

Stichprobenziehung und Validierung der Typisierungen der Untersuchungsgemeinden für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«

Stefan Fischer
Jens A. Forkel

LETHE



Hochschule Neubrandenburg
University of Applied Sciences

IMPRESSUM

Stefan Fischer, Jens A. Forkel

Stichprobenziehung und Validierung der Typisierungen der Untersuchungsgemeinden für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«

Hochschule Neubrandenburg 2014

Fachbereich Gesundheit, Pflege, Management

Brodaer Str. 2

17033 Neubrandenburg

forkel@hs-nb.de

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS		II
TABELLENVERZEICHNIS		II
1 EINLEITUNG		3
2 GRUNDGESAMTHEIT UND STICHPROBENVERFAHREN		4
2.1 Zufällige Stichprobenauswahl.....		5
2.1.1 Einfache Zufallsstichprobe		6
2.1.2 Geschichte Stichprobe		6
2.1.3 Klumpenstichprobe.....		6
2.1.4 Mehrstufige Verfahren.....		7
2.2 Bewusste bzw. systematische Auswahlverfahren		7
3 MÖGLICHKEITEN DER STICHPROBENZIEHUNG FÜR DAS PROJEKT »LEBENSQUALITÄT UND ERINNERUNG IN DÖRFlichen GEMEINSCHAFTEN (LETHE)«		8
3.1 Ziehung einer Zufallsstichprobe aus der Grundgesamtheit.....		9
3.2 Ziehung einer Zufallsstichprobe aus der nach ehemaligen Landkreisen differenzierten Grundgesamtheit		9
3.3 Ziehung einer Zufallsstichprobe auf Grundlage der validisierten Gemeindetypisierung		12
3.3.1 Methodik und Validierung der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«		12
3.3.2 Stichprobenziehung auf Grundlage der Gemeindetypisierung		24
3.4 Zusammenfassung.....		24
4 COMPUTERGESTÜTZTE ZIEHUNG EINER ZUFALLS-STICHPROBE MIT DEM PROGRAMM R		25
4.1 Das Statistikprogramm R		25
4.2 Ziehung einer Zufallsstichprobe mit dem Programm R.....		27
4.2.1 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« aus der Grundgesamtheit		29
4.2.2 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« differenziert für die Landkreise Demmin, Mecklenburg-Strelitz und Müritz		30
4.2.3 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« auf Grundlage der Gemeindetypisierung.....		33
5 ZUSAMMENFASSUNG		36
LITERATUR		36

ABBILDUNGSVERZEICHNIS

Abbildung 2.1:	Auswahlverfahren der empirischen Sozialforschung	5
Abbildung 3.1:	Georeferenzierte Darstellung der Gemeinden des Landkreises Mecklenburgische Seenplatte (mit weniger als 500 Einwohnern; Stand 06-2012) differenziert nach der Zugehörigkeit zu den Kategorien der Gemeindetypisierung.....	11
Abbildung 3.2:	Übersicht der prozentualen Punktwerte der Orte der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« für die Kategorien Wirtschaft, Raum und Soziales sowie der Gesamtscore.....	23
Abbildung 5.1:	Georeferenzierte Darstellung der Untersuchungsgemeinden des Projekts LETHE	36

TABELLENVERZEICHNIS

Tabelle 3.1:	Gemeinden des Landkreises Mecklenburgische Seenplatte (mit weniger als 500 Einwohnern; Stand 06-2012) differenziert nach der Zugehörigkeit zu den Landkreisstrukturen vor der Kreisgebietsreform in Mecklenburg-Vorpommerns 2011 (N=72).	10
Tabelle 3.2:	Variablen der Kategorie Wirtschaft der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«	12
Tabelle 3.3:	Variablen der Kategorie Raum der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«	13
Tabelle 3.4:	Variablen der Kategorie Soziales der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«	14
Tabelle 3.5:	Kategorisierung der Gemeinden der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«.....	16
Tabelle 3.6:	Übersicht der deprivierten, kompensierenden und arrivierten Gemeinden der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« und deren Ausprägungen in den Kategorien Wirtschaft (Wi), Raum (Ra) und Soziales (So).....	18
Tabelle 3.7:	Absolute Anzahl und prozentualer Anteil der intra- und interkategorialen Korrelationen der Kategorien Wirtschaft, Raum und Soziales.....	19
Tabelle 3.8:	Intrakategoriale Validisierung der Kategorie Wirtschaft und interkategoriale Validisierung der Kategorie Wirtschaft mit den Kategorien Raum und Soziales.....	20
Tabelle 3.9:	Intrakategoriale Validisierung der Kategorie Raum und interkategoriale Validisierung der Kategorie Raum mit den Kategorien Wirtschaft und Soziales	21
Tabelle 3.10:	Intrakategoriale Validisierung der Kategorie Soziales und interkategoriale Validisierung der Kategorie Soziales mit den Kategorien Wirtschaft und Raum.....	22

1 EINLEITUNG

Die Typisierungen, Stichprobenziehung und Validierung des Auswahlverfahrens der Untersuchungsgemeinden für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« erfolgte im Zeitraum vom August bis November 2013. Inhalt des Werkvertrags und des vorliegenden Abschlussberichts bildete die für die Feldphase des Projekts LETHE notwendige Auswahl der Untersuchungsorte, wobei neben der eigentlichen Ortsauswahl mittels einer Zufallsstichprobe, eine Validierung der durch Forkel (2014) vorgenommenen Kategorisierung erfolgte.

Im Vorfeld einer Beschreibung der für die Feldphase des Projekts vorausgehenden Ortsauswahl finden sich im Folgenden zunächst Erläuterungen zu verschiedenen Möglichkeiten der Stichprobenziehung, wie sie in der empirischen Sozialforschung Anwendung finden (Kap. 2). Daran anschließend werden verschiedene Möglichkeiten der Ziehung von Zufallsstichproben, welche zu einer Ermittlung der Untersuchungsgemeinden führen, diskutiert (Kap. 3). Die Definition der für die Stichprobenziehung relevanten Grundgesamtheit wurde entsprechend des Untersuchungsgebietes als die Gesamtheit der Gemeinden Mecklenburg-Vorpommerns mit einer Einwohnerzahl von weniger als 500 Einwohnern (Stand 06-2012, Stat-MV) bestimmt. Eine Berechnung des Stichprobenumfangs wurde in Bezug auf die Repräsentativität der qualitativen und quantitativen Erhebungen vorgenommen und auf $n=9$ Untersuchungsgemeinden festgelegt (Kap. 3.1).

Für die Ortsauswahl wurden mehrere Möglichkeiten in Betracht gezogen. Neben der Ziehung einer Zufallsstichprobe mit dem Umfang $n=9$ aus der Grundgesamtheit (Kap. 3.1) bestand eine zweite Möglichkeit in der Ziehung einer nach den Landkreisstrukturen vor der Kreisgebietsreform 2011 im Bundesland Mecklenburg-Vorpommern differenzierten Zufallsstichprobe (Kap. 3.2). Bei letztgenannter Möglichkeit sollte eine Ziehung von jeweils drei Untersuchungsorten aus den ehemaligen Landkreisen Demmin, Mecklenburg-Strelitz und Müritz erfolgen. Eine dritte Möglichkeit umfasst die Stichprobenziehung auf Grundlage der von Forkel (2014) vorgenommenen Gemeindetypisierung (Kap. 3.3). Ausgehend von drei Kategorien *Wirtschaftliche Handlungsfähigkeit*, *Kultur-Räumliche Bedingungen* und *Soziale Ausprägungen* wurde eine Differenzierung der Orte der Grundgesamtheit in „deprivierte“, „kompensierende“ und „arrivierte“ Gemeinden vorgenommen. Es erfolgte die Ziehung von jeweils drei Orten der neu gebildeten Kategorien (depriviert, kompensierend, arrivierte). Neben der Diskussion verschiedener Möglichkeiten der Stichprobenziehung soll auch die von Forkel (2014) vorgenommenen Kategorie- und Typenbildung diskutiert werden (Kap. 3.3.1).

Im Anschluss an die im dritten Kapitel vorgenommene Diskussion der Möglichkeiten der Stichprobenziehung wird im Anschluss daran die technische Umsetzung der Stichprobenziehung mit der Statistiksoftware **R** erläutert. Um eine Beurteilung der technischen Umsetzung zu erleichtern, erfolgt zunächst eine kurze Vorstellung der Statistiksoftware (Kap. 4.1) sowie eine anschließende allgemeine Einführung zur Stichprobenziehung mit dem Programm **R** (Kap. 4.2). Daran anschließend werden die syntaktischen Befehle für die verschiedenen diskutierten Möglichkeiten der Stichprobenziehung für das Projekt LETHE vorgestellt (Kap. 4.1.1 bis 4.1.3). Den Abschluss des vorliegenden Berichts bildet eine Zusammenfassung sowie Diskussion der Ergebnisse (Kap. 5). Neben internen Verlinkungen im vorliegenden Dokument finden sich Hyperlinks zu verschiedenen Internetseiten. Diese sind in diesem Bericht [blau](#) dargestellt.

Die technische Umsetzung der Stichprobenziehung erfolgte mit der Statistiksoftware **R** Version 3.0.1. Sämtliche syntaktischen Befehle sind im Anhang dokumentiert. Darüber hinaus sind die im Text erläuterten syntaktischen Befehle aus Gründen einer besseren Lesbarkeit in der Schriftart **Arial** und **fett** formatiert. Neben dieser Bezugnahme auf syntaktische Befehle des Programms **R** im fortlaufenden Text werden komplexere Programmierungen separat aufgeführt. Diese sind mit einem Prompt (**>**) am Zeilenanfang gekennzeichnet. Entgegen dem allgemeinen Text (Schriftart: Times New Roman, Schriftgröße: 12 Pt.) sind diese syntaktischen Befehle in der Schriftart **Arial** (Schriftgröße: 10 Pt.) formatiert und mit einer Nummerierung in eckigen Klammern (**[]**) versehen. Kommentierungen in der **R**-Syntax sind mit einer Raute (**#**) am Zeilenanfang als solche gekennzeichnet. Syntaktische Befehle, die über mehrere Zeilen hinausgehen, sind am Zeilenanfang mit einem Pluszeichen (**+**) gekennzeichnet.

2 GRUNDGESAMTHEIT UND STICHPROBENVERFAHREN

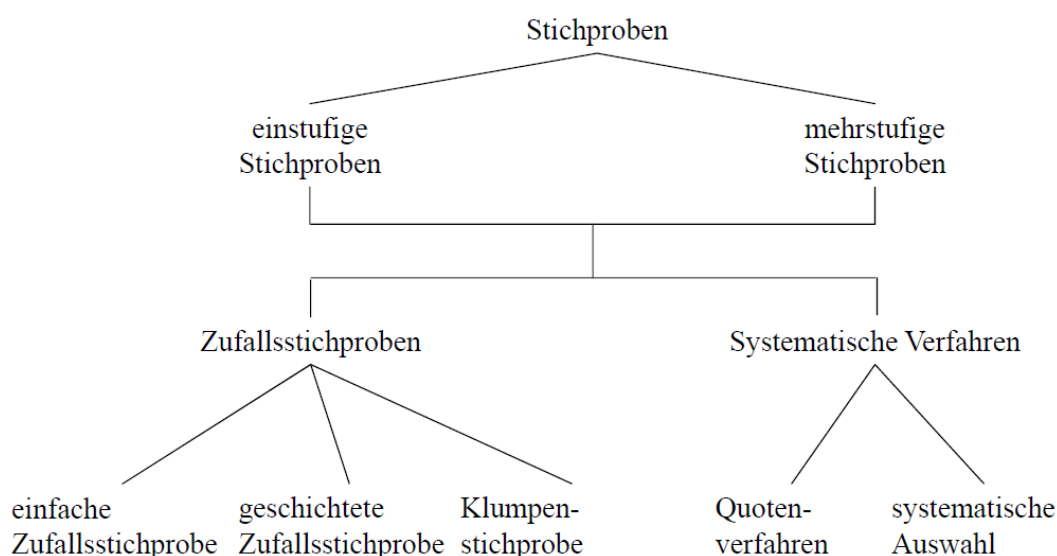
Im Vorfeld von empirischen Untersuchungen ist es unerlässlich, die für die Beantwortung einer wissenschaftlichen Fragestellung relevante „*Grundgesamtheit definieren zu können*“ (Friedrichs 1990: 128). Die „*genaue Größe einer Grundgesamtheit ist nicht immer leicht feststellbar und setzt [zudem] eine genaue Formulierung der Fragestellung voraus*“ (Mossig 2012: 2). Als Grundgesamtheit lassen sich „*alle potenziell untersuchbaren Einheiten [...], die ein gemeinsames Merkmal (oder eine gemeinsame Merkmalskombination) aufweisen*“, bezeichnen (Bortz/Schuster 2010: 79f.). Sie umfasst „*die Menge aller Individuen oder Objekte, über die eine Aussage getroffen werden soll*“ (Kauermann/Küchenhoff 2011: 5). Anders formuliert lässt sich die Grundgesamtheit als „*die Gesamtheit aller Fälle (Merkmalsträger) auf die sich eine Forschungsfrage bezieht und über die Erkenntnisse gewonnen werden sollen*“ definieren (Diaz-Bone 2006: 130). Die definierten Grundgesamtheiten müssen sich hierbei nicht zwangsläufig auf Personen beziehen, sondern können bspw. „*bei einer ökologischen Fragestellung [...] z.B. aus Planquadraten einer Fläche oder aus Seen eines Landes bestehen*“ (Kauermann/Küchenhoff 2011: 5), und „*können einen begrenzten oder theoretisch unbegrenzten Umfang aufweisen*“ (Bortz/Schuster 2010: 80).

Vollerhebungen von insbesondere großen Grundgesamtheiten bilden, da diese zumeist „*zu langwierig und zu teuer*“ sind (Atteslander 2010: 273), die Ausnahme. Da in der empirischen Forschung „*eine Betrachtung jedes einzelnen Elements der Grundgesamtheit aus Zeit- und Kostengründen [zumeist] nicht erfolgen kann*“ (Mossig 2012: 2), muss eine Auswahl einer Teilmenge der Grundgesamtheit, welche als Stichprobe oder Sample bezeichnet wird, getroffen werden (siehe u.a. Friedrichs 1990: 125; Diaz-Bone 2006: 130; Atteslander 2010: 273; Kauermann/Küchenhoff 2011: 6; Mossig 2012: 2). Durch die Bezugnahme auf Stichproben sollen dadurch, dass „*nicht alle Elemente der Grundgesamtheit, sondern nur eine Auswahl untersucht*“ wird, „*Kosten (Geld, Zeit und Energie)*“ gesenkt werden (Friedrichs 1990: 125). Unter dem Begriff der Stichprobe kann „*eine Auswahl von Elementen (n) aus der Gesamtheit aller Elemente (N), die durch ein oder mehrere gleiche Merkmale gekennzeichnet sind*“ (Friedrichs 1990: 125), verstanden werden. Demnach stellt eine Stichprobe „*eine Teilmenge aller Untersuchungsobjekte dar*“ (Bortz/Schuster 2010: 80). In der empirischen Sozialforschung stellt „*die Verwendung von Stichproben als Datengrundlage [...] den Normalfall [...] dar*“ (Diaz-Bone 2006: 130).

Ein Problem, das mit der Erhebung einer Stichprobe einhergeht, besteht darin, „*dass sich die gemessenen Variablen hinsichtlich ihrer statistischen Maßzahlen (Mittelwerte, Streuungen etc.) in Sample und Ausgangsmenge nicht zu sehr unterscheiden*“ (Atteslander 2010: 273). Diese Bedingung erscheint vor dem Hintergrund des zentralen Ziels, welches mit der Verwendung von Stichproben verfolgt werden soll, von besonderem Interesse: Die Ergebnisse, die Untersuchungen der Auswahl sollen „*Rückschlüsse von den Ergebnissen der Stichprobe auf die Eigenschaften der Grundgesamtheit*“ (Mossig 2012: 2) ermöglichen, „*sodass von der Stichprobe auf die Gesamt geschlossen werden kann*“ (Atteslander 2010: 273).¹

Grundsätzlich lassen sich die existierenden Stichprobenverfahren in zufällige und systematische Verfahren unterteilen (vgl. Abb. 2.1). Neben der einfachen sowie geschichteten Zufallsstichprobe finden Klumpen- oder Clusterstichproben Anwendung in der empirischen Sozialforschung. Bei systematischen Auswahlverfahren lassen sich das Quotenverfahren sowie die systematische Auswahl differenzieren. Zufällige sowie nichtzulässige Verfahren können ein- sowie mehrstufig konzipiert sein.²

Abbildung 2.1: Auswahlverfahren der empirischen Sozialforschung; *Quelle: Eigene Darstellung nach Atteslander (2010: 274)*



2.1 Zufällige Stichprobenauswahl

Wenngleich es sich bei einer Stichprobe lediglich um „*eine Auswahl der Fälle (Merkmalsträger) aus einer Grundgesamtheit*“ handelt (Diaz-Bone 2006: 131), sollen „*dennoch Aussagen über diese*“ getroffen werden können (Diaz-Bone 2006: 131). Um dieses gewährleisten zu können, muss – wie Diaz-Bone (2006: 131) anmerkt – „*die Stichprobe systematisch, d.h. unter Anwendung von Regeln gezogen werden.*“ Für ein grundlegendes Verständnis der unterschiedlichen Möglichkeiten, eine Zufallsauswahl zu treffen, werden im Folgenden verschiedene in der empiri-

¹ Zu den Voraussetzungen von Stichproben siehe auch Friedrichs (1990: 125). Demnach muss „(1.) *die Stichprobe [...] ein verkleinertes Abbild der Grundgesamtheit hinsichtlich der Heterogenität der Elemente und hinsichtlich der Repräsentativität der für die Hypothesenprüfung relevanten Variablen sein.* 2. *Die Einheiten oder Elemente der Stichprobe müssen definiert sein.* 3. *Die Grundgesamtheit sollte angebar und empirisch definierbar sein.* 4. *Das Auswahlverfahren muß angebar sein und Forderung (1) erfüllen*“ (Friedrichs 1990: 125).

² Neben diesen Auswahlverfahren finden sich weitere Auswahlverfahren, die in der empirischen Sozialforschung Anwendung finden können (siehe hierzu u.a. Diekmann 2009: 373-432).

schen Sozialforschung übliche Möglichkeiten der zufälligen Stichprobenauswahl erläutert. Neben der einfachen Zufallsstichprobe (Kap. 2.1.1), die eine Auswahl von Elementen für eine Stichprobe aus einer (homogenen) Grundgesamtheit ermöglicht, werden die Auswahlverfahren der geschichteten Stichprobe (Kap. 2.1.2) sowie die Klumpenstichprobe (Kap. 2.1.3) vorgestellt. Ferner finden sich Ausführungen zu mehrstufigen zufälligen Auswahlverfahren (Kap. 2.1.4).

