

Leistungsmerkmale und Möglichkeiten zeitgemäßer Tabellierung

Durchblick im Datenmeer

Tabellen gelten gemeinhin als wenig spektakulär und doch haben sie entscheidenden Einfluss: Sie dienen der Identifikation jener Ergebnisse, die in der Präsentation glänzen dürfen und so letztlich umsetzungsrelevant werden. **Knut Holzcheck** schildert, was gute Tabellierung heute leisten kann.

Sex-Appeal sagt man der Tabelle wohl kaum nach. Eher schon fristet sie in der Marktforschung ein Dasein als Mauerblümchen an einem zentralen Ort: Sie ist ja nur ein Zwischenergebnis, Basis der mühsamen Kärrnerarbeit, aus einer Studie die interessanten Aspek-

te herauszufiltern. Präsentiert werden nur ausgesuchte Zahlen. Von der Qualität dieses Filterprozesses hängt aber unter anderem die Qualität der Aussage einer Studie ab. Und deshalb ist es der Mühe wert, Tabellen zu verbessern – das war jedenfalls die Überzeugung bei der Gründung von

Gess Software im Jahre 1991. Und die wichtigsten Aspekte bei der Optimierung sind damals wie heute Lesbarkeit und arbeitsökonomische Effizienz.

Lesbarkeit

Die Lesbarkeit von Tabellen ist ein ziemlich komplexes Konstrukt. Was genau macht eine Tabelle gut oder schlecht lesbar? Wir setzen voraus, dass die Tabelle inhaltlich vernünftig ist. (Eine inhaltlich unsinnig aufgebaute Tabelle kann man durch keine Gestaltung der Welt wieder lesbar machen.) Daneben bleibt aber ein Rest, der mit Ästhetik zusammenhängt. Gut lesbare Tabellen sind meist nicht hässlich. Nun geht es hier nicht um einen Schönheitswettbewerb. Aber die gute Gestaltung einer Tabelle kann dazu dienen, dass der Leser die dargestellten Sachverhalte möglichst mühelos aufnehmen kann. Also: Was zeigt die Tabelle, sind es Prozentzahlen? Welche Zahlen gehören zu welchen Texten? Das kann das Auge am besten wahrnehmen, wenn die Abstände stimmen: nicht zu weit, nicht zu eng. Ist der Text gut lesbar umbrochen? Wird das Auge durch Linien geführt?

R Beispiel: bemerkenswerte Zahlen farblich hervorgehoben

Es gibt Tabellen, bei denen einzelne Werte besonders viel Gewicht haben. Wenn etwa eine Wohnungsbaugenossenschaft ihre Mieter nach deren Zufriedenheit befragt, dann sind hohe Prozentwerte bei negativer Bewertung ein Alarmsignal. Mit GESS tabs kann man auf solche Werte mit Farbe hinweisen. Im Beispiel werden Prozentwerte von 10 oder mehr für die Note „Ungenügend“ in Rot auf gelbem Hintergrund dargestellt.

Zusätzlich wird der Notendurchschnitt ausgegeben. Die Werte in allen Spalten werden gegeneinander mit dem unabhängigen t-Test getestet. Das Ergebnis wird mit Buchstaben zur Spaltenkennzeichnung dargestellt. Auf Wunsch kann man auch hier zur einfacheren Orientierung die Zellen, in denen ein signifikanter Unterschied auftritt, farblich kennzeichnen lassen.

Hervorheben durch Farbe und integrierter Signifikanztest (t-Test)					
Wenn Sie einmal an Ihre Wohnung in unserer Wohnungsgesellschaft denken: Wie beurteilen Sie die Qualität von					
	Hausmeister	Abrechnung	Grünanlage	Dienste	Reinigung
Col %	A	B	C	D	E
N	679	679	679	679	679
(1) sehr gut	10 %	12 %	11 %	11 %	11 %
(2) gut	25 %	22 %	20 %	23 %	25 %
(3) befriedigend	26 %	26 %	26 %	20 %	23 %
(4) ausreichend	23 %	21 %	18 %	21 %	22 %
(5) mangelhaft	12 %	10 %	12 %	9 %	13 %
(6) ungenügend	5 %	8 %	13 %	15 %	5 %
Notendurchschnitt	3.2	3.2	3,4 ABE	3,4 ABE	3,2

In der Kategorie "(6) ungenügend" werden Prozentwerte von 10% und mehr in Rot auf gelbem Hintergrund dargestellt.

Signifikant abweichende Notendurchschnitte werden blau hinterlegt
a .. z = 5%-Niveau, A .. Z = 1%-Niveau.

