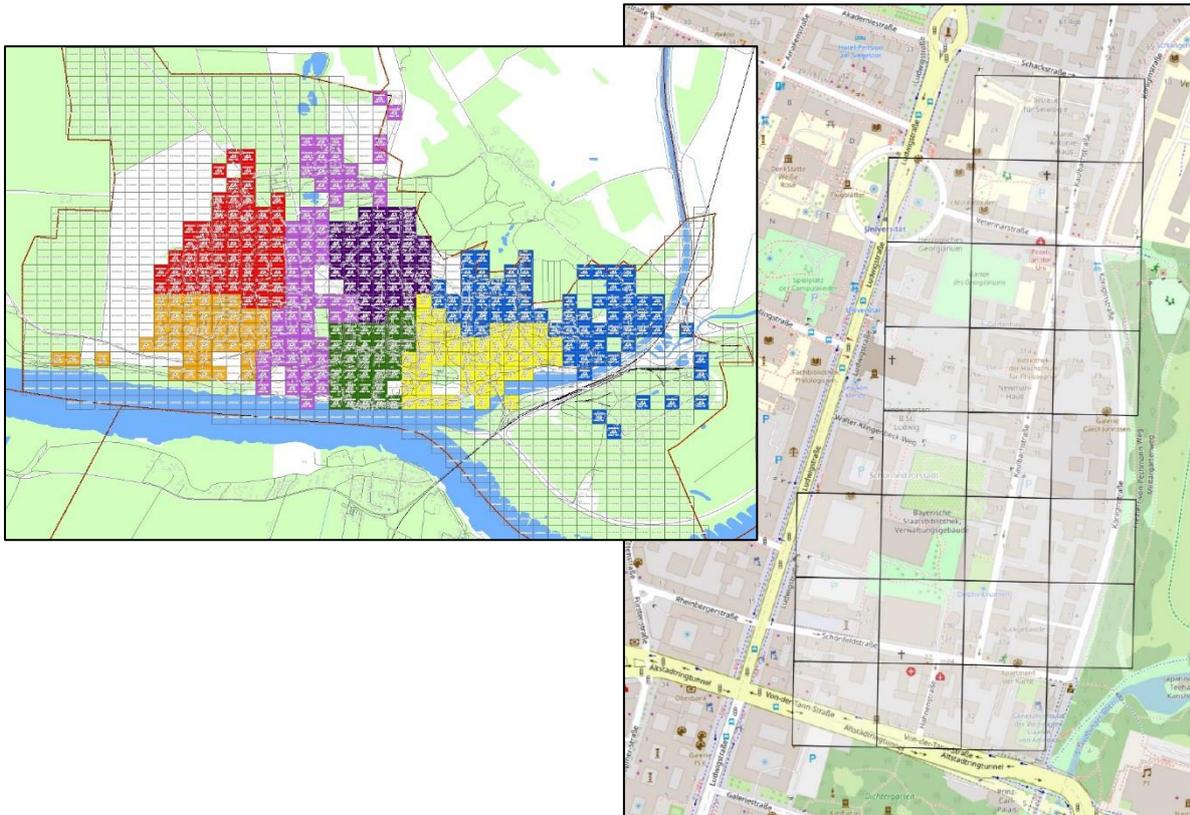


F2F-System 2023

Beschreibung der Auswahlgrundlage 2023



Erstellt für:

ADM.

ARBEITSGEMEINSCHAFT ADM-STICHPROBEN

Erstellt von:

BIK  ASCHPURWIS + BEHRENS GMBH
Markt-, Media- und Regionalforschung

Inhaltsverzeichnis

1.	Einleitung	5
2.	Aufbau des ADM-F2F-Flächenstichprobensystems 2023	7
3.	SAMPLE-POINTS (Säule 1)	9
3.1.	Anforderungen an die Point-Bildung	9
3.2.	Bildung der Sample-Points	10
3.2.1.	Einteilung der Gemeinden nach Einwohnerzahl in Gruppen für die Bearbeitung	10
3.2.2.	Gemeindefreie Gebiete	11
3.2.3.	Gemeindeguppe: unter 2.000 Einwohner	11
3.2.4.	Gemeindeguppe: 2.000 bis unter 100.000 Einwohner	17
3.2.5.	Gemeindeguppe: 100.000 Einwohner und mehr	22
3.3.	Mengengerüst aller Sample-Points im ADM-F2F-Flächenstichprobensystem 2023	26
3.4.	Bedeutungsgewichte	28
3.5.	Fortschreibung der Bedeutungsgewichte	31
3.6.	Ziehung der Institutsnetze	34
3.7.	Datensatzbeschreibungen der Sample-Point-Datei (Säule 1)	35
4.	BEGEHUNG - Beschreibung der Sample-Points über Adressen aus OSM (Säule 2)	37
4.1.	Die Abfrage der OpenStreetMap-Daten (OSM) zur Beschreibung der Points über die Geokoordinaten	38
4.2.	Auswahl der Values pro Node (Straßen- und Adressinformationen)	39
4.3.	Weiterverarbeitung der ausgelesenen OSM-Informationen	42
4.4.	Dichte der OSM-Angaben nach Gemeindegröße und Bundesland	43
4.5.	Datensatzbeschreibung der BG-Datei (Säule 2)	45
5.	STARTADRESSEN (Säule 3) Telefonbucheinträge innerhalb der Sample-Points	45
5.1.	Erstellung der Start-Adressen-Datei und Mengengerüste	46
5.2.	Postleitzahlen für 100x100m-INSPIRE-Grids	46
5.3.	Points ohne Startadressen	49
5.4.	Datensatzbeschreibung der ST-Datei (Säule 3)	50
6.	GEOKOORDINATEN (Säule 4)	51
6.1.	Was ist INSPIRE?	51
6.2.	Einwohner je Hektar	53
6.3.	Verknüpfung GeoGitter & Einwohner je Hektar	54
6.4.	Dichte-Merkmal	55
6.5.	Datensatzbeschreibung der Geokoordinaten-Datei (Säule 4)	56
6.6.	Koordinaten-Systeme	58
6.7.	Einsatz einer Geokoordinaten als Startadresse	58
6.8.	Möglichkeiten der Stichprobenziehung anhand der Geokoordinaten-Datei	59
7.	Vergleiche mit den bisherigen F2F-Systemen	60
8.	Schlussbemerkung	63
9.	Literaturverzeichnis	64

Abbildungsverzeichnis:

Abbildung 1: Entwicklung der Zahl der Interviews durch ADM-Institute 2017 – 2022 nach Befragungsart (ADM, 2023)	6
Abbildung 2: Die 4 Säulen im ADM-Flächenstichprobensystem F2F 2023	8
Abbildung 3: Beispiel für die Bearbeitung einer Gemeinde über die Zusammenfassung von Gitterzellen.....	9
Abbildung 4: Deutschland nach politischen Gemeindegrößenklassen (7 Klassen).....	10
Abbildung 5: Karte des BIK-Größenklassen (780er-Systematik).....	12
Abbildung 6: Karte der Synthetische Points im Kreis Dithmarschen (01051)	13
Abbildung 7: Synthetische Points, Point 859176 mit der maximalen Entfernung Luftlinie von 80km .	15
Abbildung 8: Sample-Point 858224 mit 17 Gemeinden	16
Abbildung 9: Kartenproduktion - Auszug Schwerin	17
Abbildung 10: Cluster zusammenfassen.....	18
Abbildung 11: Cluster aufteilen.....	18
Abbildung 12: Cluster sollen räumlich zusammenhängen	18
Abbildung 13: Begehrbarkeit eines Clusters	19
Abbildung 14: Durchschnittliche Point-Größen Bevölkerung (Zensus 2011) nach BIK-Strukturtyp in Gemeinden mit 2.000 bis unter 100.000 Einwohnern	21
Abbildung 15: Durchschnittliche Point-Größen Privathaushalte (MZ21) nach BIK-Strukturtyp in Gemeinden mit 2.000 bis unter 100.000 Einwohnern	21
Abbildung 16: GeoGitter > STT, Have their Center in	23
Abbildung 18: GeoGitter > STT, Crossed by the outline of	23
Abbildung 17: GeoGitter > STT, Overlay	23
Abbildung 19: Karlsruhe, Wolfartsweier (23), Nachbearbeitung Innerstädtische Zuordnung	24
Abbildung 20: Relevante 100m-Gitter & aktuelle Bevölkerung	29
Abbildung 21: Anzahl 100m-GeoGitter und Summe der Einwohner nach Einwohnergröße	30
Abbildung 22: Darstellung einer Neubaufläche auf der Karte mit unbewohnten GeoGittern (nicht dargestellt).....	31
Abbildung 23: Hafencity (102), starke Besiedelung seit 2011	32
Abbildung 24: Beispielsnetz mit der Verteilung von 258 Points	34
Abbildung 25: Beispiel für eine Abfrage in OSM	42
Abbildung 26: Random-Walk innerhalb eines Sample-Points.....	46
Abbildung 27: Zuordnung der PLZ zu 100x100m Gitterzellen über die geografische Schnittmengen PLZ - INSPIRE-ID.....	47
Abbildung 28: Verteilung der Startadressen pro Sample-Point nach Größe.....	48
Abbildung 29: Point 867321, Telgte, Stadt (05570044).....	49
Abbildung 30: Deutschland INSPIRE - 100km x 100km GeoGitter.....	52
Abbildung 31: Raunheim, Stadt - AUSZUG - INSPIRE: 100x100m GeoGitter inkl. OpenStreetMap & Einwohner je Hektar (Zensus 2011)	55
Abbildung 32: Dichte-Merkmal im Stadt-Land-Vergleich.....	56
Abbildung 33: Datenbeispiel der GK-Datei: Maintal, Stadt, POINT-ID: 855603.....	57
Abbildung 34: Flensburg, Stadt SP: 62420350 (1.052 Einwohner laut Zensus 2011)	59
Abbildung 35: Auswahl aller bewohnten 100m-Gitter im Umkreis von 1 km um die Düsseldorfer Königsallee (Kö)	60
Abbildung 36: Bevölkerungsentwicklung seit 2003	61
Abbildung 37: Durchschnittliche Haushaltsgröße (MZ21) nach BIK-Größenklassen	62
Abbildung 38: Durchschnittliche Haushaltsgröße (MZ21) nach politischen Größenklassen	63

Tabellenverzeichnis:

Tabelle 1: Gemeinden nach politischen Gemeindegrößenklassen, Fläche, Anzahl, Bevölkerung	10
Tabelle 2: Gemeindefreie Gebiete in Deutschland, Fläche und Anzahl.....	11
Tabelle 3: Synthetische Points nach Bundesländern, Zusammensetzung, Größe und Distanzen	14
Tabelle 4: Synthetische Points in Bundesländern nach Anzahl und Durchschnittsgrößen.....	16
Tabelle 5: Sample-Points in Gemeinden mit 2.000 bis unter 100.000 Einw. nach NUTS2-Regionen.	20
Tabelle 6: Sample-Points in Gemeinden mit mehr als 100.000 Einwohnern.....	24
Tabelle 7: Sample-Points in Bundesländern nach Anzahl und Durchschnittsgrößen	26
Tabelle 8: Sample-Points nach BIK-Gemeindegrößenklassen (10er) nach Anzahl und Durchschnittsgrößen	26
Tabelle 9: Sample-Points nach BIK-Gemeindegrößenklassen (Zusammenfassung) und NUTS2- Regionen	27
Tabelle 10: Verteilung GeoGitter laut Zensus 2011 nach Relevanz für das F2F-System 2023	29
Tabelle 11: Points mit besonders viel Neubautätigkeiten	33
Tabelle 12: Übersicht des F2F-Systems 2023 nach Instituten und Netzen	34
Tabelle 13: Säule 1-1: SP-Datei (Bildungsinformationen).....	35
Tabelle 14: Säule 1-2: SP-Datei (Sample-Point <-> AGS2212)	35
Tabelle 15: Säule 1-4: SP-Datei (Bedeutungsgewichte pro Point).....	35
Tabelle 16: Säule 1-3: SP-Datei (Instituts-Netze).....	36
Tabelle 17: OSM-Values für den Tag „highway“	41
Tabelle 18: Menge Angaben pro ausgelesener OSM-Adresse nach Bundesländern	43
Tabelle 19: Anzahl OSM-Einträge pro Gitterzelle und im Vergleich zu Wohngebäude nach Bundesland	44
Tabelle 20: Anzahl OSM-Einträge pro Gitterzelle im Vergleich zu Wohngebäuden nach BIK- Gemeindegrößenklassen	44
Tabelle 21: Säule 2: BG-Datei, Tabellen-Name: Begehungs-Beschreibung des POINTS.....	45
Tabelle 22: Mengenübersicht Startadressen pro Point nach Bundesland	47
Tabelle 23: Mengenübersicht Startadressen pro Point nach BIK-Gemeindegrößenklassen	48
Tabelle 24: Liste der Sample-Points ohne Startadressen	49
Tabelle 25: Säule 3 - ST-Datei, Startadressen im Point	50
Tabelle 26: Relevante 100m-GeoGitter	54
Tabelle 27: Unbewohnte oder geheim-zuhaltende GeoGitter	54
Tabelle 28: Dichte-Merkmal	55
Tabelle 29: Datensatzbeschreibung der Geokoordinaten-Datei (Säule 4).....	57
Tabelle 30: Datenbeispiel der GK-Datei: Maintal, Stadt, POINT-ID: 855603	57
Tabelle 31: Vergleich der F2F-Systeme nach Anzahl Points pro politischer Gemeindegrößenklasse..	61
Tabelle 32: Verteilung der Haushalte und Point-Größen nach Anzahl Haushalte pro Point	62

Abkürzungsverzeichnis

ADM = Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.

AGS = Amtlicher Gemeindeschlüssel

BG = Begehung

BMG = Bundesmeldegesetz

F2F= Face-to-Face

GS = Gebietsstand

INSPIRE = Infrastructure for Spatial Information in Europe

LBF = Laufende Bevölkerungfortschreibung der Statistischen Ämter

MZ = Mikrozensus

OSM = OpenStreetMap

PLZ = Postleitzahl

SP = Sample-Point

SS = Sachstand

ST = Startadresse

STT = Stadtteile (kommunalstatistische Einheiten) für 81 Großstädte

1. Einleitung

Die Technik der Befragung durch mündlich persönliche Interviews, sogenannten Face-to-Face-Befragungen (F2F), war bis zur flächendeckenden Telefondichte die einzige Möglichkeit, Zufallsstichproben in Deutschland durchzuführen. (Löffler, Behrens, & von der Heyde, 2014)¹ Sie basiert auf einer Einteilung der bewohnten Fläche Deutschlands in möglichst gleichstark bewohnte Teilflächen: den Stichproben-Points. Innerhalb dieser Points geht ein Interviewer, ausgehend von einer Startadresse, die durch das Befragungsinstitut bestimmt wird, in einer vorgeschriebenen Richtung durch die Straßen und Straßenabschnitte und registriert die vorgegebene Anzahl Klingelschilder der dort wohnenden privaten Haushalte. Es wird dann, meistens im Institut per systematischer Zufallsauswahl, ein Haushalt aus dieser Menge ausgewählt. Der Interviewer klingelt dann an der Tür und bittet um Teilnahme an der Befragung. Erklärt die Kontaktperson des Haushalts eine Teilnahmebereitschaft, wird noch ein weiteres Mal aus der Menge aller Haushaltsmitglieder eine Zielperson zufällig ausgewählt, mit der dann schließlich das Interview geführt wird. Weitere Haushalte im Point werden nach einer definierten Schrittweise je nach Begehungsvorschrift ausgewählt. Durch diesen mehrstufigen Auswahlprozess wird sichergestellt, dass man tatsächlich zufällig ein Interview mit einer berechenbaren Auswahlchance erhält.

In Deutschland verfolgte man diese Grundidee seit 1953. Ab 1961 nutzte man Wahlbezirke der damaligen Bundesrepublik („Flockenhaus-Emnid-Plan“) (Löffler, Behrens, & von der Heyde, 2014)². 1971 wurde dann erstmals ein sog. ADM-Mastersample erstellt, das für eine gemeinschaftliche Nutzung konzipiert wurde. Seitdem erfolgt in einem zeitlichen Abstand von rund 10 Jahren eine Aktualisierung dieses Flächenstichprobensystems (Löffler, Behrens, & von der Heyde, 2014)³. Den Auftrag für diese Arbeiten erteilt die „Arbeitsgemeinschaft Stichproben“, ein Zusammenschluss ausgewählter Mitgliedsinstitute des ADM e.V. und als außerordentliches Mitglied die Arbeitsgemeinschaft Media-Analyse e.V., vertreten durch die MMC GmbH. Der Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., abgekürzt ADM (www.adm-ev.de), vertritt als Wirtschaftsverband die Interessen der privatwirtschaftlichen Markt- und Sozialforschungsinstitute in Deutschland. Ein Methodenausschuss hat die Erstellung begleitet. Nur Arge-Mitglieder haben Zugang zur Auswahlgrundlage, eine Weitergabe an Dritte ist nach den Statuten ausgeschlossen.

Der regelmäßige Aufbau einer eigenen Flächenauswahlgrundlage ist notwendig, da der Zugriff und die Verwendung von Einwohnermeldedaten nur möglich ist, wenn die sogenannte Gruppenauskunft aus dem Melderegister im öffentlichen Interesse liegt. (vgl. Bundesmeldegesetz (BMG), Abschnitt 5, Unterabschnitt 2 (www.gesetze-im-internet.de, 2022)⁴, Stand: 19.12.2022). Da dies meist nur von wissenschaftlichen Einrichtungen, staatlichen Institutionen oder Bundesbehörden nachgewiesen werden kann, war und ist die privatwirtschaftlich finanzierte Forschung auf den Aufbau einer eigenen Auswahlgrundlage angewiesen, um für ihre Auftraggeber F2F-Untersuchungen durchführen zu können.

Die Datengrundlagen, die für eine solche Auswahlgrundlage verwendet werden können, haben sich dabei in den letzten 50 Jahren deutlich verändert. Bis 1997 hat man mit den Wahlbezirken zu einer Bundestagswahl, ihrer gedruckten Beschreibung über Straßen pro Gemeinde und der Zahl der Wahlberechtigten gearbeitet. Das System 2023 nutzt die amtliche Geodateninfrastruktur „INSPIRE“, die Geokoordinaten und Informationen zur Bevölkerung pro geografischem Gitter (100x100m) enthält.

Die Nutzung von neutralen unveränderlichen Gitterzellen und georeferenzierten Zensusergebnissen ermöglicht eine bessere Fortschreibung des Systems, das in früheren Jahren immer wieder vollständig zu hohen Kosten neu erstellt werden musste.

Daher war für die Aktualisierung der ADM-Auswahlgrundlage-F2F vor dem Hintergrund der abnehmenden Zahl an F2F-Untersuchungen (ADM, 2023)⁵ die Kostenreduzierung eine wichtige Anforderung. Darüber hinaus sollten, soweit möglich nur amtliche Daten oder frei verfügbare Daten

¹ Löffler, Behrens, von der Heyde, 2. Auflage, S. 67f.

² Ebenda, S. 68

³ Ebenda, S.85ff.

⁴ <https://www.gesetze-im-internet.de/bmg/BjNR108410013.html>

⁵ https://www.adm-ev.de/wp-content/uploads/2023/09/230914_ADM_Jahresbericht-2022_Screen-3.pdf

genutzt werden, um mögliche Lizenzgebühren zu vermeiden und Transparenz bei der Erstellung zu gewährleisten.

Standen bei der Erstellung der Auswahlgrundlage 2003 noch Straßenabschnittskordinaten auf Basis einer lizenzierten Nutzung von NavTech-Daten zur Verfügung, war im Jahr 2018 zu Beginn der Aktualisierung eine kostengünstige Nutzungsvariante nicht mehr zu erhalten. An deren Stelle treten erstmalig mit dem System 2023 Open Source Daten aus OpenStreetMap (OSM).

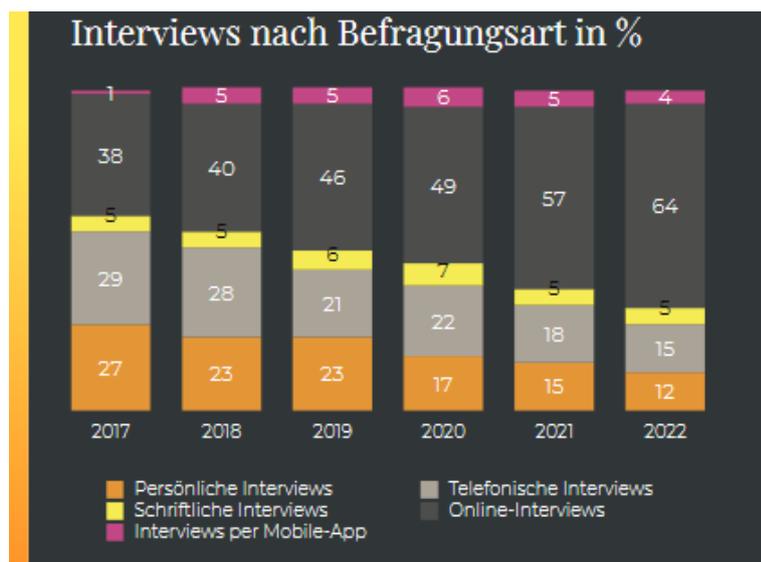
Wie die jährlich erstellte Auswertung des ADM zeigt (vgl. Abb.1), geht die Zahl der F2F-Interviews zwar zurück, da diese vor allem in der Durchführung zeitaufwändig und kostenintensiv sind, aber diese Befragungsmethode wird weiterhin genutzt.

Standardisierte persönlich-mündliche Interviews werden durch den Interviewer heutzutage in der Regel per CAPI geführt. Die Abkürzung CAPI steht für "Computer Assisted Personal Interview". Es gibt aber auch die Möglichkeit, dass der Befragte selbst anhand eines Computer-Fragebogens (CASI) antwortet, wobei der Interviewer für mögliche Rückfragen und Erläuterungen zur Verfügung steht. Man spricht dann von einem "Computer Assisted Selfadministered Interview" (CASI).

„Die persönlich-mündliche Befragung durch Interviewer oder Interviewerinnen bieten eine Reihe von Vorteilen. Dieser Befragungsmodus ermöglicht es zum Beispiel, komplexe Sachverhalte zu erheben, da der Interviewer umfassendere Erklärungen zu Fragen und Aufgaben geben kann, als dies in selbst administrierten Befragungsmodi möglich ist, außerdem kann der Interviewer zur vollständigen Beantwortung der Fragen motivieren. Auch die Durchführung von Tests, wie zum Beispiel einem Greifkrafttest oder Kompetenztests, ist im Rahmen von persönlichen Interviews möglich.

Allerdings birgt die persönlich-mündliche Befragung auch das Risiko, dass die Interviewer einen Effekt auf das Antwortverhalten der Befragten haben. Zudem sind persönliche Interviews deutlich kostenintensiver, aufwändiger und haben eine längere Feldzeit als die anderen Befragungsmodi.“ (Schröder, 2015)⁶

Abbildung 1: Entwicklung der Zahl der Interviews durch ADM-Institute 2017 – 2022 nach Befragungsart (ADM, 2023)⁷



⁶ Schröder, Jette (2015): Persönlich-mündliche Befragung. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines). DOI: 10.15465/gesis-sg_005

⁷ https://www.adm-ev.de/wp-content/uploads/2023/09/230914_ADM_Jahresbericht-2022_Screen-3.pdf

Die folgende methodische Beschreibung des Systems ist analog der Auswahlgrundlage in fünf Abschnitte gegliedert:

- Aufbau und Zusammensetzung des Gesamtsystem
- Säule 1: Sample-Points
- Säule 2: Beschreibung der Sample-Points über Straßen und Hausnummern aus OSM
- Säule 3: Startadressen innerhalb der Sample-Points
- Säule 4: Geokoordinaten der Sample-Points

2. Aufbau des ADM-F2F-Flächenstichprobensystems 2023

Grundsätzlich umfasst eine Auswahlgrundlage für die persönlich-mündliche Befragung immer die Elemente „Flächeneinheit“ und Bedeutungsgewicht. Hinzugekommen sind ab dem System 2003 auch die sog. „Startadressen“ innerhalb der Flächen. Mit diesen Startadressen kann man **zufällig** einen oder mehrere Startpunkte für die Begehung eines Stichproben-Points bestimmen, die die gesamte Point-Fläche abdecken. Man vermeidet einen räumlichen Klumpen in der Begehung eines Points. Zusätzlich lassen die Startadressen auch eine Überprüfung des Interviewers durch das beauftragende Institut zu, da diese alle mit einer Rufnummer im öffentlichen Telefonbuch eingetragen sind.

Im F2F-System 2023 kommt jetzt noch eine vierte Dimension hinzu: Die Geokoordinaten von INSPIRE-GeoGittern, aus denen sich ein Point zusammengesetzt. Dies wurde durch eine Erweiterung mit dem Zensus 2011 möglich: das Statistische Bundesamt (DESTATIS) und die Statistischen Ämter der Bundesländer haben erstmalig Einwohnerzahlen je geografischer 100x100m Gitterzelle flächendeckend für die ganze Bundesrepublik veröffentlicht (vgl. dazu Kap. 6). Damit konnte ein anderes räumliches Bezugssystem zur Definition und Beschreibung der Sample-Points als die bisherigen Straßen- und Straßenabschnitte verwendet und eine deutlich höhere Genauigkeit beim Bedeutungsgewicht der Points gewonnen werden.

Außerdem wurden für dieses System verschiedene georeferenzierte Informationen zu Flächen, Straßen und Gebäuden genutzt, die es im Web inzwischen auch dank der „Open Source Community“ frei zugänglich gibt. Für die ADM-Auswahlgrundlage-F2F 2023 wurden erstmalig Daten aus „OpenStreetMap“ (OSM) herangezogen (vgl. dazu Kap. 4).

„OpenStreetMap (OSM) wurde im Jahr 2004 ins Leben gerufen. Es entstand aus der Idee, eine freie und offene Karte der Welt zu erstellen, die von jedermann bearbeitet und genutzt werden kann. Die Gründer von OSM erkannten, dass es zu dieser Zeit keine umfassende und frei verfügbare Karte gab, die von der Gemeinschaft erstellt und aktualisiert wurde. Sie wollten eine Alternative zu den herkömmlichen, proprietären Kartenanbietern schaffen. Die Motivation hinter OSM war es, eine Karte zu schaffen, die von der Gemeinschaft gepflegt wird und auf der jeder seine eigenen Daten beitragen kann. Durch die Offenlegung der Daten und den Einsatz von Open-Source-Software ermöglicht OSM eine breite Beteiligung und fördert die Zusammenarbeit. Es hat sich zu einer der größten und detailliertesten Kartenressourcen entwickelt, die von Millionen von Menschen weltweit genutzt wird.“ (ChatGPT; Open.ai, 2024)⁸

Damit konnte für das ADM-Flächenstichprobensystems 2023 erstmalig die Geodateninfrastruktur „INSPIRE - Geographical Grid Systems“ im Zusammenhang mit den Einwohnern je Hektar aus dem Zensus 2011 zur Bildung der Sample-Points verwendet werden. Diese werden mit Daten aus OSM angereichert.

- „Geographische Gitter für Deutschland“ (GeoGitter Inspire, 100x100m) - Bundesamt für Kartographie und Geodäsie
- „Einwohner je Hektar“ (Zensus 2011) – Statistische Ämter des Bundes und der Länder
- OpenStreetMap (OSM)

⁸ ChatGPT, persönliche Kommunikation, 24. Januar 2024

Der Grundgedanke des Systems ist jedoch seit 1971 gleichgeblieben; setzt sich im Flächenstichprobensystem 2023 jetzt aber aus vier Säulen zusammen. (siehe Abb. 2)

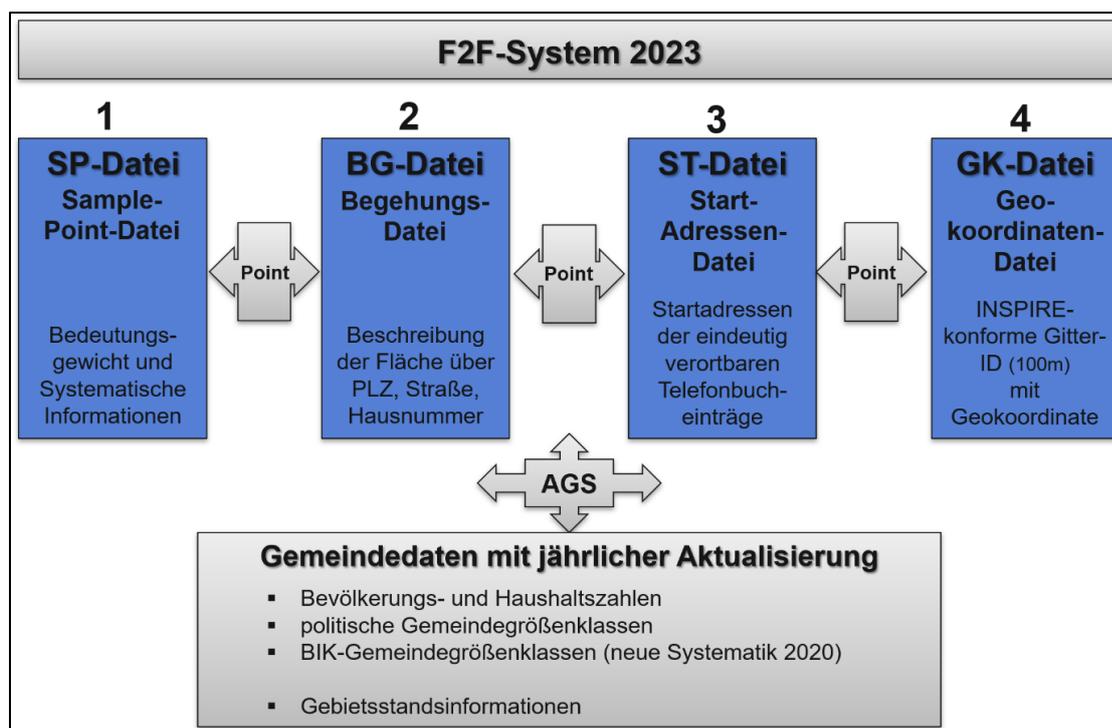
Jede Säule wird im Folgenden in einem Kapitel vorgestellt.

Um eine Aktualisierung des Systems während der mehrjährigen Nutzungsdauer zu ermöglichen, sind alle Sample-Points einer, oder bei synthetischen Points (vgl. dazu Kap. 3) mehreren Gemeinden eindeutig zugeordnet. Jede Gemeinde in Deutschland kann über den amtlichen Gemeindegeschlüssel (AGS) identifiziert werden.

„8-stelliger Schlüssel zur eindeutigen Identifizierung einer Gemeinde mit den Bestandteilen: Bundesland (2 Stellen), Regierungsbezirk (1 Stelle), Kreis (2 Stellen) und Gemeinde (3 Stellen) (DESTATIS, www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/amtlicher-gemeindegeschluessel.html)“⁹

Die Anzahl der Gemeinden (und damit der AGS) wird jährlich aktualisiert/ fortgeschrieben. Damit ist die Verwendung von aktuellen Bevölkerungs- und Haushaltszahlen, aber auch von Regionalgliederungen auf Gemeindeebene für Stichprobenallokationen oder Hochrechnungen möglich.

Abbildung 2: Die 4 Säulen im ADM-Flächenstichprobensystem F2F 2023



Die in der Markt- und Sozialforschung häufig zur Schichtung genutzten BIK-Gemeindegrößenklassen stammen aus der BIK-Regionssystematik, eine bundesweite räumliche Gliederungssystematik, die die Stadt-Umland-Beziehungen auf Gemeindeebene für Ballungsräume, Stadtregionen, Mittel- und Unterzentren darstellt. Die Vergleichbarkeit über mehrere Jahre ist nur eingeschränkt möglich, da es zwischen 2003 und 2023 aufgrund der veränderten Pendlerströme Updates dieser Systematik gab.

Eine ausführliche Beschreibung der BIK-Regionen und deren Historie ist in den Regionalen Standards 2019, Kapitel 5 nachzulesen. (GESIS, 2019)¹⁰

⁹ <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/amtlicher-gemeindegeschluessel.html>

¹⁰ https://www.asi-ev.org/standards_richtlinien/regional_standards__ausgabe_2019

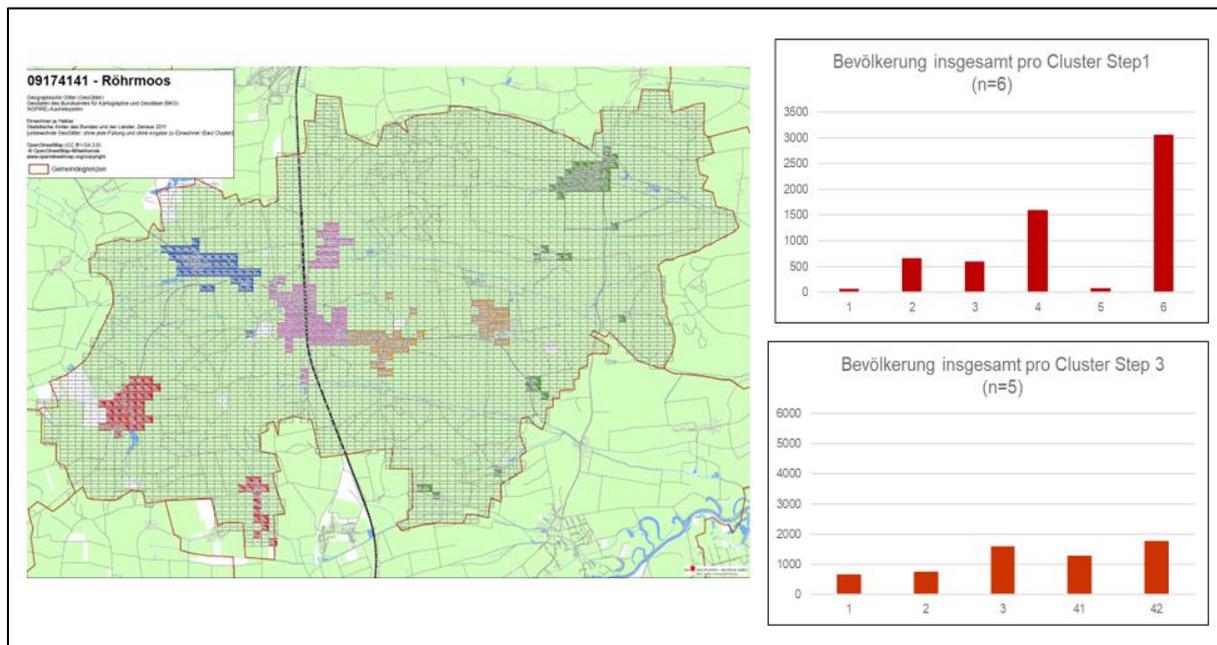
3. SAMPLE-POINTS (Säule 1)

3.1. Anforderungen an die Point-Bildung

F2F-Stichproben werden in der Regel als mehrstufige Flächenstichproben verwendet, die aus einer geschichteten Auswahlgrundlage gezogen werden. Gezogen werden Stichproben-Points, auch PSUs (Primary Sampling Units) genannt.

