

Statistische Analysen und Studien Nordrhein-Westfalen

Band 51

Impressum

Herausgeber :
Landesamt für Datenverarbeitung
und Statistik Nordrhein-Westfalen

Redaktion:
Bianca Oswald, Hans Lohmann

Preis der gedruckten Ausgabe: 3,00 EUR
Eine kostenlose PDF-Version dieser
Ausgabe finden Sie zum Download
im Publikationsservice (früher Webshop)
des LDS NRW.

Erscheinungsfolge: unregelmäßig

Bestellungen nehmen entgegen:

das Landesamt für Datenverarbeitung
und Statistik NRW,
Postfach 10 11 05,
40002 Düsseldorf,
Mauerstraße 51,
40476 Düsseldorf
Telefon: 0211 9449-2172/3516
Telefax: 0211 442006
Internet: <http://www.lds.nrw.de>
E-Mail: poststelle@lds.nrw.de

sowie der Buchhandel.

Pressestelle:
0211 9449-2521/2518

Zentraler Informationsdienst:
0211 9449-2495/2525

© Landesamt für Datenverarbeitung
und Statistik NRW, Düsseldorf, 2008

Vervielfältigung und Verbreitung,
auch auszugsweise, mit Quellenangabe
gestattet.

Bestell-Nr. Z 08 1 2008 53

ISSN 1619-506X

Inhalt

Das Forschungsdatenzentrum der Statistischen Landesämter – eine Serviceeinrichtung für die Wissenschaft	3
Sylvia Zühlke, Helga Christians und Katharina Cramer Geschäftsstelle des Forschungsdatenzentrums der Statistischen Landesämter	
Behandlungs- und Wohnortkonzept in der Diagnosestatistik	9
Diplom-Volkswirt Hans-Jürgen Treeck	
Clusteranalyse zur Identifizierung von verschiedenen Gruppen von Internetnutzern	17
Christian Hoops, Diplom-Statistiker Josef Schäfer	
Index	26

Zeichenerklärung

(nach DIN 55 301)

- 0 weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
- nichts vorhanden (genau null)
- . Zahlenwert unbekannt oder geheim zu halten
- . . . Zahlenwert lag bei Redaktionsschluss noch nicht vor
- () Aussagewert eingeschränkt, da der Wert Fehler aufweisen kann
- / keine Angabe, da der Zahlenwert nicht sicher genug ist
- x Tabellenfach gesperrt, weil Aussage nicht sinnvoll
- p vorläufige Zahl
- r berichtigte Zahl

Abweichungen in den Summen erklären sich aus dem Runden der Einzelwerte.

Das Forschungsdatenzentrum der Statistischen Landesämter – eine Serviceeinrichtung für die Wissenschaft*)

Sylvia Zühlke, Helga Christians und Katharina Cramer
Geschäftsstelle des Forschungsdatenzentrums der Statistischen Landesämter

Die statistischen Landesämter bauen mit dem Forschungsdatenzentrum der Statistischen Landesämter seit Anfang 2004 auf der Basis der Empfehlungen der „Kommission zur Verbesserung der informationellen Infrastruktur“ gemeinsam mit dem Forschungsdatenzentrum des Statistischen Bundesamtes für die Wissenschaft ein Dienstleistungsangebot der amtlichen Statistik auf. Dieses umfasst ein fachlich breites Mikrodatenangebot und verschiedene Nutzungswege, die der Wissenschaft einen komfortablen Zugang zu amtlichen Mikrodaten ermöglichen. Finanziell unterstützt werden diese Angebote im Rahmen von Projektförderungen durch das Bundesministerium für Bildung und Forschung. Im folgenden Beitrag wird das derzeitige Dienstleistungsangebot des Forschungsdatenzentrums der Statistischen Landesämter dargestellt. Außerdem werden seine weitere Entwicklung und zukünftige Herausforderungen thematisiert.

Einleitung

In ihrem Gutachten „Wege zu einer besseren informationellen Infrastruktur“ aus dem Jahr 2001 legt die „Kommission zur Verbesserung der informationellen Infrastruktur“ dar, dass die Leistungsfähigkeit der Dateninfrastruktur eine entscheidende Grundlage für eine im internationalen Maßstab innovationsfähige sozial- und wirtschaftswissenschaftliche Forschung in Deutschland ist (Kommission zur Verbesserung der informationellen Infrastruktur 2001: 15). Daher müssen Daten – auch solche, die für andere Zwecke im Rahmen staatlichen Handelns entstehen – so effizient wie möglich für wissenschaftliche Analysen genutzt werden können. Für Deutschland wird festgestellt, dass die Dateninfrastruktur diesen Herausforderungen nicht genügt. Um dies zu ändern, wird in dem Gutachten die Einrichtung von Forschungsdatenzentren, die den Zugang zu den Mikrodaten der öffentlichen Datenproduzenten erleichtern sollen, empfohlen.

Die Vorschläge der Kommission wurden im Jahr 2001 vom Gründungsausschuss des Rates für Sozial- und Wirtschaftsdaten aufgegriffen. Der

Rat hat die großen öffentlichen Datenproduzenten – darunter auch die statistischen Ämter – aufgefordert, Forschungsdatenzentren aufzubauen. Die Statistischen Ämter des Bundes und der Länder sind dieser Aufforderung nachgekommen¹⁾: Das Forschungsdatenzentrum des Statistischen Bundesamtes wurde im Oktober 2001, das der Statistischen Landesämter im April 2002 eingerichtet. Das Forschungsdatenzentrum der Statistischen Landesämter ist eine Arbeitsgemeinschaft, an der sich alle statistischen Landesämter beteiligen. Beide Forschungsdatenzentren werden als Projekte vom Bundesministerium für Bildung und Forschung (BMBF) gefördert. Die Projektförderung des Forschungsdatenzentrums der Statistischen Landesämter läuft seit Januar 2004.²⁾

Mithilfe der Projektförderung des BMBF ist es dem Forschungsdatenzentrum der Statistischen Landesämter gelungen, das Dienstleistungsangebot der amtlichen Statistik für die Wissenschaft deutlich zu verbessern und die Nutzung von amtlichen Mikrodaten zu erleichtern. Dies wurde insbesondere durch die Evaluation des Forschungsdatenzentrums der Statistischen Landesämter deutlich,

die im Oktober 2006 im Auftrag des Rates für Sozial- und Wirtschaftsdaten auf der Basis der Kriterien der Leibniz-Gemeinschaft durchgeführt wurde. Die Kommission kam zu dem Ergebnis, dass das Forschungsdatenzentrum der Statistischen Landesämter bereits ein zentraler und unverzichtbarer Bestandteil der informationellen Infrastruktur in Deutschland ist. Sie hat daher mit Nachdruck die weitere Förderung sowie die dauerhafte Etablierung der Einrichtung empfohlen.

Der folgende Abschnitt gibt einen Überblick über das bestehende Datenangebot und seine weitere Entwicklung. Anschließend werden die neuen Entwicklungen im Bereich des Datenzugangs beschrieben und das neue Aufgabengebiet der Wissensvermittlung vorgestellt. Des Weiteren wird die Nutzung des Forschungsdatenzentrums dargestellt und thematisiert, wie die Wissenschaft das Dienstleistungsangebot beurteilt. Abschließend wird ein Ausblick auf die zukünftigen Herausforderungen gegeben.

Datenangebot – verfügbare Statistiken und neue Entwicklungen

Wissenschaftliche Analysen beziehen sich häufig auf mehrere Bundesländer oder das gesamte Bundesgebiet. Solche Auswertungen wurden vor Einrichtung des Forschungsdatenzentrums der Statistischen Landesämter dadurch erschwert, dass die Mikrodaten der einzelnen Bundesländer im Regelfall dezentral nur bei dem jeweils zuständigen statistischen Landesamt gespeichert waren. Sofern länder-

*) Der Aufsatz ist in leicht gekürzter Fassung bereits erschienen in: AStA Wirtschafts- und Sozialstatistisches Archiv Heft 3 – 4/2007. – 1) Die Empfehlung wurde außerdem von der Bundesagentur für Arbeit sowie der Deutschen Rentenversicherung aufgegriffen. – 2) Die Projektförderung des Forschungsdatenzentrums des Statistischen Bundesamtes begann im Jahr 2002. Beide Forschungsdatenzentren befinden sich zurzeit in einer zweiten Förderphase, die nach den derzeitigen Planungen beim Forschungsdatenzentrum des Statistischen Bundesamtes Ende 2008 und beim Forschungsdatenzentrum der Statistischen Landesämter Ende 2009 ausläuft.

übergreifende Auswertungen durchgeführt werden sollten, war eine zeitaufwändige und kostenintensive Zusammenführung der Daten notwendig. Um den Zugang der Wissenschaft zu den Mikrodaten der statistischen Landesämter zu erleichtern und Nutzungsanfragen zeitnah und einheitlich bearbeiten zu können, hat das Forschungsdatenzentrum die Mikrodaten ausgewählter Statistiken länderübergreifend für die Wissenschaft zusammengeführt. Die Auswahl erfolgte auf der Basis einer Befragung potenzieller Nutzer/-innen (Zühlke/Hetke 2003). Im Ergebnis stehen der Wissenschaft heute fast 70 solcher Statistiken zur Verfügung. Das Datenangebot umfasst alle Bereiche der amtlichen Statistik und deckt somit den Informationsbedarf vieler verschiedener Fachdisziplinen ab. Einen Schwerpunkt bilden neben den haushalts- und personenbezogenen Befragungsdaten die wirtschaftsstatistischen Daten. Für die meisten Statistiken liegen mehrere Erhebungsjahre vor, sodass auch Entwicklungen in der Zeit untersucht werden können. Weitere Informationen zum derzeitigen Datenangebot sind unter www.forschungsdatenzentrum.de zu finden. Außerdem werden ausgewählte Datenbestände ausführlich in Kaiser/Wagner (2007), Wirth/Müller (2006) und Zühlke/Christians (2006) unter Berücksichtigung des Analysepotenzials dargestellt.

Für die Mehrheit der verfügbaren Statistiken liegen Daten bisher im Querschnitt vor. Bei den haushalts- und personenbezogenen Mikrodaten hat dies rechtliche Gründe, da eine Zusammenführung im Längsschnitt nur dann erlaubt ist, wenn dies in der einzelgesetzlichen Grundlage vorgesehen ist. Bei den umwelt- und wirtschaftsstatistischen Mikrodaten besteht hingegen die Möglichkeit einer Zusammenführung. Für diesen Bereich liegen daher bereits einige Längsschnittdaten vor. Zu nennen sind in diesem Zusammenhang die Kostenstrukturerhebung, die Krankenhausstatistik, die Umsatzsteuerstatistik, die Produktionserhebung, die Investitionserhebung

und der Monatsbericht im Verarbeitenden Gewerbe.

In der Wissenschaft besteht jedoch nicht nur ein Bedarf an Längsschnittdaten, sondern auch an der Integration von Informationen aus unterschiedlichen Statistiken, da sich hierdurch das Analysepotenzial amtlicher Mikrodaten entscheidend verbessern lässt. Dieser Aufgabe wird sich das Forschungsdatenzentrum der Statistischen Landesämter im Rahmen des Projektes „Amtliche Firmendaten in Deutschland (AFID)“ in den nächsten zwei Jahren widmen. In diesem Projekt werden wirtschafts- und umweltstatistische Mikrodaten systematisch über das Unternehmensregister miteinander und im Längsschnitt zusammengeführt, sodass flexible Auswertungen unter Berücksichtigung mehrerer Statistiken und Erhebungsjahre möglich sein werden.

Neben den Mikrodaten stellt das Forschungsdatenzentrum der Statistischen Landesämter eine umfangreiche Dokumentation des Angebots bereit, die eine sachgerechte Interpretation der Daten unterstützt. Die Informationen beziehen sich auf die Erhebungsziele und -merkmale, den Berichtskreis sowie die Erhebungsbzw. Aufbereitungsmethoden. Ergänzt werden diese Metadaten um Veröffentlichungen zu der jeweiligen Statistik, Datensatzbeschreibungen, Erhebungsbogen und Klassifikationen. Alle Informationen werden zukünftig recherchefähig in einem Metadateninformationssystem abgelegt sein, das das Forschungsdatenzentrum der Statistischen Landesämter gemeinsam mit dem Forschungsdatenzentrum des Statistischen Bundesamtes entwickelt hat. Eine vorläufige Version des Metadaten-systems ist über www.forschungsdatenzentrum.de zugänglich.

Datenzugang – neue Wege

Um die Nutzung der amtlichen Mikrodaten zu erleichtern, haben die Forschungsdatenzentren auf der Basis der bestehenden rechtlichen Regelungen³⁾

³⁾ Zu den rechtlichen Regelungen siehe Zühlke u. a. (2004).

drei Zugangswege eingerichtet: Die Scientific-Use-Files zur Off-Site-Nutzung, die Scientific-Use-Files zur Nutzung an einem Gastwissenschaftlerarbeitsplatz und die kontrollierte Datenfernverarbeitung.

Bei den Scientific-Use-Files für die Off-Site-Nutzung handelt es sich um faktisch anonymisierte Mikrodaten, die an wissenschaftliche Einrichtungen übermittelt werden dürfen. Scientific-Use-Files sind standardisierte Produkte, die einen festgelegten Merkmalsumfang haben. Solche Produkte liegen bereits für eine Reihe von haushalts- und personenbezogenen Statistiken vor. Auch für wirtschaftsstatistische Daten wurden schon erste Datenbestände erstellt. Um die Anonymität der Daten sicherzustellen, müssen bei den Scientific-Use-Files zur Off-Site-Nutzung Vergrößerungen bei der regionalen und/oder bei der fachlichen Gliederung vorgenommen werden. Die Erstellung von Scientific-Use-Files ist somit sehr aufwändig und stets mit einem – zum Teil nicht unerheblichen – Informationsverlust verbunden. Für viele wissenschaftliche Fragestellungen können sie daher unzureichend sein.

Aus diesem Grund werden zwei Wege zur On-Site-Nutzung angeboten, und zwar die Gastwissenschaftlerarbeitsplätze sowie die kontrollierte Datenfernverarbeitung. Die Gastwissenschaftlerarbeitsplätze befinden sich in den geschützten Räumen der statistischen Ämter. Aufgrund der geltenden Sicherheitsvorkehrungen dürfen die dort bereitgestellten Datenbestände mehr Informationen enthalten als die standardisierten Scientific-Use-Files. Die Ausgabe der Ergebnisse können die Wissenschaftler/-innen nur in dem statistischen Amt veranlassen, in dem sie gerade arbeiten. Dort werden die Auswertungen – vor einer Freigabe – auf Geheimhaltung geprüft. Zusätzlich wird im Vorfeld ein Nutzungsvertrag zwischen dem statistischen Amt und den Nutzer/-innen geschlossen und diese auf das Statistikgeheimnis verpflichtet. Beim Arbeiten mit den Daten am Gastwissenschaftlerarbeitsplatz können die Wis-

senschaftler/-innen auf die Unterstützung des Fachpersonals der statistischen Ämter zurückgreifen.

Der zweite Weg, Mikrodaten on-site auszuwerten besteht darin, eine kontrollierte Datenfernverarbeitung durchzuführen. Hierfür erstellen die Wissenschaftler/-innen zu ihrem Forschungsvorhaben ein Auswertungsprogramm. Der durchführende Standort des Forschungsdatenzentrums prüft das Programm und wendet es auf formal anonymisierte Mikrodaten an. Die Ergebnisse werden vor der Weitergabe an die Nutzer/-innen auf Wahrung der Geheimhaltung geprüft. Die Wissenschaftler/-innen haben also keinen direkten Kontakt mit den geheimhaltungsbedürftigen Mikrodaten. Eine weitergehende Anonymisierung der für die Auswertungen genutzten Mikrodaten ist daher nicht notwendig.

Der Datenzugang wurde durch die Einrichtung der Zugangswege deutlich vereinfacht. Von ebenso großer Bedeutung ist die regionale Präsenz des Forschungsdatenzentrums der Statistischen Landesämter. Sie stellt zum einen sicher, dass der Datenzugang an 16 Standorten im gesamten Bundesgebiet möglich ist. Dies erlaubt den meisten Wissenschaftler(inne)n die Arbeit mit den amtlichen Mikrodaten, ohne weite Wege zurücklegen zu müssen. Zum anderen steht in jedem regionalen Standort eine Ansprechpartnerin oder ein Ansprechpartner zur Verfügung, um die Nutzer/-innen bei allen Fragen zur Nutzung von Mikrodaten zu beraten.

Durch den Aufbau von Außenstellen in der unmittelbaren Nähe von wissenschaftlichen Einrichtungen wird die Erreichbarkeit des Forschungsdatenzentrums der Statistischen Landesämter weiter verbessert. Hierzu hat das Amt für Statistik Berlin Brandenburg im Rahmen eines Pilotprojektes einen Standort im Gebäude des Deutschen Instituts für Wirtschaftsforschung in Berlin eingerichtet. In einem weiteren Pilotvorhaben soll der regionale Standort Kamenz in den Räumen der Technischen Universität Dresden untergebracht wer-

den. Hiervon profitieren nicht nur die genannten Einrichtungen, sondern die gesamte ortsansässige Wissenschaft, da beide Standorte durch die örtliche Verlagerung deutlich besser erreichbar sind.

Darüber hinaus wird das Forschungsdatenzentrum der Statistischen Landesämter gemeinsam mit dem Forschungsdatenzentrum des Statistischen Bundesamtes prüfen, ob und wie die Zugangswege weiterentwickelt werden können. Eine Möglichkeit ist hier das sogenannte „echte“ Fernrechnen. Hiermit ist gemeint, dass der Datenzugang von einem speziell eingerichteten Arbeitsplatz in einer wissenschaftlichen Einrichtung über eine sichere Internetverbindung möglich ist. Solche Systeme wurden bereits in anderen europäischen Ländern wie Dänemark oder den Niederlanden etabliert.⁴⁾ Für die amtliche Statistik hat das Forschungsdatenzentrum des Statistischen Bundesamtes einen Prototyp entwickelt, der zurzeit gemeinsam mit dem Forschungsdatenzentrum der Statistischen Landesämter getestet wird.

