

Computerunterstützte Textanalysen mit TextQuest

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 6

Harald Klein

Computerunterstützte Textanalysen mit TextQuest

Eine Einführung in
Methoden und Arbeitstechniken

Rainer Hampp Verlag

München und Mering 2013

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 978-3-86618-831-0 (print)

ISBN 978-3-86618-931-7 (e-book)

SOZIALWISSENSCHAFTLICHE FORSCHUNGSMETHODEN: ISSN 1869-7151

DOI 10.1688/9783866189317

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.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Für wen dieses Buch ist – und für wen nicht	1
1.2	Die Beispieldaten	2
1.3	Historisches	3
1.4	Definitionen	5
1.4.1	Inhaltsanalyse	5
1.4.2	Computerunterstützte Inhaltsanalyse	8
1.5	Textanalytische Verfahren	12
1.5.1	Hermeneutische Textanalyse	12
1.5.2	Qualitative Datenanalyse (QDA)	13
1.6	Exkurs: Texte, Zeichensätze und Encodierungen	14
1.7	Software für computerunterstützte Inhaltsanalysen	15
2	Aufbereiten der Texte	17
2.1	Scannen von Texten	18
2.2	Fotografieren von Texten	20
2.3	Diktieren von Texten	21
2.4	Texte aus Textsatz- oder Textverarbeitungssystemen	21
2.5	Texte von Internetseiten und Websites	22
2.6	Texte in digitaler Form: Internet, PDF	23
2.7	Texte aus Statistiksoftware - Antworten auf offene Fragen	23
2.8	Texte aus Textarchiven	24
2.9	Probleme bei der Digitalisierung der Texte	24
3	Erzeugen einer Systemdatei	29
3.1	Texteinheiten und externe Variablen	29
3.1.1	Definition der Texteinheit	29
3.1.2	Definition von externen Variablen	30
3.1.3	Beispiele für die Definition von Texteinheiten und externen Variablen	32
3.2	Kennzeichnen von Texteinheiten und externen Variablen	35

3.2.1	Konventionen bei der Eingabe mit variablem Format ..	37
3.2.2	Konventionen für die Eingabe mit Spaltenformat	43
3.2.3	Generierung einer Systemdatei mit TEXTQUEST	44
4	TEXTQUEST im Überblick	47
4.1	Rundreise durch das System	47
4.2	Das Dateikonzept von TEXTQUEST	48
5	Vokabulare von Texten	51
5.1	Definition von Stichproben	54
5.2	Wörterlisten	56
5.2.1	Informationsmeldungen	58
5.2.2	Ausdruck einer Wörterliste (normale Form)	60
5.2.3	Ausdruck einer rückläufigen Wörterliste	61
5.3	Wortsequenzen	62
5.3.1	Möglichkeiten des Programms	63
5.3.2	Informationsmeldungen	64
5.3.3	Ausdruck von Wortsequenzen	65
5.4	Wortpermutationen	66
5.4.1	Möglichkeiten des Programms	67
5.4.2	Informationsmeldungen	68
5.4.3	Ausdruck von Wortpermutationen	69
5.5	Vergleiche von Vokabularen	70
5.5.1	Möglichkeiten des Programms	71
5.5.2	Informationsmeldungen	72
5.6	Kreuzreferenzen	73
5.6.1	Möglichkeiten des Programms	74
5.6.2	Informationsmeldungen	75
5.6.3	Ausdruck von Kreuzreferenzen	76
5.7	TTR-Verlauf	77
5.7.1	Möglichkeiten des Programms	77
5.7.2	Informationsmeldungen	78
5.7.3	Ausdruck eines TTR-Verlaufs	80
5.8	Sortieren von Vokabularen und Konkordanzen	81
5.9	Definition von Suchbegriffen	82
5.9.1	Zeichenketten	83
5.9.2	Angaben im Parameterfeld	85
5.9.3	Wortstammfolgen	85
6	Inhaltsanalyse	87
6.1	Erstellen und Bearbeiten eines Kategoriensystems – der Kategorienmanager	88
6.2	Standardisierte Kategoriensysteme	92
6.3	Linguistische Probleme der Kodierung	93
6.3.1	Kontext	93

6.3.2	Mehrdeutigkeit	94
6.3.3	Negation	94
6.4	Validierung eines Kategoriensystems	95
6.5	Ergebnisse einer Inhaltsanalyse	98
6.5.1	Numerische Ergebnisse	99
6.5.2	Protokolldateien	99
6.6	Durchführung einer Inhaltsanalyse	100
6.6.1	Parameter und Optionen	100
6.6.2	Interaktive Kodierung	103
6.7	Validierung der Kodierung	105
6.7.1	Kodierkontrolle	105
6.7.2	Konkordanzen	106
6.7.3	Suchbegriffe in Texteinheit	107
6.8	Weiterverarbeitung der Ergebnisse	107
7	Lesbarkeitsanalyse	109
7.1	Verarbeitung spezieller Wörterlisten	110
7.2	Definitionen	111
7.3	Möglichkeiten des Programms	112
7.4	Informationsmeldungen	113
8	Sonstige Funktionen	117
8.1	Der Dateimanager – Sichern und Wiederherstellen von Dateien	117
8.2	Betrachten der Ergebnisse	118
8.3	Ausblick	119
	Literaturverzeichnis	121
	Index	123

SOZIALWISSENSCHAFTLICHE FORSCHUNGSMETHODEN

Ingwer Borg, Patrick J.F. Groenen, Patrick Mair:

Multidimensionale Skalierung

Band 1, ISBN 978-3-86618-438-1, München u. Mering 2010, 102 S., € 19.80

Carolin Strobl: **Das Rasch-Modell.**

Eine verständliche Einführung für Studium und Praxis

Band 2, ISBN 978-3-86618-695-8, München u. Mering,
2. erw. Aufl. 2012, 131 S., € 19.80

Jost Reinecke: **Wachstumsmodelle**

Band 3, ISBN 978-3-86618-692-7, München u. Mering 2012, 111 S., € 19.80

Jürgen H.P. Hoffmeyer-Zlotnik, Uwe Warner:

Soziodemographische Standards für Umfragen in Europa

Band 4, ISBN 978-3-86618-827-3, München u. Mering 2013, 124 S., € 19.80

Michael Stegmann, Julia Werner, Heiko Müller:

Sequenzmusteranalyse. Einführung in Theorie und Praxis

Band 5, ISBN 978-3-86618-829-7, München u. Mering 2013, 88 S., € 19.80

Dieses Buch liefert aus sozialwissenschaftlicher Perspektive einen Überblick über aktuelle Längsschnittdaten und schildert relevante Analyseverfahren. Der Schwerpunkt liegt auf der Durchführung der Sequenzmusteranalyse mit der Statistiksoftware TDA, STATA und R. Dabei wird die kostenlose und frei zugängliche Software R mit ihren vielfältigen Möglichkeiten der Längsschnittdatenanalyse in den Mittelpunkt gerückt. Die ausführlichen Beschreibungen der theoretischen Grundlagen der Sequenzmusteranalyse sowie die umfangreichen Vorgehenserläuterungen zur Deskription von Sequenzen, Optimal Matching und Clusteranalyse werden durch Beispiel-Skripte und zahlreiche Abbildungen ergänzt, sodass sie analog auf eigene Daten angewendet werden können.