- Die bewohnte Fläche der Bundesrepublik soll durch möglichst **gleich große** und in der inneren Struktur **möglichst homogene** Einheiten abgebildet werden.
- Die Flächeneinheiten müssen groß genug sein, damit für die Bearbeitung dieser Sample-Points in der Laufzeit des Stichproben-Systems von mehreren Jahren eine hinreichende Menge von Haushaltsadressen zur Verfügung stehen: etwa 600 bis 700 Haushalte stellen - wie auch in den vorherigen ADM-F2F-Systemen - ein gutes durchschnittliches Mengengerüst dar.
- Je gleichmäßiger die Flächen in der Größe und räumlichen Verteilung sind, umso weniger sind Verzerrungen bei der Zufallsziehung der ersten Stufe zu erwarten, d.h., umso erwartungstreuere wird die Stichprobe sein.
- Allen Flächen soll ein primäres Bedeutungsgewicht für die Ziehung zugeordnet werden, das der Zahl der Privathaushalte in ihnen entspricht. Damit sollen die unvermeidlichen Größenunterschiede zwischen den Sample-Points bei der Stichprobenbildung adäquat berücksichtigt werden. (ADM, 2023)¹¹
- Die Point-Bildung erfolgt in Stufen: Die errechnete Clusterbildung auf Grundlage des Kriteriums der räumlichen Nähe wurde im Hinblick auf die Größe der Gitterzellen in einer zweiten Stufe anhand der Größe und der Befahr-/Begehrbarkeit überarbeitet. Dazu wurden pro Gemeinde Karten erstellt, auf denen die Gemeinden mit einer Straßenkarte und der Anzahl Einwohner pro GeoGitter abgebildet sind. Auf Basis dieser Karten wurde dann die finale Lösung erstellt und gegengeprüft. Das genaue Vorgehen ist in Abschnitt 3.2 beschrieben.

Abbildung 3: Beispiel für die Bearbeitung einer Gemeinde über die Zusammenfassung von Gitterzellen



¹¹ Stichproben-Verfahren in der Umfrageforschung, Eine Darstellung für die Praxis, Hrsg. ADM e.V., Kapitel 5.3.1, S.90

3.2. Bildung der Sample-Points

3.2.1. Einteilung der Gemeinden nach Einwohnerzahl in Gruppen für die Bearbeitung

Die Bildung der Sample-Points mit dem Bedeutungsgewicht aus den Ergebnissen des Zensus 2011 pro GeoGitter (Inspire, 100x100m) erfolgte innerhalb von Gemeindegruppen unterschiedlich. Dazu wurde im ersten Schritt die Bundesrepublik nach ihrer Einwohnergröße, der politischen Gemeindegrößenklasse, eingeteilt.

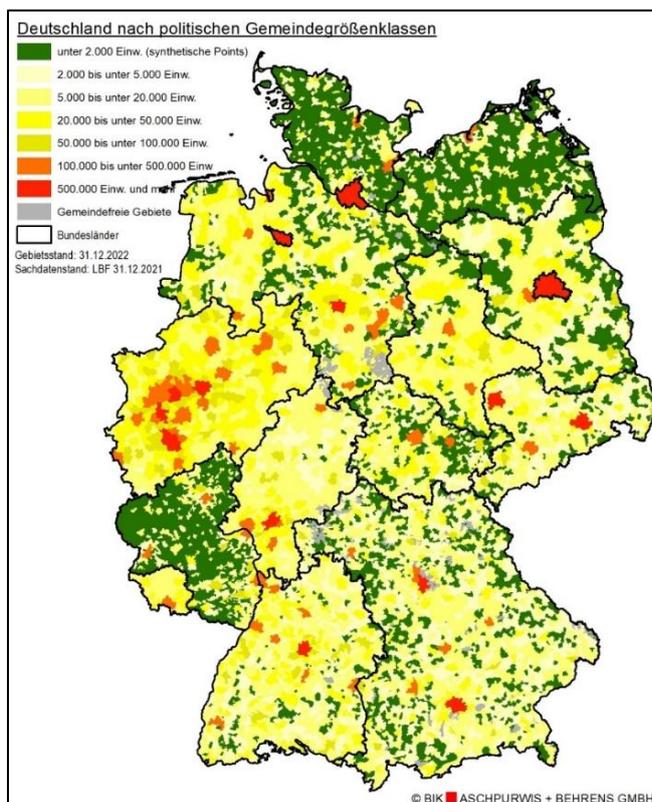
Tabelle 1: Gemeinden nach politischen Gemeindegrößenklassen, Fläche, Anzahl, Bevölkerung

Gemeindegröße ‚politisch‘ 1-7	Anzahl relevante 100m-Gitter 31.12.2019	Anzahl Gemeinden 31.12.2019		Bevölkerung Insgesamt LBF* 31.12.2018		Gemeindegruppe für die Bearbeitung
	abs.	abs.	in %	abs.	in %	
0 Gemeindefreie Gebiete	386	53		0		Gemeindefreie Gebiete
1 unter 2.000 Einw.	405.051	5.666	52,5%	4.488.075	5,4%	1 unter 2.000 Einwohner (n=5.666)
2 2.000 bis unter 5.000 Einw.	481.143	2.183	20,2%	7.050.096	8,5%	2 2.000 bis unter 100.000 Einwohner (n=5.052)
3 5.000 bis unter 20.000 Einw.	1.103.765	2.248	20,8%	22.022.667	26,5%	
4 20.000 bis unter 50.000 Einw.	558.426	511	4,7%	15.382.056	18,5%	
5 50.000 bis unter 100.000 Einw.	202.453	110	1,0%	7.436.494	9,0%	3 100.000 Einwohner und mehr (n=81)
6 100.000 bis unter 500.000 Einw.	243.455	67	0,6%	12.651.058	15,2%	
7 500.000 Einw. und mehr	183.044	14	0,1%	13.988.767	16,9%	
Gesamt	3.177.723	10.799	100%	83.019.213	100%	

*LBF= Laufende Bevölkerungsfortschreibung der Statistischen Ämter

Durch die Fortschreibung des Gebietsstands auf den 31.12.2022 und Bevölkerungsveränderungen wechseln zum Zeitpunkt der Auslieferung einige wenige Gemeinden die Gemeindegrößenklassen.

Abbildung 4: Deutschland nach politischen Gemeindegrößenklassen (7 Klassen)



3.2.2. Gemeindefreie Gebiete

Die Fläche der Bundesrepublik weist neben den administrativen Einheiten der Gemeinden auch noch sogenannte „Gemeindefreie Gebiete“ auf, die zwar als Gemeinden mit Fläche zählen, aber nicht bewohnt sind. (DESTATIS, www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/amtlicher-gemeindeschluessel.html)¹²

Tabelle 2: Gemeindefreie Gebiete in Deutschland, Fläche und Anzahl

Bundesland		Anzahl gemeindefreier Gebiete 31.12.2022	Anzahl relevante 100m-Gitter
01	Schleswig-Holstein	2	9
03	Niedersachsen	16	117
06	Hessen	4	31
07	Rheinland-Pfalz	1	14
08	Baden-Württemberg	1	1
09	Bayern	29	214
Deutschland		53	386

Einzig das Bundesland Niedersachsen führt zwei bewohnte gemeindefreie Gebiete, die damit auch als Gemeinde zählen, und in allen Gemeindestatistiken berücksichtigt werden:

- AGS 03 351 501 Loheide und AGS 03 358 501 Osterheide (Niedersachsen)¹³.

Beide Gemeinden sind dauerhaft bewohnt, gehören aber in die Fläche des Nato-Truppenübungsplatz Bergen. Aufgrund dieser Einschränkung sind nur begrenzte kommunale Einnahmequellen vorhanden und sie werden daher zur Zeit von der Bundesanstalt für Immobilienaufgaben verwaltet.

3.2.3. Gemeindegruppe: unter 2.000 Einwohner

In dieser Größenklasse bildet eine Gemeinde entweder genau einen Sample-Point oder mehrere kleine Gemeinden wurden zu einem sogenannten „**synthetischen Point**“,

- der die nötige **Größe** hat **und**
- **innerhalb einer Schichtungszelle aus Kreis x 10 BIK-Gemeindegroßenklasse** liegt, zusammengefasst.

Auf einen räumlichen Zusammenhang wurde, soweit möglich, geachtet. Die Geographischen Gitter in Zusammenhang mit den Einwohnern je Hektar (Zensus 2011) mussten für die Flächenbildung dieser Gruppe nicht herangezogen werden, da synthetische Sample-Points immer komplette Gemeindeflächen umfassen.

Um das Kriterium „liegt in einer BIK-Gemeindegroßenklasse innerhalb eines Landkreises“ für die Synthetisierung anhand aktueller Daten vornehmen zu können, wurde im Jahr 2020 zunächst die BIK-Systematik aktualisiert. Die BIK-Regionen und Verflechtungsgebiete sind eine bundesweite räumliche Gliederungssystematik, die die Stadt-Umland-Beziehungen auf Gemeindeebene für Ballungsräume, Stadtregionen, Mittel- und Unterzentren darstellt. (BIK ASCHPURWIS + BEHRENS GMBH, kein Datum)¹⁴

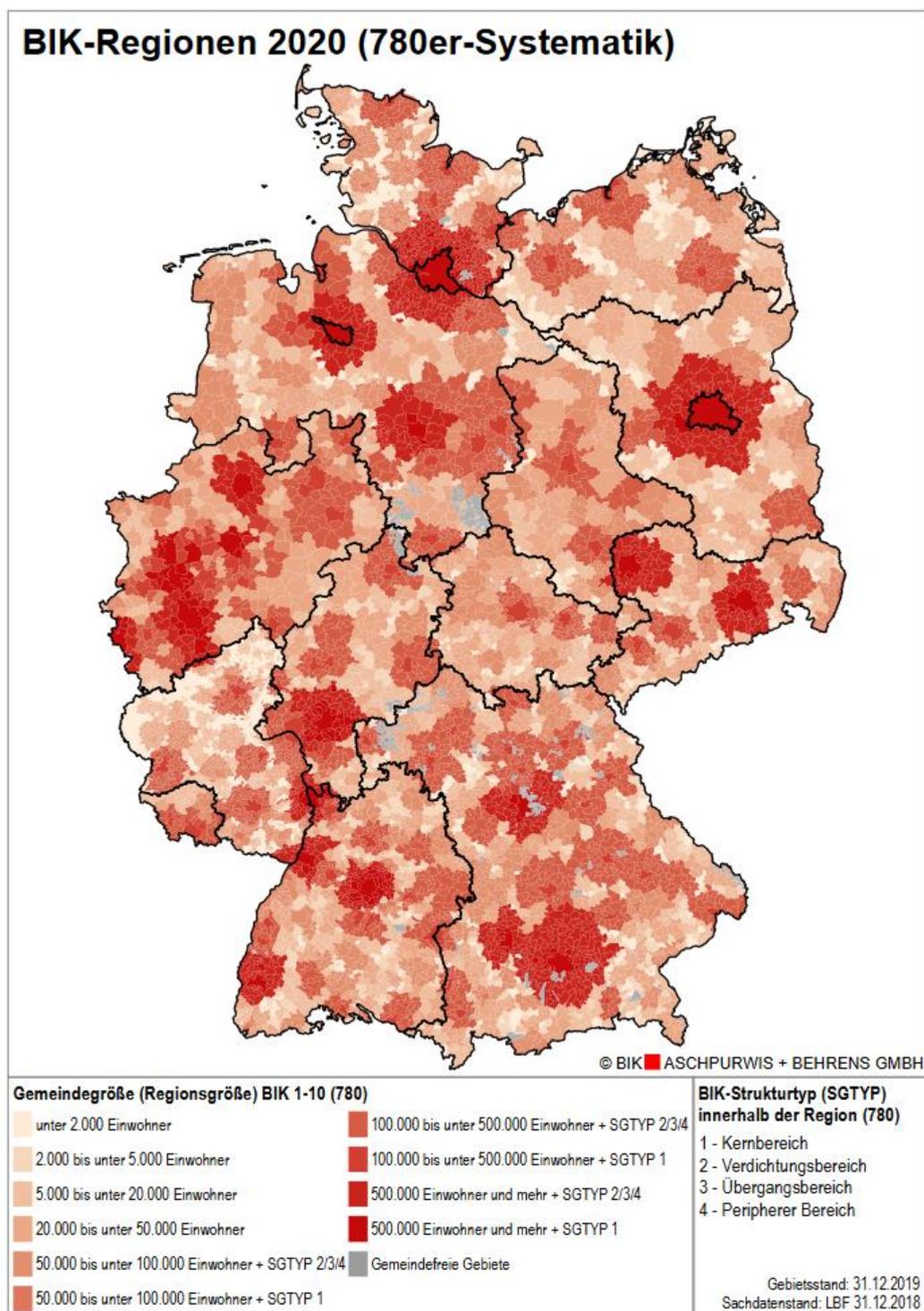
¹² <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/gemeinde-freie-gebiete.html>

¹³ https://www.statistik.niedersachsen.de/startseite/themen/gebiet_flachennutzung/gebiet_niedersachsens/gebiet-niedersachsens-191889.html

¹⁴ <https://www.bik-gmbh.de/cms/regionaldaten/bik-regionen>

Die Datengrundlage bilden die sozialversicherungspflichtig Beschäftigten, Stand 30.06.2018, und die Bevölkerung am Ort der Hauptwohnung am 31.12.2018 mit dem Gebietsstand 31.12.2019. Die BIK-Regionen werden jährlich auf Basis der LBF zum Jahresendstand aktualisiert; ebenso der Gebietsstand. Eine ausführliche Dokumentation und kartografische Darstellungen der BIK-Regionen können auf der Internetseite der BIK ■ ASCHPURWIS + BEHRENS GMBH heruntergeladen werden und sind Bestandteil der Regionalen Standards, Ausgabe 2019 (GESIS, 2019)¹⁵.

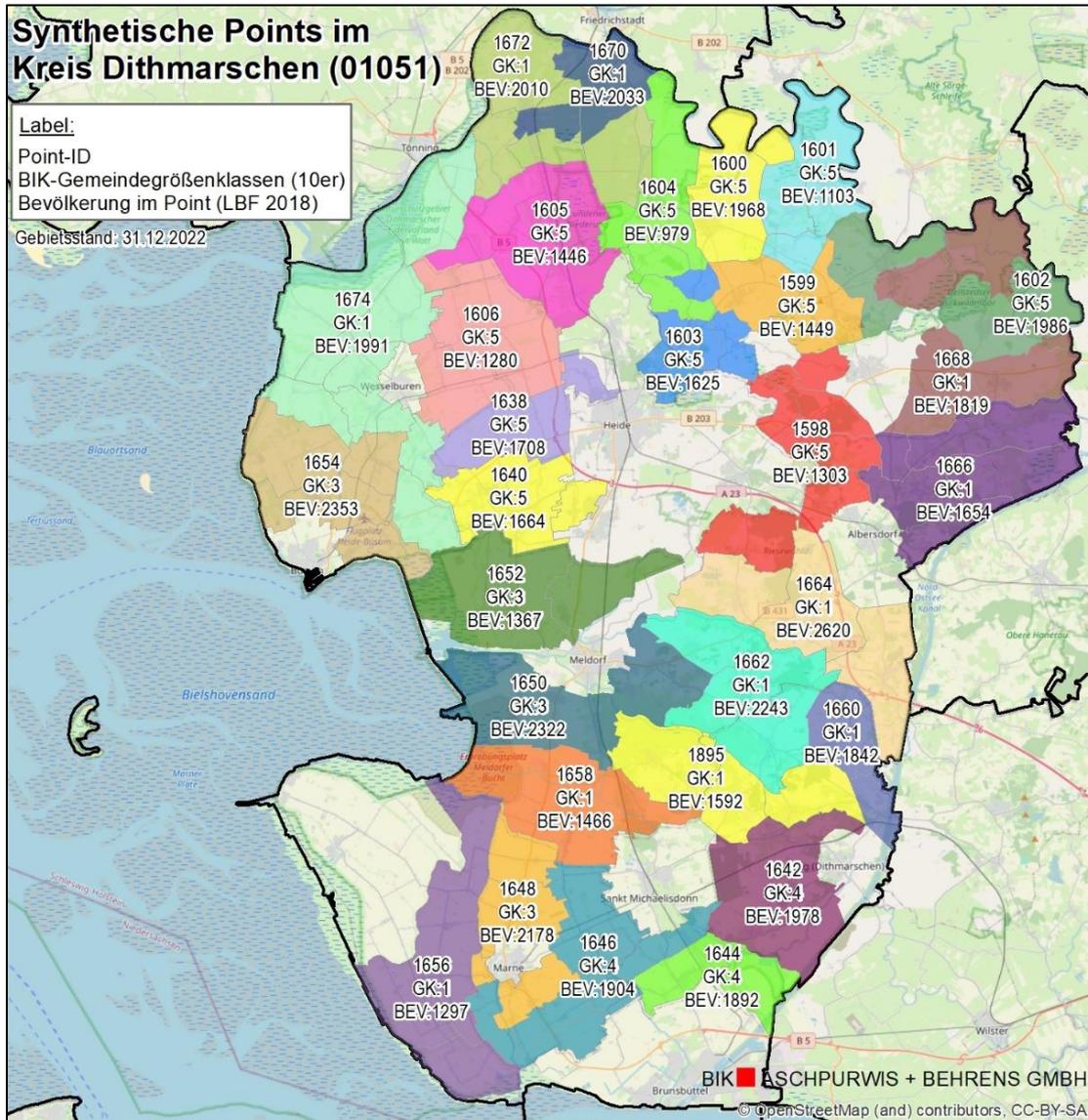
Abbildung 5: Karte des BIK-Größenklassen (780er-Systematik)



¹⁵ <https://www.asiev.org/index.php?id=50>

Für die Zusammenfassung zu den „synthetischen Points“ wurden dann je Landkreis alle AGS unter 2.000 Einwohner mit den jeweiligen BIK-Gemeindegrößenklassen, die auf Gemeindeebene trennscharf vergeben sind, dargestellt. Unter Berücksichtigung der Erreichbarkeit (Nähe, Straßen) wurden die Gemeinden dann anhand der laufenden Bevölkerungsfortschreibung (LBF) am 31.12.2018 zu synthetischen Points zusammengefasst.

Abbildung 6: Karte der Synthetische Points im Kreis Dithmarschen (01051)



Die Bedingung, dass innerhalb eines Kreises nur in der gleichen BIK-Gemeindegrößenklasse Gemeinden zusammengefasst werden konnten, führt dazu, dass in dieser Bearbeitungskategorie Points entstehen, die für die Mindestgrenze von 600 bis 700 Haushalten zum Teil zu klein sind, da einzelne Gemeinden innerhalb der festgelegten Schichtungsmerkmale nicht mehr mit weiteren Gemeinden zusammengefasst werden konnten.

Im Kreis Herzogtum Lauenburg (01053) zum Beispiel gibt es lediglich eine Gemeinde in der BIK-Gemeindegrößenklasse 3, was zu einem Sample-Point mit nur 138 Privathaushalten (Mikrozensus 2021, BIK-Haushaltsschätzmodell) führt. Der kleinste Sample-Point im ADM-Flächenstichprobensystem F2F 2023 mit 55 Privathaushalten liegt im Kreis Steinburg und beinhaltet lediglich die Gemeinde Kudensee.

Umgekehrt gab es aber auch 10 Sample-Points, die knapp oberhalb der Mindestgröße lagen. Hier handelt es sich immer um Zusammenfassungen von drei Gemeinden, bei denen eine Zuordnung einer Gemeinde dieser drei zu einem anderen Point nicht sinnvoll möglich war. Es sollte immer nur eine Gemeinde einen Synthetischen Point bilden, wenn es in der Kombination aus Kreis x 10 BIK-Gemeindegrößenklasse sonst keine weitere Gemeinde gab. Daher wurde in diesen Fällen die leichte Überschreitung der Obergrenze in Kauf genommen.

Nicht immer ließen sich direkte räumliche Zusammenfassungen von Gemeinden vornehmen:

- Es gibt Gemeinden, die in Gemeindeteile zerfallen, die räumlich getrennt sind; dies war bei synthetischen Points in Rheinland-Pfalz zum Teil der Fall.
- Zum anderen können Gemeinden der gleichen BIK-Größenklasse räumlich getrennt liegen, da es sich bei den Gemeinden um Randgemeinden unterschiedlicher Pendler-Räume (BIK-Regionen) handelt. Dann wurde der Point räumlich so dicht wie möglich gefasst, um Interviewern lange Wegstrecken zu ersparen und Fahrtkosten zu minimieren.
- Trotzdem sind für einige Points sehr große Distanzen innerhalb des synthetischen Points entstanden, um die Bildungs-Regel aus Kreis x 10 BIK-Gemeindegrößenklasse einzuhalten.
- Im Schnitt beträgt die maximale Luftlinien-Distanz im Point 11 km; in Sachsen und Brandenburg beträgt allein die Luftlinien-Distanz 17 bis 18 km. Die tatsächlichen PKW-Fahrzeiten im deutschen Straßennetz können noch deutlich länger sein.

Nach Abschluss der Point-Bildung für die ländlichen, dünn besiedelten Gemeinden sind 2.091 gemeindeübergreifende Points entstanden. Insgesamt werden 5.654 Gemeinden zum Gebietsstand 31.12.2022 durch synthetische Points abgebildet.

Tabelle 3: Synthetische Points nach Bundesländern, Zusammensetzung, Größe und Distanzen

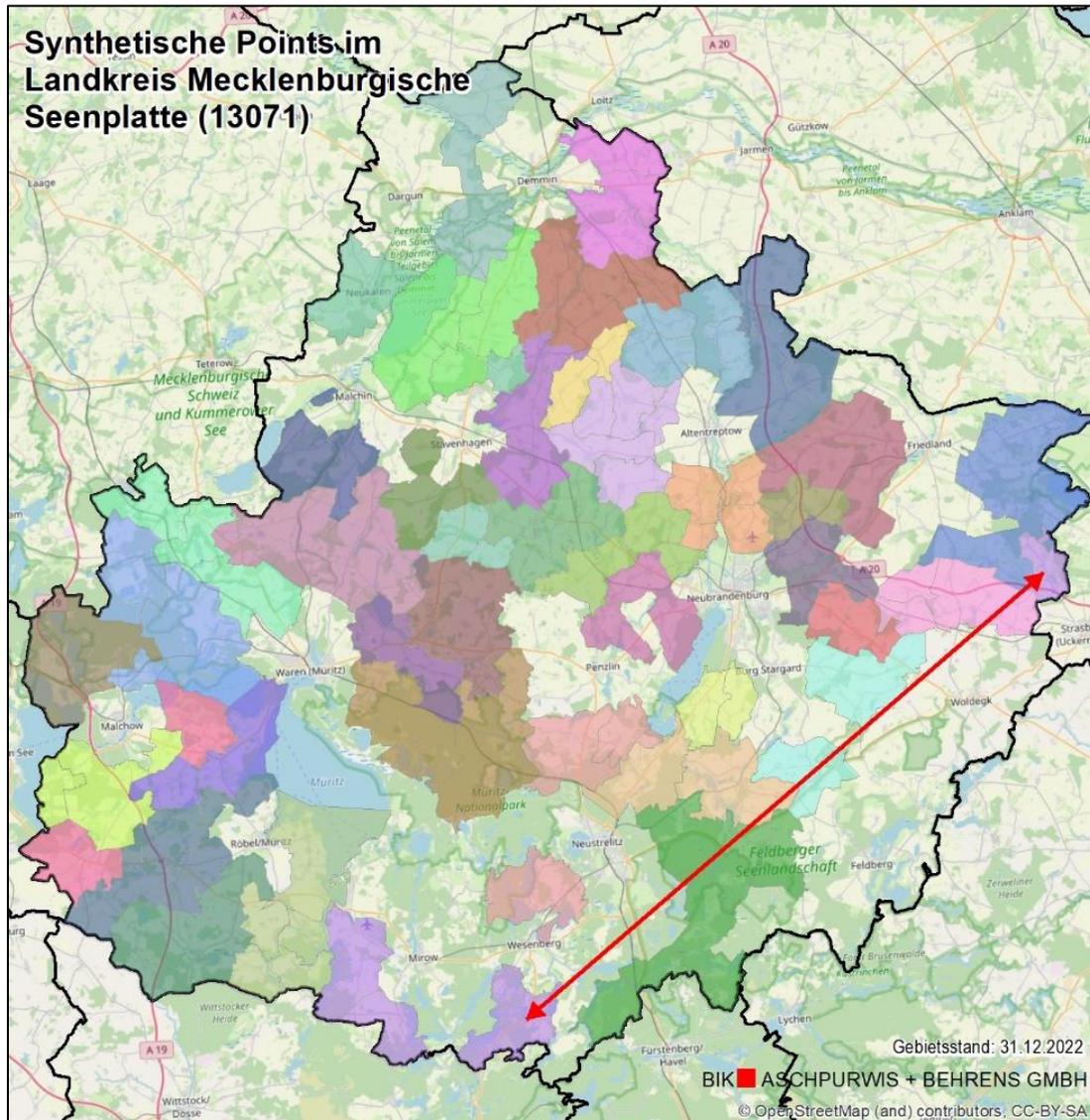
Bundesland	Anzahl synthetische Points	Anzahl Gemeinden 31.12.2022	Ø Anzahl Gem. je Point	Anzahl relevante 100m-Gitter	Ø Anzahl Gitterzellen je Point	Ø max. Distanz im Point	Größte Distanz innerhalb eines Point
01 SH	298	893	3,0	59.675	200,3	10,8 km	47,2 km
02 HH							
03 NI	190	433	2,3	49.458	260,3	13,2 km	52,2 km
04 HB							
05 NW							
06 HE	8	10	1,3	1.181	147,6	14,3 km	46,2 km
07 RP	612	1.966	3,2	76.780	125,5	7,2 km	44,3 km
08 BW	91	177	2,0	13.616	149,6	11,5 km	46,8 km
09 BY	353	736	2,1	82.974	235,1	13,2 km	57,5 km
10 SL							
11 BE							
12 BB	99	219	2,2	23.132	233,7	17,8 km	55,9 km
13 MV	210	595	2,8	47.060	244,1	15,6 km	80,7 km
14 SN	44	91	2,1	14.045	319,2	17,1 km	47,5 km
15 ST	40	81	2,0	9.309	232,7	14,8 km	40,7 km
16 TH	146	453	3,1	17.578	120,4	9,4 km	51,4 km
Deutschland	2.091	5.654	2,7	394.808	188,8	11,3 km	

*Verwendete Abkürzung siehe: <https://www.destatis.de/DE/Methoden/abkuerzung-bundeslaender-DE-EN.html>

Wieviel Fläche, und damit wie viele 100x100m-Gitter einen synthetischen Point bilden, sagt etwas über die möglichen Fahr Dauern eines Interviewers aus.

Die weiteste Entfernung zwischen Gemeinden innerhalb eines synthetischen Points liegt im Landkreis Mecklenburgische Seenplatte (13071) und misst per Luftlinie 80 km. Über die Bundesstraße 198 ist aber eine relativ gute Erreichbarkeit gewährleistet. Im Bundesland Mecklenburg-Vorpommern machen sich bei den Distanzen die durch Kreisreformen immer größer gewordenen Flächenkreise bemerkbar.

Abbildung 7: Synthetische Points, Point 859176 mit der maximalen Entfernung Luftlinie von 80km



In den Stadtstaaten Berlin, Hamburg und Bremen, aber auch in Nordrhein-Westfalen und dem Saarland gibt es keine synthetischen Points. Die meisten Stichprobeneinheiten dieser Kategorie liegen in Rheinland-Pfalz, was auf die sehr kleinräumigen Gemeindefstruktur (2.301 Gemeinden, 21,3% aller AGS Deutschlands) zurückzuführen ist. Das sieht man auch an der Menge Gemeinden, die jeweils einen Sample-Point bilden. Aber auch in den zum Teil sehr ländlich geprägten Bundesländern Schleswig-Holstein und Thüringen bilden in der Regel drei Gemeinden einen synthetischen Point.

Tabelle 4: Synthetische Points in Bundesländern nach Anzahl und Durchschnittsgrößen

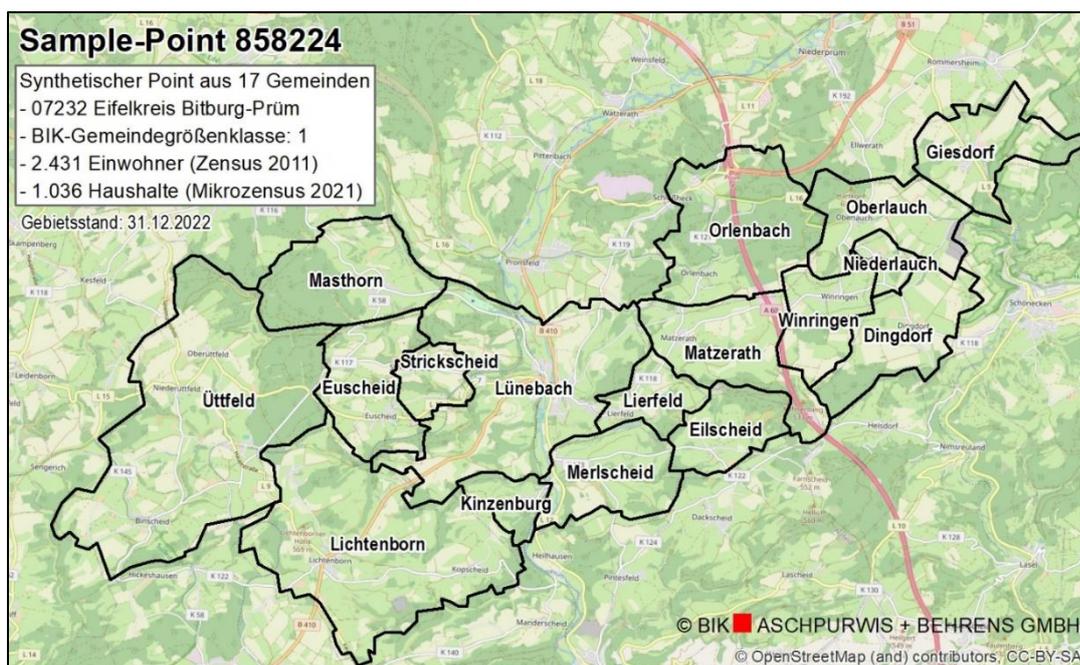
Bundesland	Anzahl synth. Points	Anzahl Gem. 31.12.2022	Ø Anzahl Gem. je Point	Anzahl relevante 100m-Gitter	Ø Anzahl Gitterzellen je Point	Ø Bev (LBF 2018) je Point	Ø Bev. (LBF 2018) je Gitterzelle	Ø HH (MZ 21) je Point	
01	SH	298	893	3,0	59.675	200,3	1.911	10,0	872
02	HH								
03	NI	190	433	2,3	49.458	260,3	2.259	10,0	1.163
04	HB								
05	NW								
06	HE	8	10	1,3	1.181	147,6	1.987	14,3	938
07	RP	612	1.966	3,2	76.780	125,5	1.877	15,4	862
08	BW	91	177	2,0	13.616	149,6	2.264	16,1	1.022
09	BY	353	736	2,1	82.974	235,1	2.816	12,7	1.272
10	SL								
11	BE								
12	BB	99	219	2,2	23.132	233,7	2.013	8,7	984
13	MV	210	595	2,8	47.060	244,1	1.890	8,6	906
14	SN	44	91	2,1	14.045	319,2	2.850	9,0	1.379
15	ST	40	81	2,0	9.309	232,7	2.569	11,4	1.167
16	TH	146	453	3,1	17.578	120,4	1.699	14,1	830
Deutschland	2.091	5.654	2,7	394.808		2.177	12,4		

Bev. = Bevölkerung/ HH= Haushalte

Die durchschnittliche Bevölkerung (LBF am 31.12.2018), anhand derer die gemeindeübergreifenden Flächen zusammengefasst wurden, ist in den Point-Summen bewusst größer als die Points in den übrigen Bildungsklassen. Zum einen wurde von den F2F-Instituten darum gebeten Sample-Points in dünn besiedelten Regionen besser etwas größer zu schneiden und zum anderen wiesen diese Gemeinden häufig zurückgehende Bevölkerungszahlen aus.