2.1.1 Einfache Zufallsstichprobe

Findet sich eine Homogenität sowie eine Vollständigkeit *„bezüglich der zu untersuchenden Merkmale [einer Grundgesamtheit] [...] kann eine reine Zufallsauswahl vorgenommen werden“* (Atteslander 2010: 274). Mittels der Grundüberlegung des *„Urnenmodells“*³ kann gewährleistet werden, dass *„jede Einheit der Grundgesamtheit die gleiche Chance [...] [aufweist], für die Stichprobe ausgewählt zu werden“* (Atteslander 2010: 274). Die Praktikabilität einer *„echten“* Urnenziehung ist bei in der Praxis zumeist großen Grundgesamtheiten nicht gegeben (vgl. u.a. Atteslander 2010: 274). Aus diesem Grund wird die für *„einen echten Zufallsprozess“* notwendige *„Zufälligkeit einer Ziehung“* (Kauermann/Küchenhoff 2011: 11) *„in der Praxis meist durch einen Zufallszahlengenerator realisiert“* (Kauermann/Küchenhoff 2011: 11). Im Rahmen einer einfachen Zufallsstichprobe, die auch als uneingeschränkte Zufallsstichprobe bezeichnet werden kann (vgl. Diaz-Bone 2006: 133), weisen alle Einheiten der Grundgesamtheiten die gleiche Chance auf, in die Stichprobe gezogen zu werden (vgl. u.a. Diaz-Bone 2006: 133).

2.1.2 Geschichtete Stichprobe

Bei einer geschichteten Stichprobe erfolgt im Vorfeld des Auswahlverfahrens *„unter Heranziehung einzelner Merkmale“* eine Unterteilung der Grundgesamtheit *„in möglichst homogene Teilgesamtheiten“* (Diaz-Bone 2006: 134). Die Grundgesamtheit wird *„anhand eines oder mehrerer Merkmale [...] [somit] in ‘Schichten‘ unterteilt“* (Friedrichs 1990: 140). Findet sich für die Merkmale der Grundgesamtheit eine hohe Heterogenität, was für die Existenz verschiedener Teilmengen spricht, *„empfiehlt sich eine Auflistung der Stichprobe, sodass in jeder Schicht eine einfache Zufallsstichprobe gezogen wird“* (Atteslander 2010: 275).

Gegenüber einfachen Zufallsstichproben besteht ein wesentlicher Vorteil geschichteter Auswahlverfahren *„darin, dass die Abbildung der Grundgesamtheit [...] in der Regel präziser gelingt“* (Diaz-Bone 2006: 134). Voraussetzung für die Ziehung einer geschichteten Auswahl sind *„Vorinformationen über die Grundgesamtheit [...], mit denen begründet werden kann, warum welche Merkmale zur Schichtung herangezogen werden können“* (Diaz-Bone 2006: 134). Entsprechend der Zufallsstichprobe findet sich bei der geschichteten Stichprobe innerhalb der einzelnen Schichten eine gleichgroße Chance der einzelnen Elemente, in die Stichprobe gezogen zu werden (Atteslander 2010: 275).

2.1.3 Klumpenstichprobe

Bei Klumpenstichproben, die auch als Clusterstichproben bezeichnet werden können (vgl. Diaz-Bone 2006: 134; Friedrichs 1990: 142; Diekmann 2009: 387), handelt es sich um Auswahlverfahren, *„die jeweils ‚Klumpen‘ von nebeneinander liegenden Einheiten in das Sample einbeziehen“* (Atteslander 2010: 275). Wie Atteslander (2010: 275) anführt, können *„die Untersu-*

³ Im Rahmen des Urnenmodells *„wird eine Kugel oder ein Zettel in einer Urne deponiert und der Inhalt gut vermisch. Im Umfang der notwendigen Stichprobe werden nun wahllos Kugeln oder Zettel aus dieser Urne herausgenommen“* (Atteslander 2010: 274).

chungseinheiten [...] z.B. in einer Liste nebeneinander stehen oder räumlich benachbart sein“. Im letztgenannten Fall (bspw. von Wohnungen in einer Straße oder einem Stadtteil) kann auch von einer Flächenstichprobe gesprochen werden (vgl. Atteslander 2010: 275).

Die Klumpen, welche Untereinheiten der Grundgesamtheit darstellen (vgl. Friedrichs 1990: 142), müssen nicht zwangsläufig eine identische Anzahl an Einheiten aufweisen (vgl. Atteslander 2010: 275). Wichtig ist dennoch, dass jedes einzelne Element lediglich einem Klumpen zugeordnet werden kann (vgl. Friedrichs 1990: 142). Die Auswahl der Klumpen erfolgt mittels einer Zufallsstichprobe (vgl. Atteslander 2010: 275). Grundsätzlich gilt, *„dass ein einzelner Klumpen [...] keine Klumpenstichprobe darstellt, sondern eine Ad-hoc-Stichprobe, bei der zufällige Auswahlkriterien praktisch keine Rolle spielen.“* (Bortz/Schuster 2010: 81). Die Bezeichnung der Stichprobenziehung als Klumpenstichprobe ist lediglich gerechtfertigt, *„wenn mehrere zufällig ausgewählte Klumpen vollständig untersucht werden“* (Bortz/Schuster 2010: 81). Wie Atteslander (2010: 275) anmerkt, *„sprechen für die Auswahl von Klumpen oder Flächen“* die Tatsachen, dass oftmals *„keine zuverlässige Liste aller Einheiten der Grundgesamtheit“* existiert sowie ferner mit der Verwendung *„geographisch begrenzte[n] Flächen [...] oft eine deutliche Zeit- und Kostenersparnis zur Folge“* hat.

2.1.4 Mehrstufige Verfahren

Mehrstufige Auswahlverfahren in der empirischen Sozialforschung stellen *„eine Kombination mehrerer Verfahren“* dar (Atteslander 2010: 276). Im Gegensatz zu einfachen Zufallsstichproben besteht das Auswahlprinzip bei mehrstufigen Zufallsstichproben *„darin, dass die Stichprobe in mehreren Schritten über [...] [verschiedene] Stufen gezogen wird“* (Diaz-Bone 2006: 133). Der Ablauf eines mehrstufigen Auswahlverfahrens kann an dem Beispiel einer Gemeindestichprobe verdeutlicht werden (siehe u.a. Diekmann 2009: 385f.): In einem ersten Schritt (Stufe 1) wird zunächst eine beliebige Anzahl von Gemeinden aus einer Gemeindefliste ausgewählt. Anschließend erfolgt für jede der ausgewählten Gemeinden im zweiten Schritt (Stufe 2) die Auswahl eines definierten Umfangs von Adressen, die Teil der Stichprobe sein sollen. Klumpenverfahren bilden *„eine besondere Form der mehrstufigen Auswahl, die dadurch gekennzeichnet ist, dass auf der letzten Stufe alle Elemente der auf der vorletzten Stufe ausgewählten Einheiten aufgenommen werden“* (Diaz-Bone 2006: 134).

2.2 Bewusste bzw. systematische Auswahlverfahren

Neben verschiedenen Verfahren der zufälligen Stichprobenauswahl, wie sie im vorangegangenen Abschnitt erläutert wurden (Kap. 2.1), können in der empirischen Sozialforschung für die Auswahl von Untersuchungselementen ferner bewusste bzw. systematische Auswahlverfahren herangezogen werden. Wenngleich diese Auswahlverfahren für die vorliegende Stichprobenziehung nicht unmittelbar von Belang erscheinen, soll im Folgenden für eine Abgrenzung der für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« verwendeten Stichprobenverfahren dennoch kurz darauf eingegangen werden.

Insbesondere in der Markt- und Meinungsforschung ist das Quotenverfahren *„stark verbreitet“* (Atteslander 2010: 276). Vorteilhaft erscheint *„eine systematische Stichprobe insbesondere dann [...], wenn die Population mit einer Ordnung oder Abhängigkeitsstruktur versehen ist“* (Kauer- mann/Küchenhoff 2011: 42). Wie auch bei einer geschichteten Stichprobe geht man bei einem Quotenverfahren von einer heterogenen Verteilung der Merkmale der Grundgesamtheit aus (vgl.

Atteslander 2010: 276). Die Anzahl der zu untersuchenden Einheiten werden hierbei entsprechend der Verteilung der Variablen in der Grundgesamtheit ausgewählt (siehe hierzu u.a. Friedrichs 1990: 134). „Die Quota-Stichprobe erlaubt [...] keinen statistisch gesicherten Schluß auf die Grundgesamtheit, die die eigentliche Auswahl“ der Untersuchungseinheiten „nicht zufällig erfolgt“ (Friedrichs 1990: 135).

3 MÖGLICHKEITEN DER STICHPROBENZIEHUNG FÜR DAS PROJEKT »LEBENSQUALITÄT UND ERINNERUNG IN DÖRFLICHEN GEMEINSCHAFTEN (LETHE)«

Neben einer Validisierung der von Forkel (2014) vorgenommenen mikrospatialen Typisierung widmet sich der vorliegende Beitrag den Möglichkeiten einer Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«, welches an der Hochschule Neubrandenburg im Zeitraum von August 2013 bis Juli 2016 durchgeführt wird. Die Grundgesamtheit der Untersuchungseinheiten wurde definiert als die Gesamtheit aller Gemeinden des Landkreises Mecklenburgische Seenplatte (Mecklenburg-Vorpommern) mit weniger als 500 Einwohnern (Stand Juni 2012). Entsprechend dieser Definition umfasst die Grundgesamtheit insgesamt 72 Gemeinden (N=72). Im Rahmen des Auswahlverfahrens wurde die Ziehung einer 12%-Stichprobe festgelegt. Ausgehend von einer Grundgesamtheit von N=72 ergibt sich bei einer 12%-Stichprobe ein Wert von 8,64. Der auf diesem Wege ermittelte Stichprobenumfang beträgt somit neun Gemeinden (n=9).

Für die Auswahl der neun Gemeinden können verschiedene Verfahren in Betracht gezogen werden:

- Im ersten Fall wäre die Ziehung einer Zufallsstichprobe aus der gesamten Grundgesamtheit der 72 Gemeinden mit 500 oder weniger Einwohnern (Stichprobenumfang: n=9) möglich. Hierbei würde eine Zufallsauswahl unabhängig von Merkmalen (bspw. Zugehörigkeit zu einem der ehemaligen Landkreise Mecklenburg-Vorpommerns) erfolgen (Kap. 3.1).
- Darüber hinaus ist die Ziehung einer Zufallsstichprobe differenziert nach der Zugehörigkeit der Gemeinden zu den ehemaligen Landkreisen Mecklenburg-Vorpommerns in Betracht zu ziehen. Hierbei würden jeweils drei Gemeinden des ehemaligen Landkreises Demmin, drei Gemeinden des ehemaligen Landkreises Mecklenburg-Strelitz sowie drei Gemeinden des ehemaligen Landkreises Müritz mittels ausgewählt werden. Im Rahmen dieser zweiten Möglichkeit würde die Ziehung von insgesamt drei Zufallsstichproben (jeweils eine Auswahl für jeden ehemaligen Landkreis) erfolgen (Kap. 3.2).⁴
- Ferner ist die Ziehung einer Zufallsstichprobe auf Grundlage der von Forkel (2014) vorgenommenen Gemeindetypisierung, soweit eine Validität des Verfahrens gegeben ist (Kap. 3.1), möglich. Entsprechend der unter Verwendung der Kategorien *Wirtschaftliche Handlungsfähigkeit* (i.F. Kategorie *Wirtschaft*), *Räumliche Bedingungen* (i.F. Kategorie

⁴ Vgl. Gesetz zur Schaffung zukunftsfähiger Strukturen der Landkreise und kreisfreien Städte des Landes Mecklenburg-Vorpommern (Kreisstrukturgesetz) vom 12. Juli 2010. In: Gesetz- und Verordnungsblatt für Mecklenburg-Vorpommern vom 28. Juli 2010, S. 366.

Raum) und *Soziale Ausprägungen* (i.F. Kategorie *Soziales*) neu gebildeten Kategorien (deprivierte, kompensierende und arrivierte Gemeinden) würden jeweils drei deprivierte, drei kompensierende sowie drei arrivierte Gemeinden ausgewählt werden (Kap. 3.3.2).

Im Folgenden werden die drei unterschiedlichen Möglichkeiten der Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« näher erläutert (Kap. 3.1 bis 3.3).

3.1 Ziehung einer Zufallsstichprobe aus der Grundgesamtheit

Eine erste Möglichkeit der Ermittlung der Untersuchungsorte für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« besteht in der Ziehung einer Zufallsstichprobe mit dem Stichprobenumfang $n=9$ aus der Grundgesamtheit der 72 Orte im Landkreis Mecklenburgische Seenplatte mit weniger als 500 Einwohnern (Stand 06-2012). Wenngleich es sich bei dieser Variante um die einfachste Möglichkeit der Ortsauswahl handelt, ist dieses Verfahren nicht geeignet, Unterschiede zwischen den Untersuchungsorten zu berücksichtigen. Da für das Projekt LETHE eine Auswahl von möglichst heterogenen Untersuchungsorten angestrebt wird, kommt die Ziehung einer Zufallsstichprobe aus der Grundgesamtheit der 72 Orte nicht in Betracht. Die technische Umsetzung dieses Verfahrens mit der Statistiksoftware **R** wird in Kapitel 4.2.1 näher erläutert.

Daher wurden alle 72 Orte des Landkreises Mecklenburgische Seenplatte mit weniger als 500 Einwohnern (Gebietsstand 30.06.2012, Statistisches Amt MV, A123) nach sozialgeografischen (z.B. Dehne, HS Neubrandenburg, 2009) und/oder sozial-historischen Typisierungsansätzen (z.B. Beetz, HS Mittweida, 2008) stratifiziert. Es ist zu erwarten, dass eine Stratifikation mit etwa drei Kategorien für die Untersuchung der hier theoretisch postulierten sozialökologischen Bedingungsfaktoren ausreichend ist. Da bereits bekannt ist, dass die Kategorien solcher Ortstypologien mit einer gewissen Abhängigkeit zur Einwohnerzahl der Orte einhergehen, werden aus jeder Kategorie bis zu drei Orte per Zufallsauswahl gezogen. Damit soll gewährleistet werden, dass zu jedem Ortstypus eine dreistellige Fallzahl realisiert werden kann. In Abhängigkeit von dieser ‚Flächenstichprobe‘ ist auf Basis von amtlichen Meldedaten (Statistisches Amt MV 2012, A123 u. A134) von rund 2.500 zu befragenden Personen auszugehen. Aus den Erfahrungen der Landgesundheitsstudie und anderen Bevölkerungssurveys wird davon ausgegangen, dass eine Ausschöpfungsquote von gut 20% bis knapp 40% erzielt wird. Daraus ergibt sich ein zu erwartender Stichprobenumfang von $n > 500$ und < 1000 , um die strukturanalytischen Perspektiven der Modellentwicklung LETHE zu erfassen und anhand von Vergleichsstudien zu überprüfen.

3.2 Ziehung einer Zufallsstichprobe aus der nach ehemaligen Landkreisen differenzierten Grundgesamtheit

Als zweite Möglichkeit einer Ortsauswahl kommt die Ziehung einer nach den ehemaligen Landkreisen des neuen Landkreises Mecklenburgische Seenplatte differenzierten Grundgesamtheit in Betracht. Hierbei werden die 72 Gemeinden mit weniger als 500 Einwohnern (Stand 06-2012) nach ihrer Zugehörigkeit zu den Landkreisen vor der Kreisgebietsreform in Mecklenburg-Vorpommern aus dem Jahr 2011 (4. September 2011) differenziert. Wie aus Tabelle 3.1 hervorgeht, wurden 25 Gemeinden dem Landkreis Demmin, 27 Gemeinden dem Landkreis Müritz sowie 20 Gemeinden dem Landkreis Mecklenburg-Strelitz zugeordnet. Anschließend erfolgte für

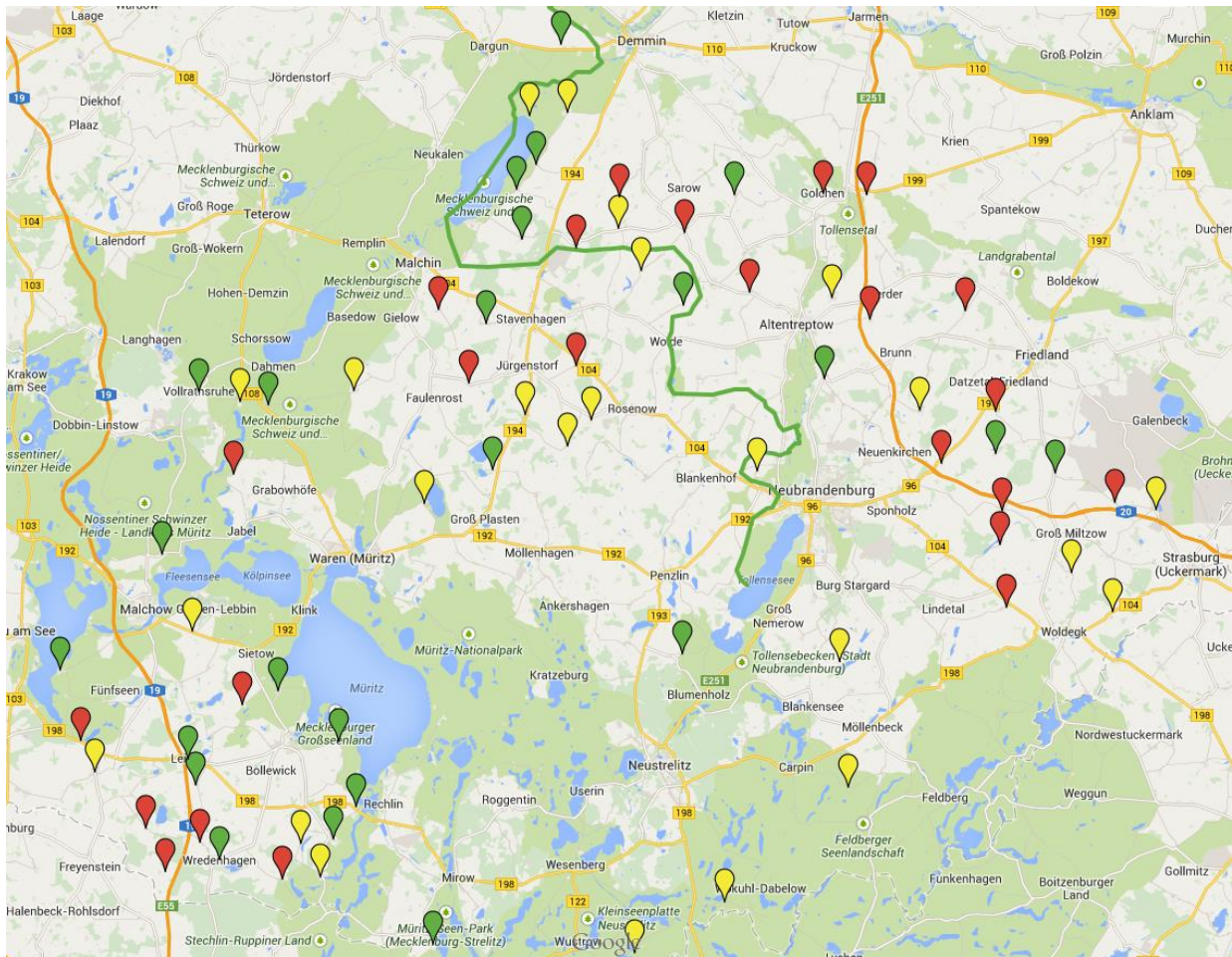
jede der drei Gruppierungen (Demmin, Müritz, Mecklenburg-Strelitz) die Ziehung einer Zufallsstichprobe, wobei für jeden der ehemaligen Landkreise eine Auswahl von drei Untersuchungsgemeinden vorgenommen wurde (vgl. Abb. 3.1).

