum Beispiel einzelne hohe Abweichungen in Bezug auf *Expenses* oder eine hohe Anzahl des Ziffernpaars „10“ bei den negativen Beträgen des Kontotyps *Revenue*, was wieder auf gerundete Beträge hinweisen könnte. Wieder liegt hier aber ein Ansatzpunkt für zukünftige Analysen.

Kontotyp / Posten	Positive Beträge		Negative Beträge	
	Ziffernpaar	Absolute Abweichung	Ziffernpaar	Absolute Abweichung
Assets	15	302.66	20	850.94
	18	270.11	15	356.40
	44	270.06	23	324.24
	60	262.00	60	316.36
	66	245.66	21	285.61
Liabilities	42	263.63	42	265.69
	63	238.74	63	252.84
	12	220.10	48	224.72
	23	216.76	61	222.10
	61	214.48	12	217.93
Expenses	20	1263.47	20	161.14
	40	551.38	24	111.80
	60	323.81	40	99.49
	42	271.21	42	87.77
	19	256.02	11	60.96
Revenue	23	90.73	10	352.58
	14	79.97	86	103.20
	19	76.98	23	102.91
	41	72.16	21	98.15
	15	64.65	65	90.82

Tabelle 7. Ziffernpaare mit der höchsten absoluten Abweichung nach Beträgen und Kontotypen.