114 synthetische Sample-Points im F2F-Stichprobensystem 2023 sind mit nur einer Gemeinde besetzt. Dies war immer dann der Fall, wenn es sonst keine weiteren Gemeinden in der gleichen Kombination aus Kreis x BIK-GGKL gab. Das Maximum bildet der Point 858224 im Eifelkreis Bitburg-Prüm, der sich aus 17 Gemeinden in der BIK-Gemeindegrößenklasse 1 (unter 2.000 Einwohne) zusammensetzt.

Abbildung 8: Sample-Point 858224 mit 17 Gemeinden



3.2.4. Gemeindegruppe: 2.000 bis unter 100.000 Einwohner

In diese Gruppe fallen zum Gebietsstand 31.12.2022 insgesamt 5.052 Gemeinden. Innerhalb dieser Gemeinden sollten möglichst gleich große, begehbare Sample-Points gebildet werden.

Dazu wurden flächendeckend die hochauflösenden Geographische Gitter in der Kachelgröße 100m (GeoGitter, BKG) mit den Veröffentlichungen des Statistischen Bundesamtes „Einwohner je Hektar“ verknüpft. Rund 2,4 Mio. 100m-Gitterzellen mit 50,9 Mio. Einwohnern (Zensus 2011) wurden für diesen Teil bearbeitet. (vgl. Kap. 7).

Die Zusammenlegung der Gitterzellen innerhalb der Gemeinden erfolgte in einer ersten Stufe automatisiert über eine Clusteranalyse. Dazu wurde in SPSS die Methode K-Means verwendet, die über die Parameter Nähe (hier Geokoordinaten der GeoGitter und die Einwohnerzahl) mit einer 10maligen Iteration eine Clusterbildung über das Merkmal der Räumliche Nähe erstellte. Der Algorithmus endete, wenn sich die Zentren der Cluster nicht mehr änderten (Wiedenbeck & Züll, 2001)¹⁶. Die so entstanden Cluster mussten nun in einer zweiten Bearbeitungsstufe auf Begehbarkeit geprüft werden, und anhand der maximalen Einwohnerzahl pro Point zusammengelegt oder geteilt werden.

Für die weitere Bearbeitung in dieser Gruppe wurden dann alle 5.052 Gemeinden mit den relevanten systematischen Informationen in einzelne Dateien abgelegt. Die 100m-GeoGitter wurden aus den vom BKG bereitgestellten Geodaten je Gemeinde exportiert. Auf Grund der Veröffentlichungen von Shape-Dateien in 100m-Gittern mussten zum Teil einzelne Gemeindeteile zusammengeführt werden.

Für diese Arbeiten der Zusammenfassung und Überprüfung von Clustern zu Sample-Points unter Beachtung von Straßenverläufen, Gleisanlagen, Brücken oder ähnlichen Hindernissen wurden im ArcGIS-Desktop-Geoinformationssystem Karten je Gemeinde produziert. Neben den 100m-GeoGittern, den Einwohnern je Hektar, den Clusterergebnissen mit der Summe der Zensus2011-Einwohner wurden den Karten die Gemeindegrenzen und diverse für die Begehung relevante Informationen aus OpenStreetMap-Shapefiles¹⁷ (Landnutzung, Straße, Schienen, Wasserflächen/-wege, Gebäude) hinzugefügt.

Abbildung 9: Kartenproduktion - Auszug Schwerin



¹⁶ Wiedenbeck, M; Züll, C. (2001), Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren, hierzu S.8

¹⁷ <https://www.geofabrik.de/data/shapefiles.html>

Anhand dieser Karten im Zusammenhang mit Excel-Arbeitsmappen wurden die Cluster je Gemeinde auf folgende Point-Bildungs-Regeln geprüft und gegebenenfalls umcodiert:

(1) Mindestens 850 bis maximal 2.000 Einwohner (Zensus 2011) je Sample-Point

Sind die rechnerisch gebildeten Cluster zu klein, wurde zusammengefasst. Sind sie zu groß, musste aufgeteilt und/ oder einzelne Teilstücke eventuell einem anderen kleinen Cluster zugeordnet werden.

Abbildung 10: Cluster zusammenfassen

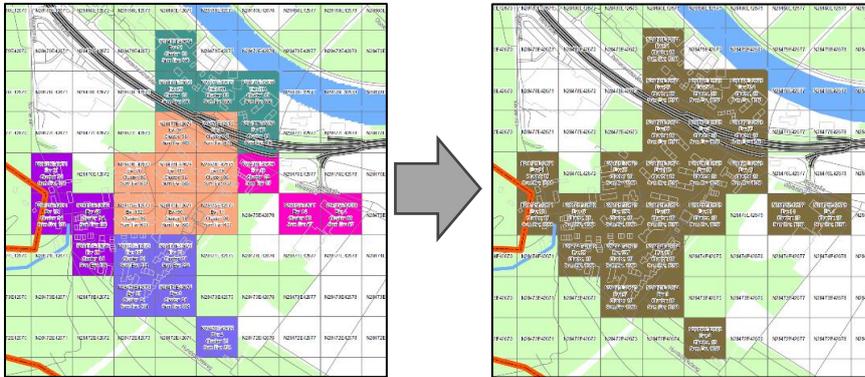
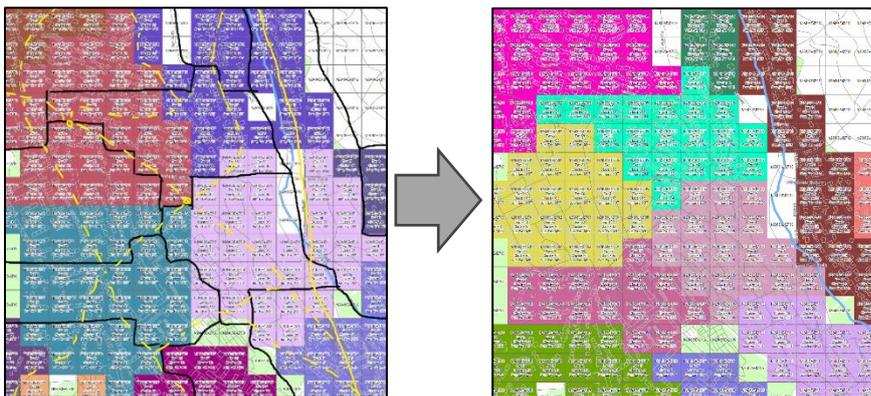


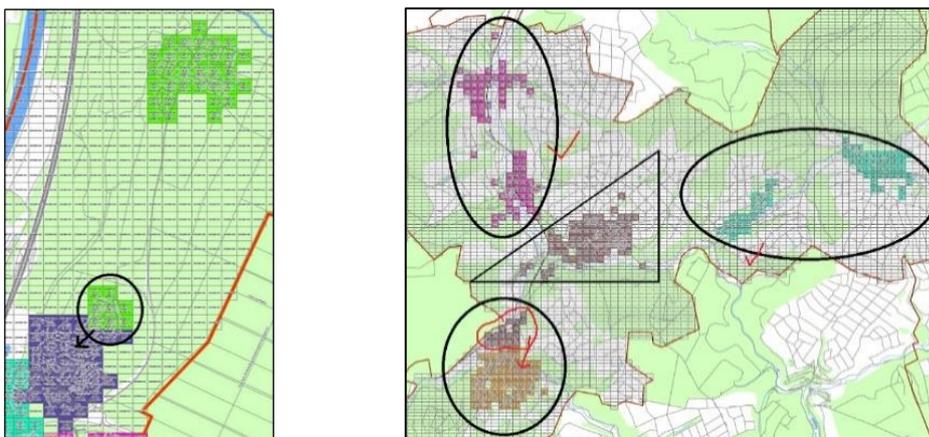
Abbildung 11: Cluster aufteilen



(2) Die Gitterzellen eines Clusters sollen räumlich zusammenhängen.

Wenn ein Point in einem Gebiet mit größeren Lücken zwischen den Siedlungen lag, wurde darauf geachtet, dass man die Geogitter im Point über eine Hauptverkehrsstraße erreichen kann (Abkürzungen, Fahrradwege und ähnliches zählen hier nicht). Dabei sollten Dörfer/ Ortschaften/ Siedlungen, wenn möglich nicht zerlegt werden.

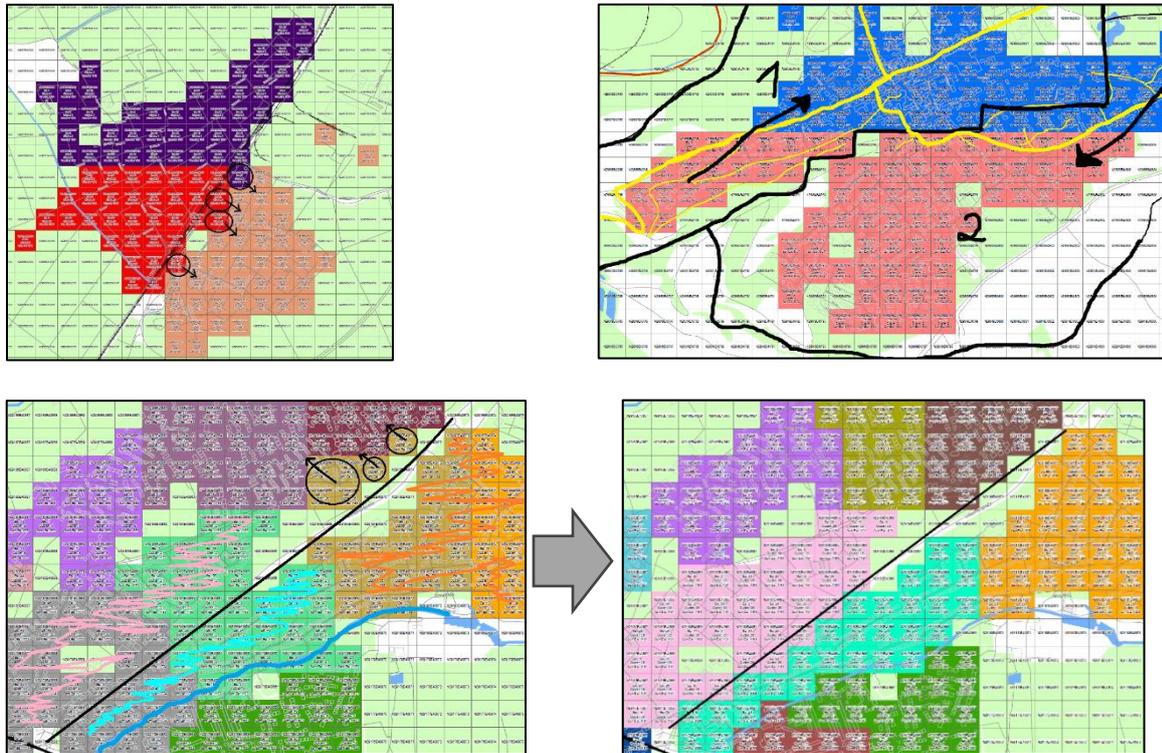
Abbildung 12: Cluster sollen räumlich zusammenhängen



(3) Die Gitterzellen innerhalb eines Clusters müssen begehbar sein.

„Begehbarkeit“ bedeutet: Man kann im Sample-Point ohne Umweg von einer Ecke zur anderen kommen, es sind Tunnel über oder unter der Eisenbahnlinie vorhanden, es gibt eine Brücke über den Fluss, es gibt eine Straße über die Autobahn, usw.. Cluster wurden möglichst unter Berücksichtigung „natürlicher Grenzen“ (Flüsse, Bahnlinien, große Straßen etc.) aufgeteilt.

Abbildung 13: Begehbarkeit eines Clusters



Der spätere **Einsatz in der Praxis** ist relevant und sollte bei der **Nachbearbeitung/ Überprüfung der Cluster im Fokus** stehen! Man muss bedenken, dass dort eine Person (Interviewer) hingeschickt wird, die nur in "ihrem" Sample-Point/ Cluster Interviews machen darf. Dabei soll sie möglichst kein anderes Cluster betreten und sich gut innerhalb des Clusters bewegen können (z.B. kein Fluss ohne Brücke o.ä.).

Am besten: klare und einfache Definitionen der Gebiete vorgeben, also zum Beispiel:

- "Interviewer 1: gehe nur in die kleine Ortschaft/ den Ortsteil/ ins Dorf XYZ"
- "Interviewer 2: Befrage nur alle Einwohner unterhalb der Bahnlinie"
- "Interviewer 3: Du darfst alle befragen die links von der Autobahn/ Straße wohnen"
- ...

Nach Überprüfung, Umstrukturierung oder Umkodierung der Sample-Points wurden alle Karten je Gemeinde anhand der korrigierten Excel-Arbeitsmappen erneut produziert und kontrolliert. Bei Bedarf wurden noch vereinzelt Gitterzellen anderen Points zugeordnet, bis alle Anforderungen an die Point-Bildung erfüllt waren. Anschließend wurden alle Daten je NUTS2-Region (DESTATIS, www.destatis.de/Europa/DE/Methoden-Metadaten/Klassifikationen/UebersichtKlassifikationen_NUTS.html)¹⁸ / ehemaligen Regierungsbezirken in Datenbanken des ADM-Flächenstichprobensystems F2F 2023 eingelesen.

Die Bearbeitungsgruppe der Gemeinden mit 2.000 bis 100.000 Einwohner erzeugte den meisten

¹⁸ https://www.destatis.de/Europa/DE/Methoden-Metadaten/Klassifikationen/UebersichtKlassifikationen_NUTS.html

Aufwand bei der Erstellung des neuen F2F-Stichprobensystems. Sie macht mit final 37.690 Sample-Points (64,4%) auch die größte Menge an Flächeneinheiten aus.

In den Stadtstaaten Berlin, Hamburg und Bremen gibt es auch in dieser Gruppe keine Points, die meisten liegen in den Bundesländern Nordrhein-Westfalen (6.663 Sample-Points), Baden-Württemberg (6.505 Sample-Points) und Bayern (6.455 Sample-Points).

Tabelle 5: Sample-Points in Gemeinden mit 2.000 bis unter 100.000 Einw. nach NUTS2-Regionen

NUTS2-Region/ ehemalige Regierungsbezirke	Anzahl Sample-Points	Anzahl Gemeinden 31.12.2022	Anzahl relevante 100m-Gitter	Ø Bevölkerung (Zensus 2011) je Point	Ø Haushalte (MZ 2021) je Point	
010	Schleswig-Holstein	1.274	211	75.103	1.402	719
020	Hamburg					
031	Statistische Region Braunschweig	675	81	43.592	1.350	652
032	Statistische Region Hannover	1.020	107	69.959	1.364	658
033	Statistische Region Lüneburg	1.115	154	92.333	1.331	664
034	Statistische Region Weser-Ems	1.471	158	145.537	1.388	650
040	Bremen					
051	Reg.-Bez. Düsseldorf	1.219	54	62.493	1.412	683
053	Reg.-Bez. Köln	1.780	94	95.227	1.389	656
055	Reg.-Bez. Münster	1.235	74	82.551	1.452	662
057	Reg.-Bez. Detmold	1.013	67	84.351	1.451	670
059	Reg.-Bez. Arnsberg	1.416	77	82.766	1.431	655
064	Reg.-Bez. Darmstadt	1.929	180	74.996	1.321	667
065	Reg.-Bez. Gießen	742	100	44.015	1.380	695
066	Reg.-Bez. Kassel	733	127	54.036	1.367	636
071	früher: Reg.-Bez. Koblenz	648	129	31.663	1.247	606
072	früher: Reg.-Bez. Trier	144	40	8.879	1.253	602
073	früher: Reg.-Bez. Rheinhessen-Pfalz	951	162	39.471	1.333	636
081	Reg.-Bez. Stuttgart	2.493	310	100.535	1.271	605
082	Reg.-Bez. Karlsruhe	1.410	193	62.589	1.283	615
083	Reg.-Bez. Freiburg	1.459	232	75.633	1.281	627
084	Reg.-Bez. Tübingen	1.143	180	66.615	1.287	629
091	Oberbayern	1.989	382	128.227	1.359	674
092	Niederbayern	777	180	81.196	1.373	668
093	Oberpfalz	606	134	51.493	1.330	639
094	Oberfranken	706	140	50.452	1.364	651
095	Mittelfranken	657	129	42.702	1.337	668
096	Unterfranken	760	168	41.991	1.311	606
097	Schwaben	960	179	62.988	1.364	681
100	Saarland	676	51	36.184	1.220	564
110	Berlin					
120	Brandenburg	1.577	193	117.595	1.330	671
130	Mecklenburg-Vorpommern	746	130	43.807	1.343	669
145	früher: Direktionsbezirk Chemnitz	890	147	64.906	1.360	645
146	früher: Direktionsbezirk Dresden	751	123	56.486	1.340	645
147	früher: Direktionsbezirk Leipzig	347	55	24.685	1.317	644
150	Sachsen-Anhalt	1.231	135	87.787	1.400	661
160	Thüringen	1.147	176	72.533	1.417	676
Deutschland		37.690	5.052	2.355.376	1.351	652

Zum Gebietstand 31.12.2022 gab es auch 12 Sample-Points, die in die politische Gemeindegrößenklasse 1 (unter 2.000 Einwohner, LBF am 31.12.2021) fallen. Das liegt daran, dass die Bevölkerung in den betroffenen Gemeinden zum Sachstand LBF 31.12.2018 (Einteilung der Gemeinden in die Bearbeitungsgruppen) leicht über der Grenze von 2.000 Einwohnern lang, nun aber durch Bevölkerungsrückgang in die untere Klasse wechselte.

Das zeigt sich auch in einer Auswertung nach den 10 BIK-Gemeindegrößenklassen: 6 Sample-Points in 3 Gemeinden ohne Verflechtungszusammenhang (Lauterecken, Stadt/ Tellau, Markt/ Teuschnitz, Stadt) sind nach der Fortschreibung der BIK-Gemeindegrößenklasse 1 zuzuordnen.

Durchschnittlich liegt die Bevölkerung (Zensus 2011) je Point in der gesamten Bundesrepublik zwischen 1.200 und 1.500 Einwohner. Die kleinsten Points in der Bearbeitung sind im bewohnten gemfr. Bezirk, Osterheide (03358501) mit 259 und 411 Zensus-Einwohnern. Diese Points wurden bewusst so klein geschnitten, da die Bevölkerung durch ein Ankunftszenrum für Asylbegehrende seit dem Zensus 2011 stark gewachsen ist. In Mühldorf am Inn, Stadt (09183128) und Feucht, Markt (09574123) wurden wegen offensichtlicher großer Neubaugebiete ebenfalls Flächen vorsätzlich sehr klein gehalten.

Abbildung 14: Durchschnittliche Point-Größen Bevölkerung (Zensus 2011) nach BIK-Strukturtyp in Gemeinden mit 2.000 bis unter 100.000 Einwohnern

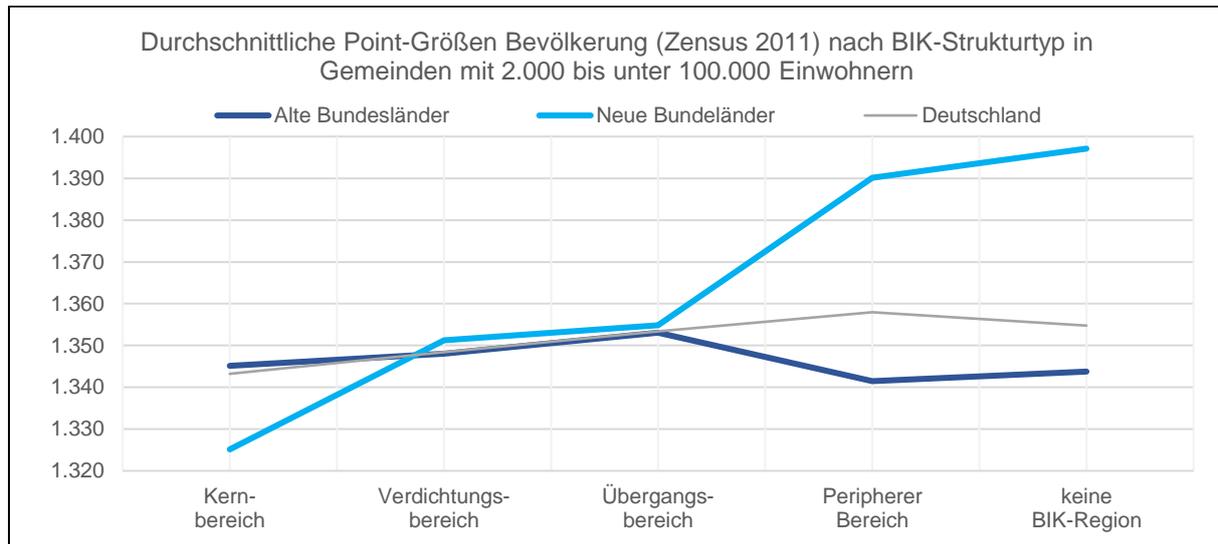
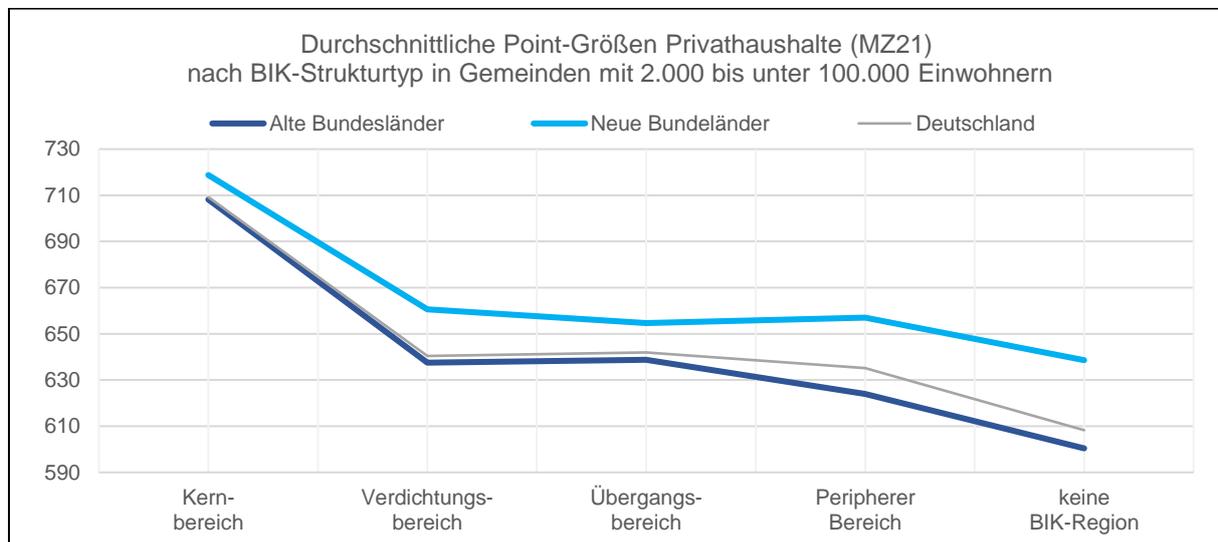


Abbildung 15: Durchschnittliche Point-Größen Privathaushalte (MZ21) nach BIK-Strukturtyp in Gemeinden mit 2.000 bis unter 100.000 Einwohnern



Generell wurde - wenn möglich - schon bei der Bildung und Prüfung der Sample-Points die demografische Entwicklung bedacht. So wurden Flächen in ländlichen Gemeinden, vor allem aber in den neuen Bundesländern eher größer geschnitten (Bevölkerungsrückgang). In Großstädten oder bei offensichtlichen Neubautätigkeiten wurden die Points eher kleiner definiert. Durch die Fortschreibung der Bedeutungsgewichte auf den Mikrozensus 2021 wurde die Diskrepanz wieder ausgeglichen.

Vergleicht man etwa die durchschnittlichen Point-Größen der Bevölkerung laut Zensus 2011 nach den 5 BIK-Strukturtypen (BIK ASCHPURWIS + BEHRENS GMBH, kein Datum)¹⁹ mit der durchschnittlichen Sample-Point-Größen der Privathaushalte (MZ21), werden gegensätzliche Verteilungen deutlich. In den Kernbereichen der BIK-Regionen wurden die Flächen eher klein geschnitten, durch einen Bevölkerungsanstieg und kleinere Haushaltsgrößen in diesem Gemeindetypen ergeben sich final aber mehr Privathaushalte je Point als in den anderen Umlandgemeinden oder gar in Gemeinden ohne Verflechtungszusammenhang.

Wichtig für den späteren Einsatz ist auch die Verteilung der Haushalts-Größen innerhalb und zwischen den Zellen.

3.2.5. Gemeindegruppe: 100.000 Einwohner und mehr

Zum Gebietstand 31.12.2022 gibt es in Deutschland 80 Gemeinden, die mehr als 100.000 Einwohner haben, für das ADM-Flächenstichprobensystems F2F 2023 wurden aber 81 Gemeinden in dieser Bildungsklasse bearbeitet. Die Stadt Cottbus fiel zum Sachstand LBF 31.12.2018 noch in die politische Gemeindegrößenklasse 6, was eine Erweiterung um innerstädtische Einteilungen ergab. Die Bevölkerung in dieser Stadt fiel mit der LBF am 31.12.2019 auf unter 100.000 Einwohner und wechselte somit in die politische Gemeindegrößenklasse 5 (50.000 bis unter 100.000 Einwohner).

Anforderung an die Point-Bildung in diesen Städten war zusätzlich die innerstädtische Einteilung, wie Stadtbezirke oder Stadtteile zu berücksichtigen.

Der vom BKG veröffentlichte Datenbestand der hochauflösenden Gitterzellen 100m erhält eine Zuordnung der Gemeinde über den Amtlichen Gemeindegrenzen (AGS). Für diese Ergänzung (Attribut „ags“) wurden die Zellen mit den Gemeinden der VG25 (harmonisierte Verwaltungsgebiete aus dem Basis-DLM) verschnitten und der AGS der flächenmäßig dominierenden Gemeinde zugeordnet. (Geodäsie)²⁰

Innerstädtisch, also innerhalb der amtlichen Gemeindegrenzen gibt es keine systematische Differenzierung.

In der ADM-Auswahlgrundlage-F2F ist dieses aber für Städte mit mehr als 100.000 Einwohnern (politische Gemeindegrößenklasse = 6+7) notwendig. Zum einen kann innerhalb kleinerer Polygone/ Flächen arbeitseffizienter geclustert werden und zum anderen sind die Abgrenzungen der Sample-Points z.T. für die Feldarbeit, Gewichtung, o.ä. trennscharf für innerstädtische statistische Einheiten erforderlich.

Dazu wurden Kartografie-Grundlagen der innerstädtischen Einteilung aller 81 Gemeinden mit mehr als 100.000 Einwohner beschafft. Damit auch Bevölkerungs- und Haushaltszahlen für eine Fortschreibung der Bedeutungsgewichte zur Verfügung stehen, wurden möglichst kommunalstatistische Einteilungen angefordert. Das war aber nicht immer ganz einfach, leider stellen bisher nur einige Kommunen alle erforderlichen Daten als Open-Data zum Download zur Verfügung. In vielen der Großstädte mussten z.B. die Städtestatistiker, die Ämter für Geoinformation/ Geodatenverarbeitung oder die Vermessungs- und Katasterämter kontaktiert werden. Außerdem unterscheiden sich die innerstädtischen Einheiten in den 80 Gemeinden: es gibt Stadtbezirke, Stadtteile, Stadtviertel, Ortsbezirke, Ortsteile, Ortschaften, Planungsräume oder statistische Bezirke/ Stadtteile. Auch die Größen der wesentlichen Regionalinformationen gehen stark auseinander, in

¹⁹ <https://www.bik-gmbh.de/cms/regionaldaten/bik-regionen>

²⁰ https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/geogitter.pdf Seite 6f

Bonn z.B. wurden nur 4 Stadtbezirke genutzt, in Erfurt hingegen standen alle Informationen für 53 Stadtteile zur Verfügung.

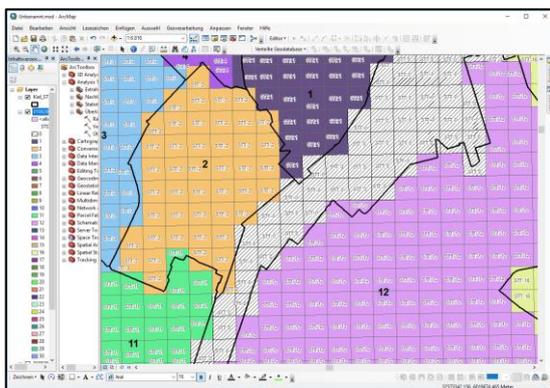
Um den INSPIRE-GeoGittern die jeweiligen kommunalstatistischen Einheiten zuzuordnen wurden die Gitterzellen je Gemeinde aus dem Gesamtbestand des BKG mit der Kartographie-Software ArcMap Desktop exportiert und mit den jeweiligen innerstädtischen Grundlagen (Shapefiles, STT) verbunden. Dieses erfolgte über eine Räumliche Verbindung der Polygone (AnalysisTools „Überlagerung/ Overlay > Räumliche Verbindung“) in 2 Schritten:

a) **HAVE THEIR CENTER IN**

Die Features in den Stadtteil-Grenzen (Verbindungs-Features) werden zugeordnet, wenn der Mittelpunkt eines GeoGitters (Ziel-Features) darin enthalten ist.

Liegt der Mittelpunkt des GeoGitters nicht innerhalb eines Stadtteils (Randgebiete) ist das zugeordnete Merkmal STT =0 (im folgenden Screenshot schraffiert).

Abbildung 16: GeoGitter > STT, Have their Center in

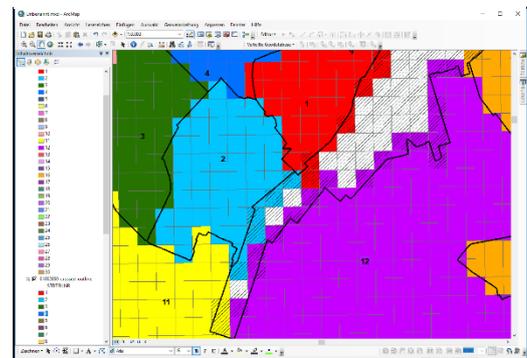
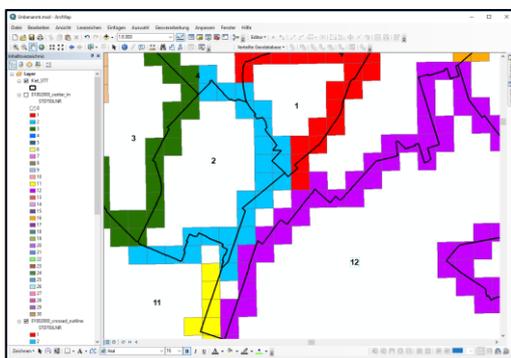


Beträgt die Wasserfläche des GeoGitters (lt. BKG, p_wasser) gleich 100% bleibt der Wert 0 (z.B. Küste/ Kieler Förde). Sofern aber Landfläche in der Gitterzelle liegt, muss dieses einem STT zugeordnet werden, um sicherzustellen die komplette bewohnte Fläche für die Bildung der Sample-Points zu berücksichtigen. Aus diesem Grund wird ein zweites Zuordnungsverfahren durchgeführt.

b) **CROSSED BY THE OUTLINE OF**

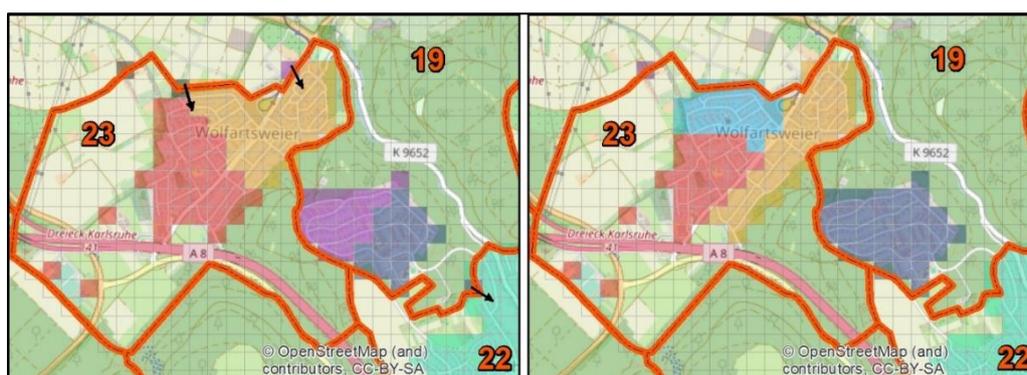
Die Features in den Stadtteil-Grenzen (Verbindungs-Features) werden zugeordnet, wenn ein GeoGitter (Ziel-Feature) durch seinen Umriss gekreuzt wird. Diese Zuordnung wurde nur berücksichtigt, wenn das erste Verfahren zu einem fehlenden Wert führte.

Abbildung 18: GeoGitter > STT, Crossed by the outline of



Vereinzelt wurden im Überarbeitungsprozess der Sample-Points Gitterzellen umcodiert. Wenn in der kartographischen Darstellung der GeoGitter und der OpenStreetMap-Daten klar ersichtlich wurde, dass die Gebäude zur benachbarten Regionaleinheit gehören, wurde die innerstädtische Einteilung händisch korrigiert. Diese Präzision war wichtig, um die Einwohnerzahlen pro Stadtteil korrekt zuzuordnen. Wies ein Stadtteil allein unter 700 Einwohner auf, wurde er mit einem anderen Stadtteil zusammengefasst.

