



Master Thesis

im Rahmen des
Universitätslehrganges „Geographical Information Science & Systems“
(UNIGIS MSc) am Interfakultären Fachbereich für GeoInformatik (Z_GIS)
der Paris Lodron-Universität Salzburg

zum Thema

„Räumliche Einflussfaktoren auf die Pendlerdynamik

Vergleich von Pendlerdaten in Niedersachsen
und Mecklenburg-Vorpommern“

vorgelegt von

B.Sc. Florian Meier
U105599, UNIGIS MSc Jahrgang 2020

Betreuer/in:
Dr. Christian Neuwirth

Zur Erlangung des Grades
„Master of Science – MSc“

Leipzig, 13.06.2023

Zusammenfassung

Die Pendeldynamik ist nicht nur ein bedeutendes Thema der Geographie und verwandter wissenschaftlicher Disziplinen, sondern oftmals auch ein Ärgernis für die Betroffenen, welches negative gesundheitliche Folgen mit sich bringt. Während sich viele Forschungen auf die Pendelströme in Großstädte hinein konzentrieren, wird die Pendeldynamik kleinere Gemeinden, welche ebenfalls Arbeitsplätze aufweisen, nur selten beleuchtet. Diese Arbeit untersucht daher exemplarisch das Pendelverhalten der Gemeinden mit weniger als 5000 Einwohnenden in den deutschen Bundesländern Mecklenburg-Vorpommern und Niedersachsen. Dabei wird der Fokus auf den Einfluss räumlicher Faktoren der Gemeinden auf die Pendeldynamik gelegt.

Als Datengrundlage werden die jährlichen Pendeldaten des Statistik-Service der Bundesagentur für Arbeit herangezogen. Als Pendelnde werden dabei sozialversicherungspflichtig Beschäftigte definiert, deren Wohnort sich von ihrem Arbeitsort auf Gemeindeebene unterscheidet, unabhängig von der tatsächlichen Anzahl an Pendelbewegungen. Aus diesen Daten wurden Pendelquoten und Pendeldistanzen berechnet. Als mögliche räumliche Einflussfaktoren werden die Flächengröße, die Bevölkerungsdichte, der Küstenanteil, die Entfernung zum nächsten Oberzentrum und der Shape-Index der Gemeinden herangezogen. Für die räumlichen Faktoren und die Pendelparameter werden deskriptive statistische Maße ermittelt und sie wurden zueinander in Korrelation gesetzt (Pearson Korrelation).

Als Ergebnis erhält man zu großem Teil nur geringe Korrelationen. Insbesondere Inseln und Küstengemeinden unterscheiden sich jedoch in ihren Pendeldaten deutlich von Binnengemeinden. Bei der Flächengröße, der Entfernung zum nächsten Oberzentrum und dem Shape-Index sind zumindest zu den Pendelquoten wahrnehmbare Korrelationen vorhanden. Zukünftige Untersuchungen könnten auch nicht-räumliche Einflussfaktoren mit einbeziehen und als Datengrundlage Mobilitätsindikatoren auf Grundlage von Mobilfunkdaten verwenden.

Abstract

Commuting to work is often connected to troubles and negative health consequences. But the topic itself is a significant part of research in geography and other related sciences. Many of them concentrate on the commuting movements into big cities. However, there are only a few how attend to smaller municipalities which also have workspaces. This thesis exemplary investigates the topic “commuting” in municipalities with a population less then 5.000 in the german states Mecklenburg-Western Pomerania and Lower Saxony. It focuses on the influence of spatial factors.

Data basis of this thesis are the yearly appearing statistics from Bundesagentur für Arbeit. To be defined as commuter in this statistic a person must have the duty of social insurance and residence and workplace must be two different municipalities. The real count of movements is unimportant. Commuting rates and distances will be calculated from this data. Area, population density, proportion of coast, distance to the next big city and shape index will be used as possible influencing spatial factors. For these and the commuting parameters descriptive statistical measures and correlations (Pearson correlation coefficient) will be calculated.

The bulk of the correlations are low. However, there are big differences in the commuting data between island or coastal municipalities and inland municipalities. For area, distance to the next big city and shape index at least the correlations to the commuting rates are perceptible. Future investigations could use non-spatial influencing factors as well as real mobility indicators based on cellular data.

Inhaltsverzeichnis

Zusammenfassung	I
Abstract	II
Abbildungsverzeichnis	IV
Tabellenverzeichnis	V
Abkürzungsverzeichnis	VI
1. Einleitung	1
2. Daten und Methodik	5
2.1 Untersuchungsgebiet.....	5
2.1.1 Mecklenburg-Vorpommern	6
2.1.2 Niedersachsen	7
2.2 Software.....	8
2.2.1 Tabellenkalkulation: Microsoft Excel.....	8
2.2.2 Datenverwaltung: PostgreSQL mit PostGIS	8
2.2.3 Geographisches Informationssystem: QGIS	8
2.3 Daten.....	9
2.3.1 Verwaltungsgrenzen	9
2.3.2 Bevölkerungszahlen	11
2.3.3 Pendelraten.....	13
2.3.4 Datenbank-Import	21
2.4 Methodik und Vorgehen	22
2.4.1 Verknüpfung der Daten und Bevölkerungszahlen-Filter	23
2.4.2 Bestimmung der räumlichen Parameter	24
2.4.3 Bestimmung der Pendelparameter	30
2.4.4 Bestimmung der statistischen Maße und Korrelationen.....	35
2.4.5 Anmerkungen zum Vorgehen in Niedersachsen.....	37
3. Ergebnisse	42
3.1 Deskriptive statistische Maße der Indikatoren.....	42
3.2 Deskriptive statistische Maße der Pendelparameter	44
3.3 Korrelation der Pendelparameter untereinander	49
3.4 Korrelation Indikatoren und Pendelparameter	53
3.5 Konkrete Gemeinden im Vergleich	59
3.5.1 Vergleich einer Binnen- und einer Inselgemeinde	59
3.5.2 Vergleich einer nah und einer entfernt zum Oberzentrum liegenden Gemeinde	61
3.5.3 Vergleich einer Gemeinde in Niedersachsen mit einer in Mecklenburg-Vorpommern	63
4. Diskussion und Fazit	65
Literaturverzeichnis	70

Abbildungsverzeichnis

1.1	<i>Verdeutlichung des möglichen Einflusses der Form eines Gebiets auf die Pendelraten</i>	4
2.1	<i>Die Lage der untersuchten Bundesländer und ihrer Landeshauptstädte</i>	5
2.2	<i>Überblick über die Verwaltungsstruktur und die zentralen Orte von M.-V.</i>	6
2.3	<i>Überblick über die Verwaltungsstruktur und die zentralen Orte von Niedersachsen</i>	7
2.4	<i>Ausschnitt des Datensatzes „Gemeinden“ aus dem WFS-Dienst „wfs_vg250“</i>	10
2.5	<i>Auszug aus dem Gemeindeverzeichnis des Statistischen Bundesamtes</i>	11
2.6	<i>Ergebnis der Aufbereitung des Gemeindeverzeichnisses für M.-V.</i>	12
2.7	<i>Auszug des Tabellenblatts „Gemeindedaten“ mit Pendelwerten aus M.-V.</i>	14
2.8	<i>Ergebnis der Aufbereitung der Pendelwerte für Mecklenburg-Vorpommern</i>	15
2.9	<i>Auszug des Tabellenblatts „Auspendler Gemeinden“ der Pendelströme von M.-V.</i>	16
2.10	<i>Auszug des Tabellenblatts „Einpendler Gemeinden“ der Pendelströme von M.-V.</i>	17
2.11	<i>Anonymisierungen in den Pendelströmen von Mecklenburg-Vorpommern</i>	17
2.12	<i>Die Auspendelströme vor und nach Ausführung des Makros „matrix_aunder“</i>	19
2.13	<i>Überblick über die Gemeinden M.-V., welche nicht untersucht werden</i>	24
2.14	<i>Kennzeichnung der für die Berechnung des Küstenanteils zu digitalisierenden Flächen</i>	26
2.15	<i>Die Küstengrenzen der zu untersuchenden Gemeinden</i>	27
2.16	<i>Lage der möglichen Zielorte zur Entfernungsmessung für die Gemeinden M.-V.</i>	28
2.17	<i>Überblick über die Gemeinden Niedersachsen, welche nicht untersucht werden</i>	39
2.18	<i>Lage der möglichen Zielorte zur Entfernungsmessung für die Gemeinden Nds.</i>	41
3.1	<i>Histogramme des Indikators „Küstenanteil“ unter Ausschluss der Binnengemeinden</i>	42
3.2	<i>Histogramme des Indikators „Entfernung Oberzentrum“</i>	43
3.3	<i>Histogramme des Pendelparameters „Einpendelquote“</i>	45
3.4	<i>Histogramme des Pendelparameters „Arbeitsplatzausstattung“</i>	46
3.5	<i>Histogramm des Pendelparameters „Auspendeldistanz“ in Niedersachsen</i>	47
3.6	<i>Ausschnitt des Streudiagramms für die Auspendelquote und -distanz in Niedersachsen</i>	50
3.7	<i>Ausschnitt des Streudiagramms für die Ein- und Auspendelquote in Niedersachsen</i>	51
3.8	<i>Streudiagramm für die Arbeitsplatzausstattung und den Independence-Index in M.-V. für die Gemeinden mit einer Arbeitsplatzausstattung von maximal 100%</i>	53
3.9	<i>Streudiagramm für die Fläche & den Idpd.-Index der Gemeinden Nds. ohne Ausreißer</i>	56
3.10	<i>Streudiagramm für den Küstenanteil & den Independence-Index der Gemeinden Nds.</i>	57
3.11	<i>Lage der Gemeinden Bothel und Spiekeroog in Niedersachsen</i>	59
3.12	<i>Die auf Kreis-, Regierungsbezirk und Länderebene aggregierten Auspendelströme der Gemeinden Bothel und Spiekeroog in Niedersachsen</i>	60
3.13	<i>Lage der Gemeinden Bütow und Wackerow in Mecklenburg-Vorpommern</i>	61
3.14	<i>Die Einpendelströme der Gemeinden Bütow und Wackerow in M.-V.</i>	62
3.15	<i>Lage der Gemeinden Lahn in Nds. und Gotthun in M.-V.</i>	63
3.16	<i>Die Auspendelströme der Gemeinden Lahn in Nds. und Gotthun in M.-V.</i>	64

Tabellenverzeichnis

2.1	<i>Überblick der Methodik</i>	22
2.2	<i>Die möglichen Zielorte zur Entfernungsmessung für die Gemeinden M.-V.</i>	28
2.3	<i>Die möglichen Zielorte zur Entfernungsmessung für die Gemeinden Nds.</i>	40
3.1	<i>Deskriptive statistische Maße der Indikatoren in M.-V. und Nds.</i>	44
3.2	<i>Anzahl der Gemeinden mit fehlenden Pendelwerte bzw. Pendelparametern</i>	45
3.3	<i>Deskriptive statistische Maße der Pendelparameter in M.-V. und Nds.</i>	47
3.4	<i>Deskriptive statistische Maße der Pendelparameter in Mecklenburg-Vorpommern und Niedersachsen unter Ausschluss der Gemeinden, die im jeweiligen Parameter zu den zwei Prozent mit den niedrigsten bzw. höchsten Werten gehören</i>	49
3.5	<i>Korrelationen ausgewählter Pendelparameter untereinander</i>	49
3.6	<i>Veranschaulichung der Abhängigkeit des max. Idpd.-Index von der ApAusstattung</i>	52
3.7	<i>Korrelation zwischen Arbeitsplatzausstattung und Idpd.-Index in den Gemeinden mit einer Arbeitsplatzausstattung bis einschließlich bzw. mindestens 100%</i>	52
3.8	<i>Korrelation zwischen den räumlichen Indikatoren und den Pendelparametern in den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens</i>	53
3.9	<i>Korrelation zwischen den räumlichen Indikatoren und den Pendelparametern in den Gemeinden M.-V. und Nds. unter Ausschluss der Gemeinden, die im jeweiligen Parameter zu den zwei Prozent mit den niedrigsten bzw. höchsten Werten gehören</i>	54
3.10	<i>Vergleich der Korrelation zwischen den räuml. Indikatoren und Pendelparametern in den Gemeinden M.-V. und Nds. vor und nach Ausschluss der Ausreißer</i>	55
3.11	<i>Werte des Korrelationskoeffizienten der Korrelationen mit unerwarteter Richtung nur für die Binnengemeinden in Mecklenburg-Vorpommern</i>	58
3.12	<i>Die Pendelparameter von Lahn und Gotthun im Vergleich</i>	63

Abkürzungsverzeichnis

ags	<i>Amtlicher Gemeindeschlüssel (Attributname)</i>
ApAusstattung	<i>Arbeitsplatzausstattung</i>
ars	<i>Amtlicher Regionalschlüssel (Attributname)</i>
BKG	<i>Bundesamt für Kartographie und Geodäsie (Deutschland)</i>
bzw.	<i>beziehungsweise</i>
ca.	<i>circa; etwa</i>
CO ₂	<i>Kohlenstoffdioxid</i>
CSV	<i>Comma-separated values (Dateiformat)</i>
DB	<i>Datenbank</i>
et al.	<i>et alia; und andere</i>
gen	<i>Geographischer Name (Attributname)</i>
geom	<i>Geometrie (Attributname)</i>
gf	<i>Geofaktor (Attributname)</i>
GIS	<i>Geographisches Informationssystem</i>
IAB	<i>Institut für Arbeitsmarkt- und Berufsforschung (Deutschland)</i>
Idpd.-Index	<i>Independence-Index</i>
ISOE	<i>Institut für sozial-ökologische Forschung (Deutschland)</i>
LTR	<i>Long-Term-Release; stabile Langzeit-Versionen von QGIS</i>
M.-V.	<i>Mecklenburg-Vorpommern</i>
MAUP	<i>Modifiable Areal Unit Problem; Problem der veränderbaren Gebietseinheit</i>
max.	<i>maximal</i>
MiD	<i>Mobilität in Deutschland (Forschungsprojekt)</i>
mind.	<i>mindestens</i>
Nds.	<i>Niedersachsen</i>
NRW	<i>Nordrhein-Westfalen</i>
NUTS	<i>Nomenclature des unités territoriales statistiques; räumliche Bezugseinheiten der amtlichen Statistik in der EU</i>
o.Ä.	<i>oder Ähnliche</i>
o.J.	<i>ohne Jahresangabe</i>
ÖPNV	<i>Öffentlicher Personennahverkehr</i>
räuml.	<i>räumlich</i>
sn_l	<i>Schlüsselnummer des Landes (Attributname)</i>
SQL	<i>Structured Query Language; Datenbanksprache</i>
stat.	<i>statistisch</i>
temp	<i>temporär</i>
u.A.	<i>und Andere</i>
UTF8	<i>8-Bit UCS (Universal Coded Character Set) Transformation Format; Zeichencodierung</i>
VBA	<i>Visual Basic for Applications; Programmiersprache in Microsoft Office</i>
vgl.	<i>vergleiche</i>
WFS	<i>Web Feature Service</i>
.xlsb	<i>Excel Binary Spreadsheet (Dateiformat)</i>
.xlsx	<i>Excel Spreadsheet XML (Extensible Markup Language) (Dateiformat)</i>
z.B.	<i>zum Beispiel</i>
z.T.	<i>zum Teil</i>

1. Einleitung

Für viele Menschen folgt jeder Arbeitstag einem identischen Muster. Sie bewegen sich vor Beginn der Arbeitszeit von ihrem Wohnort zu ihrem Arbeitsplatz und nach Beenden der Tätigkeit wieder zurück. Die Strecke zwischen diesen beiden Orten wird zweimal täglich in Anspruch genommen und ist unterschiedlich lang – zwischen wenigen Metern zwischen Wohn- und Arbeitszimmer für diejenigen, welche im Home-Office arbeiten, bis hin zu vielen Kilometern. Für ca. 1,3 Millionen Einwohnende Deutschlands liegen 2019 mehr als 150 Kilometer zwischen Wohn- und Arbeitssitz (Bundesministerium des Innern und für Heimat 2020), bei solchen Entfernungen wird auch auf andere Modelle wie einen Zweitwohnsitz in der Nähe des Arbeitsstandortes und das Fahren zum Hauptwohnsitz für das Wochenende zurückgegriffen. Je nach verwendeter Statistik bzw. Befragung beträgt der durchschnittlich zurückgelegte Arbeitsweg in Deutschland ca. 16,4 Kilometer laut Mikrozensus 2020 (Statistisches Bundesamt 2022 a) oder ca. 16,9 Kilometer laut Statistik der Bundesagentur für Arbeit im Jahr 2019, jeweils Tendenz steigend. In dünner besiedelten Gegenden beträgt letztgenannter Wert sogar über 30 Kilometer (Bundesministerium des Innern und für Heimat 2020). Das Hin- und Herbewegen zwischen Wohn- und Arbeitsort erinnert an das Pendel einer mechanischen Uhr und wird deshalb als Pendeln bezeichnet.

Die Pendeldynamik ist schon lange ein bedeutendes Thema der Geographie sowie verwandter wissenschaftlicher Disziplinen. So wurde das tägliche Pendeln bereits von Hartke (1938) als umfassender geographischer Problemkomplex dargestellt. Dieser kann „nur im größeren Problemzusammenhang diskutiert [...] [und] nur im Zusammenhang mit allen anderen Erscheinungen der sozialräumlichen Mobilität untersucht werden“ (Klingbeil 1969). Es gibt dabei unterschiedliche Definitionen des Pendelns. Die in Deutschland gebräuchlichste statistische Verwendung des Begriffs besagt, dass sich Wohn- und Arbeitsort eines sozialversicherungspflichtig Beschäftigten nicht in der gleichen Gemeinde befinden und stammt aus den jährlichen Erhebungen der Bundesagentur für Arbeit (vgl. Bundesagentur für Arbeit 2020). Da in der vorliegenden Master-Thesis mit diesen Daten gearbeitet wird, übernehme ich die beschriebene Definition trotz der Kritik daran, welche z.B. Nitschke et al. (2022) zusammenfassen und auf welche auch im Fazit dieser Arbeit eingegangen wird. Diese Definition des Pendelns und die Daten werden für die Forschung in Deutschland häufig verwendet. So hat Siedentop (2007) anhand der Daten für fünf westdeutsche Agglomerationsräume festgestellt, dass die Suburbanisierung der Arbeitsplätze nicht zwangsläufig zu weniger Pendelverkehr führt. Haas und Hamann (2008) führten einen Vergleich der Daten zwischen ost- und westdeutschen Bundesländern durch und schlussfolgerten u.A., dass sich die Pendelnden im Westen räumlich deutlich stärker verteilen. Unter Einbeziehung weiterer Daten wie z.B. der Entgelte von Ein-, Aus- und Binnenpendelnden eines Gebietes lassen sich z.B. weiterführende Verflechtungsverhältnisse im Raum Berlin aufdecken (vgl. Janke 2017) und die Entwicklung und Bedeutung des Grenzpendelns dokumentieren (vgl. Buch et al. 2020).

Je nach gewähltem Verkehrsmittel werden mit dem Pendeln unterschiedliche, jedoch meist negative Bilder assoziiert, wie von LKWs verstopfte Autobahnen und eine Straßeninfrastruktur von Städten, welche täglich an ihre Kapazitätsgrenzen gebracht wird, oder auch überfüllte, verspätete oder ausfallende Bahnen oder Busse. Die möglichen negativen gesundheitlichen Folgen des Pendelns waren Bestandteil vieler Studien, so empfinden beispielsweise Autofahrende und Nutzende des öffentlichen Nahverkehrs ihren Arbeitsweg oftmals gestresster (aufgrund der erwähnten Probleme) als Radfahrende und zu Fuß gehende, welche in der Regel kürzere Strecken zurückzulegen haben (Gatersleben und Uzzell 2007). Einem Report der Techniker Krankenkasse (Hamberger 2021) zufolge sind Pendelnde zwar allgemein nicht häufiger krankgeschrieben als Nicht-Pendelnde, haben jedoch prozentual gesehen mehr Fehltag aufgrund von Depressionen und anderen psychischen Leiden. Weitere in dem Report benannte Folgen sind u.A. Rücken- und Kopfschmerzen, Schlafstörungen, ein erhöhtes Herzinfarkt-Risiko und eine mit der Länge des Arbeitsweges steigende Unzufriedenheit der Work-Life-Balance.

Trotzdem pendelten im Jahr 2019 ca. 59,5% der sozialversicherungspflichtig Beschäftigten Deutschlands, auch dieser Wert ist im Vergleich zu früheren Auswertungen angestiegen (Bundesministerium des Innern und für Heimat 2020). Die Gründe, warum gependelt wird, klingen oft einleuchtend: Partnerschaften mit zwei berufstätigen Personen, deren Arbeitsorte in unterschiedlichen Gemeinden liegen, Umzug der Firma oder die spätere Annahme eines besseren bzw. besser bezahlten, aber weiter entfernten Jobs ohne Umzug, weil man z.B. bereits ein eigenes Haus besitzt und das soziale Umfeld nicht vollständig wechseln möchte. Viele suchen das Leben auf dem Land, möchten aber nicht auf die oft besser bezahlten Jobs in der Großstadt verzichten. Warum das Ganze jedoch eine Illusion sei und Pendelnde sich damit oft selbst betrügen, beschreiben Fischer und Book (2021).

Neben den persönlichen Auswirkungen auf die Gesundheit der Pendelnden müssen auch die Belastungen der Umwelt betrachtet werden. In Deutschland machen dienstliche Aktivitäten (Dienstreisen zählen auch dazu) knapp die Hälfte der Fahrleistung des motorisierten Personenverkehrs aus (Nobis und Kuhnimhof 2018). Einer Studie von Greenpeace (Büttner und Breitzkreuz 2020) zufolge lassen sich 18% der Emissionen des Pendelverkehrs einsparen, wenn 40% der Arbeitnehmenden pro Woche an zwei Tagen im Home-Office arbeiten. Jedoch fordert das Klimaziel 2030 eine Verringerung um 40% im Vergleich zu 1990 (Die Bundesregierung 2019). Zum Erreichen dieses Ziels werden Maßnahmen ergriffen, um den Verkehr im Allgemeinen zu verringern bzw. emissionsfreier zu gestalten, Stichpunkte dabei lauten: Förderung der Elektromobilität, CO₂-Steuer und Attraktivitätssteigerung des ÖPNV (vgl. ebd.). Doch um den zukünftigen Ausbau des Nahverkehrsnetzes sinnvoll zu gestalten, ist es zunächst erforderlich, den Ist-Zustand der Pendeldynamik zu kennen und zu verstehen, was aufgrund der Dynamik des Mobilitätsverhaltens und der Anforderungen daran ein immer wiederkehrender Prozess sein muss.

Bei der Betrachtung von Pendelströmen, also von welchen Wohn- zu welchen Arbeitsorten gependelt wird, liegt der Fokus in der Presse häufig auf diejenigen, die aus dem Umland von Großstädten oder Metropolregionen in diese hereinführen (vgl. z.B. Leipziger Zeitung 2019, Märkische Allgemeine 2021, Antenne Münster 2021). Auch in der Forschung wird das Pendeln in Städte und Metropolregionen hinein bzw. innerhalb von diesen breit beleuchtet, so haben sich z.B. Schwanen et al. (2001) mit den Unterschieden der Pendelbewegungen in den Kernstädten polyzentrischer sowie monozentrischer Siedlungsstrukturen in den Niederlanden befasst und dabei u.A. herausgefunden, dass in den Kernstädten polyzentrischer Siedlungsstrukturen häufiger der öffentliche Nahverkehr zum Pendeln genutzt wird und mit diesem auch weitere Entfernungen zurückgelegt werden als in den Kernstädten monozentrischer Strukturen. Le Néchet (2012) verglich speziell die Nutzung des öffentlichen Personennahverkehrs als Fortbewegungsmittel zum Pendeln im Großraum Paris und in der deutschen Metropolregion Rhein-Ruhr, in welcher dieser (anders als in Paris) hauptsächlich nur auf lokaler, nicht jedoch auf metropolischer Ebene genutzt wird. Weitere Forschungen in dieser Richtung, die an dieser Stelle nur aufgezählt werden sollen, gibt es z.B. für die Metropolregionen der USA (vgl. Lang und Nelson 2007), die japanische Hafenstadt Osaka (vgl. Sakanishi 2006) und die thailändische Hauptstadt Bangkok (vgl. Kaewwongwattana et al. 2015).

Jedoch bilden diese Ströme zwar einen großen Anteil, jedoch nicht die Gesamtheit des Pendelverkehrs ab. Auch außerhalb der Städte und Metropolregionen, in kleineren bzw. dünner besiedelten Gemeinden, gibt es Arbeitsplätze. Ob diese hauptsächlich von Arbeitenden mit Wohnort in der gleichen Gemeinde (Nicht- oder Binnenpendelnde) besetzt werden oder auch zu diesen im großen Umfang gependelt wird (und falls ja, wie weit?) soll im Rahmen dieser Master-Thesis untersucht werden. Dabei soll ebenfalls erforscht werden, inwiefern bestimmte räumliche Faktoren der Gemeinden einen Einfluss auf die Pendelraten haben. Europaweit und auf NUTS-Ebene 2 betrachtet, liefert eine Statistik von Eurostat folgende Erkenntnisse (vgl. Eurostat 2018):

- je größer die Fläche einer Region, desto niedriger der Pendelanteil
- je größer die Bevölkerungsdichte einer Region, desto größer der Pendelanteil
- die Topographie hat Einfluss auf den Pendelanteil, so haben z.B. Inseln und Regionen mit langem Küstenstreifen einen geringeren Pendelanteil
- Regionen in der Nähe eines großen Ballungsraums weisen einen höheren Pendelanteil auf

Mit Hilfe der Pendlerdaten der Bundesagentur für Arbeit vom Jahr 2021 soll überprüft werden, ob es auch bei den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens einen signifikanten Zusammenhang zwischen räumlichen Faktoren und der Pendelquote sowie -distanz gibt. Dabei werden alle Verwaltungseinheiten auf Gemeindeebene

untersucht, in welchen es weniger als 5000 Einwohnende gibt. Diese Zahl stellt die untere Grenze der Klassifikation einer Kleinstadt nach deutscher Reichsstatistik von 1871 dar, 674 der 726 Verwaltungseinheiten auf Gemeindeebene in Mecklenburg-Vorpommern sowie 603 der 944 Verwaltungseinheiten auf Gemeindeebene in Niedersachsen fallen darunter. Die beiden Länder eignen sich als Untersuchungsgebiet u.A. aufgrund ihrer Küstenlage und der Möglichkeit des Vergleichs eines „neuen“ und eines „alten“ Bundeslands.

Als räumliche Prädiktoren werden die vier in der Eurostat-Studie behandelten Faktoren Fläche, Bevölkerungsdichte, Topographie (hier dargestellt durch den Küstenanteil) und Entfernung zum nächsten Ballungsraum (hier Entfernung zum nächsten Oberzentrum) sowie zusätzlich der Shape-Index genutzt, welcher die Abweichung der Form einer Fläche von einem Kreis darstellt. Dessen möglicher Einfluss auf die Pendeldynamik wird in Abbildung 1.1 dargestellt.



Abbildung 1.1: Verdeutlichung des möglichen Einflusses der Form eines Gebiets auf die Pendeldaten. (Eigener Screenshot in QGIS)

Als Pendelparameter werden neben Ein- und Auspendelquote sowie Ein- und Auspendeldistanz noch zwei weitere Werte herangenommen, welche sich aus den Pendeldaten berechnen lassen. Die Arbeitsplatzausstattung ist kein Pendelparameter an sich, dient jedoch in der Auswertung der besseren Einordnung der Pendelquoten. Der Independence-Index beschreibt das Verhältnis der Anzahl von Ein-, Aus und Binnenpendelnden einer Gemeinde.

Zusammengefasst wird in dieser Master-Thesis als Leitziel untersucht, ob es in den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens mit weniger als 5000 Einwohnenden eine Abhängigkeit der Pendelparameter von räumlichen Faktoren gibt.

Dazu erfolgt eine Aufbereitung der Pendeldaten des Statistik-Service der Bundesagentur für Arbeit und eine Verknüpfung dieser mit dem Gemeindedatensatz des Bundesamtes für Kartographie und Geodäsie. Räumlicher Faktoren sowie Pendelparameter der Gemeinden werden berechnet und deskriptiv statistisch ausgewertet. Mittels einer Korrelationsanalyse soll beides in Zusammenhang gesetzt und ausgewertet werden.

2. Daten und Methodik

2.1 Untersuchungsgebiet

Als Untersuchungsgebiet für die Master-Thesis werden die deutschen Bundesländer Mecklenburg-Vorpommern und Niedersachsen herangezogen. Die beiden Länder bieten sich aufgrund ihrer Küstenlage an, da u.A. der Einfluss der topographischen Lage der Gemeinden (hier dargestellt durch den Küstenanteil) auf die Pendeldaten untersucht werden soll. Auch die geringe Anzahl größerer Städte innerhalb bzw. in unmittelbarer Umgebung des Bundeslandes (insbesondere bei Mecklenburg-Vorpommern) ist vorteilhaft, da so auch unterschiedliche Distanzen zu Oberzentren untersucht werden können, was als ein weiterer Einflussfaktor genutzt wird. Die beiden Bundesländer bieten sich auch im Vergleich zueinander an, da Mecklenburg-Vorpommern ehemals zur Deutschen Demokratischen Republik gehörte und so Unterschiede des Pendelverhaltens sowie des Einflusses auf dieses zwischen einem „alten“ und einem „neuen“ Bundesland untersucht werden können.



Abbildung 2.1: Die Lage der untersuchten Bundesländer und ihrer Landeshauptstädte.
(Eigene Darstellung erstellt mit QGIS; © GeoBasis-DE / BKG 2022; © Natural Earth 2022)

2.1.1 Mecklenburg-Vorpommern

In seiner heutigen Form existiert das Land Mecklenburg-Vorpommern seit der deutschen Wiedervereinigung 1990 und liegt im Nordosten der Bundesrepublik Deutschland. Es grenzt an die Länder Brandenburg, Niedersachsen und Schleswig-Holstein, an die Ostsee sowie an Westpommern (Polen). Mecklenburg-Vorpommern ist mit einer Fläche von 23.213 km² und einer Bevölkerungszahl von ca. 1,6 Millionen das am dünnsten besiedelte der 16 deutschen Bundesländer, die Bevölkerungsdichte beträgt 69 Einwohnende je km² (Staatskanzlei des Landes Mecklenburg-Vorpommern o.J.). Das Land ist in sechs Landkreise (fünf davon sind die flächengrößten Deutschlands) und die beiden kreisfreien Städte Schwerin (Landeshauptstadt) und Rostock gegliedert. Neben diesen beiden Städten gibt es zwei weitere Oberzentren, Neubrandenburg sowie Stralsund und Greifswald gemeinsam. Außerdem gibt es 18 Mittelzentren (Ministerium für Energie, Infrastruktur und Landesentwicklung Mecklenburg-Vorpommern 2016).