Tabelle 3.1: Gemeinden des Landkreises Mecklenburgische Seenplatte (mit weniger als 500 Einwohnern; Stand 06-2012) differenziert nach der Zugehörigkeit zu den Landkreisstrukturen vor der Kreisgebietsreform in Mecklenburg-Vorpommerns 2011 (N=72); *Quelle: (Landkreisneuordnungsgesetz – LNOG M-V), Statistik MV 06-2012, eigene Darstellung.*

Gemeinden des ehemaligen Landkreises Demmin (n=25)	Gemeinden des ehemaligen Landkreises Müritz (n=27)	Gemeinden des ehemaligen Landkreises Mecklenburg-Strelitz (n=20)
Altenhagen	Altenhof	Beseritz
Bredenfelde	Buchholz	Cammin
Breest	Bütow	Eichhorst
Briggow	Gotthun	Genzkow
Duckow	Grabow-Below	Glienke
Gnevkow	Groß Kelle	Godendorf
Golchen	Kieve	Grünow
Grammentin	Klocksין	Helpt
Grapzow	Leizen	Hohenzieritz
Grischow	Ludorf	Kublank
Gülzow	Massow	Mildenitz
Hohenbollentin	Melz	Neddemin
Kentzlin	Moltzow	Neetzka
Kittendorf	Neu Gaarz	Petersdorf
Kriesow	Penkow	Priepert
Lindenberg	Priborn	Schönbeck
Meesiger	Schwarz	Schönhausen
Pripsleben	Schwinkendorf	Staven
Ritzerow	Silz	Voigtsdorf
Röckwitz	Stuer	Zirzow
Schönfeld	Torgelow am See	
Sommersdorf	Varchentin	
Verchen	Vipperow	
Warrenzin	Vollrathruhe	
Zetemin	Wredenhagen	
	Zepkow	
	Zislow	

Eine auf diesem Wege vorgenommene Auswahl der Untersuchungsgemeinden für LETHE würde eine Ziehung von Orten aus unterschiedlichen Regionen des Landkreises Mecklenburgische Seenplatte ermöglichen. Dennoch bietet dieses Verfahren nur eingeschränkte Möglichkeiten, strukturell unterschiedliche Orte auszuwählen, da das Differenzierungsmerkmale der Stichprobenziehung (geographische Unterschiede) nicht zwangsläufig mit strukturellen Unterschieden einhergehen muss. Diese zweite Möglichkeit der Ortsauswahl für das Projekt LETHE bietet dennoch, im Gegensatz zur ersten Methode (Ziehung einer Zufallsstichprobe von n=9 aus der Grundgesamtheit der Gemeinden insgesamt), die Möglichkeit, nach einem (regionalen) Merkmal differenzierte Orte auszuwählen. Die technische Umsetzung dieses Verfahrens mit der Statistiksoftware **R** wird in Kapitel 4.2.2 näher erläutert.

Abbildung 3.1: Georeferenzierte Darstellung der Gemeinden des Landkreises Mecklenburgische Seenplatte (<500 Einwohner; Stand 06-2012) nach der Zugehörigkeit zu den Typen *arrivierte* Gemeinden (grün), *kompensierende* Gemeinden (gelb), *deprivierte* Gemeinden (rot); Darstellung mit Google Maps (2014)



3.3 Ziehung einer Zufallsstichprobe auf Grundlage der validierten mikrospatialen Typisierung der Gemeinden <500 EW im Landkreis Mecklenburgische Seenplatte

3.3.1 Methodik und Validierung der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«

Die von Forkel (2014) vorgenommene Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« umfasste insgesamt 73 Variablen für die drei Kategorien Wirtschaft (18 Variablen), Raum (23 Variablen) und Soziales (32 Variablen). Für alle 72 Gemeinden der Grundgesamtheit lagen Daten der 73 Variablen vor. Fehlende Werte einzelner Variablen waren somit nicht vorhanden.

Die 18 für die **Kategorie Wirtschaft** berücksichtigten Variablen sind in Tabelle 3.2 dargestellt. Neben den Bezeichnungen der einzelnen Indikatoren finden sich in der Darstellung ferner Angaben zu den Erhebungsjahren sowie der Art der aufgenommenen Werte (absolute bzw. relative Werte). Darüber hinaus sind die Quellen der berücksichtigten Daten vermerkt. Neben Indikatoren der Beschäftigung (u.a. sozialversicherungspflichtig Beschäftigte, Steuerpflichtige insgesamt) wurden in der Kategorie Wirtschaft verschiedene Indikatoren berücksichtigt, die eine Beurteilung der finanziellen Situation der Gemeinden ermöglichen (u.a. Schuldenstand der Gemeinden insgesamt je Einwohner und Schuldenentwicklung). Für eine Einschätzung der wirtschaftlichen Leistungsfähigkeit der Gemeinden erfolgte die Aufnahme von Indikatoren des gezahlten Steueraufkommens (u.a. zu versteuerndes Einkommen, Grundbetrag der Grundsteuer A und B, Gewerbesteuer, Gemeindeanteil an Einkommens- und Umsatzsteuer).

Tabelle 3.2: Variablen der Kategorie Wirtschaft der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«; *Quelle: Eigene Darstellung nach Forkel (2014)*

Nr.	Bezeichnung der Variable	Einheit	Jahr
1	Sozialversicherungspflichtig Beschäftigte der Einwohner Wohnort ¹	Prozent	12-2011
2	Sozialversicherungspflichtig Beschäftigte der Einwohner Arbeitsort ¹	Prozent	12-2011
3	Sozialversicherungspflichtig Beschäftigte am Arbeitsort nach fünf Wirtschaftsbereichen ²	Prozent	06-2011
4	Haushalte mit Wohngeld insgesamt ²	Anzahl	12-2009
5	Steuerpflichtige insgesamt (Lohn- und Einkommensteuerstatistik) (2004=100%) ²	Prozent	2004 zu 2007
6	Gesamtbetrag der Einkünfte ²	Prozent	2004 zu 2007
7	Gesamtbetrag der Einkünfte je Steuerpflichtigen ²	Prozent	2004 zu 2007
8	zu versteuerndes Einkommen ²	Prozent	2004 zu 2007
9	festgesetzte Einkommensteuer ²	Prozent	2004 zu 2007
10	Grundbetrag der Grundsteuer A ²	in 1000 EUR pro EFB	12-2011
11	Grundbetrag der Grundsteuer B ²	in 1000 EUR pro EFB	12-2011
12	Gewerbesteuer ²	in 1000 EUR pro EFB	12-2011
13	Gemeindeanteil an der Einkommenssteuer ²	1000 Euro	12-2011
14	Gemeindeanteil an der Umsatzsteuer ²	1000 Euro	12-2011
15	Gewerbesteuerumlage ²	1000 Euro	12-2011
16	Gewerbesteuereinnahmen (Aufkommen abzüglich Umlage) ²	1000 Euro	12-2011
17	Schuldenstand der Gemeinden insgesamt je Einwohner ²	1000 Euro	12-2011
18	Schuldenentwicklung ²	Prozent	2006 zu 2011

¹ Quelle: Statistik der Agentur für Arbeit; ² Quelle: Statistisches Landesamt Mecklenburg-Vorpommern (SIS-Online)

Die Aktualität der Daten der Kategorie Wirtschaft kann insgesamt als hoch beurteilt werden. Neben Daten aktueller Erhebungsjahre finden sich unter den wirtschaftlichen Variablen auch Indikatoren, die eine Berücksichtigung der Entwicklung der wirtschaftlichen Situation in den

Gemeinden ermöglichen (Steuerpflichtige insgesamt 2007 im Vergleich zu 2004, Schuldenentwicklung). Eine detaillierte Variablenbeschreibung findet sich bei Forkel (2014).

Die 23 für die **Kategorie Raum** berücksichtigten Variablen sind in Tabelle 3.3 dargestellt. Neben den Bezeichnungen der einzelnen Indikatoren finden sich in der Darstellung ferner Angaben zu den Erhebungsjahren sowie der Art der aufgenommenen Werte (absolute bzw. relative Werte). Darüber hinaus sind die Quellen der berücksichtigten Daten vermerkt. Neben Variablen der Bautätigkeit (u.a. Baufertigstellungen von Wohn- und Nichtwohngebäuden) und der Wohnverhältnisse (u.a. Gebäudeauslastung) finden sich in der Kategorie Raum auch Indikatoren, die eine Einschätzung der räumlichen Struktur der Gemeinden ermöglichen (Betriebs-, Erholungs-, Verkehrs-, Landwirtschafts-, Wald-, Wasserfläche). Die Berücksichtigung der Erreichbarkeiten von Ober-, Grund- und Mittelzentren lässt eine Einschätzung der Infrastruktur der 72 Orte und deren Umgebung zu. Kulturräumliche Aspekte finden in der Kategorie in Form verschiedener Variablen Berücksichtigung (u.a. Projekte der Stiftung Denkmalschutz, Kirchen/Denkmalliste, technische Denkmäler, Zeitpunkt der Ersterwähnung). Die Aktualität der Daten der Kategorie Raum kann insgesamt als hoch beurteilt werden. Eine detaillierte Variablenbeschreibung findet sich bei Forkel (2014).

Tabelle 3.3: Variablen der Kategorie Raum der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«; *Quelle: Eigene Darstellung nach Forkel (2014)*

Nr.	Bezeichnung der Variable	Einheit	Jahr
1	Baufertigstellungen von Wohn- und Nichtwohngebäuden ¹	Anzahl	12-2011
2	Gebäudeauslastung: Einwohner pro Wohngebäude ¹	Quote	12-2011
3	Anteil von Gebäude- und Freifläche an der gesamten Bodenfläche ¹	Prozent	12-2011
4	Betriebsfläche (ohne Abbauland) ¹	Prozent	12-2011
5	Erholungsfläche ¹	Prozent	12-2011
6	Verkehrsfläche ¹	Prozent	12-2011
7	Landwirtschaftsfläche ¹	Prozent	12-2011
8	Waldfläche ¹	Prozent	12-2011
9	Wasserfläche ¹	Prozent	12-2011
10	Erreichbarkeit von Oberzentrum ²	Minuten	08-2013
11	Erreichbarkeit von Grundzentrum ²	Minuten	08-2013
12	Erreichbarkeit von Mittelzentrum ²	Minuten	08-2013
13	Tourismus: geöffnete Beherbergungsbetriebe ab 9 Betten ¹	Anzahl	06-2011
14	Verkehrsunfälle (Mittelwert) ¹	Anzahl	2009 bis 2011
15	Lage in einem Landschaftsschutzgebiet ³	ja/nein	08-2013
16	Backsteinroute Punkte (Geoportal MV) ³	Anzahl	08-2013
17	Badestellen (Geoportal MV) ³	Anzahl	08-2013
18	Breitband, alle Technologien >= 6 mbit ³	Anzahl	08-2013
19	Projekte der Stiftung Denkmalschutz ⁴	Anzahl	08-2013
20	Kirchen Wiki, Kirchen in MV, Denkmalliste (Kirchenanteil pro Ortsteil ohne Vorwerke, Siedlungen) ⁵	Anzahl	08-2013
21	Denkmale insgesamt ⁵	Anzahl	08-2013
22	Technische Denkmal (Mühle, Bahnhof) ⁵	Anzahl	08-2013
23	Ersterwähnung ⁶	Jahr	08-2013

¹ Quelle: Statistisches Landesamt Mecklenburg-Vorpommern (SIS-Online).

² Quelle: Eigene Berechnungen nach Routenplaner und Definition BBR.

³ Quelle: Geoportal Mecklenburg-Vorpommern.

⁴ Quelle: Verzeichnis geförderter Projekte der Stiftung Denkmalschutz.

⁵ Quelle: Denkmalliste des Landes Mecklenburg Vorpommern 1997, Kirchen-Wiki, Landesamt für Denkmalpflege.

⁶ Quelle: Selbstdarstellungen Wikipedia, Schautafeln, Landesamt für Denkmalpflege.

Die 32 für die **Kategorie Soziales** berücksichtigten Variablen sind in Tabelle 3.4 dargestellt. Neben den Bezeichnungen der einzelnen Indikatoren finden sich in der Darstellung ferner Angaben zu den Erhebungsjahren sowie der Art der aufgenommenen Werte (absolute bzw. relative

Werte). Darüber hinaus sind die Quellen der berücksichtigten Daten vermerkt. Neben einer Vielzahl von Indikatoren der Bevölkerungsstruktur (u.a. altersbereinigte Frauenquote, Anteile verschiedener Altersklassen an der Bevölkerung, Bevölkerungsbewegung) wurden in der Kategorie Soziales verschiedene Indikatoren berücksichtigt, die eine Beurteilung der Erwerbssituation in den Gemeinden ermöglichen (u.a. sozialversicherungspflichtig Beschäftigte nach Wirtschaftszweigen, Einpendlerquote). Darüber hinaus finden sich weitere Indikatoren für die Beurteilung sozialer und sozialräumlicher Bedingungen in den Orten der Grundgesamtheit (u.a. Bildungsweisen, Wahlverhalten, Vereine, niedergelassene Ärzte). Die Aktualität der Daten der Kategorie Soziales kann insgesamt als hoch beurteilt werden. Neben Daten aktueller Erhebungsjahre finden sich unter den Variablen auch Indikatoren, die eine Berücksichtigung der Entwicklung der Verhältnisse in den Gemeinden in den zurückliegenden Jahren ermöglichen (Veränderungen der Anteile verschiedener Bevölkerungsgruppen, Bevölkerungsentwicklung insgesamt). Eine detaillierte Variablenbeschreibung findet sich bei Forkel (2014).

Tabelle 3.4: Variablen der Kategorie Soziales der Gemeindetypisierung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«; *Quelle: Eigene Darstellung nach Forkel (2014)*

Nr.	Bezeichnung der Variable	Einheit	Jahr
1	Einwohner je Quadratkilometer ¹	EW/qm	12-2011
2	altersbereinigte Frauenquote: Quotient Frauen/ Anteil an BEV 60+ ¹	Quote	12-2011
3	Kindheit und Grundschulalter: 0 bis 10-Jährige pro Einwohner ¹	Prozent	12-2011
4	Jugend und Adoleszenz: 10 bis 25-Jährige pro Einwohner ¹	Prozent	12-2011
5	Ältere und Alte: Bevölkerung im Alter von 60 Jahren und älter pro Einwohner ¹	Prozent	12-2011
6	Geborene pro Einwohner (Geborene-Kopf-Quote) ¹	Prozent	12-2011
7	Gestorbenen pro Einwohner (Gestorbene-Kopf-Quote) ¹	Prozent	12-2011
8	Zuzüge pro Einwohner ¹	Prozent	2011
9	Fortzüge pro Einwohner ¹	Prozent	2011
10	SVPB am Arbeitsort nach Wirtschaftsbereichen: Land- und Forstwirtschaft, Fischerei ¹	Prozent	06-2011
11	SVPB am Arbeitsort nach Wirtschaftsbereichen: Produzierendes Gewerbe ¹	Prozent	06-2011
12	SVPB am Arbeitsort nach Wirtschaftsbereichen: Handel, Verkehr, Gastgewerbe ¹	Prozent	06-2011
13	Erbringung von öffentlichen und privaten Dienstleistungen ¹	Prozent	06-2011
14	Einwohner in beruflicher Ausbildung ¹	Anzahl	06-2011
15	Einpendlerquote ¹	Prozent	06-2011
16	Anteil der in der Gemeinde beschäftigten ortsansässigen Bevölkerung im arbeitsfähigen Alter ¹	Prozent	06-2011
17	Erwerbslosenquote SGB II und SGB III (eigene Berechnungen) ²	Prozent	06-2011
18	Anteil der Haushalte mit Wohngeld ¹	Prozent	12-2009
19	Anteil von Kindern in Tagesstätten und Schulen an der Bevölkerung ¹	Prozent	12-2009
20	Ausländeranteil an der Bevölkerung ¹	Prozent	12-2011
21	Veränderungen des Anteils der Bevölkerung im Alter von 65 Jahren und älter ¹	Prozent	2009 zu 2011
22	Veränderungen des Anteils der Bevölkerung im Alter von unter 15 Jahren ¹	Prozent	2009 zu 2011
23	Bevölkerungsentwicklung (1990=100%) ³	Prozent	1990 zu 2000
24	Bevölkerungsentwicklung (2000=100%) ³	Prozent	2000 zu 2010
25	Wahlbeteiligung: Kommunalwahl ⁴	Prozent	2011
26	Wahlbeteiligung: Wahl zum Bundestag der BRD in Mecklenburg-Vorpommern ⁴	Prozent	09-2013
27	rechtes Wahlverhalten (NPD, PRO, REP) ⁴	Prozent	09-2013
28	Arzt in Gemeinde ⁵	ja/nein	2013
29	Vereine insgesamt ⁶	Anzahl	08-2013
30	Unternehmen, Tourismus und Handwerke ⁶	Anzahl	08-2013
31	Treffpunkte (Friseur, Markt, Gaststätte, Lokal) ⁶	Anzahl	08-2013
32	Familiendichte: Einwohner pro Wohnung ¹	Prozent	12-2011

¹ Quelle: Statistisches Landesamt Mecklenburg-Vorpommern (SIS-Online).

² Quelle: Statistik Agentur für Arbeit.

³ Quelle: Statistisches Landesamt Mecklenburg-Vorpommern und eigene Berechnungen von Forkel (2014).

⁴ Quelle: Statistisches Landesamt Mecklenburg-Vorpommern. Landeswahlleiterin, 2011, 2013.

⁵ Quelle: Gesundheitsamt des Landkreises Mecklenburgische-Seenplatte.

⁶ Quelle: Gemeinsames Registerportal der Länder: Vereins- und Handelsregister 2013, eigene Beobachtungen, Gemeindeangaben.

Nachdem die Indikatoren für alle Gemeinden vorlagen, konnte für jede Variable der Median berechnet werden, auf dessen Grundlage eine Einteilung der Gemeinden in unterdurchschnittliche und überdurchschnittliche Orte erfolgte. Wies eine Gemeinde bei einem Indikator einen unterdurchschnittlichen Wert auf (kleiner als der Median), wurde die neue Kodierung „-1“ vergeben. Lag für eine Gemeinde ein überdurchschnittlicher Wert vor (größer als der Median), wurde die neue Kodierung „+1“ vergeben. Für jede der drei Kategorien (Wirtschaft, Raum, Soziales) erfolgte die Bildung der Summen der neuen Kodierungen („-1“, „+1“). Dementsprechend konnten die Orte der Grundgesamtheit einen kategoriespezifischen Gesamtscore von +/-18 in der Kategorie Wirtschaft, +/-23 in der Kategorie Raum sowie +/-32 in der Kategorie Soziales erreichen. Ferner wurde ein Gesamtscore gebildet. Dieser konnte zwischen +/-73 liegen.