GESS mbH

Typografie

Eine gute Typografie ist in erster Linie dazu da, die Orientierung zu erleichtern, dem Auge die Inhalte einer Tabelle nahezubringen, etwa: Was ist der Fragentext? Ist die Tabelle gefiltert? Welche Antwortkategorien gibt es? Sind statistische Tests integriert? Welche Signifikanzniveaus gelten?

Für den Leser ist es hilfreich, wenn er hierbei einer unaufdringlichen typografischen Hilfestellung vertrauen kann. Der Name der Studie und vielleicht das Kapitel der Auswertung sind in großer fetter Schrift immer an der gleichen Stelle. Der Fragentext ist vollständig in einer gut lesbaren kleineren Schrift dokumentiert. Und durch die Wiederholung des formal Gleichen lernt der Leser schnell und ohne Mühe, sich auf allen Seiten zu orientieren.

Manchmal enthält eine Tabellenzelle mehrere Werte; neben einem Prozentwert zum Beispiel auch noch die absolute Zellenbesetzung. Auch hier ist es hilfreich für das Auge, wenn typografisch differenziert wird – Prozentwerte beispielsweise größer, absolute Zahlen klein und unauffällig.

Unterm Strich bleibt: Gute Typografie kann die Lesbarkeit einer Tabelle verbessern. Damit soll keine auffällige Typografie propagiert werden. Eher sind kleine, vorsichtige Anpassungen an den Zweck der Tabelle gefragt.

Grafische Elemente wie Kästen, Linien, Einrückungen oder Hintergrundfarben dienen einem doppelten Zweck: Sie leiten das Auge, und sie bieten Raum für eine unverwechselbare Gestaltung. Am plakativsten ist die Einbindung von Firmenlogos. Aber auch andere Gestaltungselemente können einen hohen Wiedererkennungswert haben. Einer unserer Kunden hat eine fette rote Linie im Tabellenkopf. Genial und unverwechselbar.

Tabellen mit Intelligenz

Integrierte Intelligenz hat viel mit Darstellung und Lesbarkeit zu tun, geht aber noch ein Stück darüber hinaus. Viele Fragestellungen, die Leser von Tabellen im Kopf haben, kann die Software bereits während der Erstellung der Tabellen vorwegnehmen, systematisch prüfen und gegebenenfalls optisch im Tabellenbild kenntlich machen.

Man kann sich das so vorstellen, als hätte die Software bereits den Farbmarker in der Hand, den ein aufmerksamer

Beispiel: interessante Abweichungen farblich hervorgehoben

Die obere Tabelle ist gut lesbar und übersichtlich. Sie ist typografisch gelungen, der Fragentext ist dokumentiert. Sie zeigt Ergebnisse aus der Wahltagsbefragung der ARD bei der Bürgerschaftswahl in Hamburg 2008. Die Tabelle wurde so ähnlich auch am Wahlabend mit GESS tabs erzeugt, wie auch die Tabellen für die 18-Uhr-Prognose der ARD.

Aber es geht übersichtlicher. Unten hat GESS tabs nur eine Kleinigkeit hinzugefügt: eine farbliche Hinterlegung als Resultat des statistischen Tests, ob die Zellenbesetzung signifikant vom Erwartungswert abweicht, den die Randverteilungen definieren. Diese Kleinigkeit bewirkt, dass man sofort sieht, wo die relevanten Abweichungen sind und in welche Richtung sie gehen. Man findet sich in der Tabelle einfach schneller zurecht.

Man sieht auf einen Blick, dass die CDU, die SPD und die Grünen mehr Anhänger unter den Frauen finden; bei der FDP, den Linken und bei den sonstigen Parteien ist dagegen der Anteil bei den männlichen Wählern höher.

Die Wahlerfolge der Grünen und der CDU sind hinsichtlich des Alters der Wähler komplementär: In allen Altersklassen unter 60 Jahren hat die CDU weniger Prozente bekommen als Total, und nur bei den Über-60-Jährigen liegen sie weit darüber. Bei den Grünen ist es exakt umgekehrt. Was sagt uns das über die Hamburger schwarz-grüne Koalitionsregierung?

Hamburg 2008: Wahlverhalten nach Geschlecht und Alter

Wem gaben Sie Ihre Stimme für die LANDESLISTE? (gelber Stimmzettel)

Sp. %	TOTAL	Geschlecht		Altersgruppen				
		männlich	weiblich	18 - 24 Jahre	25 - 34 Jahre	35 - 44 Jahre	45 - 59 Jahre	60 und älter
Basis	38632	17962	19680	2547	5475	7216	8794	13670
CDU	42,2	41,3	43,4	31,5	37,4	37,1	36,6	52,0
SPD	33,8	32,2	35,1	40,7	35,0	33,1	34,4	32,1
Grüne	9,5	8,7	10,3	11,0	12,8	15,1	12,4	3,3
FDP	4,7	5,9	3,5	4,1	5,3	4,4	4,2	5,0
Linke	6,3	7,6	5,0	5,9	5,5	6,5	9,7	4,4
Sonstige	2,6	3,4	1,7	6,2	3,6	3,2	1,8	1,7
Keine Angabe	0,9	0,9	1,0	0,5	0,5	0,6	0,9	1,5