Wissensvermittlung

Ein weiterer neuer Aufgabenschwerpunkt ist die Wissensvermittlung im Rahmen von Lehrveranstaltungen. Bisher hat das Forschungsdatenzentrum der Statistischen Landesämter dort lediglich sein Dienstleistungsangebot vorgestellt. Hierbei zeigte sich, dass dies allein nicht den Bedarf der Lehrenden abdeckt. Vielmehr besteht die Notwendigkeit, einzelne Erhebungen im Detail vorzustellen und Übungen mit den Datensätzen durchzuführen. Das Forschungsdatenzentrum will daher gemeinsam mit der Wissenschaft den Umgang mit amtlichen Statistiken stärker in der universitären Lehre etablieren sowie Konzepte für die Lehre erarbeiten, die auch von anderen Dozent(innen)

4) Einen Überblick gibt Nauenburg (2006). Im Institut für Arbeitsmarkt- und Berufsforschung fand im August 2007 ein Workshop zum Datenzugang der Wissenschaft in international vergleichender Perspektive statt. Eine Dokumentation des Workshops findet sich unter <http://fdz.iab.de/pageText.asp?PageID=126>.

eingesetzt werden können. Eine wesentliche Voraussetzung für die Wahrnehmung dieser Aufgabe ist der Ausbau des derzeitigen Angebotes an CAMPUS-Files. Bei diesen handelt es sich um vollständig anonymisierte Mikrodaten, die eigens für Lehrzwecke erstellt werden und als Download unentgeltlich im Internet zur Verfügung stehen. CAMPUS-Files sind zurzeit für den Mikrozensus 1998, die Sozialhilfestatistik 1998, die Kostenstrukturerhebung 1999, die Gehalts- und Lohnstrukturerhebung 2001, die Europäische Erhebung zur beruflichen Weiterbildung (CVTS 2) 2000 und die Lohn- und Einkommensteuerstatistik 1998 verfügbar. Wenn CAMPUS-Files für weitere Statistiken und Erhebungsjahre vorliegen, ist es möglich, die Arbeit mit amtlichen Mikrodaten in verschiedenen Studienfächern und unter der Berücksichtigung zeitlicher Veränderungen zu veranschaulichen. Ein Arbeitsschwerpunkt des Forschungsdatenzentrums wird daher künftig auch auf der Erstellung von CAMPUS-Files für weitere Statistiken und Erhebungsjahre liegen.

Nutzung des Forschungsdatenzentrums

Die Nutzung des Forschungsdatenzentrums der Statistischen Landesämter hat seit Projektbeginn kontinuierlich zugenommen. Während im ersten Halbjahr 2004 lediglich 14 Datennutzungsanträge eingingen, waren es im ersten Halbjahr 2007 bereits 63 Anträge. Insgesamt wurden zwischen Januar 2004 und Juni 2007 245 Datennutzungsanträge beim Forschungsdatenzentrum der Statistischen Landesämter gestellt. Die Nutzungsanträge beziehen sich oftmals auf mehrere Statistiken und Erhebungsjahre. Im Durchschnitt werden rund 7 Erhebungen pro Projekt benötigt. Insgesamt wurden im Rahmen der 245 Datennutzungsanträge mehr als 1 800 verschiedene Datenbestände beantragt.

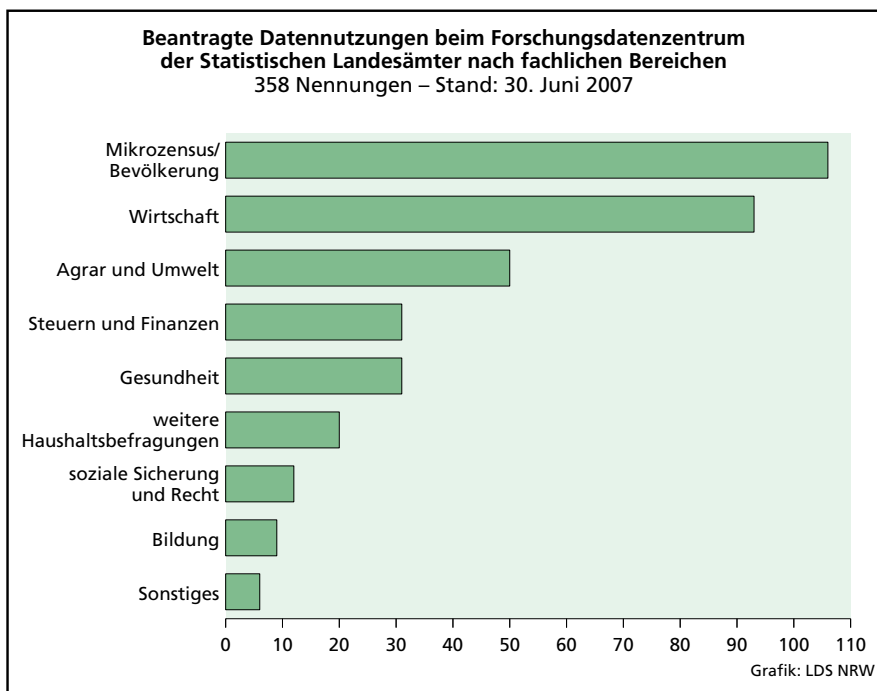
Betrachtet man die Zugangswege, so liegt der Schwerpunkt auf der On-Site-Nutzung: Zwei Drittel aller ausgewählten Nutzungswege entfallen

hierauf.⁵⁾ Dieses Ergebnis ist überraschend, da eine frühere Befragung gezeigt hat, dass seitens der Wissenschaft eindeutig die Off-Site-Nutzung präferiert wird (Zühlke/Hetke 2003). Ein Grund für die intensive On-Site-Nutzung besteht darin, dass nicht für alle Statistiken Off-Site-Files zur Verfügung stehen. Dieser Aspekt erklärt jedoch nicht allein den Umfang der On-Site-Nutzung. Ein weiterer wichtiger Grund liegt in den qualitativen Anforderungen der Forschungsprojekte. So werden für viele Forschungsprojekte, die im Forschungsdatenzentrum der Statistischen Landesämter durchgeführt werden, sehr komplexe Datenbestände aufgebaut, die auf der Basis mehrerer Statistiken und teilweise externer Informationsquellen basieren. Diese Datenbestände werden für einzelne Forschungsvorhaben entwickelt, sodass diese Daten auch zukünftig nur im Bereich der On-Site-Nutzung zur Verfügung stehen werden.⁶⁾

Die Verteilung der Nutzungsanträge auf die fachlichen Bereiche zeigt eine starke Streuung. So wurden für 106 Forschungsprojekte Daten des Mikrozensus und der Bevölkerungsstatistiken beantragt. Für 93 Projekte wurden Wirtschaftsstatistiken und für weitere 50 Projekte Agrar- und Umweltstatistiken nachgefragt. Daten aus den Bereichen „Steuern und Finanzen“ und „Gesundheit“ wurden für jeweils 31 Projekte verwendet. Die Daten aus den weiteren Haushaltsbefragungen wurden in 20 Fällen und die zur Sozialen Sicherung und Rechtspflege in 12 Fällen nachgefragt. Bildungsdaten wurden in acht Projekten ausgewertet. Insgesamt ergibt sich somit, dass das Datenangebot auch in seiner fachlichen Breite genutzt wird.

Zusammen genommen ist festzuhalten, dass durch die regionale Präsenz des Forschungsdatenzentrums der Statistischen Landesämter einerseits

5) Die Anzahl der Zugangswege ist höher als die Anzahl der Forschungsprojekte, da für ein Forschungsprojekt mehrere Zugangswege genutzt werden können. – 6) Beispiele für solche Datennutzungen sind in dem Heft 3 – 4/2007 des AStA Wirtschafts- und Sozialstatistisches Archiv, insbesondere in den Beiträgen von Görzig/Pohl (2007) und Heinbach (2007) dokumentiert.



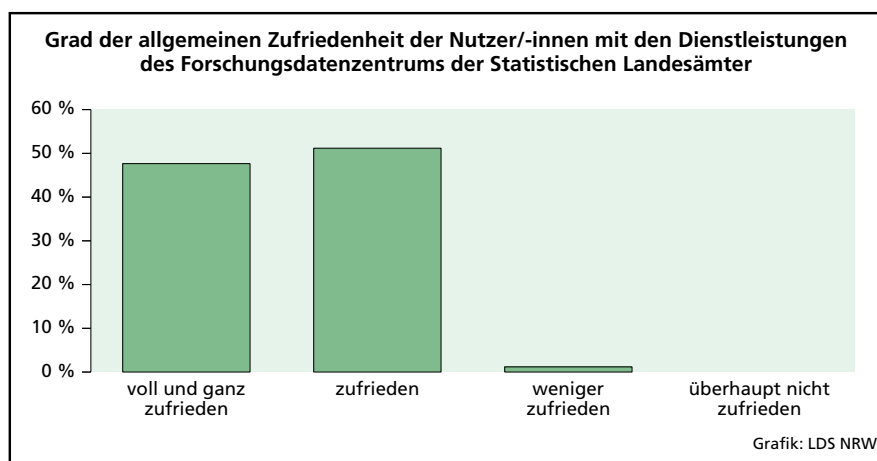
und die fachliche Breite des Datenangebotes andererseits die Nutzung von amtlichen Mikrodaten deutlich zugenommen hat.

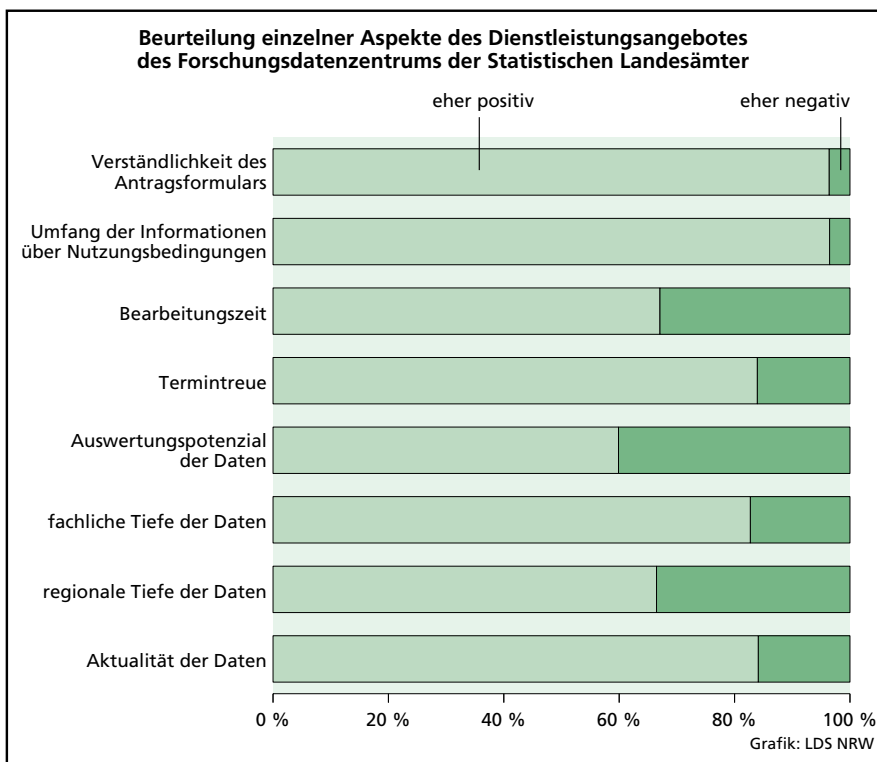
Beurteilung des Dienstleistungsangebotes

Um eine nutzergerechte Weiterentwicklung des Dienstleistungsangebotes zu gewährleisten, befragt das Forschungsdatenzentrum der Statistischen Landesämter regelmäßig seine Nutzer/-innen. Hierfür wurde mit Unterstützung des Sozialwissenschaftlichen Umfragezentrums der Universität Duisburg Anfang 2005 ein Fragebogen zur Messung der Nutzerzufriedenheit entwickelt. Erfragt wird hier insbesondere die Zufriedenheit mit der Beratung, den ausgelieferten Datenbeständen, den

Bearbeitungszeiten sowie der Durchführung der Projekte. Der Fragebogen wird seit Mitte 2005 eingesetzt. Befragt werden diejenigen Wissenschaftler/-innen, die ihr Projekt im Forschungsdatenzentrum durchführen bzw. bereits durchgeführt haben. Bislang wurden in die Befragung 162 Projekte einbezogen (Stand: Herbst 2007). Eine Rückmeldung liegt für 128 Projekte vor, sodass sich eine Rücklaufquote von 79 % ergibt. Bei einigen Projekten haben mehrere beteiligte Wissenschaftler/-innen einen Fragebogen ausgefüllt, sodass die Auswertung auf den Antworten von insgesamt 171 Fällen basiert.

Die Ergebnisse der Nutzerbefragung zeigen, dass die allgemeine Zufriedenheit der Nutzer/-innen mit den Dienstleistungen des Forschungsdatenzentrums hoch ist: So haben 48 %





der Nutzer/-innen angegeben, sie seien voll und ganz zufrieden. Weitere 51 % der Befragten sind zufrieden.

Einzelne Aspekte des Dienstleistungsangebotes werden hingegen differenziert bewertet. Um hierzu die Meinungen der Befragten zu erfassen, wurden zu einzelnen Aspekten, wie z. B. dem Antragsformular, den Nutzungsbedingungen oder den zur Verfügung gestellten Daten, Items formuliert, bei denen die Befragten angeben konnten, ob diese voll und ganz zutreffen, eher zutreffen, eher nicht zutreffen oder überhaupt nicht zutreffen. Diese Items wurden für die folgenden Darstellungen zusammengefasst, um zu verdeutlichen, ob einzelne Aspekte eher positiv oder eher negativ bewertet werden.

Aus der oben stehenden Grafik geht hervor, dass das Antragsformular sowie die Informationen über die Nutzungsbedingungen eher positiv bewertet werden. Hinsichtlich der Bearbeitungszeiten sieht das Bild gemischter aus: Das Item zielte dabei auf die Frage, ob die Bearbeitungszeiten zu zeitlichen Verzögerungen geführt haben. Bei 33 % der Nutzer/-innen traf dies zu, bei den übrigen scheint dies keine Bedeutung gehabt zu haben. Aus dem nächsten Item ist zu erkennen, dass dies nicht an

der fehlenden Termintreue liegt, denn diese wird von 84 % der Befragten positiv beurteilt. In Bezug auf die bereitgestellten Datenbestände zeigt sich, dass für 40 % der Nutzer/-innen das Auswertungspotenzial nicht vollständig zufriedenstellend war. Aus den beiden weiteren Items wird deutlich, dass hierbei der regionalen Gliederungstiefe eine größere Bedeutung zukommt als der fachlichen Gliederungstiefe.⁷⁾ Die Aktualität der Daten wird von 84 % der Befragten überwiegend positiv bewertet.

Das Forschungsdatenzentrum der Statistischen Landesämter hat sich zum Ziel gesetzt, die Wissenschaftler/-innen kompetent bei der Beantragung und der Nutzung amtlicher Mikrodaten zu beraten. Die Ergebnisse der Nutzerbefragung weisen darauf hin, dass dies in der Regel gelingt. Im Vorfeld einer Datennutzung findet zumeist ein allgemeines Beratungsgespräch statt. Für 72 % der Befragten, die ein solches Gespräch in Anspruch genommen haben, war dieses sehr hilfreich, für weitere 25 % war es hilfreich. Ledig-

7) Eine differenzierte Beurteilung des Analysepotenzials amtlicher Mikrodaten wird aus wirtschaftswissenschaftlicher Sicht von Kaiser und Wagner (2007), aus sozialwissenschaftlicher Sicht von Wirth und Müller (2006) vorgenommen.

lich 3 % der Befragten schätzten diese erste Beratung als wenig hilfreich ein. Während der Datennutzung erfolgt meist eine detaillierte fachliche und technische Beratung des Wissenschaftlers bzw. der Wissenschaftlerin. 97 % der befragten Personen fühlen sich durch ihre Ansprechpartner/-innen kompetent beraten. Die Bearbeitungszeit und Erreichbarkeit der Ansprechpartner/-innen werden ebenfalls von fast allen Befragten positiv bewertet.

Die Ergebnisse weisen insgesamt auf eine hohe Zufriedenheit der Nutzer/-innen mit dem Dienstleistungsangebot hin. Änderungsbedarf zeigt sich insbesondere bei den Bearbeitungszeiten und bei der regionalen Gliederungstiefe der Daten.

Zukünftige Herausforderungen

Das Forschungsdatenzentrum der Statistischen Landesämter versteht sich in erster Linie als Serviceeinrichtung für die Wissenschaft. Als solche fühlt es sich verpflichtet, den sich ständig wandelnden Datenbedarf der Wissenschaft aufzugreifen und das Dienstleistungsangebot kontinuierlich unter Berücksichtigung internationaler Erfahrungen weiterzuentwickeln. Dies betrifft sowohl das Datenangebot als auch die Zugangswege. Aktuelle Beispiele sind die Integration von wirtschafts- und umweltstatistischen Einzelangaben und die Verlagerung von Standorten in die unmittelbare Nähe wissenschaftlicher Einrichtungen.

Um eine nutzerorientierte Weiterentwicklung des Datenangebotes zu gewährleisten, wird das Forschungsdatenzentrum der Statistischen Landesämter zukünftig auch selber wissenschaftliche Forschungsprojekte durchführen. Im Vordergrund stehen hierbei zunächst methodische Fragen, die sich im Zusammenhang mit dem Projekt „Amtliche Firmendaten in Deutschland“ ergeben.