Die für die anschließende Kategorisierung notwendigen Daten sind in Tabelle 3.5 dargestellt. Zunächst wurden die Summen der in den Kategorien erreichten Punkte gebildet. Die Punktscores für die einzelnen Gemeinden sind in den grünen Spalten der Tabelle 3.5 aufgeführt. Die maximale Punktzahl einzelnen Kategorien (Wirtschaft, Raum, Soziales) entspricht der Anzahl der Kategorievariablen. Dementsprechend konnten die Gemeinden in der Kategorie Wirtschaft maximal einen Wert von +18 erreichen (mögliches Maximum Kategorie Raum: +23; Soziales: +32). Dies sei an einem Beispiel veranschaulicht: Die Gemeinde Altenhagen konnte in der Kategorie Wirtschaft einen Scorewert von „-11“ erreichen (grüne Spalte der Kategorie Wirtschaft), was einem prozentualen Scorewert von „-57,89%“ entspricht. Der Punktwert der Kategorie Wirtschaft liegt somit auf einer Skala von „-100,0%“ bis „+100,0%“ bei einem Wert von ca. „-58%“, was einem vergleichsweise schlechtem Ergebnis entspricht. Nachdem dieses Vorgehen für alle Orte vorgenommen wurde, erfolgte die Bildung des Medians, welcher für die Kategorie Wirtschaft „15,79“ entspricht (rote Spalte der Kategorie Wirtschaft). In einem letzten Schritt erhielten die einzelnen Orte entsprechend ihres prozentualen Punktwertes eine „-1“, wenn dieser kleiner als der Median der Kategorie war. Lag der prozentuale Punktwert über dem Median erhielten die Orte eine „+1“. Für die exemplarisch angeführte Gemeinde Altenhagen wurde ein Rang für die Kategorie Wirtschaft von „-1“ vergeben (blaue Spalte der Kategorie Wirtschaft), da der erzielte prozentuale Punktwert („-57,89“) kleiner ist, als der Median der Kategorie („15,79“).

Tabelle 3.5: Kategorisierung der Gemeinden der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«; *Quelle: Eigene Darstellung nach Forkel (2014)*

Gemeinde	Punkte Wirtschaft	% Punkte von 18 Variablen	Rang Wirtschaft	Punkte Raum	% Punkte von 23 Variablen	Rang Raum	Punkte Soziales	% Punkte von 32 Variablen	Rang Soziales	gesamt % Punkte von 73	
Altenhagen	-11	-57,89	-1	-11	-47,83	-1	6	18,75	1	-21,05	-1
Altenhof	7	36,84	1	-9	-39,13	-1	2	6,25	1	0,00	-1
Beseritz	-7	-36,84	-1	-7	-30,43	-1	0	0,00	1	-18,42	-1
Bredenfelde	1	5,26	-1	-3	-13,04	1	0	0,00	1	-2,63	-1
Breest	-5	-26,32	-1	-9	-39,13	-1	-16	-50,00	-1	-39,47	-1
Briggow	-3	-15,79	-1	-1	-4,35	1	-2	-6,25	-1	-7,89	-1
Buchholz	11	57,89	1	7	30,43	1	-10	-31,25	-1	10,53	1
Bütow	11	57,89	1	-3	-13,04	1	10	31,25	1	23,68	1
Cammin	7	36,84	1	-1	-4,35	1	-6	-18,75	-1	0,00	-1
Duckow	-11	-57,89	-1	-7	-30,43	-1	-10	-31,25	-1	-36,84	-1
Eichhorst	9	47,37	1	-3	-13,04	1	0	0,00	1	7,89	-1
Genzkow	-5	-26,32	-1	-5	-21,74	-1	-10	-31,25	-1	-26,32	-1
Glienke	-7	-36,84	-1	-7	-30,43	-1	-2	-6,25	-1	-21,05	-1
Gnevkow	11	57,89	1	-7	-30,43	-1	-2	-6,25	-1	2,63	-1
Groß Kelle	3	15,79	-1	-9	-39,13	-1	-4	-12,50	-1	-13,16	-1
Grünow	-3	-15,79	-1	3	13,04	1	2	6,25	1	2,63	-1
Gülzow	-1	-5,26	-1	-9	-39,13	-1	8	25,00	1	-2,63	-1
Helpt	3	15,79	-1	-1	-4,35	1	-4	-12,50	-1	-2,63	-1
Hohenbollentin	1	5,26	-1	-11	-47,83	-1	-8	-25,00	-1	-23,68	-1
Hohenzieritz	1	5,26	-1	9	39,13	1	8	25,00	1	23,68	1
Kentzlin	-1	-5,26	-1	-5	-21,74	-1	-10	-31,25	-1	-21,05	-1
Kieve	-5	-26,32	-1	1	4,35	1	-12	-37,50	-1	-21,05	-1
Kittendorf	-1	-5,26	-1	-11	-47,83	-1	-2	-6,25	-1	-18,42	-1
Klocksinn	-1	-5,26	-1	5	21,74	1	-14	-43,75	-1	-13,16	-1
Kriesow	5	26,32	1	-5	-21,74	-1	-10	-31,25	-1	-13,16	-1
Kublank	3	15,79	-1	-11	-47,83	-1	-12	-37,50	-1	-26,32	-1
Leizen	9	47,37	1	5	21,74	1	8	25,00	1	28,95	1
Lindenberg	5	26,32	1	-11	-47,83	-1	-2	-6,25	-1	-10,53	-1
Ludorf	15	78,95	1	5	21,74	1	12	37,50	1	42,11	1
Massow	-7	-36,84	-1	3	13,04	1	-12	-37,50	-1	-21,05	-1
Meesiger	3	15,79	-1	5	21,74	1	4	12,50	1	15,79	1
Melz	3	15,79	-1	-7	-30,43	-1	-8	-25,00	-1	-15,79	-1
Mildenitz	7	36,84	1	-5	-21,74	-1	-10	-31,25	-1	-10,53	-1
Moltzow	7	36,84	1	7	30,43	1	-2	-6,25	-1	15,79	1
Neddemin	3	15,79	-1	1	4,35	1	6	18,75	1	13,16	1
Neetzka	-7	-36,84	-1	-1	-4,35	1	-10	-31,25	-1	-23,68	-1
Neu Gaarz	-3	-15,79	-1	-5	-21,74	-1	-12	-37,50	-1	-26,32	-1
Penkow	1	5,26	-1	-9	-39,13	-1	2	6,25	1	-7,89	-1
Petersdorf	-1	-5,26	-1	-7	-30,43	-1	0	0,00	1	-10,53	-1
Priborn	1	5,26	-1	5	21,74	1	0	0,00	1	7,89	-1
Priepert	5	26,32	1	-1	-4,35	1	-6	-18,75	-1	-2,63	-1
Pripsleben	3	15,79	-1	-13	-56,52	-1	-6	-18,75	-1	-21,05	-1
Ritzerow	3	15,79	-1	-9	-39,13	-1	-4	-12,50	-1	-13,16	-1
Röckwitz	17	89,47	1	-9	-39,13	-1	14	43,75	1	28,95	1
Schönbeck	11	57,89	1	3	13,04	1	-10	-31,25	-1	5,26	-1
Schönfeld	3	15,79	-1	-11	-47,83	-1	-8	-25,00	-1	-21,05	-1
Schönhausen	3	15,79	-1	-3	-13,04	1	-6	-18,75	-1	-7,89	-1
Schwarz	9	47,37	1	7	30,43	1	-6	-18,75	-1	13,16	1
Silz	7	36,84	1	11	47,83	1	8	25,00	1	34,21	1
Sommersdorf	1	5,26	-1	-7	-30,43	-1	12	37,50	1	7,89	-1
Staven	-3	-15,79	-1	-9	-39,13	-1	-2	-6,25	-1	-18,42	-1
Stuer	-1	-5,26	-1	5	21,74	1	-12	-37,50	-1	-10,53	-1
Torgelow am See	3	15,79	-1	-7	-30,43	-1	0	0,00	1	-5,26	-1
Varchentin	3	15,79	-1	3	13,04	1	0	0,00	1	7,89	-1
Verchen	-1	-5,26	-1	-1	-4,35	1	-4	-12,50	-1	-7,89	-1
Vipperow	7	36,84	1	3	13,04	1	8	25,00	1	23,68	1
Voigtsdorf	-1	-5,26	-1	-13	-56,52	-1	-10	-31,25	-1	-31,58	-1
Median	7	15,79	-1	-5	-21,74	-1	-2	-6,25	-1	-7,89	

Tabelle 3.5 (Fortsetzung): Kategorisierung der Gemeinden der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«; *Quelle: Eigene Darstellung nach Forkel (2014)*

Gemeinde	Punkte Wirtschaft	% Punkte von 18 Variablen	Rang Wirtschaft	Punkte Raum	% Punkte von 23 Variablen	Rang Raum	Punkte Soziales	% Punkte von 32 Variablen	Rang Soziales	gesamt % Punkte von 73	
Vollrathsruehe	3	15,79	-1	1	4,35	1	-6	-18,75	-1	-2,63	-1
Warrenzin	7	36,84	1	-5	-21,74	-1	-6	-18,75	-1	-5,26	-1
Wredenhagen	1	5,26	-1	-5	-21,74	-1	0	0,00	1	-5,26	-1
Zepkow	5	26,32	1	-13	-56,52	-1	-2	-6,25	-1	-13,16	-1
Zettemin	-9	-47,37	-1	-1	-4,35	1	-8	-25,00	-1	-23,68	-1
Zirzow	-3	-15,79	-1	-5	-21,74	-1	6	18,75	1	-2,63	-1
Zislow	3	15,79	-1	1	4,35	1	2	6,25	1	7,89	-1
Godendorf	-1	-5,26	-1	7	30,43	1	-4	-12,50	-1	2,63	-1
Golchen	-7	-36,84	-1	-1	-4,35	1	-10	-31,25	-1	-23,68	-1
Gotthun	3	15,79	-1	3	13,04	1	20	62,50	1	34,21	1
Grabow-Below	-7	-36,84	-1	-9	-39,13	-1	-6	-18,75	-1	-28,95	-1
Grammentin	3	15,79	-1	1	4,35	1	0	0,00	1	5,26	-1
Grapzow	1	5,26	-1	-11	-47,83	-1	4	12,50	1	-7,89	-1
Grischow	-5	-26,32	-1	-7	-30,43	-1	-12	-37,50	-1	-31,58	-1
Median	7	15,79	-1	-5	-21,74	-1	-2	-6,25	-1	-7,89	

Im Anschluss an dieses kategoriespezifische Vorgehen (für die Kategorien Wirtschaft, Raum und Soziales) erfolgte die Sortierung der Orte der Grundgesamtheit entsprechend der Werte des prozentualen Gesamtscores. Der Gesamtscore umfasst hierbei die Summe der in den drei Kategorien erreichten Werte.

Für das im vorangegangenen Abschnitt erläuterte Beispiel der Gemeinde Altenhagen bedeutet das, dass ein Wert von „-21,05%“ errechnet wurde. Bei einem maximalen Punktwert von „+73“ (da 73 Variablen insgesamt berücksichtigt wurden), die Gemeinde auf einer Skala von insgesamt „-100,0%“ (entspricht einem Punktwert von „-73“) bis „+100,0%“ (entspricht einem Punktwert von „+73“) erreichen konnten, finden sich für den Ort Altenhagen einen Wert von „-21,05%“. Nachdem dieses Vorgehen für alle Orte der Grundgesamtheit durchgeführt wurde, wurden die Gemeinden nach ihrem prozentualen Gesamtscore sortiert. Die 24 Gemeinden mit dem geringsten Wert wurden der Kategorie der deprivierten Gemeinden sowie die 24 Gemeinden mit dem größten Wert der Kategorie der arrivierten Gemeinden zugeordnet. Die 24 mittleren Gemeinden wurden der Kategorie der kompensierenden Gemeinden zugeordnet.

Die Zuordnung der Gemeinden zu einer der neu gebildeten Kategorien (depriviert, arriviert, kompensierenden) geht aus Tabelle 3.6 hervor. Darüber hinaus finden sich in der Darstellung die gemeindespezifischen Profilbeschreibungen, die ausweisen, welches kategoriespezifische Ergebnis im Rahmen der Typisierung erreicht werden konnte.

Tabelle 3.6: Übersicht der deprivierten, kompensierenden und arrivierten Gemeinden der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« und deren Ausprägungen in den Kategorien Wirtschaft (Wi), Raum (Ra) und Soziales (So); *Quelle: Eigene Darstellung nach Forkel (2014)*

	deprivierte Gemeinden		kompensierende Gemeinden		arrivierte Gemeinden	
	Name der Gemeinde	Typ	Name der Gemeinde	Typ	Name der Gemeinde	Typ
1	Breest	Wi-Ra-So-	Klocksın	Wi-Ra+So-	Altenhof	Wi+Ra-So+
2	Duckow	Wi-Ra-So-	Kriesow	Wi+Ra-So-	Gnevkow	Wi+Ra-So+
3	Grischow	Wi-Ra-So-	Ritzerow	Wi+Ra-So-	Godendorf	Wi-Ra+So-
4	Voigtsdorf	Wi-Ra-So-	Zepkow	Wi+Ra-So+	Grünow	Wi-Ra+So+
5	Grabow-Below	Wi-Ra-So-	Lindenberg	Wi+Ra-So+	Grammentin	Wi+Ra+So+
6	Genzkow	Wi-Ra-So-	Mildenitz	Wi+Ra-So-	Schönbeck	Wi+Ra+So-
7	Kublank	Wi+Ra-So-	Schwinkendorf	Wi-Ra-So+	Eichhorst	Wi+Ra+So+
8	Neu Gaarz	Wi-Ra-So-	Petersdorf	Wi-Ra-So+	Priborn	Wi-Ra+So+
9	Golchen	Wi-Ra+So-	Stuer	Wi-Ra+So-	Sommersdorf	Wi-Ra-So+
10	Hohenbollentin	Wi-Ra-So-	Briggow	Wi-Ra+So+	Varchentin	Wi+Ra+So+
11	Neetzka	Wi-Ra+So-	Grapzow	Wi-Ra-So+	Zislow	Wi+Ra+So+
12	Zettemin	Wi-Ra+So-	Penkow	Wi-Ra-So+	Buchholz	Wi+Ra+So-
13	Altenhagen	Wi-Ra-So+	Schönhausen	Wi+Ra+So-	Neddemin	Wi+Ra+So+
14	Glienke	Wi-Ra-So+	Verchen	Wi-Ra+So-	Schwarz	Wi+Ra+So-
15	Kentzlin	Wi-Ra-So-	Torgelow am See	Wi+Ra-So+	Meesiger	Wi+Ra+So+
16	Kieve	Wi-Ra+So-	Warrenzin	Wi+Ra-So-	Moltzow	Wi+Ra+So+
17	Massow	Wi-Ra+So-	Wredenhagen	Wi-Ra-So+	Bütow	Wi+Ra+So+
18	Pripsleben	Wi+Ra-So-	Bredenfelde	Wi-Ra+So+	Hohenzieritz	Wi-Ra+So+
19	Schönfeld	Wi+Ra-So-	Gülzow	Wi-Ra-So+	Vipperow	Wi+Ra+So+
20	Beseritz	Wi-Ra-So+	Helpt	Wi+Ra+So-	Leizen	Wi+Ra+So+
21	Kittendorf	Wi-Ra-So+	Priepert	Wi+Ra+So-	Röckwitz	Wi+Ra-So+
22	Staven	Wi-Ra-So+	Vollrathruhe	Wi+Ra+So-	Gotthun	Wi+Ra+So+
23	Melz	Wi+Ra-So-	Zirzow	Wi-Ra-So+	Silz	Wi+Ra+So+
24	Groß Kelle	Wi+Ra-So-	Cammin	Wi+Ra+So-	Ludorf	Wi+Ra+So+

Für die Beurteilung der durch Forkel (2014) vorgenommenen Gemeindetypisierung erfolgte zunächst die Durchführung einer intrakategorialen und einer interkategorialen Validisierung. Hierbei wurden Korrelationsmatrizen (Berechnung der Korrelationskoeffizienten nach Pearson) der in die Gemeindetypisierung eingehenden Variablen vorgenommen. Bei einer Anzahl von drei Kategorien (Wirtschaft, Raum, Soziales) ergeben sich neun (drei intrakategorial, sechs interkategoriale) Korrelationsmatrizen. Wie aus den Tabellen 3.8 (Wirtschaft), Tabellen 3.9 (Raum) und Tabellen 3.10 Soziales ersichtlich wird, finden sich als gute zu bewertende Zusammenhänge zwischen den in die Gemeindetypisierung aufgenommenen Variablen, wobei sich im Vergleich zu den intrakategorialen Darstellungen erwartungsgemäß bei der intrakategorialen Betrachtung eine höhere Anzahl an Korrelationen finden lässt.

Dieses wird auch bei der Quantifizierung der Korrelationen bestätigt. Da die Anzahl der Variablen der einzelnen Kategorien stark differiert (Wirtschaft: 18 Variablen; Raum: 23 Variablen; Soziales: 32 Variablen), erfolgte neben der Betrachtung der absoluten Anzahl der statistisch bedeutsamen Korrelationen auch die Berechnung der Anteilswerte (vgl. Abbildung 3.7). Die höchste Anzahl der signifikanten Korrelationen findet sich bei den intrakategorialen Matrizen (Wirtschaft: 92; Raum: 116; Soziales: 142). In Relation zu der Anzahl der Zellen ergeben sich für die intrakategorialen Matrizen Wirtschaft (28,4%) und Raum (20,1%) die höchsten Anteile an signifikanten Korrelationen. Für die Kategorie Soziales findet sich mit 15,8% ein etwas geringerer Anteil an signifikanten Zusammenhängen, wobei diese Kategorie die mit Abstand größte Anzahl an Indikatoren berücksichtigt. Die geringsten Anteile an statistisch bedeutsamen Korrelationen finden sich bei den interkategorialen Korrelationsmatrizen der Kategorien Raum/Wirtschaft (9,5%) und Raum/Soziales (12,5%).