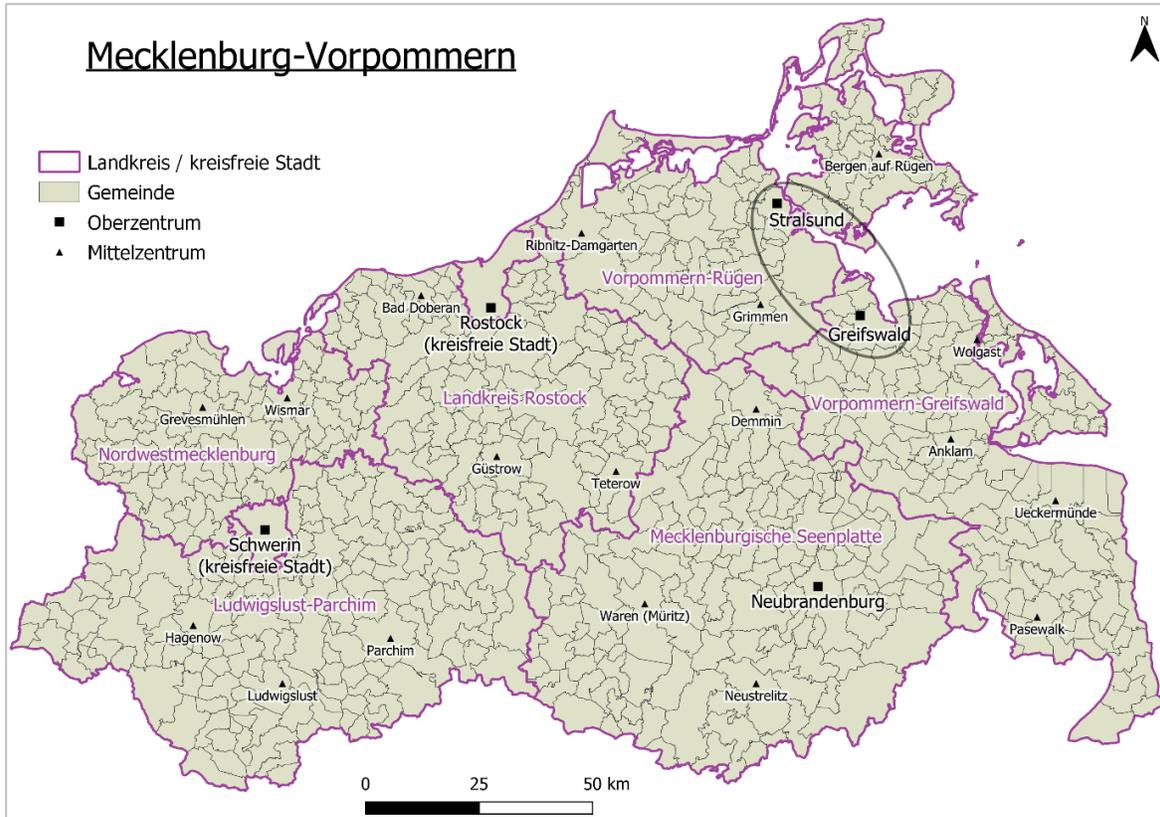


Abbildung 2.2: Überblick über die Verwaltungsstruktur und die zentralen Orte von Mecklenburg-Vorpommern. (Eigene Darstellung erstellt mit QGIS; © GeoBasis-DE / BKG 2022, © Ministerium für Energie, Infrastruktur und Landesentwicklung Mecklenburg-Vorpommern 2016)

2.1.2 Niedersachsen

Das Land Niedersachsen existiert in seiner heutigen Form seit dem Ende des 2. Weltkriegs und liegt im Nordwesten der Bundesrepublik Deutschland. Es grenzt an neun andere Bundesländer und umschließt darunter vollständig die Freie Hansestadt Bremen, die anderen Anrainer sind die Länder Schleswig-Holstein, Mecklenburg-Vorpommern, Brandenburg, Sachsen-Anhalt, Hessen und Nordrhein-Westfalen, der Freistaat Thüringen und die Freie und Hansestadt Hamburg. Auch grenzt Niedersachsen an die Nordsee und an die niederländischen Provinzen Overijssel, Drenthe und Groningen. Das Bundesland ist mit einer Fläche von 47.709 km² das zweitgrößte Bundesland und hat mit etwa 8 Millionen Einwohnenden die vierthöchste Bevölkerungszahl aller Bundesländer. Die Bevölkerungsdichte beträgt 169 Einwohnende je km², womit Niedersachsen hierbei die geringste aller westdeutschen Bundesländer hat (Niedersächsische Staatskanzlei o.J.). Das Land ist in 37 Landkreise, acht kreisfreie Städte und die Region Hannover gegliedert und besitzt elf Oberzentren sowie 88 Mittelzentren (Niedersächsisches Ministerium für Ernährung, Landwirtschaft und Verbraucherschutz 2017).

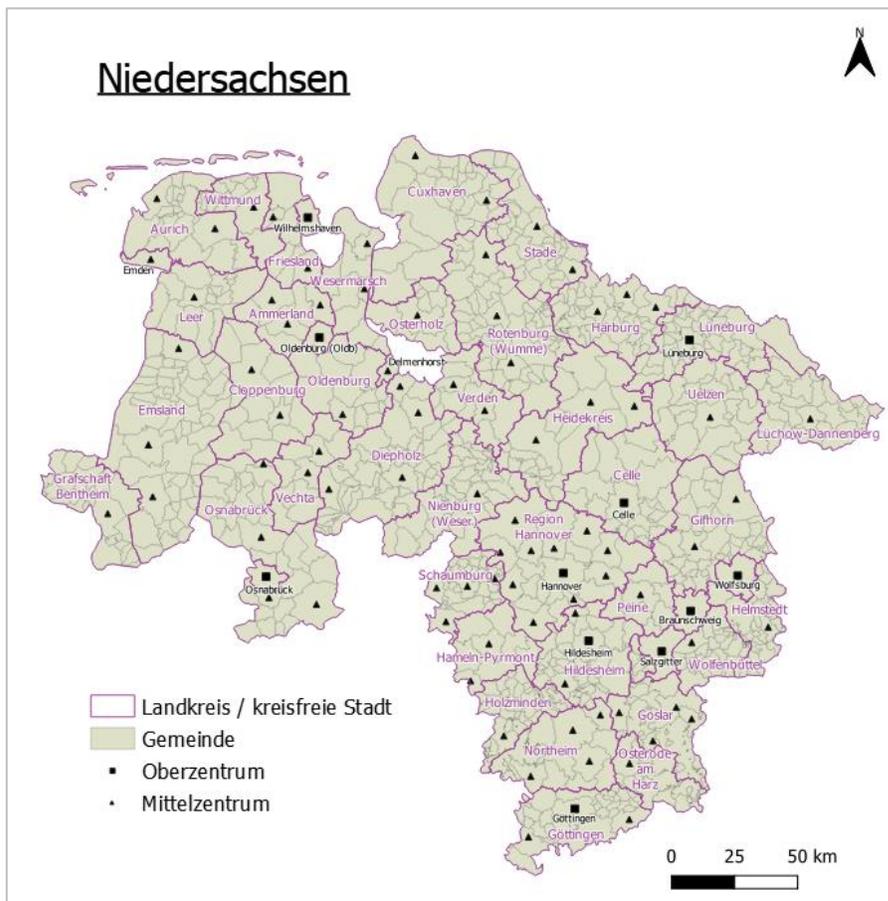


Abbildung 2.3: Überblick über die Verwaltungsstruktur und die zentralen Orte von Niedersachsen. (Eigene Darstellung erstellt mit QGIS; © GeoBasis-DE / BKG 2022, © Niedersächsisches Ministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2017))

2.2 Software

2.2.1 Tabellenkalkulation: Microsoft Excel

Alle im Rahmen dieser Master-Thesis verwendeten Daten ohne Geometrien werden im .xlsb-Format (Excel-Binärarbeitsblatt) bzw. .xlsx-Format (Excel-Arbeitsblatt) ausgegeben, was zur Anzeige bzw. zur Bearbeitung die Verwendung eines Tabellenkalkulationsprogramms wie Microsoft Excel oder Libre Office Calc notwendig macht. Ein Zwischenschritt in der Verarbeitung der Daten wird es sein, diese zur gemeinsamen Verwaltung in eine Datenbank zu importieren. Bevor das geschehen kann, müssen die Daten noch aufbereitet werden (z.B. Filter, Übertragen in ein datenbanktaugliches Format). Diese Prozesse werden im Abschnitt 2.3 beschrieben. Da dazu z.T. auch die Möglichkeiten der VBA-Programmierung benutzen werden, wird Excel in der aktuellen Version von Microsoft 365 (Stand Juli 2022), der Version 2206, genutzt. Verwendete VBA-Skripte werden bei der Aufbereitung mit angegeben.

2.2.2 Datenverwaltung: PostgreSQL mit PostGIS

Als Datenbankmanagementsystem zur Datenhaltung wird die freie Software PostgreSQL in der Version 12.1 genutzt. Um auch räumliche Daten speichern und bearbeiten zu können, ist die Erweiterung PostGIS notwendig, welche in der Version 3.0.0 vorliegt. Der Zugriff auf die Datenbank erfolgt über pgAdmin in der Version 4.16. Zur Bearbeitung und Berechnung der Daten wird die Datenbanksprache SQL genutzt. Alle verwendeten SQL-Befehle werden an den entsprechenden Stellen mit angegeben.

Die Daten beider Bundesländer werden in zwei unterschiedlichen Datenbanken gehalten. Diese Datentrennung sorgt für eine bessere Übertragbarkeit bzw. Automatisierung des Bearbeitungsprozesses, da z.B. verwendete SQL-Skripte ohne Anpassungen von Tabellennamen o.Ä. auch für andere Untersuchungsgebiete verwendet werden können. In meinem Fall heißen die beiden angelegten Datenbanken „PendlerdatenMV“ und „PendlerdatenNDS“. Zur Sicherstellung der Speicherung und Bearbeitung räumlicher Daten kann als Vorlage die automatisch mit der PostGIS-Erweiterung erstellte Datenbank „postgis_30_sample“ genutzt werden.

2.2.3 Geographisches Informationssystem: QGIS

Als geographisches Informationssystem zur weiteren räumlichen Bearbeitung und Analyse der Daten sowie zur Visualisierung der Ergebnisse und Erstellung der in dieser Master-Thesis abgebildeten Karten wird die freie Software QGIS in der LTR-Version 3.16.16 „Hannover“ benutzt. Die verwendeten Verarbeitungswerkzeuge sowie für diese benutzte Parameter werden an den entsprechenden Stellen des Vorgehens angegeben und beschrieben.

2.3 Daten

Im folgenden Abschnitt werden die in dieser Master-Thesis verwendeten Daten vorgestellt. Dabei wird auch jeweils darauf eingegangen, wie die Daten ohne Geometrien in Microsoft Excel zur Vorbereitung für den Import in die PostGIS-Datenbank aufbereitet werden bzw. wie die verwendeten WFS-Datensätze in QGIS eingebunden und von dort in die PostGIS-Datenbank importiert werden.

2.3.1 Verwaltungsgrenzen

Zur flächenhaften Darstellung der Verwaltungseinheiten auf Gemeindeebene wird der Datensatz „Gemeinden“ aus dem WFS-Dienst „Verwaltungsgebiete 1:250 000, Stand 01.01. (wfs_vg250)“ vom Bundesamt für Kartographie und Geodäsie (2022) benutzt. Der Dienst ist auf den aktuellen Stand der Verwaltungsstruktur gehalten und enthält als wichtige Attribute „gen“ (Name) und zur eindeutigen Identifizierung der Verwaltungseinheiten „ags“ (Amtlicher Gemeindeschlüssel, achtstellige Verwaltungsschlüsselnummer) und „ars“ (Amtlicher Regionalschlüssel, zwölfstellige Verwaltungsschlüsselnummer). Der Unterschied der beiden Verwaltungsschlüsselnummern liegt darin, dass in der zwölfstelligen zusätzlich der Gemeindeverband bzw. die Verwaltungsgemeinschaft verschlüsselt enthalten ist.

Der WFS-Dienst soll für die weitere Verarbeitung in der erstellten PostGIS-Datenbank abgespeichert werden. Dazu wird der Dienst zunächst in QGIS eingebunden. Das geschieht über das Bedienfeld „Browser“, in welchem WFS-Verbindungen angelegt und verwaltet sowie die Daten in QGIS eingebunden werden können. Dazu wird ausschließlich die URL des Dienstes (https://sgx.geodatenzentrum.de/wfs_vg250) benötigt. Ist die Verbindung zum Dienst aufgebaut, kann der Layer „Gemeinden“ in QGIS geladen werden.

Ebenfalls über das Bedienfeld „Browser“ können die erstellten PostGIS-Datenbanken in QGIS eingebunden werden. Dazu benötigt man als Information den benutzten PostgreSQL-Host und Port sowie die Namen der Datenbanken.

Da nur die Gemeinden von jeweils einem Bundesland für den Import in die Datenbank benötigt werden, wird auf den WFS-Layer ein Objektfilter angewendet. Mit diesem können ebenfalls die nicht benötigten Gewässerflächen (zu einer Gemeinde zugehöriges Gebiet auf der Ost- bzw. Nordsee) herausgefiltert werden, welche im Datensatz als eigene Objekte hinterlegt sind (siehe Abbildung 2.4). Der Objektfilter für die Gemeinden aus Mecklenburg-Vorpommern lautet: **„gf“ = 4 and „sn_l“ = ,13‘**

In dem Attribut „gf“ („Geofaktor“) ist numerisch hinterlegt, ob ein Objekt die Gewässerfläche (2) oder Landteilfläche (4) einer Gemeinde darstellt. Das Attribut „sn_l“ beinhaltet als Text gespeichert die Verwaltungsschlüsselnummer des Bundeslandes der Gemeinde, wobei Mecklenburg-Vorpommern die Verwaltungsschlüsselnummer 13 innehat. Für den Objektfilter der Gemeinden Niedersachsens muss dieser Wert durch 03 ersetzt werden.



Abbildung 2.4: Ausschnitt des Datensatzes „Gemeinden“ aus dem WFS-Dienst „wfs_vg250“. (Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022; © OpenStreetMap)

Die gefilterten Datensätze enthalten nun alle Gemeinden eines Bundeslandes. Da in dieser Master-Thesis nur die Gemeinden mit einer Bevölkerungszahl unter 5.000 untersucht werden, ist noch ein weiterer Filter notwendig. Da die Daten aus dem WFS-Dienst jedoch keine Bevölkerungszahlen enthalten, muss diese Information erst aus einer anderen Datenquelle entnommen werden. Die Verknüpfung von Geometrien und Bevölkerungszahlen der Gemeinden erfolgt später.

Die in QGIS integrierte „DB-Verwaltung“ kann verwendet werden, um die Gemeinden der beiden Bundesländer in ihre jeweilige Datenbank zu importieren. Dabei muss darauf geachtet werden, ein Attribut als Primärschlüssel festzulegen (hier bietet sich eine Verwaltungsschlüsselnummer an) und auch die Geometrien mit abzuspeichern. Als Name für die Tabellen verwende ich „gemeinden_geometrien“. Nach dem Import können alle Spalten abgesehen von „ags“, „geom“, „ars“ und „gen“ gelöscht werden, da diese für die Bearbeitung nicht benötigt werden.

Im Laufe der Bearbeitung wird aus dem gleichen WFS-Dienst ebenfalls der Layer „Gemeindepunkte“ benötigt. Für diesen ist kein Objektfilter notwendig, er kann vollständig in QGIS eingeladen und in die beiden Datenbanken importiert werden. Als Namen verwende ich „gemeindepunkte“, als Primärschlüssel kann erneut eine der beiden Verwaltungsschlüsselnummern („ags“ oder „ars“) dienen.

2.3.2 Bevölkerungszahlen

Für die Bevölkerungszahlen der Gemeinden wird das Gemeindeverzeichnis des Statistischen Bundesamtes mit Gebietsstand vom 31.12.2021 (Statistisches Bundesamt 2022 b) benutzt. Diese als Excel-Datei zur Verfügung gestellte Tabelle enthält alle politisch selbstständigen Gemeinden Deutschlands mit Flächengröße, Bevölkerungszahl und -dichte sowie weiteren Merkmalen. Zur eindeutigen Kennzeichnung der Verwaltungseinheiten wird der amtliche Regionalschlüssel verwendet (siehe Abbildung 2.5).

Amtlicher Regionalschlüssel (ARS)					Gemeindename	Fläche km ² 1)	Bevölkerung ²⁾			
Land	RB	Kreis	VB	Gem			insgesamt	männlich	weiblich	je km ²
Gebietsstand am 31.12.2021 (Jahr)						31.12.2021 (Jahr)	am 31.12.2021 (Jahr)			
13					Mecklenburg-Vorpommern					
13	0	00	9999		Küstengewässer einschl. Anteil am Festlandssockel					
13	0	00	9999	999	Küstengewässer einschl. Anteil am Festlandssockel	1,00	0	0	0	0
13	0	03			Rostock					
13	0	03	0000		Rostock, Hanse- und Universitätsstadt					
13	0	03	0000	000	Rostock, Hanse- und Universitätsstadt	181,36	208 400	102 559	105 841	1 149
13	0	04			Schwerin					
13	0	04	0000		Schwerin, Landeshauptstadt					
13	0	04	0000	000	Schwerin, Landeshauptstadt	130,52	95 740	45 944	49 796	734
13	0	71			Mecklenburgische Seenplatte					
13	0	71	0027		Dargun, Stadt					
13	0	71	0027	027	Dargun, Stadt	118,01	4 334	2 124	2 210	37
13	0	71	0029		Demmin, Hansestadt					
13	0	71	0029	029	Demmin, Hansestadt	82,06	10 337	4 926	5 411	126
13	0	71	0033		Feldberger Seenlandschaft					
13	0	71	0033	033	Feldberger Seenlandschaft	201,10	4 452	2 143	2 309	22
13	0	71	0107		Neubrandenburg, Vier-Tore-Stadt					
13	0	71	0107	107	Neubrandenburg, Vier-Tore-Stadt	86,12	63 043	30 468	32 575	732
13	0	71	0110		Neustrelitz, Residenzstadt					
13	0	71	0110	110	Neustrelitz, Residenzstadt	139,89	20 103	9 651	10 452	144
13	0	71	0156		Waren (Müritz), Stadt					
13	0	71	0156	156	Waren (Müritz), Stadt	159,46	21 116	9 980	11 136	132
13	0	71	5151		Demmin-Land					
13	0	71	5151	008	Beggerow	29,82	498	249	249	17
13	0	71	5151	014	Borrentin	48,28	768	386	382	16
13	0	71	5151	064	Hohenbollentin	6,54	110	51	59	17
13	0	71	5151	065	Hohenmocker	27,02	478	240	238	18
13	0	71	5151	072	Kentzlin	11,80	205	95	110	17
13	0	71	5151	076	Kletzin	27,28	726	377	349	27
13	0	71	5151	089	Lindenberg	13,08	214	114	100	16
13	0	71	5151	096	Meesiger	8,97	240	119	121	27
13	0	71	5151	112	Nossendorf	35,13	657	330	327	19
13	0	71	5151	128	Sarow	33,75	698	364	334	21
13	0	71	5151	131	Schönfeld	16,21	358	189	169	22
13	0	71	5151	136	Siedenbrünzow	27,60	482	240	242	17
13	0	71	5151	139	Sommersdorf	8,75	230	111	119	26

Die Verwaltungseinheiten auf Gemeindeebene sind die nicht grün und nicht fett geschriebenen Einträge.

Abbildung 2.5: Auszug aus dem Gemeindeverzeichnis des Statistischen Bundesamtes. (Eigener Screenshot in Excel; © Statistisches Bundesamt 2022 b)

Aus dem umfangreichen Tabellenblatt des deutschlandweiten Gemeindeverzeichnisses müssen alle Verwaltungseinheiten auf Gemeindeebene aus Mecklenburg-Vorpommern bzw. Niedersachsen mit folgenden Informationen extrahiert werden:

- amtlicher Regionalschlüssel
- Flächengröße (km²)
- Bevölkerungszahl
- Bevölkerungsdichte (je km²)

Alle anderen Zeilen und Spalten werden nicht benötigt. Das Ziel der Aufbereitung ist es, jeweils ein Tabellenblatt nur mit den benötigten Daten der Gemeinden jeweils eines Bundeslandes zu erhalten. Dazu wird wie folgt vorgegangen:

- 1.) Auflösen der verbundenen Zellen in den ersten drei Zeilen im Tabellenblatt, da diese bei der weiteren Aufbereitung hinderlich sind
- 2.) Filter in Spalte C (Verwaltungsschlüsselnummer Bundesland): Nur Anzeigen des Wertes „13“ (Mecklenburg-Vorpommern) bzw. „03“ (Niedersachsen)
- 3.) Filter in Spalte G (Verwaltungsschlüsselnummer Gemeinde): Kein Anzeigen von leeren Zellen (so werden die übergeordneten Zeilen z.B. mit den Landkreisen rausgefiltert)
- 4.) Kopieren aller gefilterten Zeilen in ein neues Tabellenblatt (falls vorhanden, ohne „Küstengewässer“)
- 5.) In diesem werden alle Spalten bis auf C-G (Verwaltungsschlüsselnummer), I (Fläche), J (Bevölkerungszahl) und M (Bevölkerungsdichte) gelöscht
- 6.) In einer neuen Spalte werden mit Hilfe der Excel-Funktion „Textkette“ die einzelnen Bestandteile der Verwaltungsschlüsselnummern zusammengesetzt
- 7.) Die Spalten mit der Flächengröße, Bevölkerungszahl und -dichte werden als Zahl formatiert, um die Tausender-Trennzeichen zu entfernen, auftretende Nachkommastellen mit Null (bei Bevölkerungszahl und -dichte) werden entfernt
- 8.) In der 1. Zeile werden Überschriften ergänzt

land	regierungsbezirk	kreis	gemeindeverband	gemeinde	flaeche	bevoelkerung	bevoelkerungsdichte	verwaltungsschluesselnummer
13	0	03	0000	000	181,36	208400	1149	130030000000
13	0	04	0000	000	130,52	95740	734	130040000000
13	0	71	0027	027	118,01	4334	37	130710027027
13	0	71	0029	029	82,06	10337	126	130710029029
13	0	71	0033	033	201,10	4452	22	130710033033
13	0	71	0107	107	86,12	63043	732	130710107107
13	0	71	0110	110	139,89	20103	144	130710110110
13	0	71	0156	156	159,46	21116	132	130710156156
13	0	71	5151	008	29,82	498	17	130715151008
13	0	71	5151	014	48,28	768	16	130715151014
13	0	71	5151	064	6,54	110	17	130715151064
13	0	71	5151	065	27,02	478	18	130715151065
13	0	71	5151	072	11,80	205	17	130715151072
13	0	71	5151	076	27,28	726	27	130715151076
13	0	71	5151	089	13,08	214	16	130715151089
13	0	71	5151	096	8,97	240	27	130715151096
13	0	71	5151	112	35,13	657	19	130715151112
13	0	71	5151	128	33,75	698	21	130715151128
13	0	71	5151	131	16,21	258	22	130715151131

Abbildung 2.6: Ergebnis der Aufbereitung des Gemeindeverzeichnisses für Mecklenburg-Vorpommern (Auszug). (Eigener Screenshot in Excel; © Statistisches Bundesamt 2022 b [bearbeitet])

Zur Vorbereitung des Imports in die PostGIS-Datenbank werden die Tabellenblätter mit dem aufbereiteten Gemeindeverzeichnis aus Excel heraus als CSV-Dateien exportiert. In diesen müssen anschließend noch die Kommata der Dezimalzahlen durch Punkte ersetzt werden (z.B. durch eine „Suchen und Ersetzen“-Funktion in einem Texteditor), da die Werte von PostgreSQL beim Import ansonsten nicht als Dezimalzahlen erkannt werden.

2.3.3 Pendeldaten

Die in dieser Master-Thesis verwendeten Pendeldaten stammen vom Statistik-Service der Bundesagentur für Arbeit. Als Pendelnde werden in diesen Daten sozialversicherungspflichtig Beschäftigte definiert, deren Wohn- sich vom Arbeitsort auf Gemeindeebene unterscheidet. Beide Angaben müssen von Arbeitgebenden gemeldet werden, wobei gewisse Unschärfen auftreten können, da einerseits keine einheitliche Regelung besteht, ob Erst- oder Hauptwohnsitz der Arbeitnehmenden angegeben werden müssen und andererseits es Arbeitgebende gibt, welche all ihre Beschäftigten aus Niederlassungen unterschiedlicher Gemeinden unter der Betriebsnummer (und damit der Gemeinde) der Hauptniederlassung melden (Statistik der Bundesagentur für Arbeit 2022 a, Tabellenblatt „Erläuterungen 1“). Aus den Angaben lassen sich für alle Gemeinden die Anzahl der sozialversicherungspflichtig Beschäftigten ermitteln, welche in dieser Gemeinde wohnen und bzw. oder arbeiten. Auch die Anzahl der Ein- und Auspendelnden sowie Binnenpendelnden jeder Gemeinde sind dadurch bestimmt.

Die Pendeldaten teilen sich auf zwei Datensätze auf. Zunächst gibt es die Pendelwerte jeder Gemeinde (sozialversicherungspflichtig Beschäftigte mit Wohn- bzw. Arbeitsort, Ein-, Aus- und Binnenpendelnde). Die Pendelströme, also die Anzahl der sozialversicherungspflichtig Beschäftigten, die von einer bestimmten Gemeinde in eine andere bestimmte Gemeinde pendeln, werden in einer zweiten Datei herausgegeben.

2.3.3.1 Pendelwerte

Die Pendelwerte werden in der Tabelle „Gemeindedaten der sozialversicherungspflichtig Beschäftigten nach Wohn- und Arbeitsort“ herausgegeben. Stichtag der Angaben ist der 30.06. jeden Jahres, die Herausgabe erfolgt am 15.01. des Folgejahres. Die Statistik steht auf der Homepage des Statistik-Service der Bundesagentur für Arbeit kostenlos zum Download bereit (Statistik der Bundesagentur für Arbeit 2022 a). In dieser Master-Thesis werden die Daten mit dem Stichtag 30.06.2021 benutzt.

Aus der Statistik kann die Anzahl der sozialversicherungspflichtig Beschäftigten mit Wohn- bzw. Arbeitsort für alle Verwaltungseinheiten auf Gemeinde- und Kreisebene (Tabellenblatt „Gemeindedaten“) sowie auf Landesebene (Tabellenblatt „Länder“) abgelesen werden. Ebenfalls ist angegeben, für wie viele Beschäftigte Wohn- und Arbeitsort gleich sind und wie viele Ein- bzw. Auspendelnde es in der jeweiligen Verwaltungseinheit gibt. Aus diesen Angaben lassen sich die meisten der in dieser Master-Thesis benutzten Pendelparameter berechnen, z.B. die Ein- und Auspendelquoten.

Weitere Tabellenblätter, z.B. mit Erläuterungen zu den verwendeten Begriffen der Statistik, ergänzen die Tabelle. Die Verwaltungseinheiten sind auf Gemeindeebene durch ihre achtstellige Verwaltungsschlüsselnummer eindeutig gekennzeichnet.

A		B		C	D	E	F	G	H	I	J	K	L	M	N	O
1 Bundesagentur für Arbeit 2 Statistik		Gemeindedaten aus der Beschäftigungsstatistik														
		zurück zum Inhalt														
		Sozialversicherungspflichtig Beschäftigte am 30.06.2021 mit Wohnort / Arbeitsort in der vg. Gemeinde														
		davon / darunter Spalte 1														
6 Schl. Nr.	7 Gemeinde/Landkreis Regierungsbezirk Land	8 Wohnort	Arbeitsort						10 Wohnort gleich Arbeitsort	11 Ein- pendler ¹⁾	12 Aus- pendler ¹⁾	13 Zahl der Betriebe				
			Männer	Frauen	Deutsche	Ausländer	u. 20 J.	20 - u 25 J.					55 J. u alter			
1	2	3	4	5	6	7	8	9	10	11	12	13				
723	13003000	Rostock, Hanse- und Universitätsstadt	81.340	42.120	39.220	76.226	5.112	1.253	6.516	17.095	93.711	58.577	35.086	22.760	5.028	
724																
725	13004000	Schwerin, Landeshauptstadt	35.831	17.897	17.934	33.735	2.096	772	2.547	8.521	50.752	25.170	25.557	10.659	2.873	
726																
727	13071001	Alt Schwerin	197	105	92	*	*	*	10	57	195	42	153	155	24	
728	13071002	Altenhagen	131	67	64	*	*	*	3	35	23	5	18	126	7	
729	13071003	Altenhof	158	83	75	154	4	*	9	37	68	11	57	147	11	
730	13071004	Altentreptow, Stadt	1.873	979	894	1.845	28	49	99	542	2.453	674	1.779	1.199	212	
731	13071005	Ankershagen, Schliemanngemeinde	197	102	95	*	*	5	4	82	96	27	69	170	15	
732	13071006	Bartow	155	86	69	*	*	3	4	42	*	15	*	140	*	
733	13071007	Basedow	252	111	141	*	*	5	10	98	95	31	64	221	22	
734	13071008	Beggerow	180	91	89	*	*	*	8	46	22	12	10	168	10	
735	13071009	Beseritz	35	16	19	35	0	0	0	14	*	*	*	*	*	
736	13071010	Blankenhof	304	144	160	304	0	7	12	94	87	12	75	292	14	
737	13071011	Blankensee	689	359	330	681	8	15	33	192	173	64	109	625	31	
738	13071012	Blumenholz	328	163	165	324	4	6	9	96	79	13	66	315	11	
739	13071013	Bollewick	255	126	129	250	5	8	10	73	154	30	123	225	30	
740	13071014	Borrentin	292	151	141	292	0	9	12	76	90	31	59	261	15	
741	13071015	Bredenfelde	75	40	35	75	0	*	9	19	*	6	*	69	*	

Abbildung 2.7: Auszug des Tabellenblatts „Gemeindedaten“ mit Pendelwerten aus Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 a)

Es kommt in dieser Statistik aus Gründen des Datenschutzes zu Anonymisierungen auf Gemeindeebene, wenn Werte unter drei vorkommen, es weniger als drei Betriebe in einer Gemeinde gibt oder ein Betrieb einer Gemeinde einen zu großen Beschäftigtenanteil auf sich vereint (Statistik der Bundesagentur für Arbeit 2022 a, Tabellenblatt „Erläuterungen 2“). In den betroffenen Zellen der Excel-Tabelle steht anstelle des tatsächlichen Wertes ein *.