In institutioneller Hinsicht steht die Frage der dauerhaften Etablierung im Vordergrund. Zu klären ist insbe-

sondere die Finanzierung eines solchen Dienstleistungsangebotes. Die statistischen Landesämter befinden sich hierzu im Gespräch mit dem Rat für Sozial- und Wirtschaftsdaten, dem Bundesministerium für Bildung und Forschung sowie den zuständigen Landesministerien. Die statistischen Landesämter werden alles daransetzen, um eine dauerhafte Einrichtung des Forschungsdatenzentrums zu ermöglichen. Sie sind hierbei jedoch weiterhin auf die Unterstützung der Wissenschaft angewiesen.

Literatur

- Görzig, Bernd; Pohl, Ramona (2007): Diversifizierungsstrategien deutscher Unternehmen. Auswertung eines Producer-Product-Panels der amtlichen Statistik. In: *AStA Wirtschafts- und Sozialstatistisches Archiv*. Heft 3 – 4/2007, S. 179 – 191
- Heinbach, Wolf Dieter (2007): Wages in wage-setting regimes with opening clauses. In: *AStA Wirtschafts- und Sozialstatistisches Archiv*. Heft 3 – 4/2007, S. 233 – 245
- Kaiser, Ulrich; Wagner, Joachim (2007): Neue Möglichkeiten zur Nutzung vertraulicher amtlicher Personen- und Firmendaten. FDZ-Arbeitspapier Nr. 21, Hrsg. von Statistische Ämter des Bundes und der Länder, Wiesbaden
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden 2001
- Nauenburg, Ricarda (2006): Automatisiertes Fernrechnen mit amtlichen Mikrodaten – aktuelle Entwicklungen. In: *Statistische Monatsschrift Berlin*, Heft 05/2006, S. 216 – 223.
- Wirth, Heike; Müller, Walter (2006): Mikrodaten der amtlichen Statistik – Ihr Potenzial in der empirischen Sozialforschung. In: Dieckmann, Andreas (Hrsg.): *Methoden der empirischen Sozialforschung*. Wiesbaden, S. 93 – 127
- Zühlke, Sylvia; Christians, Helga (2006): Datenangebot und Datenzugang im Forschungsdatenzentrum der Statistischen Landesämter. In: Landesamt für Datenverarbeitung und Statistik (Hrsg.): *Statistische Analysen und Studien Nordrhein-Westfalen*. Düsseldorf, LDS NRW, Band 29, S. 3 – 11
- Zühlke, Sylvia; Zwick, Markus; Scharnhorst, Sebastian; Wende, Thomas (2004): The research data centres of the Federal Statistical Office and the statistical offices of the Länder. In: *Schmollers Jahrbuch* 4
- Zühlke, Sylvia; Hetke, Uwe (2003): Datenbedarf und Datenzugang: Ergebnisse der ersten Nutzerbefragung des Forschungsdatenzentrums der Statistischen Landesämter. In: *Allgemeines Statistisches Archiv* 87, S. 321 – 334

Behandlungs- und Wohnortkonzept in der Diagnosestatistik

Diplom-Volkswirt Hans-Jürgen Treeck

Ärztliche Behandlungen werden in freien Praxen oder in Krankenhäusern vorgenommen. Über die Behandlungen der Patientinnen und Patienten in Arztpraxen liegen aus der amtlichen Statistik keine Informationen vor.

Anders bei den Krankenhäusern und den Rehabilitations- oder Vorsorgeeinrichtungen. Für die Patientinnen und Patienten beider Einrichtungsarten liefert die Diagnosestatistik als Teil der amtlichen Krankenhausstatistik Angaben, die eine Untersuchung über den Zusammenhang zwischen dem Behandlungsort und dem Wohnort ermöglichen.

Die nachfolgenden Ausführungen gehen zunächst auf die Grundlagen der Diagnosestatistik und ihre Einbettung in die Krankenhausstatistik ein. Anschließend werden die Zahlen der behandelten Personen in den Krankenhäusern nach dem Behandlungsort und dem Wohnort für verschiedene Gebiets-ebenen der allgemeinen Verwaltung (Regierungsbezirke, kreisfreie Städte und Kreise) in Form von sog. „Behandlungsquoten“ gegenübergestellt. Um die gebietliche Herkunft der Patientinnen und Patienten transparent zu machen, die in den einzelnen kreisfreien Städten und Kreisen behandelt worden sind, werden dann am Beispiel des Regierungsbezirks Köln Einzugsgebietstabellen für die Krankenhäuser herangezogen. Danach wird auf die Einzugsgebiete der in den Regierungsbezirken ansässigen Rehabilitations- oder Vorsorgeeinrichtungen eingegangen. Die abschließende Analyse der Behandlungsquoten für die Gesamtheit der Krankheiten in den Versorgungsgebieten der Krankenhäuser zeigt, dass die Abweichungen der Quoten geringer sind als bei den Gebieten der allgemeinen Verwaltung. Dies bedeutet, dass in den Grenzen der Versorgungsgebiete die Zahl der Patient(inn)en nach dem Behandlungsort und dem Wohnort nicht so stark differiert und somit eine wohnortnahe Versorgung der erkrankten Bevölkerung hier weitgehend sichergestellt ist.

Grundlagen der Diagnosestatistik

Die jährlich durchgeführte Krankenhausstatistik, die als Primärstatistik Angaben sowohl bei den Krankenhäusern wie bei den Rehabilitations- oder Vorsorgeeinrichtungen erhebt, ist in der Krankenhausstatistikverordnung (KHStatV) geregelt. Sie umfasst neben den Diagnosedaten auch die sog. „Grunddaten“ und die Kostendaten der Einrichtungen.

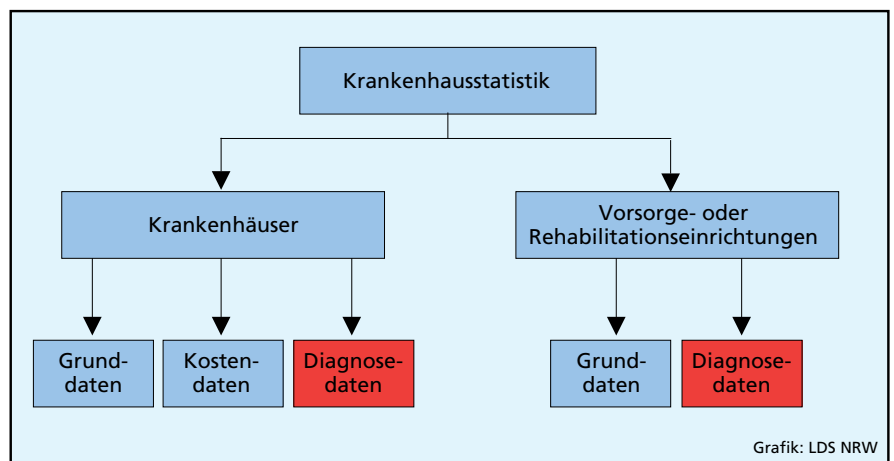
Durch den Erhebungsteil der Grunddaten sollen Grundinformationen über die Strukturen in den Krankenhäusern und Vorsorge- und Rehabilitationseinrichtungen bereitgestellt werden. Schwerpunkte für beide Einrichtungsbereiche sind zum einen die nach Fachabteilungen gegliederten Angaben über die aufgestellten

Betten, die Behandlungstage, die Patientenzu- und -abgänge sowie die Ausstattung mit medizinisch-technischen Großgeräten, zum anderen die Angaben über die Personalausstattung der Einrichtungen. Daraus können Kennzahlen über die Krankenhaushäufigkeit, die durchschnittliche

Verweildauer, die Bettennutzung usw. errechnet werden. Differenzierte Angaben zu ihren Kosten (Personal- und Sachkosten, Zinsen, Steuern usw.) sind nur von den Krankenhäusern mitzuteilen.

Bei den Grunddaten melden die Einrichtungen nur zusammengefasste Angaben für die behandelten Personen, z. B. die Zahl der aus der Behandlung Entlassenen nach den jeweiligen Fachabteilungen. Bei den Diagnosedaten sind demgegenüber Angaben für jede einzelne behandelte Person an das LDS NRW zu übermitteln, sodass für NRW ca. 5 Millionen Datensätze für die Krankenhäuser und Vorsorge- und Rehabilitationseinrichtungen vorliegen. Diese Einzelerfassung ermöglicht auch die Frage nach dem Wohnort der behandelten Person. Da die Anschriften der Einrichtungen, in denen die Behandlungen stattgefunden haben, durch die Erhebung bekannt sind, sind somit die Voraussetzungen für eine Aufbereitung der Ergebnisse sowohl nach dem Behandlungsort wie nach dem Wohnort gegeben.

Liegen Wohn- und Behandlungsort einer behandelten Person in verschiedenen Bundesländern, tauschen die betroffenen statistischen Landesämter die Angaben untereinander aus.



Mit dem Ausland besteht ein solcher Austausch nicht. Behandlungen von Personen in ausländischen Einrichtungen mit Wohnsitz in Deutschland werden den statistischen Ämtern nicht bekannt. Sie sind daher in den Zahlen nach dem Wohnort nicht berücksichtigt. Dagegen werden in deutschen Einrichtungen behandelte ausländische Personen in den Zahlen nach dem Behandlungsort erfasst. Für sie muss in der Diagnosestatistik das Wohnsitzland angegeben werden.

Bei den Angaben zu den aus vollstationärer Behandlung entlassenen Personen handelt es sich allerdings genau genommen um die behandelten Fälle. Eine Person, die innerhalb des Berichtsjahres mehrere Male in einer Einrichtung behandelt worden ist, wird mit jedem Aufenthalt in der Krankenhausstatistik als eigener Fall erfasst. Zur Gesamtzahl der behandelten Fälle gehören auch die verstorbenen Personen, die bis zu ihrem Tode vollstationär behandelt wurden.

Im Einzelnen sind in der Diagnosestatistik von den Krankenhäusern und Rehabilitations- oder Vorsorgeeinrichtungen für jede aus vollstationärer Behandlung entlassene Person folgende Angaben zu machen:

- Postleitzahl des Wohnortes (soweit im Inland wohnhaft)
- Wohnort; bei ständigem Wohnsitz im Ausland Angabe des Landes
- Geburtsdatum
- Zugangsdatum
- Abgangsdatum
- Verweildauer (Anzahl der Berechnungs- und Belegungstage)
- Sterbefall
- Hauptdiagnose nach ICD-Schlüssel
- Operation in Zusammenhang mit der Hauptdiagnose (nur Krankenhäuser)
- Fachabteilung mit der längsten Verweildauer

Als Hauptdiagnose wird die zum Zeitpunkt der Entlassung bekannte Diagnose bezeichnet, die hauptsächlich die Dauer der vollstationären Behandlung beeinflusst hat, bzw. den größten Anteil an medizinischen Leistungen verursacht hat. Sie ist von den Einrichtungen nicht mit ihrer Bezeichnung zu melden, sondern als Schlüssel. Für die Verschlüsselung ist

die Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD) zugrunde zu legen, die von der Weltgesundheitsorganisation (WHO) entwickelt worden ist. Änderungen und Ergänzungen werden etwa alle 10 Jahre vorgenommen. Ab dem Jahr 2000 gilt in der Diagnosestatistik die ICD-10. Wegen der starken Überarbeitung sind ihre Positionen nur in sehr grober Gliederung mit denjenigen der ICD-9 vergleichbar. Die Verschlüsselung ist von den Einrichtungen in der tiefstmöglichen Gliederung (ICD-10-Viersteller) vorzunehmen.

Beim Nachweis der Zahlen sind zum einen die Angaben der einzelnen Einrichtungen schützenswert und zum anderen diejenigen der einzelnen behandelten Personen. Je tiefer die Ergebnisse regional nachgewiesen werden, desto eher besteht unter Zuhilfenahme zusätzlicher Informationen der Statistik, wie die Art des Trägers, die Möglichkeit, die in einem bestimmten Gebiet zu den Daten gehörende Einrichtung zu identifizieren. Um dies zu vermeiden bzw. nicht zu viele Datenfelder geheim halten zu müssen, werden die Ergebnisse nach dem Behandlungsort in der Regel nur bis auf Kreisebene veröffentlicht. Unproblematischer sind die Angaben nach dem Wohnort, da nicht erkennbar ist, ob einzelne Personen in einem Krankenhaus innerhalb oder außerhalb der Region behandelt worden sind.

Die Diagnosedaten nach dem Behandlungsort stehen normalerweise ein Jahr nach Ablauf des Berichtsjahres zur Verfügung. Aufgrund des Länderaustausches der Angaben für die behandelten Personen mit Wohnsitz in einem anderen Bundesland können die Diagnosedaten nach dem Wohnort erst im Frühjahr des zweiten Jahres nach dem Berichtsjahr bereitgestellt werden.

Diagnosedaten gibt es gemäß Krankenhausstatistikverordnung für alle Krankenhäuser ab dem Berichtsjahr 1993. Im Laufe der Zeit wurde für gesundheitspolitische und epidemiologische Fragestellungen die Not-

wendigkeit erkannt, auch über Diagnosemeldungen der in den Vorsorge- und Rehabilitationseinrichtungen behandelten Patientinnen und Patienten zu verfügen. Daher ist die Frage nach den Diagnosen dieser Einrichtungen ab dem Berichtsjahr 2003 in das Erhebungsprogramm aufgenommen worden. Zur Entlastung der kleinen Vorsorge- und Rehabilitationseinrichtungen werden die Diagnosen jedoch nur von Häusern mit mehr als 100 Betten erfragt.

Auswirkungen des Behandlungs- und Wohnortkonzepts auf den Nachweis der Ergebnisse

1. Krankenhäuser

Durch die Möglichkeit, Krankenhäuser in anderen Gebieten als in ihren Wohngebieten aufzusuchen, werden die kranken Personen – je nach zugrunde liegendem Konzept – in unterschiedlichen Gebieten gezählt und beeinflussen damit die Höhe der regional nachgewiesenen Fälle. Die Motivation einer Patientin bzw. eines Patienten, die Gebietsgrenzen zu überschreiten, kann ganz unterschiedlicher Natur sein; z. B. kann sie/er ein Krankenhaus im Nachbargebiet wählen, weil es das nächstgelegene ist, oder sie gehen in ein weiter entferntes Krankenhaus, weil es einen „guten Ruf“ hat, eine Spezialklinik ist oder konfessionell geführt wird.

Setzt man jeweils die Zahl der in einem Gebiet behandelten Patient(inn)en zu der Zahl der Patient(inn)en, die dort ihren Wohnsitz haben, ins Verhältnis, sind die Unterschiede mit einem Blick erkennbar. Diese Relation wird hier als „Behandlungsquote“ bezeichnet.

Aus einer sehr hohen Quote kann man schließen, dass das betreffende Gebiet große überörtliche Bedeutung für die Krankenversorgung hat. Eine niedrige Quote weist darauf hin, dass viele kranke Personen Versorgungsleistungen in Krankenhäusern außerhalb des Gebietes, in dem sie wohnen, in Anspruch nehmen. Bei einer Quote von „100“ sind in ei-

nem Gebiet genauso viele Personen behandelt worden, wie dort gewohnt haben. Allerdings ist es dabei durchaus denkbar, dass ein Teil der kranken Personen von außerhalb zur Behandlung gekommen ist und im selben Umfang Patient(inn)en des Gebiets Krankenhäuser in anderen Gebieten aufgesucht haben.

Die Quoten können für die Gesamtheit der Patient(inn)en in den einzelnen Gebieten ermittelt werden, aber auch für die Patient(inn)en, die mit bestimmten Krankheiten behandelt worden sind.

Durch die Quotenbildung ist allerdings die Anzahl der Patient(inn)en nicht mehr zu erkennen. Es ist daher zu bedenken, dass es bei kleinen Fallzahlen eher zu relativ größeren Abweichungen zwischen den Zahlen nach dem Behandlungs- und dem Wohnortkonzept – und damit zu höheren von 100 abweichenden Quoten – kommen kann als bei hohen Fallzahlen.

Generell kann man bei den Krankenhäusern davon ausgehen, dass die Abweichungen der Patientenzahlen am Behandlungsort von denjenigen am Wohnort umso geringer sind, je größer die betrachtete Region ist. Das hängt damit zusammen, dass

viele Patientenwanderungen über die Grenzen von kleineren Gebieten hinweg hinsichtlich der übergeordneten Region zu innergebietlichen Patientenbewegungen werden, so dass dort Behandlungs- und Wohnregion zusammenfallen (Saldierungseffekt). Auch können sich innerhalb einer Region für einzelne Krankheiten andere Behandlungsquoten ergeben als für alle Krankheiten zusammen.

So zeigt sich bei den Behandlungsquoten für die Regierungsbezirke und das Land, dass sie bei der Gesamtheit der Krankheiten bei 100 oder nah bei 100 liegen, während bei den einzelnen Krankheitsarten in den Regierungsbezirken größere Abweichungen festzustellen sind. Dies trifft vor allem auf „Krankheiten des Auges und der Augenanhangsgebilde“ und „Angeborene Fehlbildungen, Deformitäten und Chromosomenanomalien“ zu. Gleichzeitig ist aufgrund des oben beschriebenen regionalen Saldierungseffekts erkennbar, dass die Behandlungsquoten des Landes bei den einzelnen Krankheitsarten weitgehend 100 betragen, d. h., dass in NRW wohnende Patient(inn)en auch eine ärztliche Behandlung ihrer Krankheit nahezu im eigenen Bundesland erfahren.

Noch viel stärker als bei den Regierungsbezirken kommen die Unterschiede zwischen den Zahlen der Patient(inn)en nach dem Behandlungs- und dem Wohnortkonzept bei den kleineren Gebieten der kreisfreien Städte und Kreise zum Ausdruck.