Abbildung 19: Karlsruhe, Wolfartsweier (23), Nachbearbeitung Innerstädtische Zuordnung



Anschließend wurden die Gitterzellen genau wie in der Gemeindegruppe 2 (2.000 bis unter 100.000 Einwohner) automatisiert über eine Clusteranalyse zusammengelegt und anhand der gleichen Point-Bildungs-Regeln geprüft und gegebenenfalls umcodiert.

Tabelle 6: Sample-Points in Gemeinden mit mehr als 100.000 Einwohnern

AGS2212	Gemname2212	Anzahl kommunal-statistische Einheiten	Anzahl Sample-Points	Anzahl relevante 100m-Gitter	Ø Bevölkerung (Zensus 2011) je Point	Ø Haushalte (MZ 2011) je Point
01002000	Kiel, Landeshauptstadt	29 *	175	4.097	1.348	795
01003000	Lübeck, Hansestadt	10	151	4.660	1.394	833
02000000	Hamburg, Freie und Hansestadt	94 *	1.380	28.897	1.237	709
03101000	Braunschweig, Stadt	19	176	5.049	1.379	778
03102000	Salzgitter, Stadt	7	71	2.919	1.395	629
03103000	Wolfsburg, Stadt	16	88	3.382	1.364	756
03159016	Göttingen, Stadt	14	84	2.279	1.379	724
03241001	Hannover, Landeshauptstadt	13	339	7.276	1.494	847
03254021	Hildesheim, Stadt	14	71	2.147	1.403	750
03403000	Oldenburg (Oldenburg), Stadt	8 *	115	4.700	1.368	809
03404000	Osnabrück, Stadt	23	115	4.119	1.338	826
04011000	Bremen, Stadt	20 *	414	10.358	1.310	722
04012000	Bremerhaven, Stadt	7 *	77	2.316	1.403	740
05111000	Düsseldorf, Stadt	10	393	7.608	1.492	771
05112000	Duisburg, Stadt	7	338	7.823	1.445	751
05113000	Essen, Stadt	9	382	10.026	1.483	801
05114000	Krefeld, Stadt	18 *	143	4.910	1.555	776
05116000	Mönchengladbach, Stadt	4	178	5.607	1.434	713
05117000	Mülheim an der Ruhr, Stadt	3	116	3.600	1.439	743
05119000	Oberhausen, Stadt	3	147	3.823	1.426	721
05120000	Remscheid, Stadt	4	77	2.745	1.439	708
05122000	Solingen, Klengenstadt	5	105	3.945	1.480	736
05124000	Wuppertal, Stadt	10	228	6.629	1.502	717
05162024	Neuss, Stadt	25 *	103	2.859	1.462	683
05170024	Moers, Stadt	3	73	2.586	1.429	719
05314000	Bonn, Stadt	4	218	5.465	1.403	738
05315000	Köln, Stadt	9	703	12.896	1.431	799
05316000	Leverkusen, Stadt	13	111	3.228	1.433	813
05334002	Aachen, Stadt	7	164	4.448	1.442	844
05378004	Bergisch Gladbach, Stadt	6	74	3.064	1.469	790
05512000	Bottrop, Stadt	14 *	83	2.829	1.415	815
05513000	Gelsenkirchen, Stadt	5	180	4.208	1.438	774

AGS2212	Gemname2212	Anzahl kommunalstatistische Einheiten	Anzahl Sample-Points	Anzahl relevante 100m-Gitter	Ø Bevölkerung (Zensus 2011) je Point	Ø Haushalte (MZ 2011) je Point
05515000	Münster, Stadt	6	203	6.944	1.428	706
05562032	Recklinghausen, Stadt	16 *	81	2.663	1.432	763
05711000	Bielefeld, Stadt	10	228	9.175	1.435	722
05754008	Gütersloh, Stadt	12	66	3.497	1.429	732
05774032	Paderborn, Stadt	9	98	3.899	1.454	684
05911000	Bochum, Stadt	6	247	7.051	1.468	766
05913000	Dortmund, Stadt	12	390	10.970	1.465	759
05914000	Hagen, Stadt der FernUniversität	5	124	4.181	1.517	771
05915000	Hamm, Stadt	7	122	5.186	1.444	727
05916000	Herne, Stadt	4	104	2.731	1.491	737
05970040	Siegen, Universitätsstadt	17 *	71	3.176	1.398	673
06411000	Darmstadt, Wissenschaftsstadt	9	112	2.416	1.283	674
06412000	Frankfurt am Main, Stadt	45	521	7.512	1.283	772
06413000	Offenbach am Main, Stadt	19	88	1.417	1.287	630
06414000	Wiesbaden, Landeshauptstadt	25 *	212	4.379	1.270	680
06611000	Kassel, documenta-Stadt	23	149	3.967	1.280	701
07111000	Koblenz, Stadt	29 *	84	2.365	1.284	676
07211000	Trier, Stadt	19	80	2.362	1.321	769
07314000	Ludwigshafen am Rhein, Stadt	10	120	2.299	1.312	726
07315000	Mainz, Stadt	15	152	2.936	1.318	723
08111000	Stuttgart, Landeshauptstadt	23	469	7.843	1.249	689
08121000	Heilbronn, Universitätsstadt	9	93	2.226	1.248	758
08212000	Karlsruhe, Stadt	27	231	4.523	1.253	699
08221000	Heidelberg, Stadt	15	116	2.192	1.265	713
08222000	Mannheim, Universitätsstadt	17	230	3.913	1.261	736
08231000	Pforzheim, Stadt	15	87	2.189	1.316	729
08311000	Freiburg im Breisgau, Stadt	27 *	168	3.189	1.248	801
08415061	Reutlingen, Stadt	11 *	85	2.464	1.290	725
08421000	Ulm, Universitätsstadt	17	90	2.301	1.298	688
09161000	Ingolstadt	12	94	3.100	1.330	705
09162000	München, Landeshauptstadt	25	998	15.507	1.351	795
09362000	Regensburg	18	104	2.619	1.302	808
09562000	Erlangen	9	79	2.178	1.313	749
09563000	Fürth	18	87	2.155	1.328	803
09564000	Nürnberg	10	356	7.142	1.366	801
09663000	Würzburg	13	95	2.611	1.310	732
09761000	Augsburg	17	200	3.998	1.339	815
10041100	Saarbrücken, Landeshauptstadt	4	135	4.124	1.303	704
11000000	Berlin, Stadt	96 *	2.836	39.604	1.161	694
12052000	Cottbus/Chósebuz, Stadt	17	74	2.843	1.353	692
12054000	Potsdam, Stadt	8	119	3.207	1.311	763
13003000	Rostock, Hanse- u.Uni.stadt	22 *	150	3.000	1.336	850
14511000	Chemnitz, Stadt	39	192	5.975	1.252	701
14612000	Dresden, Stadt	17 *	434	9.206	1.182	697
14713000	Leipzig, Stadt	10	424	8.199	1.187	794
15002000	Halle (Saale), Stadt	5	172	3.860	1.333	724
15003000	Magdeburg, Landeshauptstadt	31 *	171	4.409	1.334	734
16051000	Erfurt, Stadt	39 *	161	3.858	1.248	778
16053000	Jena, Stadt	16 *	83	1.937	1.275	731

* Zusammenfassung einzelner kommunalstatistischer Einheiten (Stadtteile, Ortsteile etc.), da diese zu klein/ zu wenig Einwohner.

3.3. Mengengerüst aller Sample-Points im ADM-F2F-Flächenstichprobensystem 2023

In Deutschland ergeben sich für das ADM-Flächenstichprobensystems Face-to-Face 2023 insgesamt 58.564 Sample-Points. Die folgenden Tabellen stellen die Verteilung der Sample Points in den Bundesländern, den BIK-Gemeindegroßenklassen (10er) und den NUTS2-Regionen/ ehemaligen Regierungsbezirken dar. Die Summen der Bedeutungsgewichte (Bevölkerung und Haushalte lt. Mikrozensus 2021) und die durchschnittlichen Point-Größen werden ausgewiesen.

Tabelle 7: Sample-Points in Bundesländern nach Anzahl und Durchschnittsgrößen

Bundesland	Anzahl Gemeinden 31.12.2022	Anzahl Sample-Points	Anzahl relevante 100m-Gitter	Bevölkerung (MZ 21)	Ø Bev pro Point	Haushalte (MZ 21)	Ø HH pro Point
01 SH	1.106	1.898	143.535	2.867.000	1.511	1.441.000	759
02 HH	1	1.380	28.897	1.831.000	1.327	978.000	709
03 NI	941	5.530	432.750	7.878.000	1.425	3.865.000	699
04 HB	2	491	12.674	667.000	1.358	356.000	725
05 NW	396	12.213	565.160	17.646.000	1.445	8.632.000	707
06 HE	422	4.494	193.919	6.212.000	1.382	3.057.000	680
07 RP	2.301	2.791	166.755	4.027.400	1.443	1.927.000	690
08 BW	1.101	8.165	349.828	10.981.000	1.345	5.232.000	641
09 BY	2.056	8.821	581.333	12.939.100	1.467	6.297.000	714
10 SL	52	811	40.308	968.200	1.194	476.000	587
11 BE	1	2.836	39.604	3.612.000	1.274	1.968.000	694
12 BB	413	1.795	143.934	2.492.000	1.388	1.246.000	694
13 MV	726	1.106	93.867	1.578.550	1.427	817.000	739
14 SN	419	3.082	183.502	3.977.450	1.291	2.116.000	687
15 ST	218	1.614	105.365	2.123.000	1.315	1.110.000	688
16 TH	631	1.537	95.906	2.076.450	1.351	1.083.000	705
Deutschland	10.786	58.564	3.177.337	81.876.150	1.398	40.601.000	693

Tabelle 8: Sample-Points nach BIK-Gemeindegroßenklassen (10er) nach Anzahl und Durchschnittsgrößen

BIK-Gemeindegroßenklassen (10er)	Anzahl Gemeinden 31.12.2022	Anzahl Sample-Points	Anzahl relevante 100m-Gitter	Bevölkerung (MZ 21)	Ø Bev pro Point	Haushalte (MZ 21)	Ø HH pro Point
1 unter 2.000 Einwohner (Ew.)	1.445	491	92.768	985.664	2.007	448.109	913
2 2.000 bis unter 5.000 Ew.	458	1.128	107.681	1.478.101	1.310	667.048	591
3 5.000 bis unter 20.000 Ew.	1.415	4.689	400.749	6.414.757	1.368	2.997.299	639
4 20.000 bis unter 50.000 Ew.	1.841	6.500	494.015	9.007.776	1.386	4.351.929	670
5 50.000 bis u. 100.000 Ew. - Umland	1.444	5.119	397.335	7.175.255	1.402	3.412.527	667
6 50.000 bis u. 100.000 Ew. - Kern	43	1.307	48.730	1.791.256	1.371	896.607	686
7 100.000 bis u. 500.000 Ew.-Umland	2.624	8.675	639.636	12.160.591	1.402	5.661.486	653
8 100.000 bis u. 500.000 Ew. - Kern	167	8.700	272.191	12.135.674	1.395	6.402.774	736
9 500.000 Ew. und mehr - Umland	1.117	5.555	357.723	7.956.586	1.432	3.680.142	662
10 500.000 Ew. und mehr - Kern	232	16.400	366.509	22.770.490	1.388	12.083.079	737
Gesamt	10.786	58.564	3.177.377	81.876.150	1.398	40.601.000	693

Tabelle 9: Sample-Points nach BIK-Gemeindegrößenklassen (Zusammenfassung) und NUTS2-Regionen

NUTS2-Region/ ehemalige Regierungsbezirke	BIK-Gemeindegröße - Zusammenfassung *						Gesamt	
	Land (1-4)		Umland (5+7+9)		Kerne (6+8+10)		Abs.	in%
	Abs.	in%	Abs.	in%	Abs.	in%		
010 Schleswig-Holstein	393	0,67%	603	1,03%	902	1,54%	1.898	3,24%
020 Hamburg					1.380	2,36%	1380	2,36%
031 Statistische Region Braunschweig	194	0,33%	586	1,00%	348	0,59%	1.128	1,93%
032 Statistische Region Hannover	301	0,51%	694	1,19%	474	0,81%	1.469	2,51%
033 Statistische Region Lüneburg	295	0,50%	851	1,45%	50	0,09%	1.196	2,04%
034 Statistische Region Weser-Ems	506	0,86%	946	1,62%	285	0,49%	1.737	2,97%
040 Bremen					491	0,84%	491	0,84%
051 Reg.-Bez. Düsseldorf	323	0,55%	358	0,61%	2.821	4,82%	3.502	5,98%
053 Reg.-Bez. Köln	391	0,67%	935	1,60%	1.724	2,94%	3.050	5,21%
055 Reg.-Bez. Münster	405	0,69%	621	1,06%	756	1,29%	1.782	3,04%
057 Reg.-Bez. Detmold	180	0,31%	655	1,12%	570	0,97%	1.405	2,40%
059 Reg.-Bez. Arnsberg	482	0,82%	585	1,00%	1.407	2,40%	2.474	4,22%
064 Reg.-Bez. Darmstadt	388	0,66%	740	1,26%	1.734	2,96%	2.862	4,89%
065 Reg.-Bez. Gießen	262	0,45%	360	0,61%	121	0,21%	743	1,27%
066 Reg.-Bez. Kassel	320	0,55%	328	0,56%	241	0,41%	889	1,52%
071 früher: Reg.-Bez. Koblenz	568	0,97%	270	0,46%	209	0,36%	1.047	1,79%
072 früher: Reg.-Bez. Trier	148	0,25%	121	0,21%	80	0,14%	349	0,60%
073 früher: Reg.-Bez. Rheinhes.-Pfalz	291	0,50%	566	0,97%	538	0,92%	1.395	2,38%
081 Reg.-Bez. Stuttgart	489	0,83%	1.035	1,77%	1.549	2,64%	3.073	5,25%
082 Reg.-Bez. Karlsruhe	473	0,81%	659	1,13%	952	1,63%	2.084	3,56%
083 Reg.-Bez. Freiburg	528	0,90%	712	1,22%	417	0,71%	1.657	2,83%
084 Reg.-Bez. Tübingen	460	0,79%	560	0,96%	331	0,57%	1.351	2,31%
091 Oberbayern	601	1,03%	1.016	1,73%	1.520	2,60%	3.137	5,36%
092 Niederbayern	300	0,51%	391	0,67%	124	0,21%	815	1,39%
093 Oberpfalz	187	0,32%	388	0,66%	177	0,30%	752	1,28%
094 Oberfranken	124	0,21%	425	0,73%	195	0,33%	744	1,27%
095 Mittelfranken	161	0,27%	419	0,72%	635	1,08%	1.215	2,07%
096 Unterfranken	293	0,50%	403	0,69%	224	0,38%	920	1,57%
097 Schwaben	350	0,60%	458	0,78%	430	0,73%	1.238	2,11%
100 Saarland	127	0,22%	421	0,72%	263	0,45%	811	1,38%
110 Berlin					2.836	4,84%	2836	4,84%
120 Brandenburg	639	1,09%	890	1,52%	266	0,45%	1.795	3,07%
130 Mecklenburg-Vorpommern	485	0,83%	236	0,40%	385	0,66%	1.106	1,89%
145 früher: Direktionsbezirk Chemnitz	484	0,83%	308	0,53%	309	0,53%	1.101	1,88%
146 früher: Direktionsbezirk Dresden	284	0,48%	340	0,58%	584	1,00%	1.208	2,06%
147 früher: Direktionsbezirk Leipzig	114	0,19%	216	0,37%	443	0,76%	773	1,32%
150 Sachsen-Anhalt	567	0,97%	704	1,20%	343	0,59%	1.614	2,76%
160 Thüringen	695	1,19%	549	0,94%	293	0,50%	1.537	2,62%
Deutschland	12.808	21,87%	19.349	33,04%	26.407	45,09%	58.564	100,00%

*) BIK-Gemeindegröße (Regionsgröße) 1-10 (780) / (Sachstand: LBF 2022/12)

(1-4) Gemeinden mit unter 50.000 Einwohner

(5+7+9) Gemeinden mit 50.000 Einwohner und mehr + Strukturtyp 2/3/4 (Umland)

(6+8+10) Gemeinden mit 50.000 Einwohner und mehr + Strukturtyp 1 (Kern)

3.4. Bedeutungsgewichte

Der Zensus 2011 hat neben den 100x100m-Gitterzellen auch die Anzahl der Einwohner pro Gitter ermittelt. Der Einwohnerbegriff entspricht dem Begriff „Wohnbevölkerung“. Damit sind zunächst alle Personen in Deutschland gemeint, die **eine meldefähige Adresse** haben, am Haupt- und ggf. am Nebenwohnsitz. Nur für diese Personen gibt es Auszählungen aus dem Einwohnermelderegister, das dem Zensus 2011 als Grundlage gedient hat. (Anding & Fürnrohr, 2015)²¹

„Nach dem Bundesmeldegesetz (BMG) gilt eine Person als Einwohner/-in Deutschlands, die in einer Gemeinde ihren ständigen Wohnsitz hat bzw. dort wohnberechtigt ist. Als Wohnung wird nach § 20 BMG jeder umschlossene Raum, der zum Wohnen oder Schlafen benutzt wird, bezeichnet. Hat eine Person mehrere Wohnungen im Inland, so ist nach § 21 Absatz 1 BMG eine dieser Wohnungen ihre Hauptwohnung. Die Person hat der Meldebehörde nach § 21 Absatz 4 Satz 1 BMG mitzuteilen, welche Wohnung ihre Hauptwohnung ist. Nach Absatz 2 ist dies die vorwiegend benutzte Wohnung der Person.“ (DESTATIS, Demographische Standards. eine gemeinsame Empfehlung des ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes, 2016)²²

Auf Basis dieser im Zensus 2011 ermittelten Bevölkerung erfolgte dann die sog. Laufende Bevölkerungsfortschreibung.

„Basis der Bevölkerungsfortschreibung sind die Ergebnisse des letzten Zensus (2011). Im Zensus werden alle Personen erfasst, die gemeldet sein sollten: die Zahl der gemeldeten Personen wurde dabei für die Personen korrigiert, die zwar gemeldet, aber nicht wohnhaft ('Karteileichen') oder mehrfach gemeldet sind sowie für die Personen, die wohnhaft aber nicht gemeldet sind ('Fehlbestände'). Bei der laufenden Fortschreibung der Zensusergebnisse werden Angaben der Meldebehörden über erfasste An- und Abmeldungen von Personen beim Bezug und Auszug aus einer Wohnung sowie Angaben der Standesämter über erfasste Geburten und Sterbefälle zu Personen mit Wohnung in Deutschland berücksichtigt.“

Die Zuordnung zu einer Gemeinde im Inland erfolgt nach dem Standort der alleinigen oder Hauptwohnung. Es wird die in Deutschland lebende deutsche und nicht-deutsche (ausländische) Bevölkerung erfasst. Personen, die eine Wohnung im Inland und weitere Wohnungen im Ausland haben, zählen zur Bevölkerung“ (DESTATIS, Qualitätsbericht, Fortschreibung des Bevölkerungsstandes 2022, 2023)²³

Damit kann für jede bewohnte Gitterzelle ein Bedeutungsgewicht durch die Summe der Einwohner dieser Zellen bestimmt werden.

Die Einwohnerzahl (Bevölkerung insgesamt) wird durch die statistische Geheimhaltung laut dem Statistischen Bundesamt nicht verändert. Durch die Zusammenfassung der nicht bis dünn-besiedelten GeoGitter (0-1 Einwohner, 2-3 Einwohner) ergibt sich keine signifikante Abweichung für die Abbildung der Grundgesamtheit in den relevanten Merkmalen. Vergleicht man die Einwohner in den relevanten bewohnten GeoGittern nach den 10 BIK-Gemeindegrößenklassen (780er-Systematik) (BIK ASCHPURWIS + BEHRENS GMBH, kein Datum)²⁴ mit der Verteilung der Gesamtbevölkerung laut Zensus 2011 auf Gemeindeebene sind keine Verzerrungen zu erkennen. Alle Daten wurden auf den Gebietsstand 31.12.2022 fortgeschrieben.

²¹ Anding, K.; Fürnrohr, M.: Zensus 2011, Ermittlung der Einwohnerzahlen von Bund, Ländern und Kommunen“, in: Bayern in Zahlen, 4, 2015, S.216-226

²² Statistisches Bundesamt, Statistik und Wissenschaft, Bd. 17/2016 „Demografische Standards“, S.16, https://www.destatis.de/GPStatistik/receive/DEMonografie_monografie_00003695, zuletzt abgerufen am 10.12.2018

²³ DESTATIS (2023), Qualitätsbericht, Fortschreibung des Bevölkerungsstandes, 2022, hierzu S.5

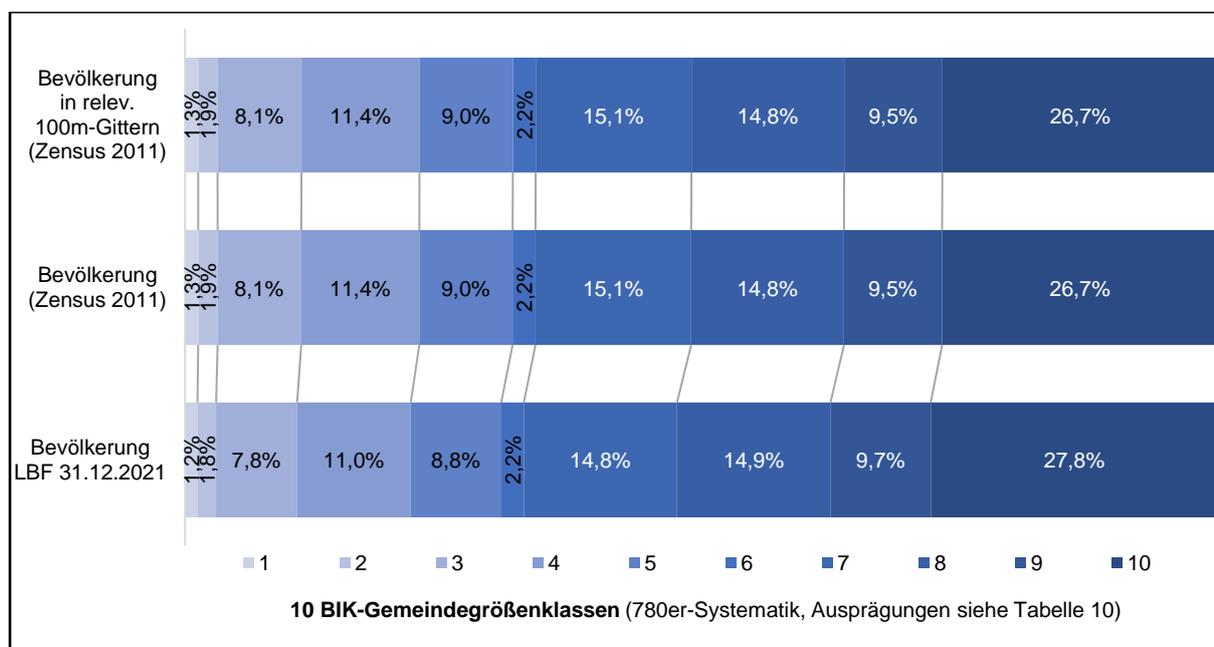
²⁴ <https://www.bik-gmbh.de/cms/regionaldaten/bik-regionen>

Tabelle 10: Verteilung GeoGitter laut Zensus 2011 nach Relevanz für das F2F-System 2023

BIK-Gemeindegrößenklassen (10er)		Anzahl Gemeinden 31.12.2022		Bevölkerung Insgesamt Zensus 2011		relevante 100m-Gitter			
						Anzahl		Bevölkerung	
						31.12.2022		Zensus 2011	
						abs.	in %	abs.	in %
1	unter 2.000 Einw.	1.445	13%	1.015.963	1,3%	92.768	3%	1.021.519	1,3%
2	2.000 bis unter 5.000 Einw.	458	4%	1.502.503	1,9%	107.681	3%	1.507.438	1,9%
3	5.000 bis unter 20.000 Einw.	1.415	13%	6.502.455	8,1%	400.749	13%	6.521.208	8,1%
4	20.000 bis unter 50.000 Einw.	1.841	17%	9.164.417	11,4%	494.015	16%	9.184.574	11,4%
5	50.000 bis u. 100.000 Einw. - Umland	1.444	13%	7.243.321	9,0%	397.335	13%	7.260.396	9,0%
6	50.000 bis u. 100.000 Einw. - Kern	43	0%	1.770.998	2,2%	48.730	2%	1.771.794	2,2%
7	100.000 bis u. 500.000 Einw. - Umland	2.624	24%	12.084.415	15,1%	639.636	20%	12.108.803	15,1%
8	100.000 bis u. 500.000 Einw. - Kern	167	2%	11.873.768	14,8%	272.191	9%	11.877.156	14,8%
9	500.000 Einw. und mehr - Umland	1.117	10%	7.617.768	9,5%	357.723	11%	7.631.423	9,5%
10	500.000 Einw. und mehr - Kern	232	2%	21.434.389	26,7%	366.509	12%	21.437.627	26,7%
Gesamt		10.786	100%	80.209.997	100%	3.177.337	100%	80.321.938	100%

Vergleicht man die Verteilung der Zensus-Bevölkerung in den relevanten 100m-GeoGittern mit der aktuellen Bevölkerung (Laufende Bevölkerungsfortschreibung (LBF) am 31.12.2021) in den 10 BIK-Gemeindegrößenklassen, zeigen sich leichte Sprünge in der Verteilung. Das ist aber nicht auf die unbewohnten oder geheim zu haltenden Gitterzellen (0-1 Einwohner (-1)) zurückzuführen, sondern auf die Bevölkerungsentwicklung der letzten 10 Jahren in den amtlichen Zahlen. Die LBF am 31.12.2021 weist 83.237.124 Einwohner aus, von denen 67,2% in BIK-Regionen mit mehr als 100.000 Einwohnern leben. Die Kernstädte sind dabei am stärksten gewachsen.

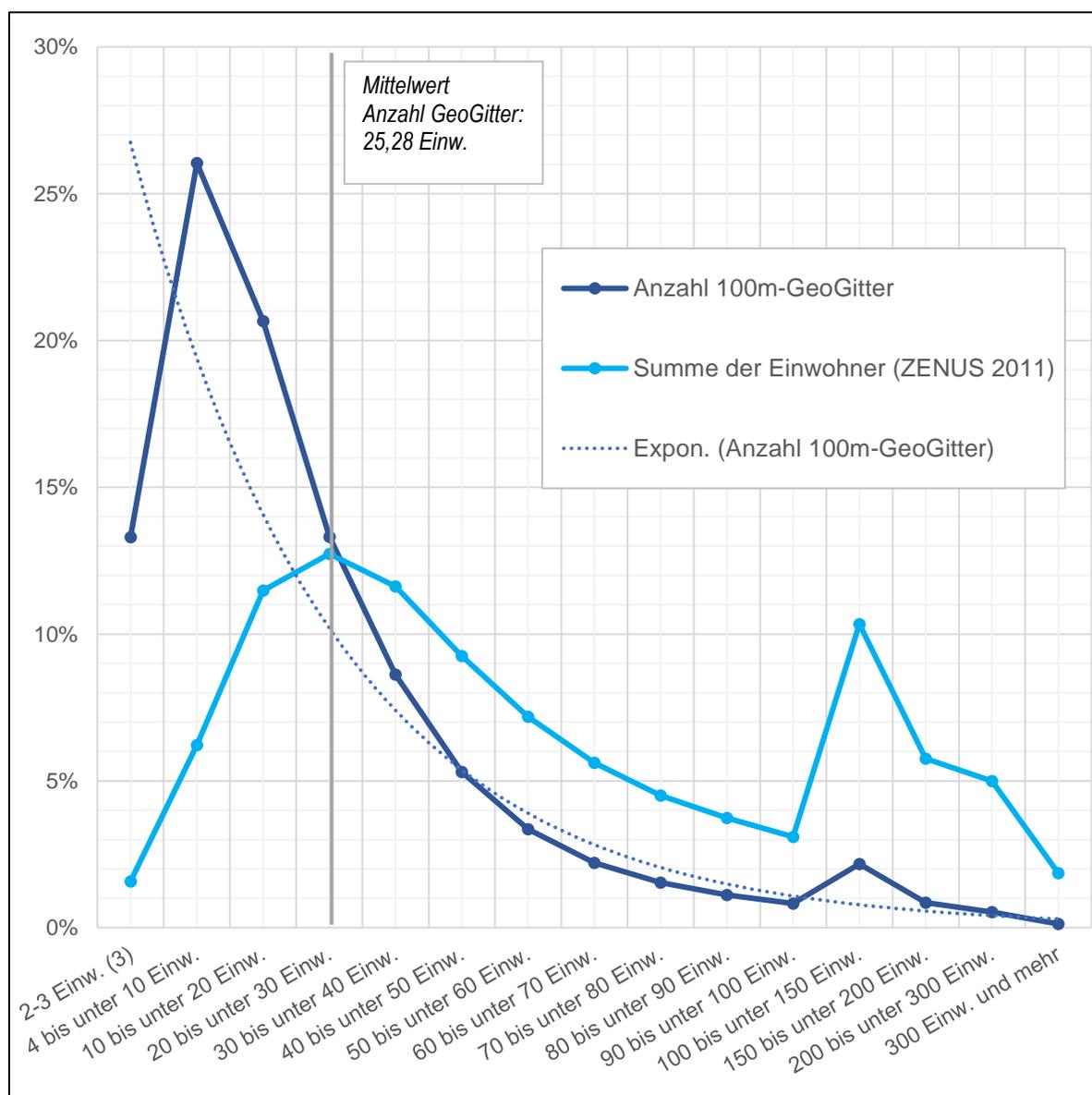
Abbildung 20: Relevante 100m-Gitter & aktuelle Bevölkerung



Für die Erstellung der ADM-Auswahlgrundlage F2F 2023, also für die Beschreibung der bewohnten Fläche der Bundesrepublik Deutschland, wurden 3.177.337 Gitterzellen mit einer Seitenlänge von 100m berücksichtigt.

In 13,3% dieser Gitterzellen leben 2-3 Personen, was aber nur 1,6% aller Einwohner laut Zensus 2011 ausmacht. Fast 830.000 Gitterzellen (26%) mit 6,2% der Gesamtbevölkerung haben 4 bis unter 10 Einwohner. Rund 50% aller Gitterzellen mit 45,1% der Zensus-Bevölkerung können in die Gruppe „10 bis unter 50 Einwohner“ eingeordnet werden. Die übrigen 12,7% 100m-GeoGitter haben mehr als 50 Einwohner. Summiert man die Bevölkerung, leben in diesen hochverdichteten Gitterzellen fast 50% aller im Zensus 2011 berücksichtigten Personen. 10,3% aller Zensus-Einwohner leben in Deutschland laut Statistischem Bundesamt in Zellen mit 100-150 Einwohnern je Hektar. Das Maximum bildet eine Gitterzelle mit 1.969 Einwohnern an der Alster in Hamburg-Mitte. Hier sind Seeleute ohne festen Wohnsitz bei ihrem Arbeitgeber, der Geschäftsstelle der Reederei Hapag-Lloyd AG, gemeldet.

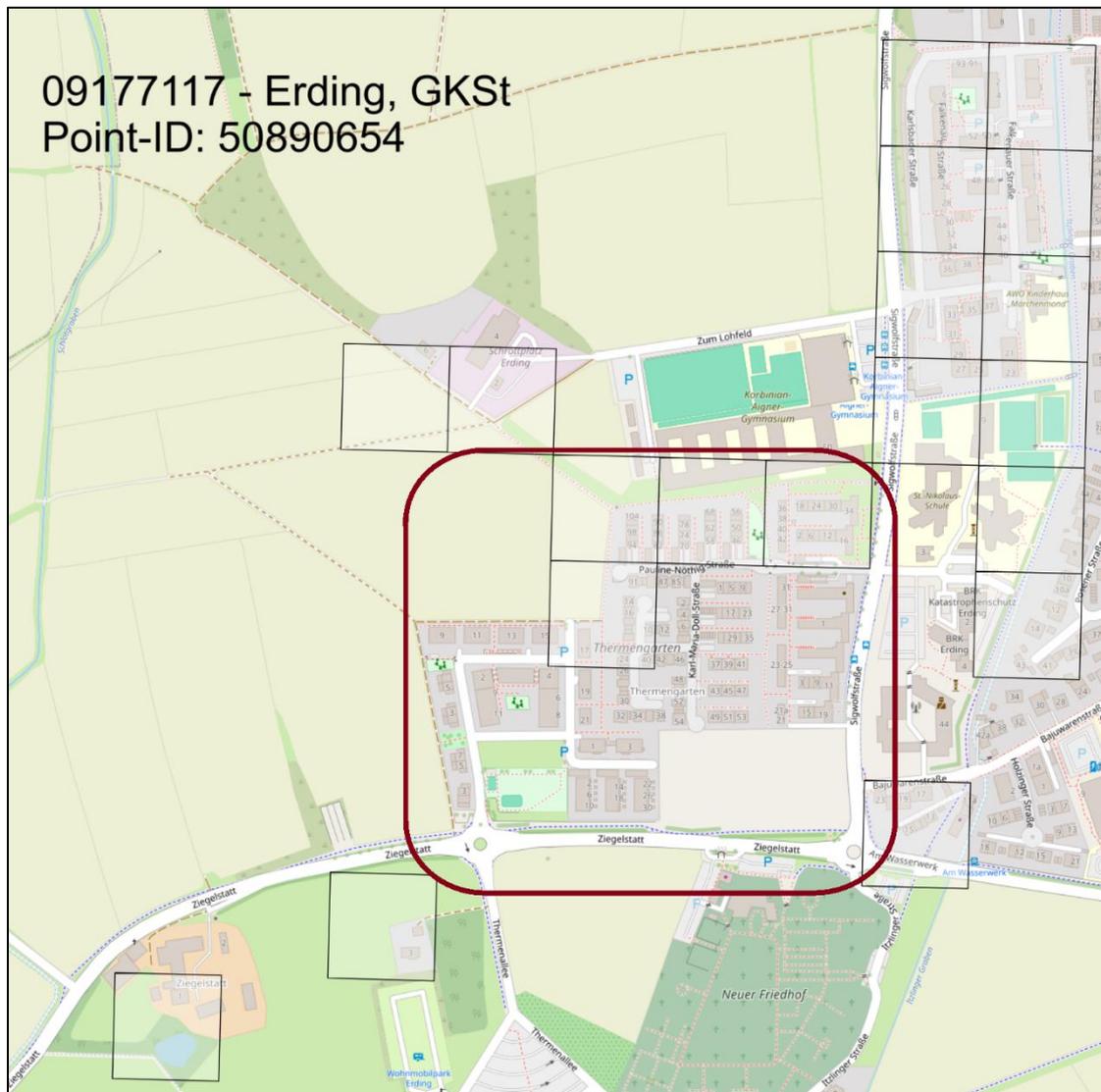
Abbildung 21: Anzahl 100m-GeoGitter und Summe der Einwohner nach Einwohnergröße



Die Berücksichtigung der Einwohnerzahl pro Gitterzelle hat noch einen weiteren Nebeneffekt: gewerblich genutzte Flächen werden nur dann im System berücksichtigt, wenn es dort auch tatsächlich gemeldete Einwohner gibt. Damit vermeidet man die Begehung im Point von reinen Gewerbegebieten, die im bisherigen System durch die alleinige Straßenlistung ohne Rücksicht auf den Einwohnerbesatz definiert waren.