Tabelle 3.7: Absolute Anzahl und prozentualer Anteil der intra- und interkategorialen Korrelationen der Kategorien Wirtschaft, Raum und Soziales; *Quelle: Eigene Darstellung nach Forkel (2014)*

	Absolute Anzahl der Korrelationen			Prozentualer Anteil der Korrelationen		
	Wirtschaft	Raum	Soziales	Wirtschaft	Raum	Soziales
Wirtschaft	92	41	87	28,4%	9,5%	16,1%
Raum	41	116	90	9,5%	20,1%	12,5%
Soziales	87	90	142	16,1%	12,5%	15,8%

Für die Kategorie Wirtschaft findet sich eine Vielzahl an intrakategorialen Korrelationen verschiedener Einkommens-, Steuer- und Beschäftigungsindikatoren (vgl. Tabelle 3.8). Signifikante Zusammenhänge finden sich bspw. zwischen der Grundsteuer B und dem Anteil sozialversicherungspflichtig Beschäftigter am Arbeitsort ($r=0,516$) sowie dem Anteil der Steuerpflichtigen insgesamt und dem zu versteuernden Einkommen ($r=0,720$). Neben dem Indikator der Grundsteuer B (signifikante Korrelation mit 11 der 17 Variablen) finden sich für den Indikator Haushalte mit Wohngeld (9 der 17 Variablen) eine hohe Anzahl signifikanter intrakategorialer Korrelation (vgl. Tabelle 3.8).

Auch bei einer interkategorialen Betrachtung findet sich eine große Anzahl an signifikanten Zusammenhängen des Indikators Grundsteuer B (signifikante Korrelationen bei 10 von 23 Variablen der Kategorie Raum und bei 11 von 32 Variablen der Kategorie Soziales). Darüber hinaus findet sich bei der interkategorialen Betrachtung der Kategorien Wirtschaft und Raum eine hohe Anzahl an signifikanten Korrelationen für den Indikator der sozialversicherungspflichtig Beschäftigten am Arbeitsort (9 von 23 Variablen). Für die Kategorie Soziales findet sich eine hohe Anzahl an signifikanten Zusammenhängen mit den Wirtschaftsindikatoren sozialversicherungspflichtig Beschäftigte am Arbeitsort (8 von 32 Variablen), Grundsteuer A und B (jeweils 11 von 32 Variablen), Gemeindeanteil an der Einkommenssteuer (13 von 32 Variablen) und Gemeindeanteil an der Umsatzsteuer (12 von 32 Variablen).

Tabelle 3.8: Intrakategoriale Validisierung der Kategorie Wirtschaft und interkategoriale Validisierung der Kategorie Wirtschaft mit den Kategorien Raum und Soziales; *Quelle: Eigene Darstellung nach Forkel (2014)*

	Kategorie Wirtschaft																		
	SVB an der EFB am Wohnort	SVB an der EFB am Arbeitsort	SVB am Arbeitsort nach Wirtschaftsbereichen	Haushalte mit Wohngeld	Steuerpflichtige insgesamt (Lohn- und Einkommenssteuerstatistik)	Gesamtbetrag der Einkünfte	Gesamtbetrag der Einkünfte je Steuerpflichtigen	zu versteuerndes Einkommen	festgesetzte Einkommensteuer	Grundsteuer A	Grundsteuer B	Gewerbesteuer	Gemeindeanteil an der Einkommensteuer	Gemeindeanteil an der Umsatzsteuer	Gewerbesteuerumlage	Gewerbesteuererinnahmen	Schuldenstand der Gemeinden je Einwohner	Schuldenentwicklung	
Wirtschaft	SVB an der EFB am Wohnort	1,000	0,198	0,056	-0,200	-259(**)	-0,115	-0,111	-0,115	-0,113	-0,050	0,034	0,051	0,113	0,096	0,055	0,048	-0,087	0,057
	SVB an der EFB am Arbeitsort	0,198	1,000	-0,087	0,200	-0,131	-0,090	-0,090	-0,089	0,040	0,228	516(**)	0,142	317(**)	323(**)	0,146	0,090	0,039	-0,041
	SVB am Arbeitsort nach Wirtschaftsbereichen	0,056	-0,087	1,000	-0,050	-0,144	-0,034	-0,030	-0,035	-0,036	-0,035	-242(**)	0,015	0,229	-0,161	0,014	-0,012	-0,016	-0,062
	Haushalte mit Wohngeld	-0,200	0,200	-0,050	1,000	454(**)	464(**)	465(**)	464(**)	345(**)	497(**)	339(**)	0,130	346(**)	304(**)	0,129	0,087	0,020	-0,088
	Steuerpflichtige insgesamt (Lohn- und Einkommenssteuerstatistik)	-259(**)	-0,131	-0,144	454(**)	1,000	718(**)	707(**)	720(**)	586(**)	254(**)	0,099	-0,107	0,087	0,064	-0,106	-0,102	-0,034	-0,043
	Gesamtbetrag der Einkünfte	-0,115	-0,090	-0,034	464(**)	718(**)	1,000	0,000(**)	0,000(**)	716(**)	286(**)	0,079	-0,018	0,049	0,064	-0,019	-0,018	0,094	-0,003
	Gesamtbetrag der Einkünfte je Steuerpflichtige	-0,111	-0,090	-0,030	465(**)	707(**)	1,000	999(**)	701(**)	283(**)	0,078	-0,012	0,051	0,068	-0,014	-0,013	0,093	-0,008	
	zu versteuerndes Einkommen	-0,115	-0,089	-0,035	464(**)	720(**)	0,000(**)	999(**)	716(**)	288(**)	0,079	-0,019	0,047	0,063	-0,020	-0,020	0,096	-0,001	
	festgesetzte Einkommensteuer	-0,113	0,040	-0,036	345(**)	586(**)	716(**)	701(**)	722(**)	1,000	326(**)	0,122	-0,080	-0,052	-0,047	-0,078	-0,089	0,125	0,161
	Grundsteuer A	-0,050	0,228	-0,035	497(**)	254(**)	283(**)	288(**)	283(**)	326(**)	1,000	258(**)	261(**)	273(**)	300(**)	263(**)	0,228	0,14	0,002
	Grundsteuer B	0,034	516(**)	-242(**)	339(**)	0,099	0,079	0,078	0,079	0,122	258(**)	1,000	287(**)	700(**)	395(**)	284(**)	240(**)	-0,060	-0,029
	Gewerbesteuer	0,051	0,142	0,015	0,130	-0,107	-0,018	-0,012	-0,019	-0,080	261(**)	287(**)	1,000	0,209	317(**)	999(**)	984(**)	-0,022	0,006
	Gemeindeanteil an der Einkommenssteuer	0,113	317(**)	-0,229	346(**)	0,087	0,049	0,051	0,047	-0,052	273(**)	700(**)	0,209	1,000	465(**)	0,209	0,203	-0,134	-0,027
	Gemeindeanteil an der Umsatzsteuer	0,096	323(**)	-0,161	304(**)	0,064	0,064	0,068	0,063	-0,047	300(**)	395(**)	317(**)	465(**)	1,000	314(**)	-289(**)	-0,066	-0,044
	Gewerbesteuerumlage	0,055	0,146	0,014	0,129	-0,106	-0,019	-0,014	-0,020	-0,078	263(**)	284(**)	999(**)	0,209	314(**)	1,000	986(**)	-0,024	0,009
	Gewerbesteuererinnahmen	0,048	0,090	-0,012	0,087	-0,102	-0,018	-0,013	-0,020	-0,089	0,228	240(**)	984(**)	0,203	289(**)	986(**)	1,000	-0,006	-0,009
	Schuldenstand der Gemeinden je Einwohner	-0,087	0,039	-0,016	0,020	-0,034	0,094	0,093	0,096	0,125	0,014	-0,060	-0,022	-0,134	-0,066	-0,024	-0,006	1,000	0,105
	Schuldenentwicklung	0,057	-0,041	-0,062	-0,088	-0,043	-0,003	-0,008	-0,001	0,161	0,002	-0,029	0,006	-0,027	-0,044	0,009	-0,009	1,000	0,105
Raum	Baufertigstellungen (Wohn- und Nichtwohnbau)	0,092	352(**)	-0,220	-0,019	-0,043	-0,055	-0,055	-0,056	-0,013	-0,035	446(**)	0,104	0,150	0,067	0,099	0,089	-0,091	-0,086
	Gebäudeauslastung: Einwohner pro Wohnfläche	0,043	-0,011	0,095	0,096	0,105	0,028	0,027	0,029	-0,017	0,077	-0,078	0,130	0,129	0,156	0,128	0,130	0,010	0,056
	Gebäude- und Freifläche	0,088	-0,067	-0,063	-0,085	0,027	-0,053	-0,054	-0,055	-0,084	-0,048	0,074	249(**)	0,123	0,075	248(**)	265(**)	-0,100	-0,001
	Betriebsfläche (ohne Abbauwand)	-0,034	-0,026	0,204	-0,052	-0,063	-0,055	-0,053	-0,057	-0,066	0,045	0,166	0,054	0,156	0,048	0,049	0,024	-0,043	0,000
	Erholungsfläche	0,162	0,134	-0,136	-0,152	-0,060	-0,113	-0,112	-0,115	-0,180	-0,226	0,186	0,053	0,058	0,123	0,054	0,068	-0,011	0,033
	Verkehrsfläche	0,106	-0,148	-0,187	-0,167	0,023	0,033	0,039	0,030	-0,110	-0,043	-0,118	0,043	-0,106	-0,060	0,046	0,055	0,021	0,058
	Landwirtschaftsfläche	0,088	-0,203	252(**)	-0,081	0,011	-0,013	-0,014	-0,013	0,043	0,190	-293(**)	0,076	-0,147	-0,040	0,072	0,073	0,116	0,142
	Waldfläche	-0,175	0,161	-0,172	0,128	-0,012	0,057	0,061	0,057	-0,055	-0,161	0,167	-0,046	-0,026	0,037	-0,046	-0,053	-0,090	-0,131
	Wasserfläche	0,104	0,166	-0,149	0,006	0,007	-0,053	-0,058	-0,052	0,038	-0,064	311(**)	-0,076	325(**)	0,026	-0,068	-0,066	-0,063	-0,048
	Erreichbarkeit von Oberzentrum	0,098	443(**)	-0,046	0,033	-0,108	-0,084	-0,084	-0,083	-0,058	-0,025	0,208	-0,014	-0,113	0,109	-0,012	-0,056	-0,081	-0,216
	Erreichbarkeit von Grundzentrum	-0,137	-0,169	-0,143	0,067	-0,004	-0,028	-0,030	-0,030	-0,020	-0,091	-0,018	-0,019	-0,095	0,021	-0,017	0,015	0,178	-0,044
	Erreichbarkeit von Mittelzentrum	0,145	-0,043	0,157	0,083	-0,037	0,192	0,198	0,192	0,146	0,091	0,036	0,119	-0,068	0,096	0,119	0,109	254(**)	0,035
	Einwohnerdichte	0,068	-0,039	-0,103	0,106	-0,057	-0,052	-0,055	-0,055	-0,160	-0,095	0,114	0,144	352(**)	0,184	0,139	0,165	-0,181	-0,155
	Tourismus: geöffnete Beherbergungsbetriebe	0,176	407(**)	-271(**)	-0,080	0,051	-0,069	-0,075	-0,067	0,047	-0,172	389(**)	-0,087	0,170	0,079	-0,082	-0,092	-0,018	-0,121
	Verkehrsunfälle	-0,120	340(**)	-0,139	0,087	-0,081	-0,003	0,003	-0,004	-0,049	375(**)	270(**)	0,200	0,217	343(**)	0,195	0,162	0,078	-0,093
	Lage in einem Landschaftsschutzgebiet	0,143	272(**)	-255(**)	-0,131	-0,021	-0,118	-0,121	-0,117	-0,064	-0,093	297(**)	-0,163	0,114	-0,001	-0,161	-0,160	0,053	-0,166
	Backsteinroute Punkte	0,196	358(**)	-0,100	286(**)	-0,052	0,051	0,051	0,054	0,172	472(**)	446(**)	0,126	270(**)	241(**)	0,132	0,110	0,150	-0,040
	Badestellen	0,152	306(**)	-0,119	-0,027	-0,005	-0,060	-0,065	-0,058	0,052	-0,007	374(**)	-0,055	261(**)	0,017	-0,047	-0,060	-0,039	-0,006
	Breitband (Technologien >=6 mbit)	-0,094	391(**)	-0,154	0,207	0,151	0,081	0,076	0,080	0,105	258(**)	295(**)	-0,102	293(**)	0,148	-0,092	-0,089	-0,077	-0,130
	Projekte der Stiftung Denkmalschutz	0,177	-0,039	-0,153	0,108	0,157	0,141	0,137	0,144	250(**)	0,217	-0,047	-0,096	0,066	0,023	-0,086	0,020	0,222	0,077
	Kirchen Wiki, Kirchen in MV, Denkmal	-0,066	0,050	0,089	0,066	-0,157	-0,060	-0,055	-0,063	-0,131	0,055	0,057	0,043	0,183	0,045	0,048	0,060	0,065	-0,207
Denkmale insgesamt	0,144	246(**)	-0,022	311(**)	-0,067	0,112	0,117	0,113	0,144	326(**)	470(**)	0,038	314(**)	0,158	0,040	0,018	-0,091	0,043	
Technische Denkmale (Mühle, Bahnhof)	-0,044	-0,120	-0,143	-0,022	0,074	0,130	0,131	0,128	0,039	-0,066	0,093	-0,022	0,105	-0,001	-0,029	-0,041	-0,095	0,083	
Ersterverwähnung	0,081	-0,117	-0,076	-0,016	-0,072	-0,010	-0,008	-0,011	-0,073	-0,099	-0,140	0,065	-0,033	-0,039	0,057	0,050	0,176	-0,109	
Soziales	altersbereinigte Frauenquote: Quotient Frau zu Mann	0,199	0,007	0,132	0,011	-0,050	-0,072	-0,073	-0,071	-0,012	0,097	-0,039	0,143	0,132	0,110	0,134	0,124	-0,054	0,096
	Kindheit und Grundschulalter: 0 bis 10-Jährige	0,121	-0,004	0,013	0,070	0,093	-0,009	-0,011	-0,010	0,014	0,065	0,017	-0,128	0,168	-0,126	-0,122	-0,115	-0,212	-0,168
	Jugend und Adoleszenz: 10 bis 25-Jährige	0,072	0,026	0,062	0,157	0,072	-0,035	-0,035	-0,032	-0,020	0,130	0,053	0,212	0,106	0,025	0,203	0,176	0,061	0,084
	Ältere und Alte: Bevölkerung im Alter von 65 Jahren und älter	-0,111	-0,009	0,069	-0,195	-0,131	-0,037	-0,037	-0,037	-0,007	-0,227	-0,041	-0,179	315(**)	-0,167	-0,171	0,150	0,159	-0,174
	Geborene pro Einwohner (Geborene-Kopf)	0,081	315(**)	0,175	378(**)	-0,164	-0,078	-0,080	-0,078	-0,007	-0,196	470(**)	-0,207	434(**)	-293(**)	-0,205	-0,180	0,007	-0,109
	Gestorbene pro Einwohner (Gestorbene-Kopf)	-0,113	-0,166	0,063	-0,196	-0,035	-0,010	-0,014	-0,010	0,072	-0,151	326(**)	-0,137	-272(**)	-0,173	-0,134	-0,123	0,059	0,008
	Zuzüge pro Einwohner	0,139	267(**)	0,118	423(**)	-0,185	-0,091	-0,090	-0,092	-0,055	256(**)	565(**)	-0,190	498(**)	-293(**)	-0,189	-0,165	0,123	-0,131
	Fortzüge pro Einwohner	0,036	324(**)	-0,136	401(**)	-0,218	-0,128	-0,130	-0,129	-0,024	-247(**)	581(**)	-0,228	523(**)	-301(**)	-0,231	-0,208	0,151	-0,157
	SVB am Arbeitsort nach Wirtschaftsbereichen	0,056	-0,087	0,000(**)	-0,050	-0,144	-0,034	-0,030	-0,035	-0,036	-0,035	-242(**)	0,015	-0,229	-0,161	0,014	-0,012	-0,016	-0,062
	SVB am Arbeitsort nach Wirtschaftsbereichen	-0,026	0,023	-345(**)	0,029	0,080	-0,105	-0,107	-0,104	-0,103	0,104	0,176	-0,117	0,175	244(**)	-0,119	-0,128	-0,039	0,213
	SVB am Arbeitsort nach Wirtschaftsbereichen	-0,127	0,110	375(**)	0,065	-0,097	-0,038	-0,038	-0,040	0,008	0,038	0,107	239(**)	0,048	0,061	237(**)	254(**)	0,221	-0,070
	Erbringung von öffentlichen und privaten Dienstleistungen	0,170	0,191	-0,164	-0,137	-0,064	-0,044	-0,051	-0,040	0,194	-0,097	0,152	-0,088	-0,043	0,010	-0,089	-0,081	-0,135	0,027
	Einpendlerquote der EWFB	255(**)	951(**)	-0,085	0,106	-0,162	-0,079	-0,078	-0,078	0,015	0,177	445(**)	0,166	287(**)	291(**)	0,168	0,117	0,032	-0,020
	Anteil der in der Gemeinde Beschäftigten	0,229</																	

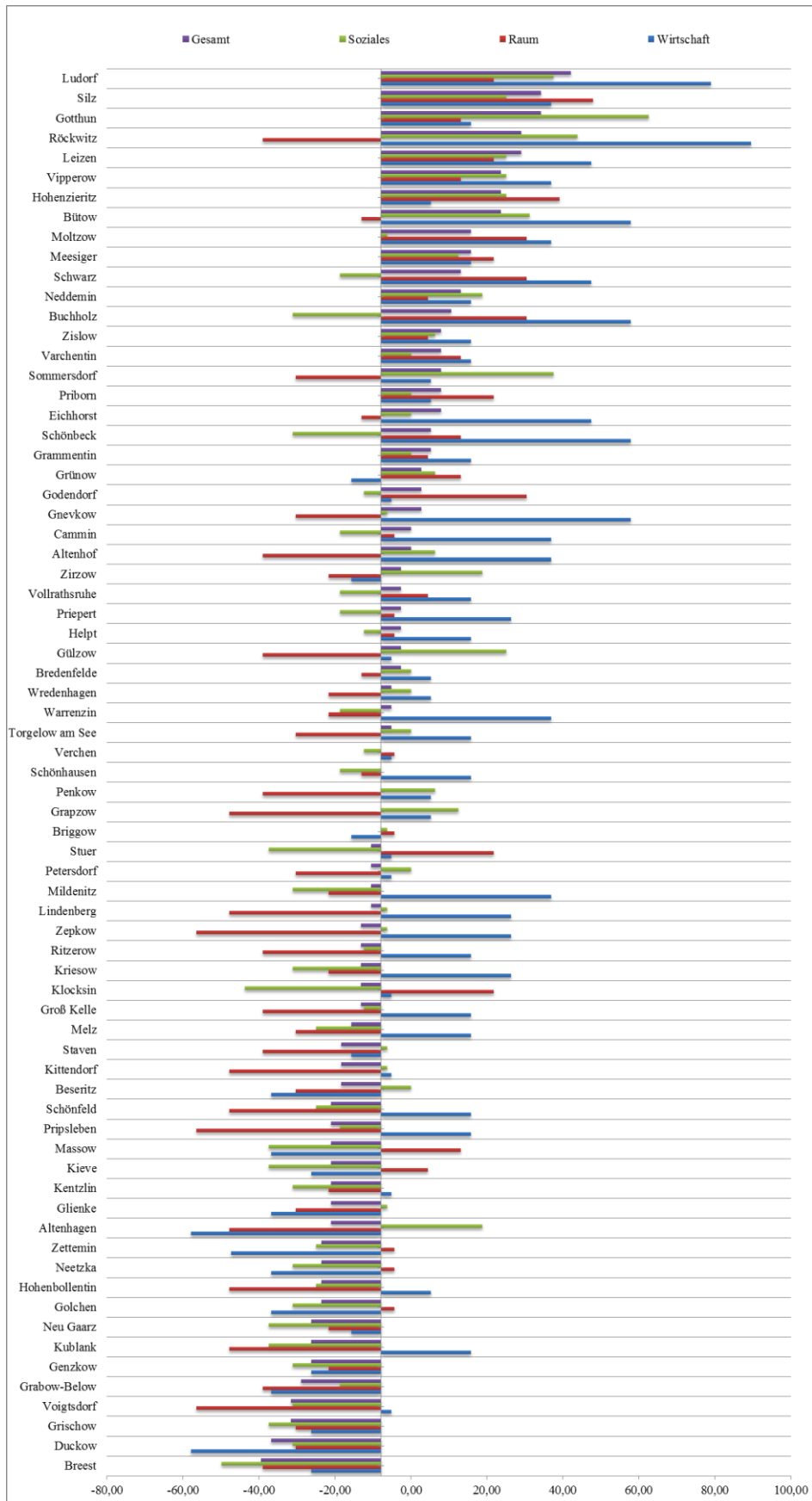
ren Anteil der Haushalte mit Wohngeld (signifikante Korrelationen bei 10 von 18 Variablen der Kategorie Wirtschaft) sowie des Indikators Treffpunkte (signifikante Korrelationen bei 8 von 18 Variablen der Kategorie Wirtschaft). Die meisten signifikanten Korrelationen mit Variablen der Kategorie Raum finden sich für die Indikatoren Unternehmen, Tourismus und Handwerke (9 von 23 Variablen), altersbereinigte Frauenquote, Einpendlerquote und Treffpunkte (jeweils 8 von 23 Variablen).