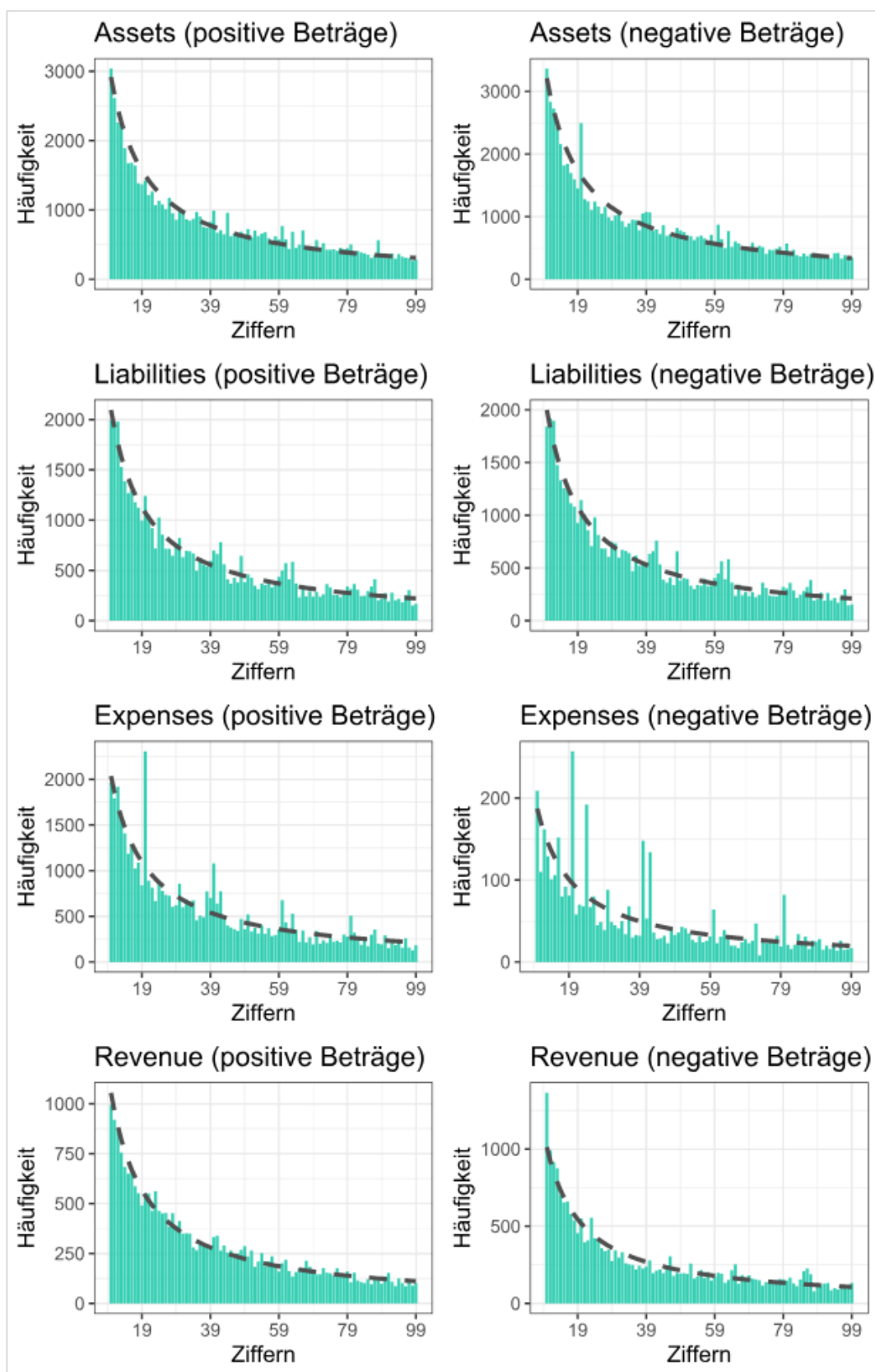


Abbildung 14. Ziffernverteilung (erste zwei Ziffern) nach positiven und negativen Beträgen und Kontotyp. Kontotyp „Equity“ aufgrund geringer Fallzahl nicht darstellt ($n = 4$).

Außerdem wurden einzelne Benford-Analysen für jeden buchenden Benutzer durchgeführt, um mögliche Manipulation durch einzelne Benutzer zu untersuchen. Die Ergebnisse dieser Analysen sind allerdings schwierig zu interpretieren, da nur die Ziffernverteilung in Bezug auf einen einzigen Nutzer (von insgesamt 48 Nutzern) eine hohe Konformität zur Benford-Verteilung aufweist. Daraus lässt sich schließen, dass eine Benford-Verteilung bezüglich der Buchungsziffern einzelner Nutzer im Vorhinein nicht anzunehmen ist und Abweichungen vermutlich keine Indizien für Manipulation darstellen.

Ähnlich verhält es sich mit den Ergebnissen der durchgeführten Last-Two-Digits-Analysen (Analyse der Verteilung der letzten beiden Stellen, um z.B. ungewöhnliche, gerundete Beträge zu entdecken): Nach grundlegender Datenexploration erscheint eine Last-Two-Digits-Analyse in Bezug auf die zu Verfügung gestellten Daten ungeeignet, da die Abweichungen zwischen den beobachteten und theoretischen Häufigkeiten so groß ist, dass im Vorhinein die Annahme „benfordverteilter“ Ziffern nicht plausibel ist.

Insgesamt führen Analysen auch die Herausforderung der Wahl eines passenden Granularitätsniveaus für Benford-Analysen vor Augen: Ist das Granularitätsniveau zu gering (z.B. bei Betrachtung des Gesamtdatensatz) könnten betrügerische Aktivitäten in der hohen Anzahl der Buchungszeilen übersehen werden. Bei zu hohem Granularitätsniveau (entspricht kleineren Datensätzen) weichen die Daten oft schon im Vorhinein von der Benford-Verteilung ab, sodass Benford-Analysen zur Untersuchung nicht geeignet sind. Benford-Analysen können die Aufdeckung von betrügerischen Aktivitäten also nicht garantieren. Die Benford-Analysen liefern aber einen wichtigen Ansatzpunkt für weitere Begutachtungen, die jedenfalls sorgsam unter Berücksichtigung von Kontext- und Domänenwissen durchgeführt werden sollten, um Auffälligkeiten bezüglich ihrer Relevanz besser einschätzen zu können. Weitere Benford-Analysen könnten dann gegebenenfalls zielgerichteter zwecks detaillierterer Untersuchung durchgeführt werden.

Neben den Benford-Analysen wurden Beträge auch bezüglich ihrer individuellen Buchungshäufigkeit betrachtet (Wie häufig wurde ein bestimmter Betrag gebucht?). Dies erfolgt unter der Annahme, dass Beträge in einer bestimmten Höhe regelmäßig gebucht werden. Tabelle 10 stellt Beträge mit der höchsten Buchungshäufigkeit dar.

Tabelle 8. Beträge mit der höchsten Buchungshäufigkeit.