Erkennbare Neubaugebiete wurden auch bereits bei der Point-Bildung, also deren Größe berücksichtigt. Diese Flächen werden vom Statistischen Bundesamt in den Zensus2011-Daten „Einwohner je Hektar“ noch als unbewohnte Gitterzellen ausgewiesen, aber in den Kartengrundlagen (OSM-Straßen/ Gebäude) oder über Kartendaten im Internet ist eine Bebauung bereits ersichtlich.

Abbildung 22: Darstellung einer Neubaufäche auf der Karte mit unbewohnten GeoGittern (nicht dargestellt)



3.5. Fortschreibung der Bedeutungsgewichte

Die Einwohner je Hektar aus dem Zensus 2011 (Statistisches Bundesamt) wurden als Bedeutungsgewicht zur Bildung der Sample-Points herangezogen. Die INSPIRE-Geogitter wurden anhand dieser Einwohnerzahlen und der Lage zu Sample-Points verdichtet.

Da der Zensus aber bereits mehr als 10 Jahre zurückliegt und sich die Bevölkerung in Deutschland durch natürliche Bevölkerungsbewegung (Geburten- und Sterbefälle) und die Wanderungstatistik über Gemeindegrenzen ständig verändert, war eine Fortschreibung des Bedeutungsgewichts je Sample-Point notwendig. Außerdem werden zur Schichtung von F2F-Stichproben häufig Privathaushalte herangezogen, auch dieses Merkmal wurde bis auf kleinräumige Ebene des Points berechnet.

Dazu wurden auf kleinster verfügbaren Regionalebene die aktuellen Bevölkerungszahlen beschafft. Für Gemeinden unter 100.000 Einwohner konnte die ADM-Gemeindedatei 2023 (Gebietsstand:

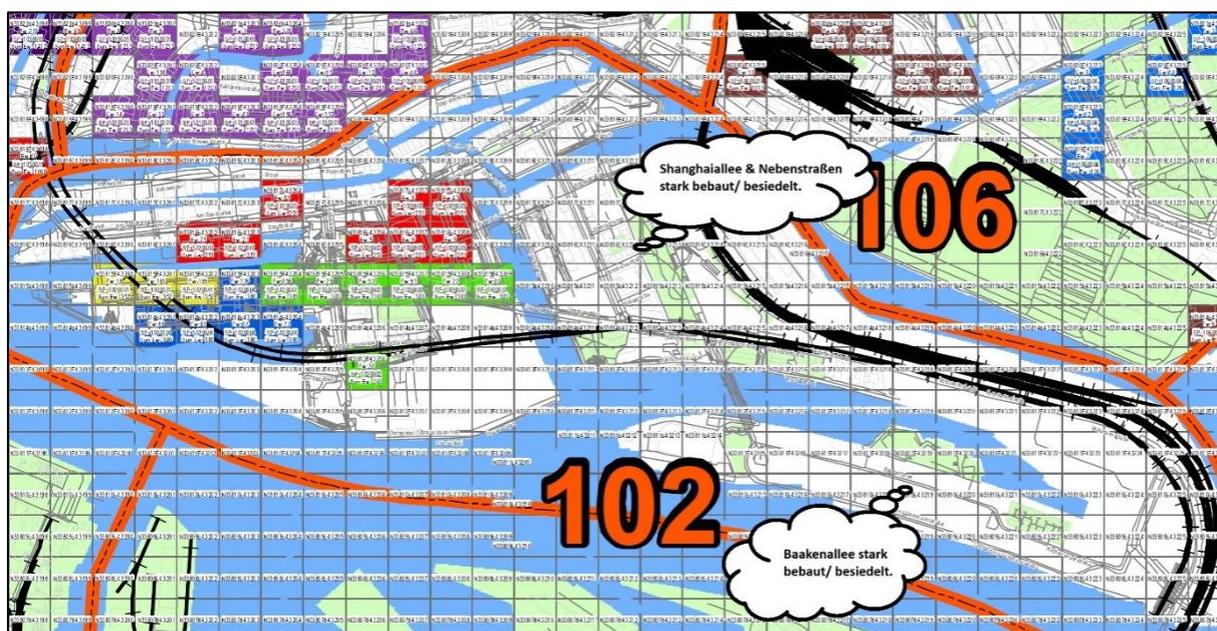
31.12.2022, Sachdatenstand: LBF am 31.12.2021/ Mikrozensus 2021) genutzt werden. Bei Großstädten mit mehr als 100.000 Einwohnern wurde zusätzlich die Städtestatistik recherchiert/ bestellt, um eine aktuelle Verteilung der Bevölkerung innerhalb der Gemeinde möglichst genau abzubilden.

Folgende Grundgesamtheiten werden in den Lieferdaten der ADM-Auswahlgrundlage-F2F 2023, angepasst an den aktuellen Rand und entsprechend der ADM-Gemeindedatei für die Sample-Points abgebildet:

- Privathaushalte gesamt (Mikrozensus 2021)
- Privathaushalte deutsch (Mikrozensus 2021)
- Bevölkerung in Hauptwohnsitzhaushalten (Mikrozensus 2021)
- Deutsche Bevölkerung in Hauptwohnsitzhaushalten ab 18 Jahre (Mikrozensus 2021)

Sample-Points, die auf Grund ersichtlicher Neubautätigkeiten bei der Point-Bildung laut Zensus 2011 vorsätzlich sehr klein gehalten wurden, mussten bei der Fortschreibung disproportional berücksichtigt werden. Als Beispiel kann der in Hamburg neu entstandene Stadtteil Hafencity (102) angeführt werden: Durch starke Bebauung seit 2011 wuchs der Stadtteil um über 360%. Der Zensus 2011 wies in der Hafencity noch 1.366 Einwohner aus, laut Auszählung aus dem Melderegister am 31.12.2021 waren hier bereits 6.302 Personen gemeldet. In Fällen wie diesen ist zu beachten, dass eine Begehung anhand der Gitterzellen-Karten oder der Begehungsdatei (Säule 2) nicht bzw. nur eingeschränkt möglich ist. Die hinterlegten Kartendaten (OSM) sind zwar immer aktuell, da diese ständig aktualisiert und abgerufen werden, aber die GeoGitter werden nicht dargestellt, da sie laut Datenquelle „Einwohner je Hektar“ (Zensus 2011, Statistisches Bundesamt) noch als unbewohnt ausgewiesen werden.

Abbildung 23: Hafencity (102), starke Besiedelung seit 2011



Folgende Sample-Points wurden mit dem Bedeutungsgewicht Zensus 2011 auf Grund von klar ersichtlichen Neubautätigkeiten kleiner geschnitten. Sie wurden disproportional fortgeschrieben und sind bei der Begehung anhand der Gitterzellen-Karten oder Begehungsdaten gesondert zu behandeln.

Tabelle 11: Points mit besonders viel Neubautätigkeiten

POINT-ID	AGS2312	Gemeindenname	STT	Stadtteil	BEV_Z2011	BEV_MZ21
737	01058009	Aukrug			738	858
768	01058050	Felde			742	890
50901241	02000000	Hamburg, Freie und Hansestadt	0102	HafenCity	380	1.140
50901243	02000000	Hamburg, Freie und Hansestadt	0102	HafenCity	251	1.012
50901245	02000000	Hamburg, Freie und Hansestadt	0102	HafenCity	383	1.149
50901247	02000000	Hamburg, Freie und Hansestadt	0102	HafenCity	352	1.145
50901942	02000000	Hamburg, Freie und Hansestadt	0106	Hammerbrook	48	1.017
50901943	02000000	Hamburg, Freie und Hansestadt	0106	Hammerbrook	242	1.098
50901944	02000000	Hamburg, Freie und Hansestadt	0106	Hammerbrook	126	1.051
50901914	02000000	Hamburg, Freie und Hansestadt	0610	Billwerder	218	1.729
33515235	03103000	Wolfsburg, Stadt	0009	Kästorf/Sandkamp	627	822
2249	03153017	Goslar, Stadt			709	861
5357	03453011	Löningen, Stadt			765	972
5358	03453011	Löningen, Stadt			907	1.017
5379	03453013	Saterland			913	1.059
5381	03453013	Saterland			902	1.147
5936	03459002	Ankum			774	1.017
5941	03459003	Bad Essen			746	962
863710	05162004	Dormagen, Stadt			877	986
863714	05162004	Dormagen, Stadt			967	1.077
33873331	06412000	Frankfurt am Main, Stadt	0003	Bahnhofsviertel	1.239	1.796
33873332	06412000	Frankfurt am Main, Stadt	0003	Bahnhofsviertel	1.370	1.779
50889828	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	943	1.116
50889829	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	929	1.104
50889832	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	961	1.137
50889846	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	942	1.116
50889870	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	964	1.141
50889874	09162000	München, Landeshauptstadt	0002	Ludwigsvorstadt - Isarvorstadt	979	1.156
50889719	09162000	München, Landeshauptstadt	0003	Maxvorstadt	904	1.136
50889731	09162000	München, Landeshauptstadt	0003	Maxvorstadt	958	1.190
50890168	09171127	Pleiskirchen			829	1.051
50890654	09177117	Erding, GKSt			730	1.079
50892975	09183128	Mühdorf a.Inn, St			239	1.331
50893269	09185149	Neuburg a.d.Donau, GKSt			665	1.101
50893276	09185149	Neuburg a.d.Donau, GKSt			930	1.085
50893281	09185149	Neuburg a.d.Donau, GKSt			819	1.067
50893766	09189159	Übersee			872	1.000
50899008	09277117	Eggllham			937	1.007
53099383	09362000	Regensburg	0006	Brandlberg - Keilberg	276	604
53099370	09362000	Regensburg	0011	Kasernenviertel	945	1.125
53099378	09362000	Regensburg	0011	Kasernenviertel	914	1.091
53099379	09362000	Regensburg	0011	Kasernenviertel	838	1.008
53099332	09362000	Regensburg	0012	Galgenberg	477	892
50899417	09373126	Freystadt, St			872	1.116
50898314	09562000	Erlangen	0007	West	264	1.272
50900870	09574123	Feucht, M			374	1.200
870721	10041516	Quierschied			765	819
871020	10044114	Rehlingen-Siersburg			750	833

3.6. Ziehung der Institutsnetze

In der Regel benutzt ein Institut für die Feldarbeit nicht das gesamte F2F-System, sondern nur einen oder mehrere Ausschnitte davon, die sogenannten Netze. Diese umfassen jeweils 258 Sample-Points und sind haushaltsproportional gezogen.

Die Berechnung der Allokation jedes Instituts-Netztes, bestehend aus 258 Points, erfolgt nach Kreis [400] x BIK-Gemeindegrößenklassen [10] proportional zu den Haushalten (MZ2021). Alle Einzel-Allokationen sind nach Cox-Rundung mit maximaler Abweichung (+/-1) pro Zelle und Schicht angepasst.

Die Ziehung der Sample-Points erfolgte gemäß Allokationstabellen zufällig innerhalb der Gemeinde/STT.

Abbildung 24: Beispielsnetz mit der Verteilung von 258 Points

		Gemeindegröße (Regionsgröße) BIK 1-10 (780)										Gesamt
		1	2	3	4	5	6	7	8	9	10	
		unter 2.000 Einw	2.000 bis unter 5.000 Einw	5.000 bis unter 20.000 Einw	20.000 bis unter 50.000 Einw	50.000 bis unter 100.000 Einw + SGTYP 2/3/4	50.000 bis unter 100.000 Einw + SGTYP 1	100.000 bis unter 500.000 Einw + SGTYP 2/3/4	100.000 bis unter 500.000 Einw + SGTYP 1	500.000 Einw und mehr + SGTYP 2/3/4	500.000 Einw und mehr + SGTYP 1	
ALTE BL	Anzahl AGS	1.210	347	999	1.252	1.107	38	2.106	154	955	210	8.378
ALTE BL	Anzahl von SP 2023	395	841	3.572	4.732	3.860	1.077	7.611	7.561	4.635	12.309	46.593
ALTE BL	Haushalte gesamt MZ21	357.900	486.792	2.255.411	3.146.746	2.556.382	728.503	4.935.555	5.553.646	3.056.901	9.183.164	32.261.000
ALTE BL	in%	0,88%	1,20%	5,56%	7,75%	6,30%	1,79%	12,16%	13,68%	7,53%	22,62%	79,46%
ALTE BL	Wert prop.	2	3	14	20	16	5	31	35	19	58	205
ALTE BL	Allokation (Points)	2	3	14	20	16	5	31	35	19	58	205
BERLIN	Anzahl AGS											1
BERLIN	Anzahl von SP 2023											2.836
BERLIN	Haushalte gesamt MZ21											1.968.000
BERLIN	in%											4,85%
BERLIN	Wert prop.											13
BERLIN	Allokation (Points)											13
NEUE BL	Anzahl AGS	237	110	411	593	353	6	502	12	162	21	2.407
NEUE BL	Anzahl von SP 2023	96	287	1.117	1.768	1.259	230	1.064	1.138	920	1.255	9.134
NEUE BL	Haushalte gesamt MZ21	90.209	180.256	741.888	1.205.183	856.145	168.104	725.931	849.128	623.241	931.915	6.372.000
NEUE BL	in%	0,22%	0,44%	1,83%	2,97%	2,11%	0,41%	1,79%	2,09%	1,54%	2,30%	15,69%
NEUE BL	Wert prop.	1	1	5	8	5	1	5	5	4	6	40
NEUE BL	Allokation (Points)	1	1	5	8	5	1	5	5	4	6	40
BRD gesamt	Anzahl AGS	1.447	457	1.410	1.845	1.460	44	2.608	166	1.117	232	10.786
BRD gesamt	Anzahl von SP 2023	491	1.128	4.689	6.500	5.119	1.307	8.675	8.699	5.555	16.400	58.563
BRD gesamt	Haushalte gesamt MZ21	448.109	667.048	2.997.299	4.351.929	3.412.527	896.607	5.661.486	6.402.774	3.680.142	12.083.079	40.601.000
BRD gesamt	in%	1,10%	1,64%	7,38%	10,72%	8,41%	2,21%	13,94%	15,77%	9,06%	29,76%	100,00%
BRD gesamt	Wert prop.	3	4	19	28	22	6	36	41	23	77	258
BRD gesamt	Allokation (Points)	3	4	19	28	22	6	36	41	23	77	258

Die Durchführung von F2F-Befragungen mit einer hohen Fallzahl erfordert in der Regel, dass mehrere Institute sich die Feldarbeit aufteilen. Für eine institutsübergreifende Zusammenarbeit ist es notwendig, die Instituts-Netze zueinander überschneidungsfrei zu ziehen, sodass nicht die gleichen Points von verschiedenen Instituten bearbeitet werden. Das gilt auch für Durchführung überschneidungsfreier Befragungswellen.

Insgesamt sind n= 67 Netze mit dem Bedeutungsgewicht „HH gesamt“ gezogen worden. Die Instituts-Netze sind zueinander überschneidungsfrei, es gibt keine Mehrfachziehungen.

Tabelle 12: Übersicht des F2F-Systems 2023 nach Instituten und Netzen

Institut	Gesamtbestand	Anzahl Netze
1 ARIS Umfrageforschung GmbH	X	4
2 IFAK-Institut GmbH & Co. KG	X	6
3 infas GmbH	X	-
4 GIM mbH	X	-
5 IPSOS GmbH	X	-
6 forsa GmbH	X	7
8 USUMA GmbH	X	2
7 Verian Group (Mantle Germany GmbH/ ehem. Kantar)	X	48 (=1)
9 GfK Germany	X	-

3.7. Datensatzbeschreibungen der Sample-Point-Datei (Säule 1)

Die Sample-Point-Datei (Säule 1) enthält alle 58.564 Sample-Points mit den dazugehörigen systematischen Informationen und Bedeutungsgewichten. Diese Daten sind aufgeteilt in 4 Tabellen, um redundante Angaben zu vermeiden und eine unkompliziertere Aktualisierung einzelner Daten zu ermöglichen. Über eine eindeutigen Point-ID (AGS für die Gemeindedaten) können alle Datentabellen miteinander verknüpft werden. Alle Sample-Point-Dateien werden für die gesamte Bundesrepublik ausgeliefert.

Tabelle 13: Säule 1-1: SP-Datei (Bildungsinformationen)

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	BevZENSUS11	Integer	Anzahl Einwohner pro Sample-Point zum Stand Zensus 2011
3	Bev2018	Integer	Bevölkerung am Ort der Hauptwohnung zum 31.12.2018 (LBF*), Zeitpunkt der Bildung der Points

LBF: Laufende Bevölkerungsfortschreibung der Statistischen Ämter am 31.12.2018

Diese Sachdatenstände wurden zur Bildung der Sample-Points herangezogen.

Die INSPIRE-Gitterzellen wurden anhand der Einwohnerzahlen des Zensus 2011(Statistisches Bundesamt) und der Lage zu Sample-Points verdichtet.

Die Zuordnung der Gemeinden in die verschiedenen F2F-Bearbeitungskategorien erfolgte anhand der Bevölkerungszahlen am 31.12.2018 (LBF), siehe dazu Kapitel 4.

- Synthetische Points, Gemeinden mit unter 2.000 Einwohner
- Sample-Points in Gemeinden mit 2.000 bis unter 100.000 Einwohner
- Sample-Points in Gemeinden mit 100.000 Einwohner und mehr

Tabelle 14: Säule 1-2: SP-Datei (Sample-Point <-> AGS2212)

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	Synth	Bool	False=kein synthetischer Point/ True=Synthetischer Point (Zusammenfassung von Gemeinden unter 2.000 Tsd. Einwohnern)
3	AGS2212	STRING	8-stelliger amtlicher Gemeindegeschlüssel (Gebietsstand: 2022/12)
4	STT	STRING	Innerstädtische Einteilungen (für Gemeinden mit mehr als 100.000 Einw.), Kommunalstatistisch verfügbare Ebene: Stadtteile, Stadtbezirke, Ortsamtsbezirke etc., möglichst entsprechend der Städtestatistik Eine Zusammenfassung der geografisch/ statistischen innerstädtischen Einteilungen für Berlin (97 Ortsteile zu 12 Bezirken) + Hamburg (104 STT zu 7 Bezirken) wird mitgeliefert, diverse weiteren übergeordneten Zusammenfassungen sind gesondert verfügbar.
5	Anz_Startadr	Integer	n = Anzahl Startadressen 0 = keine Startadressen/ Point ohne Startadressen

Diese Tabelle ist eine Schlüsseldatei mit einer n:m-Verbindung (Point-ID zu AGS 2212/ STT).

Ein Sample-Point kann mehrere Gemeinden umfassen (synthetische Points) oder eine AGS/ STT kann sich in mehrere Points aufteilen.

Tabelle 15: Säule 1-4: SP-Datei (Bedeutungsgewichte pro Point)

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	HH	Integer	Quelle: MZ2021
3	DEUHH	Integer	Quelle: MZ2021
4	BevGes	Integer	Quelle: MZ2021
5	Deuab18	Integer	Quelle: MZ2021

Tabelle 16: Säule 1-3: SP-Datei (Instituts-Netze)

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	Synth	Bool	False=kein synthetischer Point True=Synthetischer Point (Zusammenfassung von Gemeinden unter 2.000 Tsd. Einwohnern)
3	KRRNR2212	STRING	Kreiskennziffer (Gebietsstand: 2022/12)
4	KRNAME2212	STRING	Kreisname (Gebietsstand: 2022/12)
5	NUTS22212	STRING	NUTS2 Code (Gebietsstand: 2022/12), Eurostat
6	NUTSNAME2212	STRING	NUTS2 Regionen (Gebietsstand: 2022/12)
7	GKBIK10	Integer	Gemeindegröße (Regionsgröße) BIK 1-10 (780) / LBF (Sachstand: 2021/12) 1 unter 2.000 Einwohner 2 2.000 bis unter 5.000 Einwohner 3 5.000 bis unter 20.000 Einwohner 4 20.000 bis unter 50.000 Einwohner 5 50.000 bis unter 100.000 Einwohner + SGTYP 2/3/4 6 50.000 bis unter 100.000 Einwohner + SGTYP 1 7 100.000 bis unter 500.000 Einwohner + SGTYP 2/3/4 8 100.000 bis unter 500.000 Einwohner + SGTYP 1 9 500.000 Einwohner und mehr + SGTYP 2/3/4 10 500.000 Einwohner und mehr + SGTYP 1
8	Points_inZelle KreisxGKBIK10	Integer	Anzahl Sample-Points in der Schichtungszelle Kreis x GKBIK10
9	Institut	Integer	1=ARIS Umfrageforschung GMBH 2=IFAK-Institut GmbH & Co. KG 3=infas GmbH 4=GIM mbH 5=IPSOS GmbH 6=forsa GmbH 7=Verian Group (Mantle Germany GmbH/ ehem. Kantar) 8=USUMA GmbH 9=GfK Germany 99=Reserve
10	Netz	Integer	Netznummer 1-20

4. BEGEHUNG - Beschreibung der Sample-Points über Adressen aus OSM (Säule 2)

Anders als in den Vorgängersystemen bis 2010 ist mit der Umstellung der zentralen Datenquelle auf die INSPIRE-Gitterzellen und die Möglichkeiten von kartographischen Darstellungen die Bedeutung der BG-Datei zu einer Hilfsinformation für die Zuordnung der Startadressen reduziert worden. **Points sind exakt über die Geogitter und ihr Bedeutungsgewicht definiert.**

Der wesentliche Unterschied zu den bisherigen Systemen ist, dass die BG-Datei nicht mehr eine exakte Auflistung der gesamten Straßen und Straßenabschnitte umfasst, sondern die über OSM per Geo-Koordinate gewonnenen Informationen zu einem Point der Zuordnung der Startadressen dienen. Die BG-Datei ist aber nicht mehr die zentrale Säule des Systems.

Für die Erstellung der Straßeninformationen pro Point 2023 und der Zuordnung von Startadressen ist eine nutzergenerierte Datenquelle, das OpenStreetMap-Projekt (OSM), verwendet worden.

„Das OpenStreetMap-Projekt (OSM) ist eines der am häufigsten genannten Beispiele für nutzergenerierte Geodaten.“²⁵ Dabei zählt das 2004 entstandene Projekt zu der „Volunteered Geographic Information“ (VGI): die Daten werden von Nutzerinnen und Nutzern freiwillig eingetragen. Damit bildet diese Datenbank eine Teilmenge des „User Generated Content“, (UGC).

Die Entscheidung für diese Art der Nutzung fiel in der Arbeitsgemeinschaft Stichproben, nach dem es positive Testläufe auf Vollständigkeit der Quellen für verschiedene Gemeinden über ganz Deutschland hinweg gab. Rahmenbedingungen waren der Wegfall möglicher Lizenzkosten sowie Transparenz und jederzeitige Überprüfbarkeit der Quelle. Außerdem sollte ein fortschrittliches, möglichst georeferenziertes System mit Zuweisung von Koordinaten entstehen. Über die Geodaten aus OSM konnten auch Straßen- und Adressinformationen den 100x100m-Gitterzelle und dann den Sample-Points, bestehend aus GeoGittern, zugeordnet werden.

„OpenStreetMap ist ein Werkzeug, um Karteninformationen zu erstellen und zu teilen. Jeder kann zu OSM beitragen, und tausende Leute machen das jeden Tag. Nutzer zeichnen Karten auf Rechnern, statt auf Papier, aber wie sich in dieser Anleitung herausstellen wird, ist das Zeichnen einer Karte auf einem Rechner keineswegs so anders als das Zeichnen auf Papier. Wir zeichnen immer noch Linien, um Straßen, Felder und alles andere darzustellen, und wir stellen Schulen und Krankenhäuser immer noch mit Symbolen dar. Wichtig ist, dass OSM-Karten im Internet gespeichert werden, und jeder Zugriff auf sie hat, immer, umsonst. (Plennert, 2016)“²⁶

Auch verschiedene Bundesämter oder Forschungseinrichtungen haben die Qualität der OSM-Daten anerkannt und diese in unterschiedlichsten Projekten eingesetzt.²⁷

²⁵ Plennert, M. (2016): Anwendungsreif? Nutzung und Potenzial von digitalen Geodaten für Stadtforschung und Raumbewertung am Fallbeispiel OpenStreetMap“, in: STADTFORSCHUNG UND STATISTIK, 2016, Heft 1, S.27-34

²⁶ <https://learnosm.org/de/beginner/introduction/>

²⁷ Es folgen einige Berichte, Beurteilungen und Anwendungsbeispiele unabhängiger Institutionen zur Qualitätsabschätzung:

- Universität Bern „Qualität von OpenStreetMap-Daten“: https://www.digitale-nachhaltigkeit.unibe.ch/unibe/portal/fak_naturwis/a_dept_math/c_iinfamath/abt_digital/content/e90958/e490158/e900462/e977579/e977582/e980453/OpenData2020_DamianJosi_Vertiefungsartikel_ger.pdf
- Bundesamt für Kartographie und Geodäsie „Wie lässt sich die Qualität von OpenStreetMap (OSM)-Daten einschätzen?“: <https://www.bkg.bund.de/SharedDocs/Kurzmeldungen/BKG/DE/M-2023/230717-OSM.html>
- Bundesamt für Kartographie und Geodäsie „WMS TopPlusOpen“: <https://gdz.bkg.bund.de/index.php/default/wms-topplusopen-wms-topplus-open.html>
- GESIS – Leibniz-Institut für Sozialwissenschaften „Anwendungsreif? Nutzung und Potenzial von digitalen Geodaten für Stadtforschung und Raumbewertung am Fallbeispiel OpenStreetMap“: <https://www.ssoar.info/ssoar/handle/document/50501>

OSM bietet noch mehr als das: Es ist eigentlich eine Plattform, die auch Tools zur Geokodierung und zur Datenbearbeitung enthält.

Es wurden automatisiert alle relevanten Daten mit Geokoordinaten aus OSM pro Geogitter mit mindestens 2-3 Einwohnern abgefragt. Jede Gitterzelle hat eine eindeutige ID [Gitter_ID_100m], diese Gitter-ID ist ein Identifikator der Gitterzelle gemäß INSPIRE.

4.1. Die Abfrage der OpenStreetMap-Daten (OSM) zur Beschreibung der Points über die Geokoordinaten

„Die zwei wichtigsten Objekttypen in OpenStreetMap sind Nodes (also Punkte oder Knoten) und Ways (Wege). Die Attribute, die man diesen Objekten zuordnet, um zu beschreiben, um was es sich dabei handelt, werden Tags genannt. Jedes Tag wiederum besteht aus einem Schlüssel (Key) und einem Wert (Value)“. (Ramm & Topf, 2008)²⁸

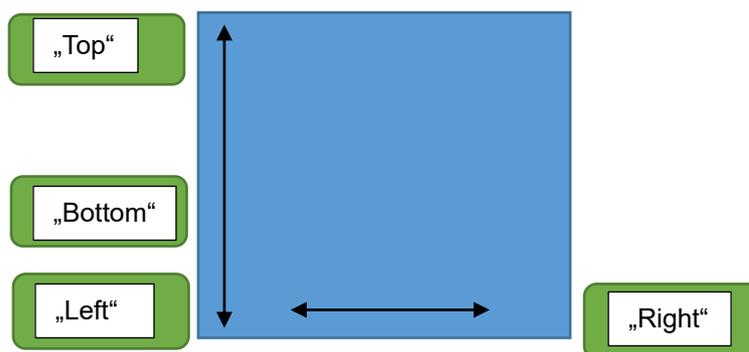
Die Eingabe von Daten erfolgt über das sog. „Tagging“. Es gibt Punkte, Linien und Flächen. Jeder Punkt hat eine Koordinate und eine Eigenschaft. Ein Objekt, z.B. eine Straße besteht aus mehreren Punkten, die eine Linie bilden, sind diese Linien geschlossen, handelt sich um eine Fläche.

Nur zu einem „Node“ (einem Knoten) gibt es eine Geokoordinate. Da wir zentral die Gitterfläche mit der Anzahl Einwohner über die Koordinate verknüpfen wollten, haben wir „Nodes“ abgefragt. Die Straßeninformationen gibt es nur zu „Ways“. Das Verbindungsglied zwischen einem Way und einem Node ist eine eindeutige Referenz-ID. Jeder Way hat mindestens 2 Knoten, maximal 2.000 Knoten.

Für die möglichste gute Beschreibung der Flächen über Straßen und Hausnummern wurden alle Sätze zu Straßen und Hausnummern, die in die vier Eck-Koordinaten einer 100x100m-Gitterzelle passen, gelistet. Zentrales Bindeglied sind die Geokoordinaten.

Zum Export wurde das Programm „OSMconvert“ genutzt, was neben der eigentlichen Abfrage über die vier Eck-Koordinaten der Gitterzellen auch Filter für die benötigten Tags und verschiedene Dateiformate als Output (u.a. auch das csv-Format) unterstützt.

Im OSM gibt es pro GeoGitter vier Punkte für die Definition, statt nur dem Koordinatenpaar der Südwestecke im Zensus 2011.



- Statistisches Bundesamt „Aufbereitung von Angaben zu Geburtsorten im Ausländerzentralregister mit OpenStreetMap“: https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2023/01/aufbereitung-angaben-geburtsort-012023.pdf?__blob=publicationFile
- Landeshauptstadt Stuttgart, Statistisches Amt „Erreichbarkeitsanalysen mit OpenStreetMap-Daten“: https://www.destatis.de/DE/Ueber-uns/Kolloquien-Tagungen/Kolloquien/2019/10_Schuetz_Praesentation.pdf?__blob=publicationFile

²⁸ Ramm, F.; Topf, J. (2008): OpenStreetMap, Kap. 5 Das Datenmodell, S.41f

Die Abfrage erfolgt über eine programmierte Schleife im OSM-Tool OSMOSIS.

```
osmosis \
  --read-xml enableDateParsing=no file=germany.osm.bz2 \
  --bounding-box top=49.5138 left=10.9351 bottom=49.3866 right=11.201 \
  --write-xml nürnberg.osm.bz2
```

4.2. Auswahl der Values pro Node (Straßen- und Adressinformationen)

Neben der Einschränkung auf die 100x100m-Flächen über die Angabe der „Bounding-Box“ wollen wir zusätzlich **nur die Informationen für die Straßen und Hausnummern pro Gitterzelle** haben. Dazu wurden nur bestimmte Keys mit ausgewählten Values genutzt.

Die Straßen und Straßenabschnitte sind in unterschiedlichen Tags vorhanden. Komplette Straßen, die in der BG-Datei an dem Hausnummernabschnitt „00000“ bis „99999“ „D“ zu erkennen sind, sind in OSM als mit dem Key „Way“ gecodet.

Es gibt aber auch einzelne Häuser mit Hausnummer; diese haben Key-Werte in der Gruppe „addr:x“.

Man hat also immer beide Keys beim Export zu berücksichtigen, um auch tatsächlich alle Straßen zu erhalten.

Der Export der OSM-Daten setzt immer auf den „Nodes“ auf. Jeder „Node“, also jeder Knoten hat eine Geokoordinate und eine eindeutige ID. Wann, von wem, ob verändert, zu welcher Uhrzeit zuletzt erfasst, interessiert uns nicht.

```
<node id="310598518" visible="true" version="8" changeset="11092640" timestamp="2012-03-25T09:10:48Z" user="Zartbitter" uid="40397" lat="54.8144513" lon="9.4719583"/>
```

Zu jeder „node_id“ gibt es weitere Attribute, die sog. „Tags“.

„Ein Attribut (engl. tag)  besteht aus einem 'Schlüssel' und einem 'Wert'. Beispielsweise wird eine Ortsstraße mit highway=residential ausgezeichnet. Schlüssel können mit Namensräumen erweitert werden, um sie zu gruppieren oder zu spezifizieren.“

In der OSM-Dokumentation werden Attribute als **schlüssel=wert** dargestellt. **Schlüssel und Werte sind freie Textfelder, die jedoch nach bestimmten Konventionen zu verwenden sind.**

- Der Schlüssel bestimmt die Art der beschriebenen Eigenschaft (z.B. highway=*). Kurz: Er macht aus einer OSM-Linie eine Straße.
- Der Wert beschreibt die Eigenschaft dann genauer (z.B. highway=motorway oder track). Kurz: Er stuft die Straße als Autobahn oder als Feldweg ein.