Aus dem umfangreichen Tabellenblatt der deutschlandweiten Gemeindedaten müssen alle Verwaltungseinheiten auf Gemeindeebene aus Mecklenburg-Vorpommern bzw. Niedersachsen mit folgenden Informationen extrahiert werden:

- achtstellige Verwaltungsschlüsselnummer
- Anzahl der sozialversicherungspflichtig Beschäftigten mit Wohnort in der Gemeinde
- Anzahl der sozialversicherungspflichtig Beschäftigten mit Arbeitsort in der Gemeinde
- Anzahl der Binnenpendelnden innerhalb der Gemeinde
- Anzahl der Einpendelnden in die Gemeinde
- Anzahl der Auspendelnden aus der Gemeinde

Alle anderen Zeilen und Spalten werden nicht benötigt. Das Ziel der Aufbereitung ist es, jeweils ein Tabellenblatt nur mit den benötigten Daten der Gemeinden jeweils eines Bundeslandes zu erhalten. Dazu wird wie folgt vorgegangen:

- 1.) Filter in Spalte A (Verwaltungsschlüsselnummer): Nur Anzeigen von Werten, welche mit „13“ beginnen (Mecklenburg-Vorpommern) bzw. welche mit „03“ beginnen (Niedersachsen) und aus acht Stellen bestehen (damit werden die aggregierten Zeilen mit den Informationen der Landkreise herausgefiltert)
- 2.) Kopieren aller gefilterten Zeilen in ein neues Tabellenblatt
- 3.) In diesem werden alle Spalten bis auf A, C und K bis N gelöscht (in den sechs verbleibenden Spalten stehen die benötigten Informationen)
- 4.) Alle Zellen, in welchen ein * steht (anonymisierter Wert), werden durch „Suchen und Ersetzen“ geleert (Suchen nach: „~*“, Ersetzen durch: leer lassen)
- 5.) In der 1. Zeile werden Überschriften ergänzt

Zur Vorbereitung des Imports in die PostGIS-Datenbank werden die Tabellenblätter mit den aufbereiteten Pendelwerten aus Excel heraus als CSV-Datei exportiert. In der erstellten CSV-Datei müssen mit Hilfe eines Texteditors alle Tausender-Trennzeichen entfernt werden, damit der Import in die Datenbank erfolgen kann.

vsn	wohntort	arbeitsort	binnenpendelnde	einpendelnde	auspendelnde
13003000	81.340	93.711	58.577	35.086	22.760
13004000	35.831	50.752	25.170	25.557	10.659
13071001	197	195	42	153	155
13071002	131	23	5	18	126
13071003	158	68	11	57	147
13071004	1.873	2.453	674	1.779	1.199
13071005	197	96	27	69	170
13071006	155		15		140
13071007	252	95	31	64	221
13071008	180	22	12	10	168
13071009	35				
13071010	304	87	12	75	292
13071011	689	173	64	109	625
13071012	328	79	13	66	315
13071013	255	154	30	123	225
13071014	292	90	31	59	261
13071015	75		6		69
13071016	200		10		190
13071017	54		0		54
13071018	138				
13071019	413	217	28	188	385
13071020	48	11	3	8	45
13071021	2.078	741	232	508	1.846
13071022	359	259	59	200	300
13071023	209	123	39	83	170
13071025	311	48	13	35	298
13071026	358	72	11	61	347
13071027	1.568	1.162	590	572	978
13071028	347	94	26	68	321
13071029	3.598	4.714	1.820	2.893	1.778

Abbildung 2.8: Ergebnis der Aufbereitung der Pendelwerte für Mecklenburg-Vorpommern (Auszug). (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit (2022 a) [bearbeitet])

2.3.3.2 Pendelströme

Die Daten zu den Pendelströmen werden vom Statistik-Service der Bundesagentur für Arbeit in der Datei „Pendlerverflechtungen der sozialversicherungspflichtig Beschäftigten nach Gemeinden“ herausgegeben. Wie bei den Pendelwerten liegt der Stichtag der Statistik jährlich auf dem 30.06., die Veröffentlichung der Daten erfolgt im Februar des Folgejahres. Auch hier liegen für diese Master-Thesis die Daten mit Stand 30.06.2021 vor. Während die Pendelströme auf Landes- und Kreisebene öffentlich zur Verfügung gestellt werden, erfolgt

die Herausgabe der Daten auf Gemeindeebene nur gegen Bezahlung und enthält dabei nur die Pendelströme in die bzw. aus den Gemeinden jeweils eines Bundeslandes.

In der Statistik werden die genauen Pendelströme aufgeschlüsselt, sprich wie viele sozialversicherungspflichtig Beschäftigte jeweils von einem bestimmten Wohnort zu einem bestimmten Arbeitsort pendeln. Die Excel-Datei enthält fünf Tabellenblätter, drei davon dienen der Information rund um die Statistik, die Daten selbst sind auf den Tabellenblättern „Auspendler Gemeinden“ und „Einpendler Gemeinden“ zu finden.

Im Tabellenblatt „Auspendler Gemeinden“ wird jede Gemeinde des Bundeslandes als Wohnort betrachtet und aufgelistet, in welchen anderen Gemeinden Deutschlands die Einwohnenden dieser ihren Arbeitsort haben. Außerdem erfolgen Aggregationen auf höhere Verwaltungsebenen (Kreis, Regierungsbezirk, Land). Die Gemeinden sind durch ihre achtstellige Verwaltungsschlüsselnummer eindeutig bestimmt.

Wohnort		Arbeitsort		Insgesamt	Männer	Frauen	Deutsche	Ausländer + Sonstige	Auszubildende
				1	2	3	4	5	6
13003000	Rostock, Hanse- und Universitätsstadt	01001000	Flensburg, Stadt	18	*	*	18	-	-
		01001	Flensburg, Stadt	18	*	*	18	-	-
		01002000	Kiel, Landeshauptstadt	90	62	28	86	4	*
		01002	Kiel, Landeshauptstadt	90	62	28	86	4	*
		01003000	Lübeck, Hansestadt	170	134	36	155	15	8
		01003	Lübeck, Hansestadt	170	134	36	155	15	8
		01004000	Neumünster, Stadt	23	17	6	*	*	*
		01004	Neumünster, Stadt	23	17	6	*	*	*
		01051	Dithmarschen	13	*	*	13	-	-
		01053	Herzogtum Lauenburg	34	23	11	34	-	-
		01054	Nordfriesland	16	13	3	16	-	*
		01055044	Scharbeutz	13	7	6	*	*	-
		01055046	Fehmarn, Stadt	26	16	10	*	*	-
		01055	Übrige Gemeinden (Kreis)	30	22	8	27	3	*
		01055	Ostholstein	69	45	24	63	6	*

Abbildung 2.9: Auszug des Tabellenblatts „Auspendler Gemeinden“ der Pendelströme von Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

Das Tabellenblatt „Einpendler Gemeindeebene“ funktioniert genau umgekehrt, die Gemeinden des Bundeslandes werden als Arbeitsort betrachtet, in welche Einwohner anderer Gemeinden Deutschlands (und zusätzlich aus dem Ausland) einpendeln. Die Aufschlüsselungen und Aggregationen erfolgen analog zu „Auspendler Gemeinden“.

	A	B	C	D	E	F	G	H	I	J
443			16051000	Erfurt, Stadt	17	8	9	17	-	-
444			16051	Erfurt, Stadt	17	8	9	17	-	-
445			16053000	Jena, Stadt	14	10	4	*	*	-
446			16053	Jena, Stadt	14	10	4	*	*	-
447			160	Übrige Kreise (Regierungsbezirk)	69	44	25	65	4	3
448			160	Thüringen	100	62	38	95	5	3
449			16	Thüringen	100	62	38	95	5	3
450			PL_00000	Polen	401	387	14	18	383	-
451			PL	Polen	401	387	14	18	383	-
452			PL	Polen	401	387	14	18	383	-
453			PL	Polen	401	387	14	18	383	-
454			RP_00000	Philippinen	26	26	-	-	26	-
455			RP	Philippinen	26	26	-	-	26	-
456			RP	Philippinen	26	26	-	-	26	-
457			RP	Philippinen	26	26	-	-	26	-
458			ZZ	Übrige Bundesländer / Staaten	32	27	5	7	25	-
459			ZZ	Einpendler aus dem Bundesgebiet	34.635	18.416	16.219	33.476	1.159	1.247
460			Z	Einpendler insgesamt	35.086	18.850	16.236	33.495	1.591	1.247
461	13004000	Schwerin, Landeshauptstadt								
462			01001000	Flensburg, Stadt	11	7	4	*	*	-
463			01001	Flensburg, Stadt	11	7	4	*	*	-
464			01002000	Kiel, Landeshauptstadt	21	15	6	*	*	*
465			01002	Kiel, Landeshauptstadt	21	15	6	*	*	*
466			01003000	Lübeck, Hansestadt	125	74	51	120	5	3

Abbildung 2.10: Auszug des Tabellenblatts „Einpendler Gemeinden“ der Pendelströme von Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

Wie bei den Pendelwerten werden auch bei den Pendelströmen Werte aus Datenschutzgründen anonymisiert, wenn diese einen gewissen Mindestwert nicht erreichen. So werden für eine Gemeinde des Bundeslandes keine Gemeinden angegeben, bei welchen die Anzahl der Einpendelnden in diese bzw. der Auspendelnden aus dieser kleiner als zehn ist, solche Gemeinden werden unter „Übrige“ zusammengefasst. Auch erfolgt eine Anonymisierung, wenn in einer Gemeinde weniger als drei Betriebe ansässig sind oder einer der ansässigen Betriebe einen zu hohen Beschäftigtenanteil auf sich vereint (Statistik der Bundesagentur für Arbeit 2022 b, Tabellenblatt „Methodische Hinweise“).

2630	13071140	Sponholz								
2631			13071021	Burg Stargard, Stadt	10	10	-	7	3	-
2632			13071035	Friedland, Stadt	10	7	3	10	-	-
2633			13071107	Neubrandenburg, Vier-Tore-Stadt	61	53	8	*	*	-
2634			13071	Übrige Gemeinden (Kreis)	52	40	12	*	*	3
263			13071	Mecklenburgische Seenplatte	133	110	23	128	5	3
263			13075	Vorpommern-Greifswald	17	17	-	17	-	-
263			130	Übrige Kreise (Regierungsbezirk)	12	12	-	*	*	-
263			130	Mecklenburg-Vorpommern	162	139	23	156	6	3
263			13	Mecklenburg-Vorpommern	162	139	23	156	6	3
264			PL_00000	Polen	73	*	*	-	73	-
264			PL	Polen	73	*	*	-	73	-
2642			PL	Polen	73	*	*	-	73	-
2643			PL	Polen	73	*	*	-	73	-
2644			ZZ	Übrige Bundesländer / Staaten	16	*	*	*	*	*
2645			ZZ	Einpendler aus dem Bundesgebiet	178	154	24	170	8	4
2646			Z	Einpendler insgesamt	251	225	26	170	81	4

Abbildung 2.11: Anonymisierungen in den Pendelströmen von Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

Die Statistik wird der Ermittlung der durchschnittlichen Distanz dienen, welche Einpendelnde in die bzw. Auspendelnde aus den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens zurücklegen. Um die beiden Tabellenblätter ebenfalls in die PostGIS-Datenbank importieren zu können, ist eine umfangreiche Aufbereitung notwendig, da diese im Gegensatz zu den anderen verwendeten Daten allein von ihrer Grundstruktur her noch nicht für eine Datenbank geeignet sind. So liegen z.B. Start- und Zielort eines Pendelstroms nie in der gleichen Zeile der Tabelle vor. Außerdem müssen sowohl die Zeilen mit aggregierten Werten (z.B. Gesamtanzahl der Auspendelnden aus einer Gemeinde in alle Gemeinden eines Landkreises) als auch mit anonymisierten Werten herausgefiltert werden. Auch die Einpendelströme aus anderen Staaten interessieren nicht, da diese nur den Staat als Wohnort der Einpendelnden angeben und keine genauere Unterscheidung auf Gemeindeebene o.Ä. erfolgt (siehe Abbildungen 2.10 und 2.11), womit die Angabe für die Untersuchung zu ungenau ist (Auspendelströme in andere Staaten sind nicht erfasst).

Im Folgenden wird das Vorgehen der Aufbereitung für die Auspendelströme beschrieben, welches analog bis auf eine kleine Änderung auch für die Einpendelströme angewendet werden kann. Folgende Informationen sollen in der Tabelle für jeden Auspendelstrom enthalten sein:

- Achtstellige Verwaltungsschlüsselnummer des Wohnortes
- Achtstellige Verwaltungsschlüsselnummer des Arbeitsortes
- Gesamtanzahl der Auspendelnden dieses Stroms

Dazu wird wie folgt vorgegangen:

- 1.) Filter in Spalte C (Verwaltungsschlüsselnummer Arbeitsort): Nur Anzeigen von Werten, welche achtstellig sind, sowie von leeren Zellen (Herausfiltern der Zeilen mit aggregierten Werten)
- 2.) Filter in Spalte E (Gesamtanzahl der Pendelnden eines Pendelstroms): Herausfiltern der Zellen, welche ein „*“ enthalten (Herausfiltern der anonymisierten Werte)
- 3.) Kopieren der Spalten A (Verwaltungsschlüsselnummer Wohnort), C und E von allen gefilterten Zeilen in ein neues Tabellenblatt (z.B. „auspendler_temp“)
- 4.) Erstellung eines weiteren Tabellenblatts (z.B. „auspendler_db“) und in diesem Definition der Spalten A und B als Text (ansonsten würde bei Verwaltungsschlüsselnummern, deren 1. Stelle eine 0 ist, diese weggeschnitten werden, weil Excel die Ziffernfolgen als Zahlen erkennt)
- 5.) Erstellung und Ausführung des VBA-Makros „matrix_auspendler“ (siehe VBA 2.1)
- 6.) In der 1. Zeile von „auspendler_db“ werden Überschriften ergänzt

Das Makro durchläuft das Tabellenblatt „auspendler_temp“ und speichert alle dort vorhandenen Informationen in einer Art Matrix ab und fügt diese dann so in das leere Tabellenblatt „auspendler_db“ wieder ein, dass Start- und Zielort sowie die Anzahl der Pendelnden eines Stroms in einer gemeinsamen Zeile abgelegt werden.

03151002			03151002	03101000	20
	03101000	20	03151002	03103000	325
	03103000	325	03151002	03151009	26
	03151009	26	03151002	03151039	16
	03151039	16	03151003	03101000	14
03151003			03151003	03103000	286
	03101000	14	03151003	03151005	14
	03103000	286	03151003	03151009	17
	03151005	14	03151004	03101000	29
	03151009	17	03151004	03103000	426
03151004			03151004	03151009	25
	03101000	29	03151004	03151039	30
	03103000	426	03151005	03101000	44
	03151009	25	03151005	03103000	740
	03151039	30	03151005	03151009	69
03151005			03151005	03151011	15
	03101000	44	03151005	03151024	12
	03103000	740			
	03151009	69			
	03151011	15			
	03151024	12			

Abbildung 2.12: Die Auspendelströme vor und nach Ausführung des Makros „matrix_auspender“. (Eigene Screenshots in Excel; © Statistik der Bundesagentur für Arbeit (2022 b) [bearbeitet])

Bei den Einpendelströmen ist zu Beginn als weiteres Filterkriterium in Spalte C notwendig, alle Zellen mit einem „_“ herauszufiltern, welche die Einpendelnden aus anderen Staaten kennzeichnen (z.B. PL_00000 = Polen).

Die Tabellenblätter „auspendler_db“ und „einpendler_db“ werden anschließend als CSV exportiert, um sie in die PostGIS-Datenbank importieren zu können.

```

'Kommentare beziehen sich auf die zuvor stehende Zeile
Sub matrix_auspendler()
  Dim matrix_auspendler(1 To 6000, 1 To 3)
  '6000 ist variabel, muss größer als Anzahl Zeilen in "auspendler_temp" sein
  Dim i, j
  i = 1
  Dim aktuelle_gemeinde_id As String
  Dim pendelgemeinde_id As String
  Dim pendler
  Worksheets("auspendler_temp").Activate
  Range("A2").Select
  '1. Zelle mit einer Verwaltungsschlüsselnummer in Spalte A
  Do Until IsEmpty(ActiveCell)
    aktuelle_gemeinde_id = ActiveCell.Value
    ActiveCell.Offset(1, 1).Select
    Do Until IsEmpty(ActiveCell)
      pendelgemeinde_id = ActiveCell.Value
      ActiveCell.Offset(0, 1).Select
      pendler = ActiveCell.Value
      matrix_auspendler(i, 1) = aktuelle_gemeinde_id
      matrix_auspendler(i, 2) = pendelgemeinde_id
      matrix_auspendler(i, 3) = pendler
      i = i + 1
      ActiveCell.Offset(1, -1).Select
    Loop
    ActiveCell.Offset(0, -1).Select
  Loop
  Worksheets("auspendler_db").Activate
  Range("A2").Select
  For j = 1 To 6000
  '6000 siehe oben, es muss der gleiche Wert sein
    ActiveCell.Value = matrix_auspendler(j, 1)
    ActiveCell.Offset(0, 1).Select
    ActiveCell.Value = matrix_auspendler(j, 2)
    ActiveCell.Offset(0, 1).Select
    ActiveCell.Value = matrix_auspendler(j, 3)
    ActiveCell.Offset(1, -2).Select
  Next j
End Sub

```

VBA 2.1: Makro „matrix_auspendler“ zum Umformen der Auspendelströme in datenbanktaugliche Form.

2.3.4 Datenbank-Import

Während die räumlichen Daten der Gemeinden und Gemeindepunkte bereits von QGIS aus in die PostGIS-Datenbank importiert wurden (siehe Abschnitt 2.3.1), steht dieser Schritt bei den vorgestellten Daten ohne Geometrie noch aus. Für beide Bundesländer wurden jeweils vier CSV-Dateien aus den Statistiken exportiert - das Gemeindeverzeichnis mit den Bevölkerungszahlen, die Pendelwerte, die Aus- und die Einpendelströme. Um diese in die PostGIS-Datenbank importieren zu können, müssen zunächst entsprechende Tabellen erstellt werden, wofür SQL genutzt wird.

<pre>CREATE TABLE gemeinden_einwohnerzahlen (land character varying(2), regierungsbezirk character varying(1), kreis character varying(2), gemeindevorband character varying(4), gemeinde character varying(3), flaeche_km real, bevoelkerung integer, bev_dichte_km real, vsn character varying(12) PRIMARY KEY); CREATE TABLE einpendelstroeme (vsn_arbeit character varying(8), vsn_wohn character varying(8), pendelnde integer, PRIMARY KEY (vsn_arbeit, vsn_wohn));</pre>	<pre>CREATE TABLE gemeinden_pendelparameter (vsn character varying(8) PRIMARY KEY, wohnort integer, arbeitsort integer, binnenpendelnde integer, einpendelnde integer, auspendelnde integer); CREATE TABLE auspendelstroeme (vsn_wohn character varying(8), vsn_arbeit character varying(8), pendelnde integer, PRIMARY KEY (vsn_wohn, vsn_arbeit));</pre>
---	--

SQL 2.1: Erstellung der Tabellen zum Import der CSV-Dateien.

Nach Ausführung der SQL-Befehle in beiden Datenbanken können die Daten aus den jeweiligen CSV-Dateien in die Datenbank-Tabellen importiert werden. Dazu kann die „Import/Export“-Funktion in pgAdmin genommen werden. Folgende Einstellungen müssen dabei benutzt werden:

- Format: csv
- Encoding: UTF8
- Header: Yes
- Delimiter: ;

Nach dem Import können noch in der Tabelle „gemeinden_einwohnerzahlen“ die Spalten mit den einzelnen Bestandteilen der Verwaltungsschlüsselnummer („land“, „regierungsbezirk“ usw.) gelöscht werden, da diese nicht benötigt werden.

2.4 Methodik und Vorgehen

In den folgenden Abschnitten wird das Vorgehen zur Beantwortung der Forschungsfrage am Beispiel der Gemeinden in Mecklenburg-Vorpommern detailliert vorgestellt. Der Prozess lässt sich für die verwendeten räumlichen Faktoren und Pendelparameter auf jedes andere Untersuchungsgebiet übertragen. Spezielle Anmerkungen zu den Daten von Niedersachsen und der Vorgehensweise bei der Bearbeitung dieser werden im letzten Unterabschnitt dieses Kapitels zusammengefasst.

Schritt	verwendete Werkzeuge	Abschnitt
Aufbereitung der Daten & Datenbank-Import	Excel (VBA-Programmierung und CSV-Export) QGIS (WFS-Einbindung und -Export) PostgreSQL (SQL und CSV-Import)	2.3.1 bis 2.3.4
Verknüpfung der Daten & Bevölkerungszahlen-Filter	PostgreSQL (SQL mit Verknüpfung)	2.4.1
Bestimmung räumlicher Parameter		2.4.2
Fläche & Bevölkerungsdichte	PostgreSQL (SQL)	
Küstenanteil	QGIS (Feldrechner, Grenze, Auflösen, Digitalisierung, Puffer, Zuschneiden, Verknüpfung)	
Entfernung Oberzentrum	QGIS (Abstand zum nächsten Knoten)	
Shape-Index	PostgreSQL (SQL)	
Bestimmung Pendelparameter		2.4.3
Ein- und Auspendelquote	PostgreSQL (SQL)	
Arbeitsplatzausstattung	PostgreSQL (SQL)	
Independence-Index	PostgreSQL (SQL)	
Ein- und Auspendeldistanz	PostgreSQL (SQL mit Verknüpfung)	
Statistische Auswertung		2.4.4
deskriptive statistische Maße (Anzahl, Minimum, Maximum, Mittelwert, Quartile, Varianz, Standardabweichung, Schiefe, Wölbung)	PostgreSQL (SQL mit Funktionen zur Schiefe- und Wölbungsberechnung)	
Korrelationsanalyse (Pearson-Produkt-Moment-Korrelation)	PostgreSQL (SQL)	

Tabelle 2.1: Überblick der Methodik.

2.4.1 Verknüpfung der Daten und Bevölkerungszahlen-Filter

Zunächst werden die Informationen der drei importierten Tabellen (Geometrien, Bevölkerungszahlen und Pendelwerte) miteinander verknüpft (die Informationen der Aus- und Einpendelströme werden an späterer Stelle verarbeitet und mit den Gemeinden verknüpft). Da in allen Tabellen mind. eine Verwaltungsschlüsselnummer der Gemeinden gespeichert ist, können diese als Verknüpfungsattribut genutzt werden.

Eine vorherige Überprüfung mit SQL stellt sicher, dass alle drei Tabellen vollständig kompatibel sind und es keine Unterschiede der enthaltenen Gemeinden gibt, z.B. aufgrund einer Gebietsreform.

```
SELECT *
FROM gemeinden_geometrien a, gemeinden_einwohnerzahlen b,
gemeinden_pendelparameter c
WHERE a.ars = b.vsn AND a.ags = c.vsn;
```

SQL 2.2: Prüfung der Kompatibilität von Geometrien, Bevölkerungszahlen und Pendelparametern

Die SELECT-Anfrage findet 726 übereinstimmende Einträge, was auch der Anzahl der Zeilen in allen drei Tabellen entspricht, wie eine Überprüfung deren Eigenschaften ergibt. Damit beziehen sich alle drei Tabellen auf den identischen Gebietsstand in Mecklenburg-Vorpommern und können ohne vorherige Bearbeitung miteinander verknüpft werden.

Während der Verknüpfung werden alle Gemeinden mit einer Bevölkerungszahl von mind. 5000 herausgefiltert, da diese nicht Untersuchungsgegenstand der Master-Thesis sind.

```
CREATE TABLE gemeinden_temp_1 AS
(SELECT a.*, b.flaeche_km, b.bevoelkerung, b.bev_dichte_km,
c.wohnort, c.arbeitsort, c.binnenpendelnde, c.einpendelnde, c.auspendelnde
FROM gemeinden_geometrien a, gemeinden_einwohnerzahlen b,
gemeinden_pendelparameter c
WHERE a.ars = b.vsn AND a.ags = c.vsn AND b.bevoelkerung < 5000);
```

SQL 2.3: Verknüpfung der drei Tabellen mit Herausfiltern der Gemeinden mit mind. 5000 Einwohnenden.

In der neu erstellten Tabelle („gemeinden_temp_1“) sind noch 674 der ursprünglich 726 Einträge vorhanden. Zu den 52 herausgefilterten Gemeinden gehören alle Ober- und Mittelzentren von Mecklenburg-Vorpommern, die übrigen Gemeinden haben fast alle die Funktion eines Grundzentrums inne. Jedoch gibt es auch Grundzentren mit weniger als 5000 Einwohnenden, die damit im Datenbestand verbleiben. Nur eine der herausgefilterten Gemeinden (Sundhagen) zählt nicht zu den Zentralen Orten des Landes (Ministerium für Energie, Infrastruktur und Landesentwicklung Mecklenburg-Vorpommern 2016).

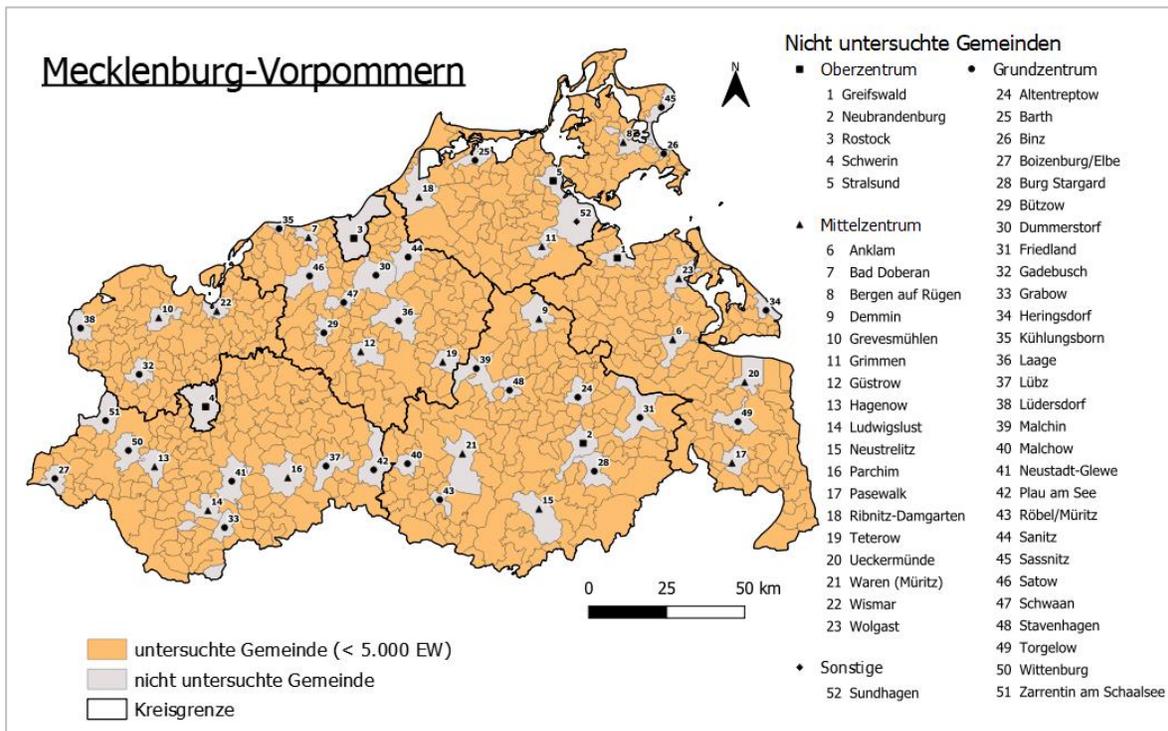


Abbildung 2.13: Überblick über die Gemeinden Mecklenburg-Vorpommerns, welche nicht untersucht werden. (Eigene Darstellung erstellt mit QGIS; © GeoBasis-DE / BKG 2022; © Statistisches Bundesamt 2022 b; © Ministerium für Energie, Infrastruktur und Landesentwicklung Mecklenburg-Vorpommern 2016)

Eine Überprüfung mit SQL ergibt die Anzahl der aufgrund der Anonymisierung der Pendelraten fehlenden Werte. Von den 674 untersuchten Gemeinden gibt es in 100 keine Angabe über die Anzahl der sozialversicherungspflichtigen Arbeitsplätze, in 35 keine Angabe über die Anzahl der Aus- und Binnenpendelnden und in 116 keine Angabe über die Anzahl der Einpendelnden.

2.4.2 Bestimmung der räumlichen Parameter

Im nächsten Schritt werden die räumlichen Werte der Gemeinden bestimmt, welche später als Einflussgrößen auf die Pendelraten untersucht werden. Die Flächengröße und Bevölkerungsdichte der Gemeinden sind eigentlich bereits durch das Gemeindeverzeichnis des Statistischen Bundesamtes vorhanden. Jedoch wird aufgrund fehlerhafter Einträge eine Neuberechnung erfolgen (siehe Abschnitt 2.4.2.3). Die anderen Werte werden im Folgenden bestimmt und der Datenbank-Tabelle hinzugefügt.

2.4.2.1 Küstenanteil (Topographische Lage)

Inseln und allgemein Gebiete mit Küstenstreifen haben laut Eurostat (2018) tendenziell einen geringeren Anteil an Pendelnden als Binnengebiete, da die Topographie erstgenannter Gebiete die Möglichkeiten des Pendelns beschränke – Flug- und Wassertransportmöglichkeiten sind weniger vorhanden bzw. können nur in geringerer Anzahl genutzt werden als Transportwege über Land. Die Verfügbarkeit von Brücken, welche Inseln mit dem Festland verbinden, können diesen Zusammenhang jedoch abschwächen.

Um die Abhängigkeit des Pendelverhaltens von der topographischen Lage der Gemeinden in Mecklenburg-Vorpommern zu untersuchen, gibt es mehrere Vorgehensweisen, letzteren Wert darzustellen. Möglich wäre z.B. eine binäre Unterscheidung der Gemeinden in Küsten- und Binnengemeinde, um anschließend zu bestimmen, ob sich die Pendelfaktoren zwischen diesen beiden Gruppen signifikant unterscheiden. Eine genauere Aussage über die Abhängigkeit der beiden Faktoren lässt sich jedoch treffen, wenn man die topographische Lage auf einer metrischen Skala darstellt. Dies geschieht hier durch den relativen Anteil der Verwaltungsgrenze einer Gemeinde, welcher eine Küste darstellt. Alle Binnengemeinden (und damit ein Großteil der untersuchten Gemeinden) werden hierbei einen Wert von 0 (0% Küste) erhalten, bei den Küstengemeinden jedoch wird es genauere Unterscheidungen geben. Bei diesem Verfahren wird nur das Meer als Küste berücksichtigt, nicht jedoch Binnengewässer, welche in Mecklenburg-Vorpommern in großer Zahl vorhanden sind und ebenfalls einen Einfluss auf die Pendeldynamik haben können.

Die Berechnung des Küstenanteils an der Verwaltungsgrenze einer Gemeinde erfolgt in mehreren Schritten mit QGS:

- 1.) Benutzung des Tools „Geometrien reparieren“ für den Gemeinde-Datensatz aus der Datenbank, um sicherzugehen, dass keine Geometriefehler vorhanden sind, welche bei späteren Verarbeitungs- oder Analyseschritten stören
- 2.) Berechnung des Umfangs der Gemeinden in Metern und Abspeichern dieses Wertes in einem neuen Attribut „umfang“ (Dezimalzahl) mit dem Feldrechner und der Funktion **\$perimeter**
- 3.) Benutzung des Tools „Grenze“, um die Umrisslinien der Flächen als Linien-Datensatz unter Beibehaltung der Sachdaten zu erstellen

Während die Gesamtlänge der Grenze jeder untersuchten Gemeinde damit bereits berechnet ist, muss noch die Länge des Küstenabschnitts dieser Grenze berechnet werden. Dazu wird ein räumlicher Vergleichsdatsatz benötigt, welcher die gesamte Landmasse in und um Mecklenburg-Vorpommern darstellt und die identische Küstenlinie wie der Gemeindedatensatz enthält. Dazu nutze ich den Datensatz „Kreis“ aus dem WFS-Dienst, aus welchem zuvor auch die Geometrien der Gemeinden genutzt wurden (Bundesamt für Kartographie und Geodäsie 2022). Dieser muss für das weitere Vorgehen dann noch um

Objekte entlang der polnischen Grenze ergänzt werden, damit der Grenzverlauf nicht im weiteren Verlauf als Küstenlinie interpretiert wird.

