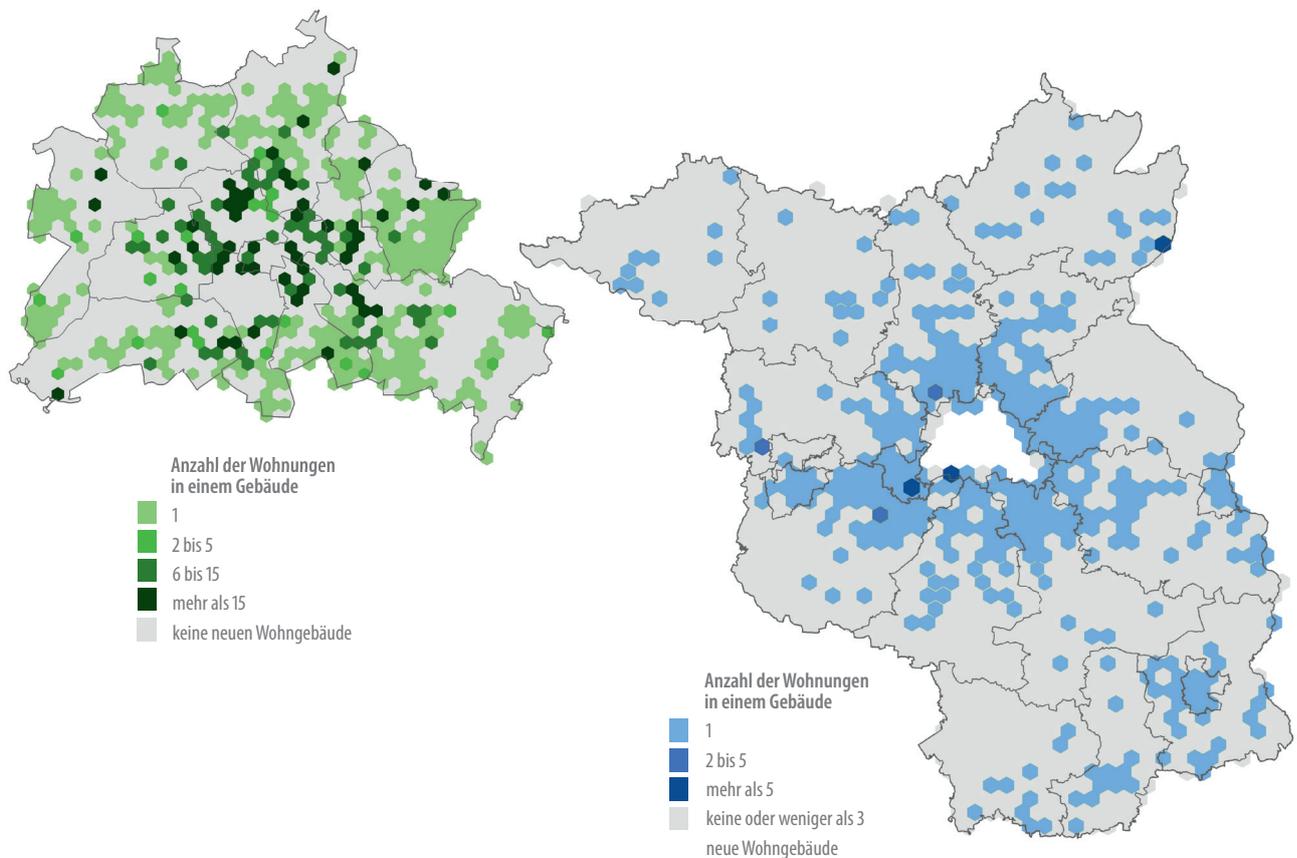


▮ **GEOREFERENZIERUNG**

Überwiegende Anzahl der Wohnungen pro fertiggestelltem Wohngebäude 2018 ...
... in Berlin im 1-km-Raster

... und in Brandenburg im 5-km-Raster



Weitere Themen: ▮ **Amtliche Statistik und COVID-19, Zensus**

Amt für Statistik Berlin-Brandenburg

**Zeitschrift für amtliche Statistik
Berlin Brandenburg**
14. Jahrgang

Herausgeber
Amt für Statistik Berlin-Brandenburg
Steinstraße 104–106
14480 Potsdam
Tel.: 0331 8173-1777

Redaktion
Nicole Dombrowski (Leitung),
Dr. Holger Leerhoff (Verantwortlicher
Redakteur i. S. d. BbgPG),
Anja Malchin,
Dr. Thomas Troegel,
Ramona Voshage
zeitschrift@statistik-bbb.de

Preis
Einzelheft EUR 6,00
ISSN 1864-5356

Satz und Gestaltung
Amt für Statistik Berlin-Brandenburg

Druck
TASTOMAT GmbH, Strausberg

© **Amt für Statistik Berlin-Brandenburg, 2020**
Auszugsweise Vervielfältigung und
Verbreitung mit Quellenangabe gestattet.

Das Amt für Statistik Berlin-Brandenburg
hat seinen Sitz in Potsdam und weitere
Standorte in Berlin und Cottbus.

Auskunft und Beratung
Steinstraße 104–106
14480 Potsdam

Telefon: 0331 8173-1777
Fax: 030 9028-4091
info@statistik-bbb.de

Zeichenerklärung

- 0 weniger als die Hälfte von 1
in der letzten besetzten Stelle,
jedoch mehr als nichts
 - nichts vorhanden
 - ... Angabe fällt später an
 - () Aussagewert ist eingeschränkt
 - / Zahlenwert nicht sicher genug
 - Zahlenwert unbekannt oder
geheim zu halten
 - x Tabellenfach gesperrt, weil
Aussage nicht sinnvoll
 - p vorläufige Zahl
 - r berichtigte Zahl
 - s geschätzte Zahl
- Abweichungen in der Summe
können sich durch Schätzungen
ergeben



Alle Ausgaben seit 2007
finden Sie auf
www.statistik-berlin-brandenburg.de

Kurzbericht

- ▮ Das AFS bekommt Zensus-Standort **3**

Entwicklungen in der amtlichen Statistik

- ▮ Ein Implementierungsprozess für Geheimhaltungsverfahren in vier Stufen **4**

Neuerscheinungen

- ▮ Indikatorenbericht der UGRdL – aktualisierte Länderergebnisse zum Umweltstatus **9**
- ▮ Dossier zur Corona-Pandemie in Berlin und Brandenburg **33**
- ▮ „Bildung in Deutschland 2020“ mit Schwerpunktthema zur Digitalisierung **49**

Statistik erklärt

- ▮ Nowcast und Übersterblichkeit **15**

Historisches

- ▮ Die Spanische Grippe im Deutschen Reich **50**

Save the date

- ▮ 40. Kongress der Deutschen Gesellschaft für Soziologie **54**
- ▮ VFS-Jahrestagung 2020 **54**
- ▮ Smart Country Convention **54**

Fachbeiträge

Gesundheitswesen

- ▮ **Sterben die Menschen in Berlin und Brandenburg mit oder an Corona?** **10**

Katrin Möbius

Fachgespräch mit Hartmut Bömermann

- ▮ **„Wie systemrelevant ist die amtliche Statistik?“** **13**

Georeferenzierung

- ▮ **Regionalanalyse auf Basis simulierter Geokoordinaten**
Gütebeurteilung des Verfahrens am Beispiel der Wahlberechtigten in Berlin **16**

Kerstin Erfurth

- ▮ **Geokodierung mit dem RBS-Geocoder** **24**

Renee Lin

- ▮ **Raumbezogene Analyse zur Wohnraumentwicklung in Berlin und Brandenburg** **28**

Nicole Jurisch

Verarbeitendes Gewerbe

- ▮ **Reurbanisierung der Industrie**
Neue Chance für die Stadt? **34**

Ramona Voshage, Martin Gornig

Zensus

- ▮ **Pretest zur Personenerhebung des Zensus 2021** **40**

Mark Hoferichter, Ralf Mädler, Natalie Lück, Laura Tomm-Demski, Ulrike Warnke, Jonas-Otto Werner



Liebe Leserinnen und Leser,

seit nunmehr vier Monaten haben wir mit dem Corona-Virus SARS-CoV-2 einen ständigen Begleiter. Zwischen der Hoffnung, dass die Beschränkungen und Hygieneregeln greifen, und dem Bangen, was neue Lockerungen und Rückreisen aus dem Sommerurlaub bringen, gestalten wir unseren Alltag. Auch das Amt für Statistik Berlin-Brandenburg (AfS) begegnet bei der täglichen Arbeit den Auswirkungen und Folgen der Corona-Pandemie. Diese sind sowohl im Miteinander der Beschäftigten an unseren Standorten spürbar als auch in den Zahlen ablesbar.

Auch deshalb beschäftigt sich ein Teil der Ausgabe 2/2020 der *Zeitschrift für amtliche Statistik Berlin Brandenburg* mit dem Zusammenhang von COVID-19 und der amtlichen Statistik. Ein Beitrag beschreibt, inwieweit sich eine Infektion mit COVID-19 auf die Todesursachenstatistik auswirkt und erklärt, woran zu erkennen ist, ob Menschen mit oder an dem Virus verstorben sind.

Im Fachgespräch stellen wir die Frage nach der Systemrelevanz der amtlichen Statistik in Krisenzeiten, versuchen unter anderem darüber aufzuklären, warum Aktualität in der amtlichen Statistik manchmal schwer umzusetzen ist und welche neuen Wege wir beschreiten, um dieser Herausforderung zu begegnen.

Einen zweiten Schwerpunkt dieser Ausgabe bildet die Georeferenzierung. Insgesamt drei Beiträge beleuchten das Thema und zeigen Anwendungsfelder auf – etwa mit der Statistik der Baufertigstellungen oder den Daten zu Wahlberechtigten in Berlin.

Die vierte industrielle Revolution steht im Mittelpunkt eines Beitrages zu strukturellen Veränderungen industrieller Raummuster aufgrund von Digitalisierung und Internationalisierung. Anhand eines Analysedatensatzes zum Verarbeitenden Gewerbe in Deutschland untersuchen die Autoren das industrielle Gründungsgeschehen in Stadt und Land.

In Vorbereitung des nächsten Zensus in Deutschland wurde zwischen November 2019 und März 2020 in Berlin und Brandenburg die Personenerhebung als Probelauf in kleinem Rahmen, ein sogenannter Pretest, durchgeführt. Über dessen Ablauf sowie die ersten Ergebnisse informiert ein weiterer Beitrag.

Zu guter Letzt folgt noch ein kleiner Hinweis auf die Statistische Bibliothek: Nach einem umfangreichen Relaunch der Website verfügt sie nun über ein modernes Layout. Eine erweiterte Suche macht es möglich, noch mehr Parameter einzubeziehen. Die gefundene Publikation kann dann mithilfe des neuen „Viewer“ direkt auf der Seite gelesen oder wie gewohnt als PDF-Datei heruntergeladen werden. Seit dem Relaunch erreichen Sie die Statistische Bibliothek unter der Adresse <https://www.statistischebibliothek.de>. Sie verfügt mittlerweile über 138 000 Dokumente, unter anderem alle Ausgaben der *Zeitschrift für amtliche Statistik Berlin Brandenburg* seit 2007 sowie weitere Publikationen des AfS. Reinschauen lohnt sich.

**Ich wünsche Ihnen eine informative Lektüre.
Bleiben Sie gesund!**

Jörg Fidorra
Vorstand des Amtes für Statistik Berlin-Brandenburg

Kurzbericht

Das AfS bekommt Zensus-Standort

von Dirk Brenneisen und Kersten Klemm



Die Senatsverwaltung für Inneres und Sport Berlin bat das Amt für Statistik Berlin-Brandenburg (AfS) seine Planungen so auszurichten, dass im Projekt Zensus 2021 auch die Erhebungsstelle Berlin örtlich, personell und organi-

satorisch im AfS integriert ist. Aufgrund des damit verbundenen personellen Umfangs erfolgte in 2019 unter wirtschaftlichen Gesichtspunkten die Suche nach einem geeigneten Dienstgebäude in Berlin. Ein Mietvertrag konnte Ende Juli 2019 für ein Bürogebäude in der Alten Rhinstraße 8 in Berlin-Marzahn unterzeichnet werden; die Objektübernahme erfolgte am 16. September 2019.

Eine Vielzahl von vorbereitenden Tätigkeiten, sowohl technisch als auch baulich, waren erforderlich, bevor der Umzug Anfang März 2020 erfolgen konnte: Neben der Berücksichtigung aller Vorgaben bezüglich Arbeits- und Brandschutz musste beispielsweise das vorhandene Schließsystem aus Sicherheitsgründen komplett ausgetauscht werden.

Der Zensus 2021 arbeitet in einer abgeschotteten Umgebung, das heißt unabhängig von den IT-Strukturen des AfS. Dafür wurde ein entsprechendes

Sicherheitskonzept mit einem Maßnahmenkatalog erarbeitet und umgesetzt. Der Brandenburgische IT-Dienstleister ZIT-BB baute eine komplett neue IT-Infrastruktur für den neuen Standort auf. Es wurden die technischen Voraussetzungen geschaffen, um das Gebäude an die Landesnetze Brandenburgs und Berlins anzubinden und den Mitarbeiterinnen und Mitarbeitern die Kommunikation mit den Kolleginnen und Kollegen an den anderen Standorten des AfS in gewohnter Weise zu ermöglichen. Dazu zählten neue Server, eine moderne Telefonanlage und die Ausstattung der PC-Arbeitsplätze.

Die Anbindung an die Landesnetze war zum 5. März 2020 abgeschlossen. Hiermit wurde sichergestellt, dass die Beschäftigten in den einzelnen Teilprojekten mit ihren Fachanwendungen, die für die tägliche Arbeit benötigt werden, arbeiten können.

In einer überaus kurzen Zeit gelang es dem ZIT-BB gemeinsam mit der Unterstützung des Ministeriums für Inneres und Kommunales Brandenburg unter Zurückstellung vieler anderer Aufgabenstellungen, das neue Dienstgebäude bezugsfähig zu machen.

Am neuen Zensus-Standort haben nun alle jetzigen und zukünftigen Projektbeschäftigten ausreichend Platz. Durch die gute Ausstattung der Arbeitsplätze sind die Kolleginnen und Kollegen auch während der derzeit herrschenden Coronapandemie weitestgehend arbeitsfähig.

Dirk Brenneisen ist IT-Referent im Projekt *Zensus 2021* des Amtes für Statistik Berlin-Brandenburg.

Kersten Klemm leitet das Projekt *Zensus 2021* im Amt für Statistik Berlin-Brandenburg.



Fotos: Sara Imam

Entwicklungen in der amtlichen Statistik

Ein Implementierungsprozess für Geheimhaltungsverfahren in vier Stufen

von Sarah Gießing, Julia Höninger, Arijana Amina Ramic und Johannes Rohde

Einleitung

Der Schutz von Einzelangaben bei der Verarbeitung und Veröffentlichung statistischer Ergebnisse hat für die amtliche Statistik einen hohen Stellenwert. Die Gewährleistung der Geheimhaltung hat daher Eingang in den Verhaltenskodex für europäische Statistiken [1] gefunden und ist national im Bundesstatistikgesetz (§ 16 Abs. 1 BStatG¹) geregelt. Entscheidungen hinsichtlich der konkreten Ausgestaltung der Geheimhaltung obliegen den jeweiligen Fachbereichen. Im verbundinternen „Handbuch zur Statistischen Geheimhaltung“ [2] sind die Methoden der statistischen Geheimhaltung ausführlich dargelegt.

Wird eine Statistik beziehungsweise ihre statistischen Produkte erstmals konzipiert oder wird das Auswertungskonzept überarbeitet, müssen zu den Erfordernissen der Fachstatistik passende Methoden zur Geheimhaltung und Regulierung des Zugangs zu vertraulichen Ergebnissen und Produkten konzipiert und eingeführt werden. Die Sicherstellung der Geheimhaltung sollte dabei ein einheitlicher Prozess im Statistischen Verbund sein und zwischen allen Anwendungsbereichen abgestimmt werden.

In diesem Beitrag wird ein Prozess dargestellt, der schrittweise zur Entwicklung eines einheitlichen Geheimhaltungsverfahrens mit möglichst automatisierter Umsetzung führt. Dieser Prozess wurde durch die Expertengruppe für Statistische Geheimhaltung im Statistischen Verbund innerhalb des Vorgehenskonzepts „Einheitliche Geheimhaltung in allen Statistiken mit Geheimhaltungsbedarf“ vorgeschlagen. Die zuständigen übergeordneten Gremien der deutschen amtlichen Statistik haben diesem Vorschlag im Oktober 2019 zugestimmt, sodass das Vorgehenskonzept seit Anfang 2020 in allen Fachbereichen im Statistischen Verbund Anwendung findet.

Dieser Prozess kann als ein Standardprozess im Statistischen Verbund aufgefasst werden und gliedert sich in vier aufeinander aufbauende Stufen, die in Abbildung a veranschaulicht sind:

Während die ersten drei Stufen (Bedarfsprüfung, Methodenwahl, Konzeptentwicklung) die eigentliche Planungsphase umfassen, verfolgt die vierte Stufe (Digitalisierung) das Ziel der praktischen Umsetzung und gegebenenfalls der Automatisierung des entwickelten Geheimhaltungsprozesses. Mit Digitalisierung ist eine möglichst weitgehend automatisierte Umsetzung gemeint. Theoretisch durchläuft jede Statistik alle vier Stufen und tritt jeweils nach Abschluss einer Stufe in die darauf folgende ein.

Um eine Entwicklungsstufe abzuschließen und damit in die folgende Entwicklungsstufe aufzurücken, müssen bestimmte Stufenziele erreicht sein. Im Folgenden werden die einzelnen Stufen und deren Ziele detailliert dargelegt.

Falls in einer amtlichen Statistik ein Geheimhaltungskonzept aktuell konzipiert oder überarbeitet wird, können Stufenziele (zum Beispiel Stufe I oder Stufe II) bereits erreicht sein. Geheimhaltungskonzepte sind regelmäßig oder im Bedarfsfall (zum Beispiel bei der Neukonzeption einer Statistik oder Überarbeitung eines Auswertungskonzepts) zu überprüfen.

Stufe I: Prüfung des Geheimhaltungsbedarfs

Gemäß § 16 BStatG besteht für Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik gemacht werden, die Pflicht zur Geheimhaltung durch die amtlichen Statistikstellen, soweit durch eine besondere Rechtsvorschrift nichts anderes bestimmt ist. Daraus ergibt sich, dass grundsätzlich alle Veröffentlichungen, die Ergebnisse einer Bundesstatistik enthalten, auf mögliche Aufdeckungsrisiken zu prüfen sind. In Ausnahmefällen enthalten die fachstatistischen Gesetze Ausnahmeregelungen oder die nachgewiesenen Sachverhalte bedürfen keiner Geheimhaltung (zum Beispiel komplexe Verrechnungen, aus denen kein Rückschluss auf Einzelangaben gezogen werden kann).

a | Darstellung des Geheimhaltungsprozesses in vier Entwicklungsstufen



¹ Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I S. 2394), das zuletzt durch Artikel 10 Absatz 5 des Gesetzes vom 30. Oktober 2017 (BGBl. I S. 3618) geändert worden ist.

Im Vorfeld der Verbreitung der Ergebnisse einer Statistik ist zu klären, ob für diese Statistik eine gesetzlich geregelte Pflicht zur Geheimhaltung beachtet werden muss. Das für die Statistik zuständige Gremium im Statistischen Verbund muss entweder kurz begründen, warum eine Geheimhaltung nicht erforderlich ist, oder ein Vorgehen für die Geheimhaltung festlegen, mit dem der gesetzlichen Pflicht zur Gewährleistung der Geheimhaltung für die betreffende Statistik nachgekommen werden kann. Dieses Vorgehen sollte gelegentlich oder anlassbezogen überprüft werden.

Insgesamt ist zu Beginn des Prozesses zur Implementierung eines Geheimhaltungsverfahrens zunächst dessen Relevanz für die Fachstatistik anhand folgender Punkte zu prüfen:

1. Es ist zu klären, ob bestimmte Merkmale oder Veröffentlichungen von der Pflicht zur Geheimhaltung ausgenommen sind. Dazu sind auch die rechtlichen Grundlagen, die die Durchführung der betreffenden Statistik anordnen, zu beachten und ggf. Expertinnen und Experten aus den Bereichen „Recht“ und „Geheimhaltungsmethodik“ zur Beratung hinzuzuziehen.
2. Falls eine Änderung der einzelgesetzlichen Rechtsgrundlage angestrebt wird (beispielsweise hinsichtlich weiterer von der Geheimhaltungspflicht ausgenommener Merkmale), müssen konkrete Schritte bis hin zur Novellierung des betreffenden Gesetzes definiert werden (zum Beispiel: Zustimmung betroffener Erhebungseinheiten beziehungsweise Stakeholder einholen, Ausarbeitung eines Zeitplans zur Gesetzesänderung). Gegebenenfalls muss eine Übergangslösung zur Sicherstellung der Geheimhaltung bis zum Inkrafttreten der Gesetzesnovellierung entwickelt werden.
3. Auch wenn eine Veröffentlichung von Ergebnissen einer Statistik nicht von allen Statistischen Ämtern erfolgt (zum Beispiel aufgrund zu geringer Fallzahlen in einem Land), sollte ein im Statistischen Verbund abgestimmtes Geheimhaltungskonzept vorliegen, da eine zukünftige Veröffentlichung in den Ländern, die bisher keine Auswertungen veröffentlichten, nicht ausgeschlossen werden kann.

Das Ziel der Entwicklungsstufe I ist erreicht, wenn durch Beschluss des für die Statistik zuständigen Gremiums geklärt ist, ob und inwiefern für eine Statistik Geheimhaltungsbedarf besteht. Ergibt die Prüfung, dass Geheimhaltungsbedarf besteht, erfolgt eine Neueinordnung der Statistik in Entwicklungsstufe II.

Auch im Falle einer im einzelstatistischen Gesetz geregelten Befreiung von der Pflicht zur Geheimhaltung sollte ein (formaler) Beschluss gefasst werden. Falls der Beschluss des zuständigen Gremiums klarstellt, dass keine Geheimhaltung erforderlich ist, muss dies möglichst detailliert begründet (zum Beispiel: Hinweis auf die betreffende gesetzliche Bestimmung oder fachlich begründete starke Aggregation in den Veröffentlichungen) und in Form eines kurzen Geheimhaltungsleitfadens dokumentiert werden. Dieses Vorgehen soll sicherstellen,

dass auch künftige Mitarbeiterinnen und Mitarbeiter der betreffenden Statistik nachvollziehen können, weshalb keine Geheimhaltung notwendig ist.

Anschließend sind keine weiteren Maßnahmen erforderlich. Wenn aufgrund von Vergrößerungen oder eines fest definierten sehr kleinen Tabellenprogramms kein Geheimhaltungsbedarf besteht, dürfen keine davon abweichenden tieferen Auswertungen veröffentlicht werden. Sind weitere Auswertungen von Seiten des Fachbereichs, durch Sonderauswertungen oder durch ein Angebot der Mikrodaten in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder angedacht, müssen die entsprechenden Anforderungen und Bedarfe im Geheimhaltungsprozess berücksichtigt werden.

Das für die Statistik zuständige Gremium kann den Geheimhaltungsbedarf jederzeit erneut prüfen.

