

Sequenzmusteranalyse

SOZIALWISSENSCHAFTLICHE FORSCHUNGSMETHODEN

herausgegeben von

Wenzel Matiaske, Martin Spieß,
Ingwer Borg, Claudia Fantapié-Altobelli, Holger Hinz,
Uwe Jirjahn, Bernhard Kittel, Manfred Kraft,
Stefan Liebig, Rainer Oesterreich, Jost Reinecke,
Kai-Uwe Schnapp, Rainer Schnell, Peter Sedlmeier,
Winfried Seidel, Gerhard Tutz, Joachim Wagner

Band 5

Michael Stegmann, Julia Werner, Heiko Müller

Sequenzmusteranalyse

Einführung in Theorie und Praxis

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-86618-829-7 (print)

ISBN 978-3-86618-929-4 (e-book)

SOZIALWISSENSCHAFTLICHE FORSCHUNGSMETHODEN: ISSN 1869-7151

DOI 10.1688/9783866189294

1. Auflage, 2013

© 2013 Rainer Hampp Verlag München und Mering
Marktplatz 5 D – 86415 Mering
www.Hampp-Verlag.de

Alle Rechte vorbehalten. Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne schriftliche Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Mikroverfilmungen, Übersetzungen und die Einspeicherung in elektronische Systeme.

∞ *Dieses Buch ist auf säurefreiem und chlorfrei gebleichtem Papier gedruckt.*

Liebe Leserinnen und Leser!

Wir wollen Ihnen ein gutes Buch liefern. Wenn Sie aus irgendwelchen Gründen nicht zufrieden sind, wenden Sie sich bitte an uns.

Vorwort

Das vorliegende Buch wendet sich an Studierende und Praktiker im Bereich der Sozialwissenschaften, die einen Einblick in die Grundlagen und in die Technik der Sequenzmusteranalyse erhalten wollen. Als Autoren ist uns dabei vor allem daran gelegen, die Leserin bzw. den Leser in die Lage zu versetzen, mit aktuellen Softwarelösungen selbst eine Sequenzmusteranalyse durchführen zu können. Deswegen haben wir uns bemüht, die einzelnen Analyseschritte, aber auch die beschriebenen Maßzahlen nicht nur zu präsentieren, sondern auch zu erklären, aber dennoch nicht mit Herleitungen zu überfrachten.

Ob dies gelungen ist, muss am Ende die geneigte Leserschaft entscheiden. Unser besonderer Dank für die kritische Durchsicht, die zahlreichen Hinweise und Anmerkungen zur Erklärung der Verfahren und der dargestellten Maßzahlen gilt an dieser Stelle ganz besonders Jan Werner.

Einleitung

„Die Menschen früherer Zeiten waren genauso schlecht wie wir. Sie wussten es nur nicht so genau.“ Dieser Satz von Tennessee Williams bringt die Intention sozialwissenschaftlicher Forschung auf den Punkt, die Frage des Verhältnisses von Gesellschaft und Individuum. Die empirische Sozialforschung kümmert sich um diese facettenreiche Fragestellung auf vielen Wegen und hat sich dazu einen methodologischen Überbau und methodischen Werkzeugkoffer geschaffen, der vielfältig sortiert ist. Es existiert eine Vielzahl an Techniken zur Erhebung und Auswertung von Daten.

Eine verstehende und erklärende Sozialwissenschaft will aber nicht nur soziale Unterschiede erkennen, sondern verstehen und Erklärungen finden. Früh war damit ein handlungsleitendes Interesse verbunden, insbesondere in der durch Durkheim (1858-1917) geprägten französischen Soziologie. Ihm war bereits klar, dass dies nicht funktioniert, wenn man die Sozialisation des Menschen außer Acht lässt. Ein Begriff, der maßgeblich von ihm geprägt wurde. Diese biografische Perspektive bringt er in seinem Werk „De la division du travail social“ zum Ausdruck:

„Zweifellos sind wir von Geburt aus nicht für einen ganz bestimmten Beruf vorgesehen; aber wir haben Neigungen und Fähigkeiten, die unsere Wahl begrenzen. Wird Ihnen nicht Rechnung getragen, so werden sie durch unsere täglichen Beschäftigungen ständig verletzt; wir leiden und suchen ein Mittel, um unser Leiden zu beenden (...)“ (Durkheim 1988: 444).