Die Tabelle 2 zeigt für die kreisfreien Städte Münster und Bonn die höchsten Behandlungsquoten. Dort übersteigt die Anzahl der behandelten Patient(inn)en diejenige der dort wohnenden Patient(inn)en um mehr als das Doppelte. An dritter Stelle steht die kreisfreie Stadt Aachen mit einer Quote von 167. Hauptgrund für die hohen Quoten dieser Städte sind die dort ansässigen Universitätskliniken mit ihren über die Stadtgrenzen hinausgehenden Einzugsgebieten. Ferner ist denkbar, dass ein Teil der in diesen Städten Studierenden bei einem Krankenhausaufenthalt ihre Heimatadresse angegeben hat. Dieser Effekt dürfte aber für die große Differenz zwischen den Zahlen nach dem Behandlungs- und Wohnort nicht ausschlaggebend gewesen sein. Von den zehn Verwaltungsbezirken mit den höchsten Behandlungsquoten sind neun kreisfreie Städte. Dazu gehören auch die Städte Essen und Düsseldorf, die ebenfalls Universitätskliniken haben. Dies macht deutlich, dass vor allem

1. Behandlungsquoten*) in den Krankenhäusern 2005 nach Krankheitsarten						
Krankheitsart	Regierungsbezirk					Nordrhein-Westfalen
	Düsseldorf	Köln	Münster	Detmold	Arnsberg	
Bestimmte infektiöse und parasitäre Krankheiten	102	100	100	101	99	100
Neubildungen	107	102	97	95	100	102
Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems	104	100	97	96	99	100
Endokrine, Ernährungs- und Stoffwechselkrankheiten	107	98	94	107	100	101
Psychische und Verhaltensstörungen	102	93	105	95	101	100
Krankheiten des Nervensystems	101	96	94	105	106	100
Krankheiten des Auges und der Augenanhangsgebilde	115	120	106	57	83	102
Krankheiten des Ohres und des Warzenfortsatzes	106	96	94	95	102	100
Krankheiten des Kreislaufsystems	103	97	94	112	97	100
Krankheiten des Atmungssystems	104	98	97	100	100	100
Krankheiten des Verdauungssystems	101	101	101	97	100	100
Krankheiten der Haut und der Unterhaut	100	98	101	91	103	99
Krankheiten des Muskel-Skelett-Systems und des Bindegewebes	98	97	99	89	110	100
Krankheiten des Urogenitalsystems	100	102	100	102	100	101
Schwangerschaft, Geburt und Wochenbett	101	103	103	97	97	100
Bestimmte Zustände, die ihren Ursprung in der Perinatalperiode haben	103	103	102	98	99	101
Angeborene Fehlbildungen, Deformitäten und Chromosomenanomalien	99	127	97	89	95	103
Verletzungen, Vergiftungen und bestimmte andere Folgen äußerer Ursachen	100	100	98	97	101	100
Insgesamt	102	100	98	99	100	100

*) Behandlungsquote = behandelte Personen am Behandlungsort dividiert durch behandelte Personen am Wohnort x 100

2. Verwaltungsbezirke mit der höchsten/niedrigsten Behandlungsquote*) 2005			
Verwaltungsbezirk (kreisfreie Stadt/Kreis)	Im Verwaltungsbezirk		Behandlungsquote
	behandelte	wohnende	
	Patient(inn)en		
Krfr. Stadt Münster	116 952	53 876	217
Krfr. Stadt Bonn	127 974	61 079	210
Krfr. Stadt Aachen	73 480	44 016	167
Krfr. Stadt Herne	60 135	43 288	139
Kreis Minden-Lübbecke	91 899	67 006	137
Krfr. Stadt Essen	192 204	140 246	137
Krfr. Stadt Hamm	61 987	45 682	136
Krfr. Stadt Bielefeld	98 210	73 599	133
Krfr. Stadt Düsseldorf	165 257	124 515	133
Krfr. Stadt Krefeld	67 128	51 061	131
...
...
Kreis Viersen	47 758	63 538	75
Krfr. Stadt Mülheim an der Ruhr	30 329	40 438	75
Kreis Coesfeld	32 075	45 638	70
Rheinisch-Bergischer Kreis	39 294	56 930	69
Kreis Gütersloh	47 600	69 358	69
Kreis Mettmann	77 056	114 004	68
Kreis Heinsberg	32 553	50 843	64
Kreis Lippe	45 684	72 013	63
Rhein-Erft-Kreis	50 511	91 344	55
Rhein-Sieg-Kreis	59 434	110 952	54

*) Behandlungsquote = behandelte Personen am Behandlungsort dividiert durch behandelte Personen am Wohnort x 100

die kreisfreien Städte eine überörtliche Versorgungsfunktion wahrnehmen.

diejenige der Patient(inn)en, die dort ihren Wohnsitz hatten.

Fast spiegelbildlich zu den kreisfreien Städten mit den hohen Behandlungsquoten handelt es sich bei den Verwaltungsbezirken mit den niedrigsten Quoten um Kreise, die zu den Einzugsgebieten der Krankenhäuser in kreisfreien Städten gehören, wie z. B. der Rhein-Erft-Kreis. Dieser Kreis zählt zu den Haupteinzugsgebieten der kreisfreien Stadt Köln. Die Zahl der dort behandelten Patient(inn)en war – ebenso wie beim Rhein-Sieg-Kreis – um über 40 Prozent kleiner als

Aus den oben dargestellten Behandlungsquoten ist nicht erkennbar, aus welchen Verwaltungsbezirken die kranken Personen zur Behandlung in einen bestimmten Kreis oder eine bestimmte kreisfreie Stadt konkret gekommen sind oder wie viele dort wohnende, kranke Personen auch dort behandelt worden sind.

Mit den Informationen der Krankenhausstatistik lässt sich auch eine Einzugsgebietsstatistik erstellen, wie in

der Tabelle 3 gezeigt wird. Um sie überschaubar zu halten, ist sie beispielhaft auf die kreisfreien Städte und Kreise des Regierungsbezirks Köln beschränkt.

Auf der Hauptdiagonalen ist erkennbar, wie viel Prozent der Patient(inn)en, die in den Krankenhäusern eines Gebiets behandelt worden sind, auch dort wohnten.

Beim Regierungsbezirk Köln zeigt sich, dass dies für die Mehrheit der kreisfreien Städte und Kreise der Fall war. Den höchsten Anteil an Behandlungen von Patient(inn)en, die in demselben Gebiet wohnten, hatte mit 90 Prozent der Kreis Heinsberg. Lediglich für die kreisfreie Stadt Bonn lag der Prozentsatz unter 50 Prozent (42 Prozent). Da Bonn eine Universitätsklinik hat, könnten besonders viele Kranke aus anderen Gebieten dort behandelt worden sein.

Von den aus „fremden“ Gebieten kommenden Patientinnen und Patienten wohnten – wie nicht anders zu erwarten – die meisten in den kreisfreien Städten und Kreisen, die an das Behandlungsgebiet angrenzen. Hier fällt wieder die Stadt Bonn auf, in deren Krankenhäuser ein Drittel der gesamten Behandlungen auf Personen mit Wohnsitz im Rhein-Sieg-Kreis entfielen. Haupteinzugsgebiet für die kreisfreie Stadt Aachen war mit einem Anteil von 22

3. Vollstationär entlassene Personen insgesamt im Regierungsbezirk Köln 2005 nach Einzugsgebieten													
Wohnort	Behandlungsort												
	krfr. Stadt Aachen	krfr. Stadt Bonn	krfr. Stadt Köln	krfr. Stadt Leverkusen	Kreis Aachen	Kreis Düren	Rhein-Erft-Kreis	Kreis Euskirchen	Kreis Heinsberg	Oberbergischer Kreis	Rhein.-Berg. Kreis	Rhein-Sieg-Kreis	Reg.-Bez. Köln
vollstationär behandelte Patient(inn)en in %													
Krfr. Stadt Aachen	52,0	0,1	0,1	–	6,0	0,6	–	0,1	0,2	–	–	0,1	4,9
Krfr. Stadt Bonn	0,1	42,0	0,5	0,1	0,1	0,1	0,3	0,2	–	0,1	0,2	4,1	6,7
Krfr. Stadt Köln	0,6	1,9	71,7	6,8	0,2	0,8	5,6	0,8	0,1	0,6	8,0	3,4	21,5
Krfr. Stadt Leverkusen	0,1	0,2	1,3	60,0	–	–	0,1	0,1	–	0,1	1,7	0,2	3,3
Kreis Aachen	21,8	0,1	0,1	0,1	77,1	4,6	0,1	0,6	2,9	–	–	0,1	7,3
Kreis Düren	6,8	0,3	0,5	0,1	7,4	77,8	0,5	1,9	1,4	–	–	0,2	6,4
Rhein-Erft-Kreis	1,0	3,4	10,7	1,5	0,6	9,8	84,8	6,2	0,8	0,1	0,5	1,7	9,6
Kreis Euskirchen	0,6	3,2	0,9	0,2	0,2	1,5	1,4	82,6	–	0,1	0,1	0,5	4,4
Kreis Heinsberg	6,8	0,1	0,1	0,1	5,4	3,2	0,1	0,1	90,4	–	0,1	–	4,6
Oberbergischer Kreis	0,1	0,8	1,4	0,7	–	–	0,1	–	–	74,5	4,5	1,0	5,8
Rhein.-Berg. Kreis	0,1	0,4	3,7	17,6	–	0,1	0,2	0,1	–	4,2	76,1	0,7	5,7
Rhein-Sieg-Kreis	0,2	33,1	2,4	0,5	0,1	0,1	4,4	3,3	–	7,7	1,9	78,0	11,9
Reg.-Bez. Köln	90,2	85,7	93,4	87,7	97,6	98,6	97,5	95,9	96,0	87,5	93,1	90,0	92,1
Nordrhein-Westfalen	93,0	88,1	97,2	98,7	98,3	99,2	98,4	96,8	99,1	96,6	99,1	91,8	95,6
Übrige Gebiete	7,0	11,9	2,8	1,3	1,7	0,8	1,6	3,2	0,9	3,4	0,9	8,2	4,4
Insgesamt	100	100	100	100	100	100	100	100	100	100	100	100	100

4. Vollstationär entlassene Personen mit der Diagnose „Krankheiten des Auges und der Augenanhangsgebilde“ im Regierungsbezirk Köln 2005 nach Einzugsgebieten													
Wohnort	Behandlungsort												
	krfr. Stadt Aachen	krfr. Stadt Bonn	krfr. Stadt Köln	krfr. Stadt Leverkusen	Kreis Aachen	Kreis Düren	Rhein-Erft-Kreis	Kreis Euskirchen	Kreis Heinsberg	Oberbergischer Kreis	Rhein.-Berg. Kreis	Rhein-Sieg-Kreis	Reg.-Bez. Köln
	vollstationär behandelte Patient(inn)en in %												
Krfr. Stadt Aachen	28,5	–	0,2	–	8,3	–	–	–	–	–	–	–	4,5
Krfr. Stadt Bonn	–	23,9	0,6	–	–	–	–	–	–	–	–	1,2	7,3
Krfr. Stadt Köln	0,2	1,5	36,9	5,3	–	–	11,4	–	–	–	47,7	9,4	18,2
Krfr. Stadt Leverkusen	–	0,3	5,5	47,4	–	–	1,0	–	–	–	1,8	1,2	2,7
Kreis Aachen	28,2	0,1	0,3	–	70,2	0,4	–	–	0,7	–	–	–	5,0
Kreis Düren	14,5	0,7	0,9	–	4,1	92,4	1,5	3,3	0,7	–	–	–	4,1
Rhein-Erft-Kreis	2,9	4,4	14,2	–	7,4	3,8	35,3	3,6	7,0	–	–	0,6	8,9
Kreis Euskirchen	1,1	7,5	1,2	–	–	0,8	2,5	84,7	–	–	–	0,6	4,2
Kreis Heinsberg	13,2	0,1	0,3	–	7,4	0,8	–	–	73,3	–	–	–	3,3
Oberbergischer Kreis	0,1	1,8	4,7	–	–	–	0,5	–	–	92,3	5,0	1,2	3,9
Rhein.-Berg. Kreis	0,1	1,0	9,5	42,1	–	0,8	–	–	–	0,9	39,4	1,8	5,2
Rhein-Sieg-Kreis	0,1	30,1	2,2	–	–	–	5,5	1,8	–	3,6	–	73,1	10,6
Reg.-Bez. Köln	89,0	71,4	76,4	94,7	97,5	99,2	57,7	93,5	81,9	96,8	94,0	88,9	78,0
Nordrhein-Westfalen	91,9	75,6	92,1	100,0	97,5	99,6	69,2	93,5	98,5	98,2	100,0	89,5	87,4
Übrige Gebiete	8,1	24,4	7,9	–	2,5	0,4	30,8	6,5	1,5	1,8	–	10,5	12,6
Insgesamt	100	100	100	100	100	100	100	100	100	100	100	100	100

Prozent der behandelten Fälle der Kreis Aachen und für die kreisfreie Stadt Leverkusen mit 18 Prozent der Rheinisch-Bergische Kreis.

Der weitaus größte Teil der in den einzelnen kreisfreien Städten und Kreisen des Regierungsbezirks Köln behandelten Personen wohnte auch in diesem Regierungsbezirk. Die Spanne reicht von 86 Prozent beim Behandlungsort Bonn bis 99 Prozent beim Kreis Düren. Noch etwas höher fallen die Prozentsätze aus, wenn man das Wohngebiet auf ganz NRW ausdehnt. Dennoch haben für die kreisfreien Städte Bonn und Aachen – bedingt durch ihre Randlage – auch Gebiete außerhalb Nordrhein-Westfalens relativ große Bedeutung.

Ein anderes Bild kann sich ergeben, wenn die Einzugsgebiete einzelner Diagnosegruppen betrachtet werden. Greift man die Augenkrankheiten heraus, wird deutlich, dass hier die Patientenwanderungen erheblich stärker ausgeprägt waren, als in der o. a. Tabelle mit den Wanderungen aller Patient(inn)en, ohne Unterscheidung nach Diagnosen. Es ist zu sehen, dass der Anteil der Augenpatienten, die in einer kreisfreien Stadt oder in einem Kreis nicht nur behandelt worden sind, sondern auch dort wohnten, meistens viel geringer war als bei der Betrachtung der Gesamtpatienten. Beim Rhein-Erft-Kreis ist die Abweichung gegenüber dem Anteil der dort wohnenden und behan-

delteten Gesamtpatienten besonders groß. Der Anteil der Augenpatienten betrug hier nur 35 Prozent, derjenige der Gesamtpatienten dagegen 85 Prozent, eine Abweichung um 50 Prozentpunkte! Aber auch beim Rheinisch-Bergischen Kreis (37 Punkte) und bei der Stadt Köln (35 Punkte) sind die Unterschiede beträchtlich. Die Wanderungen der Augenpatienten beschränkten sich nicht auf den Regierungsbezirk Köln. 22 Prozent der dort behandelten Augenpatienten kamen von außerhalb des Regierungsbezirks, bezogen auf die Gesamtpatienten waren es nur 8 Prozent.

Die starke regionale Ungleichverteilung erklärt sich in erster Linie daraus, dass viele Patient(inn)en Krankenhäuser mit Augenspezialabteilungen aufsuchen, die oft nicht in ihrem Wohngebiet liegen. Dominierend sind dabei die Universitätsstädte Bonn, Aachen und vor allem Köln. In den drei Städten zusammen konzentrierten sich 91 Prozent der rd. 19 000 im Regierungsbezirk Köln behandelten Augenpatienten. Bei den Augenbehandlungen in den anderen Gebieten handelte es sich daher nur um geringe Fallzahlen, was bei der Interpretation der Prozentzahlen zu beachten ist.

Bezogen auf ganz Nordrhein-Westfalen machten die Personen, die von außerhalb des Landes zur Behandlung nach NRW gekommen sind, nur

einen geringen Anteil an allen in NRW behandelten Patient(inn)en aus. Knapp 97 Prozent der in NRW behandelten Personen wohnten auch dort. 2,5 Prozent der behandelten Personen kamen aus den benachbarten Bundesländern Niedersachsen, Hessen und Rheinland-Pfalz. Patient(inn)en aus dem übrigen Bundesgebiet sowie aus dem Ausland – hier vor allem aus den angrenzenden Ländern Niederlande und Belgien – hatten noch nicht einmal einen Anteil von 1 Prozent an den in NRW behandelten Personen.

5. Einzugsgebiete für die 2005 in NRW behandelten Patient(inn)en		
Wohnland der in NRW behandelten Patient(inn)en	Vollstationär behandelte Patient(inn)en in NRW	
	Anzahl	%
Nordrhein-Westfalen	3 835 768	96,8
Niedersachsen	46 429	1,2
Hessen	15 510	0,4
Rheinland-Pfalz	35 668	0,9
Übriges Bundesgebiet	18 565	0,5
Ausland	12 604	0,3
darunter		
Niederlande	4 067	0,1
Belgien	2 351	0,1
Insgesamt	3 964 544	100

Wie eingangs angesprochen, steht die Zahl der in NRW wohnenden Patient(inn)en, die sich im Ausland behandeln ließen, nicht zur Verfügung, da kein Datenaustausch mit dem Ausland besteht. Somit kann keine Einzugsgebietsstatistik für die Patientenwanderungen zwischen Nordrhein-Westfalen und den übrigen Gebieten erstellt werden.

