

Es lebe die Tabelle

Zahlen vermitteln dem Betrachter Korrelationen und Unterschiede – wenn er bereit ist, diese zu lesen

Infografiken, Diagramme, Dashboards – Marktforschungsergebnisse werden immer visueller. Auch eine gute Tabelle kann helfen, die Zahlen sprechen zu lassen. Sie ist gleichzeitig ein Analyseinstrument. Moderne Tabellierungs-Software sorgt somit nicht nur für Schnelligkeit, Effizienz und Transparenz, sondern fördert Erkenntnisse, weiß **Uwe Stüwe** von Gess.

Einst war sie das Herz der Marktforschung, heute sehen viele sie als ein Relikt vergangener Zeiten: die Tabelle. Sie liefert Ergebnisse in reiner Form, indem sie die Zahlen sprechen lässt. Doch das scheint heute nicht mehr ausreichend zu sein. Aus dem Drang zur Visualisierung wurde ein Zwang: Säulen, Kurven, Torten – Hauptsache eine Grafik verpackt die nackten Zahlen. Dabei wird das Visuelle immer komplexer: bunt, interaktiv, animiert, vielschichtig. Manche Datenvisualisierungen erinnern an Kunstwerke, viele eher an Tapetenmuster. In der Flut der Infografiken und Dashboards herrscht ein optisches Rauschen, in dem die eigentliche Sache – das relevante Ergebnis – unterzugehen scheint.

Das war nicht immer so: Das populärste Standardwerk der Meinungsforschung war in den 60er Jahren das Taschenbuch „Umfragen in der Massengesellschaft“ von Elisabeth Noelle-Neumann. Dort findet man keine Grafiken, die häufigste Darstellungsform ist die Tabelle. Dabei werden die dargestellten Variablen auf zwei Ausprägungsdimensionen verdichtet. Zahlen vermitteln dem Betrachter Korrelationen und Unterschiede – wenn er bereit ist, diese zu lesen. Es scheint, dass man diese Fähigkeit den heutigen Leserinnen und Lesern nicht zutraut. Das ist schade, denn Zahlen sind die Bausteine der Interpretation einer Studie. Zahlen resultieren aus kategorialen Zuordnungen der beobachteten und gemessenen Wirklichkeit und sind immer noch die Basis für die nächste, die grafische Abstraktionsebene in Markt- und Sozialforschungstudien. Ohne Zahlen keine Grafik.

Das Daten-Tinte-Verhältnis sollte man beachten

Ohne Zweifel gibt es Grafiken, welche die Erkenntnis fördern. Über Präsentationen, Charts, Datenjournalismus und Informationsdesign wurde schon viel geschrieben. Instrukтив sind die Bücher „Wie aus Zahlen Bilder werden“ von Gerald Zelazny (immer wieder neu aufgelegt) und „Die Sprache der Zahlen“ von Hans Zeisel (leider nur

1

Vorher: Die reinen Zahlen

Hamburg 2008: Wahlverhalten nach Geschlecht und Alter

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

Sp. %	Geschlecht			Altersgruppen				
	TOTAL	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

Wir sehen, dass wir nichts sehen.

2

Nachher: Farben ermöglichen eine Analyse

Hamburg 2008: Wahlverhalten nach Geschlecht und Alter

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

Sp. %	Geschlecht			Altersgruppen				
	TOTAL	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)

noch antiquarisch erhältlich) – in beiden wird übrigens die Tabelle als angemessene Darbietungsform neben all den Grafiken gewürdigt. Den umfassendsten Blick auf das Thema Datenvisualisierung hat allerdings der Grafikdesigner und Vater des modernen Informationsdesigns Edward Tufte. Seine Bücher sind informativ, inspirierend und erhellend, allerdings ein bisschen ambivalent, was die Rolle der Tabelle betrifft. Dadurch, dass Tufte viele historische Beispiele für raffinierte bildliche Darstellung liefert, regte er auch die ausufernden Visualisierungen heutiger (Grafik-) Datenkünstler an. Auf der anderen Seite propagiert er eine Strenge, die auf jedes unnötige Zierwerk verzichtet.