- 4.) Einladen des Datensatzes „Kreis“ aus dem WFS-Dienst und Herausfiltern der Gewässerflächen (siehe Abschnitt 2.3.1) mit dem Objektfilter „**gf**“ = 4
- 5.) Benutzung des Tools „Auflösen“ auf den Kreis-Datensatz, sodass alle Objekte zu einem aggregiert werden (Verwaltungsfläche von Deutschland)
- 6.) Erstellung zweier Flächen für den aufgelösten Datensatz entlang der Grenze von Mecklenburg-Vorpommern zu Polen (z.B. mit dem Werkzeug „Spurverfolgung“ aus der Einrastwerkzeugleiste) und anschließend erneutes „Auflösen“



Abbildung 2.14: Kennzeichnung der für die Berechnung des Küstenanteils zu digitalisierenden Flächen.
(Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022; © OpenStreetMap)

- 7.) Benutzung des Tools „Grenze“ für die aufgelöste Vergleichsfläche
- 8.) Umwandlung der Umrisslinie der Vergleichsfläche in eine Flächengeometrie durch Benutzung von „Puffer“ mit einer sehr kleinen Entfernung von 0,001 Metern
- 9.) „Zuschneiden“ der Grenzlinien der Gemeinden an der gepufferten Grenzlinie der Vergleichsfläche

Als Ergebnis erhält man einen Datensatz, welcher nur noch diejenigen Teile der Verwaltungsgrenze jeder untersuchten Gemeinde erhält, welche sich an der Ostseeküste befinden. 107 Gemeinden sind demnach als Küstengemeinde einzustufen. Nun wird noch die Länge der Küstengrenzen sowie deren Anteil an der Gesamtlänge der Verwaltungsgrenzen ermittelt.

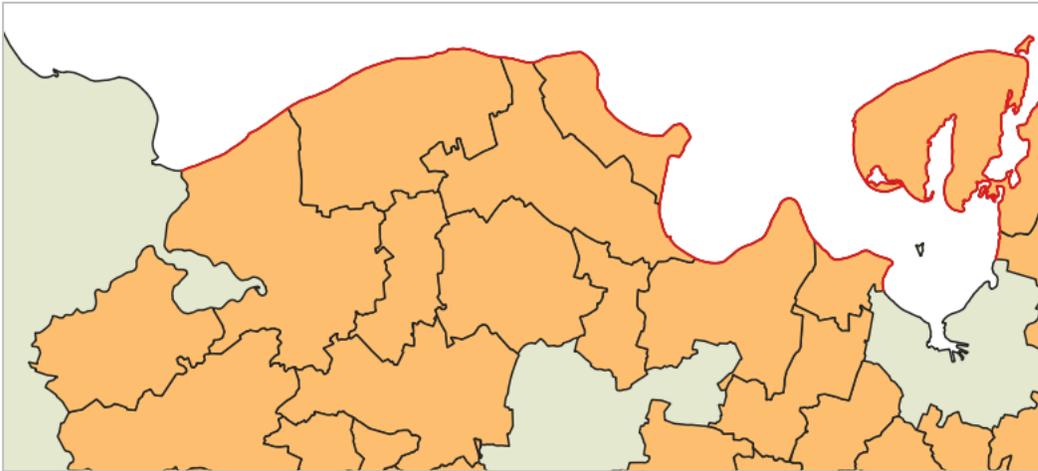


Abbildung 2.15: Die Küstengrenzen (rot) der zu untersuchenden Gemeinden (orange) (Ausschnitt).
(Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022)

- 10.) Berechnung der Länge der Küstenlinien in Metern und Abspeichern dieses Wertes in einem neuen Attribut „kueste“ (Dezimalzahl) mit dem Feldrechner und der Funktion **\$length**
- 11.) Anknüpfung der Küstenlinienlänge an den im 1. und 2. Schritt bearbeiteten Gemeindedatensatz mit dem Tool „Attribute nach Feldwert verknüpfen“
- 12.) Selektion aller Gemeinden ohne Küstenlinie mit der Attributabfrage „**kueste**“ is null und Eintragen der Zahl 0 für das Attribut in diesen mit dem Feldrechner
- 13.) Berechnung des Anteils der Länge der Küstengrenze an der Gesamtlänge der Verwaltungsgrenze und Abspeichern dieses Wertes in einem neuen Attribut „kueste_anteil“ (Dezimalzahl) mit dem Feldrechner und dem Ausdruck **100 * „kueste“ / „umfang“**
- 14.) Export des Verarbeitungsergebnisses in die Datenbank als neue Tabelle „gemeinden_temp_2“

2.4.2.2 Entfernung zu nächsten Oberzentrum (Entfernung zu Ballungsraum)

Agglomeration, Ballungsraum, Metropolregion – es gibt viele im groben Kern gleiche, sich aber in Nuancen unterscheidende Begriffe für Gebiete, welche aus einer Kernstadt (oder mehreren Kernstädten), ihren dicht bebauten Vororten sowie einem teilweise ländlichem Einzugsgebiet bestehen. Diese Gebiete bzw. insbesondere ihre Kernstädte ziehen viele Arbeitende aus den umliegenden Gemeinden an und beeinflussen damit ebenfalls die Anzahl der Pendelnden. Es soll untersucht werden, ob auch die Pendeldaten der umliegenden Gemeinden von der Nähe zu diesen Gebieten abhängig ist.

Zunächst muss definiert werden, was im Rahmen dieser Untersuchung als Äquivalent zu den Ballungsräumen betrachtet wird. Da in einer deutlich kleineren Ebene als in Eurostat (2018) geforscht wird, sollen auch die Zielorte kleindimensionierter sein als die dort

genutzten Ballungsräume. Ich habe mich für die folgenden acht Oberzentren in und um Mecklenburg-Vorpommern als mögliche nächste Zielorte der Gemeinden entschieden:

Stadt	Land / Staat	Einwohnende (ungefähr)	Distanz zu Mecklenburg-Vorpommern (ungefähr)
Schwerin	Mecklenburg-Vorpommern	95.000	-
Rostock	Mecklenburg-Vorpommern	210.000	-
Neubrandenburg	Mecklenburg-Vorpommern	65.000	-
Stralsund	Mecklenburg-Vorpommern	60.000	-
Greifswald	Mecklenburg-Vorpommern	60.000	-
Lübeck	Schleswig-Holstein	215.000	direkt an der Grenze
Lüneburg	Niedersachsen	75.000	15 km
Stettin	Polen	395.000	5 km

Tabelle 2.2: Die möglichen Zielorte zur Entfernungsmessung für die Gemeinden Mecklenburg-Vorpommerns. (Bevölkerungszahlen von Statistisches Bundesamt (2022 b), Brinkhoff (2021))

Neben den fünf Oberzentren Mecklenburg-Vorpommerns sollen damit auch drei Oberzentren nahe der Landesgrenze betrachtet werden. Weitere Orte außerhalb Mecklenburg-Vorpommerns waren entweder zu weit entfernt von der Landesgrenze, um für eine Gemeinde das nächste Oberzentrum zu sein (z.B. Hamburg oder Berlin) oder haben im Vergleich zu den anderen gewählten Orten eine viel geringere Anzahl an Einwohnenden.

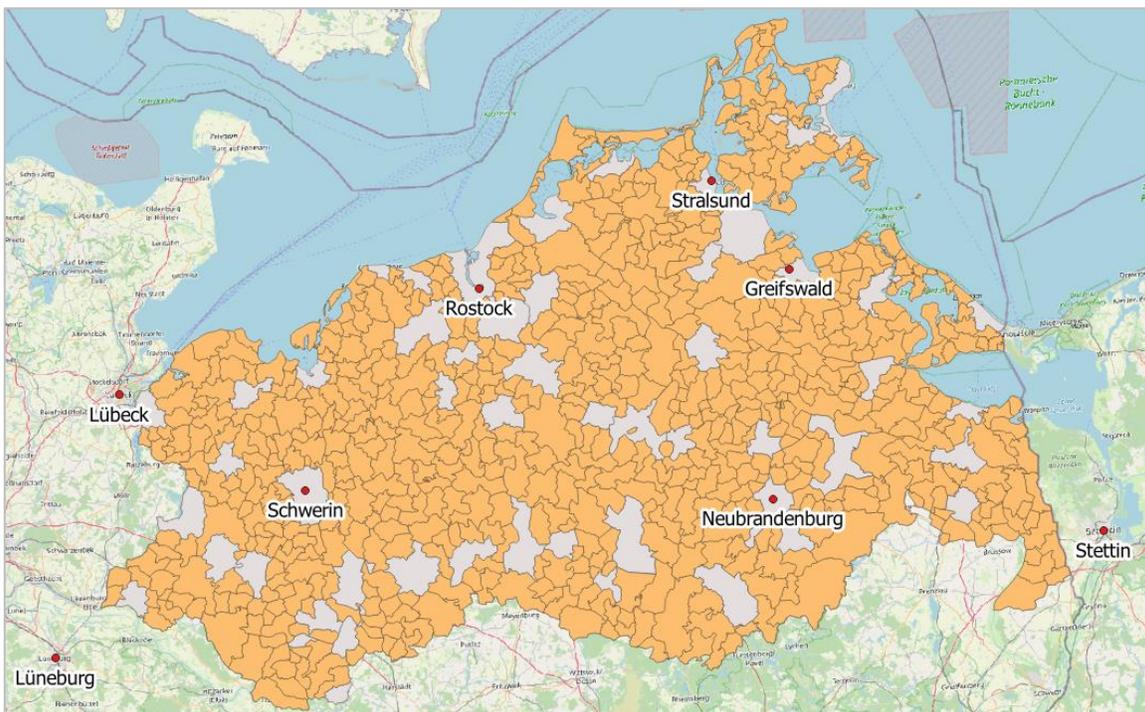


Abbildung 2.16: Lage der möglichen Zielorte zur Entfernungsmessung für die Gemeinden Mecklenburg-Vorpommerns. (Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022; © OpenStreetMap)

Zur Bestimmung der Entfernung einer jeden Gemeinde zu den acht ausgewählten Zielorten gibt es unterschiedliche Methoden. Ich habe mich gegen die effektive Distanz als Strecke oder Zeit über ein Straßennetz entschieden, da diese sich auf ein mögliches Verkehrsmittel fokussiert. Stattdessen wird die euklidische Distanz verwendet. Die Verwaltungseinheiten werden dabei als Punkt betrachtet, welche ich sowohl für die Gemeinden Mecklenburg-Vorpommerns als auch für die Zielorte aus dem Datensatz „Gemeindepunkte“ des WFS-Datensatzes vom Bundesamt für Kartographie und Geodäsie (2022) übernommen habe. Nur Stettin in Polen musste händisch als Punkt ergänzt werden.

Zur Durchführung wird in QGIS das Tool „Abstand zum nächsten Knoten (Punkte)“ mit den Gemeindepunkten als Quelllayer und den Zielortpunkten als Ziellayer benutzt. Damit werden Geometrien und Attributtabelle der Gemeindepunkte übernommen und um zwei Attribute ergänzt, der Kennung des nächsten Zielpunktes sowie der Distanz zu diesem. Diese Attribute können nun mit dem Tool „Attribute nach Feldwert verknüpfen“ dem Gemeindedatensatz angefügt werden und der ausgegebene Datensatz wird als neue Tabelle „gemeinden_final“ in die Datenbank importiert.

2.4.2.3 Shape-Index

Der Shape-Index ist eine Maßzahl zur Formbeschreibung aus der Landschaftsanalyse und beschreibt die Komplexität des Umrisses einer Fläche im Vergleich zu einer Kreisform, welche einen Shape-Index von 1 hat (Blaschke und Lang 2007). Dieser Index wird zur Formbeschreibung der untersuchten Gemeinden verwendet, da diese einen Einfluss auf die Pendelzahlen haben kann (siehe Abbildung 1.1). Der Shape-Index entspricht dem Quotienten des Umfangs einer Fläche und dem Doppelten der Quadratwurzel des Produktes von Flächengröße und der Kreiszahl π . Da der Umfang in Metern sowie die Flächengröße in Quadratkilometern bereits als Werte bei den Gemeinden vorhanden sind, lässt sich der Shape-Index mit SQL berechnen.

```
ALTER TABLE gemeinden_final
ADD COLUMN shapeindex double precision;

UPDATE gemeinden_final
SET shapeindex = umfang/(2*sqrt(flaeche_km*1000000*pi()));
```

SQL 2.4: Hinzufügen einer Spalte für den Shape-Index und Berechnung der Werte.

Eine Überprüfung der Werte bei den Gemeinden Mecklenburg-Vorpommerns hat bei zwei Gemeinden einen Wert unter 1 ergeben, was beim Shape-Index nicht möglich ist. Da der Umfang selbst mit QGIS auf Grundlage der Geometrie berechnet wurde, muss die Ursache hier bei der Flächengröße liegen, welche aus dem Gemeindeverzeichnis des Statistischen Bundesamtes stammt. Tatsächlich zeigt ein testweises Berechnen der Flächengröße auf

Grundlage der Geometrie mit QGIS Abweichungen zu den Flächenwerten des Statistischen Bundesamtes, bei einem Großteil der Gemeinden „nur“ bis zu einem Quadratkilometer, bei fünf Gemeinden jedoch über zehn Quadratkilometern, wobei die Angabe im Gemeindeverzeichnis stets die größere war. Woher die Flächenangaben des Statistischen Bundesamtes stammen, konnte nicht ermittelt werden. Für diese Untersuchung werden mit denen von QGIS bzw. PostGIS berechneten Flächengrößen weiter gearbeitet. Deshalb werden die Attribute mit der Flächengröße (und der Bevölkerungsdichte, welche durch Typumwandlung nun auch mit Nachkommastellen gespeichert wird) entsprechend mit SQL aktualisiert und danach die Update-Anweisung aus SQL 2.4 wiederholt.

```
UPDATE gemeinden_final
SET flaeche_km = ST_Area(geom)/1000000;

UPDATE gemeinden_final
SET bev_dichte_km = bevoelkerung::real / flaeche_km;
```

SQL 2.5: Aktualisierung von Flächengröße und Bevölkerungsdichte.

2.4.3 Bestimmung der Pendelparameter

Die Pendelwerte und Pendelströme des Statistik-Service der Bundesagentur für Arbeit werden zur Berechnung der Pendelparametern verwendet, welche im Folgenden vorgestellt und der Datenbanktabelle als weitere Spalten hinzugefügt werden.

2.4.3.1 Ein- und Auspendelquote

Die Anzahl der Ein- und Auspendelnden muss für eine Auswertung in Relation zu den sozialversicherungspflichtig Beschäftigten am Arbeits- bzw. Wohnort gesetzt werden. Die Einpendelquote gibt über den Anteil der von sozialversicherungspflichtig Beschäftigten bezogenen Arbeitsplätze einer Gemeinde Auskunft, welcher von Einpendelnden in diese Gemeinde belegt wird. Die Auspendelquote gibt den Anteil der sozialversicherungspflichtig Beschäftigten mit Wohnort in einer Gemeinde an, welche in eine andere Gemeinde zur Arbeit auspendelt. Bei beiden Werten ist es üblich, diese als Prozentsatz anzugeben.

Damit diese mit Nachkommastellen gespeichert werden, muss bei der Berechnung mit SQL darauf geachtet werden, Typumwandlungen der Ganzzahlen in Dezimalzahlen vorzunehmen. Um keine Division mit der Zahl Null zu erhalten, welche insbesondere bei der Zahl der sozialversicherungspflichtig Beschäftigten am Arbeitsort auftreten kann, wird diese explizit ausgeschlossen, entsprechende Ein- bzw. Auspendelquoten bleiben leer.

```

ALTER TABLE gemeinden_final
ADD COLUMN einpendelquote real;

ALTER TABLE gemeinden_final
ADD COLUMN auspendelquote real;

UPDATE gemeinden_final
SET einpendelquote = 100*einpendelnde::real / arbeitsort::real
WHERE arbeitsort > 0;

UPDATE gemeinden_final
SET auspendelquote = 100*auspendelnde::real / wohnort::real
WHERE wohnort > 0;

```

SQL 2.6: Hinzufügen von Spalten für Ein- und Auspendelquote und Berechnung der Werte.

Mit der Auspendelquote ist auch eine Aussage über die Binnenpendelquote möglich, womit der Anteil der sozialversicherungspflichtig Beschäftigten mit Wohnort in einer Gemeinde beschrieben wird, welcher in dieser Gemeinde ihre Arbeitsstelle haben. Die Binnenpendelquote wird aus der Differenz der Auspendelquote zu 100% berechnet.

2.4.3.2 Arbeitsplatzausstattung

Ein weiterer wichtiger Wert ist die Arbeitsplatzausstattung einer Gemeinde. Mit dieser wird angegeben, ob innerhalb einer Gemeinde theoretisch ausreichend (besetzte) Arbeitsplätze für alle sozialversicherungspflichtig Beschäftigten, welche in dieser wohnen, zur Verfügung stehen. Dazu wird der Quotient von der Anzahl der sozialversicherungspflichtigen Arbeitsplätze einer Gemeinde an der Anzahl der sozialversicherungspflichtig Beschäftigten mit Wohnort in dieser Gemeinde gebildet. Die Arbeitsplatzausstattung kann bei über 100% liegen, dann gibt es in der Gemeinde mehr Arbeitsplätze als Beschäftigte und sie ist auf Einpendelnde angewiesen. Liegt die Arbeitsplatzausstattung unter 100%, muss es Auspendelnde aus der Gemeinde geben.

Bei der Berechnung mit SQL müssen die identischen Hinweise wie bei der Berechnung von Ein- und Auspendelquote beachtet werden.

```

ALTER TABLE gemeinden_final
ADD COLUMN arbeitsplatzausstattung real;

UPDATE gemeinden_final
SET arbeitsplatzausstattung = 100*arbeitsort::real / wohnort::real
WHERE wohnort > 0;

```

SQL 2.7: Hinzufügen einer Spalte für die Arbeitsplatzausstattung und Berechnung der Werte.

2.4.3.3 Independence-Index

Der Independence-Index wurde von Thomas (1969) als Indikator für die Abhängigkeit einer Verwaltungseinheit von externen Pendelnden entwickelt. Er nutze diesen Indikator (neben weiteren) zur Untersuchung, ob die britischen Städte der 1950er- und 1960er - Jahre das Prinzip „self-containment“ (übersetzt in etwa „Selbstbeherrschung“) erfüllen - die Einwohnenden einer Verwaltungseinheit können in dieser u.A. leben, arbeiten und einkaufen und müssen dementsprechend für keine dieser Aufgaben sich in andere Verwaltungseinheiten begeben. Der Independence-Index hat sich als Indikator in der nationalen wie internationalen Forschung etabliert (vgl. z.B. Cervero 1996, Siedentop 2007).

Berechnet wird dieser Wert als Quotient aus der Anzahl der Binnenpendelnden und der Summe der Ein- und Auspendelnden einer Verwaltungseinheit. Während die Binnenpendelquote zumeist nur als Differenz der Auspendelquote zu 100% angegeben wird, ist der Independence-Index damit auch von der Anzahl der Einpendelnden abhängig. Der Wert beträgt 1, wenn es in einer Verwaltungseinheit gleich viele Binnenpendelnde wie Ein- und Auspendelnde gibt, und kann diesen Wert auch übersteigen, was jedoch zumeist nur in Kernstädten der Fall ist.

Auch hier müssen bei der Berechnung mit SQL die bereits genannten Hinweise berücksichtigt werden. Da es in einigen Gemeinden keine Angabe über die Anzahl der Einpendelnden gibt, würde in diesen auch der Independence-Index ungefüllt bleiben. Um mehr Ergebnisse zu erhalten, lasse ich bei diesen Gemeinden die Anzahl der Einpendelnden unberücksichtigt (es wird angenommen, es gebe keine Einpendelnden) und berechne den Independence-Index als Quotienten aus der Anzahl der Binnenpendelnden und der Anzahl der Auspendelnden. Diese Werte fehlen in deutlich weniger Gemeinden.

```
ALTER TABLE gemeinden_final
ADD COLUMN independence real;

UPDATE gemeinden_final
SET independence = binnenpendelnde::real / (einpendelnde::real + auspendelnde::real)
WHERE einpendelnde + auspendelnde > 0;

UPDATE gemeinden_final
SET independence = binnenpendelnde::real / auspendelnde::real
WHERE auspendelnde > 0 AND einpendelnde IS NULL;
```

SQL 2.8: Hinzufügen einer Spalte für den Independence-Index und Berechnung der Werte, die 2. Update-Anweisung beinhaltet die alternative Berechnung ohne Berücksichtigung der Einpendelnden für die Gemeinden, in denen deren Anzahl unbekannt ist.

2.4.3.4 Ein- und Auspendeldistanz

Zur Bestimmung der durchschnittlichen Pendeldistanz wird auf die Pendelströme zurückgegriffen. Um aus diesen die durchschnittliche zurückgelegte Distanz der Ein- bzw. Auspendelnden je Gemeinde bestimmen zu können, wird zunächst die Länge aller vorhandenen Pendelströme bestimmt. Hierbei können zunächst noch die Zeilen in den Tabellen der Ein- und Auspendelströme gelöscht werden, in welchen keine der beiden Gemeinden zu den untersuchten gehören. Dazu erfolgt ein Abgleich der Verwaltungsschlüsselnummern mit denen in „gemeinden_final“.

```
DELETE FROM einpendelstroeme
WHERE vsn_arbeit NOT IN (SELECT ags FROM gemeinden_final)
AND vsn_wohn NOT IN (SELECT ags FROM gemeinden_final);

DELETE FROM auspendelstroeme
WHERE vsn_arbeit NOT IN (SELECT ags FROM gemeinden_final)
AND vsn_wohn NOT IN (SELECT ags FROM gemeinden_final);
```

SQL 2.9: Löschen der Pendelströme, in denen keine der beiden Gemeinden zu den untersuchten gehören.

Analog zu der Bestimmung der Entfernung der Gemeinden zu dem nächstgelegenen Oberzentrum als räumlicher Einflussfaktor (siehe Abschnitt 2.4.2.2) werde ich dabei die euklidische Distanz zwischen den Gemeindepunkten des WFS-Datensatzes vom Bundesamt für Kartographie und Geodäsie verwenden. Hier erfolgt die Distanzbestimmung per SQL. In den Tabellen mit den Pendelströmen, welche bislang keine Geometrien enthalten, werden Liniengeometrien erstellt, welche als einzelnes Liniensegment zwischen den beiden Gemeinden eines Pendelstroms definiert werden. Deren Länge wird im Anschluss ebenfalls berechnet. Dazu werden die PostGIS-Funktionen ST_MakeLine und ST_Length verwendet.

```
ALTER TABLE auspendelstroeme
ADD COLUMN geom geometry(Linestring,25832);

UPDATE auspendelstroeme
SET geom = ST_MakeLine(
(SELECT geom FROM gemeindepunkte a WHERE a.ags=auspendelstroeme.vsn_wohn),
(SELECT geom FROM gemeindepunkte a WHERE a.ags=auspendelstroeme.vsn_arbeit)
);

ALTER TABLE auspendelstroeme
ADD COLUMN distanz real;
UPDATE auspendelstroeme
SET distanz = ST_Length(geom)/1000;
```

SQL 2.10: Hinzufügen der Spalten für die Linien-Geometrien der Auspendelströme sowie deren Länge und Definition der Geometrie sowie Berechnung der Werte (analoges Vorgehen für die Einpendelströme).

Die ermittelten Entfernungen jedes Pendelstroms werden mit der Anzahl der Pendelnden dieses Stroms multipliziert, um in einem weiteren Schritt diese Produkte pro Gemeinde zu addieren und durch die Gesamtanzahl der Aus- bzw. Einpendelnden dieser Gemeinde zu dividieren. Damit erhält man für jede Gemeinde die durchschnittliche Distanz, die Ein- bzw. Auspendelnde zurücklegen.

Für die Berechnung des 2. Schritts mit SQL müssen die Tabellen der Aus- bzw. Einpendelströme und der Gemeinden miteinander verknüpft werden. Dabei muss darauf geachtet werden, die korrekte Verwaltungsschlüsselnummer in den Tabellen der Pendelströme zu verwenden. Bei den Auspendelströmen ist es diejenige der Gemeinden, aus denen heraus gependelt wird („vsn_wohn“) und bei den Einpendelströmen diejenige der Gemeinden, in welche hinein gependelt wird („vsn_arbeit“).

```
ALTER TABLE auspendelstroeme
ADD COLUMN produkt real;

UPDATE auspendelstroeme
SET produkt = pendelnde * distanz;

ALTER TABLE gemeinden_final
ADD COLUMN auspendeldistanz real;

UPDATE gemeinden_final
SET auspendeldistanz =
(SELECT SUM(produkt) / SUM(pendelnde) FROM auspendelstroeme a
WHERE a.vsn_wohn = gemeinden_final.ags
);
```

SQL 2.11: Hinzufügen der Spalte für das Produkt aus Pendelnden und Distanz jedes Auspendelstroms sowie der daraus resultierenden durchschnittlichen Auspendeldistanz für jede Gemeinde mit jeweiliger Berechnung (analoges Vorgehen für die Einpendelströme).

2.4.4 Bestimmung der statistischen Maße und Korrelationen

Alle betrachteten räumlichen Einflussfaktoren und Pendelparameter sind nun bestimmt, dadurch kann im letzten Schritt die statistische Auswertung erfolgen. Für alle Indikatoren und Parameter sollen zunächst die folgenden deskriptiven statistischen Maße ermittelt werden:

- Anzahl (aufgrund der Anonymisierungen in den Pendeldaten sind nicht alle Pendelparameter für jede Gemeinde vorhanden)
- Minimum, Maximum und Mittelwert
- 1., 2. und 3. Quartil
- Varianz und Standardabweichung
- Schiefe und Wölbung

Die Berechnung der Maße wird mit SQL durchgeführt. Da es für die Schiefe und Wölbung keine in PostgreSQL bereits vordefinierte Funktionen zur Berechnung gibt, müssen diese zunächst manuell definiert werden. Dafür wird Code von Ellison (2018) verwendet, welcher die Funktion „stats_agg“ definiert, die alle Werte einer Spalte aggregiert und für diese die Anzahl (ohne Leerwerte), Minimum, Maximum, Mittelwert, Varianz, Schiefe und Wölbung berechnet (siehe SQL 2.12). Dabei werden die Definitionen der Empirische Schiefe und der Empirische Wölbung verwendet (vgl. z.B. Becker 2022). Hinzu kommen noch die PostgreSQL-Funktionen „stddev_pop“ für die empirische Standardabweichung und „percentile_disc“ zur Berechnung der Quartile (siehe SQL 2.13). Zur Darstellung der Verteilung von Indikatoren und Pendelparametern in jeweils einem Histogramm wird das QGIS-Plugin „Data Plotly“ von Ghetta (2022) verwendet.

Die Zusammenhänge zwischen Indikatoren und Pendelparametern werden mit der Pearson-Produkt-Moment-Korrelation untersucht. Diese untersucht zwei Variablen ausschließlich auf eine lineare Abhängigkeit und wird stark von Ausreißern beeinflusst (vgl. z.B. Schäfer 2016). Aus diesem Grund werden die Variablen nach einer ersten Berechnung der Korrelationen visuell in Streudiagrammen, welche ebenfalls mit dem QGIS-Plugin „Data Plotly“ erstellt werden, auf Ausreißer untersucht und diese, falls vorhanden, bei einer zweiten Berechnung ausgeschlossen. Die Verwendung nicht-parametrischer Verfahren bei mangelnder Linearität ist nicht vorgesehen. Zur Berechnung der Korrelation wird in PostgreSQL die Funktion „corr“ verwendet (siehe SQL 2.13). Es werden insbesondere die Zusammenhänge von jeweils einem Indikator (Fläche, Bevölkerungsdichte, Küstenanteil, Entfernung zu Oberzentrum, Shape-Index) zu jeweils einem Pendelparameter (Ein- und Auspendelquote sowie -distanz, Arbeitsplatzausstattung und Independence-Index) untersucht, hinzu kommen weitere Zusammenhänge einzelner Pendelparameter untereinander.

```

-- 1
create type _stats_agg_accum_type AS (
  n bigint,
  min double precision,
  max double precision,
  m1 double precision,
  m2 double precision,
  m3 double precision,
  m4 double precision
);

-- 2
create type _stats_agg_result_type AS (
  anzahl bigint,
  min double precision,
  max double precision,
  mittelwert double precision,
  varianz double precision,
  schiefe double precision,
  woelbung double precision
);

-- 3
create or replace function
_stats_agg_accumulator(_stats_agg_accum_type, double precision)
returns _stats_agg_accum_type AS '
DECLARE
  a ALIAS FOR $1;
  x alias for $2;
  n1 bigint;
  delta double precision;
  delta_n double precision;
  delta_n2 double precision;
  term1 double precision;
BEGIN
  if x IS NOT NULL then
    n1 = a.n;
    a.n = a.n + 1;
    delta = x - a.m1;
    delta_n = delta / a.n;
    delta_n2 = delta_n * delta_n;
    term1 = delta * delta_n * n1;
    a.m1 = a.m1 + delta_n;
    a.m4 = a.m4 + term1 * delta_n2
    * (a.n*a.n - 3*a.n + 3) + 6 * delta_n2 * a.m2 - 4 *
    delta_n * a.m3;
    a.m3 = a.m3 + term1 * delta_n *
    (a.n - 2) - 3 * delta_n * a.m2;
    a.m2 = a.m2 + term1;
    a.min = least(a.min, x);
    a.max = greatest(a.max, x);
  end if;
  RETURN a;
END;
'language plpgsql;

-- 4
create or replace function
_stats_agg_finalizer(_stats_agg_accum_type)
returns _stats_agg_result_type AS '
BEGIN
  RETURN row(
    $1.n,
    $1.min,
    $1.max,
    case when $1.n = 0 then null
  else $1.m1 end,
    case when $1.n = 0 then null
  else $1.m2 / nullif(($1.n - 1.0), 0) end,
    case when $1.m2 = 0 then
  null else sqrt($1.n) * $1.m3 / nullif(($1.m2 ^
  1.5), 0) end,
    case when $1.m2 = 0 then
  null else $1.n * $1.m4 / nullif(($1.m2 *
  $1.m2) - 3.0, 0) end
  );
END;
'language plpgsql;

-- 5
create aggregate stats_agg(double
precision) (
  sfunc = _stats_agg_accumulator,
  stype = _stats_agg_accum_type,
  finalfunc = _stats_agg_finalizer,
  initcond = '(0,, 0, 0, 0, 0)'
);

```

SQL 2.12: Definition der Funktion „stats_agg“ zur Berechnung von u.A. Schiefe und Wölbung der Variablen.
© Ellison (2018) [bearbeitete Ausgabenamen im 2. Schritt]

```

SELECT
(stats_agg (einpendelquote)).*,
stddev_pop (einpendelquote) AS standardabweichung,
percentile_disc (0.25) within group (order by einpendelquote) AS quartil_1,
percentile_disc (0.5) within group (order by einpendelquote) AS median,
percentile_disc (0.75) within group (order by einpendelquote) AS quartil_3
FROM gemeinden_final;

SELECT corr (shapeindex, auspendelquote) FROM gemeinden_final;

```

SQL 2.13: Bestimmung der statistischen Maße am Beispiel der Einpendelquote sowie Bestimmung der Korrelation am Beispiel von Shape-Index und Auspendelquote.