6. Vollstationär behandelte Personen in Rehabilitations- oder Vorsorgeeinrichtungen 2005 nach Einzugsgebieten

Wohnort	Behandlungsort					Nordrhein-Westfalen
	Regierungsbezirk					
	Düsseldorf	Köln	Münster	Detmold	Arnsberg	
	vollstationär behandelte Patient(inn)en in %					
Kreisfreie Städte						
Düsseldorf	13,3	3,3	0,9	1,3	1,4	2,6
Duisburg	6,7	2,1	0,6	1,4	1,8	2,0
Essen	8,5	1,8	0,7	1,9	3,8	2,9
Krefeld	2,7	0,9	0,3	0,5	0,6	0,7
Mönchengladbach	5,7	1,3	0,3	0,4	0,5	1,0
Mülheim an der Ruhr	2,5	0,7	0,3	0,6	1,1	0,9
Oberhausen	1,8	1,0	0,3	0,8	0,9	0,9
Remscheid	0,9	1,0	0,1	0,4	0,5	0,5
Solingen	1,3	1,2	0,0	0,4	0,6	0,6
Wuppertal	3,2	1,6	0,6	1,1	2,1	1,6
Kreise						
Kleve	3,7	1,6	0,4	0,8	0,9	1,2
Mettmann	8,7	2,6	0,6	1,4	1,9	2,3
Rhein-Kreis Neuss	7,7	3,2	0,5	0,9	1,1	1,8
Viersen	4,6	1,3	0,1	0,6	0,6	1,0
Wesel	3,9	1,9	0,8	1,6	1,9	1,9
Reg.-Bez. Düsseldorf	75,1	25,6	6,5	14,0	19,9	22,1
Kreisfreie Städte						
Aachen	0,7	4,1	0,2	0,2	0,3	0,8
Bonn	0,5	5,1	0,1	0,3	0,4	1,0
Köln	1,9	13,1	1,0	1,0	2,0	3,1
Leverkusen	0,8	2,5	0,3	0,3	0,7	0,7
Kreise						
Aachen	0,9	4,2	0,3	0,4	0,4	1,0
Düren	0,7	4,3	0,3	0,4	0,4	1,0
Rhein-Erft-Kreis	1,2	6,2	0,4	0,7	0,8	1,5
Euskirchen	0,3	4,2	0,2	0,2	0,2	0,8
Heinsberg	1,9	3,0	0,2	0,4	0,5	0,9
Oberbergischer Kreis	0,5	5,2	0,2	0,4	0,9	1,2
Rheinisch-Bergischer Kreis	0,7	3,6	0,2	0,4	0,7	1,0
Rhein-Sieg-Kreis	0,7	9,0	0,4	0,6	1,1	1,9
Reg.-Bez. Köln	10,9	64,4	3,8	5,2	8,4	15,0
Kreisfreie Städte						
Bottrop	0,5	0,0	0,2	0,7	0,7	0,6
Gelsenkirchen	1,1	0,2	0,7	1,5	2,4	1,5
Münster	0,2	0,1	15,3	1,2	0,5	1,2
Kreise						
Borken	0,7	0,1	2,9	1,9	1,1	1,4
Coesfeld	0,2	0,1	4,7	1,4	1,0	1,1
Recklinghausen	1,7	0,4	1,9	3,7	4,6	3,2
Steinfurt	0,3	0,1	16,3	2,4	0,7	1,8
Warendorf	0,2	0,1	11,9	1,8	1,9	1,7
Reg.-Bez. Münster	4,7	1,1	53,8	14,6	12,8	12,4
Kreisfreie Stadt						
Bielefeld	0,1	0,1	0,9	3,3	0,5	1,8
Kreise						
Gütersloh	0,2	0,1	1,1	2,5	2,5	1,9
Herford	0,1	0,0	0,4	2,7	0,2	1,5
Höxter	0,1	0,0	0,3	2,0	0,1	1,0
Lippe	0,1	0,0	0,8	4,4	0,4	2,3
Minden-Lübbecke	0,2	0,1	0,7	4,2	0,2	2,2
Paderborn	0,1	0,1	0,8	2,9	0,8	1,7
Reg.-Bez. Detmold	0,7	0,3	4,9	21,9	4,8	12,4
Kreisfreie Städte						
Bochum	1,5	0,3	0,9	1,8	3,7	2,0
Dortmund	1,2	0,4	1,8	3,3	6,2	3,3
Hagen	0,4	0,1	0,5	1,0	1,8	1,0
Hamm	0,1	0,1	2,3	1,1	1,3	0,9
Herne	0,5	0,1	0,3	0,9	1,0	0,8
Kreise						
Ennepe-Ruhr-Kreis	0,7	0,3	0,5	1,6	2,8	1,6
Hochsauerlandkreis	0,1	0,2	0,4	1,2	4,2	1,7
Märkischer Kreis	0,5	0,5	1,2	1,7	4,2	2,0
Olpe	0,1	0,5	0,2	0,4	1,7	0,7
Siegen	0,2	0,7	0,3	0,6	4,8	1,6
Soest	0,2	0,1	0,5	1,5	5,0	2,0
Unna	0,5	0,2	2,1	2,5	4,0	2,4
Reg.-Bez. Arnsberg	6,0	3,3	11,1	17,5	40,8	19,9
Nordrhein-Westfalen	97,4	94,8	80,1	73,2	86,6	81,8
Niedersachsen	0,6	0,2	15,5	13,6	3,0	8,1
Hessen	0,3	0,4	0,6	2,5	3,5	2,2
Rheinland-Pfalz	0,3	2,6	0,4	1,2	1,6	1,4
Übriges Bundesgebiet	1,2	0,7	3,3	9,3	5,1	6,2
Ausland	0,2	1,3	0,1	0,1	0,1	0,3
Insgesamt	100	100	100	100	100	100

2. Rehabilitations- oder Vorsorgeeinrichtungen

Krankenhäuser sind regional so gelegen, dass die Einwohner einen nicht zu weiten Weg haben, um eine stationäre medizinische Behandlung in Anspruch nehmen zu können. Dies ist u. a. Zweck des Krankenhausgesetzes Nordrhein-Westfalen (KHG NRW) vom Dezember 1998, in dem in § 1 festgelegt ist, dass auf der Grundlage dieses Gesetzes „eine patienten- und bedarfsgerecht gestufte wohnortnahe Versorgung der Bevölkerung durch Krankenhäuser sicherzustellen“ ist. Nach § 13, Absatz 2, ist dies Aufgabe des Krankenhausplans.

Bei Rehabilitations- oder Vorsorgeeinrichtungen verhält sich dies anders. Sie sind nicht gleichmäßig über das Land verteilt. Die dadurch verstärkt auftretenden Patientenwanderungen zwischen den Verwaltungsbezirken kommen sehr gut in den Zahlen der Diagnosestatistik zum Ausdruck, wenn auch in Kauf genommen werden muss, dass nur Diagnoseergebnisse – und damit Wohnortangaben der behandelten Personen – für die großen Einrichtungen zur Verfügung stehen. Da aber diese Einrichtungen gut 90 Prozent der gesamten behandelten Personen abdecken, dürften sich Niveau und regionale Verteilung der Zahlen durch Einbeziehung auch der kleinen Häuser nicht wesentlich ändern.

In Rehabilitations- oder Vorsorgeeinrichtungen mit mehr als 100 Betten des Regierungsbezirks Münster waren im Jahre 2005 ca. 4 700 Personen stationär untergebracht. Das waren nur 2,5 Prozent der 185 200 insgesamt in den großen Reha-Einrichtungen Nordrhein-Westfalens behandelten Kranken.

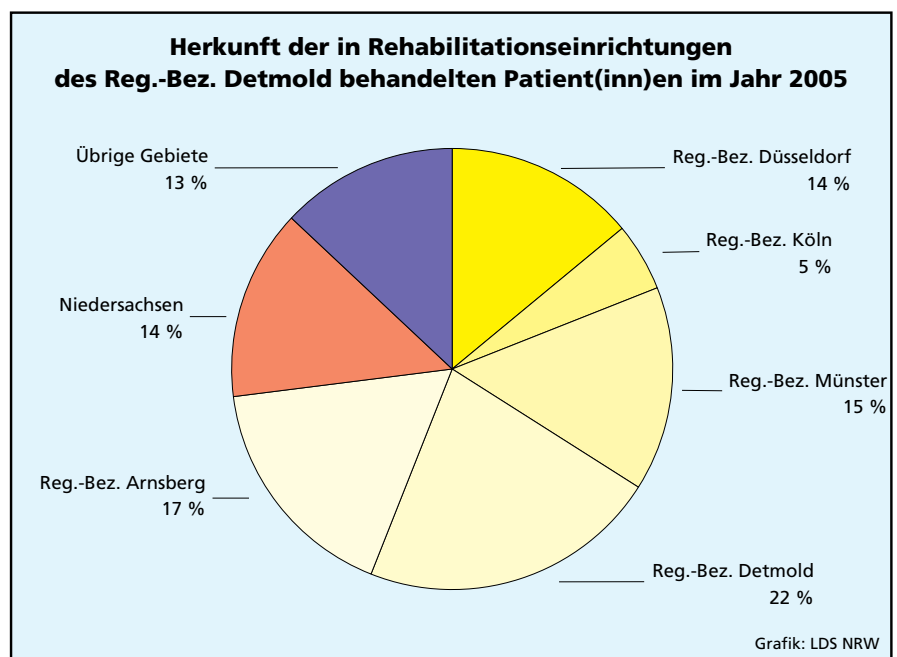
Demgegenüber hatten 22 000 Kranke mit Wohnsitz im Regierungsbezirk Münster Rehabilitations- oder Vorsorgeeinrichtungen in ganz NRW aufgesucht. Bezieht man die Herkunft der Kranken auf die kreisfreien Städte und Kreise, dann kamen die meisten der in den Einrichtungen des Regierungsbezirks Münster be-

handelten Kranken aus dem Kreis Steinfurt (16 Prozent) und der kreisfreien Stadt Münster (15 Prozent). Als Einzugsgebiet hervorzuheben ist auch das angrenzende Bundesland Niedersachsen, dessen Patient(inn)en 15 Prozent der im Regierungsbezirk Münster Behandelten ausmachten.

Ebenso wie der Regierungsbezirk Münster ist der Regierungsbezirk Düsseldorf ein „Defizit-Gebiet“, d. h., es wurden im Vergleich mit den Patient(inn)en mit Wohnort in diesem Regierungsbezirk erheblich weniger Patient(inn)en in Rehabilitations- oder Vorsorgeeinrichtungen dieses Bezirks behandelt. Nur 8 Prozent aller in NRW in Rehabilitations- oder Vorsorgeeinrichtungen untergebrachten Kranken wurden im Regierungsbezirk Düsseldorf behandelt. Im Vergleich dazu beträgt der Bevölkerungsanteil dieses Regierungsbezirks an NRW insgesamt – genauso wie sein Anteil an den dort in Krankenhäusern behandelten Patient(inn)en – ca. 30 Prozent. Wohngebiete der behandelten Patient(inn)en waren vor allem die kreisfreien Städte und Kreise des eigenen Regierungsbezirks, wie die kreisfreie Stadt Düsseldorf und der Rhein-Kreis Neuss.

Im Regierungsbezirk Köln hielt sich die Zahl der Patient(inn)en nach dem Wohnort und nach dem Behandlungsort fast die Waage. Über-

schuss-Gebiete mit erheblich mehr Behandlungen in den Rehabilitations- oder Vorsorgeeinrichtungen als dort zahlenmäßig Patient(inn)en ihren Wohnsitz hatten, waren die Regierungsbezirke Arnsberg und Detmold. Schwerpunkt der Behandlungen war der Regierungsbezirk Detmold mit seinen bekannten Heilbädern Bad Oeynhausen, Bad Salzuflen, Bad Lippspringe, Bad Meinberg und Bad Driburg. Auf diesen Regierungsbezirk entfiel die Hälfte aller in NRW vorgenommenen Behandlungen. Das war mehr als das 4-fache der Patient(inn)en, die im Regierungsbezirk Detmold wohnten und sich in einer Einrichtung innerhalb ganz Nordrhein-Westfalens hatten behandeln lassen. Einzugsgebiete der Rehabilitations- oder Vorsorgeeinrichtungen dieses Regierungsbezirks waren mit einer größeren Zahl von Patient(inn)en z. B. die Kreise Recklinghausen, Unna und Steinfurt sowie die kreisfreie Stadt Dortmund. Die Einrichtungen des Regierungsbezirks wurden im besonderen Maße auch von Patient(inn)en mit Wohnsitz außerhalb Nordrhein-Westfalens genutzt. Ungefähr jeder vierte Patient (27 Prozent) kam aus einem anderen Bundesland. Den Hauptanteil daran hatte das Nachbarland Niedersachsen mit 14 Prozent der Patient(inn)en. Das Ausland hatte als Herkunftsgebiet der Patienten (0,1 Prozent) keine Bedeutung.



3. Versorgungsgebiete der Krankenhäuser

Gebietsabgrenzungen, bei denen die Wohn- und Behandlungsregion der Krankenhauspatienten aufeinander abgestimmt sind, sind die sog. Versorgungsgebiete.

Das Land NRW ist für landesplanerische Zwecke in 16 Versorgungsgebiete eingeteilt. Diese Versorgungsgebiete wurden vor einigen Jahren von den an der Krankenhausplanung

beteiligten Institutionen (u. a. Gesundheitsministerium, Verbände der Krankenkassen, Krankenhausgesellschaft NRW) abgegrenzt. Sie setzen sich in der Regel aus mehreren kreisfreien Städten und Kreisen zusammen. Lediglich das Versorgungsgebiet 15 besteht allein aus einem Kreis (Hochsauerlandkreis). Durch die Zusammensetzung aus vollständigen kreisfreien Städten und Kreisen lassen sich die statistischen Angaben für die Versorgungsgebiete leicht aus den Ergebnissen der Kran-

kenhausstatistik, die nach den Gebieten der allgemeinen Verwaltung aufbereitet werden, errechnen.

Lt. § 13, Absatz 4 des o. a. Krankenhausgesetzes NRW ist in den Versorgungsgebieten die notwendige abgestufte Versorgung mit ortsnahen, leistungsfähigen und bedarfsgerechten Krankenhäusern zu gewährleisten.

In einem Versorgungsgebiet soll also grundsätzlich die gesamte Versorgung der dort wohnenden Bevölkerung mit Krankenhausleistungen abgedeckt werden. Daraus folgt, dass die Zahl der in diesen Gebieten behandelten Patient(inn)en nicht wesentlich von der Zahl der dort wohnenden Patient(inn)en abweichen dürfte.

Die Behandlungsquoten der Versorgungsgebiete in der Tabelle bestätigen die Forderung des KHG NRW weitgehend. Fast alle Quoten liegen zwischen 90 und 110. Eine stärkere Abweichung gibt es mit einer Quote von 116 nur noch bei Versorgungsgebiet 2 (Essen, Mülheim an der Ruhr, Oberhausen). Bei den festgestellten Abweichungen ist zu berücksichtigen, dass die kreisfreien Städte und Kreise nur ein grobes Raster zur Abgrenzung der Versorgungsgebiete bilden. Es ist zu vermuten, dass eine gemeindegrenze Abgrenzung eine noch bessere Abstimmung von Behandlungs- und Wohnregion – und damit eine noch geringere Streuung der Behandlungsquoten um 100 – bewirken würde.

7. Behandlungsquoten*) in den Versorgungsgebieten der Krankenhäuser 2005

VG - Nr.	Versorgungsgebiet	Stationär behandelte Patient(inn)en am		Behandlungsquote
		Behandlungsort	Wohnort	
1	Krfr. Städte Düsseldorf, Remscheid, Solingen, Wuppertal; Kreis Mettmann	382 966	376 200	102
2	Krfr. Städte Essen, Mülheim an der Ruhr, Oberhausen	271 539	233 985	116
3	Krfr. Stadt Duisburg, Kreise Kleve, Wesel	280 684	291 903	96
4	Krfr. Städte Krefeld, Mönchengladbach; Rhein-Kreis Neuss, Kreis Viersen	255 931	261 957	98
5	Krfr. Städte Köln, Leverkusen; Rhein-Erft-Kreis, Oberbergischer Kreis, Rheinisch-Bergischer Kreis	426 218	437 335	97
6	Krfr. Stadt Bonn, Kreis Euskirchen, Rhein-Sieg-Kreis	222 336	212 781	104
7	Krfr. Stadt Aachen, Kreise Aachen, Düren, Heinsberg	217 114	218 822	99
8	Krfr. Städte Bottrop, Gelsenkirchen; Kreis Recklinghausen	255 384	266 877	96
9	Krfr. Stadt Münster, Kreise Borken, Coesfeld, Steinfurt, Warendorf	332 739	330 280	101
10	Krfr. Stadt Bielefeld, Kreise Herford, Lippe, Minden-Lübbecke	327 912	335 121	98
11	Kreise Höxter, Paderborn	100 727	98 773	102
12	Krfr. Stadt Hamm, Kreise Soest, Unna	210 495	221 148	95
13	Krfr. Städte Bochum, Dortmund, Herne	293 712	266 625	110
14	Krfr. Stadt Hagen, Ennepe-Ruhr-Kreis, Märkischer Kreis	223 150	238 187	94
15	Hochsauerlandkreis	64 711	68 933	94
16	Kreise Olpe, Siegen-Wittgenstein	98 926	94 293	105

*) Behandlungsquote = behandelte Personen am Behandlungsort dividiert durch behandelte Personen am Wohnort x 100

Clusteranalyse zur Identifizierung von verschiedenen Gruppen von Internetnutzern

Christian Hoops*), Diplom-Statistiker Josef Schäfer

Die Nutzung von Informations- und Kommunikationstechnologien (IKT) nimmt in der heutigen Gesellschaft stetig zu. Durch E-Government oder E-Commerce lassen sich dabei viele Abläufe im Alltag vereinfachen. Im Auftrag des Statistischen Amtes der Europäischen Union findet seit 2002 jedes Jahr eine Umfrage zur Nutzung von Informations- und Kommunikationstechnologien in Privathaushalten statt. Die Erhebung wird jährlich durchgeführt, voraussichtlich bis zum Jahr 2010.

Im Rahmen eines Praktikums im Landesamt für Datenverarbeitung und Statistik NRW wurden die Daten der IKT-Erhebung des Jahres 2006 genutzt, um mittels Clusteranalysen eine Einteilung der 5 722 in Nordrhein-Westfalen befragten Personen in homogene Gruppen bezüglich ihrer Internetnutzung vorzunehmen.