Stufe II: Bestimmung der Geheimhaltungsmethode

In dieser Stufe sind verschiedene Geheimhaltungsverfahren auf ihre Eignung zum Einsatz bei der betrachteten Statistik unter Berücksichtigung der Kriterien für die Auswahl eines Geheimhaltungsverfahrens [3] zu prüfen. Zur Sicherstellung der statistischen Geheimhaltung in Tabellen kann ein informationsreduzierendes Verfahren (z. B. Zellsperre oder Vergrößerung) oder ein datenveränderndes Geheimhaltungsverfahren (z. B. die stochastische Überlagerung, Record Swapping, Rundungen oder SAFE) eingesetzt werden. Die vorgeschlagene Methode wird durch das entsprechend zuständige Gremium beschlossen. Im Rahmen von Stufe II sind folgende Aspekte zu beachten:

1. Das zuständige Gremium nimmt sich entweder selbst der Verfahrensauswahl an oder setzt eine Projektgruppe „Geheimhaltung“ ein. Die in Frage kommenden Geheimhaltungsverfahren werden gegebenenfalls unter der Hinzuziehung von Geheimhaltungsmethodikern geprüft.
2. Den Rahmen der Prüfung bilden die „Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens“ [3], welche verschiedene Geheimhaltungsverfahren anhand von Qualitätskriterien (hinsichtlich Schutz vor Aufdeckung, Qualität, Praktikabilität, Wirtschaftlichkeit) vergleichen.
3. Im Rahmen der Prüfung ist zudem zu beachten:
 - Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder sind in die Diskussion bei der Auswahl und Weiterentwicklung eines Geheimhaltungsverfahrens einzubeziehen.
 - Bei der Erstellung oder Bearbeitung fachstatistischer Geheimhaltungskonzepte sind die geheimhaltungsrelevanten Aspekte und Besonderheiten von Auswertungen auf Basis georeferenzierter Daten zu berücksichtigen.
 - Bei der Neueinführung eines datenverändernden Geheimhaltungsverfahrens müssen im Vorfeld die relevanten Nutzergruppen einbezogen werden.

4. Dem zuständigen Gremium legt die Projektgruppe „Geheimhaltung“ einen Vorschlag für ein Geheimhaltungsverfahren vor, der sich durch das Ergebnis der Prüfung auf Basis der Entscheidungskriterien zur Auswahl eines Geheimhaltungsverfahrens begründen lässt. Ggfs. kann das zuständige Gremium einen solchen Vorschlag auch selbst erarbeiten.

Das Ziel der Entwicklungsstufe II ist erreicht, wenn ein Verfahren zur Sicherstellung der Geheimhaltung durch einen Beschluss des für die Statistik zuständigen Gremiums ausgewählt wurde. Es erfolgt sodann die Neueinordnung der Statistik in die Entwicklungsstufe III.

Stufe III: Erstellung eines abgestimmten Geheimhaltungskonzepts

Die grundsätzliche Aufgabe der Entwicklungsstufe III besteht in der Entwicklung und Abstimmung eines statistikspezifischen Fachkonzepts für eine im Statistischen Verbund einheitliche, abgestimmte Geheimhaltung durch das zuständige Gremium oder eine eingesetzte Projektgruppe. Zur Dokumentation der Fachkonzepte bzw. der Geheimhaltungsmethoden und der gewählten Parameter wurde das Format des statistikspezifischen Geheimhaltungsleitfadens entwickelt.

Abhängig von der Entscheidung des zuständigen Gremiums für ein informationsreduzierendes oder ein datenveränderndes Verfahren in Stufe II sind unterschiedliche Punkte zu beachten.

Informationsreduktion

Bei einer Entscheidung für ein informationsreduzierendes Verfahren, wie beispielsweise Zellspernung oder Vergrößerung, sind folgende Aspekte zu berücksichtigen beziehungsweise festzulegen:

1. Einheitliche Regeln für die primäre Geheimhaltung werden durch Beschluss des zuständigen Gremiums definiert.
2. Das Verbreitungsprogramm im Statistischen Verbund wird abgestimmt.
3. Es erfolgt eine Koordinierung von ggfs. erforderlichen Sekundärspernungen im abgestimmten Verbreitungsprogramm, wenn:
 - Ergebnisse regelmäßig (als Standardauswertungen) sowohl für den Bund als auch auf Landesebene publiziert werden und
 - zur sekundären Geheimhaltung darin regelmäßig Tabellenfelder auf der Landesebene (oder ansonsten auf Bundesebene) gesperrt werden müssen.

Das Konzept muss zudem berücksichtigen, dass die Ergebnisse von Ländern, die gegebenenfalls keine Veröffentlichungen tätigen, nicht automatisch zur sekundären Geheimhaltung herangezogen werden können, weil entsprechende Merkmalskombinationen in diesen Ländern mitunter gar nicht oder zu schwach besetzt sein können.

4. Bei Anwendung der Zellspernung werden die Methodik zur Umsetzung der koordinierten Sekundärspernungen und die dafür genutzten Werkzeuge festgelegt.

5. Eine Abstimmung zum Umgang mit der Geheimhaltung unterhalb der Landesebene wird vorgenommen:

- Verbundeinheitliches Veröffentlichungsprogramm, zum Beispiel RegioStat
- Regional unterschiedliches Veröffentlichungsprogramm aufgrund von Länderspezifika

6. Anforderungen bezüglich einer möglichen maschinellen Umsetzung von Zellspernungen werden definiert.
7. Es erfolgt die Erstellung eines Arbeits- und Zeitplans zur Umsetzung.

Datenveränderung

Bei einer Entscheidung für ein datenveränderndes Geheimhaltungsverfahren (zum Beispiel deterministische Rundung, stochastische Überlagerung oder Mikroaggregation) sind in der Entwicklung des Geheimhaltungskonzepts folgende Aspekte zu berücksichtigen bzw. festzulegen:

1. Die Parameter des betreffenden Verfahrens werden durch Beschluss des für die Statistik zuständigen Gremiums festgelegt.
2. Weitere methodische Festlegungen (zum Beispiel Umgang mit Verhältniszahlen, Mittelwerten etc.) werden durch das zuständige Gremium beschlossen.
3. Bei der Einführung der stochastischen Überlagerung auf Basis der Cell-Key-Methode gilt: Die Vorgehensweise der Umsetzung wird durch einen Beschluss des zuständigen Gremiums festgelegt, das heißt Verwendung/Neuprogrammierung eines entsprechenden Auswertungstools, Veröffentlichung der Ergebnisse in einer Auswertungs-/Veröffentlichungsdatenbank oder in einem anderen Auswertungssystem.
4. Ein detailliertes, gegebenenfalls vorläufiges Kommunikationskonzept wird erstellt, welches Kommunikationsmaßnahmen gegenüber allen relevanten Nutzergruppen enthält. Dabei ist auch eine Abstimmung mit dem Arbeitskreis „Presse, Publikationen, Marketing“ des Statistischen Verbunds hinsichtlich der statistikübergreifenden Vereinheitlichung der Kommunikation erforderlich.
5. Es werden Anforderungen bezüglich einer maschinellen Umsetzung der datenverändernden Geheimhaltung ausgearbeitet.
6. Es erfolgt die Erstellung eines Arbeits- und Zeitplans bis zum ersten Einsatz des Verfahrens.

Das Ziel der Entwicklungsstufe III ist erreicht, wenn ein Fachkonzept zur einheitlichen Sicherstellung der Geheimhaltung inklusive eines Zeitplans zur Umsetzung entwickelt und in einem Geheimhaltungsleitfaden dokumentiert wurde sowie das für die Statistik zuständige Gremium dem Fachkonzept beziehungsweise Geheimhaltungsleitfaden zugestimmt hat. Es erfolgt sodann die Neueinordnung der Statistik in die Entwicklungsstufe IV.

Stufe IV: Aufbau und Optimierung eines automatisierten und medienbruchfreien Ablaufs der Geheimhaltung („Digitalisierung“)

Diese Entwicklungsstufe verfolgt das Ziel der technischen Umsetzung des in den Stufen I bis III geplanten Geheimhaltungsprozesses. Dabei lässt sich Stufe IV in zwei Phasen unterteilen: Zunächst erfolgt eine Prüfung durch das für die Statistik zuständige Gremium (oder eine eingesetzte Projektgruppe), ob für eine Statistik Bedarf hinsichtlich einer Automatisierung der Geheimhaltung besteht.

Falls kein Bedarf zur technischen Automatisierung beziehungsweise Optimierung festgestellt wird, ist dies durch das zuständige Statistikgremium zu beschließen. Im statistikspezifischen Geheimhaltungsleitfaden ist dann entsprechend zu erläutern, weshalb kein Bedarf hinsichtlich einer maschinellen/automatisierten Umsetzung besteht (zum Beispiel, weil der Bedarf im Rahmen einer gegebenenfalls geplanten Auswertungsdatenbank abgedeckt wird oder es wird dargelegt, welche organisatorischen Maßnahmen zur manuellen Umsetzung der sekundären Geheimhaltung festgelegt werden). In diesem Fall entfallen die im Folgenden dargelegten Stufen der zweiten Phase und Stufe IV ist abgeschlossen.

Falls weiterer Bedarf hinsichtlich einer technischen Automatisierung beziehungsweise Optimierung festgestellt wird, tritt die Statistik in die zweite Phase der Stufe IV ein: Dabei sind die weiteren zu berücksichtigenden Aspekte bei der Umsetzung und Implementierung einer technischen Lösung wiederum von der Wahl eines informationsreduzierenden oder eines datenverändernden Verfahrens abhängig.

Informationsreduktion

Bei der Entscheidung für ein informationsreduzierendes Verfahren (dies betrifft insbesondere die Wahl eines Zellsperverfahrens) sind folgende Aspekte zu berücksichtigen:

1. Die maschinelle Umsetzung der primären Geheimhaltung gemäß dem in Stufe III entwickelten Fachkonzept wird implementiert.
2. Die maschinelle Umsetzung der sekundären Geheimhaltung gemäß dem Anforderungskonzept aus Stufe III wird implementiert und gegebenenfalls optimiert.
3. Die maschinelle Übernahme der Sperrmuster in das Veröffentlichungsprogramm erfolgt gemäß dem Anforderungskonzept aus Stufe III. Es erfolgt eine Optimierung der Automatisierung, insbesondere bezüglich einer medienbruchfreien Übernahme gesetzter Sperrungen.
4. Das automatisierte Befüllen von Auskunftsdatenbanken (zum Beispiel RegioStat) wird implementiert.

Das Ziel der Entwicklungsstufe IV ist bei Einsatz eines Zellsperverfahrens erreicht, wenn das Vorgehen zur maschinellen Umsetzung der Zellsperung gemäß dem in Stufe III entwickelten Fachkonzept anwendungsreif implementiert ist. Eventuell sind Anforderungen an die Weiterentwicklung der Geheimhaltungsinstrumente beziehungsweise

an einen automatisierten und medienbruchfreien Prozess erhoben und dokumentiert. Das für die Statistik zuständige Gremium hat den erfolgreichen Einsatz des neuen Instruments mit den gewählten Parametern und für das festgelegte Set an Auswertungen abgenommen. Für Geheimhaltung in über diesen Standard hinausgehenden Auswertungen sind im Geheimhaltungsleitfaden aus Stufe III die Parameter festgelegt oder es ist ggfs. auch dafür das Instrument aus Stufe IV nutzbar.

Datenveränderung

Bei der Entscheidung für ein datenveränderndes Verfahren sind folgende Aspekte zu berücksichtigen:

1. Es erfolgt eine Nutzung beziehungsweise (Weiter-)Entwicklung von praxistauglichen Auswertungswerkzeugen mit integrierter Geheimhaltung gemäß dem in Stufe III entwickelten Fachkonzept (insbesondere bei Anwendung der stochastischen Überlagerung auf Basis der Cell-Key-Methode).
2. Eine integrierte Kennzeichnung von Ergebnissen mit großen Abweichungen zum Originalwert ist insbesondere bei Anwendung prätabularer Verfahren, falls im Fachkonzept aus Stufe III vorgesehen, technisch umgesetzt.
3. Falls Statistiken eine deterministische Rundung einsetzen: Bei Bedarf kann eine Prüfung erfolgen, ob die Voraussetzungen für einen Umstieg auf die stochastische Überlagerung auf Basis der Cell-Key-Methode gegeben sind (siehe [3] zu Vor- und Nachteilen beider Verfahren).
4. Gegebenenfalls sind Anforderungen an die Weiterentwicklung von Auswertungswerkzeugen zur Gewährleistung eines automatisierten und medienbruchfreien Geheimhaltungsprozesses zu beschreiben. Das Ziel der Entwicklungsstufe IV bei Einsatz eines datenverändernden Verfahrens ist erreicht, wenn die im in Stufe III erstellten Fachkonzept beschriebene Methode anwendungsreif implementiert ist. Gegebenenfalls wurde darüber hinaus ein automatisierter und medienbruchfreier Prozess ausgearbeitet und dokumentiert. Das für die Statistik zuständige Gremium hat die Implementierung als erfolgreich bestätigt, dem Konzept eines automatisierten und medienbruchfreien Prozesses zugestimmt und dieses an die übergeordneten zuständigen Gremien weitergegeben.

Abschluss und Dokumentation des Prozesses

Nach Abschluss der Entwicklungsstufe IV sind alle Entwicklungsstufen durchlaufen und der Prozess zur Erstellung und Implementierung eines Geheimhaltungskonzepts für eine Statistik ist abgeschlossen. Bei Bedarf kann das für die Statistik zuständige Gremium bestimmte Stufen jederzeit erneut durchlaufen. Für jede Statistik dokumentiert ein statistikspezifischer Geheimhaltungsleitfaden den erarbeiteten, im Verbund abgestimmten, einheitlichen und (soweit sinnvoll möglich) automatisierten

Geheimhaltungsprozess. Bei Statistiken, bei denen aufgrund einzelgesetzlicher Sonderregelung kein Geheimhaltungsbedarf festgestellt wurde, gibt es einen Beschluss, der die Entscheidung begründet.

Spätestens beim Eintritt einer Statistik in Stufe IV müssen die statistikspezifischen Geheimhaltungsleitfäden aktualisiert und finalisiert werden.

Sarah Gießing leitet das Referat *Statistische Geheimhaltung* des Statistischen Bundesamtes.

Julia Höninger leitet das Referat *Gesamtrechnungen, Forschungsdatenzentrum* des Amtes für Statistik Berlin-Brandenburg.

Arijana Amina Ramic ist Referentin im Stabsreferat *Standardisierung der fachstatistischen Prozesse* des Statistischen Bundesamtes.

Dr. Johannes Rohde ist Referent im Referat *Forschung, Entwicklung, FDZ der statistischen Landesämter* des Landesbetriebs Information und Technik Nordrhein-Westfalen.

Literaturverzeichnis

- [1] Amt für Veröffentlichungen der Europäischen Union (2018): Verhaltenskodex für europäische Statistiken für die nationalen statistischen Ämter und Eurostat (statistisches Amt der EU), angenommen vom Ausschuss für das Europäische Statistische System am 16. November 2017.
- [2] Gießing, Sarah; Habla, Heike; Höninger, Julia; Hoffmeister, Rita; Merz, Franz-Josef; Richter, Alexander; Scharnhorst, Sebastian; Schmidtke, Kerstin; Spies, Lydia; Tonte, Andreas; Urich, Stefanie (2018): Handbuch zur Statistischen Geheimhaltung. Stand 12. April 2018. Internes Dokument des Statistischen Verbunds.
- [3] Rohde, Johannes; Seifert, Christiane; Gießing, Sarah; Setzer, Stefanie; unter Mitarbeit von Breitenfeld, Jörg; Brings, Stefan; Höhne, Jörg; Höninger, Julia; Rothe, Patrick; Schedding-Kleis, Ulrike (2018): Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens. Version 1.0 vom 18.04.2018. Internes Dokument des Statistischen Verbunds, verfügbar im Intranet der Statistischen Ämter des Bundes und der Länder.
- [4] Rohde, Johannes; Seifert, Christiane; Gießing, Sarah (2018): Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens, erschienen in WISTA 3/2018, Wirtschaft und Statistik des Statistischen Bundesamtes, S. 90–104.

Neuerscheinung

Indikatorenbericht der UGRdL – aktualisierte Länderergebnisse zum Umweltstatus

Wie viel Primärenergie verbrauchen die Wirtschaft und die privaten Haushalte? Wie viel Erholungsfläche steht der Bevölkerung in einem Bundesland durchschnittlich zur Verfügung? Wie haben sich die CO₂-Emissionen im Vergleich zu 1990 verändert? Diese und weitere Kennzahlen der Umwelt wurden vom Arbeitskreis Umweltökonomische Gesamtrechnungen der Länder in einem Indikatorenbericht zusammengestellt. Der Bericht enthält Daten zu allen Bundesländern und umfasst ein breites Themenspektrum von A wie Abfall bis W wie Wasser. Ergänzend dazu wird die Entwicklung ausgewählter Umwelt- und Nachhaltigkeitsindikatoren in einer auf statistischen Methoden basierenden Trendanalyse dargestellt.

Seit mehr als 20 Jahren liefern die Umweltökonomischen Gesamtrechnungen der Länder (UGRdL) statistische Informationen zu umwelt- und nachhaltigkeitsrelevanten Themen wie Abfall, Energie, Fläche und Raum, Treibhausgase, Rohstoffe, Umweltschutz, Verkehr und Umwelt oder Wasser. Sie stellen im Rahmen der amtlichen Statistik die Grundlage für umweltrelevante Diskussionen und Entscheidungen in Form von Daten und Berechnungen zur Beschreibung der Wechselwirkungen zwischen Wirtschaft, privaten Haushalten und Umwelt bereit. Genutzt werden die Daten der UGRdL unter anderem in den Umwelt- und Nachhaltigkeitsberichten der Länder sowie dem Klimafolgenmonitoring.

Der aktuell veröffentlichte Bericht enthält 17 ausgewählte Umweltindikatoren. Bei der Auswahl der Indikatoren standen die globalen Ziele für nachhaltige Entwicklung der Vereinten Nationen im Fokus. Die Zuordnung der Indikatoren zu den Zielen wurde entsprechend mit Icons visualisiert. Jeder Indikator wurde darüber hinaus grafisch aufbereitet, beschrieben und seine aktuelle Entwicklung kurz erläutert.

Mit der Trendanalyse bieten die UGRdL darüber hinaus eine Methode für objektive und statistisch fundierte Aussagen zur Entwicklung von Umweltindikatoren. Im betrachteten Zeitraum zeigen beispielsweise

der Wirtschaftsindikator „Bruttoinlandsprodukt je Einwohnerin und Einwohner“ sowie die Umweltindikatoren „Pro-Kopf-Aufkommen an Haus- und Sperrmüll“ und „Anteil erneuerbarer Energien am Primärenergieverbrauch“ in fast allen Bundesländern eine positive Entwicklung. Der Pro-Kopf-Verbrauch an Primär- und Endenergie blieb dagegen meist konstant.



Der Indikatorenbericht ist als kostenloser Download unter <http://www.statistikportal.de/de/ugrdl> verfügbar. Über den QR-Code gelangen Sie direkt dahin.

Gesundheitswesen

▣ Sterben die Menschen in Berlin und Brandenburg mit oder an Corona?

von **Katrin Möbius**

In Zeiten der COVID-19-Pandemie werden unterschiedliche, teilweise kreative Theorien zu verschiedensten Lebenssachverhalten verbreitet. Eine davon lautet, dass COVID-19 als Todesursache überproportional häufig und teilweise nicht regelkonform in der amtlichen Todesursachenstatistik Verwendung findet. Der folgende Beitrag erläutert die Vorgehensweise bei der Ermittlung der Todesursache im Amt für Statistik Berlin-Brandenburg (AfS) und soll dazu beitragen, dieser Behauptung durch die transparente Darlegung der verwendeten einheitlichen Verfahren entgegenzuwirken.

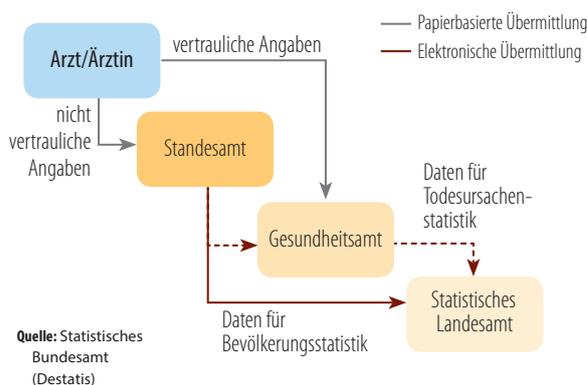
Die Todesursachenstatistik wird auf Grundlage der §§1 Absatz 1 Nr. 1c und 2 Absatz 6 Bevölkerungsstatistikgesetz¹ erhoben. Sie basiert auf der Auswertung der Leichenschauheine, welche dem zuständigen Referat des AfS spätestens am 15. des dem Sterbemonat folgenden Monats von den 18 Gesundheitsämtern des Landes Brandenburg beziehungsweise dem Zentralarchiv für Leichenschauheine (ZfL) der Gesundheitsämter Berlins übermittelt werden. Seit 2013 besteht zudem, auch auf Grundlage des §11a Absatz 1 Bundesstatistikgesetz², eine gesetzliche Verpflichtung zur elektronischen Datenlieferung (Abbildung a).

Ausschlaggebend für die Belastbarkeit der Todesursachenstatistik ist vor allem die inhaltliche Qualität des der Statistik zugrunde liegenden Leichenschauheins. Er listet alle Krankheiten auf, die zum Tode geführt haben. Dabei wird zwischen gesundheitlichen Beeinträchtigungen, die unmittelbar oder mittelbar zum Tode geführt haben, und der Krankheit, die alle anderen bedingt und als sogenanntes Grundleiden und damit als die offizielle Todesursache in die Todesursachenstatistik einfließt, unterschieden.

Die Mitarbeitenden im Bereich Todesursachenstatistik sind somit zwingend auf qualitativ hochwertige, regelkonform ausgefüllte Todesbescheinigungen sowie deren vollständige Übermittlung auf elektronischem Wege angewiesen, um das korrekte Grundleiden ermitteln zu können. Auch deshalb verpflichten die einschlägigen Bestattungsgesetze der Länder Berlin und Brandenburg die zuständige Amtsärztin/den zuständigen Amtsarzt vor Übermittlung der Leichenschauheine an das AfS zur Prüfung ihres Inhalts (vgl. beispielsweise §2 Absatz 7 DVO-Bestattungsgesetz³ des Landes Berlin).

Nach deren Eingang wird das Grundleiden und damit die Todesursache anhand der Internationalen statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD-10 WHO)⁴ ermittelt. Die Vorgaben der ICD-10 WHO sind komplex und erfordern zu ihrer Anwendung fundiertes medizinisches Wissen. Um dem gerecht zu werden, wird im AfS seit 2017 ausschließlich das internationale elektronische Kodiersystem IRIS verwendet. Dieses vom Statistischen Bundesamt in Kooperation mit dem Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) entwickelte Verfahren zur automatischen Kodierung von textlichen Ausdrücken auf Todesbescheinigungen wertet diese in standardisierter Weise nach den Regeln der ICD-10 WHO aus und verhindert somit neben fach-

a | Datenfluss in der Todesursachenstatistik



1 Gesetz über die Statistik der Bevölkerungsbewegung und die Fortschreibung des Bevölkerungsstandes (Bevölkerungsstatistikgesetz – BevStatG) vom 20. April 2013 (BGBl. I S. 826), das zuletzt durch Artikel 9 des Gesetzes vom 18. Dezember 2018 (BGBl. I S. 2639) geändert worden ist.

