



Arbeiten mit dem Demo-Datenbestand

Ausgehend vom Menüpunkt **Datenanalyse** wählen Sie den Link **interaktive Ziffernanalyse**.



Wenn Sie nicht bereits als Benutzer angemeldet sind gehen Sie bitte in den Menüpunkt **Auswertung** und melden sich als Demo-Benutzer an:

der Benutzername ist dabei **demo@dv-gutachten.de**; als Passwort geben Sie **demo** ein.

Anschließend bestätigen Sie die Eingaben durch **Absenden**. Danach haben Sie die Möglichkeit mit dem Demo-Datenbestand des Systems zu experimentieren. Zuerst müssen Sie den Datenbestand auswählen:

Auswertung anzeigen

Nutzer: demo@dv-gutachten.de Projekt: -- Bitte wählen --

Daten

Bezugsgröße

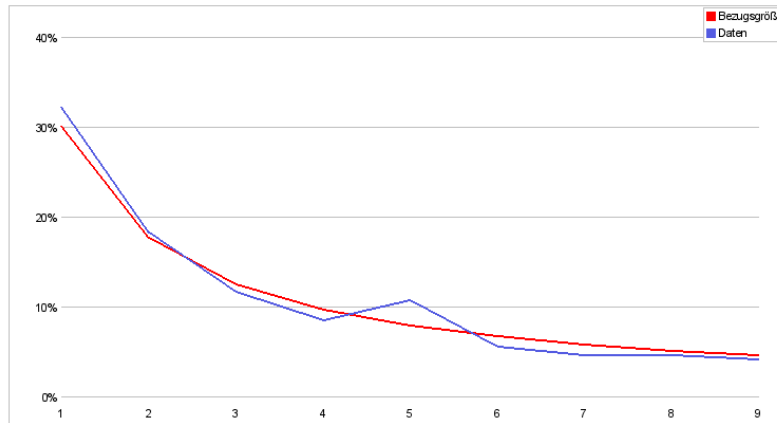
Referenz: Theorie

Er ist im Bereich globale Projekte gespeichert und trägt die Bezeichnung **Demodaten**.

Direkt nach der Auswahl erscheint der Analysebildschirm mit einer grafischen Auswertung des Gesamt-Datenbestands. Dabei werden die beobachteten Werte aus unseren Demo-Daten der zu erwartenden Verteilung aus der Theorie nach Benford gegenübergestellt.



Ziffernanalyse



Im vorliegenden Fall können wir sehen, dass die **5** häufiger auftritt, als dies nach der Theorie zu erwarten gewesen wäre. Bei Umsatzdaten kann dies daran liegen, dass die Produkte überdurchschnittlich häufig Preise haben, die mit einer 5 beginnen und oft einzeln verkauft werden. Für Verzerrungen dieser Art gibt es eine Vielzahl von Ursachen, deren Herausarbeitung Gegenstand eines entsprechenden Sachverständigengutachtens ist.

In dem zweiten Block darunter zeigt das System die numerischen Werte der beiden Verteilungen sowie die Abweichung zwischen den beiden Reihen.

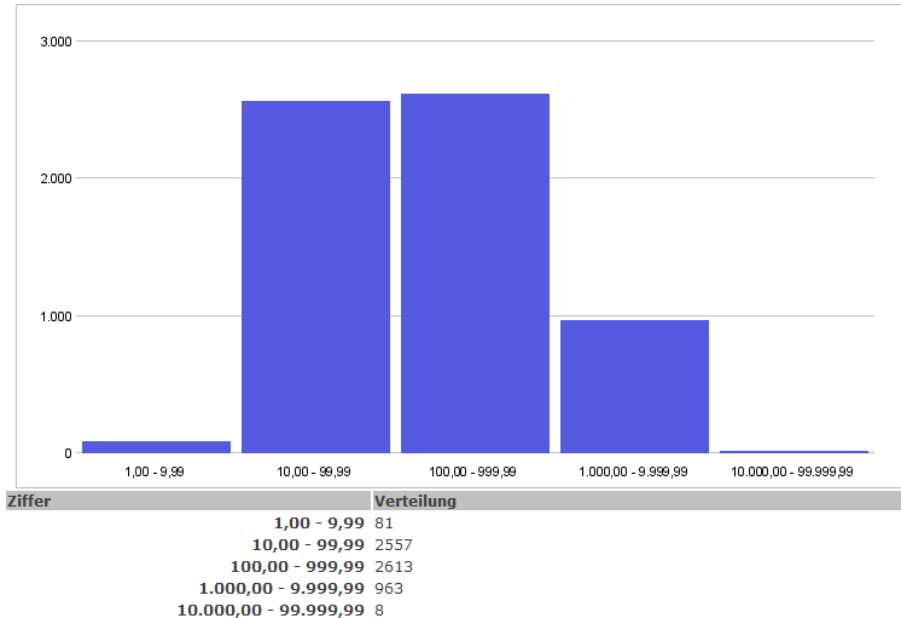
Ziffer	Verteilung	Referenz	Abweich.
Anzahl Datensätze	6222	N/A	
1	32,27%	30,10%	2,17%
2	18,19%	17,60%	0,59%
3	11,70%	12,49%	-0,79%
4	8,42%	9,69%	-1,27%
5	10,70%	7,91%	2,79%
6	5,48%	6,69%	-1,21%
7	4,60%	5,79%	-1,19%
8	4,55%	5,11%	-0,56%
9	4,08%	4,57%	-0,49%
χ^2	1,96	0,00	