Software by GESS

Hamburg 2008: Wahlverhalten nach Geschlecht und Alter

Wem gaben Sie Ihre Stimme für die LANDESLISTE? (gelber Stimmzettel)

Sp. %	TOTAL	Geschlecht		Altersgruppen				
		männlich	weiblich	18 - 24 Jahre	25 - 34 Jahre	35 - 44 Jahre	45 - 59 Jahre	60 und älter
Basis	38632	17962	19680	2547	5475	7216	8794	13670
CDU	42,2	41,3	43,4	31,5	37,4	37,1	36,6	52,0
SPD	33,8	32,2	35,1	40,7	35,0	33,1	34,4	32,1
Grüne	9,5	8,7	10,3	11,0	12,8	15,1	12,4	3,3
FDP	4,7	5,9	3,5	4,1	5,3	4,4	4,2	5,0
Linke	6,3	7,6	5,0	5,9	5,5	6,5	9,7	4,4
Sonstige	2,6	3,4	1,7	6,2	3,6	3,2	1,8	1,7
Keine Angabe	0,9	0,9	1,0	0,5	0,5	0,6	0,9	1,5

Rote Felder = signifikant unter dem Total, blaue Felder = signifikant über dem Total (0,1%-Niveau)

Software by GESS

Tabellenleser bei der Durchsicht der Tabellen anwenden würde. Drei Beispiele (siehe Kästen) verdeutlichen, wie sich bemerkenswerte Zahlen oder interessante Abweichungen hervorheben und Ergebnisse mit extern zugeführten Benchmark-Werten vergleichen lassen.

Tabellen mit Intelligenz sind Tabellen mit Mehrwert. In der täglichen Praxis enthält nur ein kleiner Anteil von Tabellen solche Ergänzungen. Es ist aber ein schönes Feature, wenn man in passenden Fällen darauf zurückgreifen kann.

Effizienz

Aber eine Tabelliersoftware muss nicht nur klare und intelligente Tabellen produzieren, sondern auch im täglichen Produktionsprozess effizient sein. Einige Beispiele illustrieren das im Folgenden anhand von GESS tabs.

Wenn zentrale Häufigkeitsverteilungen der Grundgesamtheit bekannt sind, soll die Darstellung in den Tabellen oft an diese angepasst werden. Bei GESS tabs reicht es, zur Gewichtung die gewünschten Randverteilungen anzugeben;

dies können beliebig viele sein. Die Software ermittelt dann über eine iterative Anpassung das passende Gewicht für jeden Fall. Selbstverständlich kann man Gewichte auch im Datensatz speichern, sodass sie von anderen Softwaretools, die nicht über dieses Feature verfügen, verarbeitet werden können.

Mehrfachnennungen sind das täglich Brot der Marktforschung. Für alle Formen von Mehrfachnennungen bietet GESS tabs ein einheitliches, einfaches Handling an. Man muss sich beim Tabellieren keine Gedanken mehr machen, ob eine Frage Mehrfachantworten hat. Auch für Overcodes in Fragen gibt es leicht anzuwendende Lösungen.

In Umfragen gibt es Effekte, die auf die Fragenreihenfolge zurückgehen. Dies macht es zum Beispiel erforderlich, den Befragten Produktvarianten in unterschiedlichen Reihenfolgen vorzulegen. Das nennt man Rotation. GESS tabs bietet für die Auswertung solcher Datensätze praktische Werkzeuge an, um diese unterschiedlichen Datenstrukturen möglichst effizient und fehlerfrei „ent-rotiert“ zu präsentieren.

Im Produktionsprozess eines Tabellenbandes spielen offene Fragen eine Sonderrolle, sie müssen im Auswertungsprozess erst kodiert werden. Dies führt in der Regel dazu, dass die Ergebnisse aus den offenen Antworten nicht im normalen Datensatz gespeichert sind. Mit GESS tabs können die offenen Antworten und deren Codes bequem aus separaten Dateien zugeführt werden. Oft werden offene Fragen mit sehr umfangreichen Codeplänen ausgewertet; das ergibt dann Tabellen über mehrere Seiten. Die entsprechende Tabellierlogik kann die Zählergebnisse auch hierarchisch sortiert nach Codes und Overcodes darstellen.