Tabelle 3.10: Intrakategoriale Validisierung der Kategorie Soziales und interkategoriale Validisierung der Kategorie Soziales mit den Kategorien Wirtschaft und Raum; *Quelle: Eigene Darstellung nach Forkel (2014)*

	Kategorie Soziales																	
	abnehmende Altersquote: Anteil von Kindern im Alter von 0 bis 14 Jahren pro Einwohner (10 bis 25-Jährige pro Einwohner)	Anteil von Kindern im Alter von 15 bis 19 Jahren pro Einwohner	Anteil von Kindern im Alter von 20 bis 24 Jahren pro Einwohner	Anteil von Kindern im Alter von 25 bis 29 Jahren pro Einwohner	Anteil von Kindern im Alter von 30 bis 34 Jahren pro Einwohner	Anteil von Kindern im Alter von 35 bis 39 Jahren pro Einwohner	Anteil von Kindern im Alter von 40 bis 44 Jahren pro Einwohner	Anteil von Kindern im Alter von 45 bis 49 Jahren pro Einwohner	Anteil von Kindern im Alter von 50 bis 54 Jahren pro Einwohner	Anteil von Kindern im Alter von 55 bis 59 Jahren pro Einwohner	Anteil von Kindern im Alter von 60 bis 64 Jahren pro Einwohner	Anteil von Kindern im Alter von 65 bis 69 Jahren pro Einwohner	Anteil von Kindern im Alter von 70 bis 74 Jahren pro Einwohner	Anteil von Kindern im Alter von 75 bis 79 Jahren pro Einwohner	Anteil von Kindern im Alter von 80 bis 84 Jahren pro Einwohner	Anteil von Kindern im Alter von 85 bis 89 Jahren pro Einwohner	Anteil von Kindern im Alter von 90 bis 94 Jahren pro Einwohner	Anteil von Kindern im Alter von 95 bis 99 Jahren pro Einwohner
Wirtschaft	SVB an die ERB als Wohnort	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	SVB an die ERB als Arbeitsort	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Haushalte mit Wohngeld	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Strategische Investition (Lohn- und Entgelt)	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Gesamtwert der Einkünfte je Steuerpflichtiger	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Ergebnisse der Einkommenssteuer	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Grundsicherung A	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Grundsicherung B	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Gesundheit	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Gesundheit II	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
Raum	Erreichbarkeit von Oberzentrum	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Mittelzentrum	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Untermittelzentrum	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Dienstleistungszentrum	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Metropolregion	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Großraum	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Erreichbarkeit von Metropolregion	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Erreichbarkeit von Metropolregion	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Erreichbarkeit von Metropolregion	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Erreichbarkeit von Metropolregion	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
Soziales	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	
	Verkehrsmittel	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	

Neben einer insgesamt als gut zu beurteilenden intra- und interkategorialen Validität erfolgte eine Beurteilung des Verfahrens auf Grundlage der in Abbildung 3.1 visualisierten Scorewerte der Kategorien Wirtschaft, Raum und Soziales sowie dem daraus berechneten Gesamtscore. Die Orte der Grundgesamtheit sind hierbei nach den berechneten Gesamtscores in absteigender Reihenfolge sortiert. Wie aus der gemeindeindividuellen Darstellung erkennbar wird, finden sich Gemeinden mit überdurchschnittlichen Werten sowie Gemeinden mit deutlich unterdurchschnittlichen Scorewerten. Da über die Hälfte der Gemeinden in der untersuchten Region einen unterdurchschnittlichen Gesamtscore aufweist, ist es nicht verwunderlich, dass die meisten Gemeinden einen unterdurchschnittlichen Gesamtscore aufweisen. Die ausgerechneten Gesamtscorewerte sind in der Abbildung 3.1 dargestellt. Die ausgerechneten Gesamtscorewerte sind in der Abbildung 3.1 dargestellt. Die ausgerechneten Gesamtscorewerte sind in der Abbildung 3.1 dargestellt.

Abbildung 3.2: Übersicht der prozentualen Punktwerte der Orte der Grundgesamtheit des Projekts »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« für die Kategorien Wirtschaft, Raum und Soziales sowie der Gesamtscore; *Quelle: Forkel (2014)*



Ungeachtet der – wie gezeigt werden konnte – vorhandenen augenscheinlichen Validität des für die Ortsauswahl herangezogenen Verfahrens muss grundlegend angemerkt werden, dass es sich bei den verwendeten Variablen in allen drei Kategorien jeweils um teilweise sehr unterschiedliche Indikatoren handelt. Diese starke Heterogenität erschwert es, die Bedeutung der Eignung der jeweiligen Indikatorensets grundsätzlich zu beurteilen. Bedingt durch das Ziel der Verwendung der Kategorisierung für die Auswahl der Untersuchungsgemeinden der Projektes LETHE wird diese Indikatorenheranziehung (vgl. Forkel 2014) hier dennoch als angemessen beurteilt, da das primäre Ziel dieses Verfahrens die Auswahl nicht-homogener Untersuchungsorte aus der Stichprobe ist.

3.3.2 Stichprobenziehung auf Grundlage der Gemeindetypisierung

Auf Grundlage der durch die Gemeindetypisierung vorgenommenen Differenzierung der Gemeinden der Grundgesamtheit erfolgte eine Kategorisierung der Ort in deprivierte, kompensierende und arrivierte Gemeinden (vgl. Forkel 2013). Für diese Kategorien, die jeweils 24 Gemeinden der Grundgesamtheit umfassten, erfolgte die Auswahl von jeweils drei Orten mittels der Ziehung einer Zufallsstichprobe. Die technische Umsetzung dieses Verfahrens mit der Statistiksoftware **R** wird in Kapitel 4.2.3 näher erläutert.

3.4 Zusammenfassung

Im vorliegenden Kapitel erfolgte eine Diskussion verschiedener Möglichkeiten der Ortsauswahl für das Projekt LETHE. Neben der Ziehung einer Zufallsstichprobe für die Grundgesamtheit der Gemeinden insgesamt wurde die Möglichkeit der Ziehung einer Zufallsstichprobe, die nach den ehemaligen Landkreisen Mecklenburg-Vorpommerns differenziert, diskutiert. Beide Möglichkeiten bieten keine bzw. nur geringe Möglichkeiten strukturell unterschiedliche Gemeinden auszuwählen. Die nach den Landkreisstrukturen vor der Kreisgebietsreform differenzierte Stichprobenziehung ermöglicht lediglich die Berücksichtigung eines geographischen Merkmals. Darüber hinaus erfolgte die Beschreibung einer Stichprobenziehung, die zwischen deprivierten, kompensierenden und arrivierten Gemeinden differenziert. Die Zuordnung der Gemeinden zu diesen drei Kategorien erfolgte auf Grundlage der von Forkel (2014) vorgenommenen Gemeindetypisierung. Neben einer kurzen Beschreibung methodischer Aspekte des Verfahrens erfolgte der Versuch einer Einschätzung der Validität. Die intra- und interkategoriale Beurteilung der Validität erfolgte durch die Berechnung von Korrelationsmatrizen. Auf Grundlage dieser Auswertungen sowie einer als hoch zu bewertenden Augenscheinvalidität wird das von Forkel (2014) durchgeführte Verfahren als valide erachtet. Trotz der konstatierten Heterogenität der im Rahmen der Ortstypisierung berücksichtigten Variablen konnte das primäre Ziel dieser Methodik, eine topische Grundgesamtheit zu differenzieren, erreicht werden.

4 COMPUTERGESTÜTZTE ZIEHUNG EINER ZUFALLS-STICHPROBE MIT DEM PROGRAMM R

Die technische Umsetzung der Ziehung der Zufallsstichprobe(n) für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« erfolgte mit der Statistiksoftware **R**. Im Vorfeld der Erläuterungen der Analyseschritte der Stichprobenziehung (Kap. 4.2) erfolgen zunächst einleitende und grundlegende Anmerkungen zum Statistikprogramm **R** (Kap. 4.1).

4.1 Das Statistikprogramm R

Bei dem Programm **R** handelt es sich um eine „*Statistik- und Visualisierungssoftware, die durch die R Foundation koordiniert wird*“ (Dormann 2013: XIV). Der Name „*bezeichnet [...] sowohl das Programm selbst als auch die Sprache, in der die Auswertungsbefehle geschrieben werden*“ (Wollschläger 2012: 1). Entgegen statistischen Programmen wie bspw. IBM SPSS Statistics handelt es sich bei dem Programm **R** um eine „*befehlsbasierte Statistikumgebung, d.h. anstelle einer menügeführten Auswahl von Variablen und Auswertungstechniken tritt die schriftliche Eingabe von Befehlen*“ (Manderscheid 2012: 37).

Das Programm **R** basiert auf den Ideen von **S** (vgl. u.a. Ligges 2009: 3f.), einer „*für Statistik, stochastische Simulation und Grafik entwickelte Programmiersprache*“ (Ligges 2009: 3).⁵ Wie Ligges (2009: 4) anmerkt, ist nicht genau bekannt, „*wie es zum Namen **R** kam.*“ Die beiden naheliegendsten Erklärungen besagen, dass der Name „*entstand, da Ross Ihakas und Robert Gentleman's Vornamen [Entwickler des Programms] jeweils mit **R** beginnen, oder weil man eine ‚reduced version of **S**‘ wollte*“ (Ligges 2009: 4). Seit dem Jahr 1992 kann eine „*rasante Entwicklung von **R***“ (Ligges 2009: 4) beobachtet werden.⁶

Bei dem Programm **R** handelt es sich „*um Open Source Software*“ (Ligges 2009: 5), die auf der [R-Homepage](#) „*frei verfügbar und einfach zu installieren*“ ist (Kauermann/Küchenhoff 2011: 247).⁷ Es erfolgen in regelmäßigen Abständen Aktualisierungen. Wie Dormann (2013: XVI) anmerkt, findet sich für die Software „*mit zwei major releases (2.14.0, 2.15.0, 3.0.0, ...) pro Jahr plus jeweils 2 patches (2.15.1, 2.15.2, ...)*“ ein recht hoher Aktualisierungsrhythmus. Daraus ergibt sich nicht, dass „*eine zwei Jahre alte Version [...] überholt [ist], aber neue Pakete werden für sie nicht mehr kompiliert und brauchen u.U. eine aktuelle R-Version.*“ (Dormann 2013: XVI).⁸ Nicht zuletzt aus diesem Grund, dass die für das Programm **R** verfügbaren Pakete „*fortwährend und unabhängig von der Gesamtversion **R** überarbeitet [werden], [ist] es sinnvoll [...], immer die neuesten Versionen zu verwenden.*“ (Manderscheid 2012: 27).

Die Arbeitsumgebung des Programms **R** „*besteht in der Basisversion aus der **R** Konsole, in die Befehle eingeben und Ausgaben angezeigt werden können*“ (Manderscheid 2012: 24). Der Zei-

⁵ Zu den Ursprüngen der Programmiersprache **S** siehe u.a. Ligges (2009: 3).

⁶ Zur Geschichte des Programms **R** und deren Entwicklung siehe u.a. Ligges (2009: 4); Hornik/Leisch (2002: 63ff.).

⁷ Nähere Informationen zur Installation des Programms finden sich u.a. bei Dormann (2013: XIVf.), Wollschläger (2012: 4f.), Kohn/Öztürk (2013: 3), Manderscheid (2012: 21ff.).

⁸ Durch Eingabe des Befehls **sessionInfo()** werden in **R** die jeweils installierte Version, der jeweilige Modus (64-bit bzw. 32-bit-Version) des Programms sowie die aktuell geladenen Pakete angezeigt.

lenbeginn in der **R**-Konsole wird durch den prompt (**>**) gekennzeichnet.⁹ Neben der Möglichkeit Befehle direkt in die **R**-Konsole einzugeben und mit der Enter-Taste zu bestätigen (vgl. Manderscheid 2012: 25), kann „für eine systematische Datenauswertung [...] ein Arbeiten mit Skriptdateien über einen Editor [...] sinnvoller“ (Manderscheid 2012: 25) sein. Die in dem **R**-Editor geschriebenen „Befehle lassen sich über [die Tastenkombination] **Strg** + **R** direkt aus dem Skript ausführen“ (Manderscheid 2012: 28). Wenngleich gegenüber der Verwendung eines Programmmenüs mit der Programmieren vom Befehlen in einem Editor eine große Anzahl von Vorteilen einhergeht, müssen bei der Eingabe von Befehle eine Vielzahl von Regeln beachtet werden (siehe u.a. Manderscheid 2012: 39).

Alternativ zu der Standardbenutzeroberfläche (**R** Konsole), die über kein Menü verfügt, finden sich verschiedene graphische Benutzeroberflächen (vgl. u.a. Dormann 2013: XVIIIff., Manderscheid 2012: 30-34; Wollschläger 2012: 5). Zu den am häufigsten in der Literatur verwiesenen graphische Benutzeroberflächen gehören der **R** Commander, **R**Studio sowie **JGR** (**J**ava **G**UI for **R**). Bei den graphischen Benutzeroberflächen **R**Studio und **JGR** (jaguar) handelt es sich um „Java-Programme, die **R** in eine neue Oberfläche integrieren (und deshalb auch oft als IDE, integrated development environment, bezeichnet werden)“ (Dormann 2013: XVIII). Wie von Dormann (2013: XVIII) anmerkt, laufen beide Anwendungen „nicht ganz so stabil wie das **R**-GUI [graphical user interface] selbst, sind aber trotzdem [...] zu empfehlen.“ Die graphische Benutzeroberfläche **R** Commander „ermöglicht eine rein menügesteuerte Datenauswertung ohne manuelle Befehlseingabe“ (Manderscheid 2012: 33). Neben dem Vorteil einer im Vergleich zum Programmieren in einem Editor vergleichsweise einfachen Verwendung, ist die Nutzung „auf die Menüs und deren zugrunde liegenden Funktionen beschränkt“ (Manderscheid 2012: 34).

Die Basisversion des Programms „enthält etwa 25 Pakete für die wichtigsten statistischen Verfahren und Graphiken“ (Manderscheid 2012: 25).¹⁰ Bei den Paketen handelt es sich „um Erweiterungen für **R**, die zusätzliche Funktionalität, Funktionen, Datensätze usw. bereitstellen.“ (Ligges 2009: 189). Diese müssen gesondert installiert und nach jedem Öffnen des Programms geladen werden (vgl. Manderscheid 2012: 21). Eine stetig wachsende Anzahl an Paketen findet sich bei **CRAN** (**C**omprehensive **R** **A**rchive **N**etwork). Hierbei „handelt es sich um ein Netz von Servern, das weltweit den Quellcode und die Binärdateien für diverse Betriebssysteme bereitstellt. Der zentrale Server ist unter <http://cran.r-project.org/> zu erreichen. Dort gibt es u.a. [...] Zusatzpakete zu diversen (statistischen) Verfahren sowie Publikationen, Dokumentation, FAQs, Zugriff auf Mailinglisten, Links zu anderer relevanter Software und Projekten.“ (Ligges 2009: 1). Derzeit finden sich mehr als 3.000 Zusatzpakete für das Programm **R** (Manderscheid 2012: 15), wobei auch zukünftig von einer wachsenden Anzahl auszugehen ist. Da „nicht alle installierten Pakete [...] automatisch mit dem Starten von **R** geladen [werden], da sonst das Programm zu groß und dadurch langsam werden würde“ (Manderscheid 2012: 26), müssen Pakete, die nicht Teil der Basisinstallation sind, vor jeder Analysesitzung (nach Starten des Programms) einmalig mit dem Befehl **library(Paketname)** geladen werden.

⁹ Alternativ zu der Standardeinstellung können über die Eingabe des Befehls **options(prompt=" ")** alternative Zeichen und Symbole (bspw. Leerzeichen) als Zeilenbeginn der **R**-Konsole verwendet werden.

¹⁰ Durch den Befehl **library()** werden alle installierten sowie die verfügbaren Pakete angezeigt. Eine Auflistung der installierten Pakete mit den jeweiligen Versionsnamen, Pfadangaben und Abhängigkeiten ist mit dem Befehl **library.packages()** möglich. Darüber hinaus können alle für das Programm **R** verfügbaren Pakete mit dem Befehl **available.packages()** angezeigt werden (vgl. u.a. Manderscheid 2012: 25).