Im letzten Bereich wird nun die **Verteilung** der Beträge in **Größenklassen** dargestellt. Das ist nötig weil der Einsatz der Ziffernanalyse an die Voraussetzung geknüpft ist, dass der zugrunde liegende Datenbestand logarithmisch normalverteilt ist. Die Balkengrafik gibt nun einen recht



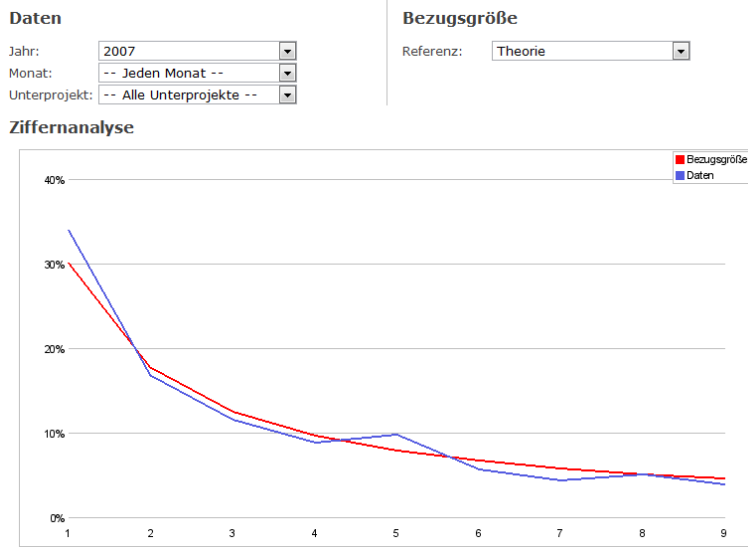
guten Eindruck davon, ob diese Voraussetzung erfüllt ist. In unserem Fall spricht nichts gegen die Durchführung der Ziffernanalyse.

Verteilung der Beträge



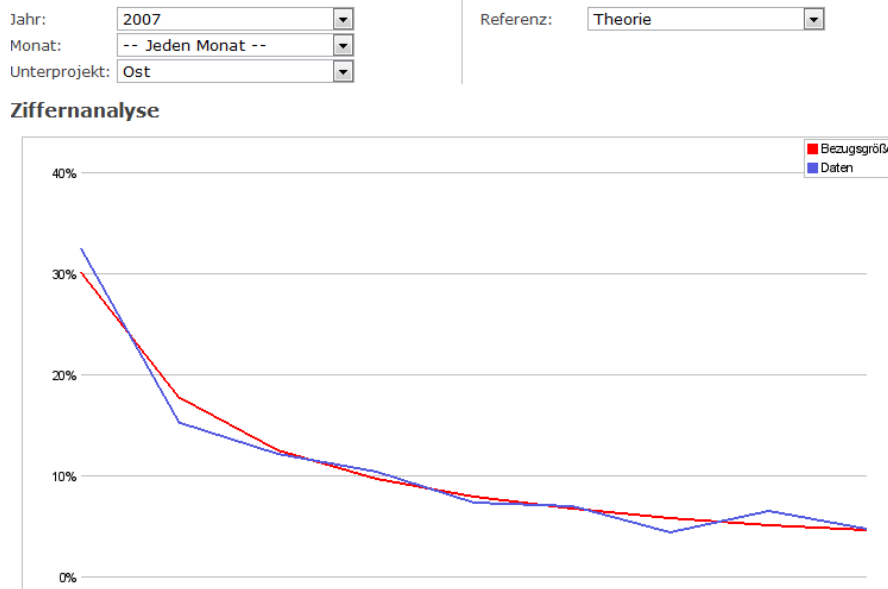
Analyse eines Teildatenbestands

bei der Gesamtbetrachtung der Daten wurde deutlich, dass die Verteilung der Ziffern eine leichte Anomalie aufweist, indem die 5 etwas zu häufig auftritt. Ist dieses in allen Perioden der Fall? Um das herauszufinden wählen wir in dem Feld **Jahr** die Periode **2007** aus. Die Abweichung bei der 5 ist hier weniger stark ausgeprägt als dies im Gesamt-Datenbestand der Fall ist.