Dabei richtete er den Blick auf die „Daten-Tinte“. Sie bezeichnet eine Maßzahl für die notwendige Tintenmenge, um reine (gezählte) Informationen sichtbar zu Papier zu bringen, im Verhältnis zur „Gesamt-Tinte“, also inklusive des „schmückenden Beiwerks“. Das Verhältnis kann zwischen 0 und 1 pendeln, wobei mit 1 eine

Der Autor



FOTO: JAROCKA/LOEBENANN

Uwe Stüve ist geschäftsführender Gesellschafter der GESS Phone & Field Marktforschung GmbH und GESS Gesellschaft für Software in der Sozialforschung mbH. Zuvor war Uwe Stüve zwölf Jahre im Mediageschäft bei internationalen Agenturen tätig, davon mehrere Jahre in der Türkei. Seine Berufslaufbahn begann der Diplom-Soziologe nach seinem Studium bei infas / MediaMetric in Hamburg.

u.stueve@gessgroup.de

reine Datendarstellung erzielt wird und man bei 0 lediglich Ornament ohne Gehalt präsentiert. Vor einigen Jahren waren im Internet sogenannte „Infografiken“ inflationär, bei denen Studienergebnisse im coolen Design illustriert wurden – etwa indem der Kaffeeverbrauch mit Kaffeebohnen-Säcken oder dampfenden Tassen veranschaulicht wurde. Hier überwogen Illustrationen, die eher von den Zahlen ablenkten – zu viel Schmuck, zu wenig Zahlen-Substanz.

Eine andere bedenkliche Entwicklung sind Dashboards – da hier die Zahlen automatisch in die vorgefertigten Grafiken und Visualisierungsformen fließen, sind sie nicht immer passend für die angezeigten Werte. Die animierte Präsentation solcher Grafiken in der digitalen Welt – sie bauen sich regelrecht vor unseren Augen auf – suggeriert eine Dynamik, die mit den Inhalten nichts zu tun hat. Nicht selten wird die Aufmerksamkeit auf jene Daten gelenkt, die sich oft ändern – was nicht mit Relevanz gleichzusetzen ist.

Neuprodukte im Vergleich zur bestehenden Produktpalette

Neue Shampoos im Benchmarktest in sechs Testdimensionen.

	Shampoo X	Shampoo XX	Shampoo RI	Shampoo T6	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

Farbliche Hervorhebungen ermöglichen Vergleich mit Benchmarks

Hier besteht die Gefahr eines unausgewogenen Daten-Tinte-Verhältnisses – auch wenn auf den digitalen Screens gar keine Tinte mehr verwendet wird.

Die Stärken der Tabelle herausstellen

Dashboards, Grafiken und manchmal sogar Illustrationen mögen ihren Platz in der Marktforschung haben. Heute wollen wir uns aber dem Stiefkind der Datenvisualisierung widmen – der Tabelle. Ihre Stärken sind:

- Sie fokussiert auf das Wichtigste – die (gezählten) Beobachtungen/Messungen.
- Sie ist keine Illustration, sondern ein Analyse-Instrument mit direktem Bezug zur Zählung.

- Sie legt auch die Produktionsweise offen – indem sie zum Beispiel Fallzahlen oder Gewichtungsfaktoren angibt.

- Sie kann wichtige ergänzende Informationen integrieren – etwa Signifikanztests, die Unterschiede und Zusammenhänge mit statistischer Sicherheit erst interpretierbar machen.