Betrag	Anzahl der verschiedenen Tage	Anzahl der Buchungen
40 000.00	42	62
63 281.42	35	59
65 044.70	19	53
63 281.40	29	37
48 800.00	24	36
61 107.14	27	36
79 522.78	30	34
89 267.62	19	34
44 460.00	27	31
85 400.00	23	30

Einer Erklärung bedarf es hier hinsichtlich des Ursprungs der Beträge. Die Regelmäßigkeit der Buchungen lässt entweder auf einen wiederkehrenden Auftrag beziehungsweise ein bestimmtes Angebot schließen oder möglicherweise auf einen Betrag, der genau unterhalb einer bestimmten Grenze liegt. Interessant sind speziell die runden Beträge, da runde Beträge je nach Datengrundlage tendenziell eher selten vorkommen. Des Weiteren wurden auch die Beträge mit der niedrigsten Buchungshäufigkeit betrachtet. Diese sind in Tabelle 9 dargestellt.

Tabelle 9. Beträge mit der niedrigsten Buchungshäufigkeit.

Betrag	Anzahl der verschiedenen Tage	Anzahl der Buchungen
76 791 234.00	1	1
76 728 144.00	1	1
76 163 758.00	1	1
76 000 000.00	1	1
75 999 810.00	1	2
69 567 614.00	1	1
69 398 231.00	1	1
68 000 004.00	2	4
68 000 000.00	2	3
41 890 220.00	2	2

Diesbezüglich ist die Höhe der selten gebuchten Beträge auffällig: Alle diese Beträge sind größer als 40 Millionen Euro. Intuitiv ist zu vermuten, dass sehr hohe Beträge selten gebucht werden, so dass hier an sich keine Auffälligkeit besteht. Aus der Sicht einer Wirtschaftsprüfung sind diese selten gebuchten Beträge aber auch die Beträge, die besonders genau überprüft werden sollten (da Buchungen abseits von Routineabläufen ein höheres Fehlerpotential aufweisen).

4.5 Prüfungen auf Unterschiede zum Vorjahr

Aufgrund der bereitgestellten Daten vom Jahr 2013 wurden die Analysen ausgeweitet, um mögliche Abweichungen zwischen den beiden Geschäftsjahren zu analysieren. In der Ganzjahresanalyse (Abbildung 15) ist zu erkennen, dass die Jahre stark voneinander abweichen. Während im Jahr 2014 im Februar, Juli, Oktober und Dezember jeweils ein starker Anstieg der gebuchten Beträge verzeichnet wurde, sind die Veränderungen zwischen den Monaten im Jahr 2013 vergleichsweise gering. Insgesamt erscheint auch die Höhe der monatlichen Beträge im Jahr 2013 geringer. Dennoch ist zu erkennen, dass in beiden Jahren die Summe der Beträge im Dezember sehr stark angestiegen ist. Dies könnte beispielsweise mit erhöhten Umsätzen im Rahmen des Weihnachtsgeschäfts in Zusammenhang stehen. Ein konsistenter saisonaler Effekt lässt sich aufgrund dieser beiden Verläufe nicht erkennen, da innerhalb eines Quartals die gebuchten Beträge immer vergleichsweise stark gestiegen und gefallen sind. Lediglich das zweite Quartal unterscheidet sich in dieser Hinsicht.