Mit der Verwendung der Software **R** geht eine Vielzahl von Vorteilen einher. Bei dem Statistikprogramm **R** handelt es sich um eine kostenlose Software, die für verschiedene Betriebssysteme (Windows, Linux, Macintosh) verfügbar ist. Das Basissystem des Programms **R** benötigt mit weniger als 50MB lediglich einen geringen Speicherplatz (vgl. Viton 2012: 2). Da es sich um ein „*freies Open Source-Programm [handelt,] wird R beständig von vielen Personen evaluiert, weiterentwickelt und den praktischen Erfordernissen angepasst. Da der Quelltext frei verfügbar ist und zudem viele Auswertungsfunktionen ihrerseits in R geschrieben sind, ist die Art der Berechnung statistischer Kennwerte vollständig transparent. Sie kann damit vom Benutzer bei Interesse analysiert, aber auch auf Richtigkeit kontrolliert werden.*“ (Wollschläger 2012: 2). Darüber hinaus findet sich mit den Mitgliedern des **R Core Teams**¹¹ sowie einer Vielzahl von freiwilligen Helfern ein „*großartiger Support*“ (Ligges 2009: 5). Da es sich bei dem Programm **R** um eine „*befehlsbasierte Statistikumgebung*“ (Manderscheid 2012: 37) handelt, wird „*eine Integration immer neuer Pakete*“ (Manderscheid 2012: 37) und damit einhergehend neuen Möglichkeiten der Datenanalyse ermöglicht. Ein wesentlicher Vorteil der „*Trennung von Eingabe und Ausgabe [in Form des R-Editors und der R-Konsole] hat den Vorteil – ähnlich wie in IBM SPSS oder Stata die Syntaxfunktion – dass Befehle und Funktionen übersichtlich organisiert, und jeder Zeit wieder ausgeführt werden können*“ (Manderscheid 2012: 28f.).

Neben vielfältigen statistischen Analysemöglichkeiten sowie umfangreichen Möglichkeiten der graphischen Visualisierung kann mit dem Programm **R** auch die Ziehung von Stichproben erfolgen. Auf die syntaktische Umsetzung der Ziehung einer Stichprobe für das Projekt LETHE wird im Folgenden näher eingegangen (vgl. Kap. 4.2).

4.2 Ziehung einer Zufallsstichprobe mit dem Programm R

Im Vorfeld der Erläuterungen der verschiedenen Möglichkeiten eine Stichprobe für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« zu ziehen, wird zunächst allgemein die technische Umsetzung der Stichprobenziehung mit dem Programm **R** anhand der syntaktischen Befehle erläutert. Anschließend wird auf die drei verschiedenen Möglichkeiten der Stichprobenziehung (Zufallsstichprobe für die Grundgesamtheit insgesamt, differenziert nach Gebietsstrukturen vor Kreisgebietsreform 2011 sowie auf Grundlage der Gemeindetypisierung) mit Hilfe der **R**-Syntax eingegangen.

Die folgenden Ausführungen zur technischen Umsetzung der Ziehung einer Zufallsstichprobe mit dem Programm **R** basieren auf den Ausführungen bei Kauermann/Küchenhoff (2011: 47-51). Das Ziehen einer Zufallsstichprobe kann mit dem Programm **R** mit der Funktion **sample(x, size, replace = FALSE, prob = NULL)** vorgenommen werden. Die Funktion **sample()** hat zur Folge, dass aus dem Vektor **x** eine Zufallsstichprobe mit dem Umfang **size** gezogen wird. Als Standardeinstellung ist das Argument **replace** (Zurücklegen gezogener Werte) auf **FALSE** (kein Zurücklegen) eingestellt. Mittels des Befehls **replace=TRUE** können die gezogenen Werte wieder zurückgelegt werden. Mit dem Argument **prob** können für die einzelnen Elemente der Grundgesamtheit (Vektor **x**) unterschiedliche Wahrscheinlichkeiten definiert werden, in die Stichprobe gezogen zu werden.¹² Als Standardeinstellung findet sich **prob=NULL**, sodass alle

¹¹ An dieser Stelle sei noch einmal auf Homepage des Projects **R** (<http://www.r-project.org/>) und den dort verfügbaren Manuels verwiesen.

¹² Mit dem Befehl **prob** können die Wahrscheinlichkeiten der Elemente der Grundgesamtheit gezogen zu werden, definiert werden. Für eine Stichprobenziehung mit dem Umfang **size=1** aus der Grundgesamtheit **x** (Vektor mit

Elemente der Grundgesamtheit eine identische Wahrscheinlichkeit aufweisen, gezogen zu werden (vgl. Kauermann/Küchenhoff 2011: 48).

```
[1]
> x <- 1:49           # Definition des Vektors x (Zahlen 1 bis 49)
> x                 # Darstellung des Vektors x
> sample(x, 6, replace=F, prob=NULL) # Zufallsstichprobe 6 aus 49 (unsortiert)
> sort(sample(x, 6, replace=F, prob=NULL)) # Zufallsstichprobe 6 aus 49 (sortiert)
```

Soll bspw. bei einer einfachen Zufallsstichprobe die Auswahl von sechs Zahlen aus einer Grundgesamtheit von 49 Zahlen erfolgen, muss zunächst ein Vektor **x** (Grundgesamt) definiert werden (**x <- 1:49**). Nach der Generierung des Vektors kann dieser mit dem Befehl **x** dargestellt bzw. vom Programm ausgewiesen werden. Die Ziehung der Zufallsstichprobe erfolgt mit dem Befehl **sample(x, 6, replace=F, prob=NULL)**, wobei **x** die Grundgesamt (Werten 1 bis 49) sowie die Zahl **6** den Stichprobenumfang darstellen. Mit dem Argument **replace=F** wird ein Zurücklegen der bereits gezogen Werte verhindert. Durch das Argument **prob=NULL** haben alle Elemente des Vektors **x** die identische Wahrscheinlichkeit gezogen zu werden. Da es sich um die Standardeinstellungen handelt, können die beiden letztgenannten Argumente (**replace** und **prob**) auch weggelassen werden. Mit dem zusätzlichen Befehl **sort()** lassen sich die Ergebnisse der Zufallsstichprobe in numerisch aufsteigender Reihenfolge darstellen.

Mit der durch den Befehl **sort(sample(x, 6, replace=F, prob=NULL))** durchgeführten Ziehung einer Zufallsstichprobe von sechs Werten aus einer Grundgesamtheit von 49 Zahlen geht das Problem einer fehlenden Reproduzierbarkeit des Ergebnisses einher. Jeder erneute Aufruf des Befehls würde eine erneute Zufallsstichprobe generieren, die unterschiedliche Werte generiert. Aus diesem Grund *„empfiehlt es sich, mit einem sogenannten „seed“, einem Startwert für den Algorithmus (in diesem Fall zum Ziehen von Zufallszahlen), zu arbeiten. Dadurch wird sichergestellt, dass bei einem erneuten Ausführen die gleiche Stichprobe gezogen wird.“* (Kauermann/Küchenhoff 2011: 252). Der Startwert des Algorithmus kann *„auf zweierlei Art festgelegt werden. Einerseits kann einfach eine beliebige Zahl als Startwert gewählt werden [...]. Andererseits kann der Startwert einmalig zufällig gezogen werden“* (Kauermann/Küchenhoff 2011: 253).

```
[2]
> start <- 20131115 # Definition des Startwertes des Algorithmus
> start <- sample(x=1:1000, size=1) # Ermittlung eines Startwertes des Algorithmus
# durch eine Zufallsstichprobe
```

Im vorliegenden Beispiel erfolgt die Definition des Objekts **start** (Startwert des Algorithmus) als Wert **20131115** bzw. mittels einer Zufallsstichprobe über dem Befehl **start <- sample(x=1:1000, size=1)**. Im zweiten Fall ist die Grundgesamtheit **x** als Vektor der Zahlen 1 bis 1.000 (**1:1000**) definiert. Der Stichprobenumfang beträgt eins (**size=1**). Die Voreinstellungen (**replace=FALSE** und **prob=NULL**) werden beibehalten.

Bei mehrmaligen Ziehen der Zufallsstichprobe darf der Befehl **start <- sample(x=1:1000, size=1)** lediglich einmal (bei der ersten Ziehung der Zufallsstichprobe) erfolgen, da ansonsten

dem Werten 1, 2 und 3) sollen unterschiedliche Wahrscheinlichkeiten der Elemente der Grundgesamtheit definiert werden, gezogen zu werden. Mit dem Befehl **sample(x=c(1:3), size=1, replace=F, prob=c(0.1, 0.1, 0.8))** wird für die ersten beiden Elemente des Vektors der Grundgesamtheit **x** (Werte 1 und 2) eine Wahrscheinlichkeit von jeweils 10% (**0.1**) sowie für das dritte Element des Vektors (Wert 3) eine Wahrscheinlichkeit von 80% (**0.8**) definiert. Die im Befehl **prob=c(0.1, 0.1, 0.8)** angegebenen Wahrscheinlichkeiten müssen der Anzahl der Werte des Vektors der Grundgesamtheit entsprechen. Beide Vektorlängen (**x=c()** und **prob=c()**) müssen identisch sein.

bei jeder erneuten Generierung der Stichprobe ein neuer mittels Zufallsstichprobe ermittelter Startwert des Algorithmus verwendet werden würde. Mit dem Befehl **set.seed(start)** „*wird der Startwert dem Programm als solcher kenntlich gemacht*“ (Kauermann/Küchenhoff 2011: 253). Nachfolgend findet sich die Syntax für das Ziehen einer Zufallsstichprobe von sechs Zahlen aus einer Grundgesamtheit von 49 Zahlen, unter Verwendung eines mittels Zufallsstichprobe ermitteltem Startwert des Algorithmus.

```
[3]
> x <- 1:49                # Definition des Vektors x (Zahlen 1 bis 49)
> x                        # Darstellung des Vektors x
> start <- sample(x=1:1000, size=1)  # Ermittlung eines Startwertes des Algorithmus
                                     # durch eine Zufallsstichprobe

> set.seed(start)
> sort(sample(x, 6, replace=F, prob=NULL))  # Zufallsstichprobe 6 aus 49 (sortiert)
```

4.2.1 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« aus der Grundgesamtheit

Als erste Möglichkeit eine Auswahl der Untersuchungsgemeinden für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« zu treffen, kommt die Ziehung einer Zufallsstichprobe aus der Grundgesamtheit (N=72) in Betracht. Im Folgenden wird die Umsetzung dieser Stichprobenziehung mit dem Programm **R** anhand der syntaktischen Befehle erläutert.

Im Vorfeld der Stichprobenziehung muss zunächst eine Definition der Grundgesamtheit (N) sowie des Stichprobenumfangs (n) erfolgen. Hierfür wird dem Objekt **N** der Wert **72 (N <- 72)** sowie dem Objekt **n** der Wert **9 (n <- 9)** zugeordnet. Anschließend werden dem Objekt **Grundgesamtheit_LETHE_ORTE** die Namen der 72 Gemeinden der Grundgesamtheit in alphabetischer Reihenfolge zugeordnet. Die Liste der alphabetischen Gemeinden der Grundgesamtheit wird mit dem Befehl **Grundgesamtheit_LETHE_ORTE <- c("Altenhagen", "Altenhof", [...], "Zirzow", "Zislow")** definiert. Der Befehl **c()** führt zu einer Kombination der 72 Gemeindefamen zu dem Objekt **Grundgesamtheit_LETHE_ORTE**.

Um eine Reproduzierbarkeit der Stichprobenziehung zu gewährleisten (identische Ergebnisse auch bei mehrmaligen Ziehungen), erfolgt eine Definition des Startwertes des Algorithmus der Stichprobenziehung. Mit dem Befehl **start <- sample(x=1:10000, size=1)** wird das Objekt **start** als ein Zufallswert aus dem numerischen Vektor **x=1:10000** (Zahlen 1 bis 10000) definiert. Die Funktion **size=1** legt fest, dass es sich bei der generierten Zufallszahl um eine Zahl handelt. Mit dem Befehl **start** wird der erzeugte Startwert des Algorithmus (zur Kontrolle) angezeigt. Durch den Befehl **set.seed(start)** wird der im vorangegangenen Schritt durch eine Zufallsziehung erzeugte Wert als Startwert des Algorithmus für die Ziehung der Zufallsstichprobe definiert. Nachdem der Startwert für den Algorithmus erzeugt wurde, muss der Wert anstelle der Befehls **start** in die Funktion **set.seed(start)** eingetragen werden. Die Ziehung der Zufallsstichprobe (Auswahl der Gemeinden aus der Grundgesamtheit insgesamt) erfolgt mit dem Befehl **Stichprobe_Gesamt <- sort(sample(x = Grundgesamtheit_LETHE_ORTE, size=n))**. Das Objekt **Stichprobe_Gesamt** (Zufallsstichprobe der Gemeinden) wird durch den Befehl **sample(x= ...)** erzeugt, wobei **size=n** den Stichprobenumfang (n=9) definiert. Der eigentliche Befehl für die Ziehung der Zufallsstichprobe (**sample(x= ...)**) wird hierbei von dem Befehle **sort(...)** umschlossen. Auf diesem Weg werden die mittels der Zufallsstichprobe gezogenen Ergebnisse (Untersuchungsorte) alphabetisch sortiert.

[4]

```

# Stichprobenziehung aus der Grundgesamtheit insgesamt (N=72)

# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N <- 72
> n <- 9
> Grundgesamtheit_LETHE_ORTE <- c("Altenhagen", "Altenhof", "Beseritz", "Bredenfelde",
+ "Breest", "Briggow", "Buchholz", "Bütow", "Cammin", "Duckow",
+ "Eichhorst", "Genzkow", "Glienke", "Gnevkow", "Godendorf",
+ "Golchen", "Gotthun", "Grabow-Below", "Grammentin",
+ "Grapzow", "Grischow", "Groß Kelle", "Grünow", "Gülzow",
+ "Helpt", "Hohenbollentin", "Hohenzieritz", "Kentzlin", "Kieve",
+ "Kittendorf", "Klocksın", "Kriesow", "Kublank", "Leizen",
+ "Lindenberg", "Ludorf", "Massow", "Meesiger", "Melz", "Mildenitz",
+ "Moltzow", "Neddemin", "Neetzka", "Neu Gaarz", "Penkow",
+ "Petersdorf", "Priborn", "Priepert", "Pripsleben", "Ritzerow",
+ "Röckwitz", "Schönbeck", "Schönfeld", "Schönhausen",
+ "Schwarz", "Schwinkendorf", "Silz", "Sommersdorf", "Staven",
+ "Stuer", "Torgelow am See", "Varchentin", "Verchen",
+ "Vipperow", "Voigtsdorf", "Vollrathsruhe", "Warrenzin",
+ "Wredenhagen", "Zepkow", "Zetemin", "Zirzow", "Zislow")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
> start <- sample(x=1:10000, size=1)
> start

# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(start)
> Stichprobe_Gesamt <- sort(sample(x = Grundgesamtheit_LETHE_ORTE, size=n))
> Stichprobe_Gesamt

```

Die neun mittels der Zufallsstichprobe ausgewählten Orte können durch den Aufruf des Objektes **Stichprobe_Gesamt** aufgerufen werden.

4.2.2 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« differenziert für die Landkreise Demmin, Mecklenburg-Strelitz und Müritz

Wie bereits erläutert wurde, erfolgte eine Zuweisung der Gemeinden der Grundgesamtheit zu einem der drei Landkreise (Demmin, Mecklenburg-Strelitz, Müritz) vor der Kreisgebietsreform Mecklenburg-Vorpommern des Jahres 2011 (vgl. Tab. 3.1). Dementsprechend lassen sich 25 Gemeinden des Landkreises Mecklenburgische Seenplatte dem (ehemaligen) Landkreis Demmin, 20 Gemeinden des Landkreises Mecklenburgische Seenplatte dem (ehemaligen) Landkreis Mecklenburg-Strelitz sowie 27 Gemeinden des Landkreises Mecklenburgische Seenplatte dem (ehemaligen) Landkreis Müritz zuordnen. Der Stichprobenumfang für jede der zu ziehenden Zufallsstichproben beträgt $n=3$ (jeweils drei Gemeinden pro ehemaligem Landkreis), sodass – wie auch bei der Stichprobenziehung für die Grundgesamtheit (insgesamt) – der Stichprobenumfang insgesamt 9 Gemeinden ($n=9$) umfasst. Für jeden der Landkreise (Demmin, Mecklenburg-Strelitz, Müritz) erfolgt die Ziehung einer Zufallsstichprobe. Im Folgenden werden die syntaktischen Befehle für die Ziehung der drei Zufallsstichproben erläutert.

Im Vorfeld der Ziehung der Zufallsstichproben für die drei Landkreise (Demmin, Mecklenburg-Strelitz, Müritz) werden zunächst die numerischen Grundgesamtheiten (**N_DM**; **N_MST**;

N_MÜR), die 25 (Demmin), 20 (Mecklenburg-Strelitz) und 27 (Müritz) Gemeinden betragen, sowie der Stichprobenumfang (**n_DM**; **n_MST**; **n_MÜR**) definiert. Anschließend erfolgt eine Definition der Grundgesamtheiten (**Grundgesamtheit_LETHE_DM**; **Grundgesamtheit_LETHE_MST**; **Grundgesamtheit_LETHE_MÜR**). Hierbei werden, wie bei der Ziehung einer Stichprobe für die Grundgesamtheit insgesamt, die Gemeinden in alphabetischer Reihenfolge gebracht. Der Befehl **c()** führt zu einer Kombination der Gemeindennamen zu den Objekten **Grundgesamtheit_LETHE_DM_ORTE**, **Grundgesamtheit_LETHE_MST_ORTE** und **Grundgesamtheit_LETHE_MÜR_ORTE**. Für jede der drei Ziehungen wird, um eine Reproduzierbarkeit der Ergebnisse zu ermöglichen, ein Startwert des Algorithmus der Stichprobenziehung ermittelt. Hierbei erfolgt die Ziehung einer Zahl aus dem Vektor 1 bis 10000 (**sample(x=1:10000, size=1)**). Nachdem die Startwerte (**start_DM**; **start_MST**; **start_MÜR**) ermittelt wurden, werden diese in die Funktion **set.seed()** eingefügt. Anschließend erfolgt die Ziehung der Zufallsstichproben mit dem Befehl **sort(sample(x = Grundgesamtheit_LETHE_DM, size=n_DM))**. Die erzeugten Stichproben sind anschließend als Objekte **Stichprobe_DM**, **Stichprobe_MST** und **Stichprobe_MÜR** verfügbar und können mit dem Aufruf des jeweiligen Objektes abgerufen werden. Die syntaktischen Befehle sind in den folgenden Syntaxbeschreibungen **5.1 bis 5.3** dokumentiert.