2.4.5 Anmerkungen zum Vorgehen in Niedersachsen

2.4.5.1 Kompatibilität der Datenbestände

Bei den Gemeinden Niedersachsens gibt es eine unterschiedliche Anzahl an Zeilen und damit Gemeinden in den importierten Gemeindedaten. Während die Datenbestände aus dem WFS-Dienst des Bundesamt für Kartographie und Geodäsie (Geometrien) sowie aus dem Gemeindeverzeichnis des Statistischen Bundesamtes (Einwohnerzahlen) übereinstimmen (964 Zeilen, alle Verwaltungsschlüsselnummern sind identisch), gibt es in der Tabelle mit den Pendelwerten des Statistik-Service der Bundesagentur für Arbeit nur 944 Zeilen. Ein Abgleich ergibt, dass es bei den Geometrien 24 Verwaltungsschlüsselnummern gibt, welche nicht in den Pendeldaten vorhanden sind, umgekehrt gibt es bei diesen vier Verwaltungsschlüsselnummern, welche bei den Geometrien fehlen.

In einem solchen Fall der nicht vollständigen Kompatibilität ist es notwendig, vor der Verknüpfung die Gründe dafür zu recherchieren, um die Tabellen entsprechend manuell anpassen zu können. Im vorliegenden Fall gibt es zwei Ursachen:

- 23 der 24 Verwaltungsschlüsselnummern, welche nicht in den Pendeldaten vorhanden sind, gehören zu sogenannten gemeindefreien Gebieten. Diese sind unbewohnt, auch im Gemeindeverzeichnis des Statistischen Bundesamtes ist ihre Bevölkerungszahl mit 0 angegeben. Damit können diese nicht in den Pendeldaten vertreten sein, sind für die Master-Thesis auch nicht relevant und werden bei der Verknüpfung der Tabellen automatisch herausgefiltert.
- Die andere Inkompatibilität hat ihre Ursache in einer Gebietsreform, welche in den Pendeldaten noch nicht berücksichtigt wurde. Die vier Gemeinden, welche zusätzlich in den Pendeldaten vorhanden sind, wurden zu einer Gemeinde zusammengefasst, welche eine neue Verwaltungsschlüsselnummer bekommen hat.

Aufgrund einer Gebietsreform dieser Art kann es notwendig sein, die Pendelraten manuell anpassen zu müssen. Da die entstandene Gemeinde in diesem Fall eine Bevölkerungszahl von über 5000 hat, fällt sie jedoch aus dem Untersuchungsbestand heraus. Deshalb ist bei den Pendelraten keine Anpassung notwendig. Die Pendelströme jedoch müssen dahingehend aktualisiert werden, da ansonsten z.B. Auspendelströme aus einer untersuchten Gemeinde in die reformierte Gemeinde nicht berücksichtigt werden würden. Zur Aktualisierung der Datenbestände ist eine Anpassung der Verwaltungsschlüsselnummern und möglicherweise eine Aufsummierung von Pendelströmen einer identischen untersuchten Gemeinde zu mehreren der ehemaligen Gemeinden notwendig.

Die SQL-Befehle zur Überarbeitung der Einpendelströme sind in SQL 2.14 beschrieben und können für die Auspendelströme mit Anpassung der Spaltenreihenfolge übernommen werden. Aufgrund des spezifischen Falls genau für den von mir verwendeten Datenstand in Niedersachsen und die etwas komplexere Vorgehensweise soll auf diese SQL-Befehle jedoch nicht weiter eingegangen werden.

```

INSERT INTO einpendelstroeme (
SELECT '03153019' AS vsn_arbeit, c.vsn_wohn, SUM(c.pendelnde) FROM
(SELECT DISTINCT a.* FROM einpendelstroeme a, einpendelstroeme b
WHERE a.vsn_wohn = b.vsn_wohn
AND NOT a.vsn_wohn IN ('03153006', '03153007', '03153009', '03153014')
AND NOT a.vsn_arbeit = b.vsn_arbeit
AND a.vsn_arbeit IN ('03153006', '03153007', '03153009', '03153014')
AND b.vsn_arbeit IN ('03153006', '03153007', '03153009', '03153014')) c
GROUP BY c.vsn_wohn);

INSERT INTO einpendelstroeme (
SELECT c.vsn_arbeit, '03153019' AS vsn_wohn, SUM(c.pendelnde) FROM
(SELECT DISTINCT a.* FROM einpendelstroeme a, einpendelstroeme b
WHERE a.vsn_arbeit = b.vsn_arbeit
AND NOT a.vsn_arbeit IN ('03153006', '03153007', '03153009', '03153014')
AND not a.vsn_wohn = b.vsn_wohn
AND a.vsn_wohn IN ('03153006', '03153007', '03153009', '03153014')
AND b.vsn_wohn IN ('03153006', '03153007', '03153009', '03153014')) c
GROUP BY c.vsn_arbeit);

DELETE FROM einpendelstroeme
WHERE vsn_arbeit IN ('03153006', '03153007', '03153009', '03153014')
OR vsn_wohn IN ('03153006', '03153007', '03153009', '03153014');

```

SQL 2.14: Aufgrund der in den Pendelraten nicht berücksichtigten Gemeindeform notwendige Überarbeitung der Einpendelströme.

Nach der Herstellung der Kompatibilität können die drei Tabellen wie im Abschnitt 2.4.1 beschrieben miteinander verknüpft werden. 603 der ursprünglich 941 Gemeinden haben eine Bevölkerungszahl unter 5000, womit bei den Gemeinden Niedersachsens absolut und insbesondere relativ gesehen weniger Gemeinden untersucht werden als in Mecklenburg-Vorpommern. Auch gibt es Landkreise, in welchen gar keine Gemeinde untersucht wird.

Erneut sind aufgrund der vorkommenden Anonymisierungen in den Pendelraten gewisse Pendelwerte nicht für alle Gemeinden vorhanden. Von den 603 untersuchten Gemeinden gibt es in 31 keine Angabe über die Anzahl der sozialversicherungspflichtigen Arbeitsplätze, in 16 keine Angabe über die Anzahl der Aus- und Binnenpendelnden und in 33 keine Angabe über die Einpendelnden.

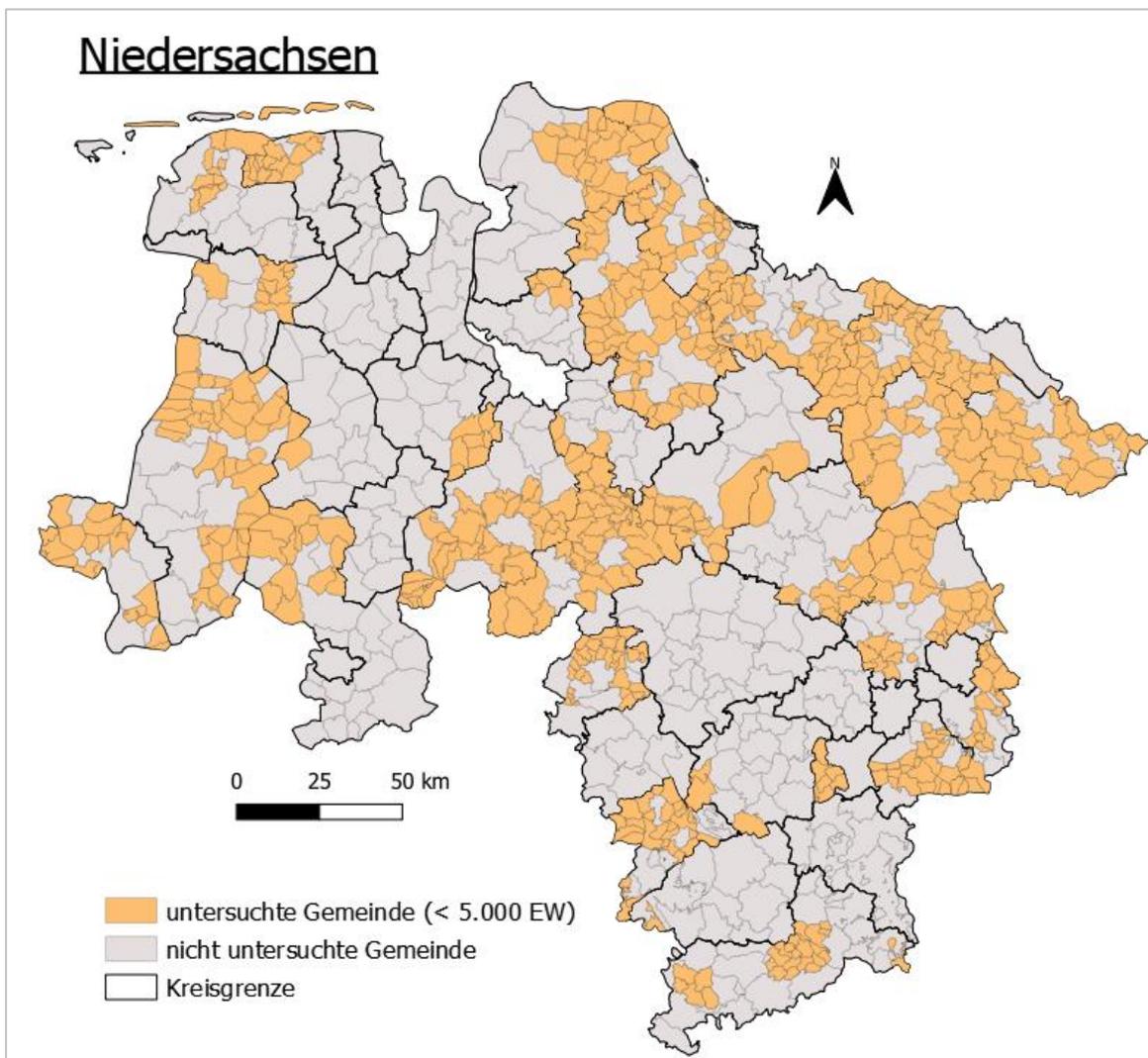


Abbildung 2.17: Überblick über die Gemeinden Niedersachsen, welche nicht untersucht werden.
(Eigene Darstellung erstellt mit QGIS; © GeoBasis-DE / BKG 2022; © Statistisches Bundesamt 2022 b)

2.4.5.2 Räumliche Parameter

Bei der Bestimmung des Küstenanteils muss analog zum Vorgehen bei den Gemeinden Mecklenburg-Vorpommerns auch mit selbst erstellten Flächen gearbeitet werden, in diesem Fall an der Grenze von Niedersachsen zu den Niederlanden.

Für die Bestimmung der Entfernung zum nächsten Oberzentrum werden folgende Orte in die Untersuchung einbezogen:

Stadt	Land / Staat	Einwohnende (ungefähr)	Distanz zu Niedersachsen (ungefähr)
Hannover	Niedersachsen	535.000	-
Braunschweig	Niedersachsen	250.000	-
Oldenburg	Niedersachsen	170.000	-
Osnabrück	Niedersachsen	165.000	-
Wolfsburg	Niedersachsen	125.000	-
Göttingen	Niedersachsen	115.000	-
Salzgitter	Niedersachsen	105.000	-
Hildesheim	Niedersachsen	100.000	-
Delmenhorst	Niedersachsen	80.000	-
Lüneburg	Niedersachsen	75.000	-
Wilhelmshaven	Niedersachsen	75.000	-
Celle	Niedersachsen	70.000	-
Hamburg	Hamburg	1.855.000	direkt an der Grenze
Bremen	Bremen	565.000	direkt an der Grenze
Bremerhaven	Bremen	115.000	direkt an der Grenze
Minden	Nordrhein-Westfalen	80.000	direkt an der Grenze
Detmold	Nordrhein-Westfalen	75.000	20 km
Rheine	Nordrhein-Westfalen	75.000	direkt an der Grenze
Groningen	Niederlande	240.000	35 km
Emmen	Niederlande	110.000	10 km
Almelo	Niederlande	55.000	10 km

Tabelle 2.3: Die möglichen Zielorte zur Entfernungsmessung für die Gemeinden Niedersachsens.
(Bevölkerungszahlen von Statistisches Bundesamt (2022 b), Brinkhoff (2023))

Damit sind alle Oberzentren von Niedersachsen sowie Delmenhorst als einziges Mittelzentrum, welches eine ähnlich große Bevölkerungszahl aufweist, als Zielort vorhanden. Hinzu kommen jeweils drei Städte aus den Niederlanden und Nordrhein-Westfalen sowie Hamburg, Bremen und Bremerhaven. Weitere von der Bevölkerungszahl ähnlich große Städte nahe an der Grenze von Niedersachsen sind nicht in der Auswahl vorhanden, da sie zu keiner Gemeinde des Landes die geringste Distanz vorweisen.

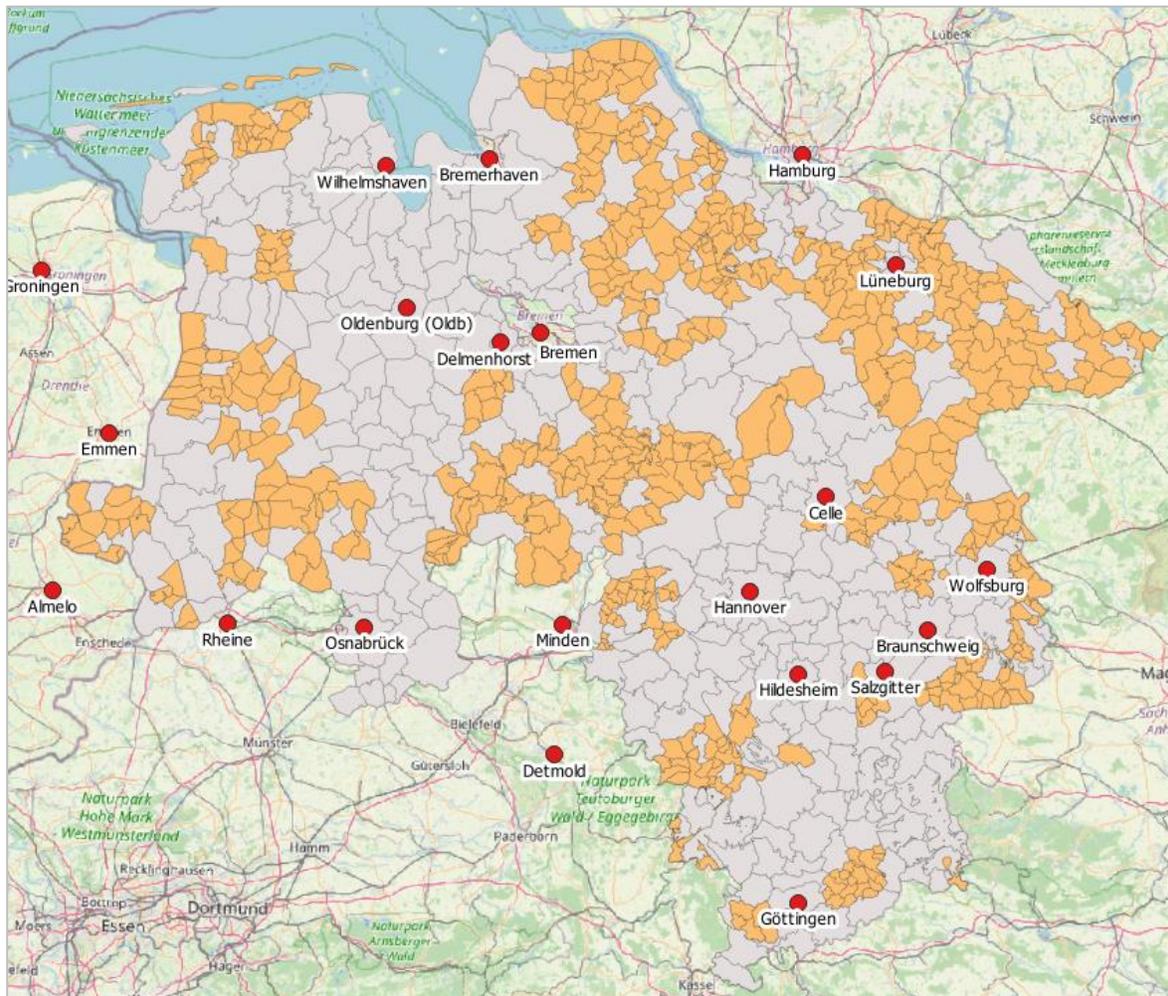


Abbildung 2.18: Lage der möglichen Zielorte zur Entfernungsmessung für die Gemeinden Niedersachsens. (Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022; © OpenStreetMap)

3. Ergebnisse

3.1 Deskriptive statistische Maße der Indikatoren

In Mecklenburg-Vorpommern wurden 674 Gemeinden untersucht, in Niedersachsen 603 Gemeinden. Die Indikatoren sind prinzipiell für alle Gemeinden vorhanden, jedoch muss bei der Auswertung beachtet werden, dass der Küstenanteil bei allen Binnengemeinden und damit dem Großteil der untersuchten Gemeinden 0% annimmt. In Mecklenburg-Vorpommern betrifft dies mit 567 Gemeinden etwa 84% der untersuchten Gemeinden, in Niedersachsen ist der Anteil mit etwa 97% (586 Gemeinden) noch höher. Viele Küstengemeinden in Niedersachsen wurden mit einer Bevölkerungszahl von mind. 5000 aus der Untersuchung ausgeschlossen, während in Mecklenburg-Vorpommern fast alle Küstengemeinden eine geringere Bevölkerungszahl aufweisen (vgl. Abbildungen 2.16 und 2.18). Auch abseits der Binnengemeinden fällt die Verteilung des Indikators in Mecklenburg-Vorpommern ausgewogener aus.

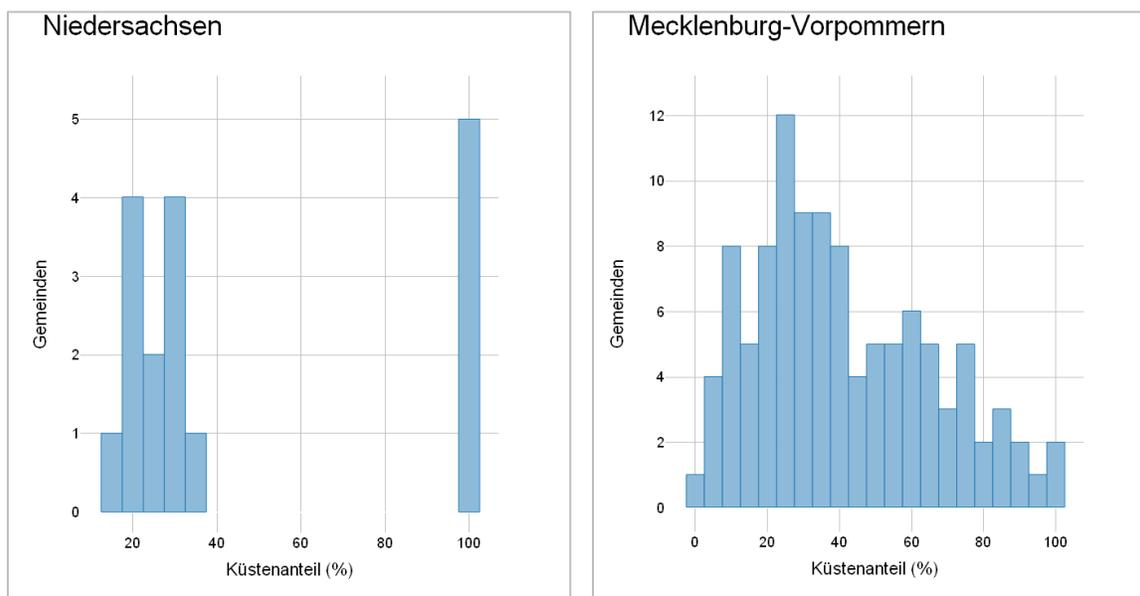


Abbildung 3.1: Histogramme des Indikators „Küstenanteil“ unter Ausschluss der Binnengemeinden. (Eigene Darstellungen erstellt mit dem QGIS-Plugin „Data Plotly“)

In Niedersachsen weist der Küstenanteil aufgrund dieser Ursache mit Abstand die höchste Schiefe (9,01) und Wölbung (88,8) unter den Indikatoren auf, während in Mecklenburg-Vorpommern die Bevölkerungsdichte (5,05) noch rechtsschiefer als der Küstenanteil (3,03) ausfällt und die Wölbung von allen fünf Indikatoren den zweitkleinsten Wert annimmt (11,8). Bei Ausschluss der Binnengemeinden ist der Küstenanteil in Mecklenburg-Vorpommern annähernd normalverteilt und weist nur noch eine leicht rechtsschiefe (0,52) und flachgipflige (2,4) Verteilung auf.

Während die Aussagekraft des Indikators in Mecklenburg-Vorpommern mit 107 Küstengemeinden unterschiedlichen Küstenanteils noch gegeben ist, sind Aussagen in Niedersachsen mit nur 17 Küstengemeinden (davon fünf Inseln) schwieriger zu treffen.

Einer Normalverteilung am nächsten kommt in beiden Bundesländern die Entfernung der Gemeinden zum nächsten Oberzentrum und damit der einzige Indikator, welcher sich nicht direkt aus der Geometrie der Gemeinden ableitet, sondern mit der Wahl der möglichen Zielorte in gewisser Weise willkürlich bestimmt wurde. Die Schiefe beträgt in Niedersachsen 0,3 und in Mecklenburg-Vorpommern 0,24 und somit ist die Verteilung in beiden Fällen nur leicht rechtsschief. Die Wölbung nimmt Werte von 2,89 (Niedersachsen) bzw. 2,5 (Mecklenburg-Vorpommern) an und kommt der Wölbung der Normalverteilung (3) damit sehr nah, die Verteilungen sind leicht flachgipflig.

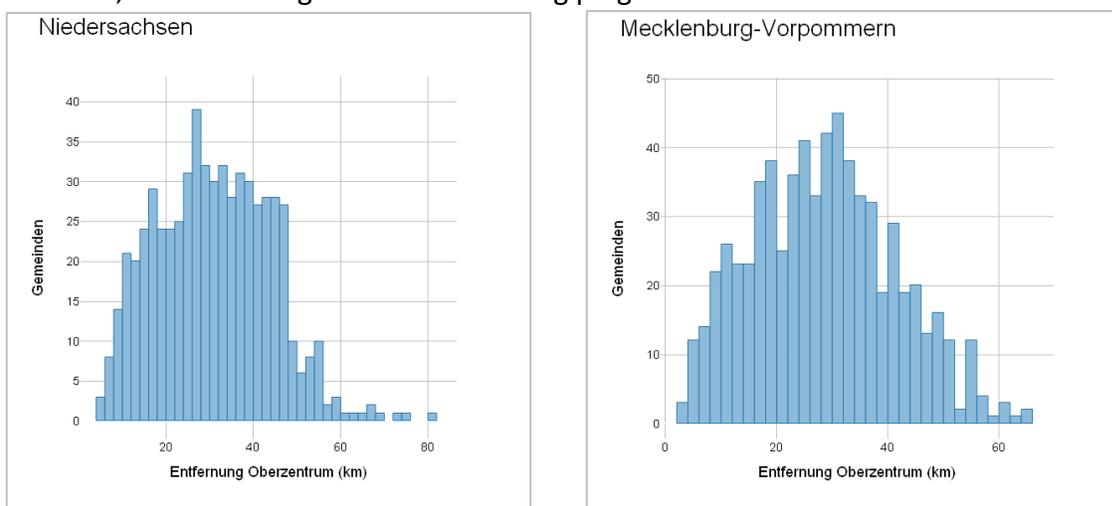


Abbildung 3.2: Histogramme des Indikators „Entfernung Oberzentrum“. (Eigene Darstellungen erstellt mit dem QGIS-Plugin „Data Plotly“)

Auch andere statistischen Maße wie Mittelwert (30,93 bzw. 28,41 Kilometer), Median (30,72 bzw. 28,33 Kilometer) und Standardabweichung (13,17 bzw. 12,83 Kilometer, jeweils zuerst der Wert von Niedersachsen) dieses Indikators sind in beiden Bundesländern ähnlich groß. Die Spannweite ist in Niedersachsen mit etwa 5 bis 81 Kilometern größer als in Mecklenburg-Vorpommern mit etwa 2,5 bis 65 Kilometern.

Die weiteren Indikatoren (Fläche, Bevölkerungsdichte und Shape-Index) sind allesamt in beiden Bundesländern rechtsschief und steilgipflig, wobei Schiefe und Wölbung bei der Bevölkerungsdichte am größten und beim Shape-Index am kleinsten sind. Alle drei Indikatoren sind in Mecklenburg-Vorpommern rechtsschiefer und steilgipfliger verteilt als in Niedersachsen. Beim Shape-Index sind in beiden Bundesländern Mittelwert und die Quartile fast identisch. Alle statistischen Maße der Indikatoren sind in Tabelle 3.1 dargestellt.

	Anzahl	Minimum	Maximum	Mittelwert	Schiefe	Wölbung	Varianz	Standardabweichung	1. Quartil	Median	3. Quartil
Fläche (km²)											
M.-V.	674	2,22	201,02	28,52	2,88	15,9	488,01	22,07	15,11	22,72	34,66
Nds.	603	1,79	177,11	25,46	2,29	12,72	351,7	18,74	12,59	20,9	32,31
Bevölkerungsdichte (je km²)											
M.-V.	674	5,86	619,66	42,86	5,05	37,44	3086	55,52	17,67	26,36	44,4
Nds.	603	8,36	1077,32	92,33	4,25	31,02	9256,61	96,13	39,69	63,8	109,53
Küstenanteil (%)											
M.-V.	674	0	100	6,51	3,03	11,8	319,6	17,86	0	0	0
Nds.	603	0	100	1,33	9,01	88,8	94,77	9,73	0	0	0
Küstenanteil (%) (ohne Binnengemeinden)											
M.-V.	107	1,9	100	40,99	0,52	2,4	602,6	24,43	22,7	35	58,7
Nds.	17	13,8	100	47,28	0,83	1,79	1257,3	34,4	22,4	29,7	100
Entfernung Oberzentrum (km)											
M.-V.	674	2,5	64,42	28,41	0,2416	2,5	164,93	12,83	18,68	28,33	36,79
Nds.	603	5,01	80,99	30,93	0,3025	2,89	173,66	13,17	20,41	30,72	40,58
Shape-Index											
M.-V.	674	1,14	4,2	1,66	2	12,2	0,0973	0,3117	1,44	1,61	1,8
Nds.	603	1,1	3,99	1,66	1,74	9,49	0,1003	0,3164	1,44	1,6	1,81

Tabelle 3.1: Deskriptive statistische Maße der Indikatoren in Mecklenburg-Vorpommern und Niedersachsen.

3.2 Deskriptive statistische Maße der Pendelparameter

Im Gegensatz zu den Indikatoren sind die berechneten Pendelparameter nicht bei allen untersuchten Gemeinden vorhanden. Das liegt an den vorkommenden Anonymisierungen der Pendeldaten des Statistik-Service der Bundesagentur für Arbeit. Die anonymisierten Werte werden in den meisten Fällen darauf hindeuten, dass es in der Gemeinde sehr wenige sozialversicherungspflichtige Arbeitsplätze gibt (womit es auch kaum Einpendelnde geben kann) und bzw. oder die Anzahl der Aus- oder der Binnenpendelnde sehr klein ist, wobei das zumeist auf die Binnenpendelnden zutreffen wird. Es gibt für diese Vermutungen jedoch keine Garantien für alle Gemeinden, weshalb es nicht korrekt wäre, die fehlenden Pendelwerte konsequent mit dem Wert „0“ gleichzusetzen. Eine einzige Ausnahme wurde, wie im Abschnitt 2.4.3.3 beschrieben, bei der Berechnung des Independence-Index gemacht, da es bei diesem keine große Auswirkung hat, ob die Anzahl der Einpendelnden 0 oder sehr klein ist, bei fehlender Anzahl an Einpendelnden dieser jedoch nicht hätte berechnet werden können. Einen Überblick über die Anzahl der Gemeinden mit fehlenden Pendelwerten bzw. -strömen und das damit verbundene geringere Vorhandensein der Pendelparameter ist in Tabelle 3.2 dargestellt. Die Anzahl der sozialversicherungspflichtig Beschäftigten mit Wohnort in der Gemeinde ist für alle Gemeinden vorhanden.

	Mecklenburg- Vorpommern	Niedersachsen
untersuchte Gemeinden	674	603
Gemeinden mit fehlenden Pendelwerten bzw. -strömen		
Arbeitsplätze	100	31
Einpendelnde	116	33
Aus- und Binnenpendelnde	35	16
Einpendelströme	276	137
Auspindelströme	6	2
Anzahl an Gemeinden ohne Pendelparameter		
Einpendelquote	117	35
Aus- und Binnenpendelquote	35	16
Einpendeldistanz	276	137
Auspendeldistanz	6	2
Independence-Index	35	16
Arbeitsplatzausstattung	100	31

Tabelle 3.2: Anzahl der Gemeinden mit fehlenden Pendelwerte bzw. fehlenden Pendelparametern.

Von allen Pendelparametern kommt die Einpendelquote einer Normalverteilung am nächsten. In Mecklenburg-Vorpommern nimmt bei dieser die Schiefe einen Wert von etwa -0,4 und die Wölbung einen Wert von 3,05 an, in Niedersachsen ist die Verteilung noch etwas linkschief (-1,39) und steilgipfliger (7,76).