2 Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I S. 2394), das zuletzt durch Artikel 10 Absatz 5 des Gesetzes vom 30. Oktober 2017 (BGBl. I S. 3618) geändert worden ist.

3 Verordnung zur Durchführung des Bestattungsgesetzes (DVO-Bestattungsgesetz) vom 22. Oktober 1980, §5 geändert, Anlagen 1 und 2 neu gefasst durch Verordnung vom 12.01.2016 (GVBl. S. 12).

4 siehe <https://www.dimdi.de/sta-tic/de/klassifikationen/icd/icd-10-gm/kode-suche/htmlgm2020/>

lichen Fehleinschätzungen aufgrund mangelnder medizinischer Kenntnisse auch die Gefahr eines ungewollten, subjektiven Einflusses der Mitarbeitenden auf die Ergebnisse. IRIS berücksichtigt durch laufende Aktualisierungen immer die aktuell gültigen Regeln der WHO und unterstützt bei der Ermittlung der zum Tode führenden Erkrankung. Ziel des Einsatzes ist darüber hinaus die verbesserte Konsistenz und Vergleichbarkeit der Todesursachenstatistik auf nationaler und internationaler Ebene und die Einführung einer multikausalen Todesursachenstatistik in absehbarer Zukunft. Im AfS werden derzeit maximal 60% der in IRIS verarbeiteten Fälle automatisch kodiert. Nicht automatisch kodierbare Fälle, beispielsweise durch Inkonsistenzen der Angaben auf dem Leichenschauschein, werden manuell nachkodiert. Jede nicht eindeutig bestimmbare Todesursache wird beim zuständigen Gesundheitsamt nachgefragt, um eine größtmögliche Qualität der Statistik sicherzustellen.

Hinsichtlich der Krankheit COVID-19 wird somit klar unterschieden, ob die Person an der Todesursache COVID-19 verstorben ist oder eine andere Krankheit zum Tode führte und COVID-19 nur zum Tod beigetragen hat.⁵

In Teil I ist von der/dem leichenschauenden Ärztin/Arzt die Kausalkette einzutragen, die von der Grunderkrankung zur unmittelbaren Todesursache führte, dabei ist in der untersten Zeile (1 c) die Grunderkrankung zu benennen. Von dort aus folgt nach oben die Ereigniskette, die schließlich in der obersten Zeile (1 a) mit der direkt zum Tode führenden Krankheit endet.

Als Beispiel dient im Folgenden ein fiktiver Todesfall, bei dem COVID-19 an unterschiedlichen Positionen des Leichenschauscheines und somit in IRIS verzeichnet ist (Abbildungen b bis d). Der Leichenschauschein besteht aus zwei Abschnitten, in die entsprechende Angaben zu Krankheiten und Todesursachen einzutragen sind. Diese werden kongruent

b | Fiktiver Fall – COVID-19 eingetragen in ... Teil I Zeile a

ermittelte offizielle Todesursache: Akuter Myokardinfarkt, nicht näher bezeichnet

c | ... Teil I Zeile c

ermittelte offizielle Todesursache: COVID-19, Virus nachgewiesen

d | ... Teil II

ermittelte offizielle Todesursache: Akuter Myokardinfarkt, nicht näher bezeichnet

⁵ Im Rahmen der aktuellen amtlichen Todesursachenstatistik, die eine monokausale Todesursache ausweist, werden lediglich die Fälle, in denen eine Person an COVID-19 verstorben ist, ausgewiesen. Mit der mittelfristig geplanten Umstellung auf eine multikausale Statistik wären später auch zusätzliche Analysen möglich, die auch die Fälle, die mit COVID-19 verstorben sind, ausweisen könnten.

in die Anwendung IRIS übernommen. In Teil II des Leichenschauzeins werden weitere wesentliche Krankheiten eingetragen, die nicht unmittelbar Teil der zum Tode führenden Kausalkette sind.

Abschließend soll an einem extremen fiktiven Fall, der die eingangs erwähnte Theorie ursprünglich untermauern sollte, deren Unrichtigkeit verdeutlicht werden: Erleidet eine Person, die mit COVID-19 infiziert ist, einen Verkehrsunfall mit tödlichen Folgen, wird keinesfalls die COVID-19-Infektion, sondern die tödliche Verletzung aufgrund des Unfalls als Todesursache festgelegt (Abbildungen e und f).

Die Todesursachenstatistik bildet für unterschiedliche nationale und internationale Organisationen das Fundament einer Todesursachenforschung und damit die Entscheidungsgrundlage für verschiedene präventive und medizinisch-kurative Maßnahmen und Strategien. Ihre hohe Qualität und internationale Vergleichbarkeit – gerade in schwierigen Situationen wie dem aktuellen Pandemie-Geschehen – ist von hoher Bedeutung. Die zuständigen Mitarbeitenden des AfS sind sich dieser hohen Verantwortung stets bewusst.

Katrin Möbius leitet das Referat *Bildung, Gesundheitswesen* des Amtes für Statistik Berlin-Brandenburg.

e | Fiktiver Verkehrsunfall – COVID-19 eingetragen in ... Teil II

The screenshot shows the IRIS software interface for a death certificate. In the 'Teil II' section, 'Covid-19 nachgewiesen' is circled in red. In the 'SIGNIERUNG' section, 'V274' is circled in red. A red arrow points from this 'V274' to the text on the right.

ermittelte offizielle Todesursache:
Verletzung des Kopfes,
Fahrer eines Motorrades
bei Zusammenstoß mit feststehen-
dem Gegenstand bei Verkehrs-
unfall verletzt

f | ... Teil I Zeile c

The screenshot shows the IRIS software interface for a death certificate. In the 'Teil I' section, row 'c', 'Covid-19 nachgewiesen' is circled in red. In the 'SIGNIERUNG' section, 'V274' is circled in red. A red arrow points from this 'V274' to the text on the right.

Fachgespräch mit Hartmut Böermann

„Wie systemrelevant ist die amtliche Statistik?“



Foto: Amt für Statistik Berlin-Brandenburg

Hartmut Böermann ist Vertreter des Vorstandes und Leiter der Abteilung *Bevölkerung und Soziales* im Amt für Statistik Berlin-Brandenburg.

Lieferketten brechen ab, Unternehmen schicken ihre Belegschaft in Kurzarbeit, Kinder können nicht mehr regulär in Kitas und Schulen gehen, Freiheitsrechte werden eingeschränkt – die Ungewissheit ist überall präsent. Wie wichtig ist die amtliche Statistik in Zeiten der Corona-Krise?

| Derzeit durchleben wir eine Krise, deren Ende wir noch nicht abschätzen können. Die gesamten Ereignisse erfordern gezielte Maßnahmen, um den Verlauf des Geschehens positiv zu beeinflussen. Die politischen Entscheider benötigen, um evidenzbasiert Alternativen abwägen zu können, Fakten, die möglichst aktuell sind. Sie stützen sich dabei auf Ergebnisse von Prognosemodellen, die wiederum auf Daten beruhen. Genauso wichtig sind Fakten für die Öffentlichkeit und die Medien, da viele Menschen durch die ergriffenen Maßnahmen existenziell betroffen waren und es voraussichtlich noch eine längere Zeit sein werden. Menschen muss sich die Sinnhaftigkeit von Eingriffen des Staates in das wirtschaftliche und soziale

Leben erschließen, damit sie diese für die Zeit der Krise mittragen können. Gelingt dies nicht, schlägt die Stunde der Populisten, die ihre eigene einfache, aber vermeintlich sinnstiftende (Verschwörungs-)Erzählung mitbringen und mit Fake-News scheinbar belegen.

Die Epidemiologie ist hier besonders gefragt, um das Virus, seine Gefährlichkeit und seine Verbreitungswege möglichst gut zu begreifen. Dies betrifft das Pandemiegeschehen in einem engeren Sinn. Die sich anschließenden Fragen zum Sterbefallgeschehen, zu den sozialen und wirtschaftlichen Folgen, wie beispielsweise Preisveränderungen, Kurzarbeit und Zahl der von Kita- und Schulschließungen Betroffenen, fallen in den Beobachtungsbereich der amtlichen Statistik.

Verlässliche, neutrale und objektive Informationen für Parlamente, Ministerien, Forschung und Öffentlichkeit zur Verfügung zu stellen, ist Aufgabe der amtlichen Statistik. Diesen Auftrag nehmen wir fortlaufend wahr, sind in einer Krise aber besonders herausgefordert.

Der französische Wissenschaftshistoriker Alain Desrosières hat in seiner Geschichte der statistischen Denkweise der vergangenen 200 Jahre sehr nachvollziehbar dargelegt, wie bei den Staatengründungen in Europa die Statistik eine der grundlegenden Institutionen war, um Informationen über die Staatsangelegenheiten zu bekommen. Für eine moderne, offene, demokratische und sozial verantwortliche Gesellschaft sind statistische Informationen

unverzichtbar. Und – wie wir heute sehen – gilt dies auch für den gesellschaftlichen Zusammenhalt, für den eine verlässliche und vertrauenswürdige Informationsbasis grundlegend ist. Gleichermäßen trifft dies auch für die Europäische Union zu, in der Eurostat im Zusammenspiel mit den Statistischen Ämtern diese Aufgabe wahrnimmt.

Während der gesamten Dauer der institutionalisierten Statistik hat es eine ständige Weiterentwicklung durch eine Differenzierung der Nachfrage nach belegbaren Fakten und durch Impulse aus der Wissenschaft, durch neue Erhebungsmethoden und Datenquellen sowie durch technische Innovationen – als eine weitere wesentliche treibende Kraft – gegeben.

Die amtliche Statistik zeichnet sich nicht zuletzt durch das breite Spektrum der von ihr produzierten Daten aus. Welche Informationen sind in der aktuellen Situation von besonderer Relevanz?

| Frühzeitig haben sich das Statistische Bundesamt und die Statistischen Ämter der Länder auf ein Minimalprogramm verständigt, das auch in einem Notbetrieb aufrechterhalten werden muss. Zu diesem Minimalprogramm zählen verschiedene Statistiken, die es ermöglichen, den wirtschaftlichen Effekt der Corona-Pandemie zu erfassen und die in die Berechnung des Bruttoinlandsprodukts (BIP) als einen zentralen Wirtschaftsindikator eingehen. Zusätzlich haben wir für Berlin und Brandenburg weitere Statistiken identifiziert, die nicht

zurückgestellt werden dürfen, um Landesaufgaben erfüllen zu können. Dazu gehören der Kommunale Finanzausgleich und der Kommunale Rettungsschirm. Von Eurostat wurden weitere Statistiken als nicht aussetzbar identifiziert, u. a. die Arbeitskräfteerhebung (Labor Force Survey), die in Deutschland eine Unterstichprobe des Mikrozensus ist.

Unsere Pandemie-Vorbereitungen waren und sind darauf abgestellt, diese Erhebungen auch bei einer stärkeren Betroffenheit durch das Infektionsgeschehen durchzuführen.

Das AfS hat ein Corona-Dossier auf seiner Internetseite veröffentlicht. Wer ist die Zielgruppe?

| Mit der Veröffentlichung haben wir das Ziel verfolgt, Fragen, die in der Öffentlichkeit gestellt werden, mit statistischen Fakten zu beantworten. Wir haben also die Außensicht, den Fragenden, den Antwort-Suchenden, den Kunden und die Kundin in den Fokus genommen. Unsere Standardveröffentlichungen haben ein konkretes Programm und feste Veröffentlichungstermine. Die sehr dynamisch verlaufende Entwicklung der Pandemie hat dieses Schema gesprengt. Wir wollen schnell sein und viele erreichen, die nach Informationen suchen.

Welche Daten sind in dem Dossier zu finden?

| Wir sind ausgegangen von den Fragen, die sich in der Krise stellen, und haben dann die Statistiken durchmustert, ob sich diese Fragen anhand unserer Statistiken beantworten lassen. Die Themen erstrecken sich über die Bereiche Gesundheit, Gesellschaft und Wirtschaft. Um dem Anspruch nach größtmöglicher Aktualität zu entsprechen, wird das Dossier laufend aktualisiert und um aktuell verfügbare Informationen ergänzt.

Neu ist, dass wir auch Daten veröffentlichen, die noch nicht den statistischen Erstellungs- und Qualitätssicherungsprozess vollständig durchlaufen haben, aber die auf drängende Fragen eine erste Antwort zulassen. Von den Leserinnen und Lesern fordert das ein höheres Maß an Statistical Literacy; eine Kompetenz, die in unserem Bildungssystem leider zu wenig gefördert wird. Hier werden wir stärker als Vermittler wirken.

Zahlreiche amtliche Daten sind nur mit großem zeitlichem Abstand verfügbar. Zum Beispiel stammen die neuesten Strukturdaten zu Krankenhäusern aus dem Jahr 2017. Wie kommt es dazu? Kann das im Statistischen Verbund geändert werden?

| Mit dieser Frage sprechen Sie einen wichtigen Punkt an. Die Krankenhausstatistik ist bereits in Normalzeiten planungsrelevant, in der jetzigen Situation könnte sie wichtige Strukturdaten liefern und auch zur Klärung der Frage beitragen, ob Bettenkapazitäten zu stark bzw. überhaupt abgebaut wurden. Leider gibt es keine aktuellen Daten, da es bei der Anpassung der Software an die erheblich erweiterte Statistik Verzögerungen gegeben hat. Eine wichtige Botschaft für uns ist, dass solche Verzögerungen nicht akzeptabel sind und wir besser werden müssen, um unserem Anspruch gerecht zu werden.

Im Zusammenhang mit den Schutzmaßnahmen gegen die Corona-Pandemie wurde das föderale System in Deutschland verschiedentlich – positiv wie negativ – thematisiert. Die amtliche Statistik in Deutschland ist ebenfalls föderal organisiert. Welche Vorteile hat die föderale Struktur in der Statistik?

| Das in der Corona-Krise wieder hervorgeholte Argument des „Flickenteppichs“ trifft für die amtliche Statistik nicht zu, da wir im Verbund der Ämter des Bundes und

der Länder methodisch einheitlich arbeiten. Den Vorteil einer föderalen Struktur sehe ich in der Nähe zu den Nachfragenden in den Ländern und zu den Auskunftgebenden. Wir haben zur Landespolitik, der Verwendungsseite der Statistiken, eine enge Arbeitsbeziehung. Wir handeln aber – was ich betonen möchte – diskriminierungsfrei: Unsere Daten sind für alle verfügbar und werden veröffentlicht, sodass wir keine Informationsasymmetrie zwischen Exekutive, Parlament und Öffentlichkeit haben. Beim Erhebungsprozess kommen uns die räumliche Nähe und die Kenntnis regionaler Besonderheiten zugute. Ohne vertiefte Kenntnisse der regionalen Besonderheiten sind Ergebnisse der Statistiken oft nur schwer richtig einzuordnen. Ein großes Plus für den Datenschutz sind die nicht in einem zentralen Silo gespeicherten Daten, die für potenzielle Angriffe ein hoch attraktives Ziel wären und die den möglichen Schaden und Vertrauensverlust maximierten.

Eine föderale Organisation führt zu höheren Kommunikationsaufwänden, die durch digitale Formen der Zusammenarbeit deutlich vereinfacht werden können. Hier hat die Corona-Krise bereits als Katalysator gewirkt.

Inwieweit wirkt sich Corona auf die Statistikproduktion im AfS und die Arbeit im Verbund aus?

| Wie bereits geschildert, haben wir ein Notfallprogramm unter den Statistischen Ämtern des Bundes und der Länder abgestimmt. Vor besondere Herausforderungen sahen sich auch viele Auskunftgebende gestellt. Wir hatten daher ein wertschätzend-motivierendes Schreiben versandt und die Mahnverfahren zeitweilig ausgesetzt.

Beschleunigt hat sich durch Corona die digitale Transformation der Arbeits- und Kommunikationsformen im AfS und im Verbund. Um die räumliche Distanzierung in den Büros zu gewährleisten und für die Aufrechterhaltung des Betriebs im Notfall wurde die Ausstattung mit Telearbeitsmöglichkeiten drastisch erhöht. Statt Präsenzsitzungen gab und gibt es im AfS, wie im Verbund, Telefon- bzw. Videokonferenzen. Vorhaben, die vorher nur langsam vorankamen, wurden plötzlich in kürzester Zeit umgesetzt, weil alle Akteure in- und außerhalb des AfS mitzogen. Der Kulturwandel im AfS (Stichwort Vertrauenskultur, Agilität) hat damit an Momentum gewonnen. Neu genutzte Arbeitsformen und der Gewinn an Zeitsouveränität werden auch nach der Krise bleiben.

Die Corona-Krise hat viele Signale generiert, die wir aufgenommen haben und die wir weiter aufnehmen müssen. Wir haben uns vorgenommen, das AfS auf dem Weg der digitalen Transformation und des Kulturwandels zu öffnen, flexiblen Arbeitsstrukturen voranzubringen, damit die amtliche Statistik auch in Zukunft ihren Rang behaupten kann.

Statistik erklärt: Nowcast und Übersterblichkeit

Sowohl die COVID-19-Fallzahlen zu Todesfällen und Infektionen, welche vom Robert Koch-Institut auf Grundlage der Daten der Gesundheitsämter zusammengestellt werden, als auch Sterbefälle im Allgemeinen sind, obwohl sie zu einem bestimmten Zeitpunkt bereits erfasst sind, für die Ausweisung zu diesem Zeitpunkt häufig unvollständig. Generell besteht die Erwartung, dass nach einer Datenerfassung der gegenwärtige Stand bekannt ist und Voraussagen ausschließlich für künftige Zeitpunkte notwendig sind. Warum sind bei den Fallzahlen zu Verstorbenen und COVID-19-Infizierten, im Gegensatz zu anderen Daten, die zu einem bestimmten Zeitpunkt vorliegenden Fälle aber zu genau diesem Zeitpunkt nicht vollständig bekannt?

Das Phänomen, dass Prognosen nicht für einen zukünftigen Zeitpunkt erstellt werden, sondern für eine Zeit, die bereits eingetreten ist, ist nicht nur in der derzeitigen Corona-Krise relevant. Insbesondere Daten, die einen gewissen Prozess durchlaufen, bis sie auswertbar sind, erfordern Prognosen für den aktuellen Zeitpunkt. Dabei muss es sich nicht ausschließlich um den aktuellen Tag handeln, sondern es stehen auch Tage oder Wochen der jüngeren Vergangenheit im Fokus. Zu genau diesen Daten gehören die Fallmeldungen des Robert Koch-Instituts sowie die derzeit stark nachgefragten Sterbefallzahlen. Da Sterbefälle von Ärzten mittelbar über die Standesämter an die Statistischen Ämter der Länder gemeldet werden, kann ein sogenannter Meldeverzug entstehen, welcher in einer Zeitreihe durch das Sinken des Graphen am aktuellen Rand deutlich wird. Diese Abnahme darf dabei keineswegs als Rückgang der Sterbefallzahlen selbst interpretiert werden, da durch den Meldeverzug die Daten noch nicht vollständig vorliegen.

Um diese Art der Fehlinterpretation zu vermeiden, kann es sinnvoll sein, Daten für aktuelle Zeitpunkte zu schätzen. Dadurch besteht auch kurzfristig die Möglichkeit, das Niveau zu erkennen, auf dem sich die Zahlen momentan befinden. Solche Schätzungen werden oft mit dem Begriff „Nowcast“ bezeichnet. Dieser Fachausdruck ist eine Kontraktion der Wörter „Now“ (jetzt) und „Forecast“ (Vorhersage). Dahinter verbergen sich statistische Verfahren, welche Prognosen für die Zukunft erstellen. Zum Teil können diese Methoden auch für einen Nowcast verwendet werden und helfen damit, aktuelle, aber aus verschiedenen Gründen noch unvollständige Daten zu prognostizieren; um das obige Beispiel aufzugreifen: etwa den bekannten Meldeverzug auf Grundlage bekannter Daten auszugleichen.

Unter dem Begriff der **Übersterblichkeit** in einer bestimmten Zeitspanne wird im Allgemeinen eine Erhöhung der Sterbefallzahlen eines bestimmten Zeitabschnitts gegenüber dem Vergleichszeitraum der Vorjahre verstanden. Für die Bestimmung des Vergleichsmaßes werden die Sterbefallzahlen von beispielsweise fünf Vorjahren verwendet und über diese Jahre ein Mittelwert gebildet. In den letzten Jahren gab es in Berlin einen leichten Aufwärtstrend bei den Sterbefallzahlen. Dieser wird jedoch nicht als Übersterblichkeit interpretiert.

Mit der Übersterblichkeit werden nur kurzfristige Abweichungen vom „Normalzustand“ beschrieben, wobei der Normalzustand durch die Fallzahlen der letzten Jahre definiert wird. Klassische Beispiele in diesem Zusammenhang sind Grippe- oder Hitzewellen, welche zu einer temporären Häufung von Sterbefällen führen und sich in Graphen durch starke Spitzen im zeitlichen Verlauf widerspiegeln.

Georeferenzierung

Regionalanalyse auf Basis simulierter Geokoordinaten Gütebeurteilung des Verfahrens am Beispiel der Wahlberechtigten in Berlin

von Kerstin Erfurth

Für Daten mit geografischem Bezug eignen sich Kartendarstellungen zur Visualisierung, um einen einfachen Zugang zu komplexen Informationen zu erhalten. Insbesondere die Verteilung verschiedener Bevölkerungsgruppen und die Identifikation von Hotspots stellen ein für Planungszwecke bedeutendes Interesse dar. Diese Arbeit beschäftigt sich mit der Bewertung des neuen Kernelheaping-Verfahrens gegenüber anderen in der Praxis gängigen Verfahren zur kartografischen Dichteschätzung von Daten. Dazu wurde ein praxisnahes Szenario mit den Daten der Wahlberechtigten in Berlin geschaffen, in welchem unter kontrollierten Bedingungen Vergleiche durchgeführt werden können. Es konnte gezeigt werden, dass das Kernelheaping-Verfahren in der Lage ist, qualitativ bessere Ergebnisse zu erzielen als die bisher verwendeten Standardverfahren.

Dieser Beitrag ist ein Nachdruck aus: Statistisches Bundesamt, (2020): WISTA, Heft 2/2020, S. 25.

1 Einführung

Die Erhebung und Interpretation von Daten ist zu einem zentralen Thema in der modernen Informationsgesellschaft geworden. Ein einfacher Zugang zu komplexen Informationen wird durch geeignete Visualisierungen ermöglicht, wobei die gewählte Darstellungsmethode einen wesentlichen Einfluss auf die Interpretation der Daten haben kann. Für Daten mit geografischem Bezug eignen sich insbesondere Kartendarstellungen, welche unter anderem mit farbigen Symbolen, Grenzlinien oder Flächen angereichert werden, um deren Raumbeziehungen und ihre relativen Verhältnisse leicht verständlich zu machen. Dabei spielen auch Diskretisierung¹, Kategorisierung und die verwendeten Farbabstufungen eine große Rolle, da der visuelle Eindruck durch deren Wahl stark beeinflusst werden kann. Im Umgang mit aggregierten Daten wird der Ansatz der Vorverarbeitung zu einer Schlüsseltechnik für gute Ergebnisse. Um umfassende Informationen über alle interessierenden Geokoordinaten zu erhalten, wird ein neuer nicht parametrischer Ansatz zur Dichteschätzung namens Kernelheaping evaluiert.