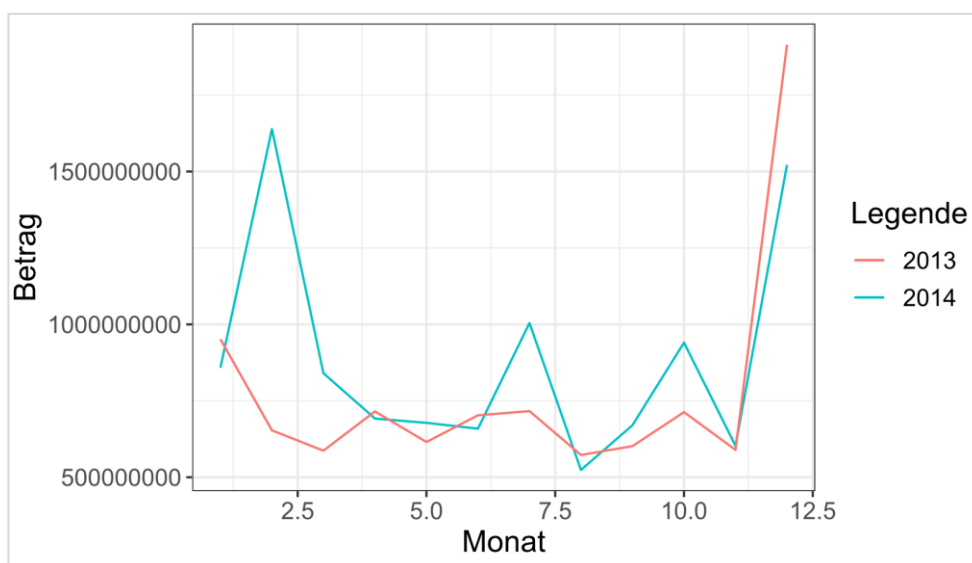


Abbildung 15. Verlauf der summierten Beträge pro Monat.

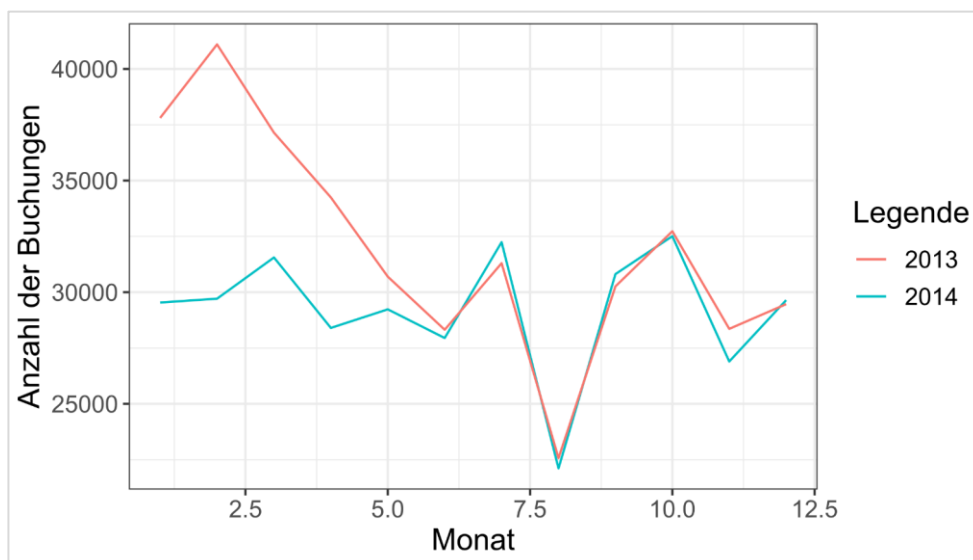


Abbildung 16. Verlauf der Anzahl an Buchungen pro Monat.

Beim Vergleich der Jahre 2013 und 2014 in Bezug auf die Anzahl der Buchungen ist in Abbildung 16 zu erkennen, dass im Jahr 2014 insgesamt weniger Buchungen durchgeführt wurden. Dennoch erscheint die Summe der Beträge (Abbildung 15) gegenüber dem Jahr 2013 höher. In diesem Sinne kann man sagen, dass im Jahr 2014 zwar insgesamt weniger gebucht wurde, aber die Beträge der Buchungen höher waren. Aus Abbildung 17 lassen sich darüber hinaus Abweichungen der Buchungsanzahl in Bezug einzelne Konto entnehmen (dargestellt sind die zehn Konten mit den höchsten Abweichungen zum Vorjahr). Dabei fallen vier Konten durch besondere hohe Abweichungen auf. Weitere Details diesbezüglich sind in Tabelle 10 dargestellt.