Die Analyse erfolgte zum einen unmittelbar als Clusteranalyse¹⁾ mit Bestimmung der optimalen Anzahl verschiedener Personengruppen, die hier fünf beträgt. Entsprechend ihrer Ausprägungen in den relevanten Merkmalen werden die identifizierten Gruppen als „Power-User“, „Wissbegierige“, „Entertainment-User“, „Basis-User“ und „Gelegenheits-User“ bezeichnet. Zum anderen wurde eine Faktorenanalyse¹⁾ vorgeschaltet, deren orthogonal rotierte Faktoren per definitionem nicht mehr voneinander abhängig sind. Aufgrund ihrer Ladungen in den einzelnen Merkmalen wurden vier Hauptfaktoren gebildet und als „Unterhaltungsfaktor“, „Handelsfaktor“, „Bildungsfaktor“ und „Erreichbarkeitsfaktor“ bezeichnet. Bezüglich dieser vier Faktoren wurden die befragten Personen abermals in homogene Gruppen eingeteilt. Diesmal ergab die Analyse eine optimale Anzahl von sechs verschiedenen Personengruppen. Neben den Gruppen aus der ersten Analyse wurde dabei auch die Gruppe der „Kommunikations-User“ gebildet. Die Gruppe der „Entertainment-User“ ist dagegen nicht mehr identifiziert worden. Zudem gab es hierbei nicht nur eine, sondern zwei Gruppen von „Power-Usern“, die sich vor allem durch ihr Interesse an Bildung voneinander unterscheiden.

Ziel der Clusteranalyse war es, Gruppen von Personen zu identifizieren, die bezüglich ihres Internetnutzungsverhaltens innerhalb der Gruppe homogen, aber zwischen den Gruppen heterogen sind. Zunächst wurde die Unabhängigkeit der für die Gruppenbildung relevanten Variablen getestet²⁾ und die befragten Personen wurden mittels einer Two-Step-Clusteranalyse auf verschiedene Gruppen verteilt³⁾.

*) Christian Hoops ist Student der Statistik an der Universität Dortmund. Der Beitrag ist im Rahmen eines Praktikums im LDS NRW entstanden. – 1) Vgl. Lehrbücher zu multivariaten statistischen Verfahren, z. B. Bühl, Zöfel, 1998. – 2) Die Unabhängigkeit zweier Merkmale wurde getestet durch einen Unabhängigkeitstest und ihre Abhängigkeit quantifiziert mit dem Phi-Koeffizienten. – 3) Mit Hilfe des Bayes-Kriterium nach Schwarz sowie dem Log-Likelihood-Distanzmaß wird die optimale Anzahl der Cluster berechnet (vgl. „Twostep-Cluster“, SPSS Version 14.01).

Vor der Durchführung der Faktorenanalyse wurde anhand des Kaiser-Meyer-Olkin-Kriterium⁴⁾ geprüft, ob die Variablen dafür geeignet sind. Die Faktorenanalyse lieferte voneinander unabhängige Faktoren, die für eine weitere Clusteranalyse verwendet wurden. Die Ergebnisse beider Analysen wurden miteinander verglichen. Die folgende Übersicht zeigt die für die Bildung der Cluster verwendeten Merkmale.

Übersicht über die in der Analyse verwendeten Variablen

- Häufigkeit der Internetnutzung
- Senden und Empfangen von E-Mails

4) Das Kriterium dient zur Überprüfung der Korrelationsmatrix vor Durchführung einer Faktorenanalyse. Es gibt Auskunft darüber, ob mit dieser Stichprobe eine Faktorenanalyse sinnvoll ist. Werte größer als 0,7 sind nach Kaiser und Rice ziemlich gut (vgl. Backhaus et al., 2000, S. 269).

- Chatten/Besuch von Foren
- Suche nach Information über Waren und Dienstleistungen
- Nutzung von Internet-Radio/Internet-Fernsehen
- Spielen/Herunterladen von Spielen oder Musik
- Herunterladen von Software
- Lesen/Herunterladen von Internet-Zeitungen, Magazinen
- Internet-Banking
- Verkauf von Waren oder Dienstleistungen
- Schul- und Hochschulausbildung
- Weiterbildung aus privatem Interesse
- Weiterbildung aus beruflichem Interesse

Der Personen-Datensatz enthält die Ausprägungen von insgesamt 5 722 Personen, von denen 4 112 (71,9 %) das Internet genutzt und 4 097 die oben genannten Fragen beantwortet haben. Nur diese wurden in die Analysen einbezogen.

Die gebildeten Cluster wurden nach den Merkmalen Geschlecht, Alter, Familienstand, Schulbildung und Erwerbsstatus ausgewertet.

Die Daten wurden zwischen Mai und Juli 2006 als Quotenstichprobe erhoben. Basis der Auswahl stellte die Dauerstichprobe dar, ein Access-Panel der amtlichen Statistik, das aus ausscheidenden Mikrozensus-Haushalten rekrutiert wird. Darüber hinaus wurde auf weitere Haushalte zurückgegriffen, die bereits an einer anderen Erhebung wie z. B. der „Einkommens- und Verbrauchsstichprobe“ oder den „Laufenden Wirtschaftsrechnungen“ teilgenommen haben.

Als Quotierungsmerkmale wurden Geschlecht, Alter, Haushaltstyp, die soziale Stellung des Haupteinkommensbezieher und das Haushaltsnettoeinkommen verwendet. Zur

Hochrechnung erfolgt eine Anpassung dieser Merkmale an die Verteilung des Mikrozensus. Die Bildung der Faktoren und Cluster erfolgt an der gegebenen Stichprobe ohne Berücksichtigung der Hochrechnung. Anteilswerte werden dagegen hochgerechnet dargestellt. Die hochgerechneten Anteile entsprechen damit in der Regel nicht den Anteilen der ungewichteten Fallzahlen.

Die Clusteranalyse wird hier als ein exploratives Verfahren eingesetzt, das nach einer Gruppen-Struktur in den Daten sucht. Als Ergebnis werden Gruppen identifiziert, in denen sich Elemente befinden, die sich innerhalb der Gruppe bzgl. vorher festgelegter Merkmale ähneln und zwischen den Gruppen möglichst verschieden sind.

Es wird hier ein Two-Step-Verfahren angewandt, das die Bestimmung einer optimalen Clusteranzahl ermöglicht. Gegenüber hierarchischen Verfahren hat die Two-Step-Clusteranalyse außerdem den Vorteil, dass Objekte im Laufe der Analyse die Gruppen wechseln können. Dies führt, vom Endergebnis her betrachtet, zu einer besseren Zuordnung als bei hierarchischen Methoden. Ferner gilt die Two-Step-Clusteranalyse als recht robust gegenüber Verletzungen sowohl der Verteilungsannahme (Normal- oder Multinomialverteilung) als auch der Annahme der Unabhängigkeit der Variablen⁵⁾.

Das Verfahren errechnete eine optimale Anzahl von 5 Clustern, deren Größe in Tabelle 1 dargestellt ist. Am kleinsten ist Cluster 1 ausgefallen. Er beinhaltet nur gut halb so viele Personen wie der größte Cluster 5, der hochgerechnet rund ein Viertel aller

1. Verteilung der hochgerechneten Personen nach Clustern		
Cluster	Clustergröße	Hochgerechnete Anteile
	Anzahl	%
1	592	16,2
2	878	20,8
3	690	15,8
4	897	22,7
5	1 040	24,5

5) Vgl. „SPSS 13.0 Base Benutzerhandbuch“, uni-hamburg.de, S. 531.

2. Internetnutzung und -aktivitäten nach Clustern						
Internetnutzung Internetaktivität	Cluster					Internetnutzer insgesamt
	1	2	3	4	5	
Anzahl						
Personen	592	878	690	897	1 040	4 097
Prozent						
Internetnutzung						
jeden Tag oder fast jeden Tag	96,2	76,6	76,4	57,5	15,7	60,4
mindestens einmal in der Woche	3,8	21,5	21,7	42,0	35,0	26,7
mindestens einmal im Monat	–	1,8	1,2	0,4	35,5	9,4
weniger als einmal im Monat	–	–	0,8	–	13,8	3,5
Internetaktivitäten						
Senden und Empfangen von E-Mails	99,7	98,7	98,2	100,0	41,1	84,9
Chatten/Besuch von Foren	76,7	41,1	88,2	19,6	20,2	44,3
Suche nach Informationen über Waren und Dienstleistungen	99,5	93,3	76,1	99,8	50,0	82,4
Nutzung von Internet-Radio/Internet-Fernsehen	62,9	11,8	32,9	8,6	3,2	20,5
Spielen/Herunterladen von Spielen und Musik	87,0	3,5	82,7	3,5	12,7	31,7
Herunterladen von Software	84,4	38,7	53,1	17,4	5,7	35,4
Lesen/Herunterladen von Internet-Zeitungen oder -Magazinen	58,5	41,6	22,5	16,4	9,1	27,6
Internet-Banking	89,2	52,5	2,1	53,2	10,6	40,3
Verkauf von Waren und Dienstleistungen	71,6	28,8	18,3	34,3	6,4	29,8
Schul- oder Hochschulausbildung	30,2	24,8	47,4	0,7	16,7	21,8
Weiterbildung aus privatem Interesse	65,7	85,3	35,1	6,3	15,0	39,9
Weiterbildung aus beruflichem Interesse	50,6	63,3	11,2	2,5	10,3	26,2

bei der Clusteranalyse berücksichtigten Personen enthält.

Tabelle 2 zeigt die Häufigkeit der Internetnutzung sowie der genannten Aktivitäten für die fünf erzeugten Cluster.

Personen des Clusters 1 gehen zum größten Teil „jeden Tag oder fast jeden Tag“ ins Internet. Die meisten Aktivitäten werden prozentual häufiger genutzt als von den Personen der übrigen Cluster. Deshalb wäre eine mögliche Bezeichnung dieses Clusters die Gruppe der „Power-User“.

Personen des zweiten Clusters nutzen das Internet vorrangig zum „Senden und Empfangen von E-Mails“ und sind stärker als andere Gruppen an der privaten und beruflichen Weiterbildung interessiert. Eine zutreffende Bezeichnung für die Personen in Cluster 2 könnte der „Wissbegierige“ sein.

Bei den Befragten des dritten Clusters wurden vor allem die Aktivitäten „Chatten/Besuch von Foren“,

„Spielen/Herunterladen von Spielen und Musik“ und „Schul- oder Hochschulausbildung“ weitaus häufiger ausgeübt als bei den übrigen Personen. Zudem werden sie bei der „Nutzung von Internet-Radio/Internet-Fernsehen“ nur noch von den „Power-Usern“ übertroffen. Daher könnte ein User in diesem Cluster als „Entertainment-User“ bezeichnet werden, der das Internet vor allem zur Unterhaltung durch Foren, Chats, Spiele, Musik, Radio oder Fernsehen benutzt.

Die Internetnutzung der Personen des Clusters 4 beschränken sich neben dem „Senden und Empfangen von E-Mails“ vor allem auf das „Suchen nach Informationen über Waren und Dienstleistungen“, auf das „Internet-Banking“ sowie auf den „Verkauf von Waren und Dienstleistungen“ – allerdings ohne die „Power-User“ bei diesen Aktivitäten zu übertreffen. Wenn auch diese Tätigkeiten im Business-Bereich liegen, könnte diese Gruppe angesichts der insgesamt unterdurchschnittlichen Internet-Aktivitäten als „Basis-User“ bezeichnet werden.

3. Soziodemografische Merkmale der befragten Personen								
Merkmal	Cluster					Personen ohne Internetzugang	Selten-Nutzer	Befragte Personen insgesamt
	1	2	3	4	5			
	%							
Männlich	73,2	53,6	55,1	45,5	46,1	38,9	45,1	48,6
Weiblich	26,8	46,4	44,9	54,5	53,9	61,1	54,9	51,4
Ledig	52,7	39,7	78,5	19,5	41,7	13,1	28,8	33,9
Verheiratet	41,3	48,5	17,7	72,0	48,3	57,9	56,4	51,2
Verwitwet	0,4	1,9	0,5	1,6	3,0	18,9	5,5	7,0
Geschieden	5,7	9,8	3,4	6,8	7,0	10,1	9,3	7,9
Alter von ... bis ... Jahren								
10 – 15	2,3	1,5	26,9	0,8	19,2	1,7	8,3	7,3
16 – 24	21,6	12,6	40,5	5,7	10,9	1,0	6,9	11,4
25 – 34	29,3	21,3	11,4	17,7	10,2	3,5	15,6	13,1
35 – 44	27,0	25,8	10,8	32,8	20,8	7,0	22,8	18,8
45 – 54	11,6	21,9	6,5	23,8	18,1	11,7	15,6	15,5
55 – 64	6,6	11,3	2,7	13,5	12,0	19,3	16,9	13,0
65 – 74	1,2	4,4	1,3	4,6	7,2	30,3	10,2	12,3
75 und mehr	0,4	1,2	–	1,0	1,6	25,6	3,7	8,6
Hauptschul- bzw. Volksschulabschluss	18,2	19,9	25,6	29,1	40,9	69,4	45,0	41,4
Realschul- bzw. gleichwertiger Abschluss	30,9	27,8	45,8	35,2	31,1	21,3	38,0	29,6
Fachhochschulreife	13,4	12,8	12,3	14,8	12,6	4,2	6,4	10,1
Abitur	37,5	39,4	16,3	21,0	15,4	5,2	10,6	18,9
Erwerbstätige	67,6	70,1	48,1	75,6	57,6	20,9	52,0	50,5
Arbeitslose	8,9	7,5	8,0	3,9	9,6	4,6	7,7	6,5
Schüler/-innen	3,6	3,5	25,3	0,3	4,3	0,4	1,5	3,8
Studierende	9,7	5,6	5,9	0,3	1,1	0	0,2	2,6
Übrige Personen	10,2	13,3	12,7	19,9	27,5	74,0	38,5	36,6

Personen des fünften Clusters nutzen das Internet eher selten und liegen bei allen genannten Aktivitäten unter dem Durchschnitt aller Befragten. Hinzuzufügen sind als Vergleichsgruppen noch die Personen, die die Frage „Haben Sie schon einmal das Internet genutzt?“ mit „Nein“ beantwortet haben sowie diejenigen, die zwar schon mal das Internet genutzt haben, aber das letzte Mal „vor mehr als drei Monaten“. Erstere werden von nun an „Personen ohne Internetzugang“ genannt, Letztere „Selten-Nutzer“.

Tabelle 3 zeigt die unterschiedliche Verteilung der soziodemografischen Merkmale für die befragten Personen.

Der typische „Power-User“ (Cluster 1) ist männlich, ledig, zwischen 16 und 44 Jahren und hat einen hohen Schulabschluss. Sowohl Erwerbstätige als auch insbesondere Studierende sind in dieser Gruppe überproportional vertreten. Der „Wissbegierige“ (Cluster 2) teilt mit dem Power-User den hohen Schulabschluss, ist aber im Durchschnitt etwas älter als der Power-User und entspricht bei

Geschlecht und Familienstand eher dem Durchschnitt aller Befragten. Die „Entertainment-User“ (Cluster 3) sind vor allem jung, mehr als zwei Drittel sind jünger als 25 Jahre und weisen dementsprechend einen hohen Anteil an Schülerinnen und Schülern auf. Der „Basis-User“ (Cluster 4) ist in der Regel verheiratet, im Alter zwischen 35 und 54 Jahren und erwerbstätig. Der Frauenanteil liegt über dem Durchschnitt. Bei den „Gelegenheits-Usern“ in Cluster 5 sind sowohl die Kinder zwischen 10 und

15 Jahren als auch die mittleren Altersklassen (35 – 54 Jahre) überproportional vertreten. Dagegen sind in diesem Cluster die Personen ab 75 Jahren sowie Studierende relativ selten vorhanden.

In der Gruppe der Personen ohne Internetzugang sind mehr als die Hälfte der Befragten 65 Jahre oder älter. Mehr als zwei Drittel der Personen in dieser Gruppe haben lediglich einen Hauptschul- bzw. Volksschulabschluss, nur gut 20 % sind erwerbstätig. Die „Selten-Nutzer“ sind in der Verteilung der untersuchten Merkmale den Personen in der Gesamtstichprobe recht ähnlich. Unterrepräsentiert sind allerdings die Studierenden, Schülerinnen und Schüler sowie Personen im Alter von 75 oder mehr Jahren. Personen mit niedrigen oder mittleren Schulabschlüssen sind in dieser Gruppe leicht überrepräsentiert.

Um unkorrelierte Variablen zu erhalten, was streng genommen Voraussetzung zur Anwendung einer Clusteranalyse ist, sowie um eine Reduzierung der Zahl der Variablen zu erreichen, wird nun der Clusteranalyse eine Faktorenanalyse vorgeschaltet. Durchgeführt wird die Faktorenanalyse hier in Form einer Hauptkomponentenanalyse⁶⁾. Die Idee dabei ist eine Reduzierung der ursprünglichen Vielzahl von Merkmalen auf wenige Faktoren oder Hauptkomponenten, die so gebildet werden, dass sie un-

6) Die Hauptkomponentenanalyse ist eine Methode zur Extraktion von Faktorladungen und Kommunalitäten.