Die Programmierschnittstelle schließlich erlaubt es, mit Macros (und mit Macros, die ihrerseits Macros aufrufen) eigene Prozeduren zu schreiben, die viele Tabellieraufgaben erheblich ökonomischer gestalten. Hierzu trägt auch die flexible Übergabe von Parametern bei.

Flexible Ein- und Ausgabeformate

GESS tabs kann viele verschiedene Datenformate verarbeiten, normale ASCII-Datensätze ebenso wie die verschiedenen Formen von Quantum-Dateien, und Datensätze im CSV-Format ebenso wie dBase-Dateien oder SPSS-SAV-Files. Im Standardfall werden die Tabellen in Postscript gesetzt und lassen sich als PDF-Dateien weiterverarbeiten. Alternativ können die Ergebnisse auch als HTML-Tabellen ausgegeben oder über

R Beispiel: Vergleich mit externen Benchmark-Werten

Darstellungen von Veränderungen und statistische Tests beziehen sich in aller Regel auf vergleichbare Daten aus derselben Studie. GESS tabs kann aber Zellen auch nach dem Ergebnis eines Tests einfärben, ob der Prozentwert signifikant von einem extern vorgegebenen Wert abweicht, der als Benchmark dienen soll.

Neuprodukte im Vergleich zur bestehenden Produktpalette

Neue Shampoos im Benchmarktest in sechs Testdimensionen.

	Shampoo X	Shampoo XX	Shampoo RI	Shampoo To	Shampoo F	Shampoo A	Benchmark
Basis 100%	789	787	787	787	787	785	1675
Top Two Box Glanz	59,1 %	59,0 %	57,5 %	60,3 %	60,0 %	50,2 %	57,6 %
Top Box Glanz	33,6 %	33,9 %	33,0 %	35,4 %	33,5 %	27,1 %	27,1 %
Top Two Box Fülle	20,5 %	18,8 %	25,3 %	27,2 %	18,8 %	21,0 %	25,3 %
Top Box Fülle	1,5 %	1,1 %	1,7 %	1,7 %	1,7 %	1,1 %	3,9 %
Top Box Verpackung	38,6 %	27,4 %	32,0 %	35,9 %	36,3 %	32,7 %	33,9 %
Top Box Qualität	48,1 %	33,9 %	44,1 %	41,3 %	48,8 %	41,1 %	42,3 %

Farben:
Rot = signifikant unterhalb des Benchmarks / Grün = signifikant oberhalb des Benchmarks (5% Niveau)
Blau: Benchmark Werte

Software by GESS

Nehmen wir an, ein Hersteller vermarktet eine Produktpalette, die bereits in vorangegangenen Studien getestet wurde. Wenn diese Palette erweitert werden soll und hierfür weitere Produkte mit denselben Fragen getestet werden, stellt sich natürlich die Frage, wie die neuen Produkte im Vergleich zu den bereits eingeführten bestehen. In diesem (fiktiven) Beispiel wurden die neuen Produkte gegen einen Vergleichswert aus mehreren vorangegangenen Studien getestet, der in der letzten Spalte dokumentiert ist. Man kann die Farben hierbei ebenso frei wählen wie das Signifikanzniveau.

die OLE-Schnittstelle an Excel übergeben werden. Hierbei können die Tabellen auch sofort in Excel-Standard-Graphiken umgesetzt werden.

Gleichzeitig wird eine ganze Reihe von statistischen Tests angeboten, die in die Tabellen integriert werden. Im Laufe der Jahre ist auch die Menge an statistischen Maßzahlen gewachsen: Es gibt mittlerweile mehr als 85, und man kann neben Häufigkeiten und Prozentzahlen auch solche Exoten wie einen Mittelwert unter Ausschluss von beispielsweise fünf Prozent der Extremwerte oder das 75-Prozent-Perzentil ausgeben lassen.

Für viele unserer Kunden ist schließlich vor allem die Integration der verschiedenen GESS-Softwareprodukte wichtig: Die Befragungssoftware (etwa für CATI und Online) und die Erfassungstools sind nahtlos mit GESS tabs verzahnt.

Auch die beste Software kann die Herstellung von Tabellen nicht zu einem primär lustbetonten Handeln machen, aber vielleicht etwas weniger stressig und etwas befriedigender. Ein bisschen Spaß kann beim Tabellieren dann gut aufkommen. Nur Sex-Appeal wäre wohl doch zu viel verlangt. ■



Knut Holzcheck hat die GESS Software GmbH 1991 gegründet und ist Autor von GESS tabs. Er ist Mitgründer von dimap und seit 1997 verantwortlich für die Softwareentwicklung bei der Wahlberichterstattung von Infratest-dimap. Zuvor hat Holzcheck als empirischer Rechtssoziologe und in der Wahlberichterstattung bei infas gearbeitet. www.gessgroup.de