

Gruppierung von Personen zu Marktsegmenten



Florian Keusch

Generalsekretär der Österreichischen
Werbewissenschaftlichen
Gesellschaft, Wien

✉ florian.keusch@wu-wien.ac.at

Mit Hilfe der Clusteranalyse lassen sich Konsumenten oder Marken und Unternehmen, die sich in bestimmten Kriterien besonders ähnlich sind, zu Gruppen zusammenschließen. Folgende Punkte sind zu beachten:

1. Anhand welcher Kriterien lassen sich die Objekte zu Clustern zusammenfassen?
2. Wie vielen Gruppen sollen idealer Weise gebildet werden?
3. Wie lassen sich die einzelnen Segmente gezielt ansprechen?

Foto: ©Kyoshi Takahase Segurido - Fotolia.com

Ziel der Clusteranalyse ist es, ähnliche Objekte (Personen, Marken, Unternehmen) so zu vereinen, dass die Objekte innerhalb einer Gruppe möglichst homogen und die Gruppen untereinander möglichst heterogen sind. Damit ist dieses multivariate Verfahren ein ideales Instrument für die Segmentierung von Märkten.

Fallbeispiel: Einstellung zu Versicherungen

Bei einer vor kurzem durchgeführten Studie zum Thema Versicherungen wurden von Dezember 2006 bis Jänner 2007 insgesamt 369 Personen mit Schwerpunkt Ostösterreich mittels Face-to-Face-Interviews befragt. Unter anderem wurden die Auskunftspersonen gebeten neun verschiedene Aussagen zum Thema Versicherungen anhand einer sechsstufigen Skala von 1 „stimme überhaupt nicht“ zu bis 6 „stimme völlig zu“ zu beurteilen.

Erster Schritt ist die Wahl des richtigen Proximitäts-(Ähnlichkeits-)maßes, das vom vorliegenden Datenniveau abhängt. Da die Daten anhand einer 6-stufigen Ratingskala erhoben wurden, wird metrisches Skalenniveau unterstellt und zur Beurteilung der Ähnlichkeit zweier Elemente ihre räumliche Distanz zueinander erhoben. Eine sehr kleine Distanz zwischen den Elementen weist auf große Ähnlichkeit hin und umgekehrt. Hier wird die Euklidische Metrik als Ähnlichkeitsmaß verwendet, da somit die kürzeste Entfernung zwischen zwei Punkten zueinander beschrieben wird. Die Euklidische Distanz lässt sich mit folgender Formel berechnen:

$$d_{kl} = \sqrt{\sum_{r=1}^R (x_{kr} - x_{lr})^2}$$

mit

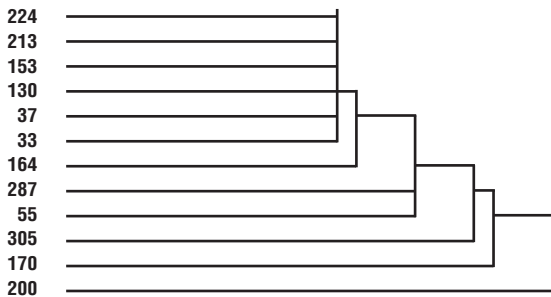
d_{kl} : Distanz der Punkte k, l

x_{kr}, x_{lr} : Koordinaten der Punkte k, l auf der r-ten Dimension ($r=1, 2, \dots, R$)

In einem zweiten Schritt wird ein Fusionierungsalgorithmus zur Bildung von Gruppen mit weitgehend übereinstimmenden Eigenschaften anhand der Ähnlichkeitswerte gewählt. Zuvor wird geprüft, ob statistische Ausreißer, die sich von den übrigen Objekten durch eine vollkommen anders gelagerte Kombination der Merkmalsausprägungen auszeichnen, die Clusterbildung negativ beeinflussen. Zur Identifikation dieser Ausreißer wird der Single-Linkage-Algorithmus (Nächstgelegener Nachbar/Nearest Neighbour) gewählt, der in einem ersten Schritt jene beiden Objekte zu einem Cluster vereint, welche die geringste Distanz zueinander aufweisen. Mit Hilfe eines Dendrogramms werden die Schrittfolgen und ihre jeweiligen Ereignisse graphisch verfolgt. Zur Identifikation von statistischen Ausreißern ist das untere Ende des Dendrogramms von Bedeutung, da hier jene Elemente hinzugefügt werden, die den größten Abstand zu allen anderen Objekten aufweisen. Auf den ersten Blick können in **Abbildung 1** fünf Personen als Ausreißer identifiziert werden (200, 170, 305, 55, 287).