4. Eigenwerte und rotierte Summe der quadrierten Ladungen für jede Komponente						
Hauptkomponente	Anfängliche Eigenwerte			Rotierte Summe der quadrierten Faktorladungen		
	insgesamt	Varianz	Varianz kumuliert	insgesamt	Varianz	Varianz kumuliert
	Faktor	%		Faktor	%	
1	3,124	24,028	24,028	2,121	16,312	16,312
2	1,698	13,062	37,091	1,857	14,283	30,595
3	1,247	9,594	46,684	1,643	12,640	43,234
4	1,034	7,956	54,640	1,483	11,406	54,640
5	0,850	6,535	61,175			
6	0,772	5,935	67,109			
7	0,723	5,559	72,669			
8	0,678	5,214	77,883			
9	0,626	4,817	82,699			
10	0,608	4,674	87,373			
11	0,568	4,371	91,744			
12	0,544	4,188	95,931			
13	0,529	4,069	100,000			

5. Rotierte Komponentenmatrix				
Variable	Hauptkomponente			
	Unterhaltungsfaktor	Handelsfaktor	Bildungsfaktor	Erreichbarkeitsfaktor
Häufigkeit der Internetnutzung	0,134	0,161	0,094	0,777
Senden und Empfangen von E-Mails	0,084	0,228	0,069	0,783
Chatten/Besuch von Foren	0,657	-0,191	0,042	0,289
Suche nach Informationen über Waren und Dienstleistungen	-0,034	0,583	0,111	0,281
Nutzung von Internet-Radio/Internet-Fernsehen	0,625	0,285	0,101	-0,054
Spielen/Herunterladen von Spielen und Musik	0,784	0,024	-0,046	0,053
Herunterladen von Software	0,530	0,388	0,198	0,109
Lesen/Herunterladen von Internet-Zeitungen oder -Magazinen	0,279	0,351	0,403	-0,076
Internet-Banking	-0,012	0,682	0,151	0,167
Verkauf von Waren und Dienstleistungen	0,242	0,575	0,025	0,112
Schul- oder Hochschulausbildung	0,473	-0,478	0,337	0,130
Weiterbildung aus privatem Interesse	0,123	0,018	0,795	0,140
Weiterbildung aus beruflichem Interesse	-0,048	0,144	0,795	0,064

korreliert sind und bei den untersuchten Daten einen möglichst hohen Anteil der Gesamtvarianz der Merkmale erklären, das heißt auf sich vereinigen können.

Die Anzahl der Faktoren wird durch das Kaiser-Kriterium⁷⁾ bestimmt (auch Eigenwert-Kriterium genannt), welches alle Faktoren mit einem Eigenwert⁸⁾ größer als 1 extrahiert. Wie Tabelle 4 zu entnehmen ist, werden 4 Faktoren extrahiert, die insgesamt 54,64 % der gesamten Varianz erklären.

Um unkorrelierte Faktoren zu erhalten, wird ein orthogonales Rotationsverfahren, die Varimax-Rotation⁹⁾ gewählt. Die rotierte Komponentenmatrix zeigt an, welche Variable wie hoch welchen Faktor lädt (vgl. Tabelle 5). Charakteristisch für den ersten Faktor sind hohe Werte der Variablen „Chatten/Besuch von Foren“, „Spielen/Herunterladen von Spielen und Musik“ und „Nutzung von Internet-Radio/Internet-Fernsehen“. Da diese Aktivitäten zum größten Teil der Unterhaltung dienen, nennen wir den ersten Faktor folglich „Unterhaltungsfaktor“.

7) Nach dem Kaiser-Kriterium werden lediglich Faktoren mit Eigenwerten größer als 1 zugelassen, da Faktoren mit einem kleineren Wert nicht mehr Varianz erklären als eine Variable allein. – 8) Der Eigenwert eines Faktors gibt hier an, wie viel Varianz der Faktor für alle Variablen erklärt. – 9) Rotation, welche die Varianz der Ladungsquadrate maximiert, so dass Faktoren in einigen Variablen hoch und in den restlichen Variablen möglichst niedrig laden (vgl. Fahrmeier et al., 1996, S. 679 ff.)

Auf dem zweiten Faktor weisen die Variablen „Suche nach Informationen über Waren und Dienstleistungen“, „Internet-Banking“ und „Verkauf von Waren und Dienstleistungen“ hohe Werte auf. Deshalb erhält dieser Faktor den Namen „Handelsfaktor“. Auf dem dritten Faktor laden vor allem die Variablen „Weiterbildung aus privatem Interesse“ und „Weiterbildung aus beruflichem Interesse“ hoch sowie – wenn auch nicht nur auf diesem Faktor – die Variable „Schul- oder Hochschulausbildung“. Dies ist der Grund dafür, den dritten Faktor „Bildungsfaktor“ zu nennen. Auf dem vierten Faktor weisen die Variablen „Häufigkeit der Internetnutzung“ und „Senden und Empfangen von E-Mails“ hohe Werte auf. Darum wird der vierte Faktor als „Erreichbarkeitsfaktor“ bezeichnet.

Die nachfolgend beschriebene Clusteranalyse erfolgt nicht mehr unmittelbar mit den größtenteils wechselseitig korrelierten ursprünglichen Variablen, sondern mit den per definitionem unkorrelierten Faktoren. Um die optimale Clusteranzahl zu bestimmen, wird abermals das

6. Verteilung der Personen auf die Cluster		
Cluster	Clustergröße	Hochgerechnete Anteile
	Anzahl	%
A	847	21,6
B	527	10,5
C	870	20,5
D	792	18,7
E	610	16,8
F	451	11,8

Bayes-Kriterium nach Schwarz berechnet. Als optimale Clusteranzahl wurde hierbei sechs ermittelt.

Tabelle 6 zeigt nur relativ geringe Unterschiede in der Clustergröße. Die Mittelwerte und Standardabweichungen der vier Faktoren in jedem Cluster sind in Tabelle 7 angegeben. Der größte Mittelwert ergibt sich beim Erreichbarkeitsfaktor in Cluster D. Er ist mit 1,739 sehr groß. Da die Variablen „Häufigkeit der Internetnutzung“ und „Senden und Empfangen von E-Mails“ mit einem Wert von jeweils fast 0,8 stark positiv auf dem vierten Faktor laden (siehe Tabelle 5), haben Personen mit einem hohen Erreichbarkeitsfaktor tendenziell auch einen hohen Wert bei den Variablen „Senden und Empfangen von E-Mails“ und „Häufigkeit der Internetnutzung“. Da diese Variablen mit 1 („ja“) und 2 („nein“) bzw. mit 1 („jeden Tag oder fast jeden Tag“ oder „mindestens einmal in der Woche“) und 2 („mindestens einmal im Monat“ oder „weniger als einmal im Monat“) codiert sind, bedeutet dies, dass Personen mit einem hohen Wert bei Faktor 4 eher nicht das Internet zum „Senden und Empfangen von E-Mails“ in den letzten drei Monaten genutzt haben und zudem eher angaben, das Internet nur „mindestens einmal“ oder sogar „weniger als einmal im Monat“ in den letzten drei Monaten genutzt zu haben. Zusammenfassend ist zu sagen, dass die Personen in Cluster D in

7. Hochgerechnete Mittelwerte und Standardabweichungen der Faktoren in den Clustern								
Cluster	Unterhaltungsfaktor		Handelsfaktor		Bildungsfaktor		Erreichbarkeitsfaktor	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
A	+0,738	+0,332	-0,256	+0,674	+0,826	+0,251	-0,582	+0,248
B	-0,526	+0,737	+1,494	+0,485	+0,028	+0,784	-0,687	+0,466
C	+0,688	+0,535	-0,255	+0,767	-1,087	+0,672	-0,395	+0,417
D	+0,375	+0,598	+0,340	+0,714	+0,323	+0,615	+1,739	+0,833
E	-0,966	+0,676	-0,652	+0,816	+0,889	+0,460	-0,294	+0,442
F	-1,390	+0,637	-0,785	+0,814	-1,083	+0,636	+0,089	+0,395

den letzten drei Monaten vor der Befragung also eher schlecht im Internet erreichbar waren. Es zeigt sich, dass dieser Cluster viele Personen aus der Gruppe der „Gelegenheits-User“ beinhaltet, welche nur sehr selten das Internet benutzen.

Cluster A hat einen relativ großen Mittelwert beim „Bildungsfaktor“ und beim „Unterhaltungsfaktor“. Dies bedeutet aufgrund der Codierung von „1“ für „ja“ und „2“ für „nein“, dass Personen in diesem Cluster das Internet seltener zur Unterhaltung oder auch Bildung nutzen als die Grundgesamtheit. Dieser neue Cluster entspricht zum größten Teil der Gruppe der „Basis-User“ aus der unmittelbaren Clusteranalyse. Des Weiteren sind für den ersten Cluster die Standardabweichungen in den vier Faktoren sehr gering. Dies bedeutet, dass dieser Cluster bezüglich der vier Faktoren homogener ist als die übrigen Cluster. Der Mittelwert für Cluster B hat beim „Handelsfaktor“ einen sehr hohen positiven und beim „Erreichbarkeitsfaktor“ sowie beim „Unterhaltungsfaktor“ einen stark negativen Wert. Die Personen in diesem Cluster benutzen das Internet also nur ganz selten zum „Handeln“, aber häufig zur Unterhaltung und für die Erreichbarkeit. Cluster C hat einen großen Mittelwert beim „Unterhaltungsfaktor“ und einen großen negativen Mittelwert beim „Bildungsfaktor“. Daher beinhaltet dieser Cluster viele Personen aus der Gruppe der „Wissbegierigen“. Bei Cluster E ist ein stark negativer Mittelwert beim „Unterhaltungsfaktor“ vorhanden und ein stark positiver Mittelwert beim „Bildungsfaktor“. Cluster F hat beim „Unterhaltungsfaktor“, „Handelsfaktor“ und „Bildungsfaktor“ einen stark negativen Wert.

Die Profile der neu gebildeten Cluster A bis F (vgl. Tabellen 8 und 9) lassen sich stark vereinfacht wie folgt beschreiben.

Cluster A:
durchschnittliche Nutzungsfrequenz, Nutzung vor allem für das Empfangen und Senden von E-Mails sowie zur Information über Waren, hoher Anteil von verheirateten Personen

8. Internetnutzung und -aktivitäten nach Clustern							
Internetnutzung Internetaktivität	Cluster						Internetnutzer insgesamt
	A	B	C	D	E	F	
	%						
Internetnutzung							
jeden Tag oder fast jeden Tag	55,1	60,8	73,2	11,3	83,6	92,6	60,4
mindestens einmal in der Woche	44,9	37,3	25,2	23,4	16,1	6,5	26,7
mindestens einmal im Monat	–	1,5	1,6	46,8	0,3	0,9	9,4
weniger als einmal im Monat	–	0,4	–	18,6	–	–	3,5
Internetaktivitäten							
Senden und Empfangen von E-Mails	100,0	94,4	98,2	26,6	98,2	99,3	84,9
Chatten/Besuch von Foren	18,0	84,4	30,9	17,5	74,9	78,9	44,3
Suche nach Informationen über Waren und Dienstleistungen	88,4	47,7	96,2	59,6	94,2	98,0	82,4
Nutzung von Internet-Radio/Internet-Fernsehen	2,1	14,7	6,9	4,5	45,0	73,8	20,5
Spielen/Herunterladen von Spielen und Musik	3,4	46,8	6,0	12,5	75,3	83,5	31,7
Herunterladen von Software	16,0	19,1	36,0	6,9	64,3	88,4	35,4
Lesen/Herunterladen von Internet-Zeitungen oder -Magazinen	13,7	8,2	38,5	10,9	28,2	77,1	27,6
Internet-Banking	44,7	1,8	55,5	11,9	54,2	65,2	40,3
Verkauf von Waren und Dienstleistungen	26,6	2,6	30,6	8,3	51,7	61,2	29,8
Schul- oder Hochschulausbildung	1,0	76,6	18,4	11,3	12,5	46,5	21,8
Weiterbildung aus privatem Interesse	2,7	47,8	84,9	14,6	14,1	91,8	39,0
Weiterbildung aus beruflichem Interesse	–	12,5	69,2	9,6	4,0	69,3	26,2

mittleren Alters (35 – 54 Jahre), überwiegend erwerbstätig

hoher Anteil an Schülerinnen und Schülern sowie Studierenden

Cluster B:
durchschnittliche Nutzungsfrequenz, Nutzung vor allem für das Chatten bzw. für Foren sowie für die schulische Bildung, größtenteils unter 25 Jahren sowie überwiegend weiblich,

Cluster C:
durchschnittliche bis überdurchschnittliche Nutzungsfrequenz, Nutzung vor allem für das Empfangen und Senden von E-Mails sowie zur Information über Waren und Inter-

9. Soziodemografische Merkmale der Internetnutzer nach Clustern							
Merkmal	Cluster						Internetnutzer insgesamt
	A	B	C	D	E	F	
	%						
Männlich	47,1	40,0	52,3	46,0	62,7	76,6	48,6
Weiblich	52,9	60,0	47,7	54,0	37,3	23,4	51,4
Ledig	21,3	86,6	34,7	38,3	49,4	63,9	33,9
Verheiratet	69,7	8,3	53,3	51,2	44,3	32,6	51,2
Verwitwet	1,8	0,9	1,4	3,1	0,9	1,2	7,0
Geschieden	7,2	4,3	10,6	7,4	5,5	2,3	7,9
Alter von ... bis ... Jahren							
10 – 15	1,6	39,2	0,4	17,5	8,9	4,5	7,3
16 – 24	6,5	43,7	8,3	8,5	22,4	29,1	11,4
25 – 34	16,5	4,9	21,2	10,7	22,9	27,1	13,1
35 – 44	31,1	3,8	29,1	21,3	25,5	22,2	18,8
45 – 54	24,4	3,3	23,6	19,7	12,6	8,4	15,5
55 – 64	13,9	3,3	11,1	13,3	5,9	6,3	13,0
65 – 74	4,9	1,8	5,1	7,3	1,8	1,4	12,3
75 und mehr	1,2	–	1,1	1,8	–	1,0	8,6
Hauptschul- bzw. Volksschulabschluss	31,4	18,8	17,9	42,2	31,5	14,4	41,4
Realschul- bzw. gleichwertiger Abschluss	34,7	46,7	28,9	30,1	38,0	28,5	29,6
Fachhochschulreife	14,2	10,3	13,1	13,4	12,1	15,1	10,1
Abitur	19,8	24,2	40,0	14,3	18,4	41,9	18,9
Erwerbstätige	74,9	28,5	73,8	57,7	64,3	65,8	50,5
Arbeitslose	4,4	6,9	6,2	10,3	9,2	8,6	6,5
Schüler/-innen	0,5	41,0	1,2	2,8	6,6	6,3	3,8
Studierende	0,7	10,5	4,0	0,3	4,7	11,0	2,6
Übrige Personen	19,5	13,0	14,8	29,0	15,2	8,2	36,6

10. Verteilung der Personen in Cluster nach beiden Analysen						
Einteilung nach der anhand der Faktoren durchgeführten Clusteranalyse		Einteilung nach der unmittelbar anhand der Merkmale durchgeführten Clusteranalyse				
		1	2	3	4	5
A	Anzahl	–	27	15	709	96
	% ¹⁾	–	3,3	2,1	83,1	11,5
B	Anzahl	1	77	316	–	133
	% ¹⁾	0,1	13,4	58,4	–	28,2
C	Anzahl	44	700	9	83	34
	% ¹⁾	5,6	80,9	1,1	9,1	3,3
D	Anzahl	–	7	10	7	768
	% ¹⁾	–	1,0	1,4	0,6	97,1
E	Anzahl	207	29	267	98	9
	% ¹⁾	35,6	4,8	41,5	16,6	1,5
F	Anzahl	340	38	73	–	–
	% ¹⁾	76,2	8,7	15,1	–	–

1) hochgerechnete Anteile in %

net-Banking sowie schulische und private Bildung, hohes Bildungsniveau, überwiegend erwerbstätig

Cluster D:

eher seltene Nutzung, oft ohne Senden und Empfangen von E-Mails, eher niedriges Bildungsniveau

Cluster E:

häufige Nutzung mit großer Vielfalt an Aktivitäten – wenn auch weniger ausgeprägt als Cluster F, eher männlich, Konzentration auf das Alter zwischen 16 und 45 Jahren, viele Parallelen zu Cluster F, im Vergleich dazu aber niedrigeres Bildungsniveau und weniger Computer-Nutzung für das Lesen von Zeitungen sowie für Bildungszwecke

Cluster F:

häufige Nutzung mit großer Vielfalt an Aktivitäten insbesondere auch für berufliche und private Weiterbildung, hierin liegt der Hauptunterschied zu Cluster E, ähnlich Cluster E überwiegend männlich, Konzentration auf das Alter zwischen 16 und 45 Jahren, aber im Vergleich dazu wesentlich höheres Bildungsniveau

Im Vergleich zur unmittelbar durchgeführten Clusteranalyse zeigen sich viele Parallelen. Tabelle 10 macht die Zusammenhänge zwischen den beiden unterschiedlich ermittelten Clusterzuordnungen deutlich. Zu erkennen ist, dass Cluster A weitgehend identisch ist mit den zuvor ermittelten „Basis-Usern“ (Cluster 4), Cluster C mit den „Wissbegierigen“ (Cluster 2) und Cluster D aus den „Gelegenheitsnutzern“ (Cluster 5) besteht. Die Strukturen dieser Cluster bleiben

also weitgehend erhalten. Die Namen können daher auch zur Beschreibung der entsprechenden Cluster der zweiten Analyse beibehalten werden.

Die Gruppe der „Power-User“ (Cluster 1) verteilt sich dagegen auf die neuen Cluster E und F und diejenige der „Entertainment-User“ (Cluster 3) auf die Cluster B und E. Die neuen Cluster B, E und F lassen sich damit nicht eindeutig einer Gruppe der zuvor beschriebenen unmittelbaren Clusteranalyse zuordnen. Die Merkmalsausprägungen der Personen in diesen Clustern werden im Folgenden unmittelbar gegenübergestellt.