Tabellen sind das Handwerkszeug des Markt- und Sozialforschers, sie sollten der Ausgangspunkt jeder Analyse und Interpretation sein. Sicherlich müssen Ergebnisse für unterschiedliche Publikumssegmente verschieden präsentiert werden. Wenn Experten miteinander kommunizieren, sind Tabellen dafür eine effiziente und transparente Art und Weise, vorausgesetzt, sie sind gut gemacht. Edward Tufte's Konzept der „Daten-Tin-

te“ kann hier wieder helfen, gute Tabellen zu produzieren.

Denn: Tabellen mit redundanten Informationsgehalten, einem Daten-Tinte-Verhältnis weit unter 1, dienen nicht der Lesbarkeit. Der Kern des Informationsgehaltes der Daten muss dem Betrachter leicht und offen entgegentreten. Diese Darstellungsform der gesammelten Daten ist im Marktforschungsprozess gerade in der Zwischenphase wichtig, bevor dem Auftraggeber die Ergebnisse und Interpretationen geliefert werden. Hier erhält die Datenquelle erstmals lesbare Formen, Entitäten werden gruppiert, nehmen Gestalt an. Der Hypothesenbildung folgt so ein erster Interpretationsprozess, gegebenenfalls ergänzt durch multivariate Analysen. In dieser Phase sind schlichte Datentabellen hilf-

reich. Mit den angemessenen statistischen Maßzahlen angereichert ermöglichen sie den Blick auf Neues und das Wesentliche.

Nützliche Software für Tabellen

Hier wird der Nutzen einer guten Tabellierungs-Software sichtbar. Der Anwender muss alle erforderlichen Darstellungen bestimmen können, um die Datenanalyse zu ermöglichen: Aufbaudimensionen, gewünschte Maßzahlen, flexibel notwendige Parameter und statistische Tests unterstützen die Les- und Interpretierbarkeit der Daten. So produzierte Tabellen enthalten alle relevanten Werte oder führen auffäl-

weichung vom Erwartungswert. Sofort wird der Blick des Forschers auf die Auffälligkeiten gelenkt. Jetzt sind Interpretationen sogar mit einer entsprechenden statistischen Sicherheit formulierbar.

Hervorhebungen und Benchmarks helfen

Wie dargestellt, unterstützt eine smarte Tabellierungs-Software Forschende bei systematischen Analysen und liefert auf gezielte Fragen korrekte und gut lesbare Antworten gleich mit. Während der Tabellierung können aber nicht nur statistische Tests durchgeführt und als optische Ergänzung im Tabellenbild kenntlich gemacht werden,

Tabellierungs-Software können Gewichtungen schnell und effizient durchgeführt werden. Die auftretenden Stichprobenfehler erfordern eine Anpassung an die Grundgesamtheit. In unserem Beispiel reicht es, die gewünschten Randverteilungen anzugeben. Das Programm ermittelt mit dieser Vorgabe über eine iterative faktorielle Anpassung die notwendigen Gewichte für jeden Fall. Selbstverständlich kann man diese Gewichte im Datensatz speichern, so dass sie auch von anderen Softwaretools, die nicht über dieses Feature verfügen, verarbeitet werden können (Abbildung 4).

Die Gewichtungsvorgaben von 50/50 bei dem Merkmal Geschlecht für alle Altersklas-

Col % Phys%C PhysDC	N	Alter			
		18-29 Jahre	30-49 Jahre	50-64 Jahre	65 und älter
N	1007	201	201	201	403
Geschlecht					
männlich	403	50 % 44 6	50 % 45 5	50 % 53 -3	25 % 43 -18
weiblich	604	50 % 56 -6	50 % 55 -5	50 % 48 2	75 % 57 18

4

Gewichtungseffekte
einfach sichtbar machen

lige, signifikante Werte mit Farbe oder intelligent unterstützenden Symbolen vor Augen. Unser erstes Beispiel, ein Vergleich von zwei Tabellen mit gleichen Werten, verdeutlicht diese Unterschiede, dabei haben wir unsere Tabellierungs-Software GESStabs verwendet.