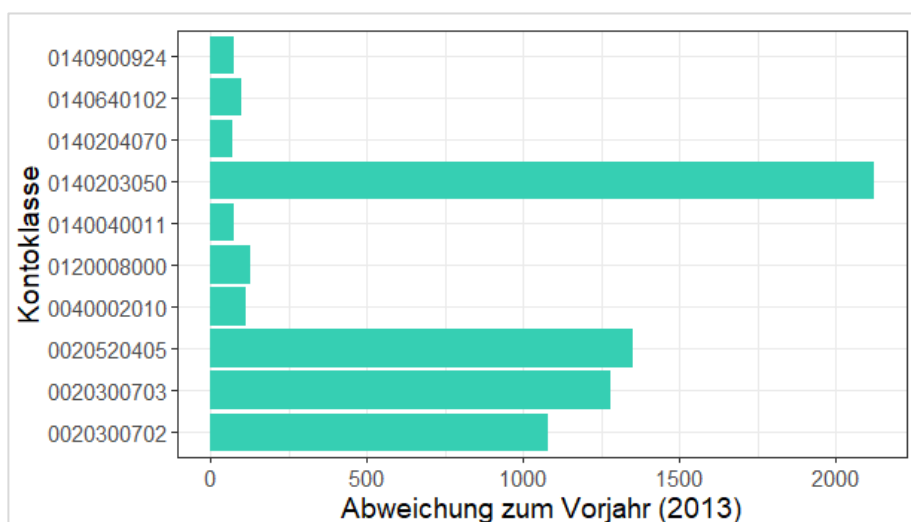


Abbildung 17. Konten mit den höchsten Abweichungen in der Buchungsanzahl im Vergleich zum Vorjahr.

Tabelle 10. Konten mit den höchsten Abweichungen in der Buchungsanzahl im Vergleich zum Vorjahr.

Konto	Anzahl der Buchungen (2014)	Anzahl der Buchungen (2013)	Abweichung in %
0140203050	733	33	2 121
0020520405	623	43	1 348
0020300703	303	22	1 277
0020300702	224	19	1 078
0120008000	39	17	129
0040002010	32	15	113



0140640102	60	30	100
0140900924	79	45	75
0140040011	135	78	73
0140204070	24	14	71

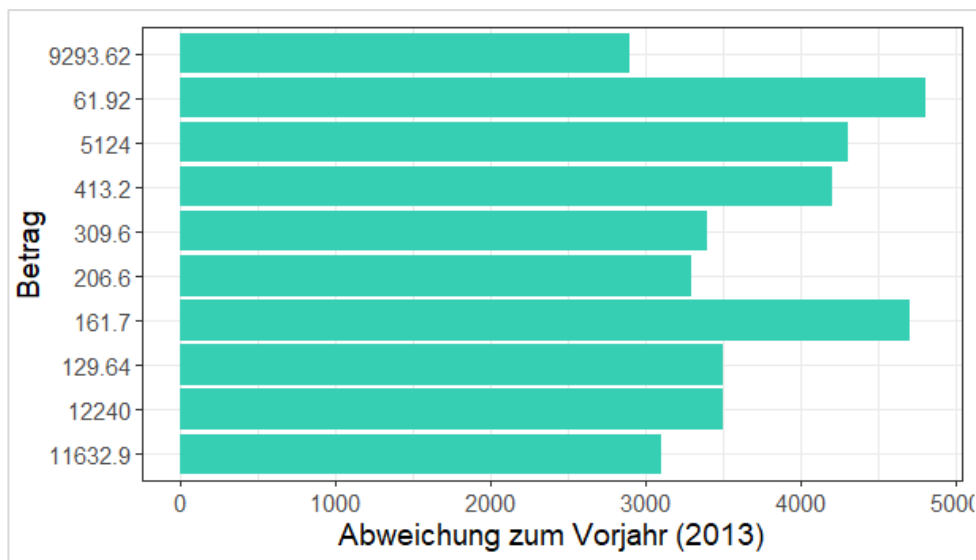


Abbildung 18. Beträge mit den höchsten Abweichungen in der Anzahl im Vergleich zum Vorjahr.

Unter der Annahme, dass auch Beträge in einer bestimmten Höhe regelmäßig gebucht werden, wurden auch die Buchungshäufigkeit in Bezug auf Beträge verglichen. Entsprechend stellt Abbildung 18 Beträge dar, die im Jahr 2013 selten und im Jahr 2014 häufig gebucht wurden (für weitere Details siehe Tabelle 11). Diese Abweichungen könnte möglicherweise ein Anzeichen für eine neue betrügerische Routine sein oder ein neues Produkt oder Produktkombination hat sich etabliert.

Tabelle 11. Beträge mit den höchsten Abweichungen (Buchungshäufigkeit) im Vergleich zum Vorjahr.