Wie in Tabelle 11 zu sehen ist, nannten die Personen in Cluster B zu 81,9 % die Aktivität „Chatten/Besuch von Foren“. Dieser Wert ist fast doppelt so hoch wie derjenige der Internetnutzer insgesamt. Zudem haben die Personen in der Gruppe B zu 80,6 % das Internet „für Schul- oder Hochschulausbildung“ genutzt. Im Verhältnis zur Gesamtstichprobe relativ selten wurden die Aktivitäten „Herunterladen von Software“, „Lesen/Herunterladen von Internet-Zeitungen oder Magazinen“, „Internet-Banking“, „Verkaufen von Waren und Dienstleistungen“ und „Weiterbildung aus beruflichem Interesse“ ausgeübt. Da sowohl die Aktivität „Senden und Empfangen von E-Mails“ als auch „Chatten/Besuch von Foren“ häufig angegeben wurde, werden die Personen in Cluster B als „Kommunikations-User“ bezeichnet. Der neue Cluster E hat bei fast jeder Aktivität im Vergleich zur Gesamtstichprobe erhöhte Werte. Lediglich für „Schul- oder Hochschulausbildung“, „Weiterbildung aus privatem Interesse“ sowie für „Weiterbildung aus beruflichem Interesse“ scheint der User in Cluster E das Internet seltener ge-

11. Internetnutzung und -aktivitäten nach Clustern – Vergleich der Cluster 1 und 3 mit den Clustern B, E und F						
Internetnutzung Internetaktivität	Cluster					Internetnutzer insgesamt
	3 (Entertainment)	B (Kommunikation)	1 (Power-User)	E (Power-User ohne Bildungsinteresse)	F (Power-User mit Bildungsinteresse)	
	%					
Internetnutzung						
jeden Tag oder fast jeden Tag	76,4	60,8	96,2	83,6	92,6	60,4
mindestens einmal in der Woche	21,7	37,7	3,8	16,1	6,5	26,7
mindestens einmal im Monat	1,2	1,5	–	0,3	0,9	9,4
weniger als einmal im Monat	0,8	0,4	–	–	–	3,5
Internetaktivitäten						
Senden und Empfangen von E-Mails	98,2	94,4	99,7	98,2	99,3	84,9
Chatten/Besuch von Foren	88,2	81,9	76,7	74,9	78,9	44,3
Suche nach Informationen über Waren und Dienstleistungen	76,1	47,7	99,5	94,2	98,0	82,4
Nutzung von Internet-Radio/ Internet-Fernsehen	32,9	14,7	62,9	45,0	73,8	20,5
Spielen/Herunterladen von Spielen und Musik	82,7	46,8	87,0	75,3	83,5	31,7
Herunterladen von Software	53,1	19,1	84,4	64,3	88,4	35,4
Lesen/Herunterladen von Internet-Zeitungen oder -Magazinen	22,5	8,2	58,5	28,2	77,1	27,6
Internet-Banking	2,1	1,8	89,2	54,2	65,2	40,3
Verkauf von Waren und Dienstleistungen	18,3	2,6	71,6	51,7	61,2	29,8
Schul- oder Hochschulausbildung	47,4	76,6	30,2	12,5	46,5	21,8
Weiterbildung aus privatem Interesse	35,1	47,8	65,7	14,1	91,8	39,0
Weiterbildung aus beruflichem Interesse	11,2	12,5	50,6	4,0	69,3	26,2

12. Verteilung der Personen in Cluster nach beiden Analysen							
Einteilung nach der unmittelbar anhand der Merkmale durchgeführten Clusteranalyse		Einteilung nach der anhand der Faktoren durchgeführten Clusteranalyse					
		A	B	C	D	E	F
1	Anzahl	–	1	44	–	207	340
	% ¹⁾	–	0	7,1	–	37,0	55,8
2	Anzahl	27	77	700	7	29	38
	% ¹⁾	3,4	6,8	80,0	0,9	3,9	5,0
3	Anzahl	15	318	9	10	267	73
	% ¹⁾	2,9	38,8	1,5	1,6	44,0	11,3
4	Anzahl	709	–	83	7	98	–
	% ¹⁾	79,0	–	8,3	0,5	12,3	–
5	Anzahl	96	133	34	768	9	–
	% ¹⁾	10,2	12,1	2,7	74,0	1,0	–

1) hochgerechnete Anteile in %

nutzt zu haben als die Personen in der gesamten Quotenstichprobe. Deshalb wäre eine mögliche Bezeichnung dieser Personen „Power-User ohne Bildungsinteresse“. Beim Vergleich der Cluster E und F fällt auf, dass vor allem bei den Aktivitäten „Lesen/Herunterladen von Internet-Zeitungen oder Magazinen“, „Schul- und Hochschulausbildung“, „Weiterbildung aus privatem Interesse“ und „Weiterbildung aus beruflichem Interesse“ Cluster F höhere Werte aufweist. Deshalb werden die User in diesem Cluster „Power-User mit Bildungsinteresse“ genannt. Beim Vergleich der Cluster 3 und B fällt auf, dass die größten Unterschiede bei den Aktivitäten „Spielen/Herunterladen von Spielen und Musik“ und „Herunterladen von Software“ liegen. Der vorherige Cluster der „Entertainment-User“ ist in der zweiten Clusteranalyse nicht mehr nachweisbar, die dazugehörigen Personen verteilen sich vor allem auf die neuen Cluster der „Kommunikations-User“ und der „Power-User ohne Bildungsinteresse“ (vgl. Tabelle 12).

Zusammenfassung

Kernstück der Analyse ist die Durchführung einer Two-Step-Clusteranalyse, die Personengruppen zusammenfasst, die eine ähnliche Internetnutzung aufweisen. Diese Gruppenbildung wird anhand der Häufigkeit der Internetnutzung sowie der in den

drei Monaten vor der Befragung im Internet zu privaten Zwecken durchgeführten Aktivitäten vorgenommen. Ergebnis dieser unmittelbaren Clusteranalyse sind fünf verschiedene Cluster, deren Objekte man wegen ihrer Merkmalsausprägungen als „Power-User“, „Wissbegierige“, „Entertainment-User“, „Basis-User“ und „Gelegenheits-User“ bezeichnen kann.

In einer weiteren Analyse werden die für die Clusteranalyse relevanten Variablen mittels einer Faktorenanalyse zu Faktoren gebündelt, die bedingt durch das Rotationsverfahren voneinander unabhängig sind. Mittels einer Hauptkomponentenanalyse wurden vier Faktoren extrahiert, der „Unterhaltungsfaktor“, der „Handelsfaktor“, der „Bildungsfaktor“ und der „Erreichbarkeitsfaktor“. Mit diesen Faktoren wurde abermals eine Clusteranalyse durchgeführt, die nun sechs Cluster erzeugt.

Trotz unterschiedlicher Vorgehensweise kommen beide Analysen zu weitgehend gleichen Ergebnissen. Die unmittelbar aus den Ausgangsdaten durchgeführte Clusteranalyse wird damit vom Ergebnis her durch das methodisch fundiertere aber aufwändigere Verfahren einer Clusteranalyse mit vorgeschalteten Hauptkomponentenanalyse bestätigt. Damit erweist sich dieses Verfahren als robust gegenüber einer Verletzung der Unabhängigkeitsannahme in Bezug auf die verwendeten Ausgangs-

variablen. Die einzigen nennenswerten Unterschiede bestehen in der Herausbildung eines „Entertainment-User“ bei der einen anstatt eines „Kommunikations-Users“ bei der anderen Klassifikation sowie in der Bildung von ein oder zwei Clustern von „Power-Usern“. Der bei der Hauptkomponentenanalyse extrahierte „Bildungsfaktor“ führt bei der auf den Faktoren aufbauenden Clusteranalyse zu einer Teilung der Gruppe der „Power-User“ im Hinblick auf ein unterschiedliches Interesse an einer Nutzung des Internets zur persönlichen oder beruflichen Weiterbildung.

Literatur und weitere Quellen

Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; Weiber, Rolf (2000). *Multivariate Analysemethoden – Eine anwendungsorientierte Einführung*. 9. Auflage, Springer, Berlin, Heidelberg

Bühl, Achim; Zöfel, Peter (1998). SPSS für Windows Version 7.5: *Praxisorientierte Einführung in die moderne Datenanalyse*. 4. Auflage, Addison-Wesley, Bonn Heidelberg

Fahrmeir, Ludwig; Hamerle, Alfred; Tutz, Gerhard (1996). *Multivariate statistische Verfahren*. 2. Auflage, Walter de Gruyter, Berlin, New York

„Erläuterungen zum PHI-Koeffizienten“, quantlet.com, 22.09.2007 http://www.quantlet.com/mdstat/scripts/estat_zko/phi/estat/bpreview/003_phikoeffizient.html

„SPSS 13.0 Base Benutzerhandbuch“, uni-hamburg.de, 7. 9. 2007 <http://www1.uni-hamburg.de/RRZ/Software/SPSS/ManualsGer.130/BaseUsersGuide130.pdf>



Index 2006 – 2008

Ausgaben des Jahres 2008

- Band 51**
Z 08 1 2008 53
3,00 EUR
- Das Forschungsdatenzentrum der Statistischen Landesämter
– eine Serviceeinrichtung für die Wissenschaft**
Sylvia Zühlke, Helga Christians und Katharina Cramer – Geschäftsstelle des Forschungsdatenzentrums der Statistischen Landesämter
- Behandlungs- und Wohnortkonzept in der Diagnosestatistik**
Diplom-Volkswirt Hans-Jürgen Treeck
- Clusteranalyse zur Identifizierung von verschiedenen Gruppen von Internetnutzern**
Christian Hoops, Diplom-Statistiker Josef Schäfer
- Band 50**
Z 08 1 2008 52
3,00 EUR
- Kleinräumige Berechnung der Kaufkraft: Ein Kaufkraftmodell für die Gemeinden in NRW**
Dr. Marco Scharmer, Jan Pollmann
- Regionalisierung der Arbeitskosten
– ein statistisches Angebot von wirtschaftlichem Interesse?**
Diplom-Volkswirt, Diplom-Kaufmann Hermann Marré
- Band 49**
Z 08 1 2008 51
3,50 EUR
- Auswirkungen des demografischen Wandels
Modellrechnung für den Bereich Pflegebedürftigkeit – Aktualisierte Ergebnisse**
Dr. Kerstin Ströker

Ausgaben des Jahres 2007

- Band 48**
Z 08 1 2007 59
3,00 EUR
- LEBEN IN EUROPA 2005: Armutsgefährdung und Lebensbedingungen in NRW**
Diplom-Sozialwissenschaftlerin Karin Habenicht
- Neue Zuwanderergruppen in NRW
– sozialstrukturelle Merkmale und Arbeitsmarktintegration**
Dr. Wolfgang Seifert
- Band 47**
Z 08 1 2007 58
3,00 EUR
- Bildungsreport NRW 2007: Informationen zu ausgewählten Bildungsbereichen**
Bettina Lander M. A.
- Band 46**
Z 08 1 2007 57
3,50 EUR
- Cross-SIS: Zum Aufbau einer Web-GIS-Anwendung
für den Bereich Statistik auf europäischer Ebene**
Diplom-Soziologin Kerstin Schmidtke, MPH
- Methodeneffekte bei Online-Erhebungen:
Ergebnisse aus den laufenden Wirtschaftsberechnungen privater Haushalte**
Diplom-Soziologin Anke Gerhardt, Urszula Sikorski
- Band 45**
Z 08 1 2007 56
3,00 EUR
- Wertschöpfungskette im Wandel – von der Landwirtschaft zum Verbraucher**
Dr. Guido Recke
- Entwicklungstendenzen im Gemüseanbau**
Dr. Guido Recke
- Band 44**
Z 08 1 2007 55
3,00 EUR
- Auswirkungen der Revision 2005 der Volkswirtschaftlichen Gesamtrechnungen
auf die Indikatorenbildung in den Umweltökonomischen Gesamtrechnungen der Länder**
Dr. Olivia Martone, Michael Hinz
- Schätzverfahren zur Bodenversiegelung: UGRdL-Ansatz**
Diplom-Geografin Britta Frie, Diplom-Volkswirt Ralph Hensel
- Band 43**
Z 08 1 2007 54
3,50 EUR
- Die Einkommenssituation der privaten Haushalte in den Gemeinden
Nordrhein-Westfalens**
Dr. Marco Scharmer
- Zusammenhang und Entwicklung von Wirtschaftswachstum, Investitionen
und Beschäftigung im Produzierenden Gewerbe und Dienstleistungsbereich
in Nordrhein-Westfalen**
PD Dr. Jörg-Peter Schräpler

noch: **Ausgaben des Jahres 2007**

Band 42
Z 08 1 2007 53
2,50 EUR
Modellierung von Szenarien zur zukünftigen Entwicklung der Bevölkerung mit Migrationshintergrund in NRW
Dr. Kerstin Ströker

Auf dem Weg in das Berufsleben: Auszubildende im dualen Ausbildungssystem
Bettina Lander M. A.

Band 41
Z 08 1 2007 52
2,50 EUR
Gesundheit und Erwerbsbeteiligung
Diplom-Statistikerin Maria Frese, Norbert Merschmann

Aufwendungen und Investitionen in Informations- und Kommunikationstechnologien 2004
Heike Schrankel, Diplom-Statistiker Josef Schäfer

Band 40
Z 08 1 2007 51
2,50 EUR
Materialeinsatz und Rohstoffverbrauch in Nordrhein-Westfalen
Dr. Leontine von Kulmiz

Ausgaben des Jahres 2006

Band 39
Z 08 1 2006 63
2,50 EUR
Gesundheit und Arztbesuche – Erste Ergebnisse aus EU-SILC – LEBEN IN EUROPA 2005
Diplom-Sozialwissenschaftlerin Karin Habenicht

Statistische Quellen mit Gesundheitsdaten für NRW
Diplom-Volkswirt Hans-Jürgen Treeck

Band 38
Z 08 1 2006 62
8,50 EUR
**Auswirkungen des demografischen Wandels
Aktualisierte Ergebnisse der Modellrechnungen für die Bereiche Haushalte,
Erwerbspersonen und Pflegebedürftigkeit**
Diplom-Soziologin Kerstin Schmidtke, MPH; Diplom-Mathematiker Paul Berke

Band 37
Z 08 1 2006 61
2,50 EUR
Wie Ausländer und Deutsche Unternehmen in NRW gründen wollen
Diplom-Volkswirtin Doris Blechinger

Verbesserung der Ergebnisqualität bei der Fortschreibung des Bruttoinlandsprodukts
Dr. Marco Scharmer

Band 36
Z 08 1 2006 60
7,00 EUR
Auswirkungen des demografischen Wandels – Teil II
Diplom-Soziologin Kerstin Schmidtke, MPH; Diplom-Statistikerin Diana Andrä

Band 35
Z 08 1 2006 59
2,50 EUR
**Zur Möglichkeit der Regionalisierung privater Einkommen
auf die Gemeinden Nordrhein-Westfalens**
Dr. Marco Scharmer

**Einkommensverteilung und Armutsrisikoquoten im Bund-Länder-Vergleich
– Analysen mit dem Mikrozensus 2003**
Diplom-Sozialwissenschaftlerin Karin Habenicht, Diplom-Sozialwissenschaftler Alfred Hullmann, Yvonne Bergmann

Band 34
Z 08 1 2006 58
2,50 EUR
**Strukturinformationen zum Arbeitsmarkt 3. Quartal 2005
Erwerbsverhalten älterer Menschen**
Diplom-Sozialwissenschaftlerin Karin Habenicht

Band 33
Z 08 1 2006 57
2,50 EUR
Auswirkungen des demografischen Wandels – Kurzfassung – Teil II
Diplom-Soziologin Kerstin Schmidtke, MPH

Band 32
Z 08 1 2006 56
3,00 EUR
Bildungsreport NRW 2006: Informationen zu ausgewählten Bildungsbereichen
Bettina Lander M. A.

Die Arbeitsmarktintegration von jungen Erwachsenen 2004
Dr. Wolfgang Seifert

Band 31
Z 08 1 2006 55
3,00 EUR
**Vorausberechnung der Bevölkerung in den kreisfreien Städten und Kreisen
Nordrhein-Westfalens 2005 bis 2025/2050**
Dr. Kerstin Ströker

Band 30
Z 08 1 2006 54
3,00 EUR
Bedeutung der Türkei für den Außenhandel des Landes Nordrhein-Westfalen
Diplom-Volkswirt Nils Radmacher-Nottelmann

noch: **Ausgaben des Jahres 2006**

Band 29
Z 08 1 2006 53
3,00 EUR

**Datenangebot und Datenzugang im Forschungsdatenzentrum
der Statistischen Landesämter**

Dr. Sylvia Zühlke, Diplom-Sozialwissenschaftlerin Helga Christians

**Das Studienkontenmodell an den Hochschulen in Nordrhein-Westfalen:
Auswirkungen auf die Studierendenzahlen**

Birgit Kempf, Anne Schopp, Anja Gehrcken

Nachfrage nach Dienstleistungen 2003

Heike Schrankel, Diplom-Statistikerin Katharina Götz

**Die Arbeitsmarktintegration von jungen Erwachsenen
mit Migrationshintergrund in NRW**

Dr. Wolfgang Seifert

Band 28
Z 08 1 2006 52
3,00 EUR

**Gründungen und Schließungen in Deutschland
– Ein Konzept zur demografischen Auswertung des Unternehmensregisters**

Diplom-Kauffrau Ingrid Kaack

Strukturelle Entwicklung der Landwirtschaft von 1971 bis 2003

Diplom-Volkswirt Hartmut Kopmeier

Wirtschaftliche Aspekte der Studierenden 1996 und 2004

Lars Kuhlmeier, Dr. Wolfgang Seifert

Verwendung von Verwaltungsdaten für konjunkturstatistische Zwecke

Diplom-Kauffrau Ingrid Kaack

**Personal und Finanzen der öffentlich bestimmten Fonds, Einrichtungen,
Betriebe und Unternehmen (FEU) in privater Rechtsform im Jahr 2003**

Diplom-Volkswirt Peter Emmerich

Band 27
Z 08 1 2006 51
3,00 EUR

**Der Weg zu einem ersten registergestützten Zensus
in Deutschland im Jahre 2011**

Diplom-Volkswirt Helmut Eppmann, Diplom-Statistiker Josef Schäfer

**Ergänzende Strukturmerkmale zur Erwerbstätigenrechnung
– Unterscheidung nach „Vollzeit-“ und „Nicht-Vollzeiterwerbstätigen“**

Diplom-Ingenieur Hans Menge, Miriam Sperl

**Wirtschaftskraft und Wirtschaftsentwicklung
in den kreisfreien Städten und Kreisen Nordrhein-Westfalens**

Dr. Marco Scharmer