Am Besten zur Auffindung homogener Marktsegmente ist die Ward Methode geeignet. Ziel dieses Verfahrens ist es, Gruppen so zu bilden, dass ein vorgegebenes Heterogenitätsmaß (Varianzkriterium oder Fehlerquadratsumme) durch die Hinzunahme eines weiteren Objekts am wenigsten vergrößert wird. Die Fehlerquadratsumme (quadrierte

Abb. 1: Ausschnitt aus dem Dendrogramm mit Single Linkage Verfahren



Euklidische Distanz) ist am Startpunkt des Clusterungsprozesses gleich Null und nimmt mit jedem Fusionierungsschritt überproportional zu. Für eine Gruppe g wird sie mit folgender Formel berechnet:

$$V_g = \sum_{k=1}^{K_g} \sum_{j=1}^J (x_{k,jg} - \bar{x}_{jg})^2$$

mit

$x_{k,jg}$: Beobachtungswert der Variablen j ($j = 1, \dots, J$) bei Objekt k (für alle Objekte $k = 1, \dots, K_g$ in Gruppe g)

\bar{x}_{jg} : Mittelwert der Beobachtungswerte der Variablen j in Gruppe g

Zur Beantwortung der Frage, wie viele Cluster ausgewählt werden sollen, kann das Elbow-Kriterium herangezogen werden. Dazu wird die Entwicklung der Fehlerquadratsumme in Abhängigkeit von der Clusterzahl in einem Diagramm aufgetragen. Ein überproportionaler Anstieg der Summe der Fehlerquadrate im Übergang zwischen zwei

Clusterlösungen lässt sich durch einen Knick im Graphen ableiten. In diesem Fall ist der Knick bei einer 3-Clusterlösung zu finden, die daher als optimale Lösung gewählt wird). Neben diesem Auswahlkriterium sollten auch immer sachlogische Überlegungen in Bezug auf Clustergröße und Aussagefähigkeit der einzelnen Gruppen angestellt werden (► vgl. Abbildung 2).

Zur Benennung der drei Cluster werden diese, anhand der Mittelwerte bei den neun Statements, in einem Polaritätenprofil verglichen. In acht der neun Statement gibt es statistisch signifikante Unterschiede zwischen den Gruppen, nur die Aussage „Ich würde mich als sportliche(n) AutofahrerIn bezeichnen“ trennt nicht signifikant zwischen den Clustern (Varianzanalyse; $F=1,020$; $p>0,1$) (► vgl. Abbildung 3).

Mit 50% der befragten Personen sind die „Sicherheitsbewussten“ das größte Segment. Im Vergleich zu den anderen Gruppen ist für sie Sicherheit ein wichtigeres Ziel im Leben (MW: 5,16), sie fühlen sich am besten versichert (5,10), sind in Versicherungsfragen am besten informiert (4,23) und halten es für sehr sinnvoll, möglichst viele Versicherungen abzuschließen (3,74). Der Aussage „Für Versicherungen interessiere ich mich nur, wenn ich von anderen darauf aufmerksam gemacht werde“ stimmen sie eher nicht zu (1,99). 26% der Auskunftspersonen fallen in den Cluster der „Risikoaffinen“. Sie fühlen sich verhältnismäßig am schlechtesten versichert (4,35) und finden es weniger sinnvoll, möglichst viele Versicherungen abzuschließen (2,83), ihnen ist Sicherheit aber auch ein weniger wichtiges Ziel im Leben (3,84). Trotzdem zeigen sie großes Interesse für Versicherungen (1,86).

Abb. 2: Entwicklung der Fehlerquadratsumme in Abhängigkeit von der Zahl der gebildeten Cluster (Elbow-Kriterium)