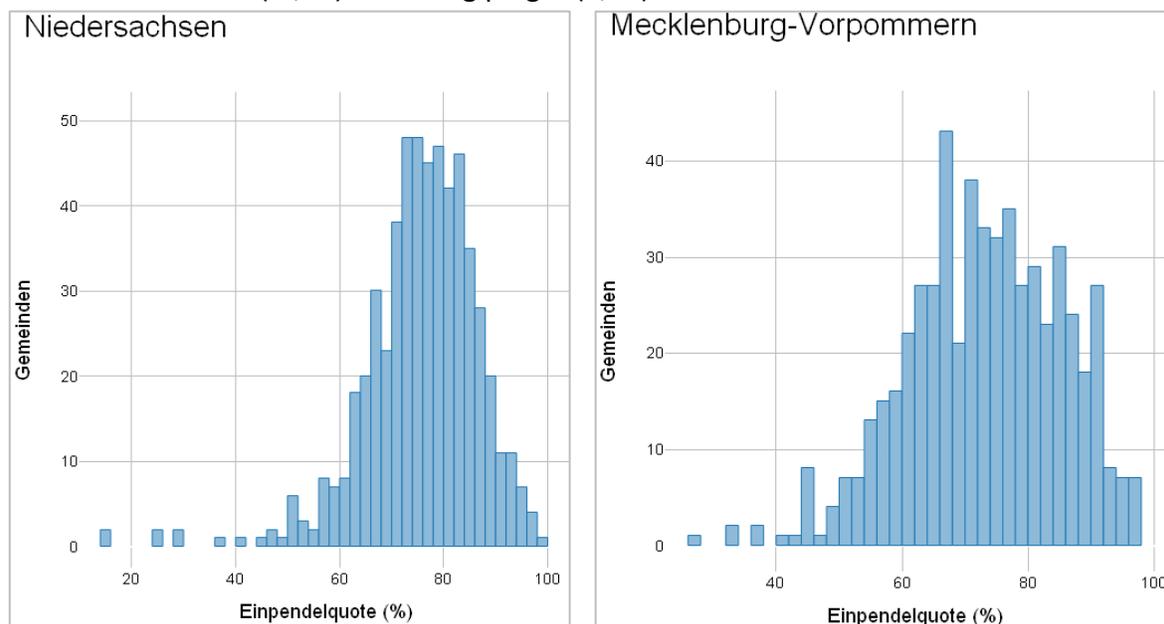


Abbildung 3.3: Histogramme des Pendelparameters „Einpendelquote“. (Eigene Darstellungen erstellt mit dem QGIS-Plugin „Data Plotly“)

Auffällig ist, dass auch alle anderen Pendelparameter in Niedersachsen eine höhere Schiefe (im Betrag) und Wölbung aufweisen als in Mecklenburg-Vorpommern. Die einzige Ausnahme davon ist die Arbeitsplatzausstattung, welche in Mecklenburg-Vorpommern rechtsschief (4,13 zu 2,39) und steilgipfliger (31,5 zu 11,61) ausfällt.

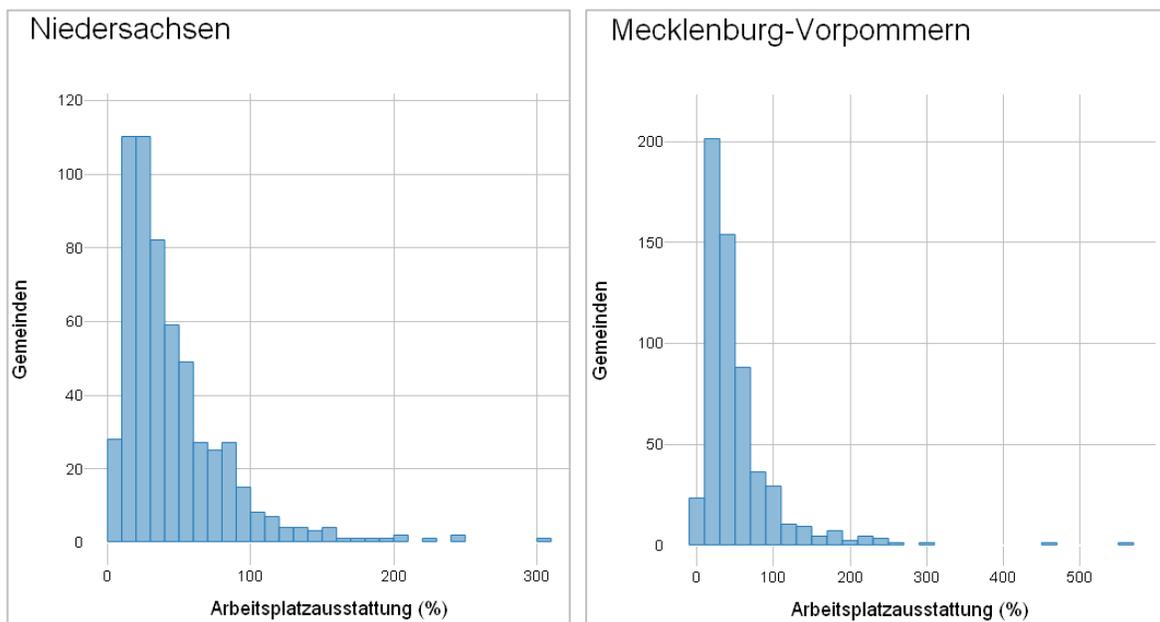


Abbildung 3.4: Histogramme des Pendelparameters „Arbeitsplatzausstattung“. (Eigene Darstellungen erstellt mit dem QGIS-Plugin „Data Plotly“)

Eine weitere Auffälligkeit ist das Vorhandensein von starken Ausreißern nach oben bei allen Pendelparametern abgesehen von der Ein- und Auspendelquote, bei welchen es dafür starke Ausreißer nach unten gibt. Auch diese fallen, genau wie Schiefe und Wölbung, in Niedersachsen bei allen Pendelparametern mit Ausnahme der Arbeitsplatzausstattung extremer aus als in Mecklenburg-Vorpommern. Die Ausreißer können einen großen Einfluss auf die statistischen Maße wie z.B. Schiefe und Wölbung haben (vgl. z.B. Grellmann o.J.). Alle statistischen Maße der Pendelparameter sind in Tabelle 3.3 dargestellt.

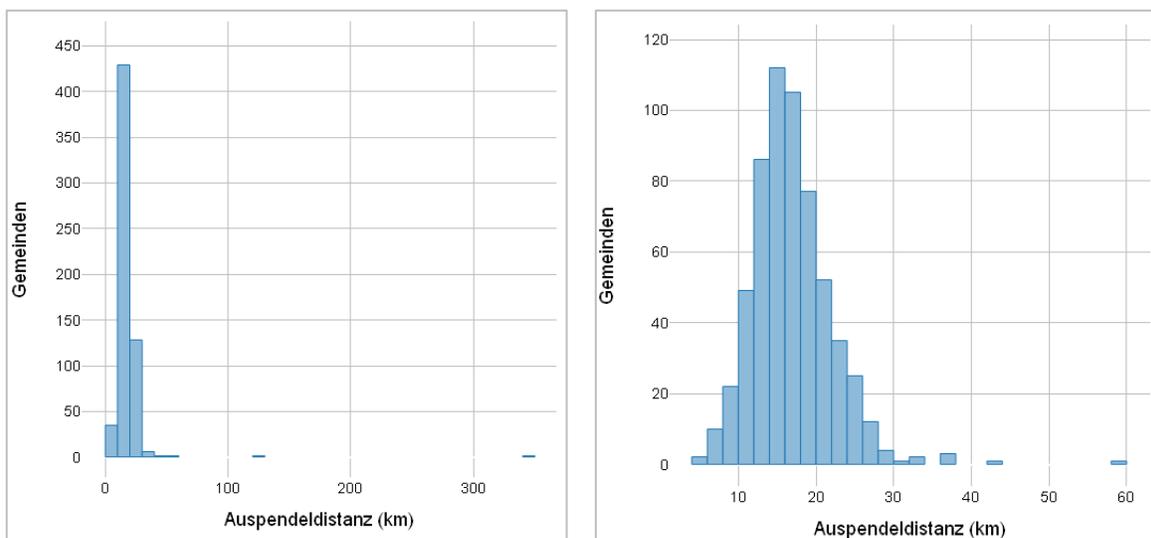


Abbildung 3.5: Histogramm des Pendelparameters „Auspendelndistanz“ in Niedersachsen. Links sind alle Werte dargestellt, rechts wurden die Gemeinden mit einem Wert von mind. 100 Kilometern entfernt. (Eigene Darstellungen erstellt mit dem QGIS-Plugin „Data Plotly“)

	Anzahl	Minimum	Maximum	Mittelwert	Schiefe	Wölbung	Varianz	Standardabweichung	1. Quartil	Median	3. Quartil
Einpendelquote (%)											
M.-V.	557	26,67	97,88	73,04	-0,3962	3,05	152,33	12,33	65,35	72,96	82,69
Nds.	568	14,11	98,76	75,52	-1,39	7,76	124,48	11,15	70,15	76,42	82,77
Auspendelquote (%)											
M.-V.	639	28,49	100	88,82	-2,12	9,16	87,41	9,34	86,19	91,66	94,89
Nds.	587	14,29	99,58	90,2	-4,44	32,43	85,48	9,24	87,6	92,31	95,47
Einpendeldistanz (km)											
M.-V.	398	2,14	120,23	11,16	7,06	74,52	79,72	8,92	7,19	9,84	12,64
Nds.	466	2,04	302,37	11,83	13,78	233,43	260,51	16,12	7,58	9,8	12,54
Auspendeldistanz (km)											
M.-V.	668	2,2	87,77	16,4	2,51	20,45	50,09	7,07	11,94	15,32	19,72
Nds.	601	5,66	340,24	17,63	17,94	382,1	218,78	14,78	13,5	16,45	19,59
Independence-Index											
M.-V.	639	0	0,9366	0,092	3,26	25,54	0,0062	0,0786	0,0437	0,0707	0,1145
Nds.	587	0,0042	2,79	0,0907	9,97	112,78	0,0389	0,197	0,0391	0,0622	0,095
Arbeitsplatzausstattung (%)											
M.-V.	574	0	563,91	51,16	4,13	31,5	2530,05	50,26	22,84	38,08	60,61
Nds.	572	1,01	301,29	45,37	2,39	11,61	1404,97	37,45	20,51	34,41	57,53

Tabelle 3.3: Deskriptive stat. Maße der Pendelparameter in Mecklenburg-Vorpommern und Niedersachsen.

Um die Werte der beiden Bundesländer noch besser miteinander vergleichen zu können, wird die Berechnung der statistischen Maße wiederholt, dabei jedoch jeweils die unteren zwei Prozent (Ein- und Auspendelquote) bzw. die oberen zwei Prozent (Ein- und Auspendeldistanz, Independence-Index und Arbeitsplatzausstattung) der Werte herausgefiltert. Dazu wird die SQL-Anweisung zur Berechnung der statistischen Maße (siehe SQL 2.13) um die Bedingung aus SQL 3.1 ergänzt. Das Ergebnis ist in Tabelle 3.4 abgebildet.

```
... WHERE arbeitsplatzausstattung <=
(SELECT percentile_disc (0.98) within group (order by arbeitsplatzausstattung)
FROM gemeinden_final);

... WHERE einpendelquote >=
(SELECT percentile_disc (0.02) within group (order by einpendelquote)
FROM gemeinden_final);
```

SQL 3.1: Ergänzung zur Berechnung der statistischen Maße (siehe SQL 2.13) mit Herausfiltern der höchsten bzw. niedrigsten zwei Prozent an Werten am Beispiel von Arbeitsplatzausstattung und Einpendelquote.

Am Vergleich der statistischen Maße mit und ohne der obersten oder untersten zwei Prozent wird der Einfluss der Ausreißer sehr deutlich. Die Schiefe und Wölbung aller Verteilungen liegt deutlich näher an den Werten einer Normalverteilung, z.B. hat sich bei der Auspendeldistanz in Niedersachsen die Schiefe von 17,94 auf 0,28 und die Wölbung von 382,1 auf 2,78 verringert. Einzige Ausnahme ist hierbei die Wölbung der Einpendelquote aus Mecklenburg-Vorpommern, welche mit den Ausreißern nur leicht steilgipflig ist (3,05), ohne dieser jedoch flachgipflig ausfällt (2,34).

Im Vergleich der beiden Bundesländer fällt unter Ausschluss der Gemeinden mit den Werten der obersten oder untersten zwei Prozent folgendes auf:

- in den Gemeinden Niedersachsens liegen Mittelwert und Median bei Ein- und Auspendelquote leicht über denen der Gemeinden Mecklenburg-Vorpommerns
- die statistischen Maße der Ein- und Auspendeldistanz sind in beiden Bundesländern annähernd gleich, jedoch ist die Spannweite (Differenz von Maximum und Minimum) der Einpendeldistanz in Niedersachsen und die der Auspendeldistanz in Mecklenburg-Vorpommern größer (jeweils sowohl ein kleineres Minimum als auch ein größeres Maximum)
- der Independence-Index ist allgemein sehr niedrig, in Mecklenburg-Vorpommern sind Mittelwert und Median minimal größer als in Niedersachsen
- auch bei der Arbeitsplatzausstattung liegen Mittelwert und Median der Gemeinden Mecklenburg-Vorpommerns über denen von Niedersachsen

	Anzahl	Minimum	Maximum	Mittelwert	Schiefe	Wölbung	Varianz	Standardabweichung	1. Quartil	Median	3. Quartil
Einpendelquote (%)											
M.-V.	546	44,73	97,88	73,72	-0,1095	2,34	130,9	11,43	65,56	73,39	82,9
Nds.	557	48	98,76	76,38	-0,3108	3,04	86,58	9,3	70,52	76,73	83
Auspindelquote (%)											
M.-V.	627	61,68	100	89,55	-1,4	4,62	59,06	7,68	86,73	91,75	94,94
Nds.	576	71,19	99,58	91,09	-1,18	4,07	34,09	5,83	88,23	92,45	95,49
Einpendeldistanz (km)											
M.-V.	391	2,14	30,57	10,27	1,06	4,93	20,27	4,5	7,1	9,8	12,57
Nds.	457	2,04	32,67	10,31	1,17	5,94	17,05	4,12	7,57	9,71	12,33
Auspendeldistanz (km)											
M.-V.	655	2,2	32,6	15,88	0,5	2,95	32,22	5,67	11,91	15,12	19,36
Nds.	589	5,66	28,62	16,58	0,28	2,78	19,23	4,38	13,43	16,27	19,4
Independence-Index											
M.-V.	627	0	0,3249	0,0858	1,4	4,82	0,0037	0,0612	0,0434	0,0689	0,1102
Nds.	576	0,0042	0,231	0,0702	1,1	3,99	0,0018	0,0423	0,0386	0,0614	0,091
Arbeitsplatzausstattung (%)											
M.-V.	563	0	206,47	46,51	1,78	6,86	1211,17	34,77	22,43	37,7	59,76
Nds.	561	1,01	156,3	42,16	1,3	4,58	863,1	29,35	20,24	34,12	56,23

Tabelle 3.4: Deskriptive statistische Maße der Pendelparameter in Mecklenburg-Vorpommern und Niedersachsen unter Ausschluss der Gemeinden, die im jeweiligen Parameter zu den zwei Prozent mit den niedrigsten (Ein- und Auspendelquote) bzw. höchsten Werten (restliche Parameter) gehören.

3.3 Korrelation der Pendelparameter untereinander

Ein weiterer Teilaspekt der Untersuchung ist die Frage, ob ein Zusammenhang einzelner Pendelparameter untereinander erkennbar ist. Einige der Parameter hängen von den gleichen Werten ab, so werden z.B. sowohl die Ein- bzw. Auspendelquote als auch der Independence-Index von der Anzahl der Ein- bzw. Auspendelnden beeinflusst, weshalb diese Korrelationen nicht untersucht werden. Auch andere Konstellationen habe ich ausgeschlossen, wenn die Parameter zwei unterschiedliche Aspekte betrachten, z.B. Einpendelquote und Auspendeldistanz. Verblieben sind die in Tabelle 3.5 dargestellten fünf untersuchten Korrelationen.

Korrelationskoeffizient zwischen den Pendelparametern...	... in Mecklenburg-Vorpommern	... in Niedersachsen
Einpendelquote ↔ Einpendeldistanz	0,0203	0,0399
Auspindelquote ↔ Auspendeldistanz	-0,357	-0,4614
Einpendelquote ↔ Auspendelquote	0,3395	0,4468
Einpendeldistanz ↔ Auspendeldistanz	0,1297	0,069
Arbeitsplatzausstattung ↔ Independence-Index	0,2688	0,2181

Tabelle 3.5: Korrelationen ausgewählter Pendelparameter untereinander.

Auffällig ist der große Unterschied sowohl zwischen den Konstellationen Einpendelquote \leftrightarrow Einpendeldistanz und Auspendelquote \leftrightarrow Auspendeldistanz als auch zwischen den Konstellationen Einpendelquote \leftrightarrow Auspendelquote und Einpendeldistanz \leftrightarrow Auspendeldistanz. Während der Zusammenhang zwischen Quote und Distanz in beiden Bundesländern bei Einpendelnden kaum vorhanden ist, ist insbesondere in Niedersachsen eine deutlichere negative Korrelation bei den Auspendelnden erkennbar. Je höher die Auspendelquote ausfällt, desto geringer wird die durchschnittliche Auspendeldistanz.

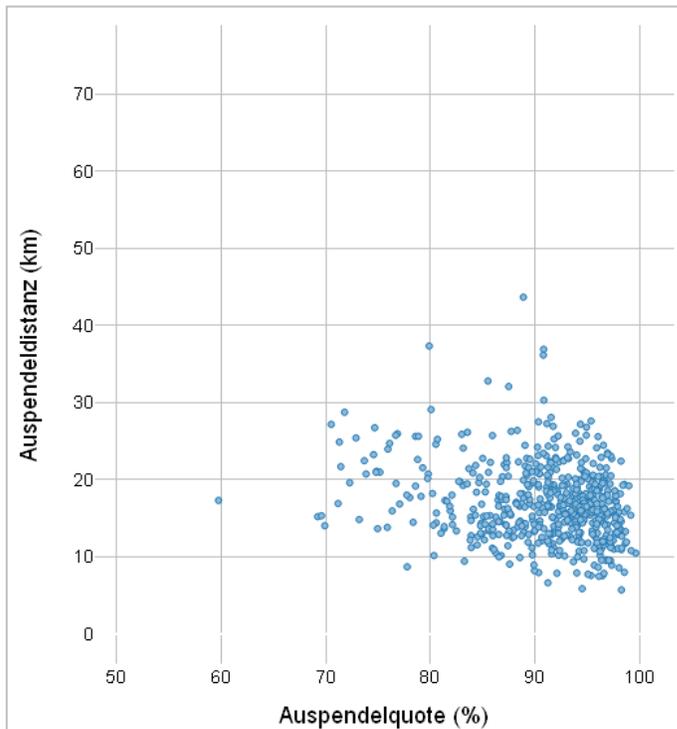


Abbildung 3.6: Ausschnitt des Streudiagramms für die Auspendelquote und -distanz in Niedersachsen. Die Ausreißer (Quote < 50%, Distanz > 50 Kilometer) sind nicht dargestellt.
(Eigene Darstellung erstellt mit dem QGIS-Plugin „Data Plotly“)

Ähnlich verhält es sich mit den Korrelationen der Ein- und Auspendelquoten bzw. -distanzen jeweils zueinander. Während die Distanzen kaum korrelieren, ist zwischen den Quoten ein deutlicherer Zusammenhang ersichtlich, wieder insbesondere in Niedersachsen. Dabei wächst mit steigender Einpendelquote auch die Auspendelquote und umgekehrt.

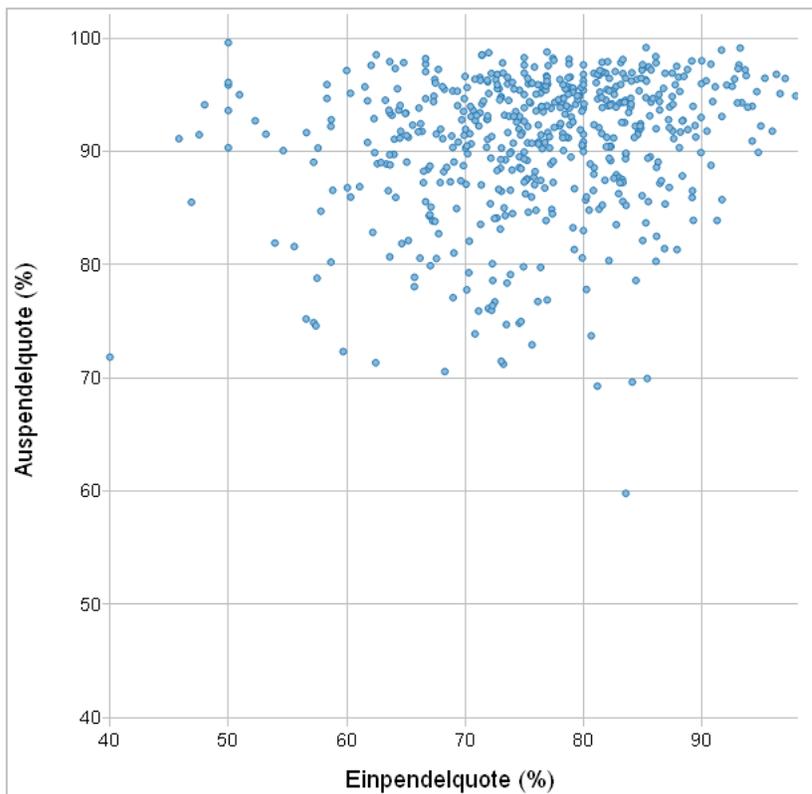


Abbildung 3.7: Ausschnitt des Streudiagramms für die Ein- und Auspendelquote in Niedersachsen. Die Ausreißer (Quoten < 40%) sind nicht dargestellt.
(Eigene Darstellung erstellt mit dem QGIS-Plugin „Data Plotly“)

Etwas genauer soll der Zusammenhang zwischen Arbeitsplatzausstattung und Independence-Index untersucht werden. Die Korrelationskoeffizienten sind mit 0,2688 in Mecklenburg-Vorpommern und 0,2181 in Niedersachsen für die gesamten Datensätze eher gering. Jedoch soll untersucht werden, ob die Korrelation größer wird, wenn nur die Gemeinden bis bzw. ab 100% Arbeitsplatzausstattung untersucht werden. Hintergrund ist die Abhängigkeit des theoretisch maximal möglichen Independence-Indexes von der Arbeitsplatzausstattung, welcher bei der maximal möglichen Anzahl von Binnenpendelnden bei gegebener Anzahl an Arbeitenden und Arbeitsplätzen in einer Gemeinde erreicht wird. Eine kleiner werdende Arbeitsplatzausstattung unter 100% bewirkt eine höhere Anzahl an Auspendelnden, während bei einer wachsenden Arbeitsplatzausstattung über 100% die Gemeinde abhängig von Einpendelnden wird. Beide Fälle verursachen einen kleineren Independence-Index, z.B. kann dieser nur bei einer Arbeitsplatzausstattung zwischen 50% und 200% theoretisch Werte von 1 annehmen, mit sich 100% annähernder Arbeitsplatzausstattung divergiert der theoretisch mögliche Independence-Index nach Unendlich.

Arbeitsplatz- ausstattung (%)	10	25	50	75	90	100	120	150	200	300	500
Maximal möglicher Independence-Index	1/9	1/3	1	3	9	∞	5	2	1	0,5	0,25

Tabelle 3.6: Veranschaulichung der Abhängigkeit des maximalen Independence-Index von der Arbeitsplatzausstattung.

Die Ausgabe der Korrelationskoeffizienten für die Teilmengen der Gemeinden erfolgt durch eine Ergänzung der in SQL 2.13 vorgestellten Abfrage.

```
SELECT corr (arbeitsplatzausstattung, independence)
FROM gemeinden_final
WHERE arbeitsplatzausstattung <= 100;

... WHERE arbeitsplatzausstattung >= 100;
```

SQL 3.2: Bestimmung der Korrelation zwischen Arbeitsplatzausstattung und Independence-Index nur für die Gemeinden mit einer Arbeitsplatzausstattung bis einschließlich bzw. mindestens 100%.

Korrelationskoeffizient Arbeitsplatzausstattung ↔ Independence-Index in Gemeinden...		
	... <= 100% Arbeitsplatzausstattung	... >= 100% Arbeitsplatzausstattung
Mecklenburg-Vorpommern	0,5188 (523)	-0,3097 (51)
Niedersachsen	0,3273 (533)	-0,1897 (40)

Tabelle 3.7: Korrelation zwischen Arbeitsplatzausstattung und Independence-Index in den Gemeinden mit einer Arbeitsplatzausstattung bis einschließlich bzw. mindestens 100%. In Klammern steht jeweils die Anzahl der Gemeinden.

Der Unterschied in der Stärke des Zusammenhangs wird zwischen den beiden Bundesländern bei dieser Betrachtung deutlicher sichtbar. In Mecklenburg-Vorpommern ist die Korrelation zwischen den beiden Pendelparametern stärker als in Niedersachsen, in beiden Bundesländern ist die positive Korrelation bei den Gemeinden mit maximal 100% Arbeitsplatzausstattung stärker als die negative Korrelation der Gemeinden mit mindestens 100% Arbeitsplatzausstattung. Dabei muss aber beachtet werden, dass die Anzahl der Gemeinden in letzterer Gruppe geringer ist.

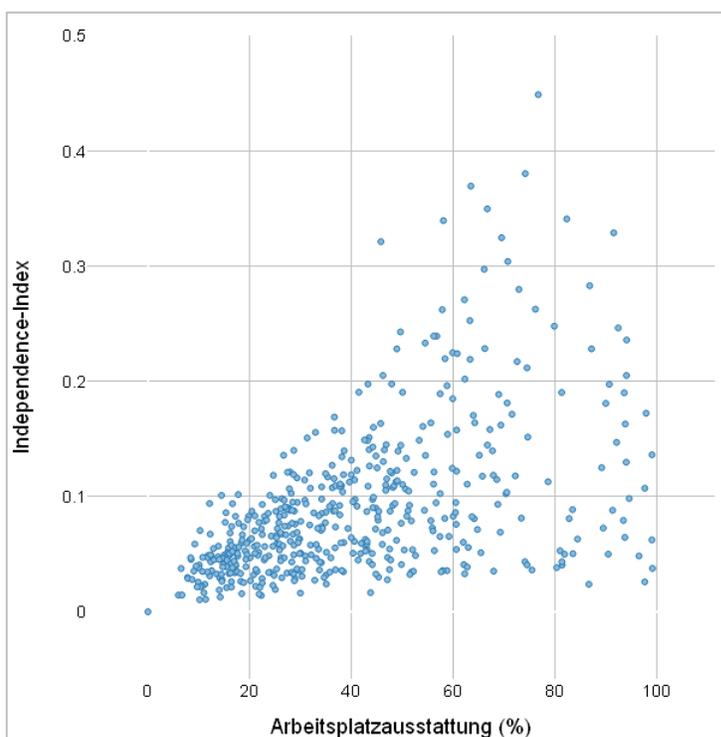


Abbildung 3.8: Streudiagramm für die Arbeitsplatzausstattung und den Independence-Index in Mecklenburg-Vorpommern für die Gemeinden mit einer Arbeitsplatzausstattung von maximal 100%. (Eigene Darstellung erstellt mit dem QGIS-Plugin „Data Plotly“)

3.4 Korrelation Indikatoren und Pendelparameter

Die Untersuchung des Zusammenhangs zwischen den räumlichen Indikatoren und den Pendelparametern ist der hauptsächliche Fokus dieser Master-Thesis. Mit der Korrelationsanalyse wird untersucht, ob ein signifikanter Einfluss der Indikatoren auf die Pendelparameter vorliegt. Die ersten Berechnungen unter Verwendung aller untersuchten Gemeinden ergibt die in Tabelle 3.8 dargestellten Werte.

	Einpendel- quote	Auspindel- quote	Einpendel- distanz	Auspindel- distanz	Arbeitsplatz- ausstattung	Independence- Index
Fläche	-0,3491	-0,3297	0,0851	0,3653	0,0030	0,3888
	-0,2273	-0,2331	0,0286	0,0713	0,1126	0,0493
Bevölkerungs- dichte	0,1395	-0,2483	0,0467	0,0616	0,2366	0,1586
	0,1717	0,0087	-0,0367	-0,0572	0,1329	-0,0224
Küstenanteil	-0,1263	-0,4355	0,0401	-0,0093	0,1945	0,4038
	-0,4577	-0,7130	0,0007	0,5985	0,1712	0,8608
Entfernung Oberzentrum	-0,3198	-0,2742	-0,0586	0,1231	0,0198	0,2732
	-0,1588	-0,2655	-0,0256	0,0687	0,1333	0,1826
Shape-Index	-0,1714	-0,3378	0,0732	0,1663	0,1064	0,3825
	-0,0326	-0,1144	0,0026	0,0169	0,0690	0,1272

Tabelle 3.8: Korrelation zwischen den räumlichen Indikatoren und den Pendelparametern in den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens. Absolutwerte von mindestens 0,3 sind blau hinterlegt.

Es sind hier zu großen Teilen nur geringe oder sehr geringe Korrelationen, wenn überhaupt, erkennbar. Absolutwerte der Korrelation von mehr als 0,5 gibt es nur beim Küstenanteil in Niedersachsen, bei der Bevölkerungsdichte sowie der Arbeitsplatzausstattung erreicht keine Korrelation Absolutwerte von mehr als 0,25 und bei der Einpendeldistanz sind alle Korrelationen kleiner als 0,1.

Wie im Abschnitt 3.2 soll eine Wiederholung der Berechnung unter Ausschluss der Ausreißer stattfinden. Für die Pendelparameter erfolgt der Ausschluss analog zum genannten Abschnitt, für die räumlichen Indikatoren erfolgt ein Ausschluss der oberen zwei Prozent nur bei der Fläche, der Bevölkerungsdichte und dem Shape-Index. Beim Küstenanteil und der Entfernung zum Oberzentrum wird kein Ausschluss vorgenommen, da bei diesen Indikatoren keine starken Ausreißer (nach oben oder unten) vorhanden sind.

```
SELECT corr (flaeche_km, einpendelquote) FROM gemeinden_final
WHERE einpendelquote >=
  (SELECT percentile_disc (0.02) within group (order by einpendelquote)
   FROM gemeinden_final)
AND flaeche_km <=
  (SELECT percentile_disc (0.98) within group (order by flaeche_km)
   FROM gemeinden_final);
```

SQL 3.3: Bestimmung der Korrelationen zwischen räumlichen Indikatoren und Pendelparametern unter Ausschluss der Gemeinden mit Ausreißern am Beispiel der Fläche und der Einpendelquote.

	Einpendel- quote	Auspandel- quote	Einpendel- distanz	Auspandel- distanz	Arbeitsplatz- ausstattung	Independence- Index
Fläche	-0,3643	-0,3436	0,1180	0,3095	0,0202	0,3822
	-0,3381	-0,4853	0,1693	0,3069	0,1541	0,5151
Bevölkerungs- dichte	0,2056	-0,1845	0,1228	0,0066	0,3651	0,0763
	0,1766	0,0712	-0,0309	-0,0870	0,0662	-0,1068
Küstenanteil	-0,0801	-0,3518	0,0205	-0,0276	0,2634	0,3035
	-0,1161	-0,1962	0,0493	0,0504	0,2135	0,1367
Entfernung Oberzentrum	-0,3287	-0,2874	-0,1025	0,0512	0,0359	0,3159
	-0,1077	-0,2565	-0,0986	0,0500	0,1904	0,2602
Shape-Index	-0,1193	-0,2463	-0,0007	0,1589	0,1004	0,1928
	-0,0010	0,0386	-0,0116	0,0269	-0,0238	-0,0254

Tabelle 3.9: Korrelation zwischen den räumlichen Indikatoren und den Pendelparametern in den Gemeinden Mecklenburg-Vorpommerns und Niedersachsens unter Ausschluss der Gemeinden, die im jeweiligen Parameter zu den zwei Prozent mit den niedrigsten (Ein- und Auspendelquote) bzw. höchsten (sonstige Pendelparameter sowie Fläche, Bevölkerungsdichte und Shape-Index) Werten gehören. Absolutwerte von mindestens 0,3 sind blau hinterlegt.