Schauen wir auf eine typische Tabelle mit älteren Daten einer Wahl zur Hamburger Bürgerschaft (Abbildung 1). Sie ist schon recht vorbildlich, da gut lesbar und übersichtlich. Der spaltenweise Vergleich der Prozentzahlen nach Geschlecht und Alter führt mit einiger Akribie zu Auffälligkeiten, welche dann vertieft betrachtet werden könnten.

Mit etwas Ausdauer lassen sich weitere Auffälligkeiten finden. Aber schon bald beginnen Zweifel über die Aussagekraft gefundener Unterschiede. Ein Signifikanztest wird benötigt. Warum mit diesem nicht schlicht die signifikanten Tabellenzellen farblich kennzeichnen? Das haben wir in der zweiten Tabelle durchgeführt – in der verwendeten Tabellierungs-Software können Signifikanztests leicht ausgewählt und berechnet werden (Abbildung 2). In dieser Tabelle ist ein dreistufiger Signifikanztest hinterlegt (auf 5 Prozent, 1 Prozent und 0,1-Prozent-Niveau), je kräftiger die Farbe, umso sicherer ist dies eine signifikante Ab-

ebenso lassen sich Vergleiche darstellen.

Wie hilfreich die Hervorhebung von Werten ist, soll ein kurzes Beispiel verdeutlichen: Angenommen, eine Überschreitung der Grenze von 10 Prozent in der Kategorie „Unzufrieden“ soll dem Kunden Handlungsbedarf signalisieren, dann kann es hilfreich sein, wenn solche Zahlenwerte beim Ausdruck der Tabelle farblich gekennzeichnet werden. Man kann z. B. alle kritischen Zellen farblich hinterlegen oder die Zahlenwerte in einer auffälligen Farbe drucken oder anzeigen. Eine verwandte Fragestellung ist der Vergleich mit Werten, die für den Kunden wichtig sind, die aber nicht in der tabellierten Studie erhoben wurden. Hier sollte es möglich sein, fix hinterlegte Durchschnittswerte mit den Ergebnissen einer aktuell zu analysierenden Studie zu vergleichen. In unserem Beispiel kann man mit einem frei wählbaren Farbcode die Performance anzeigen (siehe Abbildung 3).

Gut proportioniert: Die Gewichtung

Bei vielen repräsentativen Umfragen sind Gewichtungen der Stichprobe notwendig. Eine besondere Herausforderung sind Wahlprognosen, denn alles muss schnell gehen, aber verlässlich sein. Mit einer guten

sen wird in dieser Tabelle dokumentiert. Sehr gut lassen sich zudem die Veränderungen der Rohdaten und der gewichteten Daten in dieser Tabelle aufzeigen. In der Altersklasse 18 bis 29 Jahre erhielten die Männer ein um 6 Prozent höheres Gewicht, denn in den Rohdaten waren nur 44 Prozent Männer in dieser Altersklasse vorhanden. Aber auch komplexere, verschachtelte Gewichtungsvorgaben lassen sich bequem vorgeben und durchführen.

Die Lesbarkeit von Tabellen ist, wenn man anfängt, darüber nachzudenken, ein ziemlich komplexes Konstrukt. Was genau macht eine Tabelle gut oder schlecht lesbar? Voraussetzung ist eine inhaltlich vernünftige Tabelle, bei der die Daten methodisch korrekt erhoben und zugeordnet wurden. Denn eine inhaltlich unsinnige Tabelle kann durch keine Gestaltungs-korrektur der Welt wieder lesbar werden. Eine gute Gestaltung einer Tabelle dient jedoch dazu, dem Leser die dargestellten Datenstrukturen, die Sachverhalte, möglichst mühelos zu veranschaulichen. Damit kann sie mehr sein als nur ein Zwischenschritt im Analyse-Prozess, dank moderner Tabellierungs-Software. Wir alle sollten das Lesen von Tabellen wieder zu einer Kardinaltugend der Markt- und Sozialforschung machen. ■