Der neue Ansatz wird statistisch mit einer zugrunde liegenden bekannten realen Dichte von Wahlberechtigten in Berlin auf verschiedenen Aggregationsebenen bewertet und quantitativ den Standardverfahren vergleichend gegenübergestellt. Dafür konnte auf anonymisierte Adressdaten des Amtes für Statistik Berlin-Brandenburg zugegriffen werden.

Es wird gezeigt, dass das Kernelheaping von den untersuchten Verfahren die beste Möglichkeit bietet, lokale Aggregate unabhängig von Verwaltungsgrenzen, zum Beispiel Wahlbezirken, zu behandeln.

Es ermöglicht geografische Kartendarstellungen mit exakten Geokoordinaten, welche als „Wähler je Pixel“ interpretiert werden können, auch wenn die exakten Geokoordinaten für Wahldaten ursprünglich nicht zur Verfügung stehen. Dies dient neben der Visualisierung von Verteilungen auch der Identifikation von Hotspots interessierender Personengruppen.

2 Datengrundlage „Adressdichte“

Als Analysegrundlage werden anonymisierte Einzeldaten zur Anzahl der Wahlberechtigten in Berlin auf Adressebene genutzt. Sie stellt eine Grundgesamtheit dar, da alle Personen erhoben wurden, die im Dezember 2016 ihren Hauptwohnsitz in Berlin gemeldet hatten. Der kritische Aspekt ist demzufolge nicht der Informationsverlust durch die Ziehung einer Stichprobe, von welcher auf die Grundgesamtheit geschlossen werden soll. Der Fokus liegt auf dem Informationsverlust durch die Aggregation der Daten. Es werden prinzipiell alle Personen erfasst, ihre räumlichen Geokoordinaten werden jedoch gerundet beziehungsweise aggregiert. Daher ist die Anwendung insbesondere für amtliche Daten interessant, beispielsweise für die in dieser Arbeit untersuchten Wahlberechtigtenzahlen.

Voraussetzung für nicht parametrische Ansätze für Dichteschätzungen ist die „Glattheit“ der zu schätzenden Dichte (Fahrmeir et al. 1996). Aus diesem Grund wird über die Adressdaten eine minimale Kerndichte als Glättungsprozedur gelegt. Auf diese

¹ Diskretisierung beschreibt die Zerlegung stetiger, räumlicher Flächen in kleine Abschnitte oder Punkte.

Weise wird eine geglättete Version der Originaldaten erzeugt, welche nicht nur die Voraussetzungen für Dichteschätzungen erfüllt, sondern auch bessere Visualisierungsmöglichkeiten liefert. Wird der Volumeninhalt unter der bivariaten Dichte-Kurve auf Eins normiert, entsteht die „Adressdichte“ der Wahlberechtigten. Diese Adressdichte ist die Ausgangsbasis der Untersuchungen und stellt eine sehr realitätsnahe Datengrundlage einer „wahren Dichte“ dar.

Um zu einer numerisch verarbeitbaren Datenbasis zu gelangen, muss darüber hinaus eine Diskretisierung in Pixel durchgeführt werden. Auch zu Darstellungszwecken ist dieser Schritt unvermeidbar. Dabei wird jedem Pixel ein Dichtewert zugeordnet, sodass unter Berücksichtigung der Pixelgröße der Flächeninhalt unter allen Pixeln ebenfalls Eins ergibt. Die diskretisierte Version der Adressdichte beschreibt demzufolge wieder eine Dichte.

Diese Vorverarbeitungsschritte liefern die finale Adressdichte auf Pixelbasis. Dabei wurde in den Untersuchungen die Größe der Pixel so gewählt, dass sie etwa einem Hektar Landfläche entspricht. Bei einer Umrechnung des Dichtewertes auf Wahlberechtigte entsteht eine leicht zu interpretierende Größe „Wahlberechtigte je Hektar“. Diese Adressdichte wird in allen Berechnungen als wahre Dichte angenommen und genutzt, um Fehlerterme und Gütekriterien zu berechnen. Alle untersuchten Dichteschätzungen werden auf das gleiche Raster zurückgeführt, um eine Vergleichbarkeit herzustellen (Abbildung a).

3 Aggregationslevel

Um Berechnungen auf Basis unterschiedlicher Aggregationsstufen kontrolliert durchführen zu können, werden die ursprünglichen Wahlberechtigtenzahlen zunächst auf acht vorab gewählte regionale Bezugssysteme kumuliert. Auf diese aggregierten Daten werden die verschiedenen Verfahren angewendet und anschließend verglichen. Je mehr Polygonzüge ein Aggregationslevel besitzt, desto detaillierter ist es und desto geringer ist der Informationsverlust im Vergleich zu den Originaldaten. In der Untersuchung wurden die zwölf Berliner Bezirke (BEZ), 60 Prognose-räume (PRG), 96 Ortsteile (ORT), 138 Bezirksregionen (BZR), 192 Postleitzahlen (PLZ), 447 Planungsräume (PLR), 660 Briefwahlbezirke (BWB) und 1779 Urnenwahlbezirke (UWB) verwendet.

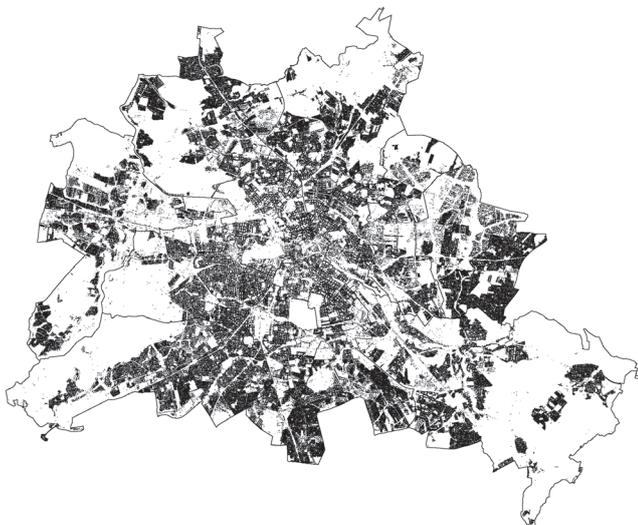
Diese Strukturen sind zum Teil natürlich gewachsen und liegen nicht in einem regelmäßigen Raster oder Gitter. Teilweise sind die Flächen unterschiedlich groß mit stark variierenden Einwohnerzahlen. Derartige Inhomogenitäten treten sehr häufig bei administrativen Grenzen auf. Mit der Betrachtung verschiedener Aggregationsstufen kann abhängig von den gewählten Verfahren festgestellt werden, ab welchem Grad des Informationsverlusts in den Daten eine Anwendung überhaupt noch sinnvoll sein kann.

4 Verfahren

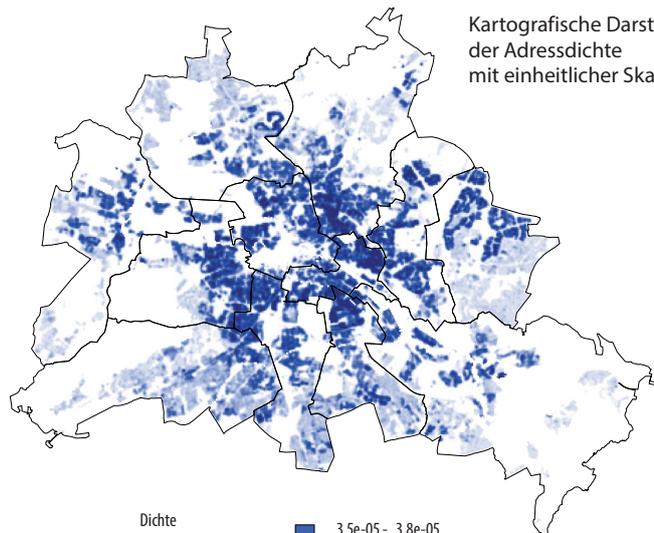
Die Verfahren, welche zentral verglichen werden, sind Choroplethenkarten, klassische (naive) Kerndichteschätzung und das Kernelheaping-Verfahren für simulierte Geokoordinaten.

a | Umrechnung auf die finale Adressdichte auf Pixelbasis

Verteilung der Berliner Adressen über das Land mit exakten Geokoordinaten



Kartografische Darstellung der Adressdichte mit einheitlicher Skala



Dichte	
>3.4e-06	3.5e-05 - 3.8e-05
3.4e-06 - 6.3e-06	3.8e-05 - 4.1e-05
6.3e-06 - 9.2e-06	4.1e-05 - 4.4e-05
9.2e-06 - 1.2e-05	4.4e-05 - 4.7e-05
1.2e-05 - 1.5e-05	4.7e-05 - 5.0e-05
1.5e-05 - 1.8e-05	5.0e-05 - 5.3e-05
1.8e-05 - 2.1e-05	5.3e-05 - 5.6e-05
2.1e-05 - 2.4e-05	5.6e-05 - 5.8e-05
2.4e-05 - 2.7e-05	5.8e-05 - 6.1e-05
2.7e-05 - 2.9e-05	6.1e-05 - 6.4e-05
2.9e-05 - 3.2e-05	6.4e-05 - 6.7e-05
3.2e-05 - 3.5e-05	6.7e-05 - 7.0e-05
	>7.0e-05

Quellen: Anonymisierte Adressdaten des Amtes für Statistik Berlin-Brandenburg, Stichtag 31.12.2016

Die im Fokus stehende Methode ist dabei der Kernelheaping-Algorithmus. Er erzeugt iterativ eine Kerndichte und berücksichtigt damit die Tatsache, dass regional aggregierte Daten vorliegen. Dieser Algorithmus ist eine Anwendung eines Stochastic-Expectation-Maximization(SEM)-Algorithmus, bei dem der stochastische Teil als geschichtete Stichprobe von Geokoordinaten aus der Dichte des vorherigen Schrittes realisiert wird. Eine Schicht entspricht dabei einem Polygon. Nach einigen Iterationen konvergiert der Algorithmus zu einer Dichte mit Schätzwerten für alle Geokoordinaten. Für eine erfolgreiche Anwendung ist eine feste, aber frei wählbare Anzahl von Iterationen erforderlich. In der Analyse wird dieser Parameter systematisch modifiziert, um Unterschiede in den Ergebnissen aufzuzeigen und optimale Parametereinstellungen zu finden.

Zusätzlich werden klassische Choroplethenkarten auf die Datensätze angewendet. Diese sind Standard für die Visualisierung von aggregierten Datensätzen in der amtlichen Statistik. Für Choroplethenkarten sind keine genauen Geokoordinaten erforderlich. Ihr größter Nachteil ist die homogene Farbe innerhalb der regionalen Einheiten. In der Regel sind die Farbkategorien auf etwa fünf Ebenen begrenzt, was einen erheblichen Informationsverlust bedeutet. Darüber hinaus haben regionale Einheiten im Allgemeinen nicht die gleiche Größe. Daher kann die Interpretation von Choroplethenkarten über die Flächengrößen, welche am attraktivsten ist, zu irreführenden Schlussfolgerungen führen. In der Analyse wird gezeigt, dass einfache Flächennormalisierungen bereits helfen, realistischere Kartendarstellungen zu erstellen.

Für Vergleichszwecke wird zudem ein naives Kerndichte-Schätzverfahren (kernel density estimation – KDE) verwendet. Es handelt sich um eine gängige Glättungstechnik, mit der eine Dichte für aggregierte Daten geschätzt werden kann. Ein wesentlicher Nachteil dieses Verfahrens ist die komplexe Aufgabe, geeignete Glättungsparameter (Bandbreiten) zu finden.

Für alle Methoden wurden systematisch Parameter angepasst, um optimale Einstellungen aufzuzeigen und die Techniken hinsichtlich ihrer Robustheit zu bewerten.

5 Diskretisierung

Wie oben erwähnt, wurden alle Dichten in gleicher Art und Weise diskretisiert, um eine einheitliche Berechnungsgrundlage zu schaffen. Dabei wurde für alle Methoden eine identische Rastergröße (Gridsize) festgelegt, welche eine interpretierbare Pixelgröße liefert. Aus diesem Grund ist die Gridsize selbst zunächst eine nicht intuitiv krumme Zahl. Berlin besitzt eine Ost-West-Ausdehnung von etwa 45784 Metern. Sollen Pixel mit einer Fläche von einem Hektar entstehen, so ist eine Gridsize von 458 Pixel in horizontaler Richtung sinnvoll. Die Nord-Süd-Spanne Berlins beträgt etwa 37739 Meter. Da die entstehenden Pixel quadratisch sein sollen, ergibt sich eine Gridsize in vertikaler Ausrichtung von 378 Pixel. Alle finalen Karten besitzen demzufolge eine Auflösung von 458 x 378 Pixel.

6 Bewertungskriterien

Um schließlich Vergleiche auf der Grundlage quantitativer Kriterien mithilfe der realen Daten zu ermöglichen, wurde der mittlere quadratische Fehler (mean squared error – MSE) berechnet. Dieser basiert auf dem Bias und der Varianz. Alle drei Werte können in einem ersten Schritt pixelweise erhoben werden, weil die Ergebnisse in derselben Art und Weise diskretisiert vorliegen und die wahre Dichte (Adressdichte) bekannt ist. Da sich eine Auswertung auf Pixelbasis gut für eine grafische Darstellung eignet, aber einen Vergleich zwischen den Verfahren erschwert, werden in einem zweiten Schritt über alle Pixel gemittelte Werte bestimmt. Für die Choroplethenkarten und die naive Kerndichteschätzung fehlt durch die deterministische Berechnung der zufällige Anteil, weswegen es für diese Schätzungen keine Verfahrensvarianz gibt. Dennoch werden die entsprechenden Bewertungskriterien für die Choroplethenkarten und die naive Kerndichteschätzung analog ermittelt. Die Besonderheit dabei ist jedoch, dass die Verfahrensvarianz mit Null in die finale Kriterienberechnung eingeht.

Das Kernelheaping-Verfahren muss zur Bestimmung des Bias, der Varianz und des MSE mehrfach ausgeführt werden, da durch den stochastischen Anteil das Ergebnis jeder Ausführung leicht variiert. Mit der für die Analysen implementierten Erweiterung des Verfahrens um unabhängige Berechnungsketten können diese direkt parallel innerhalb des Kernelheaping-Pakets berechnet werden. Auf diese Weise entstehen mehrere voneinander unabhängige Markov-Ketten² je Pixel. Es wird eine initiale „Burnin“-Phase definiert, deren Iterationsergebnisse im späteren Verlauf verworfen werden. Danach folgen „Sample“-Iterationen, deren Dichteschätzungen in das Endergebnis einfließen. Um für eine Kette zu einem Ergebnis für ein Pixel zu gelangen, wird über alle Iterationen der Mittelwert gebildet. So entsteht für jede Markov-Kette eine separate Verteilung, welche einen Mittelwert und eine Varianz aufweist.

Dieses Prinzip lässt sich entsprechend auch über alle Ketten anwenden, um zu einem Gesamtmittelwert zu gelangen, der dem finalen Schätzergebnis je Pixel entspricht. Darüber hinaus ergibt sich eine Verfahrensvarianz, die durch die mehrfache Durchführung des Verfahrens entsteht und empirisch ermittelt werden kann.

7 Ergebnisse

In den Analysen wurden alle Verfahren für alle Aggregationsstufen berechnet und kartografisch dargestellt. Die folgenden Grafiken zeigen ausgewählte Karten, in denen immer Bezug auf die Dichtewerte der einzelnen Pixel genommen wird. Auf eine Umrechnung der „Dichtewerte je Pixel“ auf „Wahlberechtigte je Hektar“ wird an dieser Stelle verzichtet. Diese Skalierung dient einer intuitiven Interpretation, ist jedoch für Vergleichsuntersuchungen nicht notwendig.

² Markov-Ketten sind besondere stochastische Prozesse, bei denen der zukünftige Zustand eines Prozesses nur durch den aktuellen Zustand bedingt wird. Er wird nicht durch vergangene Zustände beeinflusst.

7.1 Choroplethenkarten

Die Berechnung von Choroplethenkarten basiert zunächst auf der einfachen Idee, dass Wahlberechtigte innerhalb eines Bezirks je nach Aggregationslevel aufsummiert werden und sich die regionale Zuordnung ausschließlich auf diesen Bezirk bezieht. So ergibt sich je Bezirk eine Gesamtzahl an Wahlberechtigten. Bei dieser Darstellung geht der exakte räumliche Bezug eines jeden Wahlberechtigten verloren. Es kann keiner Person eine genaue Geokoordinate zugeordnet werden, stattdessen nur noch ein Bezirk.

Um die Choroplethenkarte als nicht parametrische Dichteschätzmethode in die Vergleichsuntersuchungen aufnehmen zu können, muss nach der Aggregation der Wahlberechtigten eine Normierung stattfinden. So ist die Choroplethenkarte gegenüber der Originaldarstellung zwar anders skaliert, das Verhältnis zwischen den Daten bleibt jedoch erhalten. Damit hinterlässt eine Karte den gleichen optischen Eindruck.

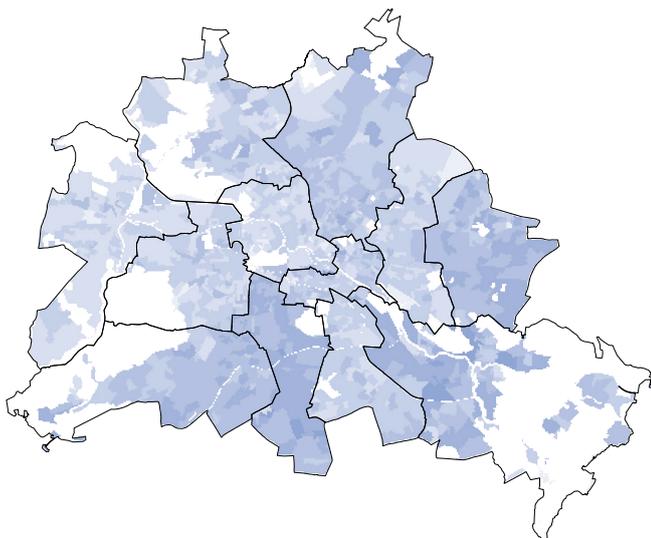
Es gibt einen weiteren Ansatz, die Choroplethenkarte zu normieren, um zu einer Dichte zu gelangen. Bei einer einfachen Normierung wird mithilfe der Gesamtzahl der Wahlberechtigten aller Bezirke skaliert. Um die inhomogenen Flächengrößen der verschiedenen Bezirke zu berücksichtigen, wird die Anzahl der Wahlberechtigten je Bezirk zunächst durch die jeweilige Bezirksgröße geteilt und anschließend die Gesamtheit auf Eins normiert. Auf diese Weise entsteht eine Dichte, welche einen Wahlberechtigtenanteil je Flächeneinheit (hier Pixel) zulässt. So werden die verschiedenen Bezirksgrößen berück-

sichtigt und es entsteht ein adäquater Vergleich zu Kerndichten. Beide Verfahren wurden in die Analysen einbezogen.

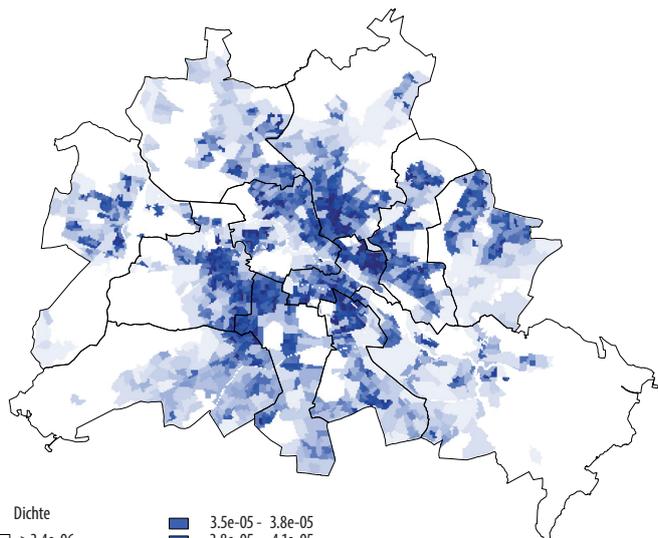
Die Schätzergebnisse der Choroplethenkarten zeigen die folgenden Abbildungen b und c. Bei Abbildung b handelt es sich (abgesehen von der Normierung) um eine klassische Darstellung der Verteilung von Wahlberechtigten auf Urnenwahlbezirke. Die simple Choroplethenkarte hat einen auffallend geringen Kontrast. Die Flächen werden insbesondere für die Wahlbezirke so geschnitten, dass ähnliche Wählerzahlen entstehen. Auf diese Weise wird der Bearbeitungsaufwand je Bezirk vergleichbar gehalten. Bei der Visualisierung der Wahlberechtigten entsteht in der Farbgebung daher ein homogenes Gesamtbild. Die Flächengröße bleibt dabei unberücksichtigt. Dies ist das Hauptproblem der simplen Choroplethenkarte.

Aus diesem Grund ist in Abbildung c mit der flächennormierten Choroplethenkarte eine Anpassung an diese Problematik dargestellt. Dieses Vorgehen liefert eine differenziertere Darstellung der Wahlberechtigtdichte. Dennoch gibt es keine Abstufungen innerhalb eines Bezirks, was insbesondere an den Bezirksgrenzen sichtbar wird. Im Ortsteil Moabit in Berlin-Mitte stößt beispielsweise eine sehr helle Fläche auf eine sehr dunkle Fläche. Dieser Effekt spiegelt die Realität nur unzureichend wider. Obwohl die Flächengröße einbezogen wird, sind die Informationen nur flächenbezogen abgebildet. Dem gegenüber stehen die punktbezogenen Schätzungen, die eine deutlich genauere Auflösung der Ergebnisse erlauben.

b | Kartografische Darstellung der Ergebnisse für die simple Choroplethenkarte



c | Kartografische Darstellung der Ergebnisse für die flächennormierte Choroplethenkarte



Dichte	
>3.4e-06	3.5e-05 - 3.8e-05
3.4e-06 - 6.3e-06	3.8e-05 - 4.1e-05
6.3e-06 - 9.2e-06	4.1e-05 - 4.4e-05
9.2e-06 - 1.2e-05	4.4e-05 - 4.7e-05
1.2e-05 - 1.5e-05	4.7e-05 - 5.0e-05
1.5e-05 - 1.8e-05	5.0e-05 - 5.3e-05
1.8e-05 - 2.1e-05	5.3e-05 - 5.6e-05
2.1e-05 - 2.4e-05	5.6e-05 - 5.8e-05
2.4e-05 - 2.7e-05	5.8e-05 - 6.1e-05
2.7e-05 - 2.9e-05	6.1e-05 - 6.4e-05
2.9e-05 - 3.2e-05	6.4e-05 - 6.7e-05
3.2e-05 - 3.5e-05	6.7e-05 - 7.0e-05
	>7.0e-05

Dargestellt ist die Verteilung von Wahlberechtigten auf Urnenwahlbezirke in Berlin zum Stichtag 31.12.2016.

7.2 Naive Kerndichteschätzung

Für die Berechnungen der naiven Kerndichteschätzungen wurde das in R implementierte Paket `ks` von Tarn Duong (Duong, 2017) verwendet. Dieses entspricht in seiner Funktionsweise den gängigen, in Geoinformationssystemen verfügbaren Plugins zur Kerndichteschätzung. Bei der naiven Kerndichteschätzung handelt es sich im Gegensatz zur Choroplethenkarte um ein punktbezogenes Verfahren zur Dichteschätzung.