Hier ein paar Beispiele, wie Schlüssel und Werte verwendet werden:

- highway=residential: Ein Attribut mit dem Schlüssel highway und dem Wert residential. Üblicherweise an einer Linie verwendet, um eine Ortsstraße zu beschreiben.
- name=Zur Linde definiert den Namen eines Objektes. z.B. von Straßen, Geschäften, Kneipen, ...
- maxspeed=50 hier gibt der Wert als Zahl die erlaubte Geschwindigkeit an.
- addr:street=* und addr:city=* sind zwei Schlüssel mit dem Namensraum addr. Damit können zusammengehörige Werte gruppiert werden, wie z.B. die Teile einer Adresse.

Nur bei Bedarf sollten mehrere, per Semikolon getrennte, Werte gleichzeitig verwendet werden.“ (OpenStreetMap, <https://wiki.openstreetmap.org/wiki>, kein Datum)²⁹

„Die Elemente, aus denen alle Karten bestehen, sind die geometrischen Grundelemente und die Relationen. Ein Daten-Grundelement ist bei OpenStreetMap entweder ein Punkt (engl. node), eine Linie (engl. way) oder eine Relation. Flächen haben kein eigenes Datenelement. Sowohl den geometrischen Grundelementen als auch den Relationen können bestimmte Attribute (engl. tags)

²⁹ <https://wiki.openstreetmap.org/wiki/DE:Attribut>

zugewiesen werden. Diese Attribute bestehen immer aus einem Schlüssel (engl. key) sowie einem Wert (engl. value).“ (OpenStreetMap, <https://wiki.openstreetmap.org/wiki/DE:Elemente>, kein Datum)³⁰

Kleinste Einheit bei OSM ist also der sog. „Node“, in deutschen „Knoten“ . Zu diesem Knoten gibt es jeweils eine Koordinate mit Längen- und Breitengrad.

Das nächste Element sind die sog. „Ways“, in Deutsch „Linien“ . Diese Linien haben mindestens zwei Knoten.

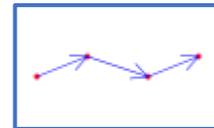
„**Linien**“ (engl. ways) in OSM entsprechen Linienzügen, also mehreren aufeinanderfolgenden Geradenabschnitten. Die Stellen, an denen sich aufeinanderfolgende Geradenabschnitte berühren, sind OSM-Punkte. Eine Linie kann aus 2 bis maximal 2000 Einzelpunkten bestehen. Mit Linien werden Straßen, Flüsse, Eisenbahnlinien und vieles mehr modelliert.

Linien werden mit Attributen versehen. Diese werden der gesamten Linie, nicht den einzelnen OSM-Punkten zwischen den Geradenabschnitten zugeordnet. Beispiel: Oberflächenbeschaffenheit und zulässige Höchstgeschwindigkeit einer Autobahn, Bundesstraße oder Landstraße. Linien müssen in jedem Fall einen Schlüssel mit einem dazu passenden Wert besitzen.

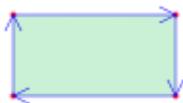
Linien haben auch eine Richtung. Diese beschreibt je nach zugewiesenem Attribut beispielsweise die Richtung einer Einbahnstraße oder die Fließrichtung eines Gewässers. Oft ist die Richtung aber auch unbedeutend wie z.B. bei einer Grenze.

Linien können auch nachträglich in kleinere Teilstücke unterteilt werden, wenn unterschiedliche Attribute hinzugefügt werden sollen, zum Beispiel ein Abschnitt mit „Zone 30“. Ebenso können mehrere Teilstücke nachträglich zu einer langen Linie zusammengefügt werden.

Punkte einer bestimmten Linie können gleichzeitig auch Punkte anderer Linien sein. Wenn sich Wege treffen (z.B. an einer Kreuzung), sollte der jeweilige Endpunkt zu beiden Wegen gehören. Nur dann gelten die Linien als verbunden. Anfangs- und Endpunkt einer Linie können auch identisch sein, so dass die Linie also geschlossen ist, z.B. bei einem Kreisverkehr.“



Eine Unterform der Linien stellen im OSM die Flächen dar . Sie werden als „geschlossene Linie“ interpretiert. „Eine Fläche (engl. area) ist eine geschlossene Linie mit besonderen Attributen. Eine Fläche ist also eine Sonderform des Linien-Elements und hat kein eigenes data primitive in der Datenbank. Bei einer geschlossenen Linie sind Anfangs- sowie Endpunkt identisch. Zur Fläche wird so ein Gebilde, sobald es explizit als solche ausgewiesen wird (durch das Attribut area=yes (en)) oder durch Attribute, die dies implizit bewirken (z.B. building=yes oder landuse=forest). Flächen werden auf den meisten Karten gesondert dargestellt.“



Aussparungen in Flächen werden als Multipolygon (Sonderform der Relation) gezeichnet. Beispielsweise eine Waldlichtung oder eine Insel. Sind Flächen so groß oder so detailliert, dass durch die 2000-Punkte-Beschränkung ihr Umfang nicht mehr mit einer einzigen Linie dargestellt werden kann, wird ebenfalls ein Multipolygon benutzt.

Der einzige Unterschied zwischen einer „normalen“ geschlossenen Linie und einer Fläche besteht in den Attributen. Beispielsweise sind eine Ringstraße oder ein Kreisverkehr keine Flächen (highway=tertiary + junction=roundabout), ein See oder ein Wald (natural=water, landuse=forest) aber schon. In seltenen Fällen kann eine geschlossene Linie auch beides sein, also eigenständige Linie und Begrenzung einer Fläche. Ein Beispiel wäre ein Kreisverkehr, der gleichzeitig Begrenzung der eingeschlossenen Grasfläche ist.“

Das dritte Element, die „Relation“, ist für unsere Zwecke nicht relevant . „Eine Relation gruppiert Punkte, Linien oder auch andere Relationen. Sie beschreibt Beziehungen zwischen Objekten oder deren Teile. Diese Objekte sind die Mitglieder (engl. members) einer Relation und haben dort eine

³⁰ <https://wiki.openstreetmap.org/wiki/DE:Elemente>

festgelegte Rolle. Eine Relation kann beliebig viele Objekte umfassen. Beispiele sind ein Wan (OpenStreetMap, <https://wiki.openstreetmap.org/wiki/DE:Elemente>, kein Datum) derweg aus der Gesamtheit seiner Wegstücke oder ein Wald aus seinen Holzungen und Lichtungen.“³¹

Grundsätzlich interessierten nur Adress-Informationen aus OSM. Beim Export wurden daher die Daten auf die folgenden Merkmale gefiltert, denen wir über die Geokoordinate dann einen Point zugewiesen haben:

```
<tag k="addr:city" v="Flensburg" />
<tag k="addr:country" v="DE" />
<tag k="addr:housenumber" v="90" />
<tag k="addr:postcode" v="24944" />
<tag k="addr:street" v="Twedter Mark" />
```

Es gibt in OSM über tausend unterschiedlichste Tags. Man kann sich über die Website <https://wiki.openstreetmap.org/wiki/DE:Taginfo> eine Liste und Hilfsfunktionen zur korrekten Benutzung der Tags laden. Hier wird auch schon ein erstes Problem bei der Nutzung von OSM deutlich. Man sollte die Eingrenzung der Tags auf nur einen Begriff nicht zu eng machen.

Aber: neben den Adress-Informationen wurde zusätzliche der Tag „highway“ genutzt, der ebenfalls einen Straßennamen enthalten kann.

Ob es sich um eine normale Ortsstraße „innerorts“ oder schon „außerorts“ handelt, ist noch klar zu deuten, aber ob es eine „Anliegerstraße“ oder ein „verkehrsberuhigter Bereich“ ist, kann man im Zweifelsfall nur noch über die Straßenschilder erkennen. Es sind nur Straßen „innerorts“ wichtig. Das können sein:

Tabelle 17: OSM-Values für den Tag „highway“

Innerorts	k = v	name
Hauptstraßen	highway=residential	name=[Straßenname]
Hauptstraße mit Übergang von einer Straßenklasse in eine andere (z.B. Bundesstraße wird Ortsstraße)	highway=unclassified	name=[Straßenname]
Bundesstraße innerorts	highway=primary	name=[Straßenname]
Durchgangsstraßen/	highway=secondary	name=[Straßenname]
Anliegerstraßen	highway=tertiary	name=[Straßenname]
Wohnstraßen (verkehrsberuhigt)	highway=living_street	name=[Straßenname]
Fahrradstraßen	highway=* + bicycle_road=yes	name=[Straßenname]
Fußwege (der Unterschied, ob nur „ausgeschildert“ oder „ausgeschildert, Anlieger frei“ ist hier egal)	highway=path+foot=designated highway=footway + foot=designated	name=[Straßenname]
Fußgängerzone (der Unterschied, ob für Radfahrer und Lieferverkehr frei oder nicht ist hier egal)	highway=pedestrian	name=[Straßenname]

Nicht interessant sind etwa die maximal erlaubte Geschwindigkeit (maxspeed=30), der Straßenbelag, die mögl. Sperrung für Kraftfahrzeuge, die Kennzeichnung von Einbahnstraßen (onway=yes) oder das Vorhandensein eines Bürgersteigs.

Um nur die Adressinformationen zu erhalten, wurden mithilfe des OSM-Tools OSMOSIS zunächst nur alle Sätze mit dem „tag k="addr:" und „highway“ genutzt. Abgefiltert wurden Autobahnen und Autobahnauffahrten.

Mit dem Befehlsparametern „—tf accept“ oder „—tf reject“ schließt man bestimmte Arten von „Ways“ aus, im o.g. Beispiel sind das die Autobahnen und Autobahn-Auf-/Abfahrten.

Zusätzlich sind auch die Informationen zur Bebauung mit Häusern interessant. In der BG-Datei sind nur Straßen und Hausnummern aufgenommen, wenn sie bewohnt sind. Das erklärt sich damit, dass

³¹ <https://wiki.openstreetmap.org/wiki/DE:Elemente>

es nur darum geht, Privathaushalte und Privatpersonen für eine Befragung zu finden. **Daher sind sog. „Point-of-Interest“ (POI) nicht Gegenstand des F2F-Systems.** POIs sind z.B. Kirchen, Kindergärten, Schulen, Gemeindezentren, Rathäuser, Marktplätze, Supermärkte etc.

Adressinformationen kann es per Definition nur für Punkte  und Flächen (geschlossene Linien)  geben. Ein Satz in OSM für eine konkrete Adresse sieht z.B. so aus:

```
\brandenburg-latest.osm.pbf --all-to-nodes --drop-relations -
b=14.561587514168709,51.847613027555695,14.88362425733277,52.04906952422888315
--csv-headline
--csv="@id @lon @lat
highway name
addr:postcode
addr:city
addr:street
addr:housenumber"
-o=test_xxx.csv
```

4.3. Weiterverarbeitung der ausgelesenen OSM-Informationen

Nach dem Auslesen der Informationen mussten die csv-Dateien pro 100x100m-Gitterzelle noch um leere Sätze bereinigt werden (Geo-Koordinate ohne weitere Info). Ebenso um Sätze, die nur die Art der Straße beinhalten, aber keine PLZ, Straßennamen oder HHNR. Das waren Tags ohne weitere Differenzierung wie z.B. „track“ (Wanderwege), „turning-circle“, „service“, „residential“, „path“ und ähnliche. Innerhalb dieses Schritts verringerte sich die Menge der ursprünglich ausgelesenen Sätze um 85-90 %.

In einer zweiten Bereinigungsstufe erfolgte dann die Zusammenfassung der GeoGitter zu einem Point. Auch hier gab es noch einmal Doppeleinträge zu bereinigen, wenn die Straßen nicht hausnummernscharf getrennt waren.

Nicht bereinigt wurde die Schreibweise der Straßennamen aus OSM, hier erfolgte eine Übernahme 1:1. Die Vollständigkeit der Straßenangaben schwankt stark über ganz Deutschland verteilt, da sich die Angaben auf die freiwillige Mitarbeit der „OSM“-Community stützt.

Zusätzlich wurde als Qualitätscheck des Auslese-Vorgangs für jedes zehnte 100x100m-Gitter eine Auslesestatistik angefordert, eine sog. „-out statistics“ der pbf-Datei. Hier gab es keinerlei Auffälligkeiten.

Abbildung 25: Beispiel für eine Abfrage in OSM

```
„timestamp min: 2008-12-31T13:53:18Z
timestamp max: 2019-03-19T17:19:44Z
lon min: 9.4165997, lon max: 9.4171801
lat min: 54.7532633, lat max: 54.7532682
nodes: 2, ways: 1, relations: 8
node id min: 323292781, node id max: 323862177
way id min: 71781307, way id max: 71781307
relation id min: 27014, relation id max: 2958987
keyval pairs max: 19
keyval pairs max object: relation 27014
noderefs max: 70
noderefs max object: way 71781307
relrefs max: 143
relrefs max object: relation 27014“
```

Eine Besonderheit beim Auslesen ist zu vermerken: für den Point 85122654, bestehend aus der ehemals selbständigen Gemeinde Passow bei Schwedt sind keine OSM-Daten gesondert ausgelesen worden. Für die Gemeinde selbst gibt es auch keine vollständigen Startadressen, da die vier privaten Einträge der Gemeinde alle keine Straßenangabe und Hausnummern aufweisen. Zur Bildung der finalen Points wurden die Gitterzelle mit den OSM-Informationen verdichtet. Doppelte Straßenangaben ohne Hausnummer wurden entfernt, sodass nur eindeutige Straßen und ggf. Hausnummern pro Point auf Basis von OSM enthalten sind.

Die Adressangaben aus OSM sind zu 90% vollständig aus PLZ, Ort, Straße und Hausnummer, was ein Anspielen der Startadressen aus Geokoordinaten plus Adressangaben erleichtert hat.

Tabelle 18: Menge Angaben pro ausgelesener OSM-Adresse nach Bundesländern

Bundesland	Anzahl OSM-Adresseinträge in der BG-Datei	dv. PLZ+Ort + Str+HHNR		Komplette Adressen pro Point	Unvollst. nur 3 Angaben		Unvollst. nur 2 Angaben		Unvollst. nur 1 Angabe	
		abs.	in%		abs.	in%	abs.	in%	abs.	in%
01 SH	513.998	460.384	89,60%	242,6	50.724	9,90%	2.220	0,40%	670	0,10%
02 HH	215.752	197.119	91,40%	142,8	13.972	6,50%	1.537	0,70%	3.124	1,40%
03 NI	1.761.151	1.586.257	90,10%	286,8	163.340	9,30%	9.349	0,50%	2.205	0,10%
04 HB	105.041	96.986	92,30%	197,5	7.070	6,70%	278	0,30%	707	0,70%
05 NW	4.519.021	4.248.273	94,00%	347,8	257.696	5,70%	3.368	0,10%	9.684	0,20%
06 HE	1.275.143	1.160.711	91,00%	258,3	106.547	8,40%	6.080	0,50%	1.805	0,10%
07 RP	862.839	773.603	89,70%	277,2	80.570	9,30%	7.740	0,90%	926	0,10%
08 BW	2.658.692	2.433.124	91,50%	298	219.243	8,20%	2.684	0,10%	3.641	0,10%
09 BY	2.234.627	1.955.947	87,50%	221,7	238.342	10,70%	25.485	1,10%	14.853	0,70%
10 SL	292.567	272.798	93,20%	336,4	19.588	6,70%	19	0,00%	162	0,10%
11 BE	293.572	268.876	91,60%	94,8	19.148	6,50%	189	0,10%	5.359	1,80%
12 BB	416.597	370.980	89,10%	206,7	40.229	9,70%	4.712	1,10%	676	0,20%
13 MV	343.734	313.439	91,20%	283,4	24.932	7,30%	5.173	1,50%	190	0,10%
14 SN	702.816	623.416	88,70%	202,3	58.614	8,30%	18.512	2,60%	2.274	0,30%
15 ST	338.235	292.667	86,50%	181,3	33.000	9,80%	11.627	3,40%	641	0,20%
16 TH	339.270	297.717	87,80%	193,7	37.693	11,10%	3.167	0,90%	693	0,20%
Deutschland	16.873.055	15.352.597	91,00%	262,2	1.370.708	8,10%	102.140	0,60%	47.610	0,30%

Die BG-Datei enthält nicht mehr wie in den bisherigen Systemen nur Straßen und Straßenabschnitte! Die Daten wurden auch nicht verdichtet, sondern es sind alle, den Point-Fläche zugewiesenen Adressen der nutzergenerierte Datenquelle „OpenStreetMap-Projekt (OSM)“ im Original enthalten.

Auch z.B. der Gemeindegemeinde – als redundante Information ist hier nicht enthalten, kann aber bei Bedarf über die AGS eindeutig zugeordnet werden. Innerstädtische Einteilung (inkl. Bezeichner der Stadtteile/ Ortsteile) bei Städten mit mehr als 100.000 Einwohner können über die PointID zugespielt werden.

Da es sich bei OSM um eine nutzergenerierte Open-Source- Datenquelle (Daten werden von Nutzerinnen und Nutzern freiwillig eingetragen und stehen kostenfrei zur Verfügung) handelt, sind nicht allen Felder (Nodes, Attributen, Elementen) vollständig und korrekt ausgefüllt. Es wurden hier aber bewusst alle verfügbaren Informationen je Gitterzelle bzw. je Sample-Point ausgelesen und beibehalten. Je mehr Informationen zur Verfügung stehen, desto mehr Telefonbuch-Einträge (Startadressen) können zugeordnet werden bzw., desto mehr Sekundärstatistik kann zum Beispiel über einen String-Abgleich angespielt werden.

Wenn die Daten dieser Säule überhaupt noch zur Begehung genutzt werden sollen, dann wird dieses als Ergänzung zur kartografischen Darstellung empfohlen. Sie ist nicht mehr, wie in früheren Systemen, zur Auflistung gedacht.

4.4. Dichte der OSM-Angaben nach Gemeindegröße und Bundesland

Die Dichte der OSM-Adressen, und damit die Güte als Zuordnungsmerkmal, lässt sich am ehesten mit der Anzahl Wohngebäude gegenprüfen. Je nach Bebauungsstruktur der Flächen gibt es mehr oder weniger Adressen. Ein Hochhaus mit vielen Einwohnern hat nur eine Adresse, dicht bebaute Gitterzellen haben viele Adressen.

Die Aufteilung nach Bundesland zeigt dann allerdings ein diverses Bild. Sachsen-Anhalt, Brandenburg und Schleswig-Holstein als Flächenstaaten weisen rechnerisch die höchste Eintragsdichte auf. Es pendeln alle Werte um 1, was auf eine gute Quellenlage hinweist.

Tabelle 19: Anzahl OSM-Einträge pro Gitterzelle und im Vergleich zu Wohngebäude nach Bundesland

Bundesland	Anzahl OSM-Adresseinträge in der BG-Datei	Anzahl relevante 100m-Gitter	Ø- Anzahl OSM-Adresseinträge pro Gitterzelle	Bestand an Wohngebäuden zum 31.12.2019	Ø- Anzahl OSM-Adresseinträge pro Wohngebäude
01 SH	513.998	143.535	3,58	826.121	1,61
02 HH	215.752	28.897	7,47	254.661	1,18
03 NI	1.761.151	432.750	4,07	2.264.496	1,29
04 HB	105.041	12.674	8,29	140.391	1,34
05 NW	4.519.021	565.160	8	3.901.200	0,86
06 HE	1.275.143	193.919	6,57	1.398.939	1,1
07 RP	862.839	166.755	5,17	1.193.920	1,39
08 BW	2.658.692	349.828	7,6	2.441.501	0,92
09 BY	2.234.627	581.333	3,84	3.087.722	1,38
10 SL	292.567	40.308	7,26	306.447	1,05
11 BE	293.572	39.604	7,41	329.115	1,12
12 BB	416.597	143.934	2,9	677.505	1,63
13 MV	343.734	93.867	3,66	397.738	1,16
14 SN	702.816	183.502	3,83	830.313	1,18
15 ST	338.235	105.365	3,21	578.747	1,71
16 TH	339.270	95.906	3,54	532.161	1,57
Deutschland	16.873.055	3.177.337	5,31	19.160.977	1,14

Die Siedlungsdichte ist in den stark verdichteten Regionen in den Kernen in den BIK-Gemeindegrößenklassen 6,8 und 10 sehr hoch. Das zeigt die Menge an OSM-Einträgen der Adressen zur Anzahl der Gitterzellen. Betrachtet man aber die Menge der Wohngebäude im Verhältnis zu den OSM-Adressen, dreht sich das Bild. Die Vermutung, dass es in kleineren Gemeinden im ländlichen Raum mehr Angaben zu den Wohnadressen gibt, kann man mit einem Vergleich der Anzahl Wohngebäude mit dem OSM-Adressen nach Erstellung des Systems als bestätigt ansehen.

Tabelle 20: Anzahl OSM-Einträge pro Gitterzelle im Vergleich zu Wohngebäuden nach BIK-Gemeindegrößenklassen

BIK-Gemeindegrößenklassen (10er)	Anzahl Einträge BG-Datei	Anzahl relevante 100m-Gitter	Ø- Anzahl OSM-Adresseinträge pro Gitterzelle	Bestand an Wohngebäuden zum 31.12.2019	Ø- Anzahl OSM-Adresseinträge pro Wohngebäude
1 unter 2 Tsd. Ew.	257.763	92.768	2,78	370.382	1,44
2 2 bis unter 10 Tsd. Ew.	359.442	107.681	3,34	479.204	1,33
3 10 bis unter 20 Tsd. Ew.	1.618.687	400.749	4,04	1.913.674	1,18
4 20 bis unter 50 Tsd. EW	2.136.559	494.015	4,32	2.521.081	1,18
5 50 bis unter 100 Tsd. Ew./ Rand	1.697.249	397.335	4,27	2.070.445	1,22
6 50 bis unter 100 Tsd. Ew./ Kern	367.684	48.730	7,55	385.472	1,05
7 100 bis unter 500 Tsd. Ew./ Rand	3.013.125	639.636	4,71	3.587.200	1,19
8 100 bis unter 500 Tsd. Ew./ Kern	2.217.126	272.191	8,15	2.229.619	1,01
9 500 Tsd. Ew. und mehr/ Rand	1.876.436	357.723	5,25	2.234.384	1,19
10 500 Tsd. Ew. Und mehr/ Kern	3.328.987	366.509	9,08	3.369.516	1,01
Gesamt	16.873.058	3.177.337	5,31	19.160.977	1,14

4.5. Datensatzbeschreibung der BG-Datei (Säule 2)

Tabelle 21: Säule 2: BG-Datei, Tabellen-Name: Begehungs-Beschreibung des POINTS

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	AGS2212	STRING	8-stelliger amtlicher Gemeindegeschlüssel (Gebietsstand: 2022/12) Dieses Feld ist redundant, aber bei Synthetischen Points notwendig, um eine Zuordnung der Straßeninformation zu einer Gemeinde eindeutig herzustellen.
3	PLZ	STRING	PLZ, Postleitzahl, die PLZ ist hier lt. OSM an die Straße und Hausnummer einer Gemeinde gekoppelt. Innerhalb eines Points können daher mehrere Postleitzahlen vorkommen. (Quelle: OSM), die nachträgliche Ergänzung der PLZ erfolgt ggf. über eine Teil-PLZ, die auf den letzten Stellen Punkte aufweist.
4	Ort	STRING	Orts-/Ortsteilname (Quelle: OSM), gehört zur postalischen Anschrift, ist nicht immer mit der AGS identisch
5	Straße	STRING	Straßenname (Quelle: OSM)
6	Hausnummer	STRING	Hausnummernangabe lt. OSM (ggf. Zusätze)

5. STARTADRESSEN (Säule 3) Telefonbucheinträge innerhalb der Sample-Points

Als dritte Säule des F2F-Systems werden aus dem öffentlich zugänglichen Telefonbuch vollständige Einträge aus „Vorname, Nachname, Straße, Hausnummer, PLZ und Ortsangaben“ über einen Abgleich der Zeichenketten mit den OSM-Daten zu den 100x100m-Gitterzellen der Samplepoints gespielt. Eine direkte Georeferenzierung der Einträge war zum Entstehungszeitpunkt des neuen Systems noch zu teuer.

Zweck dieser Startadressen seit dem System von 2003 ist es, eine Zufallsauswahl des Startpunktes der Begehung bereits im Institut zu wählen, und dies nicht dem Interviewer zu überlassen.

Grundsätzlich handelt es sich um eine „Klumpenauswahl“: aus einer begrenzten Anzahl Points werden in der Regel mehrere Adressen innerhalb eines Points ausgewählt. Dieses Vorgehen erfolgt aus ökonomischen Gründen.

HÄDER/HÄDER (Häder & Häder, 2022)³² nennen den „Random-Walk“, also die Begehung eines Points mit einer Adressauflistung, um aus dieser Liste dann eine Adresse für die Haushaltsbefragung zur Ermittlung der Zielperson zu wählen, als eine „kompliziert zu kontrollierende“ Methodik.

Zweck von Startadressen ist es, **zufällig** zu einem Zielhaushalt zu gelangen, in dem dann eine Befragung einer ebenfalls zufällig ausgewählten Zielperson stattfindet.

Für die Begehung eines Sample-Points starten die Interviewer an einer Startadresse, gehen dann einen nach Begehungsvorschrift festgelegte Strecke ab und listen alle Adressen auf. Dann kann nach drei Varianten verfahren werden:

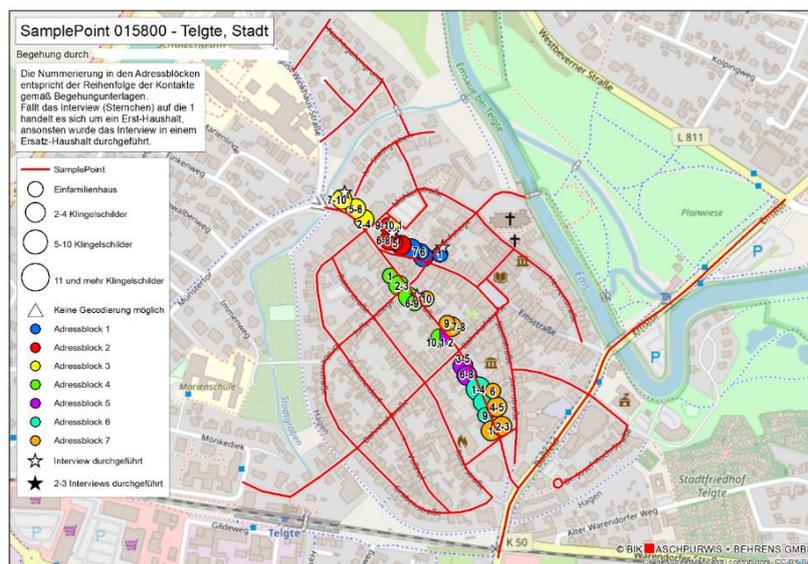
- Ein Adressauflistungsblatt wird an das Institut geschickt, und dort eine Auswahl getroffen und an den Interviewer zurückgemeldet. In der Praxis gibt es dieses Verfahren kaum noch.
- Der Interviewer darf so lange den Random-Walk fortsetzen bis er die geforderte Anzahl an Interviews für diesen Point zusammen hat. Dabei besteht aber immer die Gefahr, dass leicht zu erreichende Personen in der Stichprobe überrepräsentiert sind.
- Der Interviewer erhält eine Höchstzahl an zu kontaktierenden Haushalten. Nur bei diesen Haushalten darf er versuchen, zu einem Interview im Point zu gelangen. Erreicht er innerhalb

³² Baur, N.; Blasius, J. (Hrsg.) (2022): Handbuch Methoden der empirischen Sozialforschung, hierzu Kap. 27, Stichprobenziehung in der quantitativen Sozialforschung, S. 425 ff.

dieser Menge die geforderte Anzahl Interviews nicht, dann fallen die nicht erreichten Interviews in diesem Point aus, ggf. muss ein Ersatz-Point gezogen werden.

Nur eine aufwendige Georeferenzierung der Sample Points mit deren Straßen und Straßenabschnitten und aller gelisteten Adressen ließen bisher eine Darstellung dieses „Random Walks“ zu.

Abbildung 26: Random-Walk innerhalb eines Sample-Points



Die Stufe der Adressermittlung und Erhebung kann künftig besser nachvollzogen werden, die Laufwege können anhand der Geokoordinate der 100x100m Gitterzellen oder unter Umständen sogar über Koordinaten der Gebäude protokolliert werden. Über weitere Daten zur Sozialstruktur (Helbig, 2023)³³ eines Sample-Points, die sich über Geokoordinaten/ die INSPIRE-IDs verknüpfen lassen, kann ein Point, der u.U. ausgefallen ist, zusätzlich beschrieben werden.

Zur Kontrolle der Feldarbeit dienen häufig Nachfragen per Anruf oder Postkarte, ob ein Interview geführt wurde, und wenn ja, wie lange es gedauert hat. Um die Startpunkte genau zu bestimmen, werden als Startadressen nur vollständige Adressen zugelassen oder Geo-Koordinaten, wo Startadressen fehlen. (siehe auch Punkt 5.3)

5.1. Erstellung der Start-Adressen-Datei und Mengengerüste

Ausgangsbasis bilden die im öffentlichen Telefonbuch eingetragenen vollständigen Adresseinträge. Diese sind öffentlich zugänglich, und dürfen daher genutzt und weiterverarbeitet werden.

Zum Zusammenführen war die Schreibweise der Straßennamen entscheidend. Orts- sowie Straßennamen wurden regularisiert.

Es wurden alle Witespace-Charakter (u.a. Leerzeichen, Tab-Stopps, Bindestriche) entfernt, die Bezeichnung Str wurde einheitlich in „strasse“ umgeformt, pl .in platz, alle Sonderzeichen und Umlaute vereinheitlicht in eine Schreibweise.

5.2. Postleitzahlen für 100x100m-INSPIRE-Grids

Um für alle bewohnten GeoGitter mindestens die PLZ zu haben, wurden über eine räumliche Verbindung allen 100x100m INSPIRE-Grids Postleitzahlen zugewiesen. Das war notwendig, um die

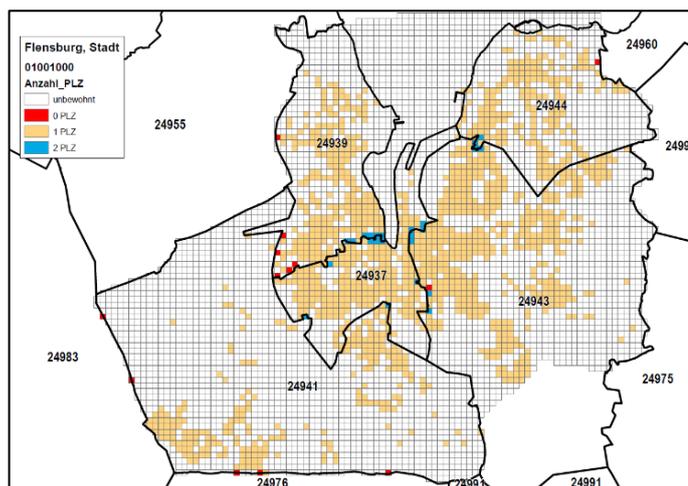
³³ Helbig, Marcel (2023), Kleinräumige Ungleichheiten in den deutschen Städten, WZB – Wissenschaftszentrum Berlin für Sozialforschung. Datenfile Version 1.0.0

Telefonbucheinträge über die Adressen zuzuordnen und um pro Point mindestens die PLZ als Info zu haben, wenn die Anzahl Startadressen zu gering ist.

Über eine automatisierte Zuordnung der 100x100m INSPIRE-Grids zu den Postleitzahlengebieten (geografische Schnittmengen) konnten 99% der GeoGitter eindeutig eine PLZ zugeordnet werden. Bei rund 8.000 Gitterzellen konnten keine PLZs gefunden werden, meist handelte es sich hier um „Randflächen“, die nicht eindeutig verbunden werden konnten. Diese „Lücken“ wurden aber durch die Zusammenfassung zu Sample Points komplett aufgehoben.

Abbildung 27: Zuordnung der PLZ zu 100x100m Gitterzellen über die geografische Schnittmengen PLZ - INSPIRE-ID

PLZ	INSPIRE - Grid 100x100m	
	Anzahl	in%
Keine PLZ	8.004	0,30%
1 PLZ	2.665.324	99,02%
2 PLZ	16.432	0,61%
3 PLZ	1.956	0,07%
4 PLZs	20	0,00%
Gesamt	2.691.736	100,00%



Danach lief die 1. Zuordnung über einen direkten Match von PLZ, Ortsname, Straße und Hausnummer. Der 2. Match erfolgte dann über AGS des Ortes, PLZ, Straße und Hausnummer. Für den 3. Match erfolgte die Zuordnung **pro Point** auf Eindeutigkeit bei Straße dann in AGS und PLZ, wenn die Straße nur in einem Point lag.

Allerdings hat sich die Menge der im Telefonbuch eingetragenen Personen seit dem Jahr 2003 von rund 30 Mio. auf nur noch rund 10 Mio. verringert. Als komplette Startadressen konnten wir 8.5 Mio. Einträge verwenden. Rund 1,8 Mio. ließen sich nicht als Startadresse nutzen, da gewerblich oder unvollständige OSM-Adressen. Da sich die Menge der zu nutzenden Adressen so stark verringert hat, wurde im Gegensatz zu den vorherigen Systemen auf eine Reduzierung der Startadressen auf 10% pro Point verzichtet.