Der Ausschluss der Gemeinden mit Ausreißern hat nicht in allen Fällen zu einer erhöhten Korrelation zwischen den räumlichen Indikatoren und den Pendelparametern geführt. Bei 29 von 60 Werten erhöht sich der Absolutwert der Korrelation um mindestens 0,01 gegenüber der Berechnung ohne Ausschluss, bei 24 Werten wird der Absolutwert der Korrelation jedoch um mindestens 0,01 kleiner. Dabei gibt es deutliche Unterschiede zwischen den einzelnen Indikatoren und Parametern, wie Tabelle 3.10 zeigt. Bei Korrelationsberechnungen mit der Fläche und der Entfernung zum nächsten Oberzentrum sowie mit der Einpendeldistanz und der Arbeitsplatzausstattung führt ein Ausschluss der Ausreißer zumeist zu stärkeren Korrelationen, während bei Korrelationsberechnungen mit dem Küstenanteil und dem Shape-Index meist das Gegenteil der Fall ist.

	Einpendel- quote	Auspendel- quote	Einpendel- distanz	Auspendel- distanz	Arbeitsplatz- ausstattung	Independence- Index
Fläche						
Bevölkerungs- dichte						
Küstenanteil						
Entfernung Oberzentrum						
Shape-Index						

Tabelle 3.10: Vergleich der Korrelation zwischen den räumlichen Indikatoren und Pendelparametern in den Gemeinden Mecklenburg-Vorpommerns (oben) und Niedersachsens (unten) vor und nach Ausschluss der Ausreißer. Bei grün hinterlegten Zellen ist der Absolutwert der Korrelation bei Ausschluss der Ausreißer um mindestens 0,01 gestiegen, bei gelb hinterlegten Zellen um mindestens 0,01 gesunken.

Allgemein betrachtet sind die Korrelationen gering. Nach Ausschluss der Ausreißer erreicht nur die Korrelation zwischen der Fläche und dem Independence-Index in Niedersachsen einen Wert von über 0,5 (0,5151). Hier ist auch der stärkste Anstieg der Korrelation nach dem Ausschluss erkennbar, so hat die gleiche Korrelation mit dem gesamten Datensatz nur einen Wert von 0,0493 aufgewiesen.

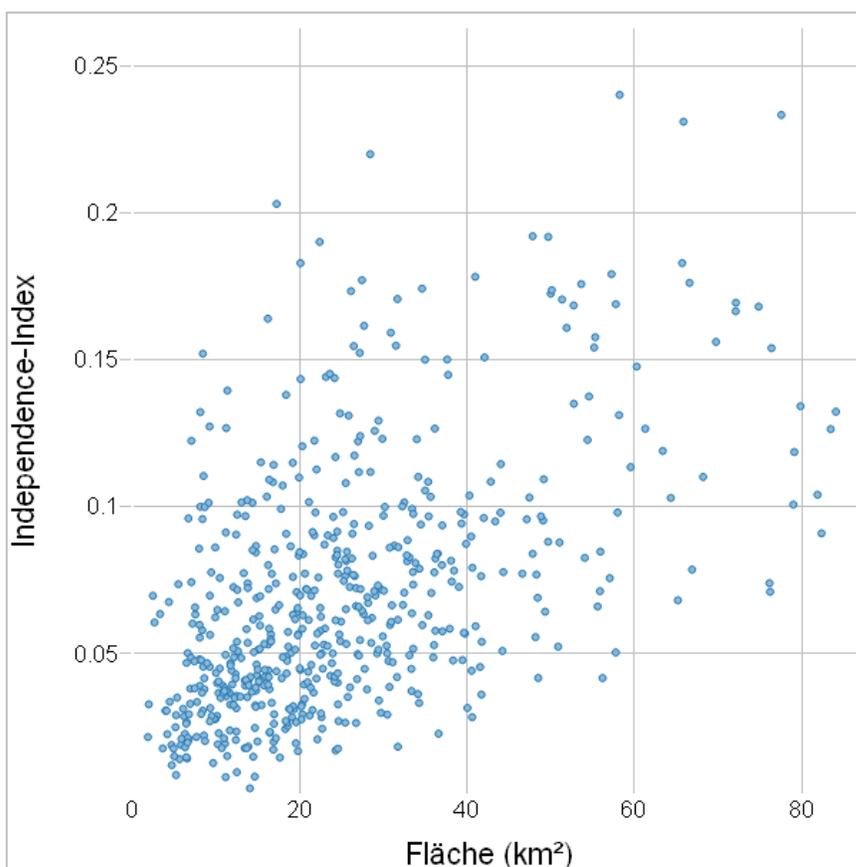


Abbildung 3.9: Streudiagramm für die Fläche und den Independence-Index der Gemeinden Niedersachsen ohne Ausreißer. (Eigene Darstellung erstellt mit dem QGIS-Plugin „Data Plotly“)

Die bereits angesprochene vergleichsweise hohen Korrelationen zwischen dem Küstenanteil und insgesamt drei Pendelparametern in Niedersachsen vor Ausschluss der Ausreißer muss erneut unter Berücksichtigung der Tatsache betrachtet werden, dass 586 von 603 Gemeinden einen Küstenanteil von 0% aufweisen. Betrachtet man die entsprechenden Streudiagramme (siehe Abbildung 3.10), wird insbesondere der Unterschied zwischen Binnen- und Inselgemeinden (100% Küstenanteil, fünf Gemeinden) deutlich. So nimmt z.B. der Independence-Index bei Binnengemeinden mit einer Ausnahme maximal Werte von 0,37 an, während die fünf Inselgemeinden Werte zwischen 1,25 und 2,8 aufweisen. Ähnliche Auffälligkeiten treten auch mit der Ein- und Auspendelquote sowie der Auspendeldistanz aus. Dadurch verringern sich hier die Korrelationen, wenn die Ausreißer ausgeschlossen werden. Die Gemeinden Mecklenburg-Vorpommerns, bei welchen die Verteilung zwischen Küsten- und Binnengemeinden etwas ausgewogener ausfällt, bestätigen die Zusammenhänge nur bei der Auspendelquote (-0,4355) und dem Independence-Index (0,4038), jedoch schwächer.

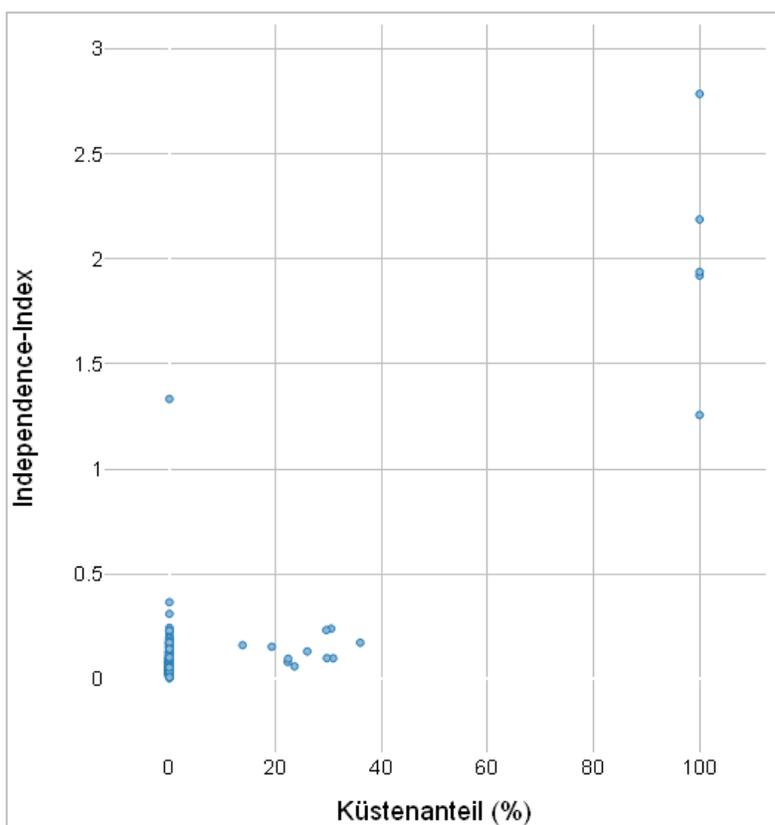


Abbildung 3.10: Streudiagramm für den Küstenanteil und den Independence-Index der Gemeinden Niedersachsens. (Eigene Darstellung erstellt mit dem QGIS-Plugin „Data Plotly“)

Es lässt sich sagen, dass Ein- und Auspendelquote insgesamt deutlich stärkere Korrelationen aufweisen als Ein- und Auspendeldistanz. Bei der Einpendeldistanz hat die höchste vorhandene Korrelation (Flächengröße; Niedersachsen) einen Wert von 0,1693. Bei der Auspendeldistanz hat neben dem Küstenanteil in Niedersachsen zumindest auch die Flächengröße in beiden Bundesländern eine wahrnehmbare Korrelation (jeweils etwa 0,3). Bei der Ein- und Auspendelquote ist zusätzlich eine stärkere Korrelation mit den räumlichen Indikatoren „Entfernung Oberzentrum“ und (in Mecklenburg-Vorpommern) „Shape-Index“ ersichtlich. Die Korrelationen der Auspendelquote sind dabei zusammengefasst größer als die der Einpendelquote. Auch der Independence-Index zeigt ähnliche Korrelationen wie die Auspendelquote. Bei allen drei Pendelparametern ist die Korrelation in Mecklenburg-Vorpommern zusammengefasst größer als in Niedersachsen.

Die geringste Korrelation bei den räumlichen Indikatoren sind bei der Bevölkerungsdichte vorhanden, mit einem Höchstwert von -0,2483 (Auspendelquote; Mecklenburg-Vorpommern). Die höchsten Korrelationen (neben dem bereits angesprochenen Küstenanteil) sind bei der Flächengröße und der Entfernung zum nächsten Oberzentrum ersichtlich, wobei auch hier Unterschiede zwischen den beiden Bundesländern vorhanden sind, in Mecklenburg-Vorpommern nehmen die Korrelationen zumindest für die

vollständigen Datensätze höhere Werte an. Werden die Ausreißer ausgeschlossen, bewirkt die Flächengröße der Gemeinden in Niedersachsen höhere Korrelationen als der in Mecklenburg-Vorpommern.

Die Richtung der Korrelationen zwischen den Indikatoren Fläche, Bevölkerungsdichte, Küstenanteil und Entfernung zum Oberzentrum und den Pendelquoten entspricht mit Ausnahme eines Wertes der Annahme anhand von Eurostat (2018). So wirken sich eine größere Fläche, ein größerer Küstenanteil und eine höhere Entfernung zum nächsten Oberzentrum negativ auf die Pendelquoten aus, eine höhere Bevölkerungsdichte positiv. Einzig die Auspendelquote in Mecklenburg-Vorpommern hat mit der Bevölkerungsdichte eine negative Korrelation (-0,2483). Der Shape-Index hat, anders als vermutet, einen negativen Einfluss auf die Pendelquoten, die Korrelation mit der Auspendelquote in Niedersachsen ist einzig bei Ausschluss der Ausreißer positiv, mit 0,0386 aber auch nur sehr gering.

Es soll überprüft werden, ob die angesprochene unerwartete Richtung einiger Korrelationen (Shape-Index in beiden Bundesländern, Bevölkerungsdichte in Mecklenburg-Vorpommern) mit anderen Indikatoren zusammenhängt. So sind in Mecklenburg-Vorpommern neun der zehn Gemeinden mit dem höchsten Shape-Index sowie acht der zehn Gemeinden mit der höchsten Bevölkerungsdichte Küstengemeinden. Die negative Korrelation könnte entsprechend durch den Einfluss des Indikators Küstenanteil stammen, welche stärker wirkt als eine eigentlich positive Korrelation mit dem Shape-Index. Eine einfache Überprüfung dieser Möglichkeit ist die Berechnung der Korrelation zwischen den Indikatoren und den Pendelquoten nur für die Binnengemeinden, welche aufgrund der Datenlage beim Küstenanteil nur für Mecklenburg-Vorpommern durchgeführt wurde.

```
SELECT corr (bev_dichte_km, auspendelquote) FROM gemeinden_final
WHERE kueste_anteil = 0;
```

SQL 3.4: Bestimmung der Korrelationen zwischen räumlichen Indikatoren und Pendelparametern nur für die Binnengemeinden am Beispiel der Bevölkerungsdichte und der Auspendelquote.

	Korrelationskoeffizient zwischen...
Bevölkerungsdichte ↔ Auspendelquote	-0,0569
Shape-Index ↔ Einpendelquote	-0,2261
Shape-Index ↔ Auspendelquote	-0,1191

Tabelle 3.11: Werte des Korrelationskoeffizienten der Korrelationen mit unerwarteter Richtung nur für die Binnengemeinden in Mecklenburg-Vorpommern.

Auch für die Binnengemeinden liegen nach wie vor die negativen Korrelationen vor, so dass ein potentieller Störfaktor durch den Küstenanteil sich zumindest für die Richtung der Korrelation nicht bestätigt hat. Mit anderen räumlichen Indikatoren sind keine ähnlichen Zusammenhänge erkennbar. Damit bleibt es bei den genannten Korrelationen bei der unerwarteten Richtung.

3.5 Konkrete Gemeinden im Vergleich

Um den möglichen Einfluss der räumlichen Indikatoren und Unterschiede zwischen den Bundesländern zu verdeutlichen, werden im Folgenden konkrete Gemeinden miteinander verglichen, welche sich jeweils am gegenüberliegenden Ende des Wertebereich eines Indikators befinden, während sich die restlichen Indikatoren ähneln, sowie Gemeinden, welche sich in allen räumlichen Indikatoren ähneln, aber in unterschiedlichen Bundesländern liegen.

3.5.1 Vergleich einer Binnen- und einer Inselgemeinde

Als erstes Beispiel werden eine Binnen- und eine Inselgemeinde aus Niedersachsen miteinander verglichen. Die Gemeinde Bothel ist eine Binnengemeinde, gehört zum Landkreis Rotenburg (Wümme) und liegt etwa 45 Kilometer östlich von Bremen. Die Gemeinde Spiekeroog ist eine Nordseeinsel, gehört zum Landkreis Wittmund und liegt etwa 40 Kilometer nordwestlich von Wilhelmshaven. Die Gemeinden ähneln sich nicht nur in ihrer Entfernung zum nächsten Oberzentrum, sondern auch in ihrer Fläche (16,8 bzw. 18,6 km²) und ihrem Shape-Index (1,55 bzw. 1,59). Die Bevölkerungsdichte unterscheidet sich etwas stärker (140 bzw. 45 Einwohnende je Quadratkilometer), beide Werte liegen jedoch im unteren Bereich der Spannweite dieses Indikators.



Abbildung 3.11: Lage der Gemeinden Bothel und Spiekeroog in Niedersachsen.
(Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022)

Bothel hat eine Einpendelquote von 71% und eine Auspendelquote von 87%. Beide Werte liegen nahe des 1. Quartils des jeweiligen Pendelparameters. In Spiekeroog betragen Ein- und Auspendelquote nur 28% bzw. 29%, damit weist die Gemeinde jeweils den fünftniedrigsten Wert dieser Parameter auf - nur die anderen vier Inselgemeinden liegen

noch unter diesen Werten. Auch der Independence-Index liegt dadurch weit auseinander. In Bothel liegt dieser bei 0,11 und in Spiekeroog bei 1,26. Verursachend für diese Unterschiede kann auch die Arbeitsplatzausstattung sein, welche in Spiekeroog annähernd 100% beträgt, was jedoch nicht zwangsläufig zur Folge hat, dass die Zahl bzw. der Anteil der Binnenpendelnden hoch ist. In Bothel beträgt die Arbeitsplatzausstattung 46% und liegt damit etwa mittig zwischen Median und 3. Quartil.

Für Spiekeroog sind in den Pendeldaten keine konkreten Einpendelströme auf Gemeindeebene hinterlegt, weshalb keine durchschnittliche Einpendeldistanz für diese Gemeinde ermittelt werden konnte. Aggregiert lässt sich festhalten, dass von den 123 Einpendelnden auf die Insel 31 ebenfalls im Landkreis Wittmund wohnen, während 43 aus Niedersachsen, 18 aus Nordrhein-Westfalen, 27 aus dem restlichen Bundesgebiet und vier aus dem Ausland stammen. Die durchschnittliche Distanz wird damit vergleichsweise hoch liegen. Die durchschnittliche Auspendeldistanz ist mit etwa 340 Kilometer angegeben. Das liegt daran, dass der einzige hinterlegte Auspendelstrom auf Gemeindeebene nach Bonn führt, 25 der 126 Auspendelnden arbeiten dort. Die anderen Auspendelnden sind mit ihrem Arbeitsplatz im Landkreis Wittmund (17), in sonstigen Gemeinden in Niedersachsen (30), in sonstigen Gemeinden in Nordrhein-Westfalen (21) und in weiteren Bundesländern (33) registriert. Auch unter Berücksichtigung dieser Zielorte ist die durchschnittliche Auspendeldistanz vergleichsweise hoch. In Bothel liegen beide Distanzen bei durchschnittlich etwa 15 Kilometern. Mehr als zwei Drittel der Auspendelnden verbleiben im gleichen Landkreis, über 85% in Niedersachsen und etwa 10% sind mit ihrem Arbeitsplatz in Hamburg oder Bremen registriert. Die Anteile und Entfernungen unterscheiden sich damit deutlich von denen in Spiekeroog. Bei den Einpendelströmen sieht die Verteilung ähnlich aus.

03357006 Bothel			03462014 Spiekeroog		
02	Hamburg	43	03462	Wittmund	17
032	Statistische Region Hannover	13	034	Statistische Region Weser-Ems	42
03357	Rotenburg (Wümme)	644	03	Niedersachsen	47
03358	Heidekreis	59	05314	Bonn, Stadt	25
03361	Verden	67	053	Reg.-Bez. Köln	26
033	Statistische Region Lüneburg	792	057	Reg.-Bez. Detmold	10
03	Niedersachsen	814	05	Nordrhein-Westfalen	46
04	Bremen	46	ZZ	Übrige Bundesländer	33
05	Nordrhein-Westfalen	12	Z	Auspendler insgesamt	126
ZZ	Übrige Bundesländer	23			
Z	Auspendler insgesamt	938			

Abbildung 3.12: Die auf Kreis-, Regierungsbezirk und Länderebene aggregierten Auspendelströme der Gemeinden Bothel und Spiekeroog in Niedersachsen.

(Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

3.5.2 Vergleich einer nah und einer entfernt zum Oberzentrum liegenden Gemeinde

Der nächste Vergleich beinhaltet zwei Gemeinden in Mecklenburg-Vorpommern, welche sich in ihrer Entfernung zum nächstgelegenen Oberzentrum deutlich unterscheiden. Die Gemeinde Bütow im Landkreis Mecklenburgische Seenplatte liegt etwa 57 Kilometer Luftlinie südwestlich von Neubrandenburg, wobei die wahre Strecke aufgrund der Müritz und des Müritz-Nationalparks, welche sich beide in dieser Linie befinden, noch deutlich größer wird. Die Gemeinde Wackerow im Landkreis Vorpommern-Greifswald hingegen befindet sich zweieinhalb Kilometer nordwestlich von Greifswald. Beide Gemeinden sind Binnengemeinden und ähneln sich in ihrer Flächengröße (26 bzw. 32 km²) und ihrem Shape-Index (1,52 bzw. 1,44), die Bevölkerungsdichten unterscheiden sich absolut wie beim vorherigen Beispiel etwas stärker (17 bzw. 47 Einwohnende je Quadratkilometer), jedoch liegen beide Werte im unteren Bereich der Spannweite dieses Parameters.

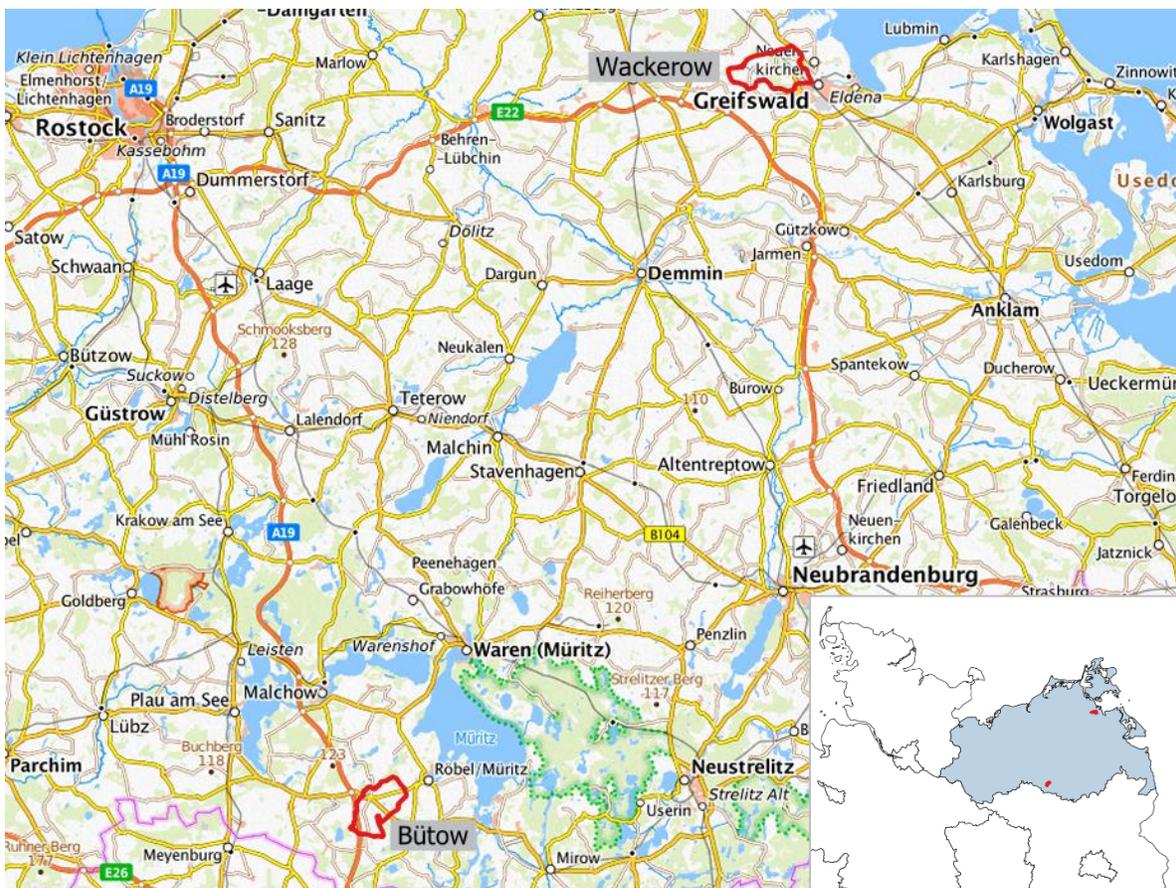


Abbildung 3.13: Lage der Gemeinden Bütow und Wackerow in Mecklenburg-Vorpommern.
(Eigener Screenshot in QGIS; © GeoBasis-DE / BKG 2022)

Bütow hat eine Einpendelquote von 67% und eine Auspendelquote von 81%. Die Werte liegen beide nahe des 1. Quartils dieser Parameter. Die Arbeitsplatzausstattung mit 59% liegt am 3. Quartil und der Independence-Index mit 0,15 über diesem. In Wackerow sind Ein- und Auspendelquote mit 76% bzw. 97% höher als in Bütow, die Einpendelquote liegt

knapp oberhalb des Medians und die Auspendelquote über dem 3. Quartil. Dadurch fällt der Independence-Index mit 0,03 sehr klein aus, der Wert liegt unterhalb des 1. Quartils. Verursachend dafür ist auch die sehr geringe Arbeitsplatzausstattung von 14%, auf 430 sozialversicherungspflichtig Beschäftigte kommen nur 62 besetzte Arbeitsplätze. Auch die durchschnittlichen Ein- und Auspendeldistanzen unterscheiden sich. In Wackerow werden durchschnittlich 2,5 Kilometer zum Einpendeln (die drittkleinste Entfernung aller Gemeinden in Mecklenburg-Vorpommern) und 7,7 Kilometer zum Auspendeln benötigt, auch letzterer Wert liegt unter dem 1. Quartil. Die kleine Einpendeldistanz ist darauf zurückzuführen, dass der einzige nicht anonymisierte Einpendelstrom auf Gemeindeebene aus Greifswald stammt, 20 der 47 Einpendelnden fahren aus dem Oberzentrum in diese Gemeinde. Von den übrigen Einpendelnden stammen weitere 13 aus dem gleichen Landkreis, elf aus Mecklenburg-Vorpommern und drei aus anderen Bundesländern. Damit wird ersichtlich, dass die korrekte durchschnittliche Einpendeldistanz etwas höher liegt, als der Wert vermuten lässt. Auch in Bütow gibt es nur einen nicht anonymisierten Einpendelstrom auf Gemeindeebene, elf der 83 Einpendelnden stammen aus Eldetal (6,7 Kilometer Entfernung), außerdem stammen 53 aus dem gleichen Landkreis, elf aus Mecklenburg-Vorpommern und acht aus anderen Bundesländern. Dadurch ist eine Aussage über die tatsächliche durchschnittliche Einpendeldistanz nur eingeschränkt möglich, jedoch sind die Einpendelnden deutlich weiter verstreut und legen größere Strecken zurück, als dies in Wackerow der Fall ist. Für die durchschnittliche Auspendeldistanz ist ähnliches zu beobachten, beide Werte liegen unter dem 1. Quartil, jedoch ist die Distanz in Wackerow mit 7,7 Kilometern nochmal kleiner als in Bütow mit 10,8 Kilometern. In Wackerow arbeiten 265 der 415 Auspendelnden im nahegelegenen Greifswald, in Bütow streuen auch die Auspendelströme deutlich stärker in viele unterschiedliche Gemeinden.

13071023 Bütow			13075141 Wackerow		
13071175	Eldetal	11	13075039	Greifswald, Universitäts- und Hansestadt	20
13071	Übrige Gemeinden (Kreis)	53	13075	Übrige Gemeinden (Kreis)	13
13071	Mecklenburgische Seenplatte	64	13075	Vorpommern-Greifswald	33
130	Übrige Kreise (Regierungsbezirk)	11	130	Übrige Kreise (Regierungsbezirk)	11
130	Mecklenburg-Vorpommern	75	130	Mecklenburg-Vorpommern	44
13	Mecklenburg-Vorpommern	75	13	Mecklenburg-Vorpommern	44
ZZ	Übrige Bundesländer / Staaten	8	ZZ	Übrige Bundesländer / Staaten	3
ZZ	Einpendler aus dem Bundesgebiet	83	ZZ	Einpendler aus dem Bundesgebiet	47
Z	Einpendler insgesamt	83	Z	Einpendler insgesamt	47

Abbildung 3.14: Die Einpendelströme der Gemeinden Bütow und Wackerow in Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

3.5.3 Vergleich einer Gemeinde in Niedersachsen mit einer in Mecklenburg-Vorpommern
Abschließend werden zwei Gemeinden verglichen, welche in allen räumlichen Indikatoren ähnliche Werte aufweisen, von welchen jedoch eine in Mecklenburg-Vorpommern und eine in Niedersachsen liegt. Die Gemeinde Gotthun liegt im Landkreis Mecklenburgische Seenplatte an der Müritz und etwa 47 Kilometer südwestlich von Neubrandenburg, die Gemeinde Lahn liegt im Landkreis Emsland etwa 49 Kilometer östlich von Emmen in den Niederlanden. Beide Gemeinden sind Binnengemeinden und ähneln sich ebenfalls in ihrem Shape-Index (1,22 zu 1,14). Ihre Fläche (10 bzw. 21 km²) und Bevölkerungsdichte (32 bzw. 41 Einwohnende je Quadratkilometer) unterscheiden sich etwas, alle Werte liegen jedoch im unteren Bereich der Spannweite der Indikatoren.



Abbildung 3.15: Lage der Gemeinden Lahn in Niedersachsen (links) und Gotthun in Mecklenburg-Vorpommern.
(Eigene Screenshots in QGIS; © GeoBasis-DE / BKG 2022)

Die Pendelparameter der beiden Gemeinden werden in Tabelle 3.12 gegenübergestellt. Die niedersächsische Gemeinde Lahn hat eine niedrigere Arbeitsplatzausstattung, eine höhere Ein- und Auspendelquote und damit einen niedrigeren Independence-Index als Gotthun in Mecklenburg-Vorpommern. Genau wie die Pendelquoten sind auch die durchschnittlichen Ein- und Auspendeldistanzen in Lahn größer.

	Lahn (Niedersachsen)	Gotthun (Mecklenburg-Vorpommern)
Arbeitsplatzausstattung (%)	40,2	67,8
Einpendelquote (%)	80,9	76,3
Auspendelquote (%)	92,3	83,9
Independence-Index	0,06	0,12
Einpendeldistanz (km)	6	4,5
Auspendeldistanz (km)	10,8	7,8

Tabelle 3.12: Die Pendelparameter von Lahn und Gotthun im Vergleich.

In Gotthun verbleiben 102 der 120 Auspendelnden im Landkreis Mecklenburgische Seenplatte, sechs pendeln in andere Kreise Mecklenburg-Vorpommerns und zwölf in andere Bundesländer aus. In Lahn ist der Anteil der Auspendelnden in den eigenen Landkreis mit 291 von 361 Auspendelnden etwas kleiner, 56 sozialversicherungspflichtig Beschäftigte pendeln in andere Kreise Niedersachsens und 14 in andere Bundesländer aus. Für die Einpendelströme gilt das gleiche, jedoch muss dabei auch berücksichtigt werden, dass der Landkreis Mecklenburgische Seenplatte fast doppelt so groß ist wie der Landkreis Emsland und auch mehr als doppelt so viele Gemeinden beinhaltet, weshalb eine allgemeine Äußerung zu den Pendeldistanzen schwierig zu treffen ist.