Durch den Aggregationsprozess entsteht ein Informationsverlust von Geokoordinaten. Da diese jedoch für eine Schätzung benötigt werden, muss für das Verfahren initial jeweils ein beliebiger Punkt innerhalb eines Bezirks gewählt werden. Für die in den Analysen durchgeführten Berechnungen werden die geografischen Mittelpunkte der Polygone genutzt. Darauf wird eine einfache Glättung angewendet. Darüber hinaus bleibt eine wesentliche Entscheidungsmöglichkeit innerhalb des Verfahrens die Wahl der Bandbreite.

Für den finalen Verfahrensvergleich wurden zwei verschiedene Varianten für die Wahl der Bandbreite einbezogen. Zum einen wurde die Bandbreite optimal anhand des Datensatzes selbst gewählt. Dabei gilt die Bandbreite als optimal, welche im Vergleich zur wahren Adressdichte den minimalen MSE liefert. Dies ist nur möglich, wenn eine wahre Dichtekarte bekannt ist. Zum anderen – da die wahre Dichte in der Praxis im Allgemeinen nicht zur Verfügung steht – wurde der Plugin-Selektor nach Wand und Jones (Wand/Jones, 1994) des R-Pakets `ks` genutzt. Dieser soll eine optimale Bandbreite liefern. Dabei entste-

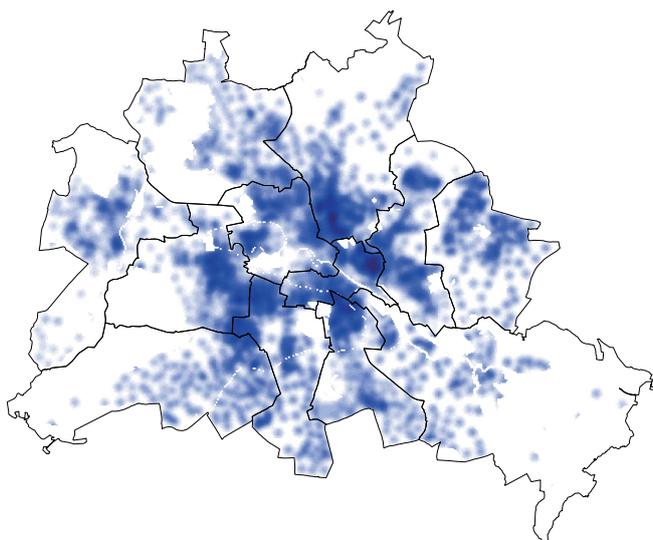
hen auf jedem Aggregationslevel zwei verschiedene optimale Bandbreiten und entsprechend auch zwei verschiedene Dichteschätzungen, die sich in ihrer Erscheinung und Qualität unterscheiden.

Grafik d zeigt die mittels der wahren Dichte berechnete naive Kerndichteschätzung mit optimaler Bandbreite. Dieses Ergebnis sieht optisch vielversprechend aus. Auffallend sind jedoch die vielen „Punktwolken“, welche durch die Wahl der geografischen Mittelpunkte entstehen. Die Bandbreite für die gesamte Karte wird einheitlich berechnet, daher ist sie in den größeren Randbezirken tendenziell zu klein gewählt, sodass Dichtepunkte nur im Zentrum der einzelnen Randbezirke auftreten.

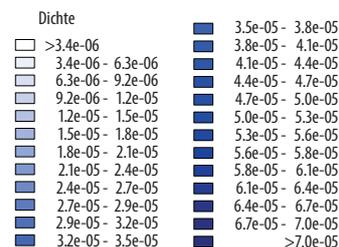
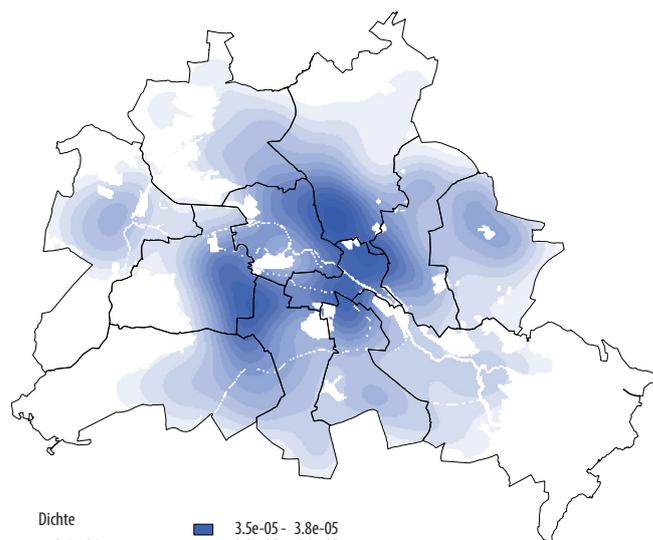
Eventuell wird die flächennormierte Choroplethenkarte des vorherigen Abschnitts als optisch ansprechender empfunden, da diese scharfe Kanten besitzt, an denen sich das Auge festhalten kann. Die harten Kanten, wie sie beispielsweise in Moabit zu finden sind, geben jedoch kein wirklichkeitsnahes Szenario wieder. Im Vergleich zur wahren Dichte wirkt die naive Kerndichteschätzung daher bereits realistischer. Es ist jedoch zu beachten, dass für Grafik d die wahre Dichte zur Bandbreitenwahl genutzt wurde. Diese steht in der Praxis nicht zur Verfügung.

Das Schätzergebnis unter Verwendung eines Plugin-Verfahrens zur Bandbreitenwahl ist in Grafik e dargestellt. Hier fällt auf, dass viele Detailinformationen verloren gehen. Dies ist selbst auf Ebene der Urnenwahlbezirke ersichtlich, obwohl die Daten mit 1779 Messwerten vergleichsweise genau erhoben wurden. Bei der Anwendung der naiven Kerndichteschätzung auf Stichproben für Rückschlüsse auf eine

d | Kartografische Darstellung für die naive Kerndichteschätzung mit optimaler Bandbreite



e | Kartografische Darstellung für die naive Kerndichteschätzung mit Plugin-Bandbreite



Dargestellt ist die Verteilung von Wahlberechtigten auf Urnenwahlbezirke in Berlin zum Stichtag 31.12.2016.

Grundgesamtheit wäre dieses Ergebnis plausibel, da die Grundstruktur der Verteilung prinzipiell gut wiedergegeben wird. In diesem Anwendungsfall jedoch entsteht durch die großflächigen Außenbezirke eine zu große Bandbreite, was zu dem starken Glättungseffekt führt.

7.3 Kernelheaping-Verfahren

Für die Berechnungen des Kernelheaping-Verfahrens wurde aus dem Paket kernelheaping von Marcus Groß (Groß, 2017) die Methode dshapebiv verwendet. Der Algorithmus durchläuft die folgenden Schritte:

Initialisierung

1. Berechnung der Anzahl der Pixel und ihrer zugehörigen Geokoordinaten basierend auf der Gridsize, um ein Raster zu erhalten, welches alle Bezirke der Aggregationsstufe abdeckt (Bounding-Box)
2. Berechnung einer naiven Kerndichteschätzung f_0 für alle Pixel mit den auf die jeweiligen Mittelpunkte aggregierten Daten mit einer initialen Bandbreite, basierend auf der Größe der Bounding-Box und der Anzahl der Bezirke

Iteration

Führe Schritte 3 bis 6 für eine festgelegte Anzahl von Iterationen aus: $t = 1, \dots, i$

3. Zufällige Ziehung von j_R Pixeln mit Zurücklegen für jeden Bezirk R basierend auf der Dichte f_t . Dabei entspricht die Anzahl der gezogenen Pixel dem zugehörigen Absolutwert der initial aggregierten Daten.

4. Berechnung der „optimalen“ Bandbreite nach Wand und Jones auf Basis der gezogenen Pixel
5. Berechnung einer naiven Kerndichteschätzung f_{t+1} auf Basis der gezogenen Pixel mit „optimaler“ Bandbreite
6. Nullsetzung der Dichtewerte aller Pixel außerhalb von Polygonen und ausmaskierten Pixel

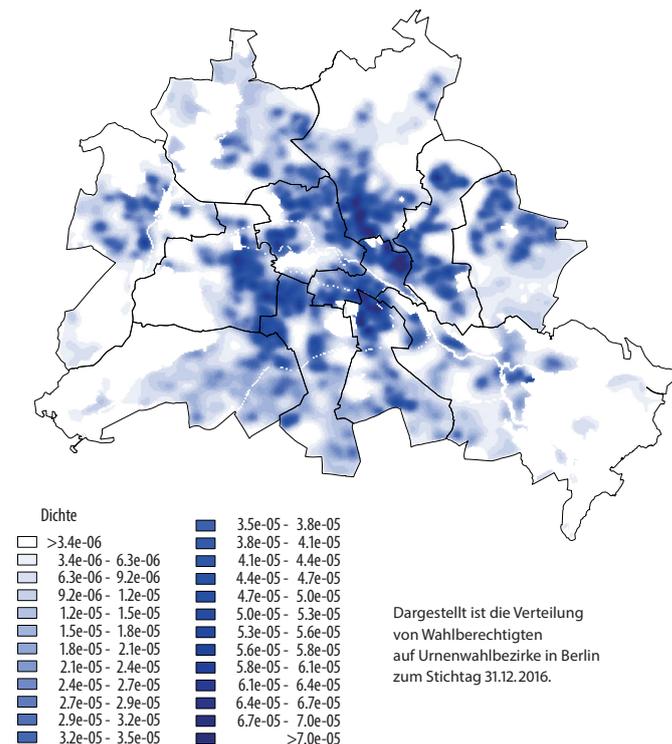
Resultat

7. Berechnung eines Dichte-Mittelwertes für alle Pixel über die Dichten f_b, \dots, f_{i-1}

Zwei wesentliche Entscheidungskriterien innerhalb des Verfahrens sind die Anzahl an Iterationen i und die Länge der Burnin-Phase b , welche das Verfahren durchlaufen soll. Um den Einfluss dieser Parameter auf das Endergebnis abschätzen zu können, wurden die Schätzungen für verschiedene Iterationszahlen durchgeführt und verglichen. So kann die Differenz zur wahren Adressdichte abhängig von der Iterationszahl beziehungsweise von der Laufzeit des Verfahrens bestimmt werden.

Ein Ergebnis des Kernelheaping-Verfahrens, als weiteres punktbezogenes Schätzverfahren, visualisiert Grafik f ebenfalls für die Ebene der Urnenwahlbezirke. Es ist gut sichtbar, dass unter Verwendung des Kernelheaping-Verfahrens deutlich weniger Details verloren gehen als bei der naiven Kerndichteschätzung, insbesondere unter Verwendung der Plugin-Methode. Zudem gelingt es dem Verfahren besser, die flächenbezogene Inhomogenität der einzelnen Bezirke auszugleichen. Dadurch werden sowohl die kleinen Bezirke im Zentrum als auch die großen Flächen am Stadtrand gut geschätzt. Beispielsweise sticht das Märkische Viertel im Bezirk

f | Kartografische Darstellung für die Kerndichteschätzung mit dem Kernelheaping-Verfahren



Reinickendorf im Norden Berlins als eine Art „Fragezeichen“ in der wahren Dichte auf Urnenwahlbezirksebene gut erkennbar hervor (siehe Grafik a). In den naiven Kerndichtekarten ist dieses Detail nicht gleichermaßen gut ausgeprägt.

An dieser Stelle sei darauf hingewiesen, dass die bessere Detailwiedergabe ohne eine explizite Angabe einer Bandbreite erreicht wird. Dies ist ein wesentlicher Vorteil gegenüber der naiven Kerndichteschätzung. Zudem konnte in der Arbeit gezeigt werden, dass die Wahl der Anzahl an Iterationen nur eine geringe Auswirkung auf das Ergebnis hat.

Grafik g zeigt, dass mit zunehmender Iterationszahl der MSE weitgehend stabil bleibt, während gleichzeitig die Varianz abnimmt. Allerdings spielt die Varianz im Vergleich zu den Größenordnungen des Bias kaum eine Rolle.

8 Fazit

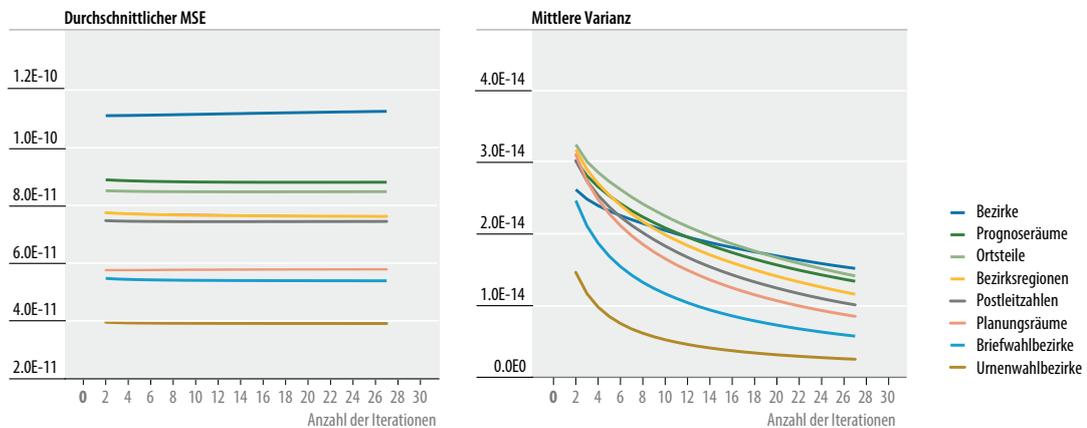
Wie oben bereits beschrieben, wird mit der Berechnung des MSE je Pixel (MSE-Karten) nur ein pixelweise gültiges Maß für die Qualität einer Schätzung wiedergegeben. Auf dieser Grundlage kann jedoch nicht die ganzheitliche Güte der Schätzergebnisse objektiv beurteilt werden. Daher wird für jede Karte der durchschnittliche MSE über alle Pixel berechnet. So ergibt sich ein globales Maß für die Qualität einer

Schätzung, mit dessen Hilfe die Verfahren verglichen werden können.

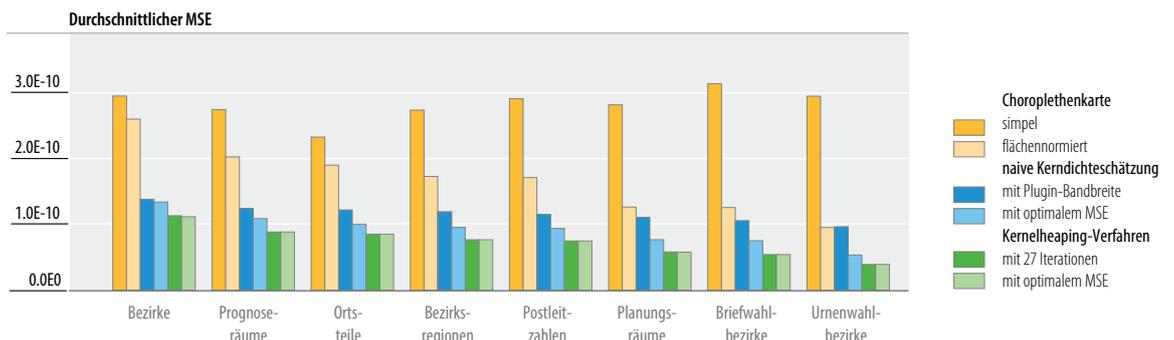
Dieser durchschnittliche MSE über alle Pixel wird in Grafik h noch einmal – abhängig vom Aggregationslevel – für alle betrachteten Verfahren in einem Diagramm gegenübergestellt. Dabei sind die Verfahren farblich gruppiert. Da die Skala der Y-Achse auf null gesetzt ist, lässt sich das Verhältnis der MSE-Werte zwischen den Verfahren ablesen. Es zeigt sich beispielsweise, dass das Kernelheaping-Verfahren gegenüber der naiven Kerndichte auf Ebene der Urnenwahlbezirke einen um etwa 25% geringeren durchschnittlichen MSE aufweist. Ebenfalls ersichtlich ist, dass die flächennormierte Choroplethenkarte auf der Aggregationsstufe der Urnenwahlbezirke nur etwa ein Drittel des MSE der simplen Choroplethenkarte besitzt.

Bei den einfachen Choroplethenkarten fällt auf, dass bei der simplen Schätzung eine genauere Aggregation nicht gleichermaßen zu einer besseren Schätzung führt. Intuitiv betrachtet, sollte sich mit detaillierteren Daten auch eine Besserung der Schätzergebnisse einstellen. Diese besitzen jedoch mit einer besseren Datenerhebung keinen messbaren Informationsgewinn. Die flächennormierte Choroplethenkarte hingegen weist solche Verbesserungen auf.

g | Direkter Vergleich des mittleren quadratischen Fehlers (MSE) und der mittleren Varianz für das Kernelheaping-Verfahren, abhängig von der Iterationszahl für acht verschiedene Aggregationslevel



h | Zusammenfassender Vergleich der mittleren quadratischen Fehler (MSE) aller betrachteten Verfahren für alle Aggregationslevel



Für die naive Kerndichteschätzung werden mit optimaler Bandbreite zwar zunächst solide Ergebnisse erzielt, diese beruhen jedoch auf der Kenntnis der wahren Dichte, welche in der Praxis nicht zur Verfügung steht. Es ist ebenfalls erkennbar, dass sich mit einem höheren Detailgrad der Aggregationslevel auch ein geringerer durchschnittlicher MSE einstellt. Wird die naive Kerndichte ohne Kenntnis der wahren Dichte berechnet und ein Plugin-Bandbreitenselektor genutzt, können die Ergebnisse gegenüber dem Kernelheaping-Verfahren nicht überzeugen. Auf Ebene der Urnenwahlbezirke liegt die naive Kerndichte mit der flächennormierten Choroplethenkarte gleichauf.

Insgesamt lässt sich feststellen, dass das Kernelheaping-Verfahren über alle Aggregationsstufen hinweg das beste Ergebnis in Hinblick auf den minimalen durchschnittlichen MSE liefert. Die Iterationszahl spielt dabei eine untergeordnete Rolle für die Qualität des Schätzergebnisses. Demzufolge zeigt sich, dass das Kernelheaping-Verfahren auf Realdaten praxistauglich und robust einsetzbar ist.

Neben der quantitativen Bewertung ist auch optisch eine detailliertere Dichteschätzung erkennbar. Durch das punktbasierte Schätzverfahren werden einzelnen Pixeln des Bildes separate Dichtewerte zugeordnet. Dadurch gelingt es mit dem Kernelheaping-Verfahren, feine Strukturen in stark besiedelten Gebieten herauszuarbeiten, während gleichzeitig in großflächigen Arealen homogen geschätzt wird. Die einfache Kerndichteschätzung, als weiteres punktbasiertes Schätzverfahren, konnte hier keine gleichwertigen Ergebnisse erzielen, da mit einer fix gewählten Kernelgröße nicht adäquat auf die inhomogenen Aggregationen eingegangen werden kann. Die Choroplethenkarten geben die wahre Dichte am schlechtesten wieder, haben jedoch durch ihre Eigenschaft, insbesondere auf detaillierten Aggregationsleveln optisch ansprechende, scharfe Konturen zu bilden, ihre Vorteile.

Das Kernelheaping-Verfahren zeigt sich robust hinsichtlich der Parameterwahl. Auch wenn für die hier verwendete Adressdichte optimale Parameter ermittelt werden konnten, liefert das Kernelheaping auch bei vordefinierten, pauschalen Parametern ähnlich gute Resultate. Dies ist ein weiteres, wesentliches Unterscheidungsmerkmal gegenüber den einfachen Kerndichteschätzungen, welche mit den in dieser Arbeit verwendeten Plugin-Bandbreiten teilweise unbrauchbare Ergebnisse lieferten.

Ein Nachteil des Kernelheapings sind die hohen Ressourcenanforderungen. Sie sind im Vergleich zu den Choroplethen- und einfachen Kerndichtekarten sehr viel umfangreicher. Ein weiterer potenzieller Nachteil ist die nicht deterministische Berechnung. Durch den stochastischen Anteil entsteht eine gewisse Unsicherheit des Ergebnisses. Es konnte jedoch gezeigt werden, dass diese Varianz für die betrachteten Adressdaten im Vergleich zum Bias um den Faktor 103 geringer ist und damit in der Praxis keine Rolle spielt. Im Hinblick auf Geheimhaltungsaspekte könnte sich diese Unsicherheit wiederum als Vorteil herauskristalisieren. Damit stellt das Kernelheaping-Verfahren eine praxistaugliche Alternative für die Dichteschätzung für Wahldaten dar.



Kerstin Erfurth hat Statistik (M. Sc.) an der Humboldt-Universität zu Berlin studiert und das Zertifikat *European Master in Official Statistics (EMOS)* erlangt. Im Rahmen des EMOS-Programms absolvierte sie einen Forschungsaufenthalt im Amt für Statistik Berlin-Brandenburg. Aus

dieser Kooperation ist ihre Masterarbeit zum Thema „Gütebeurteilung und Einsatz simulierter Geokoordinaten bei der regionalen Analyse zur Bundestagswahl 2017“ entstanden, für die sie mit dem Gerhard-Fürst-Preis 2019 ausgezeichnet wurde. Sie ist Referentin in der Stabsstelle *Statistische Methoden und Grundsatzfragen* des Amtes für Statistik Berlin-Brandenburg.

Literaturverzeichnis

- Duong, Tarn, 2017: Kernel Smoothing. <https://cran.r-project.org/web/packages/ks/index.html>, R package Version 2.0.
- Fahrmeir, Ludwig; Hamerle, Alfred; Tutz, Gerhard, 1996: *Multivariate statistische Verfahren*. 2. Auflage. Berlin/New York.
- Groß, Marcus; Rendtel, Ulrich, 2016: Kernel Density Estimation for Heaped Data. In: *Journal of Survey Statistics and Methodology*. Jahrgang 4. Ausgabe 3/2016, S. 339 ff.
- Groß, Marcus, 2017: Kernelheaping: Kernel Density Estimation for Heaped and Rounded Data. <https://cran.r-project.org/web/packages/Kernelheaping/index.html>, R package Version 2.0.
- Groß, Marcus, 2016: *Messfehlermodelle für die Survey-Statistik und die Wirtschaftsarchäologie*. Dissertation. Freie Universität Berlin. [Zugriff am 9. März 2020.] Verfügbar unter: <https://refubium.fu-berlin.de/handle/fub188/9385>
- Groß, Marcus; Rendtel, Ulrich; Schmid, Timo; Schmon, Sebastian; Tzavidis, Nikos, 2017: Estimating the density of ethnic minorities and aged people in Berlin: multivariate kernel density estimation applied to sensitive georeferenced administrative data protected via measurement error. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. Jahrgang 180. Ausgabe 1/2017, S. 161 ff.
- Härdle, Wolfgang; Müller, Marlene; Sperlich, Stefan; Werwatz, Axel, 2004: *Nonparametric and Semiparametric Models*. Springer Series in Statistics. Berlin/Heidelberg.
- Silverman, Bernard W., 1986: *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability 26. London.
- Stadtportal, BerlinOnline. Berlin Open Data. 2017: Geometrien und Wahl- sowie Wahlstrukturdaten stehen unter der offenen Lizenz CC-BY und sind über das Open Data Portal des Landes Berlin verfügbar. [Zugriff am 9. März 2020.] Verfügbar unter: <https://daten.berlin.de>
- Wand, Matt P.; Jones, Chris, 1994: *Multivariate plug-in bandwidth selection*. In: *Computational Statistics*. Jahrgang 9. Ausgabe 2/1994, S. 97 ff.