Betrag	Anzahl der Buchungen (2014)	Anzahl der Buchungen (2013)	Abweichung in %
61.92	49	1	4 800
161.7	48	1	4 700
5 124	44	1	4 300
413.2	86	2	4 200
129.64	72	2	3 500
12 240	36	1	3 500
309.6	35	1	3 400
206.6	68	2	3 300
11 632.9	32	1	3 100
9 293.62	30	1	2 900

Bei genauerer Betrachtung der Daten sind aber nicht nur die Beträge mit der größten Abweichung interessant, sondern auch Beträge mit der geringsten Abweichung zum Vorjahr. Die in Tabelle 12 enthaltenen Daten zeigen die zehn Beträge mit der geringsten Abweichung in der Buchungshäufigkeit. Derartige Beträge könnten auf eine einmalige, doppelt ausgestellte Rechnung hindeuten.



Tabelle 12. Beträge mit den geringsten Abweichungen (Buchungshäufigkeit) Vergleich zum Vorjahr.

Betrag	Anzahl der Buchungen (2014)	Anzahl (2013)	Abweichung in %
5 440.00	37	39	5.13
5 400.00	22	20	10.00
4 800.00	63	64	1.56
3 840.00	22	23	4.35
2 600.00	29	27	7.41
2 400.00	78	77	1.30
2 300.00	21	23	8.70
1 720.00	25	23	8.70
1 440.00	40	42	4.76
1 292.76	23	22	4.55

Des Weiteren ergänzt die folgende Analyse die Untersuchung aus Abschnitt 4.2, indem sie die Verteilungen des Belegdatums auch mit dem Vorjahr (2013) vergleicht. Tabelle 13 stellt nochmals die Tage mit der höchsten Anzahl an Buchungen (Belegdatum) für das Jahr 2014 dar.

Tabelle 13. Tage mit der höchsten Anzahl an Buchungen (Jahr 2014).

Datum	Anzahl der Buchungen
2014-01-31	2 546
2014-02-28	2 262
2014-03-31	2 423
2014-04-30	2 099
2014-06-30	1 809
2014-09-30	2 522
2014-10-31	2 928
2014-12-31	3 457