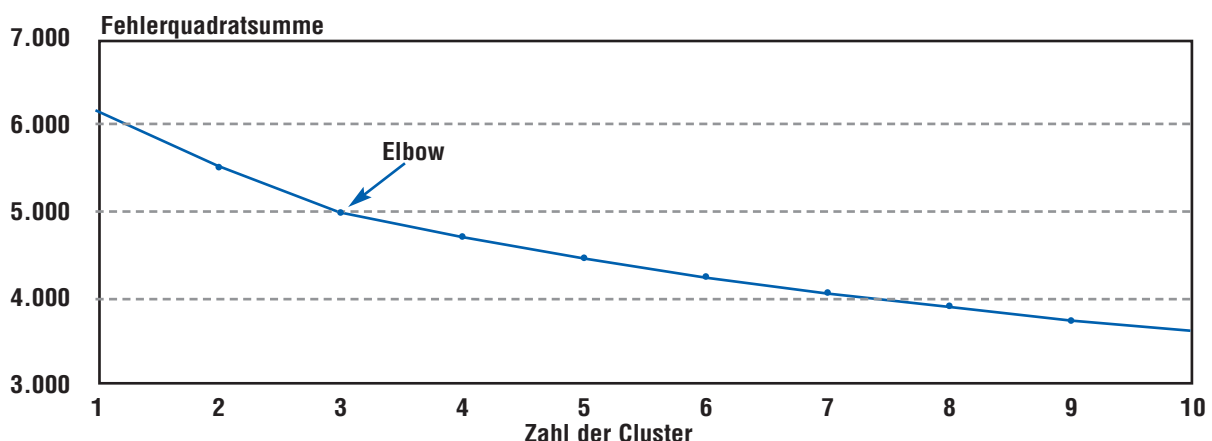
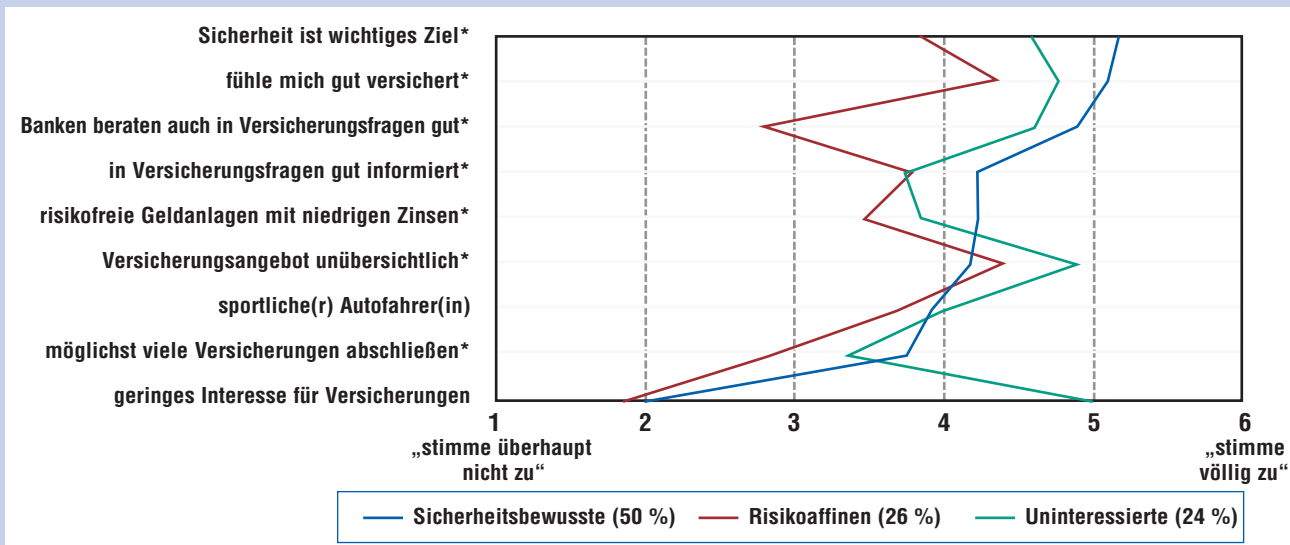


Abb. 3: Drei Marktsegmente, die sich in ihrer Einstellung zu Versicherungen unterscheiden (* sign. Unterschied; $p < 0,05$)

n = 364, Erhebungszeitraum: 12/2006 - 1/2007

Die „Uninteressierten“ sind mit knapp unter einem Viertel der Auskunftspersonen (24%) die kleinste Gruppe. Sie halten das Angebot an Versicherungsleistungen für eher unübersichtlich (4,88) und zeigen das geringste Interesse für Versicherungen (4,99). In den anderen Kriterien liegt diese Gruppe in ihrer durchschnittlichen Einstellung häufig zwischen den beiden anderen Clustern.

Um ausgewählte Segmente mit einer zielgruppenadäquaten Kommunikation ansprechen zu können, lassen sich die einzelnen Gruppen anhand von Unterschieden in z.B. demographischen und psychographischen Variablen, Medianutzung und anderen Verhaltensmerkmalen näher beschreiben. So ist beispielsweise in der Gruppe der „Risikoaffinen“ der Anteil der Männer mit über 66% deutlich höher als in den anderen Gruppen (Chi-Quadrat-Test; Pearson=15,307; $df=2$; $p < 0,001$) und es gibt deutlich mehr Personen mit einem Universitäts- bzw. FH-Abschluss (Chi-Quadrat-Test; Pearson=21,891; $df=6$; $p < 0,001$). Keine

Unterschiede lassen sich zwischen den einzelnen Clustern in Bezug auf das Alter (Varianzanalyse; $F=1,040$; $p > 0,1$), den Familienstand (Chi-Quadrat-Test; Pearson=8,395; $df=6$; $p > 0,1$) und die Haushaltsgröße der Gruppenmitglieder (Varianzanalyse; $F=0,971$; $p > 0,1$) feststellen. Um die Gruppe der „Risikoaffinen“ gezielt anzusprechen, sollte diese Merkmale sowohl bei der Kreation der Kommunikationsmaßnahmen als auch in der Mediaplanung berücksichtigt werden.

Literaturempfehlung:

Eckey, H.-F./Kosfeld, R./Rengers, M. (2002): Multivariate Statistik. Grundlagen – Methoden – Beispiel. Wiesbaden: Gabler.

Fantapé Altobelli, C. (2007): Marktforschung. Methoden – Anwendungen – Praxisbeispiele. Stuttgart: Lucius & Lucius.

Rohr, R./Swietek, P. (in Arbeit): Einflussfaktoren auf den Entscheidungsprozess beim Abschluss von Versicherungen und der Beitrag des Versicherungsmaklers. Diplomarbeit an der Wirtschaftsuniversität Wien.

Besuchen Sie uns im Internet:

www.dwg-online.net

Deutsche Werbewissenschaftliche Gesellschaft