Georeferenzierung

Geokodierung mit dem RBS-Geocoder

von **Renee Lin**

Der Geoservice des Amtes für Statistik Berlin-Brandenburg (AfS) pflegt die Geodaten, das heißt die Koordinaten für Berliner Adressen, Straßen, Blöcke und verschiedene Raumbezüge wie Wahlbezirke oder Einschulbereiche. Der Berliner Straßenschlüssel wird ebenso vom AfS vergeben. Im Fortschreibungsverfahren werden Adressen sowie deren Zuordnungen zu Blockseiten, Blöcken, Straßenabschnitten und Straßen in das Regionale Bezugssystem (RBS) aufgenommen. Basierend auf diesen Daten wurde im AfS ein Geokodierungsverfahren implementiert und ein RBS-Geocoder mit einer grafischen Benutzeroberfläche entwickelt.

Geokodierung

Raumbezogene Analysen haben in den letzten Jahren mehr an Bedeutung gewonnen und werden in fast allen Bereichen des öffentlichen Lebens, der Wirtschaft und in vielen Wissenschaftsdisziplinen benötigt. Voraussetzung für eine solche Analyse ist die Transformation von Sachdaten in ein räumliches Bezugssystem. Dieses Verfahren der Geokodierung erfolgt mit einer Software, dem sogenannten Geocoder. Dieser besteht aus miteinander verzahnten Operationen, Algorithmen und Datenquellen (Goldberg et al. 2008), die zwischen eingegebener Adresse und georeferenzierten Daten vermitteln (Schulte et al. 2010). Im Geokodierungsprozess wird eine Anschrift in eine Position auf der Erdoberfläche umgewandelt und dieser dann raumbezogene Informationen, zum Beispiel in welchem Ortsteil die eingegebene Anschrift liegt, zugewiesen. Der komplette Ablauf besteht aus mehreren Schritten, bei denen innerhalb einer Referenzdatenbank die zutreffendste Koordinate zur Adresse gesucht wird. Die resultierenden Datensätze werden als geografische Objekte mit Attributen ausgegeben, die zu Zwecken der Zuordnung oder für räumliche Analysen verwendet werden können. In der amtlichen Statistik werden nachgeokodierte Daten zur Bearbeitung und Auswertung sowie zur Ergebnisdarstellung und Visualisierung verwendet.

Referenzdaten

Referenzdaten spielen eine wichtige Rolle im Geokodierungsprozess. Der Umfang und die Aktualität der Referenztabellen beeinflussen die Ergebnisqualität erheblich. Referenztabellen im RBS-Geocoder bestehen aus den folgenden drei Teilen:

- aktuelle Adresse
- historische Adresse
- Objekt

Für die interne Nutzung sowie für die Bereitstellung der Daten, zum Beispiel dem Landesamt für Bürger- und Ordnungsangelegenheiten (LABO) oder der Senatsverwaltung für Stadtentwicklung und Wohnen Berlin, wird monatlich eine Datensicherung und Koordinatentransformation im RBS durchgeführt. Für historische Adressen werden die gelöschten Adressen von der Datenbank ausgelesen. Die Objekttable ist eine Sammlung von Bahnhöfen, Plätzen und Flughäfen ohne Hausnummerzuordnung. X/Y-Koordinaten des Objekts sind Mittelpunkte der Straßen oder Objekte; eventuell ist keine Zuordnung zu einigen Gebieten (zum Beispiel Block) möglich. Auch die Referenzdaten für den RBS-Geocoder werden monatlich aktualisiert.

Für die Referenztabellen werden nicht nur Straßennamen, Hausnummer, Postleitzahl und Koordinaten aufgenommen, sondern auch die Zuordnungen der verschiedenen Raumbezüge integriert, wie Planungsraum, Verkehrszelle, LAEA-Gitterzelle¹ usw. (Tabelle 1).

RBS-Geocoder

Der RBS-Geocoder ist eine lokale Java-Anwendung im AfS, die die Zuordnung von Koordinaten und Raumbezügen zu einer Liste von Adressen erlaubt. Abbildung a zeigt die Benutzeroberfläche des RBS-Geocoders. Der komplette Ablauf der Geokodierung besteht von der Eingabe bis zur Ausgabe aus vier Teilen (Abbildung b). Zu Beginn wird eine Adressenliste mit den notwendigen Angaben für

¹ LAEA (Lambert azimuthal equal-area) ist eine Kartenprojektion, in der die gesamte Kugeloberfläche wiedergegeben werden kann. LAEA-Gitterzellen basieren auf dieser flächen-

treuen Azimutalprojektion. Als europaweit einheitliche geografische Gitter werden sie in quadratischen Zellgrößen vom 100 m bereitgestellt. Sie dient der Darstellung und Analyse statistischer Sachverhalte.

Straßenname, Postleitzahl und Hausnummer (inkl. Hausnummerzusatz) aus einer csv-Datei hochgeladen und der Stand der Referenzdaten sowie eine Ausgabevorlage erfasst. Danach werden Sonderzeichen und Abkürzungen in den Adressen normalisiert. In diesem Vorgang werden fehlende Datensätze von der Eingabedatei ausgefiltert. Daraufhin folgt der dritte Schritt, der den Algorithmus der Adresszuordnung umfasst. Alle Schritte des

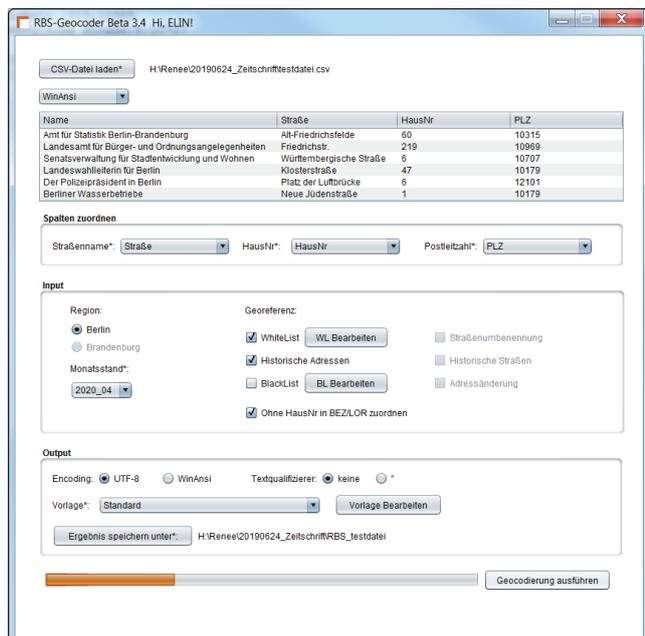
Prozesses der Adresszuordnung werden in folgender Reihenfolge ausgeführt:

1 | Übersicht über die Berliner Raumbezüge des RBS-Geocoders

Attribut	Beschreibung
AGB	Amtsgerichtsbezirk
AKZ	Aktive Zentren
ARB	Arbeitsamt
ATY	Adresstyp
AWK	Abgeordnetenhauswahlkreise
BEZ	Bezirk
BEZA	Bezirk (alt)
BLK	Blocknummer
BLSFolgeNr	Blockseitefolgenummer
BWB	Briefwahlbezirk
BWK	Bundestagswahlkreise
EGM	Erhaltungsgebiete Milieuschutz
EGS	(EGS) Erhaltungsgebiet städtebauliche Eigenart
ESA	Einschulbereich aktuelles Schuljahr
ESP	Einschulbereich
EUF	EU-Fördergebiet 2000–2006
EWK	Einwohnerkategorie
FIN	Finanzamt
GAF	Gemeinschaftsaufgabe Fördergebiet 2007–2013
GeoGrid_	Gitter-ID im Koordinatensystem LAEA Europe
100m_ID	(EPSG: 3035)
GRS	Großsiedlung
GRW	Fördergebiet Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ 2014–2020
KBA	Teilverkehrszellen (Kraftfahrt-Bundesamt)
LOR	Lebensweltlich Orientierte Räume (Planungsraum)
OT	Ortsteil
OTA	Ortsteil alt
OWT	Ost-West politische Teilung
PKB	Polizeikontaktbereich
QM	Quartiersmanagement neu
QMA	Quartiersmanagement (historisch)
SAN	Sanierungsgebiet
SDS	städtebaulicher Denkmalschutz
STG	Statistisches Gebiet
STRNr	Straßennummer
STU	Stadtumbau
SVE	Spielplatzversorgungseinheit
TVZ	Teilverkehrszelle
TVZA	Teilverkehrszellen bis 12/2012
UR2	EU-Gemeinschaftsinitiative (Urban II)
URA	Senstadt Monitoring (UrbanAudit)
UWB	Urnenwahlbezirk (Stimmbezirk)
VBWB	Volksentscheid Briefwahlbezirk
VKZ	Verkehrszelle
VKZA	Verkehrszellen bis 12/2012
VUWB	Volksentscheid Urnenwahlbezirk (Abstimmbezirke)
WOL	Wohnlage
WSG	Wasserschutzgebiet
X_25833	X-Koordinate im ETRS89/UTM Zone 33N (EPSG: 25833)
X_3035	X-Koordinate im LAEA (EPSG: 3035)
Y_25833	Y-Koordinate im ETRS89/UTM Zone 33N (EPSG: 25833)
Y_3035	Y-Koordinate im LAEA (EPSG: 3035)

1. Die gesuchte Adresse wird über eine Referenzliste (Blacklist) zu einer benutzerdefinierten Adresse zwingend erforderlich zugeordnet. Das heißt, dass Nutzende ein bestimmtes Ergebnis steuern können. Die Liste kann zukünftig von den Nutzenden selbst gepflegt werden.
2. Hier wird eine gesuchte Adresse mit Hausnummernbereich erkannt und eine eindeutige Hausnummer berechnet. Ein Treffer wird für die erste Hausnummer bzw. die zweite Hausnummer erzielt. Der Algorithmus unterscheidet aufgrund der Berliner Nummernvergabe mit dem gebräuchlichen Zickzack-System (Jacobs 2016) zwischen geraden, ungeraden und fortlaufenden Hausnummern. Bei geraden/ungeraden Bereichen zählt der Algorithmus in Zwischenschritten, bei fortlaufenden Bereichen in einem Schritt. Hausnummerzusätze (einzelner Buchstabe) werden nur hochgezählt, wenn in beiden Bereichen Hausnummerzusätze existieren und sie der Reihenfolge des Alphabets entsprechen.
3. Die gesuchte Adresse wird einer aktuellen RBS-Adresse zugeordnet.
4. Die gesuchte Adresse wird einer historischen RBS-Adresse zugeordnet.
5. Die gesuchte Adresse wird einem Bahnhof, einem Platz oder einem Flughafen zugeordnet.
6. Die gesuchte Adresse wird über eine Referenzliste (Whitelist) zugeordnet. In der Whitelist werden Straßennamen oder Adressen mit typischen Rechtschreibfehlern, Abkürzungen, Straßenumbenennungen usw. oder Objektbezeichnungen (zum Beispiel Rotes Rathaus) zu aktuellen RBS-Adressen zugeordnet. Die Liste ist nicht vollständig und kann von den Nutzenden selbst gepflegt werden.

a | Benutzeroberfläche des RBS-Geocoders



II-Daten, Baufertigstellungen, Unternehmensregisterdaten, Wahllokale und Spielhallen. Etwa 270 000 Firmenadressen aus dem Berliner Unternehmensregister werden regelmäßig über den RBS-Geocoder geokodiert und danach für die Qualitätsprüfung und statistische Analyse verwendet. Das Ergebnis liegt im Durchschnitt bei 99,8 % gefundenen Adressen. Die errechnete mittlere Laufzeit beträgt 12,5 Minuten. Abbildung c zeigt, dass Qualität und Plausibilität der Geokodierung im Zeitraum von November 2018 bis Oktober 2019 nahezu konstant geblieben sind. Auch für große Datenmengen, wie die Kfz-Daten mit 1,5 Mill. Adressen, hat sich der RBS-Geocoder bewährt und liefert zuverlässige Ergebnisse.

Fazit

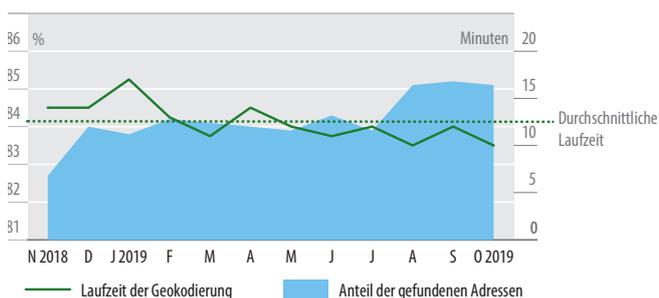
Die Geokodierung ist ein viel diskutiertes Thema. Sie gewinnt nicht nur in der Forschung zunehmend an Aufmerksamkeit, sondern wird auch in § 14 des E-Government-Gesetzes² ausdrücklich gefordert. Der Geoservice des AfS pflegt derzeit fast 400 000 Berliner Adressen sowie umfangreiche historische Daten. Auf Basis dieser vollständigen und amtlichen Referenzdaten wurde der RBS-Geocoder im AfS für die Geokodierung entwickelt. Folgende Kernpunkte und Funktionen sind im RBS-Geocoder berücksichtigt:

1. Vielfältige Informationen zu den Anschriften: Das Ergebnis liefert nicht nur die zu den Anschriften zugehörigen Geokoordinaten, sondern zusätzlich die raumbezogenen Informationen (zum Beispiel Block, LOR, LAEA-Gitterzelle usw.).
2. Berücksichtigung der historischen Adresse und von Objekten: Die gesuchte Adresse kann einer historischen RBS-Adresse sowie einem Bahnhof, Flughafen oder Platz zugeordnet werden.

3. Unterschiedliche Referenzdatenstände stehen zur Verfügung: Die Monatsstände werden rückwirkend bis Juni 2018 angeboten und regelmäßig am Monatsanfang für den vergangenen Monat aktualisiert.
 4. Nachbearbeitung über Whitelist und Blacklist: Die gesuchte Adresse kann über individuelle Referenzlisten einer aktuellen RBS-Adresse zugeordnet werden. Damit haben Nutzende die Möglichkeit, die Ergebnisse zu verbessern.
 5. Betrachtung der Adressenbesonderheit: Für die gebräuchliche Zickzack-Hausnummerierung in Berlin unterscheidet der Algorithmus zwischen geraden, ungeraden und fortlaufenden Hausnummern.
 6. Nutzerorientierung: Über eine Nutzerverwaltung werden Whitelist, Blacklist und Ausgabevorlage Fachbereichen zugeordnet und gespeichert.
 7. Nachvollziehbares Ergebnisprotokoll: Alle Eingabeparameter, Daten der Nutzenden, Ausführungsdatum und -uhrzeit, die Gesamtlaufzeit, Ausgabepfad sowie die Ergebnisse mit einer statistischen Information werden protokolliert.
- Bisher wurde ein Geokodierungsprozess für alle Berliner Adressen aufgebaut und für die praktische Anwendung der RBS-Geocoder realisiert. Als nächste Schritte sind eine Visualisierungsfunktion zur Darstellung und Verortung, eine Weiterentwicklung zur Web-Anwendung und die Erweiterung auf Brandenburger Adressen geplant. Anschließend kann der RBS-Geocoder auch externen Kunden zur Verfügung gestellt werden.

Renee Lin ist Referentin im Geoservice des Amtes für Statistik Berlin-Brandenburg.

c | Anteil der gefundenen Adressen und Laufzeit der Geokodierung im RBS-Geocoder



Literaturverzeichnis

- Goldberg, D. W.; Swift, J. N.; Wilson, J. P. (2008): Geocoding Best Practices: Reference Data, Input Data and Feature Matching. Researchgate. Online unter: https://www.researchgate.net/publication/239919647_Geocoding_Best_Practices_Reference_Data_Input_Data_and_Feature_Matching [Abgerufen am 26.02.2020].
- Jacobs, St. (2016): Wie das Chaos bei den Berliner Hausnummern entstand, Online unter: <https://www.tagesspiegel.de/berlin/datenanalyse-wie-das-chaos-bei-den-berliner-hausnummern-entstand/13426854.html> [Abgerufen am 26.02.2020].
- McDonald, Y. J.; Schwind, M.; Goldberg, D. W.; Lampléy, A.; Wheeler, C. M. (2017): An analysis of the process and results of manual geocode correction. *Geospat Health*, 12(1): 526. doi:10.4081/gh.2017.526.
- Schulte, B.; Lippmann, F.; Schweikart, J. (2010): Geokodierung mit Webkartendiensten – Möglichkeiten, Unterschiede und Grenzen. In: Strobl, J., Blaschke, T., Griesebner, G. (Hrsg): *Angewandte Geoinformatik 2010*. Beiträge zum 21. AGIT-Symposium Salzburg. Heidelberg: Wichmann, 773–778.
- Xu, S.; Flexner, S.; Carvalho, V. (2012): Geocoding Billions of Addresses: Toward a Spatial Record Linkage System with Big Data. In conjunction with the seventh International Conference on Geographic Information Science 2012 (GIScience 2012).

² Gesetz zur Förderung der elektronischen Verwaltung (E-Government-Gesetz – E-GovG) vom 25. Juli 2013 (BGBl. I S. 2749), das zuletzt durch Artikel 15 des Gesetzes vom 20. November 2019 (BGBl. I S. 1626) geändert worden ist.

Georeferenzierung

▣ Raumbezogene Analyse zur Wohnraumentwicklung in Berlin und Brandenburg

von **Nicole Jurisch**

Im Zuge der Umsetzung der Europarichtlinie INSPIRE werden immer mehr amtliche Statistiken geokodiert und somit für raumbezogene Analysen nutzbar gemacht. Auch für die Statistik der Baufertigstellungen liegen nun geokodierte Daten vor, die eine kleinräumigere Darstellung der Wohnraumentwicklung ermöglichen. Anhand von Rasterzellen in unterschiedlicher Auflösung kann veranschaulicht werden, wo und wie im Jahr 2018 in den einzelnen Regionen Berlins und Brandenburgs gebaut wurde und welche Art von Wohnungen vorwiegend entstanden.

Laut Statistischem Bundesamt lebten 2018 in Deutschland ungefähr 7,4% der Bevölkerung in überbelegten Wohnungen.[1] Das bedeutet, dass mehr Personen in einer Wohnung leben, als tatsächlich Räume zur Verfügung stehen. Knapper Wohnraum und steigende Mieten sind Gründe, warum Menschen in solch beengten Verhältnissen wohnen. Diesem Missverhältnis soll in Berlin und Brandenburg unter anderem mit vermehrtem Wohnungsbau begegnet werden. So entstanden im Jahr 2018 in der gesamten Metropolregion 7960 (Berlin: 2101, Brandenburg: 5859) neue Wohngebäude und 27934 (Berlin: 16551, Brandenburg: 11383) neue Wohnungen.

Doch Familien, Singles und Paare legen je nach beruflicher und familiärer Situation unterschiedliche Maßstäbe bei der Wohnungssuche an. Gleichzeitig bieten die bereits stark verdichteten Stadtzentren nur wenig Raum für neue Wohngebäude und auch in verschiedenen Regionen Brandenburgs, insbesondere im Berliner Umland, steigen die Preise für Immobilien stetig.

Die verschiedenen Herausforderungen in der Stadtentwicklung erfordern auf lokaler Ebene recht unterschiedliche Lösungsansätze und führen zu einer sehr vielfältigen Entwicklung der Bautätigkeit. Mit der räumlichen Analyse zur Statistik der Baufertigstellung werden die spezifischen Muster der Wohnraumentwicklung in den urbanen sowie in den ländlichen Regionen sichtbar.

Wo wurde gebaut?

Neue Wohnungen entstanden 2018 in nahezu allen Bezirken Berlins und den Gemeinden Brandenburgs (Abbildungen a, b). Insbesondere in den Bezirken Mitte (2404) und Treptow-Köpenick (2526) wurden viele neue Wohnungen gebaut. Im Land Brandenburg verzeichneten vor allem die Gemeinden im Berliner Umland eine stärkere Bautätigkeit als die Regionen in größerer Entfernung zu den Ballungszentren. So fanden sich mit 7339 etwa 65% aller in Brandenburg neu gebauten Wohnungen im Ballungszentrum rund um die Bundeshauptstadt und mit 2109 immerhin 19% allein in der Landeshauptstadt Potsdam, gefolgt von Cottbus mit 509 (knapp 5%) neuen Wohnungen.

Für viele Familien sind Wohnlage und Preis einer Mietwohnung oder Immobilie ein wichtiges Auswahlkriterium. Insbesondere für die Mietpreise in Berlin ist der Berliner Mietspiegel ein wichtiger Wegweiser. Dieser orientiert sich an der ermittelten Wohnlage, welche die Wertigkeit der Lagegegebenheiten im Vergleich zu anderen Adressen im Stadtgebiet widerspiegelt. [2] Sie wird für jede Aktualisierung des Mietspiegels neu ermittelt. Abbildung c zeigt die überwiegende Wohnlage der 2018 in Berlin neu entstandenen Wohnungen und ihre räumliche Verteilung im Ein-Kilometer-Raster. Wohnungen in guter Wohnlage entstanden vor allem im Südwesten der Stadt in den Bezirken Steglitz-Zehlendorf und Charlottenburg-Wilmersdorf. Insgesamt befan-

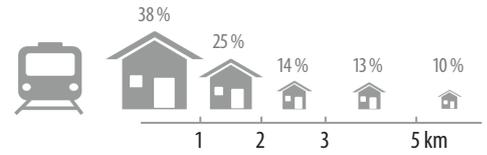
den sich jedoch nur 12,4% aller neuen Wohnungen in guter Wohnlage, während 36,6% in mittlerer und 51,0% in einfacher Wohnlage gebaut wurden (Abbildung c).

Im Land Brandenburg sind die Nähe zu öffentlichen Verkehrsmitteln und die Anbindung an die Ballungszentren entscheidend für die Entwicklung von Wohnraum. Insbesondere für die vielen Berufspendlerinnen und -pendler im sogenannten „Speckgürtel“ Berlins ist eine schnelle und unkomplizierte Verkehrsverbindung in das Stadtzentrum mit S-Bahn oder Regionalexpress ein entscheidender Faktor für die Wahl des Wohnorts. So entstanden 2018 in Brandenburg 38,4% aller Wohnungen im Umkreis

von weniger als einem Kilometer um einen Bahnhof, weitere rund 38,9% der neu gebauten Wohnungen liegen ein bis drei Kilometer von der nächsten Bahnstation entfernt.

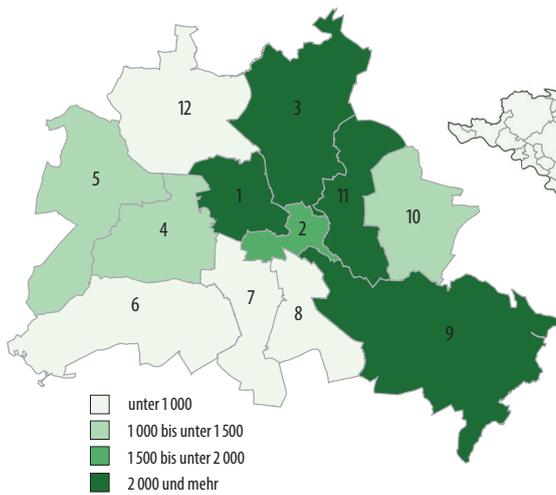
Was wurde gebaut?