[5.1]

```
# Stichprobenziehung für den (ehemaligen) Landkreis Demmin (DM)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_DM <- 25
> n_DM <- 3
> Grundgesamtheit_LETHE_DM_ORTE = c("Altenhagen", "Bredenfelde", "Breest", "Briggow",
+ "Duckow", "Gnevkow", "Golchen", "Grammentin",
+ "Grapzow", "Grischow", "Gülzow", "Hohenbollentin",
+ "Kentzlin", "Kittendorf", "Kriesow", "Lindenberg",
+ "Meesiger", "Pripsleben", "Ritzerow", "Röckwitz",
+ "Schönfeld", "Sommersdorf", "Verchen", "Warrenzin",
+ "Zettemin")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
> start_DM <- sample(x=1:10000, size=1)
> start_DM

# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(start_DM)
> Stichprobe_DM <- sort(sample(x = Grundgesamtheit_LETHE_DM_ORTE, size=n_DM))
> Stichprobe_DM
```

[5.2]

```

# Stichprobenziehung für den (ehemaligen) Landkreis Mecklenburg-Strelitz (MST)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_MST <- 20
> n_MST <- 3
> Grundgesamtheit_LETHE_MST_ORTE = c("Beseritz", "Cammin", "Eichhorst", "Genzkow",
+ "Glienze", "Godendorf", "Grünow", "Helpt",
+ "Hohenzieritz", "Kublank", "Mildenitz", "Neddemin",
+ "Neetzka", "Petersdorf", "Priepert", "Schönbeck",
+ "Schönhausen", "Staven", "Voigtsdorf", "Zirzow")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
> start_MST <- sample(x=1:10000, size=1)
> start_MST

# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(start_MST)
> Stichprobe_MST <- sort(sample(x = Grundgesamtheit_LETHE_MST_ORTE, size=n_MST))
> Stichprobe_MST

```

[5.3]

```

# Stichprobenziehung für den (ehemaligen) Landkreis Müritz (MÜR)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_MÜR <- 27
> n_MÜR <- 3
> Grundgesamtheit_LETHE_MÜR_ORTE = c("Altenhof", "Buchholz", "Bütow", "Gotthun",
+ "Grabow-Below", "Groß Kelle", "Kieve", "Klocksinn",
+ "Leizen", "Ludorf", "Massow", "Melz",
+ "Moltzow", "Neu Gaarz", "Penkow", "Priborn",
+ "Schwarz", "Schwinkendorf", "Silz", "Stuer",
+ "Torgelow am See", "Varchentin", "Vipperow",
+ "Vollrathruhe", "Wredenhagen", "Zepkow", "Zislow")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
> start_MÜR <- sample(x=1:10000, size=1)
> start_MÜR

# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(start_MÜR)
> Stichprobe_MÜR <- sort(sample(x = Grundgesamtheit_LETHE_MÜR_ORTE, size=n_MÜR))
> Stichprobe_MÜR

```

4.2.3 Stichprobenziehung für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« auf Grundlage der Gemeindetypisierung

Wie bereits beschrieben wurde (vgl. Kap. 3.3.2), kommt für die Auswahl der Untersuchungsgemeinden für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)« eine Stichprobenziehung auf Grundlage der von Forkel (2014) vorgenommenen Gemeindetypisierung in Betracht. Wie bei der nach den ehemaligen Landkreisstrukturen differenzierten Stichprobenziehung (Demmin, Mecklenburg Strelitz, Müritz) findet sich auch bei der dritten Möglichkeit der Ortsauswahl eine Differenzierung der Gemeinden in drei Kategorien. Entsprechend der Gemeindetypisierung wird zwischen deprivierten, kompensierenden und arrivierten Gemeinden unterschieden.

Im Vorfeld der Programmierung erfolgt zunächst durch den Befehl **rm(list = ls())** die Löschung des gesamten Workspace. Die Kontrolle des leeren Workspace erfolgt mit dem Befehl **ls()**. Im zweiten Schritt erfolgt die Definition der für die Stichprobenziehungen notwendigen Objekte. Die Anzahl der Elemente der drei Kategorien (**N_Kat**) sowie die zu ziehenden Stichprobenumfänge (**n_Kat**) sind für die drei Kategorien identisch. Darüber hinaus werden für die drei Kategorien (depriviert, kompensierend, arriviert) Objekte definiert, die die Namen der Gemeinden der Kategorien enthalten. Hierfür werden die Orte in alphabetischer Reihenfolge sortiert und mittels des Befehls **c()** zu den Objekten **Orte_Kat_neg**, **Orte_Kat_mittel** sowie **Orte_Kat_pos** kombiniert.

Im dritten Schritt der Ziehung der Zufallsstichproben erfolgte die Ermittlung der Startwerte der Algorithmen. Hierbei wurde mittels des Befehls **start <- sample(x=1:6072, size=1)** das Objekt **start** als ein Zufallswert aus der Zahlenfolge 1 bis 6072 (**x=1:6072**) definiert.

Für die Kategorie deprivierte Gemeinden wurde ein Startwert des Algorithmus der Stichprobenziehung von 4839 ermittelt, welcher in die Funktion **set.seed()** eingesetzt wird. Anschließend erfolgt die Ziehung der Zufallsstichprobe für die Kategorie deprivierte Gemeinden mit dem Befehl **sample_Kat_neg <- sort(sample(x = Orte_Kat_neg, size=n_Kat_neg))**, wobei mit der Funktion **sample()** die Ziehung einer Stichprobe aus der Grundgesamtheit (**x**) **Orte_Kat_neg** mit dem Umfang (**size**) von **n_Kat_neg** erfolgt. Für die Kategorie der deprivierten Gemeinden wurden die Orte Beseritz, Voigtsdorf und Zettemin ausgewählt.

[6.1]

```
# Stichprobenziehung für die Kategorie deprivierte Gemeinden (neg)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_Kat_neg <- 24
> n_Kat_neg <- 3
> Orte_Kat_neg <- c("Altenhagen", "Beseritz", "Breest", "Duckow", "Genzkow", "Glienke",
+                 "Golchen", "Grabow-Below", "Grischow", "Groß Kelle",
+                 "Hohenbollentin", "Kentzlin", "Kieve", "Kittendorf", "Kublank", "Massow"
+                 "Melz", "Neetzka", "NeuGaarz", "Pripsleben", "Schönfeld", "Staven"
+                 "Voigtsdorf", "Zettemin")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
# start <- sample(x=1:6072, size=1)
# start
# [1] 4839
```

```
# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(4839)
> sample_Kat_neg <- sort(sample(x = Orte_Kat_neg, size=n_Kat_neg))
> sample_Kat_neg
[1] "Beseritz" "Voigtsdorf" "Zettemin"
```

Für die Kategorie kompensierenden Gemeinden wurde ein Startwert des Algorithmus der Stichprobenziehung von 4024 ermittelt, welcher in die Funktion **set.seed()** eingesetzt wird. Anschließend erfolgt die Ziehung der Zufallsstichprobe für die Kategorie kompensierenden Gemeinden mit dem Befehl **sample_Kat_mittel <- sort(sample(x = Orte_Kat_mittel, size=n_Kat_mittel))**, wobei mit der Funktion **sample()** die Ziehung einer Stichprobe aus der Grundgesamtheit (**x**) **Orte_Kat_mittel** mit dem Umfang (**size**) von **n_Kat_mittel** erfolgt. Für die Kategorie der kompensierenden Gemeinden wurden die Orte Helpt, Verchen und Zirzow ausgewählt.

[6.2]

```
# Stichprobenziehung für die Kategorie kompensierende Gemeinden (mittel)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_Kat_mittel <- 24
> n_Kat_mittel <- 3
> Orte_Kat_mittel <- c("Bredenfelde", "Briggow", "Cammin", "Grapzow", "Gülzow", "Helpt",
+ "Klocksın", "Kriesow", "Lindenberg", "Mildenitz", "Penkow",
+ "Petersdorf", "Priepert", "Ritzerow", "Schönhausen", "Schwinkendorf",
+ "Stuer", "Torgelow am See", "Verchen", "Vollrathsrue", "Warrenzin",
+ "Wredenhagen", "Zepkow", "Zirzow")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
# start <- sample(x=1:6072, size=1)
# start
# [1] 4024
# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(4024)
> sample_Kat_mittel <- sort(sample(x = Orte_Kat_mittel, size=n_Kat_mittel))
> sample_Kat_mittel
[1] "Helpt" "Verchen" "Zirzow"
```

Für die Kategorie arrivierte Gemeinden wurde ein Startwert des Algorithmus der Stichprobenziehung von 438 ermittelt, welcher in die Funktion **set.seed()** eingesetzt wird. Anschließend erfolgt die Ziehung der Zufallsstichprobe für die Kategorie arrivierte Gemeinden mit dem Befehl **sample_Kat_pos <- sort(sample(x = Orte_Kat_pos, size=n_Kat_pos))**, wobei mit der Funktion **sample()** die Ziehung einer Stichprobe aus der Grundgesamtheit (**x**) **Orte_Kat_pos** mit dem Umfang (**size**) von **n_Kat_pos** erfolgt. Für die Kategorie der deprivierten Gemeinden wurden die Orte Eichhorst, Gnevkow und Priborn ausgewählt.

[6.3]

```
# Stichprobenziehung für die Kategorie arrivierte Gemeinden (positiv)
# Schritt1: gesamten Workspace löschen und Kontrolle des leeren Workspace
> rm(list = ls())
> ls()

# Schritt 2: Definition relevanter Objekte
> N_Kat_pos <- 24
> n_Kat_pos <- 3
> Orte_Kat_pos <- c("Altenhof", "Buchholz", "Bütow", "Eichhorst", "Gnevkow", "Godendorf",
+                 "Gotthun", "Grammentin", "Grünow", "Hohenzieritz", "Leizen", "Ludorf",
+                 "Meesiger", "Moltzow", "Neddemin", "Priborn", "Röckwitz", "Schönbeck",
+                 "Schwarz", "Silz", "Sommersdorf", "Varchentin", "Vipperow", "Zislow")

# Schritt 3: Ermittlung des Start Wertes des Algorithmus
# start <- sample(x=1:6072, size=1)
# start
# [1] 438

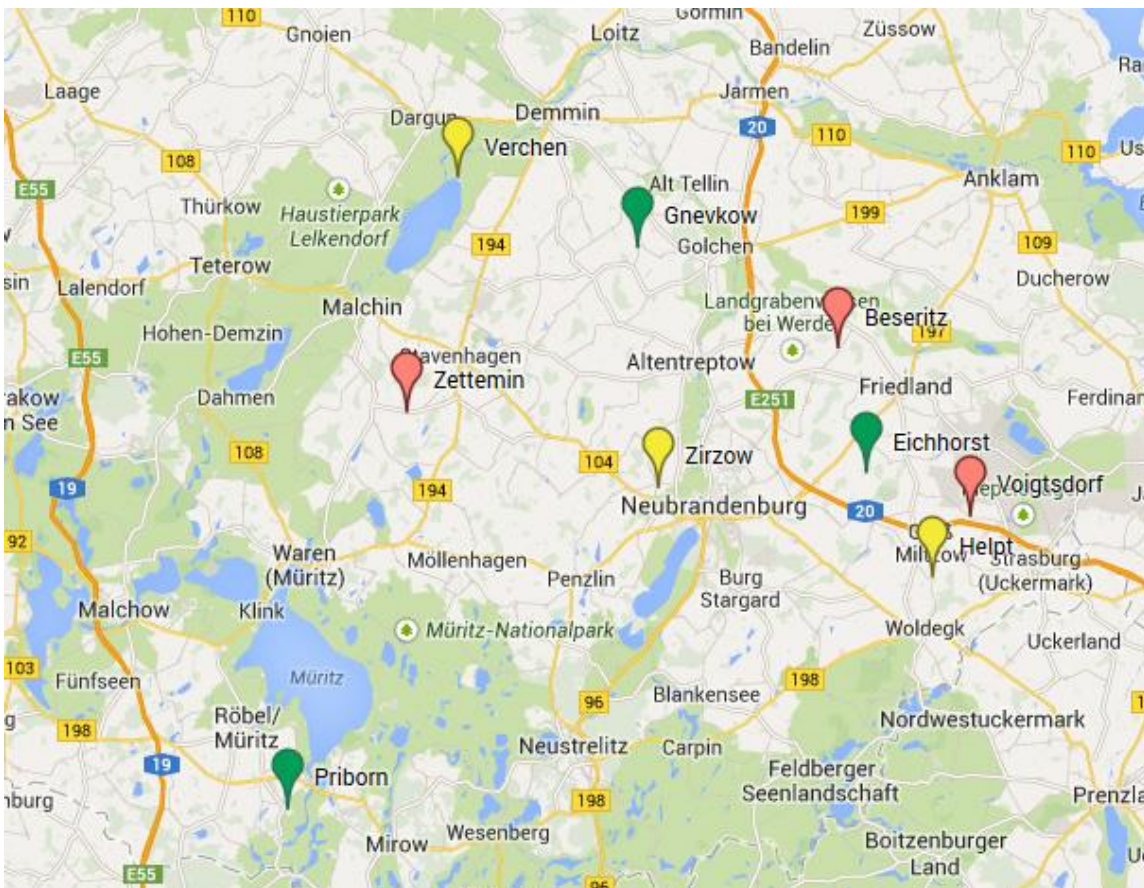
# Schritt 4: Ziehung der Zufallsstichprobe
> set.seed(438)
> sample_Kat_pos <- sort(sample(x = Orte_Kat_pos, size=n_Kat_pos))
> sample_Kat_pos
[1] "Eichhorst" "Gnevkow" "Priborn"
```

5 ZUSAMMENFASSUNG

Im Anschluss an Erläuterungen zu verschiedenen in den Sozialforschungen angewendeten Methoden der Stichprobenziehung erfolgte im vorliegenden Bericht eine Diskussion unterschiedlicher Möglichkeiten der Ortsauswahl für das Projekt »Lebensqualität und Erinnerung in dörflichen Gemeinschaften (LETHE)«. In diesem Zusammenhang erfolgte die Einschätzung der von Forkel (2014) vorgenommenen Typisierung der Orte der Grundgesamtheit. Insgesamt wird die Validität des Verfahrens als gut bewertet. Trotz der konstatierten Heterogenität der im Rahmen der Ortstypisierung berücksichtigten Variablen konnte das primäre Ziel dieser Methodik, eine topische Grundgesamtheit zu differenzieren, erreicht werden. Die Differenzierung der Orte der Grundgesamtheit in deprivierte, kompensierende und arrivierte Gemeinden ermöglicht eine Auswahl von strukturell unterschiedlichen Untersuchungsorten. Als Positiv erscheint hierbei insbesondere die Vielfältigkeit der aufgenommenen Variablen und regionalen Kontextfaktoren. Die Umsetzung der Ziehung der drei Zufallsstichproben (jeweils eine Stichprobenziehung für jede der neu gebildeten Kategorien) erfolgte mit der Statistiksoftware **R**. Um eine Reproduzierbarkeit der Ziehung der Zufallsstichprobe zu gewährleisten, erfolgte im Vorfeld die Festlegung eines kategoriespezifischen Startwertes des Algorithmus der Stichprobenziehung. Als Untersuchungsgemeinden für das Projekt LETHE wurden folgende Orte ausgewählt:

- Kategorie deprivierte Gemeinden: Beseritz, Voigtsdorf, Zetemin
- Kategorie kompensierende Gemeinden: Helpt, Verchen, Zirzow
- Kategorie arrivierte Gemeinden: Eichhorst, Gnevkow, Priborn

Abbildung 5.1: Georeferenzierte Darstellung der Ortsauswahl; *Quelle: google maps (2013)*



LITERATUR

- Atteslander, Peter (2010):** Methoden der empirischen Sozialforschung. 13., neu bearbeitete und erweiterte Auflage. Erich Schmidt Verlag: Berlin.
- Baier, Thomas/ Neuwirth, Erich (2007):** Excel :: COM :: R. *Computational Statistics* 22(1): 91–108
- Böker, Fred/ Sperlich, Stefan/ Zucchini, Walter (2013):** Statistikübungen für Bachelor und Masterstudenten. Ein Arbeitsbuch mit einer Einführung in R. 2. Auflage. Springer-Verlag: Berlin, Heidelberg.
- Bortz, Jürgen/ Schuster, Christof (2010):** Statistik für Human- und Sozialwissenschaftler. 7., vollständig überarbeitete und erweiterte Auflage. Springer-Verlag: Berlin, Heidelberg.
- Diaz-Bone, Rainer (2006):** Statistik für Soziologen. UVK-Verl.-Ges.:Konstanz.
- Diekmann, Andreas (2009):** Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. 20. Auflage. Rowohlt Taschenbuchverlag: Reinbek bei Hamburg.
- Dormann, Carsten F. (2013):** Parametrische Statistik. Verteilungen, maximum likelihood und GLM in R. Springer-Verlag: Berlin, Heidelberg.
- Duller, Christine (2007):** Einführung in die Statistik mit EXCEL und SPSS. Ein anwendungsorientiertes Lehr- und Arbeitsbuch. 2., überarbeitete Auflage. Physica-Verlag: Heidelberg.
- Eckstein, Peter P. (2012):** Statistik für Wirtschaftswissenschaftler. Eine realdatenbasierte Einführung mit SPSS. 3., überarbeitete und erweiterte Auflage. Gabler Verlag: Wiesbaden.
- Fahrmeir, Ludwig/ Künstler, Rita/ Pigeot, Iris/ Tutz, Gerhard (2010):** Statistik: Der Weg zur Datenanalyse. 7., neu bearbeitete Auflage. Springer: Berlin.
- Friedrichs, Jürgen (1990):** Methoden empirischer Sozialforschung. 14. Auflage. Westdeutscher Verlag: Opladen.
- Forkel, Jens A. (2014 - im Ersch.):** Mikrospatiale Typisierungen für die Sozial- und Raumwissenschaften. Neubrandenburg.
- Häder, Michael (2010):** Empirische Sozialforschung. Eine Einführung. 2., überarbeitete Auflage. VS Verlag für Sozialwissenschaften. Wiesbaden.
- Handl, Andreas (o.J.a):** Eine kleine Einführung in R. Universität Bielefeld.
- Handl, Andreas (o.J.b):** Einführung in die Statistik mit R.
- Hornik, Kurt/ Leisch, Friedrich (2002):** Vienna and R: Love, marriage and the future. In: Dutter, Rudolf (Hrsg.): Festschrift 50 Jahre Österreichische Statistische Gesellschaft. Austrian Statistical Society: Wien: S. 61-70.
- Kaya, Maria/ Himme, Alexander (2007):** Möglichkeiten der Stichprobenbildung. In: Albers, Sönke/ Klapper, Daniel/ Konradt, Udo/ Walter, Achim/ Wolf, Joachim (Hrsg.): Methodik der empirischen Forschung. 2., überarbeitete und erweiterte Auflage. Gabler: Wiesbaden. S. 79-88.

- Kauermann, Göran/ Küchenhoff, Helmut (2011):** Stichproben. Methoden und praktische Umsetzung mit R. Springer-Verlag: Berlin, Heidelberg.
- Kohn, Wolfgang/ Öztürk, Riza (2013):** Statistik für Ökonomen. Datenanalyse mit R & SPSS. 2., überarbeitete Auflage. Springer-Verlag: Berlin Heidelberg.
- Ligges, Uwe (2009):** Programmieren mit R. 3., überarbeitete und erweiterte Auflage. Springer-Verlag: Berlin, Heidelberg.
- Manderscheid, Katharina (2012):** Sozialwissenschaftliche Datenanalyse mit R. Eine Einführung. VS Verlag für Sozialwissenschaften: Wiesbaden.
- Mossig, Ivo (2012):** Stichproben-Stichprobenauswahlverfahren Stichprobenumfang. Beiträge zur Wirtschaftsgeographie und Regionalentwicklung. Nr. 1-2012. herausgegeben von Prof. Dr. Ivo Mossig. Institut für Geographie. Universität Bremen.
- Plank, Andreas (2010):** Grafiken und Statistik in R. Skript zum Umgang und zur multivariaten Datenanalyse.
- Wollschläger, Daniel (2012):** Grundlagen der Datenanalyse mit R. Eine anwendungsorientierte Einführung. 2. Auflage. Springer-Verlag: Berlin, Heidelberg.