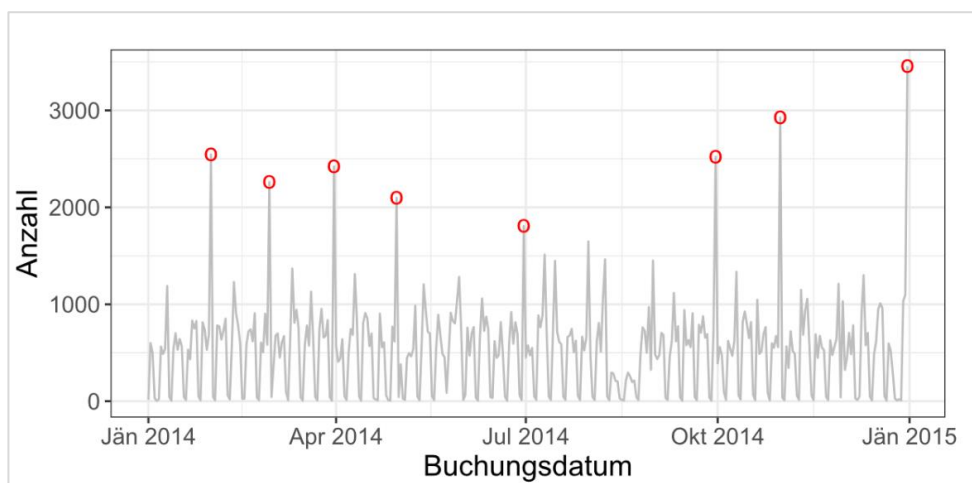


Abbildung 19. Anzahl der Buchungen pro Tag

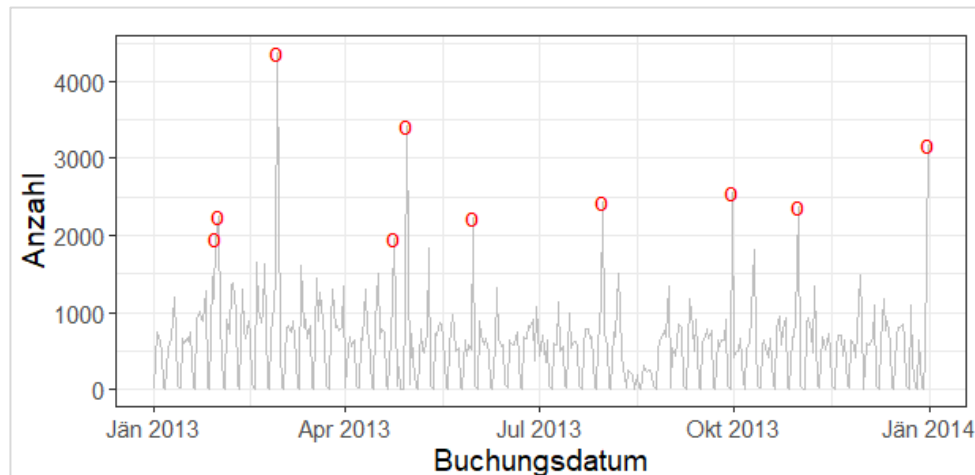


Abbildung 20. Anzahl der Buchungen pro Tag (2013).

Wie in Abbildung 19 und Abbildung 20 zu erkennen häufen sich Buchungen jeweils am Monatsende (für beide Jahre). Das kann verschiedene Erklärungen haben. Beispielweise können die Häufungen in Zusammenhang mit regelmäßigen monatlichen Abbuchungen wie Gehaltszahlungen an Mitarbeiter stehen. Jedenfalls zeigt der Vergleich zwischen dem Jahr 2014 und 2013, dass Buchungen am Ende des Monats als eher unverdächtig zu bewerten sind (unter der Annahme, dass die Daten aus dem Jahr 2013 umfassend geprüft wurden).

5 ZUSAMMENFASSUNG

Zusammenfassend wurde in Bezug auf den zu Verfügung gestellten Datensatz unter anderem folgende Auffälligkeiten identifiziert, die auf fehlerhafte Daten, Manipulation und / oder betrügerische Aktivitäten hinweisen:

- Journaleinträge mit unvollständigen Zeilennummern
- Journaleinträge mit Eintragsdatum vor Belegdatum
- Buchungszeilen mit einem Erfassungsdatum aus dem Jahr 2013
- Auffälligkeiten in Bezug auf die Höhe, Ziffern oder Häufigkeit von bestimmten Beträgen
- Einträge mit Beleg- und/oder Erfassungsdatum am Wochenende
- Auffällige Abweichungen zum Vorjahr
- Kombinationen dieser Auffälligkeiten in bestimmten Buchungszeilen

Es ist darauf hinzuweisen, dass es sich bei den gefundenen Auffälligkeiten lediglich um Anhaltspunkte für fehlerhafte Daten beziehungsweise betrügerische Aktivitäten handelt. Darüber hinaus ist die umfassende Prüfbarkeit der Daten aufgrund fehlender Angaben eingeschränkt (z.B. fehlen Buchungstext und Uhrzeit). Es wurden in diesem Sinne keine „handfesten“ Beweise für betrügerische Aktivitäten gefunden, sondern Ansatzpunkte für tiefergehende Untersuchungen. Die Bewertung und Interpretation der gefundenen Auffälligkeiten sollten sorgsam unter Berücksichtigung von Kontext- und Domänenwissen erfolgen (z.B. Details des Buchungssystems, betriebliche Abläufe). Entsprechend empfiehlt es sich die gefundenen Auffälligkeiten mit dem Management beziehungsweise fachkundigen Mitarbeitern der ABC-Gesellschaft gemeinsam zu betrachten und zu besprechen. Auf dieser Grundlage können dann gegebenenfalls ausgewählte Aspekte noch genauer untersucht werden.



6 REFERENZEN

- Association of Certified Fraud Examiners (n.d.). Using Benford's law to detect fraud. Abgerufen von https://www.acfe.com/uploadedFiles/Shared_Content/Products/Self-Study_CPE/Using-BenfordsLaw_2018_final_extract.pdf.
- Nigrini, M. J., & Miller, S. J. (2009). Data diagnostics using second-order tests of Benford's law. *Auditing: A Journal of Practice & Theory*, 28(2), 305-324.
- Nigrini, M. J. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection* (Vol. 586). John Wiley & Sons.
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.