03454027 Lahn		13071045 Gotthun	
03453011	Lönigen, Stadt	13071124	Röbel/Müritz, Stadt
03453	Übrige Gemeinden (Kreis)	13071156	Waren (Müritz), Stadt
03453	Cloppenburg	13071	Übrige Gemeinden (Kreis)
03454019	Haselünne, Stadt	13071	Mecklenburgische Seenplatte
03454023	Hüven	130	Übrige Kreise (Regierungsbe:
03454024	Klein Berßen	130	Mecklenburg-Vorpommern
03454026	Lähden	13	Mecklenburg-Vorpommern
03454035	Meppen, Stadt	ZZ	Übrige Bundesländer
03454041	Papenburg, Stadt	ZZ	Auspendler in das Bundesge:
03454047	Sögel	Z	Auspendler insgesamt
03454057	Werite, Stadt		
03454	Übrige Gemeinden (Kreis)		
03454	Emsland		
034	Übrige Kreise (Regierungsbezirk)		
034	Statistische Region Weser-Ems		
03	Übrige Regierungsbezirke (Bundesl		
03	Niedersachsen		
ZZ	Übrige Bundesländer		
ZZ	Auspendler in das Bundesgebiet		
Z	Auspendler insgesamt		

Abbildung 3.16: Die Auspendelströme der Gemeinden Lahn in Niedersachsen und Gotthun in Mecklenburg-Vorpommern. (Eigener Screenshot in Excel; © Statistik der Bundesagentur für Arbeit 2022 b)

4. Diskussion und Fazit

Die Ergebnisse der Untersuchung lassen sich wie folgt zusammenfassen:

1. **Die Ein- und Auspendelquoten sind auf Gemeindeebene durchweg sehr hoch.** Mittelwert und Median der Einpendelquote liegen in Mecklenburg-Vorpommern beide bei etwa 73% und in Niedersachsen bei etwa 75,5% bzw. 76,4%. Von den 674 untersuchten Gemeinden in Mecklenburg-Vorpommern konnte in 557 Gemeinden aus der Datengrundlage eine Einpendelquote berechnet werden, in 20 davon liegt diese unter 50%. In Niedersachsen liegt für 568 der 603 untersuchten Gemeinden der Wert vor, in zwölf davon liegt dieser unter 50%. Bei der Auspendelquote sind die Werte noch höher, Mittelwert und Median betragen hier 88,8% bzw. 91,7% (Mecklenburg-Vorpommern) sowie 90,2% bzw. 92,3% (Niedersachsen). Auspendelquoten unter 50% gibt es in Mecklenburg-Vorpommern in vier Gemeinden (639 vorhandene Werte), in Niedersachsen in sechs Gemeinden (587 vorhandene Werte).
2. **Es gibt demzufolge kaum Binnenpendelnde, dementsprechend fällt auch der Independence-Index zumeist niedrig aus.** In Niedersachsen gibt es bei diesem sechs Werte über eins, der siebthöchste liegt bei 0,37 und damit schon deutlich niedriger. In Mecklenburg-Vorpommern gibt es keine Werte größer als eins, der höchste liegt bei 0,94 und die vier nächsthöchsten zwischen 0,37 und 0,47 (der Independence-Index ist bei allen Gemeinden vorhanden, bei welchen auch die Auspendelquote berechnet werden konnte). In 442 Gemeinden Mecklenburg-Vorpommerns und in 463 Gemeinden Niedersachsens liegt der Independence-Index bei einem Wert unter 0,1.
3. **Die niedrigen Werte beim Independence-Index liegen nicht an einer zu niedrigen (oder in wenigen Gemeinden zu hohen) Arbeitsplatzausstattung.** Auch mit einer Arbeitsplatzausstattung von 10%, welche in Mecklenburg-Vorpommern nur 23 und in Niedersachsen nur 28 Gemeinden unterschreiten, ist ein Independence-Index von maximal 0,11 möglich. Ein- und Auspendelquote sind in fast allen Gemeinden deutlich höher, als die Zahl an wohnhaften sozialversicherungspflichtig Beschäftigten sowie Arbeitsstellen bedingt.
4. **Die Richtung der Korrelationen zwischen den räumlichen Indikatoren und den Pendelquoten entspricht mit einer Ausnahme der auf NUTS-Ebene 2.** Nur die Korrelation zwischen Bevölkerungsdichte und Auspendelquote in Mecklenburg-Vorpommern ist negativ, obwohl eine positive angenommen wurde. Bei der Fläche, dem Küstenanteil und der Entfernung zum nächstgelegenen Oberzentrum treten die erwarteten negativen Korrelationen mit Ein- und Auspendelquote auf.
5. **Anders als erwartet treten zwischen dem Shape-Index und den Pendelquoten negative Korrelationen auf.** Das Ergebnis wiederholt sich ebenfalls, wenn die Korrelationsberechnung nur für die Binnengemeinden durchgeführt wird, so dass eine Beeinflussung durch den Küstenanteil auf die Richtung der Korrelation nicht vorliegt.

6. **Die berechneten Korrelationen zwischen den räumlichen Indikatoren und den Pendelparametern sind zu großen Teilen sehr gering bis nicht vorhanden.** Das ist nach Hinkle et al. (1990) der Fall, wenn der berechnete Korrelationskoeffizient einen Wert kleiner gleich 0,3 annimmt, was unabhängig vom Ausschluss der Ausreißer bei 47 der 60 Korrelationen der Fall ist. Ohne Ausschluss der Ausreißer weisen zehn Werte eine geringe Korrelation auf (0,3 bis kleiner gleich 0,5), ein Wert eine mittlere Korrelation (0,5 bis kleiner gleich 0,7; Küstenanteil und Auspendeldistanz in Niedersachsen) und zwei Werte eine hohe Korrelation (0,7 bis kleiner gleich 0,9; Küstenanteil und Auspendelquote bzw. Independence-Index in Niedersachsen). Mit Ausschluss von Ausreißern ergeben sich bei zwölf Werten geringe und bei einem Wert eine mittlere Korrelation (Fläche und Independence-Index in Niedersachsen).
- a. **Ausschließlich sehr geringe bzw. keine Korrelation kann bei der Bevölkerungsdichte nachgewiesen werden.** Die Anzahl der Einwohnenden einer Gemeinde im Verhältnis zur Größe hat also keinen erkennbaren Einfluss auf die Anteile bzw. zurückgelegten Distanzen der Ein- und Auspendelnden.
 - b. **Der Shape-Index weist in Mecklenburg-Vorpommern geringe Korrelationen zur Auspendelquote (-0,34) und zum Independence-Index auf (0,38),** wenn die Ausreißer nicht ausgeschlossen werden. Auch bei den anderen Pendelparametern ist der Einfluss der formbezogenen Maßzahl in Mecklenburg-Vorpommern größer als in Niedersachsen.
 - c. **Die Flächengröße weist eine geringe Korrelation zu Ein- und Auspendelquote, Independence-Index sowie Auspendeldistanz auf.** Ohne Ausschluss der Ausreißer nur in Mecklenburg-Vorpommern, mit Ausschluss in beiden Bundesländern und beim Independence-Index fällt die Korrelation in Niedersachsen knapp in den mittleren Bereich (0,52).
 - d. **Die Entfernung zum nächsten Oberzentrum weist eine geringe Korrelation zur Einpendelquote (0,33) und zum Independence-Index (0,32) in Mecklenburg-Vorpommern auf,** letztgenanntes nur nach Ausschluss der Ausreißer.
 - e. **Der Küstenanteil weist in Mecklenburg-Vorpommern eine geringe Korrelation zur Auspendelquote (-0,44) und zum Independence-Index (0,4) auf. In Niedersachsen sind diese Korrelationen hoch (-0,71 bzw. 0,86), zusätzlich gibt es eine geringe zur Einpendelquote (-0,46) und eine mittlere zur Auspendeldistanz (0,6).** Werden in Niedersachsen die Ausreißer bei den Pendelparametern ausgeschlossen, fallen die Korrelationen in den sehr geringen Bereich. Dort sind die Werte aufgrund der Datenlage beim Küstenanteil (nur 17 der 603 Gemeinden sind Küstengemeinden) nicht sehr aussagekräftig. Trotzdem bestätigt sich, dass die Lage einer Gemeinde das Pendelverhalten beeinträchtigt und Unterschiede zwischen Binnen-, Küsten- und Inselgemeinden bestehen.

Trotz aller Einschränkungen, welche die verwendeten Daten mit sich bringen, konnte das Pendelverhalten in den Gemeinden abseits der Städte gut dargestellt werden. Dass es viele Auspendelnde in die größeren Städte gibt, ist keine Überraschung. Unerwarteter sind da die ebenfalls konstant hohen Einpendelquoten. Etwa drei von vier Arbeitsplätze in den Gemeinden mit weniger als 5000 Einwohnenden werden von Einpendelnden besetzt. Damit bestätigt sich auch für diese Ebene, was Siedentop (2007) bereits für westdeutsche Agglomerationsräume gezeigt hat - auch wenn Arbeitsplätze im suburbanen oder ländlichen Raum zur Verfügung stehen, bewirkt das keinen geringeren Pendelverkehr. Es entsteht eine große Struktur an Personen, welche in diesen Gegenden „kreuz und quer“ pendeln, um ihre jeweiligen Arbeitsstellen zu erreichen.

Die Untersuchung hat gezeigt, dass sich die Erkenntnisse auf NUTS-Ebene 2 nur bedingt auf die Verwaltungsebene der Gemeinden in Deutschland übertragen lassen, zumindest für die Gemeinden Mecklenburg-Vorpommerns und Niedersachsens mit weniger als 5000 Einwohnenden unter Verwendung der Daten des Statistik-Service der Bundesagentur für Arbeit. Die Einflüsse der räumlichen Faktoren, zumindest der Flächengröße, der Entfernung zum Oberzentrum und des Küstenanteils, lassen sich zwar teilweise erkennen, sind aber aufgrund des allgemein sehr hohen Pendelaufkommens nicht die entscheidenden Faktoren, ob in einer Gemeinde niedrige oder hohe Pendelquoten vorherrschen. Einzig die Insellage einer Gemeinde verursacht einen spürbaren Rückgang des Pendelns aufgrund der deutlich höheren Kosten, die bei einer Insel zur Überquerung der Gemeindegrenzen notwendig sind.

Die Vorgehensweise der Untersuchung lässt sich gut auf andere Untersuchungsgebiete übertragen, so dass einer Wiederholung z.B. für andere Bundesländer theoretisch möglich ist. Einzig der Küstenanteil als räumlicher Einflussfaktor fällt in einem Großteil dieser weg, kann aber durch andere Faktoren ersetzt werden, welche die topographische Lage einer Gemeinde bzw. die Verbindung zu anderen Gemeinden beschreibt. Dazu können z.B. natürliche Bedingungen wie die Hangneigung zählen, aber auch Variablen der Verkehrsinfrastruktur. Wesentlich andere Ergebnisse sollten auch in den anderen Bundesländern Deutschlands nicht erwartet werden, ein Vergleich zur Gemeindeebene anderer Staaten jedoch bietet sich an. Die Begrenzung der Gemeinden auf diejenigen mit maximal 5000 Einwohnenden sollte überdacht werden, wenn dadurch zu viele Gemeinden aus der Untersuchung herausfallen, wie es in Niedersachsen der Fall ist. Je nach Bundesland bzw. Untersuchungsgebiet sind andere Obergrenzen der Bevölkerungszahl praktikabel. Bei der Obergrenze von 10.000 Einwohnenden wären in Niedersachsen 733 (statt 603) der 944 Gemeinden in die Untersuchung gefallen. Der Ausschluss von bevölkerungsreichen Städten erscheint jedoch weiterhin sinnvoll, da diese zu Ausreißern in den Statistiken führen würden. Beachtet werden muss außerdem, dass die untersuchten Gebietseinheiten ähnliche Flächengrößen aufweisen sollten, um einem Modifiable Areal Unit Problem (MAUP) vorzubeugen. Bei sehr großen, mglw. aus mehreren Ortsteilen zusammengesetzten

Gemeinden, wie sie auch in den untersuchten Bundesländern vereinzelt vorkommen, müssen teilweise größere Distanzen und mehrere zur gleichen Gemeinde gehörende Orte überwunden werden, um nach Definition zu pendeln. Zur Vermeidung dieser Problematik können generierte Gebiete (z.B. Quadrate, Hexagone) statt der Verwaltungseinheiten genutzt werden, jedoch müssten dafür die Pendelraten auf diese transformiert werden.

Das Fazit für die Verwendung der Pendelraten des Statistik-Service der Bundesagentur für Arbeit fällt gemischt aus. Wie in der Einleitung bereits aufgeführt, werden in diesen Statistiken als Pendelnde alle sozialversicherungspflichtig Beschäftigten bezeichnet, deren Arbeitsort sich vom Wohnort (auf Gemeindeebene) unterscheidet, wobei es unerheblich ist, wie oft bzw. ob überhaupt gependelt wird (Bundesagentur für Arbeit 2020). Deshalb muss man sich bei der Arbeit mit den Statistiken bewusst machen, was genau diese aussagen. Es wird nicht die Anzahl der tatsächlichen Bewegungen zwischen den Gemeinden erfasst, sondern nur die Anzahl der Personen, welche in einer Gemeinde wohnen und deren Arbeitsstätte sich in einer anderen Gemeinde befindet. Ob die Arbeitnehmenden an jedem Arbeitstag pendeln, wöchentlich z.B. zu einem Zweitwohnsitz nahe der Arbeitsstätte oder nur vereinzelt bzw. gar nicht (z.B. Home-Office), ist für die Statistik irrelevant. Auch werden nur sozialversicherungspflichtig Beschäftigte erfasst, wodurch Verbeamtete, Selbstständige und ausschließlich geringfügig Beschäftigte nicht von der Statistik erfasst werden (Industrie- und Handelskammer zu Düsseldorf 2022). So werden z.B. von den etwa 1,6 Millionen Einwohnenden Mecklenburg-Vorpommerns knapp 618.000 als sozialversicherungspflichtig beschäftigt erfasst (Statistik der Bundesagentur für Arbeit 2022 a, Tabellenblatt „Länder“).

Die reinen Pendelwerte (Anzahl der sozialversicherungspflichtig Beschäftigten mit Wohnort bzw. Arbeitsort pro Gemeinde, Anzahl der Ein-, Aus und Binnenpendelnden) haben sich trotz der besprochenen Einschränkungen gut verwenden lassen, auch die Anonymisierungen waren hier nicht sehr störend. Damit konnten zwar die Pendelparameter wie Einpendelquote nicht für alle Gemeinden bestimmt werden, aber die Datendichte ist ausreichend. Anders sieht es bei den Pendelströmen aus. Hier sorgen die Anonymisierungen bzw. Aggregationen auf höhere Verwaltungsebenen nicht nur für fehlende Werte bei den Pendeldistanzen, sondern die vorhandenen sind auch sehr ungenau. Da Pendelströme auf Gemeindeebene nur aufgezählt werden, wenn mind. 10 Personen diesen nutzen, werden viele Pendelströme, gerade bei den Einpendelströmen, durch die Aggregation so ungenau, dass die Pendeldistanz nicht sinnvoll verwendet werden kann. Ein Extrembeispiel dafür ist die in Abschnitt 3.5.1 besprochene Gemeinde Spiekeroog, bei welcher eine durchschnittliche Auspendeldistanz von etwa 340 Kilometern ermittelt wird, weil der einzig auf Gemeindeebene angegebene Auspendelstrom nach Bonn führt. Genauso tritt das für viele Gemeinden im genau umgekehrten Fall auf, Pendelströme in die Nachbargemeinden sind angegeben, aber diejenigen in weiter entfernte Gemeinde sind oftmals zu dünn besetzt und nur aggregiert vorhanden - eine zu kleine Pendeldistanz ist die

Folge. Hier wäre es von Vorteil, den vollständigen Datensatz auf Gemeindeebene ohne Anonymisierungen und Aggregationen zur Verfügung zu haben, denn die Aussagekraft der Ein- und Auspendeldistanz ist so zu ungenau für eine sinnvolle Verwendung. Auch die mit den Pendeldistanzen ermittelten Korrelationen sind dadurch nicht gut verwendbar.

Eine dynamische Alternative stellt die Verwendung von Mobilitätsindikatoren auf Grundlage von Mobilfunkdaten dar. Ein Beispiel sind die vom Unternehmen Teralytics aufbereiteten und zur Verfügung gestellten Daten, welche von Mobilfunkgeräten stammen, die aufgrund einer Bewegung die Funkzelle wechseln. Die Aufbereitung und Verwendung solcher Daten für Untersuchungen wie diese ist mit einem deutlich größerem Aufwand verbunden, so muss z.B. zunächst erkannt werden, unter welchen Voraussetzungen es sich wirklich um eine Arbeitspendel-Bewegung und nicht um eine Freizeit-Bewegung handelt. Mit den Resultaten werden sich dafür aber auch deutlich bessere Aussagen über das Pendelverhalten machen lassen können. Ein Vergleich der statischen Pendeldaten der Bundesagentur und der dynamischen Mobilitätsdaten von Teralytics oder anderen Anbietern kann aufzeigen, inwiefern erstere auch Mobilität darstellen.

In dieser Master-Thesis wurde sich auf die Korrelation von räumlichen Faktoren der Verwaltungsstrukturen zu den Pendelparametern konzentriert. Natürlich lässt sich das Pendelverhalten auch in Korrelation zu vielen anderen möglichen Einflussfaktoren setzen. Das können Faktoren sein, welche die Menschen einer Verwaltungseinheit selbst beschreiben, z.B. das Alter oder das Einkommen, aber auch subjektiv abfragbare Punkte wie die Pendelbereitschaft oder die Verbundenheit zum Wohnort bzw. zur Arbeitsstelle. Als weiteres Teilgebiet können Faktoren wie die durchschnittlichen Grundstückspreise hinzukommen. Auch die statistische Auswertung kann weiter ausgebaut werden. In dieser Untersuchung wurde ausschließlich auf lineare Korrelation untersucht, jedoch wird diese Methode durch die teilweise auftretende Heteroskedastizität (unterschiedliche Streuungen der Y-Werte für verschiedene X-Werte, vgl. z.B. Abbildung 3.8) eingeschränkt. Mögliche Erweiterungen des Verfahrens wären eine Regression mit Bootstrapping oder eine Transformation der Variablen (Regorz 2021), aber auch die Verwendung anderer Regressionsverfahren, z.B. eine nichtparametrische Regression.

Das Themenfeld ist weit und lässt sich durch einzelne Arbeiten oder Untersuchungen nicht abdecken, denn schließlich ist das tägliche Pendeln nach Hartke (1938) ein umfassender geographischer Problemkomplex, welcher „nur im größeren Problemzusammenhang diskutiert [...] [und] nur im Zusammenhang mit allen anderen Erscheinungen der sozialräumlichen Mobilität untersucht werden“ (Klingbeil 1969) kann.

Literaturverzeichnis

Monografien, Sammelwerke, Zeitschriftenaufsätze

- Blaschke, Thomas; Lang, Stefan (2007): *Landschaftsanalyse mit GIS*. Stuttgart.
- Buch, Tanja; Carstensen, Jeanette; Hamann, Silke; Otto, Anne; Seibert, Holger; Sieglén, Georg (2020): *Grenzpendler aus dem Ausland. Immer mehr Beschäftigte in Deutschland mit ausländischem Wohnort*. In: *IAB-Kurzbericht* (9/2020). Nürnberg.
- Cervero, Robert (1996): *Jobs-Housing Balance Revisited. Trends and Impacts in the San Francisco Bay Area*. In: *Journal of the American Planning Association* (62. Jahrgang, Heft 4). S. 492-511. Chicago.
- Gatersleben, Birgitta; Uzzell, David (2007): *Affective Appraisals of the Daily Commute. Comparing Perceptions of Drivers, Cyclists, Walkers, and Users of Public Transport*. In: *Environment and Behavior* (39. Jahrgang, Heft 3). S. 416-431. Surrey.
- Haas, Anette; Hamann, Silke (2008): *Pendeln – ein zunehmender Trend, vor allem bei Hochqualifizierten. Ost-West-Vergleich*. In: *IAB-Kurzbericht* (6/2008). Nürnberg.
- Hartke, Wolfgang (1938): *Das Arbeits- und Wohnortsgebiet im Rhein-Mainischen Lebensraum. Untersuchungen über Grundlagen der Kultur- und Wirtschaftsgeographie und ihren Raumbegriff am besonderen Beispiel der Pendelwanderung*. In: *Rhein-Mainische Forschungen* (18). Frankfurt am Main.
- Hinkle, Dennis E.; Jurs, Stephen G.; Wiersma, William (1990): *Applied Statistics for the Behavioral Sciences*. In: *Journal of Educational Statistics* (1/1990). Washington DC.
- Janke, Ralf-Peter (2017): *Entgelte von Pendlern – Aussagen zur Mobilität von Beschäftigten am Beispiel des Berliner Umlandes*. Nürnberg.
- Kaewwongwattana, Peerakan; Phimolsathien, Thepparat; Pimdee, Paitoon (2015): *Determinants Of Consumer Decision Making Of A Common Ticketing System In Bangkok's Metropolitan Commuter Transportation Systems*. In: *The Journal of Applied Business Research* (31. Jahrgang, Heft 6). S. 2025-2038. Littleton.
- Klingbeil, Detlev (1969): *ZUR SOZIALGEOGRAPHISCHEN THEORIE UND ERFASSUNG DES TÄGLICHEN BFRUFSPENDELNS(sic!)*. In: *Geographische Zeitschrift* (57. Jahrgang, Heft 2). S. 108–131. Stuttgart.
- Lang, Robert E.; Nelson, Arthur C. (2007): *Beyond the Metroplex: Examining Commuter Patterns at the „Megapolitan“ Scale*. Cambridge.
- Le Néchet, Florent (2012): *Approche multiscalaire de la mobilité domicile-travail en Île-de-France et dans la région Rhin-Ruhr*. In: *Cahiers de géographie du Québec* (56. Jahrgang, Heft 2). S. 405-426. Québec.
- Nitzschke, Luca; Quentin, Paula; Kanisius, Fabian; Schluckebier, Kai; Burlon, Nora Sofie; Buscher, Jost; Deffner, Jutta; Bruns, André; Stein, Melina; Mühlhans, Heike; Othengrafen, Frank; Joost, Jan-Marc (2022): *Pendeln verstehen: Status quo, Forschungsstand und Perspektiven*. In: *ISOE-Materialien Soziale Ökologie* (67). Frankfurt am Main.
- Sakanishi, Akiko (2006): *Commuting patterns in the Osaka metropolitan area: A GIS-based analysis of commuter rail passengers*. In: *review of urban & regional development studies* (18. Jahrgang, Heft 1). S. 41-59. Tokio.

- Schäfer, Thomas (2016): Methodenlehre und Statistik. Einführung in Datenerhebung, deskriptive Statistik und Inferenzstatistik. Wiesbaden.
- Schwanen, Tim; Dieleman, Frans M.; Dijst, Martin (2001): Travel behaviour in Dutch monocentric and policentric urban systems. In: Journal of Transport Geography (9. Jahrgang, Heft 3). S. 173-186. Amsterdam.
- Siedentop, Stefan (2007): Auswirkungen der Beschäftigungssuburbanisierung auf den Berufsverkehr. Führt die Suburbanisierung der Arbeitsplätze zu weniger Verkehr? In: Informationen zur Raumentwicklung (10. Jahrgang, Heft 2). S. 105-124. Bonn.
- Thomas, Ray (1969): London's New Towns. A Study of Self-contained and Balanced Communities. In: Political & Economic Planning (Ausgabe 510). London.

Internetquellen

- Antenne Münster (2021): Pendlerstatistik in NRW: Die meisten Menschen pendeln in die großen Städte. Münster. Online abrufbar unter <https://www.antennemuenster.de/artikel/pendlerstatistik-in-nrw-die-meisten-menschen-pendeln-in-die-grossen-staedte-1156064.html>. (zuletzt aufgerufen am 31.08.2022)
- Becker, Martin (2022): Symmetrie- und Wölbungsmaße. In: Deskriptive Statistik und Wahrscheinlichkeitsrechnung. Saarbrücken. Online abrufbar unter: <https://www.lehrstab-statistik.de/online/Deskriptive-WR/Vorlesung/3-6-symmetrie-und-w%C3%B6lbungsma%C3%9F.html>. (zuletzt aufgerufen am 26.02.2023)
- Brinkhoff, Thomas (2021): City Population. Polen - Größere Städte. Oldenburg. Online abrufbar unter <http://www.citypopulation.de/de/poland/cities/>. (zuletzt aufgerufen am 13.02.2023)
- Brinkhoff, Thomas (2023): City Population. Niederlande - Verwaltungsgliederung. Oldenburg. Online abrufbar unter <http://www.citypopulation.de/de/netherlands/admin/>. (zuletzt aufgerufen am 13.02.2023)
- Bundesagentur für Arbeit (2020): Methodische Hinweise zum Thema Beschäftigung. Nürnberg. Online abrufbar unter <https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Methodik-Qualitaet/Methodische-Hinweise/BST-Meth-Hinweise/BST-Meth-Hinweise-Nav.html>. (zuletzt aufgerufen am 23.07.2022)
- Bundesministerium des Innern und für Heimat (2020): Deutschlandatlas - Pendeldistanzen und Pendlerverflechtungen. Berlin. Online abrufbar unter <https://www.deutschlandatlas.bund.de/DE/Karten/Wie-wir-uns-bewegen/100-Pendlerdistanzen-Pendlerverflechtungen.html>. (zuletzt aufgerufen am 11.08.2022)
- Büttner, Lisa; Breitkreuz, Anna (2020): Arbeiten nach Corona. Warum Homeoffice gut fürs Klima ist. Berlin. Online abrufbar unter <https://www.greenpeace.de/publikationen/Warum-Homeoffice-gut-fuers-Klima-ist.pdf>. (zuletzt aufgerufen am 11.09.2022)
- Die Bundesregierung (2019): Klimaschutzprogramm 2030. Berlin. Online abrufbar unter <https://www.bundesregierung.de/breg-de/themen/klimaschutz/klimaschutzprogramm-2030-1673578>. (zuletzt aufgerufen am 11.09.2022)

- Eurostat (2018): Statistiken zu Pendlerverflechtungen auf regionaler Ebene. Luxemburg. Online abrufbar unter https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Statistics_on_commuting_patterns_at_regional_level/de&oldid=401615. (zuletzt aufgerufen am 04.09.2022)
- Fischer, Konrad; Book, Simon (2021): Pendler betrügen sich selbst. Düsseldorf. Online abrufbar unter <https://www.wiwo.de/erfolg/beruf/arbeitsweg-pendler-betruegen-sich-selbst/20560038.html>. (zuletzt aufgerufen am 26.08.2022)
- Grellmann, Martin (o.J.): Schiefe in der Statistik. Dresden. Online abrufbar unter <https://martin-grellmann.de/schiefe-in-der-statistik>. (zuletzt aufgerufen am 18.03.2023)
- Hamberger, Beatrice (2021): Pendeln kostet Zeit – und Nerven. Hamburg. Online abrufbar unter <https://www.tk.de/techniker/magazin/life-balance/balance-im-job/pendeln-kostet-zeit-und-nerven-2048874?tkcm=ab>. (zuletzt aufgerufen am 26.08.2022)
- Industrie- und Handelskammer zu Düsseldorf (2022): Beschäftigung (sozialversicherungspflichtig). Düsseldorf. Online abrufbar unter <https://www.ihk.de/duesseldorf/standort/regionalwirtschaft/wirtschaftsdaten/beschaefigung-sozialversicherungspflichtig-2593156>. (zuletzt aufgerufen am 24.07.2022)
- Leipziger Zeitung (2019): Zahl der Berufspendler nach Leipzig strebt der 100.000er-Marke entgegen. Leipzig. Online abrufbar unter <https://www.liz.de/politik/kassensturz/2019/10/Zahl-der-Berufspendler-nach-Leipzig-strebt-der-100-000er-Marke-entgegen-300091>. (zuletzt aufgerufen am 31.08.2022)
- Märkische Allgemeine (2021): Immer mehr Brandenburger pendeln nach Berlin. Hannover. Online abrufbar unter <https://www.maz-online.de/brandenburg/immer-mehr-brandenburger-pendeln-nach-berlin-l65MCN6Z6X6XO3A4EG25WNOZ7A.html>. (zuletzt aufgerufen am 31.08.2022)
- Ministerium für Energie, Infrastruktur und Landesentwicklung Mecklenburg-Vorpommern (2016): Landesraumentwicklungsprogramm Mecklenburg-Vorpommern. Schwerin. Online abrufbar unter <https://www.regierung-mv.de/serviceassistent/download?id=1576266>. (zuletzt aufgerufen am 17.07.2022)
- Niedersächsische Staatskanzlei (o.J.): Niedersachsen in Zahlen. Hannover. Online abrufbar unter https://www.niedersachsen.de/startseite/land_leute/das_land/zahlen_fakten/niedersachsen-in-zahlen-20094.html. (zuletzt aufgerufen am 21.01.2023)
- Niedersächsisches Ministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2017): Landes-Raumordnungsprogramm Niedersachsen 2017. Hannover. Online abrufbar unter <https://www.ml.niedersachsen.de/download/157519>. (zuletzt aufgerufen am 21.01.2023)
- Nobis, Claudia; Kuhnimhof, Tobias (2018): Mobilität in Deutschland – MiD Ergebnisbericht. Bonn, Berlin. Online abrufbar unter http://www.mobilitaet-in-deutschland.de/pdf/MiD2017_Ergebnisbericht.pdf. (zuletzt aufgerufen am 11.09.2022)
- Regorz, Arndt (2021): Regressionsvoraussetzungen: Homoskedastizität (Varianzhomogenität). Bochum. Online abrufbar unter http://www.regorz-statistik.de/inhalte/tutorial_regression_homoskedastizitaet.html. (zuletzt aufgerufen am 11.06.2023)

- Staatskanzlei des Landes Mecklenburg-Vorpommern (o.J.): Mecklenburg-Vorpommern. Daten und Fakten. Schwerin. Online abrufbar unter <https://www.mecklenburg-vorpommern.de/ueber-das-land-zum-leben/daten-und-fakten>. (zuletzt aufgerufen am 17.07.2022)
- Statistisches Bundesamt (2022 a): Mikrozensus 2020 - Erwerbstätige nach Stellung im Beruf, Entfernung, Zeitaufwand und benutztem Verkehrsmittel für den Hinweg zur Arbeitsstätte 2020 in %. Wiesbaden. Online abrufbar unter <https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Erwerbstaetigkeit/Tabellen/pendler1.html?nn=250572>. (zuletzt aufgerufen am 08.08.2022)

Daten, Software und mehr

- Bundesamt für Kartographie und Geodäsie (2022): WFS Verwaltungsgebiete 1:250 000, Stand 01.01. (wfs_vg250). Datensatz „Gemeinden“ und „Gemeindepunkte“. Frankfurt am Main. Online herunterladbar unter <https://gdz.bkg.bund.de/index.php/default/wfs-verwaltungsgebiete-1-250-000-stand-01-01-wfs-vg250.html>. (zuletzt aufgerufen am 17.11.2022)
- Ellison, Chucky (2018): PostgreSQL Stats Aggregate (stats_agg). New York. Online abrufbar unter <https://github.com/ellisonch/PostgreSQL-Stats-Aggregate>. (zuletzt aufgerufen am 26.02.2023)
- Ghetta, Matteo (2022): Data Plotly. In: QGIS Python Plugins Repository. Pontedera. Online abrufbar unter <https://plugins.qgis.org/plugins/DataPlotly/>. (zuletzt aufgerufen am 26.02.2023)
- Statistik der Bundesagentur für Arbeit (2022 a): Tabellen „Gemeindedaten der sozialversicherungspflichtig Beschäftigten nach Wohn- und Arbeitsort“, Stichtag 30.6.2021. Nürnberg. Online herunterladbar unter https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?nn=1523064&topic_f=beschaeftigung-sozbe-gemband (Eintrag „2021 – Deutschland, Länder und Kreise“). (zuletzt aufgerufen am 26.07.2022)
- Statistik der Bundesagentur für Arbeit (2022 b): Sozialversicherungspflichtig Beschäftigte – Pendlerverflechtungen nach Gemeinden, Stichtag 30.6.2021. Hannover.
- Statistisches Bundesamt (2022 b): Gemeindeverzeichnis. Alle politisch selbstständigen Gemeinden mit ausgewählten Merkmalen am 30.09.2022 (3. Quartal 2022). Wiesbaden. Online herunterladbar unter <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugQ/AuszugGV3QAktuell.html>. (zuletzt aufgerufen am 08.10.2022)

Für Abbildungen mit dem verkürzten Verweis „© OpenStreetMap“ gilt:

© OpenStreetMap contributors

www.openstreetmap.org

Lizenz: www.openstreetmap.org/copyright