Das Bevölkerungswachstum und die große Nachfrage nach mehr Wohnraum stellen insbesondere Großstädte wie Berlin vor neue Herausforderungen. Während es in den Randbezirken oft noch ausreichend Raum für neue Bauprojekte gibt, müssen bei städtebaulichen Planungen



Neu fertiggestellte Wohnungen 2018 ...

a | ... in Berlin

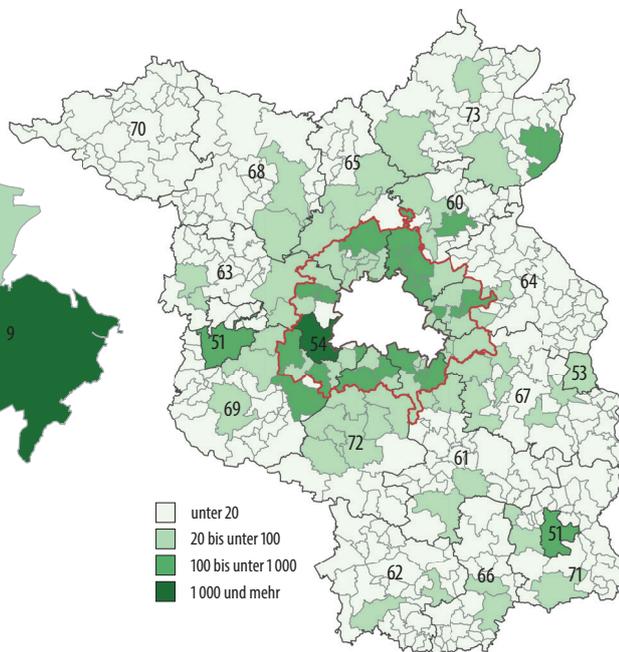


- ☐ unter 1 000
- ☐ 1 000 bis unter 1 500
- ☐ 1 500 bis unter 2 000
- ☐ 2 000 und mehr

Berlin

- 1 Mitte
- 2 Friedrichshain-Kreuzberg
- 3 Pankow
- 4 Charlottenburg-Wilmersdorf
- 5 Spandau
- 6 Steglitz-Zehlendorf
- 7 Tempelhof-Schöneberg
- 8 Neukölln
- 9 Treptow-Köpenick
- 10 Marzahn-Hellersdorf
- 11 Lichtenberg
- 12 Reinickendorf

b | ... im Land Brandenburg



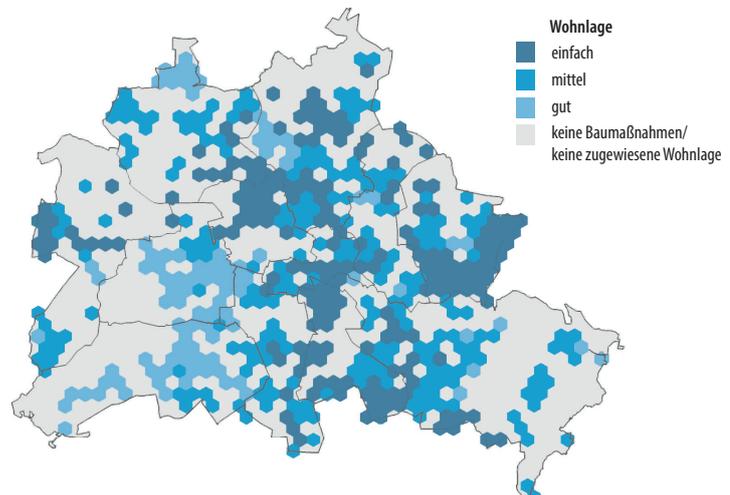
- ☐ unter 20
- ☐ 20 bis unter 100
- ☐ 100 bis unter 1 000
- ☐ 1 000 und mehr

Brandenburg

- 51 Brandenburg an der Havel
- 52 Cottbus
- 53 Frankfurt (Oder)
- 54 Potsdam
- 60 Barnim
- 61 Dahme-Spreewald
- 62 Elbe-Elster
- 63 Havelland
- 64 Märkisch-Oderland
- 65 Oberhavel
- 66 Oberspreewald-Lausitz
- 67 Oder-Spree
- 68 Ostprignitz-Ruppin
- 69 Potsdam-Mittelmark
- 70 Prignitz
- 71 Spree-Neiße
- 72 Teltow-Fläming
- 73 Uckermark

c | Überwiegende Wohnlage der 2018 neu fertiggestellten Wohnungen in Berlin

im 1-km-Raster



Wohnlage

- ☐ einfach
- ☐ mittel
- ☐ gut
- ☐ keine Baumaßnahmen/
keine zugewiesene Wohnlage

im Stadtzentrum freie Bauflächen effektiv genutzt werden. Die wenigen freien Flächen werden in der Regel mit mehrgeschossigen Wohnhäusern und möglichst vielen Wohnungen bebaut.

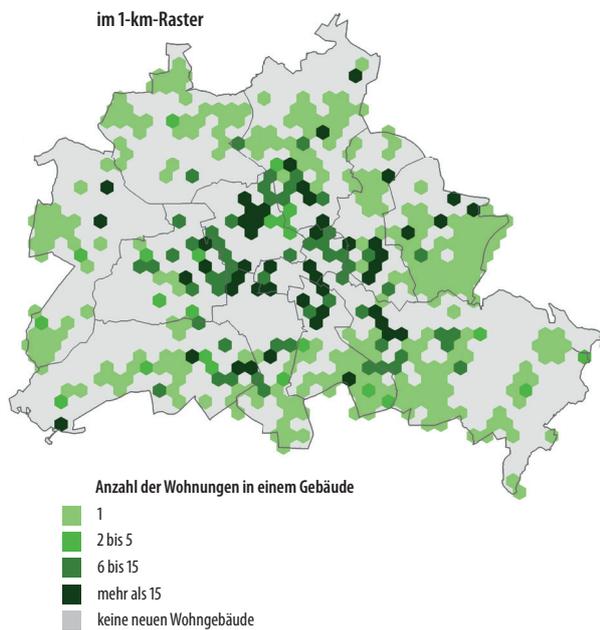
Im Jahr 2018 entstanden in den zentralen Bezirken Berlins vorwiegend größere Wohngebäude mit mehr als zehn Wohnungen. Abbildung d zeigt die überwiegende Anzahl von Wohnungen je Baufertigstellung in den Rasterzellen. In den Bezirken Mitte (90,6%) und Friedrichshain-Kreuzberg (80,5%) war die überwiegende Zahl der Baufertigstellungen mit mehr als zehn Wohnungen ausgestattet. Fast

11% der neu entstandenen Wohngebäude in Mitte bieten sogar mindestens 40 Parteien neuen Wohnraum. Bemerkenswert ist demgegenüber die Anzahl kleinerer Häuser mit nur einer Wohnung je Baufertigstellung (Einfamilienhäuser) in den Randbezirken. Mit immerhin 65,3% aller neu fertiggestellten Wohngebäude dominierten solche Einfamilienhäuser 2018 den Wohnungsbau in Berlin, wobei sie mit 1 371 nur 9,6% aller Wohnungen in neuen Wohngebäuden ausmachten.

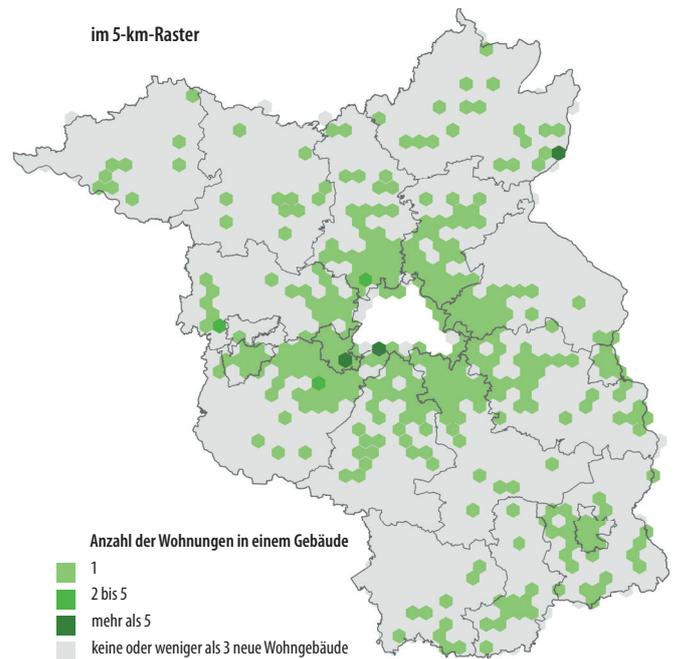
Ganz anders sah das Bild im Land Brandenburg aus (Abbildung e). Auch hier gab es in den kreisfreien

Überwiegende Anzahl der Wohnungen pro fertiggestelltem Wohngebäude 2018 ...

d | ... in Berlin

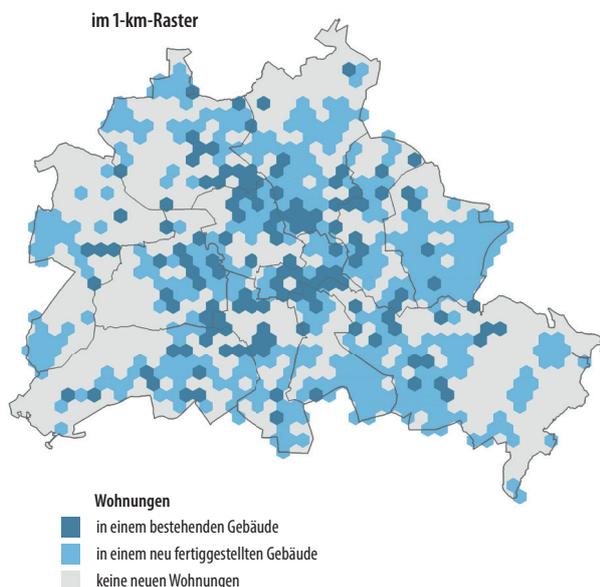


e | ... im Land Brandenburg

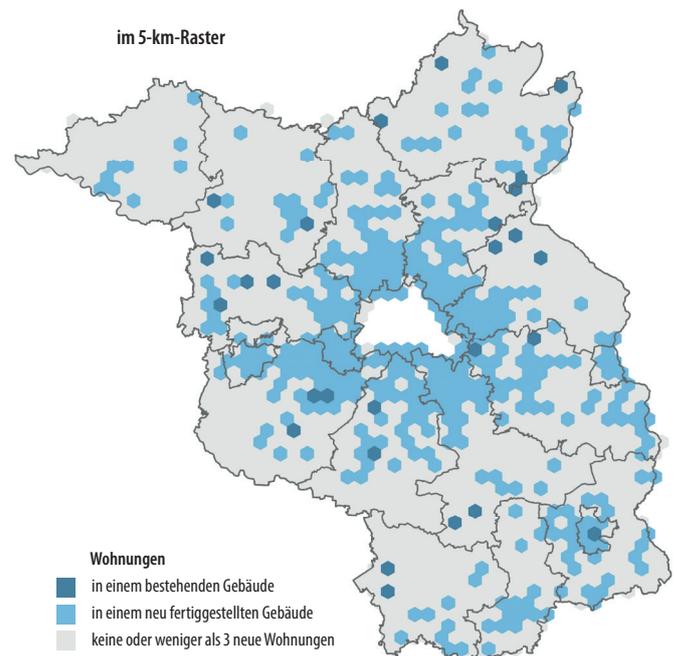


Überwiegende Art der Fertigstellung 2018 ...

f | ... in Berlin



g | ... im Land Brandenburg



Städten Potsdam (15,4%) und Cottbus (9,8%) sowie in Teltow (21,0%) neu fertiggestellte Wohngebäude mit mehr als zehn Wohnungen je Baufertigstellung, allerdings war der Anteil mit 2,4% an allen Baufertigstellungen im Bundesland eher unbedeutend. Die Karte mit einem Fünf-Kilometer-Raster verdeutlicht die große Zahl an Einfamilienhäusern in Brandenburg, die sogar einen prozentualen Anteil von 89,4% der neuen Wohngebäude erreichen und 50,9% aller Wohnungen in neuen Wohngebäuden bereitstellen.

Die limitierten Bauflächen in den Innenstädten führen nicht nur zum Bau größerer Wohnhäuser, sondern befördern auch den Aus- und Umbau bereits bestehender Wohngebäude. So entstanden im Jahr 2018 insgesamt 2224 (13,4%) der neuen Wohnungen in Berlin durch Baumaßnahmen an bereits bestehenden Wohngebäuden. Die Mehrzahl dieser Wohnungen befindet sich im Zentrum der Stadt, wo Bauflächen rar sind und alte Mietshäuser die Möglichkeit zum Wohnflächenausbau bieten oder Nichtwohngebäude zu Wohngebäuden umgebaut werden können (Abbildung f). Besonders in den Bezirken Mitte (424 Wohnungen) und Pankow (388 Wohnungen) wurden bestehende Wohngebäude ausgebaut, um neuen Wohnraum zu schaffen.

Mit 1095 Wohnungen (9,6% aller neuen Wohnungen) wurden im Land Brandenburg wesentlich seltener bestehende Gebäude um- oder ausgebaut. Lediglich in der kreisfreien Stadt Potsdam (197 Woh-

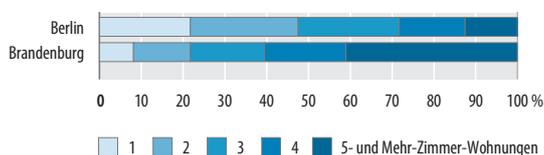
nungen) und in der Gemeinde Werder (Havel) (137 Wohnungen) wurden im größeren Umfang bestehende Gebäude für den Bau von neuen Wohnungen genutzt (Abbildung g).

Wie wurde gebaut?

Die Ansprüche an eine neue Wohnung sind sehr individuell. Während Singles und Paaren oft ein bis zwei Zimmer zum Leben genügen, brauchen Familien mit Kindern erheblich mehr Raum. Doch große, familiengerechte Wohnungen sind in Berlin selten: Ein Blick auf die Verteilung der verschiedenen Wohnungstypen offenbart, dass der Trend zu kleineren Wohnungen auch 2018 anhielt. Mit einer Anzahl von 7 868 machten Wohnungen mit maximal zwei Zimmern knapp die Hälfte (47,5%) aller neuen Wohnungen in Berlin aus. Immerhin 24,2% (4 006) waren 3-Zimmer- und 15,8% (2 615) 4-Zimmer-Wohnungen. Wohnungen mit mehr als vier Zimmern (12,5%; 2 062) wurden 2018 in Berlin nur wenige gebaut. Diese wenigen größeren Wohnungen wurden vorwiegend am Stadtrand gebaut, während kleinere Single-Wohnungen eher im Stadtzentrum entstanden (Abbildung h).

Auch hinsichtlich der Größe neu entstandener Wohnungen zeichnet sich für Brandenburg ein gegensätzliches Bild zu Berlin ab, welches mit der Vielzahl an neuen Einfamilienhäusern korrespondiert. Insgesamt entstanden hier 6 896 (60,6%) neue Wohnungen mit mindestens vier Zimmern und nur 2 466 (21,7%) kleinere Wohnungen mit ein bis zwei Zimmern. Ähnlich wie in Berlin sieht die Wohnungsentwicklung in der Landeshauptstadt aus, wo 44,5% der neu entstandenen Wohnungen 1- bzw. 2-Zimmer-Wohnungen waren. Ausgeglichen war die Entwicklung hingegen im Berliner Umland: Sowohl kleine Wohnungen für Singles als auch größere Wohneinheiten für Familien entstanden hier rund um die Großstadt.

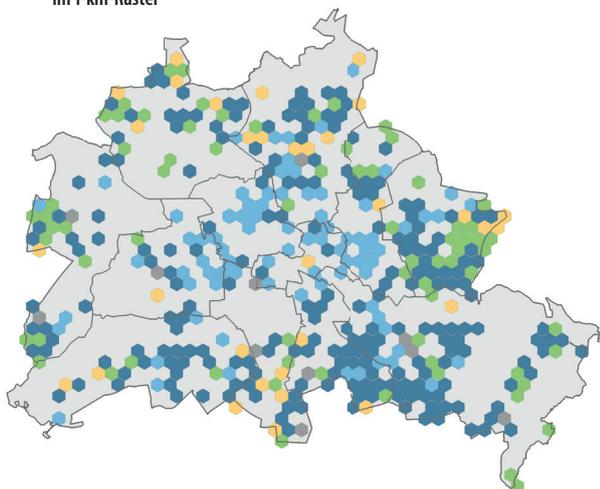
h | Anteil der 1- bis Mehr-Zimmer-Wohnungen 2018 in Berlin und im Land Brandenburg



Überwiegende primäre Energiequelle der fertiggestellten Wohngebäude 2018 ...

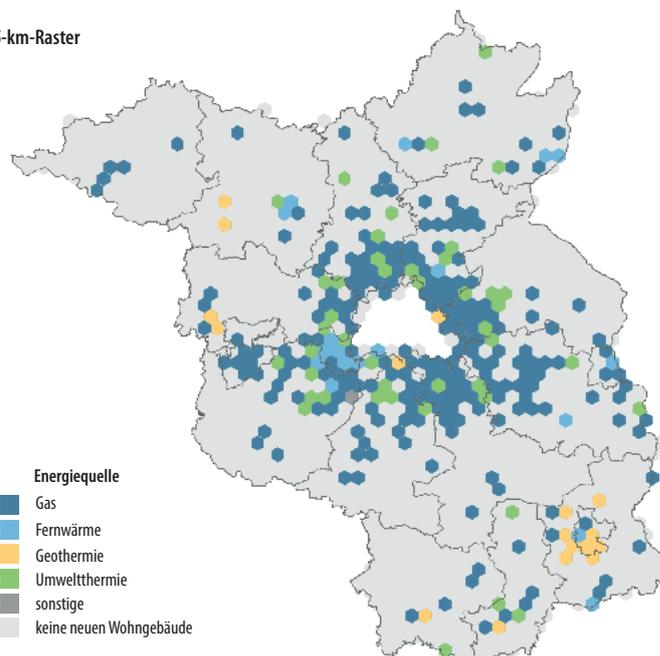
i | ... in Berlin

im 1-km-Raster



j | ... im Land Brandenburg

im 5-km-Raster



Ein interessanter Aspekt bei der Fertigstellung von neuen Wohngebäuden ist die Frage nach der primären Energiequelle für die Heizung. Auch hierbei gab es 2018 klare Unterschiede zwischen Stadtzentrum und Außenbereich sowie den Brandenburger Gemeinden. Im Berliner Zentrum wie auch in der kreisfreien Stadt Potsdam war Fernwärme die überwiegende Quelle für Heizenergie (Abbildungen i, j). Im Bezirk Mitte wurde in den neuen Wohngebäuden zu 95,3 % mit Fernwärme geheizt, gefolgt von Charlottenburg-Wilmersdorf mit 77,1 % Fernwärmeheizungen. Auch in Potsdam war die Fernwärme mit 46,7 % die am stärksten genutzte Energiequelle. In den Randbezirken Berlins sieht das Bild etwas bunter aus, denn ein Großteil der Berliner Neubauten wird hier mit Gas beheizt und platziert den Energieträger mit einem Anteil von 48,1 % aller Neubauten auf Platz eins in der Berliner Gesamtliste. Daneben gibt es auch eine größere Anzahl an Gebäuden mit Geothermie und Umweltthermie (5,4 %; 16,9 %) als regenerative Energieformen, welche auch in vielen Häusern Brandenburgs (13,0 %; 25,7 %) zum Einsatz kommen. Insbesondere in Cottbus war der Anteil dieser Heiztechniken in den neu gebauten Wohnhäusern mit 42,5 % Geothermie und 11,2 % Umweltthermie sehr hoch. An erster Stelle stand in Brandenburg jedoch die Gasheizung als primäre Energiequelle für 51,5 % aller neu fertiggestellten Wohngebäude.

Der notwendigen Flexibilität und Kreativität bei der effizienten Nutzung freier Flächen wird auch mit einer in den letzten Jahren zunehmenden Nutzung von Nichtwohngebäuden, wie alten Industriehallen, Supermarktdächern und Gewerbegebäuden, begegnet. Auf diesem Wege sind in Berlin 2018 bereits 155 neue Wohnungen entstanden und es bleibt abzuwarten, wie neue Trends das Wohnen auch in den nächsten Jahren verändern. Die Statistik der Baufertigstellungen in Verbindung mit geografischen Koordinaten bleibt also ein spannendes Analysefeld auf der Suche nach Lösungen für knappen Wohnraum.

Nicole Jurisch ist Sachbearbeiterin in der Stabsstelle *Querschnittsanalysen und Digitale Transformation* des Amtes für Statistik Berlin-Brandenburg.

Quellenverzeichnis

- [1] <https://www.destatis.de/Europa/DE/Thema/Bevoelkerung-Arbeit-Soziales/Soziales-Lebensbedingungen/Ueberbelegung.html>
- [2] <https://www.stadtentwicklung.berlin.de/wohnen/mietspiegel/de/wohnlagen.shtml>

Ein großer Dank gilt der Kollegin Brit Boche aus dem Referat *Baugewerbe, Bautätigkeit* für die fachliche Unterstützung bei der Analyse.

Neuerscheinung

▣ Dossier zur Corona-Pandemie in Berlin und Brandenburg

Für eine realistische Einschätzung der wirtschaftlichen und sozialen Lage sind verlässliche Zahlen unabdingbar. Diese in hoher Qualität bereitzustellen, ist eine der Kernaufgaben der amtlichen Statistik. Angesichts der aktuellen Herausforderungen durch die globale Ausbreitung des SARS-CoV-2-Virus und der dadurch verursachten Erkrankung COVID-19 im Frühjahr 2020 ist auch die amtliche Statistik in besonderer Weise gefordert. Das Amt für Statistik Berlin-Brandenburg (AfS) hat – jeweils orientiert an konkreten Fragestellungen zur Corona-Situation – statistische Daten nach den drei Themenbereichen Gesundheit, Gesellschaft und Wirtschaft aufgeschlüsselt.

Auf der einen Seite ist das Bedürfnis nach zuverlässigen Informationen zum gesellschaftlichen und wirtschaftlichen Geschehen mit Bezug auf die Corona-Pandemie enorm; auf der anderen Seite werden gerade in Krisenzeiten sehr aktuelle Zahlen benötigt, die aber in vielen Fällen noch nicht in der Qualität vorliegen, wie sie dem Selbstverständnis der amtlichen Statistik entspricht.

Das Dossier zur Corona-Pandemie in Berlin und Brandenburg stellt relevante Strukturdaten mit Bezug zur Corona-Pandemie zusammen und soll der Zielgruppe so die aufwendige Recherche im Gesamtdatenbestand des AfS ersparen oder – falls über die in diesem Dossier zusammengetragenen Informationen hinausgehende Angaben benötigt werden – ihr diese durch gezielte Verweise auf das Datenangebot zumindest erleichtern.