Ist es also die Straße, in der wir gelebt haben, die bestimmt, welchen Kindergarten wir besuchen? Ist es der Kindergarten, der bestimmt, in welche Schule wir gehen? Ist es die Schule, die uns in den Beruf führt? Oder wird all das von Beginn durch grundlegende Lebenschancen festgelegt, welche die Richtung unseres Weges durch die Institutionen und das Erwerbsleben determinieren? Fragen, die sich die Sozialwissenschaften schon immer gestellt haben. Die Sequenzmusteranalyse bzw. das Optimal Matching in Verbindung mit typisierenden Verfah-

ren liefert dazu einen Untersuchungsansatz, der den Grad der Verschiedenheit von Lebensläufen ins Zentrum rückt. Im Mittelpunkt des Forschungsinteresses steht der Lebensweg. Die unterschiedlichen Lebenswege sollen als Ganzes analysiert und eventuelle Lebenswegmuster entwickelt werden, um soziale Konstellationen und Unterschiede nicht nur zu erkennen, sondern auf die Biografie zurückspiegeln zu können. Die Biografie oder der Lebensweg sind dabei als Spiegelbild der Sozialisation zu verstehen.

Für diese Art der empirischen Analyse werden detaillierte Informationen über die Biografien benötigt. Dass diese für quantitative Verfahren vorliegen, ist allerdings keine Selbstverständlichkeit. Mit den EDV-technischen Möglichkeiten wuchsen das Interesse und die Möglichkeiten, solche Daten zu sammeln und auszuwerten. Eine Pionierstellung in Deutschland nimmt dabei das 1984 begründete Sozioökonomische Panel ein, das weite Verbreitung in den Sozialwissenschaften gefunden hat. Eine andere Quelle sind die Daten der Sozialversicherung, insbesondere der Rentenversicherung, die seit einigen Jahren verfügbar sind. Diese Versicherungsverläufe halten tagesgenaue biografische Informationen bereit. Und aktuell erfährt der Aufbau von Panels einen großen Zuspruch in den Sozialwissenschaften.

Das vorliegende Buch setzt sich zum Ziel eine bestimmte Herangehensweise der Analyse von Längsschnittdaten zu erörtern, die so genannte Sequenzmusteranalyse, insbesondere in Form des Optimal-Matching-Verfahrens. Damit die Leserin bzw. der Leser in die Lage versetzt werden kann, die Grundlagen dieses Verfahrens nicht nur zu verstehen, sondern selbst anwenden zu können, gliedert sich das Buch in vier Teile. Zu Anfang werden kurz die Möglichkeiten der Analyse von Längsschnittdaten skizziert und die besondere Form der Aufbereitung von Sequenzdaten besprochen. Darauf aufbauend werden die deskriptiven Analysemöglichkeiten behandelt. Dem folgen die Darlegung des Optimal Matching Verfahrens und die Schilderung der darauf aufbauenden Typisierung mittels Clusteranalyse. Ein besonderes Augenmerk wird darauf gerichtet, dass neben dem Verständnis der Verfahren auch die technischen Durchführungswege erklärt werden. Der Schwerpunkt der Darstellung liegt dabei auf den Analysemethoden und Maßzahlen, die in den gängigen Softwarelösungen implementiert sind. In den einzelnen Kapiteln wird dann ausführlich erklärt, wie das konkrete Vorgehen mit R zu bewerkstelligen ist. Eingangs wird dazu auch ein Abriss über die drei wesentlichen Softwarelösungen (TDA, STATA, R) gegeben und beschrie-

ben, wie jeweils vorzugehen ist. Abgeschlossen wird das Buch mit einer kurzen Auseinandersetzung zu den Möglichkeiten und den Grenzen des Verfahrens.