Tabelle 22: Mengenübersicht Startadressen pro Point nach Bundesland

Bundesland	Anzahl Points	Anzahl Startadressen	Ø Anzahl pro Point
01 Schleswig-Holstein	1.898	234.797	123,7
02 Hamburg	1.380	100.488	72,8
03 Niedersachsen	5.530	868.038	157,0
04 Bremen	491	44.114	89,8
05 Nordrhein-Westfalen	12.213	1.835.465	150,3
06 Hessen	4.494	700.582	155,9
07 Rheinland-Pfalz	2.791	487.396	174,6
08 Baden-Württemberg	8.165	1.370.735	167,9
09 Bayern	8.821	1.478.575	167,6
10 Saarland	811	127.403	157,1
11 Berlin	2.836	158.729	56,0
12 Brandenburg	1.795	196.109	109,3
13 Mecklenburg-Vorpommern	1.106	133.834	121,0
14 Sachsen	3.082	367.254	119,2
15 Sachsen-Anhalt	1.614	190.118	117,8
16 Thüringen	1.537	202.929	132,0
Gesamt	58.564	8.496.566	145,1

Die Quelle der Einträge des öffentlichen Telefonbuchs hat allerdings auch Begrenzungen:

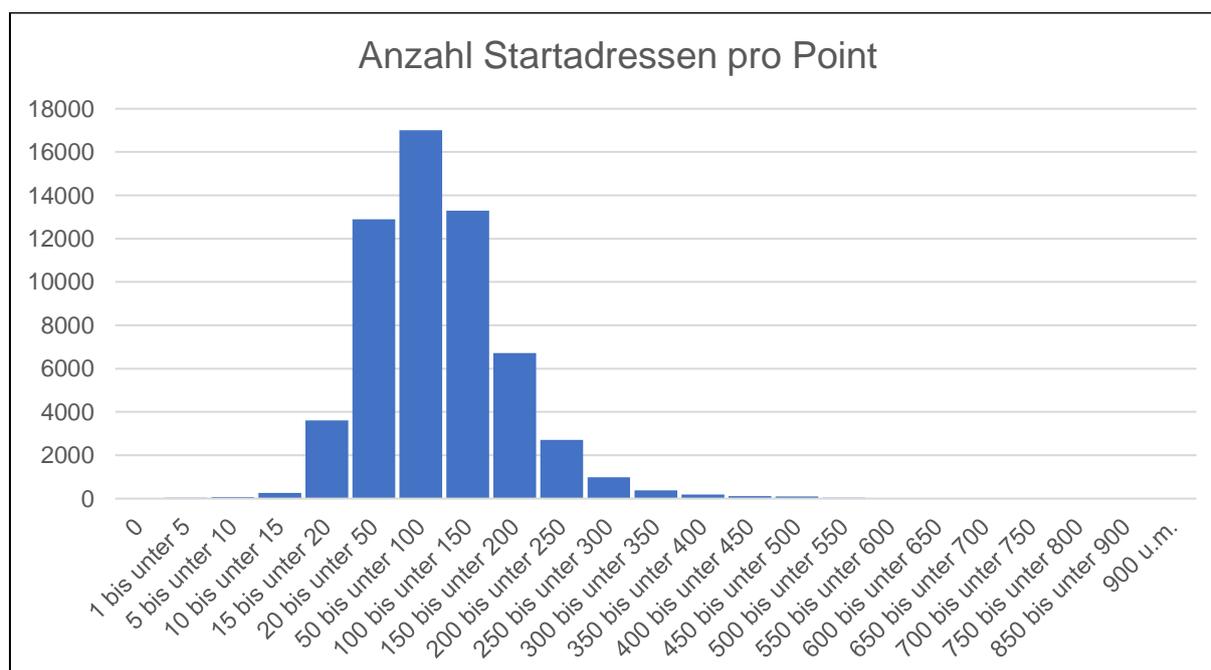
- In den Stadtstaaten (Berlin, Hamburg, Bremen) sind deutlich weniger Startadressen zu finden als in den Flächenländern. Das hat seinen Grund in der geringeren Festnetzichte.
- In Schleswig-Holstein gab es in den letzten Jahren einen massiven Kabelnetzausbau, der zu vielen Wechseln bei Festnetz-Nutzern von der Telekom zu Glasfaser-Anbietern führte. Damit sank die Menge der Einträge im hauptsächlich von der Telekom gespeisten Telefonbuch.
- Nicht deutschsprachige Vor- und Nachnamen, die über Onomastik erkannt werden können, sind ebenfalls nur noch zu geringen Teilen im Telefonbuch eingetragen. Daher sind Startadressen aus dieser Bevölkerungsgruppe unterproportional im F2F-System vorhanden.
- Das Telefonbuch umfasst inzwischen vor allem schon lange existierende Festnetzanschlüsse, die ohne Wechsel der Rufnummer seit Ende der Eintragspflicht 1990 bestehen. Damit werden vor allem ältere Personen erreicht.

Tabelle 23: Mengenübersicht Startadressen pro Point nach BIK-Gemeindegroßenklassen

BIK-Gemeindegroßenklassen (10er)	Anzahl Points	Anzahl Startadressen	Ø Anzahl pro Point
1 unter 2 Tsd. Ew.	491	155.663	317,0
2 2 bis unter 5 Tsd. Ew.	1.128	209.692	185,9
3 5 bis unter 20 Tsd. Ew.	4.689	858.050	183,0
4 20 bis unter 50 Tsd. Ew.	6.500	1.119.345	172,2
5 50 bis unter 100 Tsd. Ew./ Rand	5.119	874.901	170,9
6 50 bis unter 100 Tsd. Ew./ Kern	1.307	168.496	128,9
7 100 bis unter 500 Tsd. Ew/ Rand	8.675	1.536.022	177,1
8 100 bis unter 500 Tsd. Ew./ Kern	8.700	1.030.601	118,5
9 500 Tsd. Ew. +/- Umland	5.555	864.963	155,7
10 500 Tsd. Ew. +/- Kern	16.400	1.678.833	102,4
Gesamt	58.564	8.496.566	145,1

Auch bei der Verteilung der Startadressen pro Point kann man klare Unterschiede erkennen, die auf dem unterschiedlichen Eintragsverhalten im Telefonbuch basieren. In der BIK-Gemeindegroßenklasse 6 sind viele Städte in den neuen Bundesländern zu finden, deren Bevölkerung unterdurchschnittlich eingetragen ist. Ebenso sieht man wieder die Unterschiede beim Kern und Umland der Großstädte (Kerne: BIK-GGKL 8+10, Umland: BIK-GGKL 7+9).

Abbildung 28: Verteilung der Startadressen pro Sample-Point nach Größe



Alle diese Argumente machen deutlich, wie wichtig die Erweiterung des Systems um Gitterzellen mit Geokordinaten ist, über die per Karte eine Begehung gestartet und nachverfolgt werden kann.

Abbildung 29: Point 867321, Telgte, Stadt (05570044)



5.3. Points ohne Startadressen

Erstmalig im F2F-System gibt es aus Mangel an eindeutigen Startadressen auch 33 Points, die keine Startadresse aufweisen (vgl. Tab. 19).

Die Gründe liegen entweder

- in einem Informationsmangel beim OSM, dort fehlen Hausnummern
- oder es gibt nur gewerbliche Einträge im Telefonbuch
- oder es liegen keinerlei Straßeninformationen über den Gemeindeteil in OSM vor, ebenso nicht im Telefonbuch. Dies kam einmal vor: Im neu eingemeindeten Ortsteil Passow der Gemeinde Schwedt/Oder (Point-ID: 85122654). Für diesen Point gibt es in der BG und der ST-Datei keinen Satz. Er definiert sich über die Menge Bevölkerung (SP-Datei) und die Gitterzellen (GK-Datei).

Tabelle 24: Liste der Sample-Points ohne Startadressen

POINT-ID	Synth.	AGS2312	STT	Anzahl Startadressen	Begründung
50902513	f	02000000	504	0	keine eindeutige Straßenabgrenzung mit Hausnummern bei OSM
6585	f	03102000	1	0	keine eindeutige Straßenabgrenzung mit Hausnummern bei OSM
4635	f	03359010		0	nur 1 OSM Adresse, dazu nur gewerbliche Einträge, KORREKT
5771	f	03457013		0	nur 2 OSM Adressen, Hafestraße, nur gewerbliche Einträge
15531991	f	04012000	12	0	zu wenig passende OSM Adressen
15531992	f	04012000	12	0	nur 4 OSM Adressen, ohne Hausnummernzusatz
33873461	f	06412000	19	0	nur 1 OSM Adresse, dazu nur gewerbliche Einträge, KORREKT
854365	f	06414000	16	0	keine passende OSM Adresse
854408	f	06414000	7	0	keine passende OSM Adresse
869683	f	07339030		0	keine passende OSM Adresse
34507933	f	08221000	15	0	keine passende OSM Adresse
53099393	f	09362000	13	0	keine passende OSM Adresse
50899439	f	09373147		0	keine passende OSM Adresse, keine Hausnummerneinschränkung

6. GEOKOORDINATEN (Säule 4)

Für die Darstellung und Analyse statistischer Sachverhalte ist die Verwendung von regelmäßigen Gitterzellen als Träger semantischer Informationen ein etabliertes Verfahren. Durch die gleichmäßige Verteilung und einheitliche Größe der Gitterzellen lassen sich statistische Sachverhalte zeitstabil und vergleichbar auswerten. (Geodäsie)³⁴

Das ADM-Flächenstichprobensystems Face-to-Face 2023 wurde zur Bildung der Sample-Points um die Geodateninfrastruktur „INSPIRE - Geographical Grid Systems“ und somit um Geokoordinaten erweitert.

Das ermöglicht nicht nur die einheitliche Bearbeitung und Definition der Sample-Points (siehe Säule 1 – SAMPLE-POINTS) für die gesamte Bundesrepublik, sondern auch eine Anreicherung der Auswahlgrundlage um georeferenzierte Sekundärstatistik und es erleichtert die Aktualisierung der Bedeutungsgewichte (z.B. Zensus 2022).

6.1. Was ist INSPIRE?

INfrastructure for **SP**atial **InfoR**mation in **EU**rope (INSPIRE) ist das Vorhaben der Europäischen Union (EU) für eine gemeinsame Geodateninfrastruktur in Europa. Sie will damit die grenzübergreifende Nutzung von Geodaten in Europa erleichtern, insbesondere um gemeinschaftliche umweltpolitische Entscheidungen zu unterstützen. Das Europäische Parlament und der Rat verabschiedeten dazu die Richtlinie 2007/2/EG. Diese trat am 15. Mai 2007 in Kraft und wurde inzwischen von den Mitgliedstaaten in nationales Recht umgesetzt.

Das Bundesamt für Kartographie und Geodäsie (BKG) stellt sogenannte Geographische Gitter mit Anforderungen an INSPIRE Grids und an nationale Gittersysteme bereit. Die Daten stehen als GeoPackage, im CSV-Format sowie als ESRI-Shapefile im Open-Data-Bereich des BKG zur Verfügung (<https://gdz.bkg.bund.de/index.php/default/open-data.html>).

Geographische Gitter (GeoGitter) - Auszug

Geodaten des Bundesamtes für Kartographie und Geodäsie (BKG)
https://www.bkg.bund.de/SharedDocs/Produktinformationen/BKG/DE/P-2020/200310_GeoGitter.html

- **Inhalt:** Geographische Gitter verschiedener Kachelgrößen für die Darstellung und Analyse statistischer Sachverhalte.
 > inklusive Amtlichem Gemeindeschlüssel (für Gitter 100m und 250m) der flächenmäßig dominierenden Gemeinde
- **Gebiet:** Bundesrepublik Deutschland
- **Räumliche Gliederung:** INSPIRE-Kachelsystem (INSPIRE-Referenzsystem LAEA, Lambertsche flächentreue Azimutalprojektion Ellipsoid GRS80, Datum ETRS89)
- **Auflösung:** 100m, 250m, 500m, 1km, 5km, 10km, 100km
- **Datenformate:** CSV, Shape

Die hochauflösenden Gitterzellen 100m und 250m erhalten neben grundlegenden Informationen wie den Anteil der Fläche Land/ Wasser eine Zuordnung der Gemeinde über den Amtlichen Gemeindeschlüssel (AGS) der flächenmäßig dominierenden Gemeinde.

Über den Identifikator der Zelle gemäß INSPIRE (INSPIRE- ID, Beispiel: 100mN32670E45141) lassen sich die Koordinaten im **INSPIRE Referenzsystem LAEA**, Lambertsche flächentreue Azimutalprojektion Ellipsoid GRS80, Datum ETRS89 („INSPIRE-Koordinaten“) ablesen und

³⁴ https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/geogitter.pdf, Seite 4

Ergebnisse aus dem Zensus 2011 „Einwohner je Hektar“ (Statistisches Bundesamt, Destatis) können anspielt werden.

Abbildung 30: Deutschland INSPIRE - 100km x 100km GeoGitter



6.2. Einwohner je Hektar

Der Zensus 2011 umfasste erstmalig eine sogenannte „Georeferenzierung“. Im Rahmen der flächendeckenden Gebäude- und Wohnungszählung für Deutschland wurden alle Anschriften geocodiert/ georeferenziert. Zu dieser Datei aller Anschriften wurden dann die Informationen aus den Melderegistern und der erwerbstatistischen Register der Bundesagentur für Arbeit hinzugegestellt.

Jede Anschrift ist über ihre Geokoordinate einer Gitterzelle mit einer Seitenlänge von 100m zugeordnet. Dabei ist zu berücksichtigen, dass die Anschrift die kleinste regionale Zuordnung im Zensus ist. Fällt eine Anschrift durch ihre geografische Lage in mehrere Gitterzellen, so werden die Zensuswerte nur an einer Gitterzelle ausgewiesen. Die Zuordnung der Zensuswerte zu einer Gitterzelle ergibt sich immer aus der Geokoordinate, die der Anschrift zugewiesen ist. Hauptquelle für die Geokoordinaten und damit für die Zuordnung der Anschriften zu Gitterzellen sind die Angaben des Bundesamtes für Kartografie und Geodäsie. Liegen die Ergebnisse je Anschrift vor, wird unter Verwendung der Gitterzellennummer der aggregierte Wert ermittelt. (www.zensus2011.de/DE/Infothek/Begleitmaterial_Ergebnisse/Begleitmaterial_node.html, kein Datum)³⁵

Auswertbar sind diese Daten seit einer Änderung des Bundesstatistikgesetz (BstatG) zum 01.08.2013, erste Veröffentlichungen wurden 2015 online gestellt.

Damit gibt es erstmalig einheitliche Flächen, die flächendeckend über Geokoordinaten beschrieben werden und denen amtliche Bevölkerungszahlen laut Zensus 2011 zugewiesen sind.

Einwohner je Hektar

Statistisches Bundesamt, Destatis

<https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html>

- Basis der Auswertung sind Ergebnisse des Zensus 2011 mit Stichtag 9. Mai 2011.
- Die Zensusergebnisse wurden für diese Auswertung in INSPIRE-konforme 100x100m-Gitterzellen zusammengefasst.
- Die Informationen des Zensus 2011 zu Personen, Haushalten, Familien, Gebäuden und Wohnungen liegen flächendeckend vor, wobei jeder Datensatz einer Anschrift zugeordnet ist. Jede Anschrift wiederum ist über ihre Geokoordinate einer Gitterzelle mit einer Seitenlänge von 100m zugeordnet.
- Die Ergebnisse des Zensus 2011 werden vor ihrer Veröffentlichung mit einem datenverändernden Verfahren geheim gehalten. Dieses Verfahren kann in Einzelfällen zu größeren Abweichungen zwischen dem originalen Wert und dem veröffentlichten, geheim gehaltenen ausgezählten Wert führen.
- Die Einwohnerzahl (Bevölkerung insgesamt) wird durch die statistische Geheimhaltung nicht verändert. Lediglich Gitterzellen mit einzelnen Personen werden wie Gitterzellen ohne Personen ausgewiesen und Gitterzellen mit nur zwei Personen wie Gitterzellen mit drei Personen.

© Statistisches Bundesamt, Wiesbaden 2015

Auswertungen dieser Daten ergaben, dass von den insgesamt 38,17 Mio. Gitterzellen mit einer Seitenlänge von 100m rund 3,18 Mio. Gitter für die Abbildung der bewohnten Fläche der Bundesrepublik Deutschland relevant sind.

Auf Grund der statistischen Geheimhaltung werden Gitterzellen mit einzelnen Personen vom Statistischen Bundesamt wie Gitterzellen ohne Einwohner ausgewiesen (-1), diese Flächen müssen für alle weiteren Betrachtungen wie unbewohnte Gitter behandelt werden. Gitterzellen mit nur zwei Personen werden wie Gitterzellen mit drei Personen veröffentlicht (3), hier wird für die Erstellung der ADM-Auswahlgrundlage-F2F 2023 immer mit 3 Einwohner gerechnet.

³⁵https://www.zensus2011.de/SharedDocs/Downloads/DE/Pressemitteilung/DemografischeGrunddaten/Datensatzbeschreibung/Bevoelkerung_100m_Gitter.xlsx?__blob=publicationFile&v=2

Tabelle 26: Relevante 100m-GeoGitter

Bevölkerung in Gitterzellen	Anzahl 100m-Gitter	
	abs.	in %
0-1 Einwohner (-1)	32.608.117	85,4%
2-3 Einwohner (3)	423.013	1,1%
4 Einwohner und mehr	2.754.710	7,2%
kein Bevölkerungswert laut Destatis für Gitterzellen*	2.388.027	6,3%
Gesamt	38.173.867	100,0%

*kein Bevölkerungswert laut Destatis für Gitterzelle: Diese Gitterzellen sind zwar in den Geodaten des BKG enthalten, aber nicht im Zensusdatensatz ‚Einwohner je Hektar‘ vom Statistischen Bundesamt.

Unbewohnte oder geheim zuhaltende Gitterzellen (0-1 Einwohner (-1)) werden bei der Erstellung des Flächenstichprobensystems nicht weiter berücksichtigt. Diese Flächen sind durch eine vielfältige Nutzung geprägt. Den Großteil machen landwirtschaftliche Nutzflächen, Wälder und Gehölze, Seen, Flüsse, Kanäle und nahe Küstengewässer aus, aber vor allem in und nahe den besiedelten Gebieten können verschiedenste Nutzungsarten der Gitterzellen erkannt werden: Gewerbefläche, Verkehrswege, Flughäfen, ehemalige Militäranlagen, Sport-, Freizeit und Erholungsflächen, Friedhöfe oder Kleingartenanlagen.

6.3. Verknüpfung GeoGitter & Einwohner je Hektar

Über den Identifikator der Zelle gemäß INSPIRE (INSPIRE- ID) wurden die Geographische Gitter (GeoGitter, BKG) mit den Veröffentlichungen des Statistischen Bundesamtes „Einwohner je Hektar“ verknüpft.

Neben den unbewohnten oder geheim zuhaltenden Gitterzellen können 2.796.114 GeoGitter (100m) im weiteren Prozess nicht berücksichtigt werden, da sie entweder im Datensatz vom Statistischen Bundesamt nicht enthalten waren, gemeindefreien Gebieten zuzuordnen sind oder vom BKG nicht regional verortet wurden.

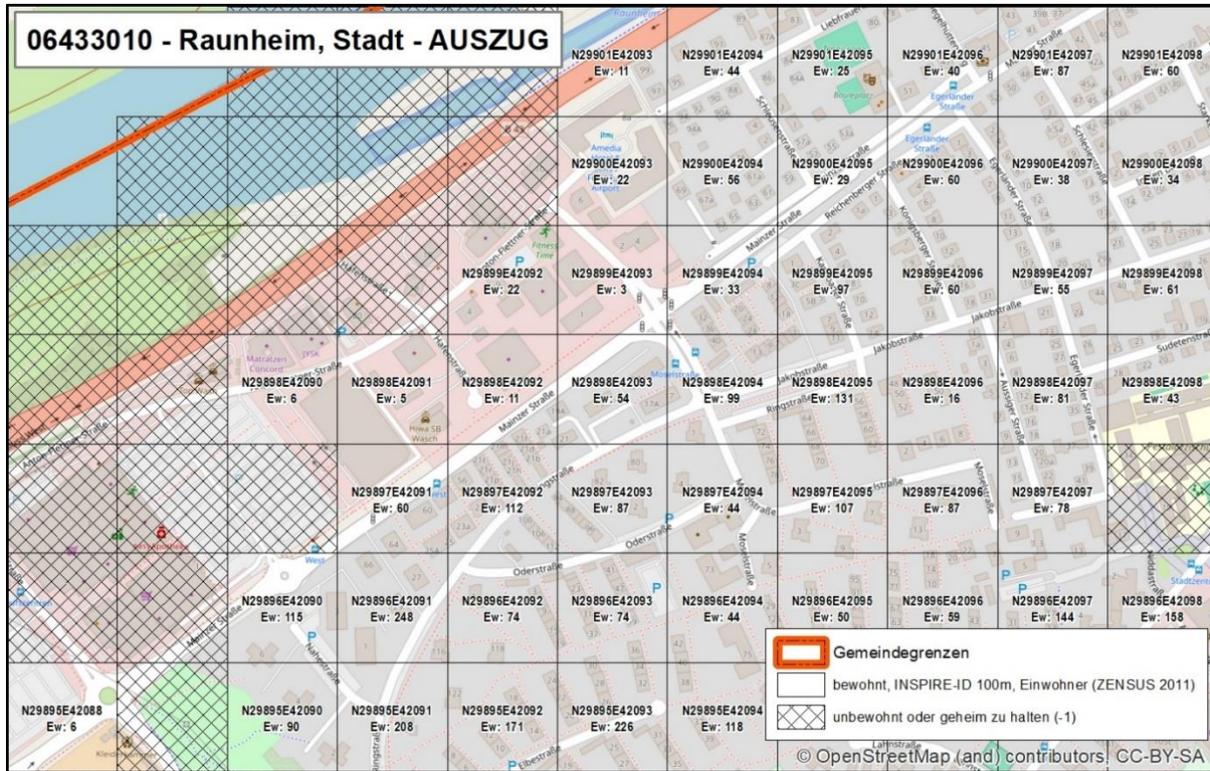
Tabelle 27: Unbewohnte oder geheim-zuhaltende GeoGitter

Bundesland*	Anzahl Gemeinden 31.12.2022		Fläche in km ²		Anzahl 100m-Gitter		Anzahl relevante 100m-Gitter		Anzahl unbewohnte/ geheime 100m-Gitter	
	abs.	in %	abs.	in %	abs.	in %	abs.	in %	abs.	in %
01 SH	1.106	10,3%	15.705	4,4%	1.564.667	4,4%	143.535	4,5%	1.421.132	4,4%
02 HH	1	0,0%	755	0,2%	75.155	0,2%	28.897	0,9%	46.258	0,1%
03 NI	941	8,7%	46.583	13,2%	4.661.895	13,2%	432.750	13,6%	4.229.145	13,1%
04 HB	2	0,0%	419	0,1%	40.312	0,1%	12.674	0,4%	27.638	0,1%
05 NW	396	3,7%	34.112	9,6%	3.414.082	9,7%	565.160	17,8%	2.848.922	8,8%
06 HE	422	3,9%	20.789	5,9%	2.078.800	5,9%	193.919	6,1%	1.884.881	5,9%
07 RP	2.301	21,3%	19.852	5,6%	1.986.234	5,6%	166.755	5,2%	1.819.479	5,7%
08 BW	1.101	10,2%	35.673	10,1%	3.570.953	10,1%	349.828	11,0%	3.221.125	10,0%
09 BY	2.056	19,1%	68.369	19,3%	6.831.770	19,3%	581.333	18,3%	6.250.437	19,4%
10 SL	52	0,5%	2.570	0,7%	258.022	0,7%	40.308	1,3%	217.714	0,7%
11 BE	1	0,0%	891	0,3%	89.124	0,3%	39.604	1,2%	49.520	0,2%
12 BB	413	3,8%	29.654	8,4%	2.966.944	8,4%	143.934	4,5%	2.823.010	8,8%
13 MV	726	6,7%	23.294	6,6%	2.315.831	6,5%	93.867	3,0%	2.221.964	6,9%
14 SN	419	3,9%	18.450	5,2%	1.848.104	5,2%	183.502	5,8%	1.664.602	5,2%
15 ST	218	2,0%	20.464	5,8%	2.055.580	5,8%	105.365	3,3%	1.950.215	6,1%
16 TH	631	5,9%	16.202	4,6%	1.620.280	4,6%	95.906	3,0%	1.524.374	4,7%
Deutschland	10.786	100,0%	353.783	100,0%	35.377.753	100,0%	3.177.337	100,0%	32.200.416	100,0%

*Verwendete Abkürzung siehe: <https://www.destatis.de/DE/Methoden/abkuerzung-bundeslaender-DE-EN.html>

Durch die Verbindung dieser amtlichen Daten lässt sich die bewohnte Fläche der Bundesrepublik kartographisch darstellen. Layert man zusätzlich die OpenStreetMap (OSM) Grundkarten, können detaillierte Strukturen, wie Wasser, Straßen, Bahnlinien oder Gebäude aufgezeigt werden.

Abbildung 31: Raunheim, Stadt - AUSZUG - INSPIRE: 100x100m GeoGitter inkl. OpenStreetMap & Einwohner je Hektar (Zensus 2011)



6.4. Dichte-Merkmal

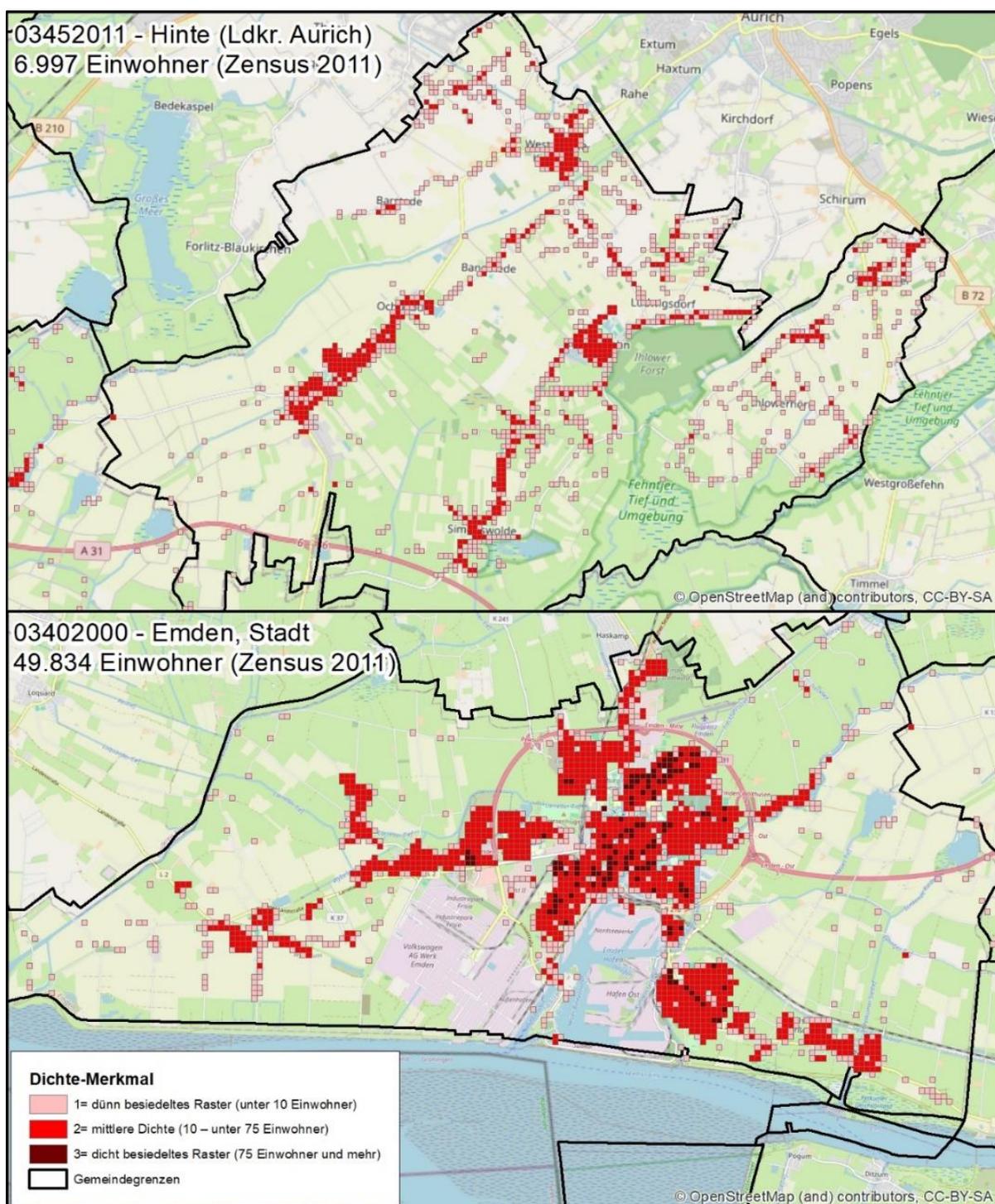
Die BIK ■ ASCHPURWIS + BEHRENS GMBH hat anhand der Siedlungsstrukturen ein Dichte-Merkmal festgelegt. Die Ausprägung 1 beschreibt dünn besiedelte Gitterzellen mit weniger als 10 Einwohnern. GeoGitter mit mittlerer Dichte weisen 10 bis unter 75 Einwohner auf, in diese Gruppe fallen zum Beispiel Siedlungen aus Einfamilienhäusern oder Reihenhaussiedlungen. Dicht besiedelte Gitterzellen (75 Einwohner und mehr) sind in der Regel mit Wohnblocks/ Hochhäusern bebaut.

Tabelle 28: Dichte-Merkmal

Dichte-Merkmal	Anzahl 100m-GeoGitter		Summe Einwohner (Zensus 2011)	
	abs.	in%	abs.	in%
1= dünn besiedeltes GeoGitter (unter 10 Einwohner)	1.250.421	39,4%	6.266.052	7,8%
2= mittlere Dichte (10 – unter 75 Einwohner)	1.725.687	54,3%	48.437.116	60,3%
3= dicht besiedeltes GeoGitter (75 Einwohner und mehr)	201.229	6,3%	25.618.770	31,9%
Gesamt	3.177.337	100,0%	80.321.938	100,0%

©

Abbildung 32: Dichte-Merkmal im Stadt-Land-Vergleich



6.5. Datensatzbeschreibung der Geokoordinaten-Datei (Säule 4)

Die Geokoordinaten-Datei (Säule 4) enthält alle bewohnten Geographischen Gitter (lt. „Einwohner je Hektar“ - Statistische Ämter des Bundes und der Länder, Zensus 2011) in der Auflösung 100m mit Zuordnung zum amtlichen Gemeindegrenzen der flächenmäßig dominierenden Gemeinde und zum Sample-Point. Diese 100m-INSPIRE-Gitterzellen wurden für Gemeinden mit 100.000 Einwohnern und mehr (politische Gemeindegrößen-klassen 6-7) zusätzlich mit innerstädtischen Einteilungen (möglichst amtliche statistische Einheiten, wie Stadtbezirke/ Stadtteile/ Ortsämter o.ä.) verbunden.