Danach wollen wir herausfinden, ob sich alle Unternehmensteile in 2007 gleich verhalten. Dies steuern wir über die Auswahl im Feld **Unterprojekt**. Als Ergebnis erhalten wir die Information, dass im Unternehmensteil **Ost** die Ziffernanalyse Ergebnisse bringt, die nach der Theorie zu erwarten sind. Im Unternehmensteil **West** beobachten wir dagegen besonders starke Abweichungen.



Auf diese Art können wir beliebige Teile des Datenbestands analysieren und so herausfinden, ob beobachtete Anomalien sich besonders stark in einzelnen Perioden oder Unternehmensteilen häufen. Hier müsste dann gegebenenfalls eine Analyse der einzelnen Geschäftsvorfälle vorgenommen werden.

Vergleich von Teildatenbeständen

bisher haben wir als Referenzgröße ausschließlich die Verteilungswerte der Theorie verwendet. Unser System bietet auch die Möglichkeit Teildatenbestände direkt miteinander zu vergleichen und so interne Anomalien zu finden. Das ist insbesondere dann sinnvoll, wenn durch Verzerrungen (etwa Preisgestaltung s.o.) der Datenbestand insgesamt eine andere Verteilung aufweist, als dies die Theorie erwarten lässt. Durch den internen Vergleich haben wir jetzt die Möglichkeit herauszufinden, wo besonders starke Abweichungen vorliegen. Dabei kann zum einen der Gesamt-Datenbestand als Referenzgröße dienen:



Daten

Jahr: 2007
Monat: -- Jeden Monat --
Unterprojekt: Ost

Bezugsgröße

Referenz: Auswahl
Jahr: -- Alle Jahre --
Unterprojekt: -- Alle Unterprojekte --

darüber hinaus können auch bei der Bezugsgröße andere Perioden und/oder andere Unternehmensteile ausgewählt werden. In der Regel ist es allerdings sinnvoll, sich hier schrittweise vorzuarbeiten und beispielsweise zuerst einen Zeitvergleich vorzunehmen und diesen dann noch durch die Auswahl von **Unterprojekt** zu verfeinern.

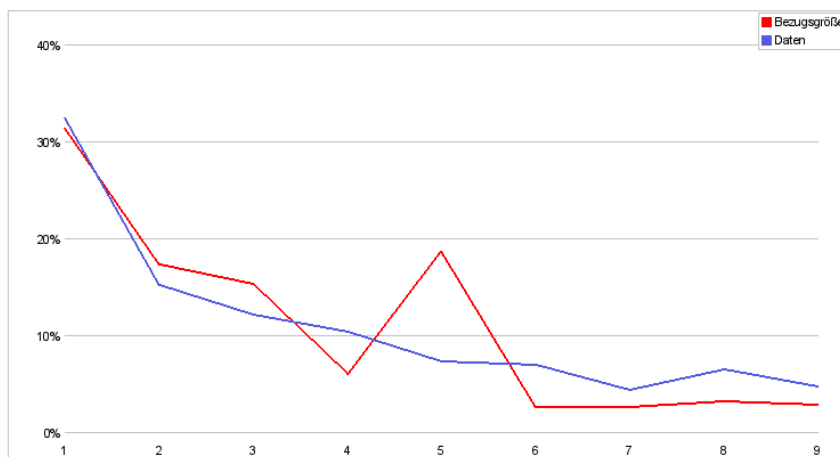
Daten

Jahr: 2007
Monat: -- Jeden Monat --
Unterprojekt: Ost

Bezugsgröße

Referenz: Auswahl
Jahr: 2008
Monat: -- Jeden Monat --
Unterprojekt: West

Ziffernanalyse



Zum Abschluss ihrer Sitzung melden Sie sich ab (**Abmelden**)

Sie haben nach einer erneuten Anmeldung unter Ihrer eigenen E-Mail-Adresse dann die Möglichkeit einen echten Datenbestand hochzuladen und in der gezeigten Form zu analysieren. Nach dem Importvorgang erhalten Sie umgehend eine Mail, die Ihnen auf Basis eines Chi-Quadrattests die Information gibt, ob Ihr Gesamt Datenbestand einen ersten Validitätstest bestanden hat.

Viele Spaß bei der Analyse Ihrer Datenbestände!

Büro Göttingen

Henri-Dunant Str. 36 • D-37075 Göttingen
Tel. (0551) 230 60 • Fax. (0551) 5083 9686
Steuernummer: 20/118/08736

Büro Hannover

Siegesstr. 4 • D-30175 Hannover
Tel. (0511) 5106176
USt-Id: DE 208852267

Bankverbindung

Sparkasse Göttingen
BLZ 26050001
Kto. 56016272