Tabelle 29: Datensatzbeschreibung der Geokoordinaten-Datei (Säule 4)

Pos.	Feldname	Typ	Inhalt
1	PointID	Integer	Eindeutige Kennung des Sample-Points (8- Stellen)
2	INSPIRE-ID	STRING	Identifikator der Zelle gemäß INSPIRE* (ID, Beispiel: 100mN32670E45141), Gitterauflösung: 100m Basis: Geographische Gitter für Deutschland, BKG
3	AGS2212	STRING	8-stelliger amtlicher Gemeindegeschlüssel (Gebietsstand: 2022/12)
4	STT	STRING	Innerstädtische Einteilungen (für alle Städte über 100.000 Einwohnern), Kommunalstatistisch verfügbare Ebene: Stadtteile, Stadtbezirke, Ortsamtsbezirke etc., möglichst entsprechend der Städtestatistik Eine Zusammenfassung der geografisch/ statistischen innerstädtischen Einteilungen für Berlin (97 Ortsteile zu 12 Bezirken) + Hamburg (104 STT zu 7 Bezirken) wird mitgeliefert, diverse weiteren übergeordneten Zusammenfassungen sind gesondert verfügbar.
5	Y_SW	Double	Y-Koordinate der Süd-West-Ecke der Zelle (WGS 84)**
6	X_SW	Double	X-Koordinate der Süd-West-Ecke der Zelle (WGS 84)**

*Der Identifikator einer Zelle (**INSPIRE-ID**) setzt sich zusammen aus:

- der Größe der Zelle (Beispiel: 100m)
- dem Symbol „N“ für die Koordinatenachse in Richtung „North“
- dem Quotienten aus der Koordinate der linken unteren Ecke in Nordrichtung (in Metern) und der größten Zehnerpotenz, durch die sich die Zellengröße (in Metern) ganzzahlig dividieren lässt
- dem Symbol „E“ für die Koordinatenachse in Richtung „East“
- dem Quotienten aus der Koordinate der linken unteren Ecke in Ostrichtung (in Metern) und der größten Zehnerpotenz, durch die sich die Zellengröße (in Metern) ganzzahlig dividieren lässt

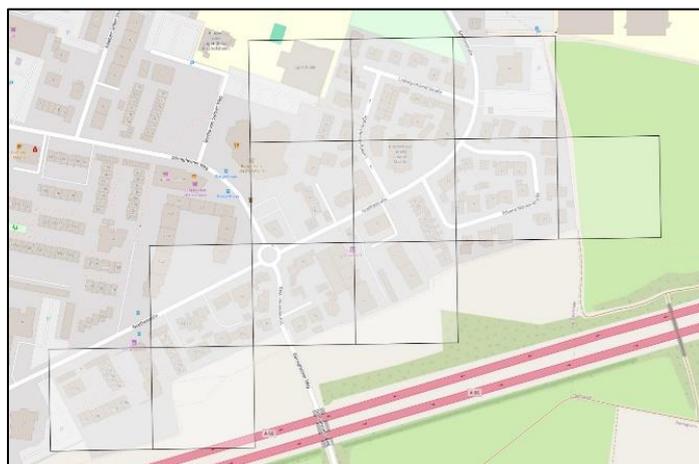
** WGS84- Koordinaten

Die Koordinaten aller bewohnten 100m-Gitterzellen (Süd-West-Ecke) wurden ins „World Geodetic System 1984 (WGS 84)“ übersetzt (8 Nachkommastellen, Bsp.: 54.80417561/9.415489109)

Tabelle 30: Datenbeispiel der GK-Datei: Maintal, Stadt, POINT-ID: 855603

PointID	INSPIRE-ID	AGS2212	STT	X_SW (WGS84)	Y_SW(WGS84)
855603	100mN30044E42360	06435019		8.81089178	50.14566361
855603	100mN30044E42361	06435019		8.81229051	50.14567817
855603	100mN30045E42361	06435019		8.81226811	50.14657714
855603	100mN30045E42362	06435019		8.81366687	50.14659169
855603	100mN30045E42363	06435019		8.81506562	50.14660622
855603	100mN30046E42362	06435019		8.81364450	50.14749067
855603	100mN30046E42363	06435019		8.81504328	50.14750520
855603	100mN30046E42364	06435019		8.81644206	50.14751971
855603	100mN30046E42365	06435019		8.81784085	50.14753421
855603	100mN30047E42362	06435019		8.81362212	50.14838964
855603	100mN30047E42363	06435019		8.81502093	50.14840417
855603	100mN30047E42364	06435019		8.81641974	50.14841868

Abbildung 33: Datenbeispiel der GK-Datei: Maintal, Stadt, POINT-ID: 855603



6.6. Koordinaten-Systeme

Aus der INSPIRE-ID können die Koordinaten im **INSPIRE Referenzsystem LAEA**, Lambertsche flächentreue Azimutalprojektion Ellipsoid GRS80, Datum ETRS89 („INSPIRE-Koordinaten“) abgelesen werden. Das European Terrestrial Reference System 1989 (ETRS89) ist ein geodätisches, kartesisches Referenzsystem mit erdzentriertem, erdfestem System (ECEF), das auf die eurasische Platte fixiert ist und hauptsächlich für die europäischen Gebiete verwendet wird.

Beispiel INSPIRE Referenzsystem LAEA [INSPIRE-ID: 100m N 30044 E 42360]

- x_sw (ETRS89-LAEA): 3004400
- y_sw (ETRS89-LAEA): 4236000

Die BIK ■ ASCHPURWIS + BEHRENS GMBH hat die „INSPIRE-Koordinaten“ in das weltweit definierte Raumbezugssystem **World Geodetic System 1984 (WGS84)** umgerechnet, auf dem alle gängigen Satellitenpositionierungssysteme basieren.

Beispiel World Geodetic System 1984 (WGS84) [INSPIRE-ID: 100m N 30044 E 42360]

- x_sw (WGS84): 8.81089178 (Längengrad/ longitude)
- y_sw (WGS84): 50.14566361 (Breitengrad/ latitude)

Das System erzeugt eine geozentrische und konsistente Basislinie mit einer Genauigkeit von bis zu plus oder minus einem Meter; die WGS84-Koordinaten ortsfester Punkte ändern sich aufgrund der Verschiebung der Kontinentalplatten jährlich um mehrere Zentimeter.

Die Umrechnung von Koordinaten eines Punktes in ein anderes Koordinatensystem bezeichnet man als Koordinatentransformation. Unterschiedliche geodätische Koordinatensysteme sind im Allgemeinen unterschiedlich im Erdkörper gelagert. Die Koordinaten für eine Position ändern sich abhängig von dem Datum und dem Sphäroid, auf dem die Koordinaten basieren. So kann eine Umrechnung von Y-Koordinaten in Abhängigkeit von X- Koordinaten (oder andersherum) variieren. Die Breitengrad- und Längengradwerte werden in Dezimalgrad bzw. das Dezimalformat angegeben. Lassen Sie sich Koordinaten z.B. bei OpenStreetMap oder Google Maps anzeigen, werden Ihnen diese im Dezimalformat ausgegeben.

6.7. Einsatz einer Geokoordinaten als Startadresse

Die Geokoordinaten-Datei kann auch zur Ziehung von Startadressen für die Begehung der Sample-Points genutzt werden. Diese Datei beschreibt die gesamte bewohnte Fläche der Bundesrepublik und basiert auf amtlichen Veröffentlichungen. Anders als die OSM-Daten, die auf Grund der freiwilligen Erfassung durch die OpenStreetMap-Community regional einen unterschiedlichen Detaillierungsgrad veröffentlichen (Openstreetmap, kein Datum)³⁶ oder gar der Telefonbucheinträge, die in Säule 3 mit rund 8,5 Mio. vollständigen Adressen nur noch 21% der Haushalte repräsentieren, ist in der Geokoordinaten-Datei die Grundgesamtheit gänzlich abgebildet.

Die - als Startpunkt innerhalb des Sample-Points - gezogene Geokoordinate im Dezimalformat kann in alle gängigen Online-Kartendienste zur Positionsbestimmung und Routenberechnung eingegeben werden. Der Standort kann also als Straßenkarte oder als Luft- bzw. Satellitenbild im Institut oder auf dem Tablet/ Smartphone angezeigt werden.

³⁶ <https://www.openstreetmap.de/faq/> (Wie vollständig sind die Daten?)

Abbildung 34: Flensburg, Stadt SP: 62420350 (1.052 Einwohner laut Zensus 2011)



Da es hier aber keine Adresse mit Namen mehr gibt, müssen neue Regeln für die Ermittlung des ersten begehbaren Haushalts festgelegt werden. So könnte eine Anweisung zum Beispiel lauten: „Gehe zu dem privat bewohnten Gebäude rechts, dessen Hauseingang der gezogenen Koordinate am nächsten liegt. Leben in diesem Haus mehrere Privathaushalte, starte mit der 3. Klingel rechts-oben.“

6.8. Möglichkeiten der Stichprobenziehung anhand der Geokoordinaten-Datei

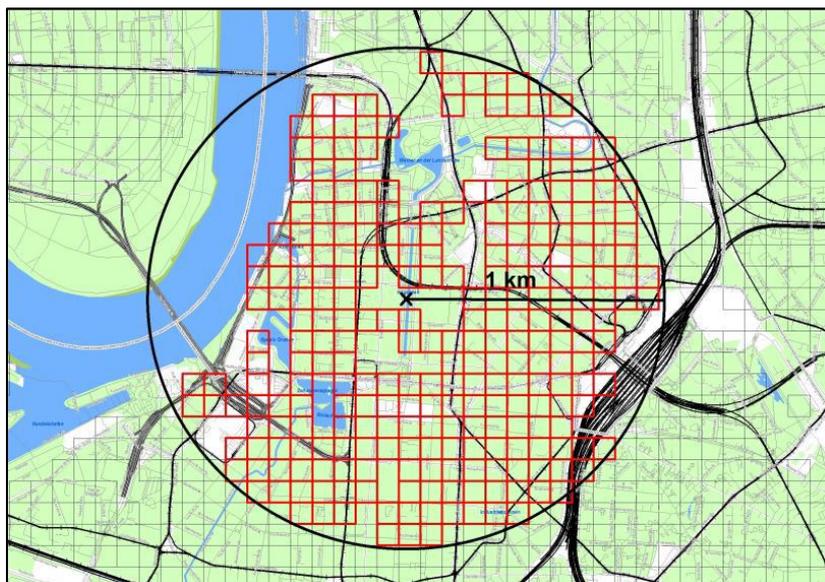
Neben der Nutzung der Geokoordinaten der Gitterzellen als Erweiterung für Startadressen oder Anreicherung um sekundärstatistische Daten (INSPIRE) ergeben sich neue Möglichkeiten der Stichprobenziehung anhand dieser Datei.

Für georeferenzierten Stichproben können anhand der räumlichen Gebietsgliederungseinheiten (AGS, STT, INSPIRE-GeoGitter) zum Beispiel OpenStreetMap-Daten oder andere verfügbare (Geo-)daten / Open-Data an den Datenbestand angespielt werden. So können verschiedenste Stichprobenmodelle realisiert werden; auch eine Integration dieser statistischen und geographischen Informationen in Datenbestände ist möglich.

Denkbar zum Beispiel wären die Ziehung von INSPIRE GeoGittern mit reiner Zufallsauswahl, geschichteter Auswahl, Auswahl von am dichtest bewohnten Gitterzellen innerhalb einer Gemeinde, etc.:

- Ziehung bewohnter 100m-Gitter/ Geokoordinaten zur Ermittlung des Sample Points.
- Ziehung bewohnter 100m-Gitter/ Geokoordinaten mit Einwohnerzahl je ha ≥ 200 (i.d.R. Hochhäuser).
- Ziehung bewohnter 100m-Gitter/ Geokoordinaten mit POI (Supermärkten, Schulen u.ä.)
- Ziehung bewohnter 100m-Gitter/ Geokoordinaten in beliebigem Umkreis um z.B. einen Unternehmenssitz, einen POI, um sonstige Geokoordinaten.
- ...

Abbildung 35: Auswahl aller bewohnten 100m-Gitter im Umkreis von 1 km um die Düsseldorfer Königsallee (Kö)



7. Vergleiche mit den bisherigen F2F-Systemen

Bis 2003 wurde das Flächenstichproben-System auf Grundlage der kleinsten administrativen Einheiten, der Wahlbezirke, jeweils zum Stand einer bestimmten Bundestagswahl gebildet. Man konnte pro Wahlbezirk die Liste des zur Fläche des Wahlbezirks (WBZ) gehörenden Straßen und Straßenabschnitte in der Begehungsdatei abrufen, die den Point in der Fläche definierte und als Bedeutungsgewicht die Anzahl Wahlberechtigter (Deutsche ab 18 Jahren).

Dieses Vorgehen bezog sich auf die verfügbaren Daten, bedeutete aber einen enormen Zeit- und Kostenaufwand, da die Wahlbezirksverzeichnisse pro Gemeinde (!) erhoben und bezahlt werden mussten. Häufig gab es diese nur als Ausdruck, so dass dann noch eine aufwendige manuelle Erfassung erfolgen musste. Je stärker sich andere Formen – kostengünstigere - der Erhebungstechnik durchsetzen, umso wichtiger wurde es, das Verfahren der Definition und Abgrenzung umzustellen. Dazu kam der Aspekt, dass die Zahl der Wahlberechtigten ab 18 Jahre immer weniger die Deutschen insgesamt und die Nicht-deutsche Bevölkerung abbildet.

2003 wurde daher erstmalig nicht mehr mit Wahlbezirken gearbeitet, sondern mit kleinräumigen administrativen Gliederungen, für die Daten der gesamten Bevölkerung vorlagen. Konstante Untergliederungen sollten den Aktualisierungsaufwand im Rahmen halten. Die Verwendung eines für alle F2F-Institute einmalig zur Nutzung lizenzierten Straßenverzeichnis mit Straßenabschnitts-Koordinate (NavTech-Bestand) erlaubte eine andere Form der Sample-Point-Bildung. Zusätzlich wurde die sog. „Startadressen-Datei“ eingeführt, die 10% aller im Telefonbuch eingetragener Adressen pro Sample-Point beinhaltet, so dass eine Begehung über mehrere Startadressen im Point erfolgen konnte. Eine eindeutige Zuordnung der Points zu den amtlichen Gemeinden erlaubte eine Aktualisierung der Sachdaten, die 2010 und 2014 erfolgte.

„Für die Aktualisierung des Systems 2010 wurde geprüft, ob die Daten über Gebäude der Kataster- und Vermessungsämter eine Alternative zu den kommunalen Informationen und den digitalen Geosystemen darstellen, [...] Diese Daten waren jedoch noch nicht flächendeckend für alle Bundesländer verfügbar.“ (Heckel & Hofmann, 2014)³⁷

Mit der erstmaligen Datenermittlung der Einwohner je 100x100m-Gitterzelle im Rahmen des Zensus 2011 und der Veröffentlichung dazugehöriger Geoinformationen konnte jetzt eine andere Form der Flächendefinition gewählt werden.

³⁷ Heckel, C, Hoffmann, O. (2014), „Das ADM-Stichproben-System (F2F) ab 1997“, S. 97

Tabelle 31: Vergleich der F2F-Systeme nach Anzahl Points pro politischer Gemeindegrößenklasse

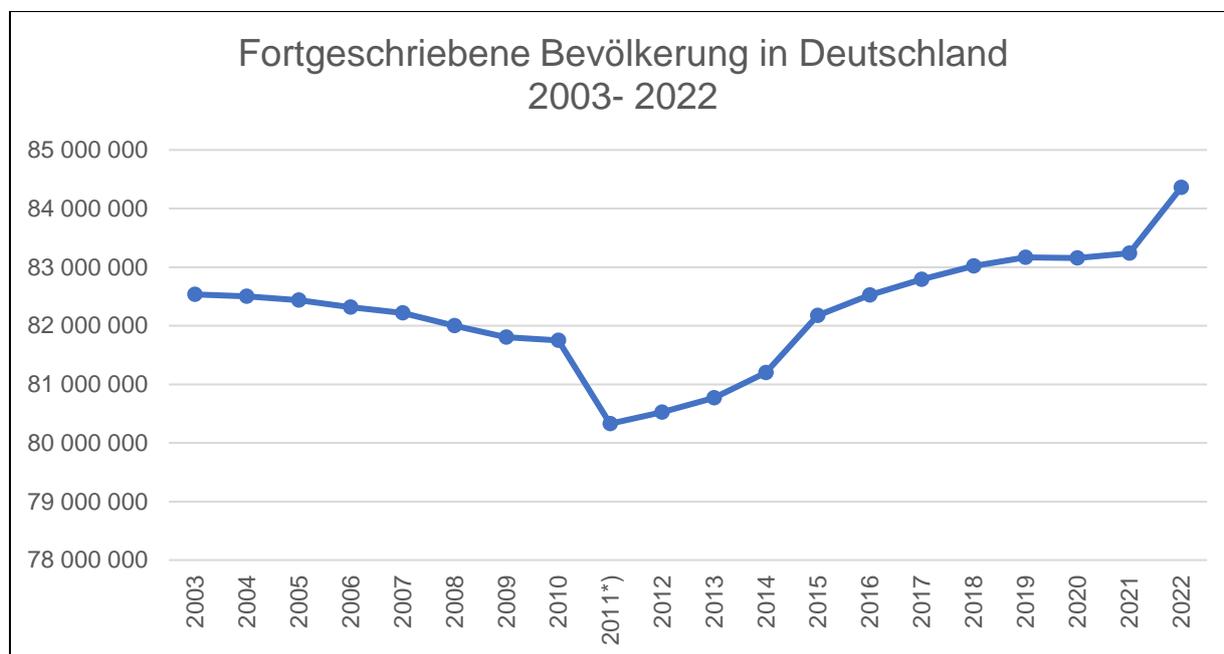
Anzahl Points	Point-Typ	GKPOL 1	GKPOL 2	GKPOL 3	GKPOL 4	GKPOL 5	GKPOL 6	GKPOL 7	Summe pro Typ	Gesamt
2010 GS 2010, SS 2009	Gemeinde /-teil	2.254	4.973	13.720	9.911	4.815	7.806	8.121	51.600	52.947
	Gemeinde-übergreifend	1345	2	0	0	0	0	0	1.347	
2015 GS 2014, SS 2013	Gemeinde /-teil	2.228	4.832	14.155	9.827	4.958	7.838	7.803	51.641	52.929
	Gemeinde-übergreifend	1.285	3	0	0	0	0	0	1.288	
2023 GS 2022, SS 2021	Gemeinde /-teil	24	5.255	16.131	10.852	5.428	8.744	10.039	56.473	58.564
	Gemeinde-übergreifend	2.074	17	0	0	0	0	0	2.091	

GKPOL: 7 Gemeindegrößenklassen (politisch)

Als „Gemeindeübergreifend“ sind hier die sog. „synthetischen Points“ bezeichnet, die aus Zusammenfassungen von Gemeinden unter 2.000 Einwohnern innerhalb des gleichen Kreises und der gleichen BIK-Gemeindegrößenklasse gebildet wurden.

Der Zensus 2011 wies – als Folge von Korrekturen der seit der Volkszählung 1987 fortgeschriebenen Bevölkerungszahlen - rund 1,5 Millionen Einwohner weniger aus, und damit verbunden auch geringere Haushaltszahlen. Dies ist gut am Sprung der Anzahl Haushalte gesamt von System 2010 zu 2015 zu erkennen, es sank auch die Zahl der Sample-Points.

Abbildung 36: Bevölkerungsentwicklung seit 2003



Quelle: DESTATIS, Statistischer Bericht, Bevölkerungsfortschreibung 2022, EVAS Nr 12411, Die den Bevölkerungsdaten zugrunde liegenden Meldungen der Meldebehörden enthalten zahlreiche Melderegisterbereinigungen, die infolge der Einführung der persönlichen Steuer-Identifikationsnummer durchgeführt worden sind. Die Ergebnisse sind mit dem jeweiligen Vor- bzw. Folgejahr eingeschränkt vergleichbar. *) ab 2011: Ergebnisse der Bevölkerungsfortschreibung auf Grundlage des Zensus 2011.

Die Ergebnisse ab Berichtsjahr 2016 sind aufgrund methodischer Änderungen und technischer Weiterentwicklungen nur bedingt mit den Vorjahreswerten vergleichbar. Des Weiteren existieren Einschränkungen in der Genauigkeit der Ergebnisse, die u.a. aus Problemen bei der melderechtlichen Erfassung von Schutzsuchenden resultieren.

Der Anstieg der Points passt zum Zuwachs an Bevölkerung und Haushalten, wie in der obenstehenden Grafik dargestellt. Die optimale Point-Größe sollte 700 Haushalte sein. Bei den Gemeinden in der politischen gemeindegrößenklasse 1 (unter 2.000 Einwohner) war der Wunsch der Institute, die Sample-Points eher größer zu schneiden, da die Begehung dieser vor allem synthetischen Points längere Anfahrtswege und das „Tür-Klingeln“ bei misstrauischen Anwohnern häufig den Ausfall eines ganzen Dorfes bedeuten kann. Dafür sollten in Großstädten (politische Gemeindegrößenklasse 6+7, mehr als 100.000 Einwohner) rund 750 Haushalte einen Point bilden, hier sind viele 1-Personen-Haushalte zu finden.

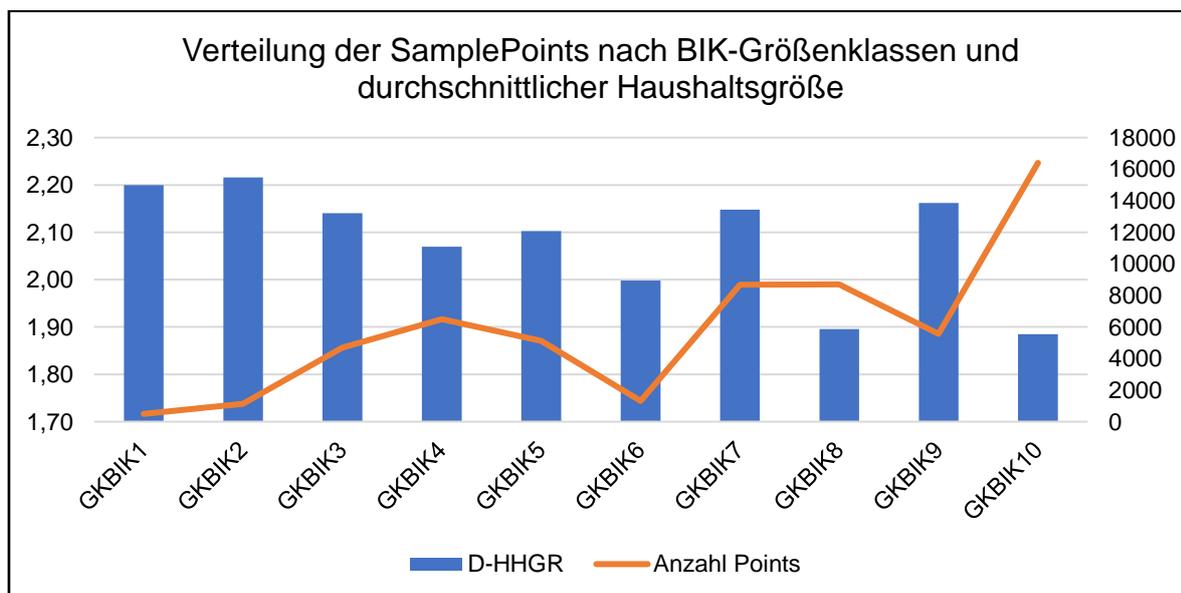
Tabelle 32: Verteilung der Haushalte und Point-Größen nach Anzahl Haushalte pro Point

Anzahl Points	Anzahl Haushalte	GKPOL 1	GKPOL 2	GKPOL 3	GKPOL 4	GKPOL 5	GKPOL 6	GKPOL 7	Gesamt
2010 GS 2010, SS 2009	Anzahl HH	2.221.560	3.385.870	9.805.541	7.123.145	3.481.801	6.346.955	7.154.828	39.519.700
	Ø- Anz. HH	617	681	715	719	723	813	881	746
	Minimum	22	48	66	62	229	44	43	22
	Maximum	1.582	1.361	1.435	1.475	1.559	1.863	1.869	1.869
2015 GS 2014, SS 2013	Anzahl HH	2.122.583	3.259.847	9.27.496	7.091.755	3.605.178	6.434.300	6.930.200	39.371.359
	Ø- Anz. HH	604	674	701	722	727	821	888	744
	Minimum	22	45	65	58	232	43	42	22
	Maximum	1602	1375	1816	1756	1549	1872	1848	1871
2023 GS 2022, SS 2021	Anzahl HH	2.060.624	3.231.630	10.352.804	7.274.701	3.732.341	6.507.900	7.441.000	40.601.000
	Ø- Anz. HH	982	613	642	670	688	744	741	693
	Minimum	55	351	352	359	372	356	367	55
	Maximum	2478	2437	1162	1176	1147	1195	1266	2478

GKPOL: 7 Gemeindegrößenklassen (politisch)

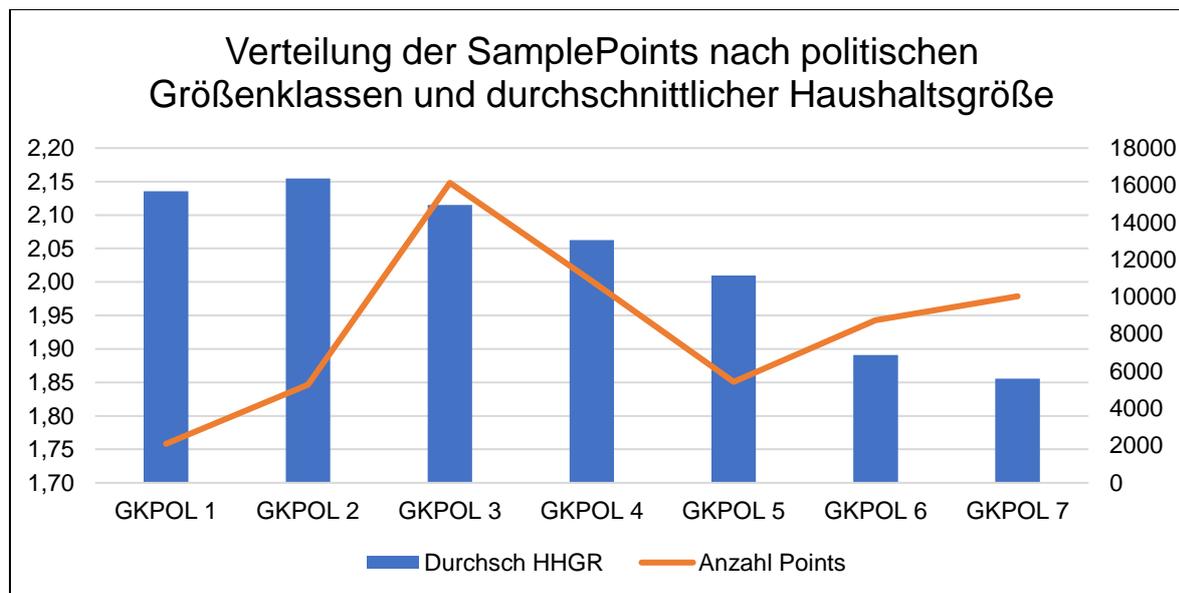
Beim Einsatz des F2F-Stichprobensystems ist die Verwendung der BIK-Gemeindegrößenklassen (BIK ASCHPURWIS + BEHRENS GMBH, kein Datum)³⁸ zu bevorzugen, die diese Systematik die Sample-Points in Bezug auf ihre durchschnittliche Haushaltsgröße besser trennt. Da bei einer Verwendung der politischen Gemeindegrößenklassen (Anzahl Einwohner in der Gemeinde) außer Acht gelassen wird, in welchem Verflechtungs-zusammenhang eine Gemeinde liegt, erhält man in den unteren, kleinen Gemeindegrößen (GKPOL 1-3) eine große undifferenzierte Zahl. Vergleicht man dies mit der Verteilung nach den BIK-Gemeindegrößenklassen, rücken kleine Gemeinden in das Randgebiet von Kernen oder Umland, und weisen entsprechend kleinere oder größer Haushalte aus.

Abbildung 37: Durchschnittliche Haushaltsgröße (MZ21) nach BIK-Größenklassen



³⁸ <https://bik-gmbh.de/regionaldaten/bik-regionen/>

Abbildung 38: Durchschnittliche Haushaltsgröße (MZ21) nach politischen Größenklassen



8. Schlussbemerkung

Anders als in den Vorgängersystemen ist ein Point im System 2023 jetzt definiert durch eine Fläche, bestehende aus 100x100m-Gitterzellen und dem Bedeutungsgewicht der Fläche. Die genaue Abbildung der Fläche, hinterlegt mit Straßen und Hausnummern erfolgt über die Gitterkoordinaten, die mit OSM-Kartendaten verknüpft und damit immer aktuell abgebildet werden. INSPIRE-Koordinaten bleiben konstant, die Bebauung in der Gitterzelle verändert sich, wird aber durch die aktuellen Kartendaten angezeigt. **Daher ist die Begehungsdatei nicht mehr ! das zentrale Element des Systems.** Auch können Koordinaten der Gitterzellen als Startadressen genutzt werden. Durch Geokodierung gelisteter Adressen und des realisierten Interviews ergeben sich im Nachgang der Begehung eines Points deutlich bessere Kontrollmöglichkeiten, z.B. ob ein Interviewer aus dem Point gelaufen ist.

Die Beschreibung des Sample-Points über die INSPIRE -GeoGitter (mit Geokoordinaten) erlaubt die Kartenansicht des Points, die der Interviewer über einen Link auch auf seinem Smartphone oder Tablet aufrufen kann. Dies erleichtert die Begehung.

Die Erweiterung der Auswahlgrundlage um die Koordinaten der 100x100m Gitterzellen und somit der Points erlaubt es, über die INSPIRE-IDs oder die Koordinaten (-bereiche) andere Informationen zu den untersuchten Flächen hinzuzuspielen. Sekundärstatistische Datenanalysen, über die Befragung hinausgehend werden einfacher. Beispiele wären das soziale Umfeld im Point, aber auch zurückgelegte Wegstrecken oder die Entfernung zu Geschäften, Arztpraxen, Behörden, Schulen und Kindergärten. (Beuthner, Breuer, & Jünger, 2021)³⁹

Damit bietet das neue ADM-Flächenstichprobensystems Face-to-Face 2023 eine andere, erweiterte Praxisanwendung als die bisherigen Systeme. Neben dem Verzicht auf lange Straßenlisten wird auch eine Nachkontrolle über die geokodierten Interviews mit Uhrzeit über den Laufweg ermöglicht.

³⁹ Beuthner, Christoph, Johannes Breuer, and Jünger, Stefan (2021). Data Linking - Linking survey data with geospatial, social media, and sensor data. Mannheim, GESIS - Leibniz Institute for the Social Sciences (GESIS Survey Guidelines). DOI: 10.15465/gesis-sg_en_039

9. Literaturverzeichnis

- ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (2023). *Jahresbericht 2022*.
- Anding, K., & Fürnrohr, M. (2015). Zensus 2011, Ermittlung von Einwohnerzahlen von Bund, länder und Kommunen. *Bayern in Zahlen*, 4.
- Beuthner, C., Breuer, J., & Jünger, S. (2021). *Data Linking - Linking survey data with geospatial social media and sensor data*. Mannheim: GESIS Survey Guidelines.
- BIK ASCHPURWIS + BEHRENS GMBH. (kein Datum). *BIK ASCHPURWIS + BEHRENS GMBH*. Von <https://bik-gmbh.de/regionaldaten/bik-regionen/> abgerufen
- ChatGPT; Open.ai. (2024).
- DESTATIS. (2016). *Demographische Standards. eine gemeinsame Empfehlung des ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes* (Bd. 17). Wiesbaden.
- DESTATIS. (2023). *Qualitätsbericht, Fortschreibung des Bevölkerungsstandes 2022*. Wiesbaden.
- DESTATIS. (kein Datum). www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/amtlicher-gemeindeschluessel.html. Abgerufen am 11. 04 2024 von <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/amtlicher-gemeindeschluessel.html>
- DESTATIS. (kein Datum). www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/gemeinde-freie-gebiete.html. Abgerufen am 11. 04 2024 von <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Glossar/gemeinde-freie-gebiete.html>
- DESTATIS. (kein Datum). www.destatis.de/Europa/DE/Methoden-Metadaten/Klassifikationen/UebersichtKlassifikationen_NUTS.html. Abgerufen am 11. 04 2024 von https://www.destatis.de/Europa/DE/Methoden-Metadaten/Klassifikationen/UebersichtKlassifikationen_NUTS.html
- Geodäsie, B. f. (kein Datum). https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/geogitter.pdf. Abgerufen am 11. 04 2024 von https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/geogitter.pdf
- GESIS. (2019). *Regionale Standards 2019, 3.Auflage, GESIS Schriftenreihe*. Wiesbaden: GESIS.
- Häder, M., & Häder, S. (2022). Stichprobenziehung in der quantitativen Sozialforschung, Kap. 27. In N. Baur, & J. Blasius, *Handbuch Methoden der empirischen Sozialforschung, 3. Auflage*. Wiesbaden: Springer VS.
- Heckel, C., & Hofmann, O. (2014). Das ADM-Stichproben-System (F2F) ab 1997. In A. A.-u. e.V., *Stichproben-Verfahren in der Umfrageforschung, Eine Darstellung für die Praxis, 2. Auflage*. Wiesbaden: Springer VS.
- Helbig, M. (2023). *Kleinräumige Ungleichheiten in den deutschen Städten, Datenfile Version 1.0.0*. Berlin: WZB - Wissenschaftszentrum Berlin für Sozialforschung.
- Löffler, U., Behrens, K., & von der Heyde, C. (2014). Die Historie der ADM-Stichproben. In A. A.-u. e.V., *Stichproben-Verfahren in der Umfrageforschung, Eine Darstellung für die Praxis, 2. Auflage* (S. 67 f). Frankfurt am Main: Springer VS.
- LSN Niedersachsen. (kein Datum). www.statistik.niedersachsen.de. Abgerufen am 11. 04 2024 von https://www.statistik.niedersachsen.de/startseite/themen/gebiet_flachennutzung/gebiet_niedersachsens/gebiet-niedersachsens-191889.html
- OpenStreetMap. (kein Datum). <https://wiki.openstreetmap.org/wiki/>. Abgerufen am 11. 04 2024 von <https://wiki.openstreetmap.org/wiki/DE:Attribut>
- OpenStreetMap. (kein Datum). <https://wiki.openstreetmap.org/wiki/DE:Elemente>. Abgerufen am 11. 4 2024 von <https://wiki.openstreetmap.org/wiki/DE:Elemente>
- Openstreetmap. (kein Datum). <https://www.openstreetmap.de/faq/>. Abgerufen am 11. 4 2024 von <https://www.openstreetmap.de/faq/>

Plennert, M. (2016). Anwendungsreif? Nutzung und Potenzial von digitalen Geodaten für Stadtforschung und Raumbewertung am Fallbeispiel OpenStreetMap. *STADTFORSCHUNG UND STATISTIK*, 1.

Ramm, F., & Topf, J. (2008). *OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten*. Berlin.

Schröder, J. (2015). Persönlich-mündliche Befragung. *GESIS Survey Guidelines*(5).

Wiedenbeck, M., & Züll, C. (2001). Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren. *ZUMA How-to-Reihe*(Nr. 10). Mannheim: ZUMA Zentrum für Umfragen, Methoden und Analysen, How-to-Reihe Nr. 10.

www.gesetze-im-internet.de. (19. 12 2022). Abgerufen am 11. 4 2024 von www.gesetze-im-internet.de/bmg/BJNR108410013.html

www.zensus2011.de/DE/Infothek/Begleitmaterial_Ergebnisse/Begleitmaterial_node.html. (kein Datum). Abgerufen am 11. 4 2024 von https://www.zensus2011.de/DE/Infothek/Begleitmaterial_Ergebnisse/Begleitmaterial_node.html