Weiterhin werden im Dossier in Form von Sonderauswertungen sehr aktuelle, teilweise noch vorläufige Zahlen zum aktuellen Geschehen zusammengestellt, die sich im regulären Datenangebot des AfS noch nicht finden. Dort, wo diese Zahlen mit Unsicherheiten verbunden sind, wird explizit darauf hingewiesen, um den Nutzenden eine Einordnung der Belastbarkeit der Angaben zu ermöglichen.

Dem Anspruch nach größtmöglicher Aktualität entsprechend wird das Dossier laufend aktualisiert und um weitere Informationen ergänzt.

statistik Berlin Brandenburg

DOSSIER ZUR CORONA-PANDEMIE

in Berlin und Brandenburg

Stand: 10. Juni 2020 (wird fortlaufend aktualisiert)



Gesundheit
Seite 3

- Krankenhauspersonal
- Krankenhausbetten
- Sterbefälle ↻
- Todesursachen
- Rauchen und Adipositas
- Pflege ↻



Gesellschaft
Seite 20

- Kinderbetreuung
- Schulen
- Hochschulen ☆
- Wohnraum ☆
- Haushalte mit Kindern und Senioren
- Bevölkerung ↻
- Pendler
- Verkehrsunfälle ☆



Wirtschaft
Seite 41

- Gesamtwirtschaft ☆
- Unternehmensregister
- Gewerbemeldungen
- Insolvenzen ☆
- Verarbeitendes Gewerbe ☆
- Energie
- Außenhandel
- Preise ☆
- Einzel- und Kfz-Handel ☆
- Gastgewerbe
- Tourismus ↻
- Luftverkehr ↻
- Verdienste ☆
- Selbstständige
- Saisonarbeiter



Das Corona-Dossier des AfS ist verfügbar unter:
<https://www.statistik-berlin-brandenburg.de/corona/corona.asp>
 Über den QR-Code gelangen Sie direkt dahin.

Verarbeitendes Gewerbe

Reurbanisierung der Industrie

Neue Chance für die Stadt?

von **Ramona Voshage** und **Martin Gornig**

Mit der Digitalisierung steht die Industrie vor einem großen Umbruch, der auch die räumlichen Standortmuster verändert. Um Hinweise auf die Richtung der Strukturveränderungen zu gewinnen, analysieren die Autoren das industrielle Gründungsgeschehen. Spitzenreiter sind dabei großstädtische Regionen und hierbei insbesondere bislang weniger erfolgreiche Städte wie Berlin, Leipzig und Dresden sowie die Metropolen an Rhein und Ruhr.

Dieser Beitrag ist ein Nachdruck aus: Bundesinstitut für Bau-, Stadt- und Raumforschung (2020): Informationen zur Raumentwicklung, Heft 6/2019, S. 94.

Die Städte wurden mit der Industrialisierung groß. Sie stieß die Wachstumsprozesse traditioneller Zentren an und sorgte dafür, dass neue Städte entstanden. Industrialisierung und Urbanisierung sind historisch gesehen untrennbar miteinander verbunden (vgl. Croon 1963). Merkmale wie eine besondere Produktionsweise (Fordismus, Taylorismus), eine spezifische Sozialstruktur (Arbeiterschaft) und daran ausgerichtete räumliche Strukturen (spezifische Gewerbezone und Wohngebiete) prägten die räumliche Gliederung der modernen Stadt. Vor allem in den Städten wurden Technologien und Produkte entwickelt, die wiederum das Wachstum der Industrie vorantrieben. Mit der zunehmenden Automatisierung und wachsender internationaler Konkurrenz in den 1970er und 1980er Jahren verschwand die Industrie durch Betriebsverlagerungen oder Betriebsschließungen jedoch vielfach aus den großen Städten. Die räumlichen Muster in Deutschland und in vielen anderen traditionellen Industrieländern prägten mehr und mehr Suburbanisierungsprozesse. Flächenintensive ebenerdige Produktionsorganisation, flexible LKW-basierte Logistikstrukturen und Umweltschutzauflagen vertrieben die Industrie zunehmend aus den Zentren (vgl. Siebel 2005).

Digitalisierung verändert industrielle Wachstumsmuster

Mit der Digitalisierung steht nun die Industrie wiederum vor großen strukturellen Veränderungen. In Deutschland wird dieser Umbruch als vierte industrielle Revolution und Industrie 4.0 bezeichnet. Neue digitale Kundenbeziehungen, datengetriebene Steuerungsprozesse, sensorgesteuerte Robotergenerationen oder additive Fertigungstechnologien ermöglichen es, industrielle Pro-

duktionsprozesse und Produkte zu denken und umzusetzen (vgl. Hüther 2016). Damit verbunden sind auf der einen Seite enorme Potenziale zur Rationalisierung, die ganze Produktionsschritte und Berufsgruppen überflüssig machen (vgl. IAB 2015). Auf der anderen Seite ergeben sich völlig neue Absatzpotenziale, beispielsweise durch kleinstserielle Fertigung und Sofortproduktionen (vgl. PwC 2014).

Die Bedeutung von räumlichen Kostenunterschieden oder von Raumüberwindungskosten – also Kosten, die beim Transport von Waren, Dienstleistungen oder Informationen von Ort A nach Ort B entstehen – verändert sich ebenfalls. Entsprechend verändert die Digitalisierung der Industrie auch die Standortanforderungen und die Raumstruktur der Industrie auf viele unterschiedliche Arten. Das gilt für die internationale Arbeitsteilung, aber insbesondere auch für die Stadt-Land-Beziehungen in Deutschland. Gerade hinsichtlich räumlicher Muster in Deutschland gibt es sowohl Szenarien mit zunehmend polarisierenden Entwicklungsmustern zwischen digitalen Hochburgen und abgehängten Regionen als auch konvergente Prozesse mit abnehmenden regionalen Unterschieden (vgl. Bertelsmann Stiftung, SNV 2016). Ähnlich sieht es bei den Entwicklungspotenzialen innerhalb von Stadtregionen aus: Hoch produktive Fertigungen im Hightech-Sektor können die Stadtgesellschaften beispielsweise infolge sich spezialisierender Anforderungen an berufliche Qualifikationen und Freisetzung von Berufsgruppen mit niedrigem Qualifikationsgrad weiter spalten. Es bietet sich aber gleichzeitig auch die Chance, über digitale Technologien lokale handwerkliche Produktionen in bislang benachteiligten Stadtquartieren zu revitalisieren (vgl. Läßle 2016).

In welche Richtung entwickeln sich die industriellen Raummuster tatsächlich? Wie verändert sich die Beschäftigungssituation bei wachsenden und neu gegründeten Industriebetrieben – gerade auch verglichen mit den regionalen Mustern bei schrumpfenden und geschlossenen Betrieben?

Beschäftigungstrend in deutschen Agglomerationen negativ

Verglichen mit anderen traditionellen Industrieregionen schneidet Deutschland bei der industriellen Entwicklung insgesamt gut ab. Mit über 20% liegt der Beschäftigungsanteil weit höher als in den meisten anderen OECD-Ländern (vgl. Gornig/Schiersch 2016). An den regionalen Entwicklungsmustern innerhalb Deutschlands hat sich allerdings bislang wenig geändert. Dies gilt vor allem für die Industrie in den Agglomerationen. Als Agglomerationen werden hier die Verflechtungsregionen von Kernstädten mit mehr als 500 000 Einwohnern betrachtet. Die räumliche Abgrenzung erfolgt approximativ anhand der jeweiligen Raumordnungsregionen.

Die Raumordnungsregionen sind funktionalräumliche, bundesweit vergleichbare Analyseregionen für Zwecke der Raubeobachtung und Politikberatung (vgl. BBSR 2019). In den Fällen von Hamburg und Bremen wurden auch die angrenzenden Raumordnungsregionen mit einbezogen.

Die industrielle Gesamtbeschäftigung in den so abgegrenzten Agglomerationen ist weiter gesunken. Die Zahl der Industriearbeitsplätze in den übrigen Regionen blieb hingegen nahezu konstant (Abbildung a). Der Gesamttrend sagt allerdings nichts darüber aus, wie sich die Beschäftigung auf Unternehmensebene entwickelt hat (vgl. Davies/Haltiwanger/Schuh 1996; OECD 1987). Saldierte Veränderungen im Bestand ergeben sich aus ganz unterschiedlichen Prozessen. Arbeitsplätze entfallen, wenn Betriebe ihre Tätigkeit einstellen oder ihren Personalbestand reduzieren. Zugleich entstehen immer wieder neue Arbeitsplätze: Zum einen, weil Betriebe neu gegründet werden und zum anderen, weil sie ihren Personalbestand erhöhen. Im Saldo dieser vier Prozesse ergibt sich zwar die oben angesprochene gesamte Veränderung, doch die bloße Information über diese Bestandsveränderung verdeckt die tatsächlichen Bewegungen.

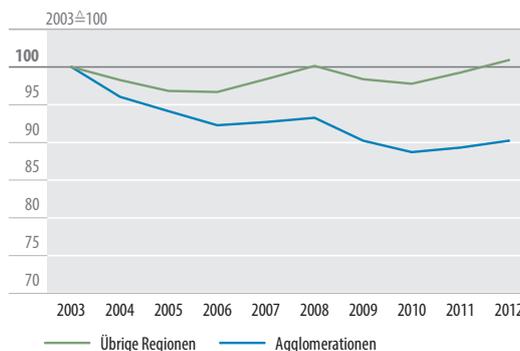
Informationen über diese verschiedenen Dynamiken liefern wichtige Hinweise gerade auch auf räumliche Strukturveränderungen. An welchen Orten Betriebe wachsen oder gegründet werden, wo sie schrumpfen oder schließen, könnte auf künftige Standortanforderungen der Industrie hinweisen. Um solche differenzierten Informationen über die Entwicklung der Beschäftigtenzahl zu erhalten, analysieren die Autoren in diesem Beitrag die Arbeitsplatzdynamik mit dem Messkonzept des Job-Turnover (vgl. Strotmann 1999; Gerlach/Wagner 1995). Bei diesem Konzept wird mit Einzeldaten der Betriebe kategorisiert, wie viele Arbeitsplätze in einer Periode hinzugekommen und wie viele weggefallen sind. Die Grundgesamtheit der Betriebe des Verarbeitenden Gewerbes verändert sich, sobald Betriebe neu

hinzukommen und sobald Betriebe ausscheiden (Abgänge). Es ergeben sich vier verschiedene relevante Bruttoströme: Neue Betriebe schaffen durch ihren Zugang Arbeitsplätze, Abgänge sind hingegen mit einem Arbeitsplatzabbau verbunden. Bestehende Betriebe lassen sich ferner in folgende Gruppen untergliedern: Bestandsbetriebe, die ihre Beschäftigtenzahl erhöhen, gelten als wachsende Betriebe. Analog dazu zählen Bestandsbetriebe, die Beschäftigte abbauen, zu den schrumpfenden Betrieben.

Möglich ist eine derartige Analyse allerdings nur, wenn die vorliegenden Daten eine Gruppierung aller Betriebe in schließende, schrumpfende, neu gegründete und wachsende Betriebe erlauben. Entscheidend für die Aussagekraft ist zudem, dass alle relevanten Betriebe erfasst werden. Anzustreben ist daher eine Totalerhebung. In Deutschland ist die amtliche Statistik des Verarbeitenden Gewerbes vergleichsweise gut ausgebaut und hat für viele Analysen nahezu den Charakter einer Vollerhebung. Ungeachtet dessen handelt es sich nicht um ein integriertes, aufeinander abgestimmtes System amtlicher Statistiken. Dementsprechend müssen sich komplexere Analysen stets an unterschiedlichen Statistiken aus diesem Segment bedienen.

Der vorliegenden Auswertung liegt ein Mikrodatsatz für das Verarbeitende Gewerbe Deutschlands zugrunde. Die Grundlage für den Datensatz bildet das AFiD-Panel Industriebetriebe der Statistischen Ämter des Bundes und der Länder (vgl. Malchin/Pohl 2007; Malchin/Voshage 2009), das die Industriebetriebe für die Jahre 1995 bis 2002 komplett erfasst. Mit dem ersatzlosen Wegfall der Kleinbetriebserhebung ab 2003 entstand hier eine Dokumentationslücke. Angaben des Unternehmensregisters der Statistischen Ämter des Bundes und der Länder (vgl. FDZ 2012) schlossen diese Lücke. Mit geeigneten Schätzverfahren wurden die fehlenden Betriebe mit weniger als 20 tätigen Personen ergänzt (vgl. Voshage 2019). Dieser so gebildete Analysedatsatz, der Informationen zu den Raumordnungsregionen enthält, deckt das Verarbeitende Gewerbe in Deutschland für den Zeitraum 2003 bis 2012 erstmals komplett ab.

a | Entwicklung der Zahl der Beschäftigten nach Agglomerationen im Verarbeitenden Gewerbe, 2003 bis 2012



Quellen: Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder, AFiD-Panel Industriebetriebe und AFiD-Panel Unternehmensregister, eigene Berechnungen

Die Analyse zeigt, dass in den Agglomerationen im gesamten Beobachtungszeitraum mehr Arbeitsplätze verloren gingen als in den übrigen Regionen – sowohl durch schrumpfende als auch durch „abgehende“ Industriebetriebe (Abbildung b). In beiden Regionstypen dominierte jedoch maßgeblich der Arbeitsplatzabbau durch schließende Betriebe.

Durch wachsende Betriebe gewannen die Agglomerationen und die übrigen Regionen hingegen ähnlich viele Arbeitsplätze hinzu. In Bezug auf die Beschäftigungsgewinne durch neue Betriebe kehrt sich das regionale Muster sogar um: Durch neue Industriebetriebe gewannen die Agglomerationen mit durchschnittlich 4,2% spürbar mehr Arbeitsplätze als die übrigen Regionen mit 3,8%.

Zusammenfassend lassen sich die gesamten Beschäftigungsverluste der Industrie überwiegend auf schließende Betriebe zurückführen. Insbesondere die verlorenen Arbeitsplätze in den Agglomerationen ergaben sich vorrangig aus „abgehenden“ Betrieben. Für neue Industriebetriebe scheinen die Entwicklungsbedingungen in den Agglomerationen hingegen besonders günstig (Gornig/Werwatz 2018). Ausgehend davon, dass die neuen (digitalen) Rahmenbedingungen am ehesten junge Betriebe in ihrem Standortverhalten beeinflussen, würde das zu neuen Standortanforderungen führen. Aber wo genau in Deutschland entsteht welche Art von neuen Industrieunternehmen? Als Gründungen gelten hier neue rechtlich selbstständige Betriebe im Verarbeitenden Gewerbe laut Gewerbeanmeldungen.

Industrielle Gründungsintensität in den Agglomerationen hoch

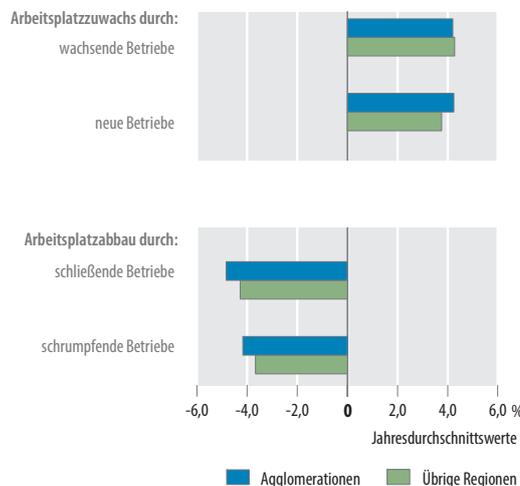
Abbildung c weist die industriellen Gründungen im Durchschnitt der Jahre 2012 bis 2016 bezogen auf die Beschäftigten im Verarbeitenden Gewerbe zu Beginn des Beobachtungszeitraums aus. Im Fokus ste-

hen damit unterschiedliche Gründungsintensitäten, um mögliche Dynamiken in den Standortmustern zu identifizieren. Dabei lässt sich zunächst festhalten, dass die Gründungsintensität als gewichteter Durchschnitt in den Agglomerationen mit 80 Gründungen je 100 000 Beschäftigten im Jahr um fast 40% höher liegt als in den übrigen Regionen der Bundesrepublik.

Gleichzeitig unterscheiden sich aber die Gründungsintensitäten auch stark zwischen den Agglomerationen. In Berlin werden bezogen auf die bisherige Industriebeschäftigung jährlich fast viermal so viele Unternehmen im Verarbeitenden Gewerbe gegründet wie in den Nicht-Agglomerationen. Ebenfalls überdurchschnittliche Gründungsintensitäten weisen die Regionen Leipzig/Dresden, das Rhein-Main-Gebiet und die Metropolen an Rhein und Ruhr auf. Von den bislang so wachstumsstarken süddeutschen Agglomerationen besitzt nur München auch eine sehr starke industrielle Gründungsintensität. In der Region Stuttgart entstehen hingegen bezogen auf die bisherige Beschäftigung unterdurchschnittlich wenig neue Industrieunternehmen.

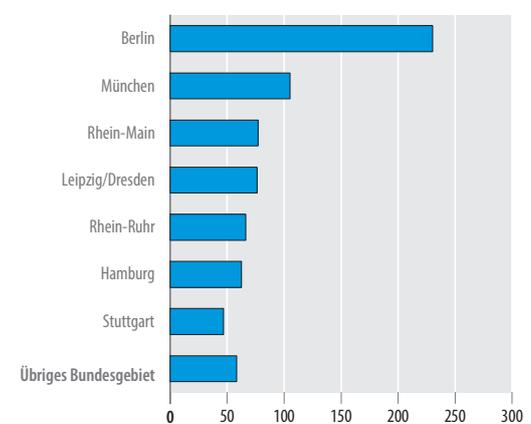
Insgesamt deutet das Gründungsgeschehen darauf hin, dass sich die industriellen Wachstumsmuster im Zuge der Digitalisierung nicht nur zugunsten der großen Städte verschieben, sondern sich auch zwischen den Agglomerationen neue Wachstumshierarchien herausbilden könnten. Um abzuschätzen, ob die neuen Wachstumsmuster wesentlich durch eine neue digitale Hightech-Industrie oder auch einfachere digitale konsumorientierte Industrien getragen werden, lassen sich die Gründungen den verschiedenen Technologiebereichen zuordnen (vgl. Eurostat 2017; die Bereiche Spitzentechnologie und hochwertige Technologie wurden dabei zum Bereich Hightech zusammengefasst).

b | Bruttoströme der Arbeitsplatzdynamik nach Agglomerationen im Verarbeitenden Gewerbe, 2003/2004 bis 2011/2012



Quellen: Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder, AFID-Panel Industriebetriebe und AFID-Panel Unternehmensregister, eigene Berechnungen

c | Betriebsgründungen im Verarbeitenden Gewerbe – Durchschnitt 2012 bis 2016 je 100 000 Beschäftigten



Quellen: Gewerbeanzeigenstatistik, AFID-Panel Industriebetriebe und AFID-Panel Unternehmensregister, eigene Berechnungen

Abbildung d setzt die Gründungsintensitäten der betrachteten Agglomerationen jeweils in Bezug zur durchschnittlichen Entwicklung in der übrigen Bundesrepublik.

Die Agglomerationen ziehen vor allem neue Betriebe aus dem Bereich der Lowtech-Industrien wie Nahrungsmittel-, Möbel- und andere Konsumgüterproduktionen an. In Berlin ist die Gründungsintensität fünfmal und in München dreimal so hoch wie in den Nicht-Agglomerationen. In Hamburg und im Rhein-Main-Gebiet übersteigt die Gründungsintensität in den Lowtech-Industrien den Referenzwert immerhin um rund 70%. Aber auch in Stuttgart, Leipzig, Dresden und den Metropolen an Rhein und Ruhr ist die Gründungsintensität überdurchschnittlich. Dieses Ergebnis deutet darauf hin, dass sich mit der Digitalisierung gerade für eher traditionelle Konsumgüterindustrien neue Entwicklungspotenziale in den Städten eröffnen.

Gleichzeitig weisen die Agglomerationen mit ihrer ausgebauten Forschungsinfrastruktur insgesamt auch im Bereich der Hightech-Industrien wie Fahrzeugbau, Maschinenbau, Elektrotechnik oder Pharmaproduktion eine deutlich höhere Gründungsintensität auf als die übrigen Regionen in Deutschland. Spitzenreiter ist hier wiederum die Bundeshauptstadt: Bezogen auf die Ausgangsbeschäftigung werden in Berlin mehr als viermal so viele Hightech-Unternehmen gegründet wie in den Nicht-Agglomerationen. Verglichen damit ist die Gründungsintensität im Mediumtech-Bereich, wozu insbesondere die Chemie und andere Grundstoffindustrien zählen, relativ gering. Besonders ausgeprägt ist die Polarisierung im Gründungsgeschehen in Berlin. Ein ähnliches Schema weisen aber auch Hamburg, das Rhein-Main-Gebiet und die Region Rhein-Ruhr auf. In der Region Stuttgart fällt

die Gründungsintensität hingegen gerade bei den Hightech-Industrien, in denen die Region derzeit so erfolgreich ist, unterdurchschnittlich aus.

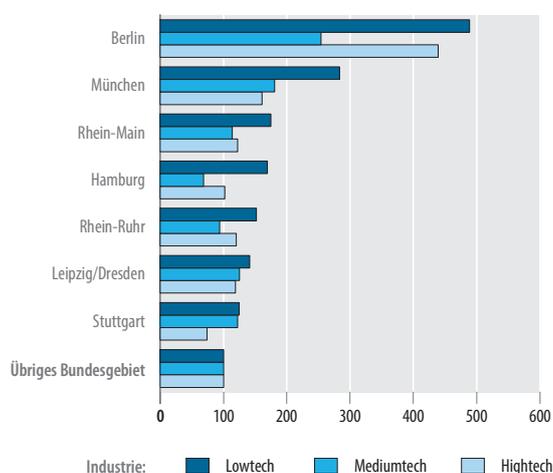
Hochschulen und Konsumenten als Anziehungspunkte

Was aber macht die großen Städte nun so attraktiv für industrielle Gründungen? Ein genauerer Blick auf die Standorte der Gründungen innerhalb der Städte bringt Antworten. In diesem Fall ist ein Blick auf die größte deutsche Stadt mit der absolut wie relativ höchsten Gründungsdynamik – nämlich Berlin – möglich. Im Forschungsdatenzentrum des Amtes für Statistik Berlin-Brandenburg wurden dazu die Mikrodaten des Unternehmensregisters ausgewertet (vgl. Klare 2017). Um den Datenschutz zu wahren, wurden die Unternehmensstandorte 60 statistischen Planungsräumen in der Stadt zugeordnet (vgl. Geoportal Berlin, 2010). Für die Jahre 2013 und 2014 ließen sich auch die Gründungen diesen Räumen zuordnen.

Die industriellen Gründungen in Berlin weisen dabei eine hohe räumliche Konzentration auf. Bei den Gründungen im Bereich der Hightech-Industrie entfallen rund ein Drittel der 155 Gründungen auf sechs Standorte (Abbildung e). Zu den meisten Gründungen im Hightech-Bereich kam es in der westlichen Innenstadt im Bezirk Charlottenburg. Dort befindet sich der Hauptcampus der Technischen Universität Berlin. Weiter östlich sitzt unter anderem die Beuth Hochschule für Technik. Zu den bevorzugten Gründungsstandorten zählt auch Adlershof im Südosten Berlins, in dem sich ein Campus der Humboldt-Universität zu Berlin und ein großes Technologiezentrum befinden (vgl. Handrich/Pavel/Proske 2008).

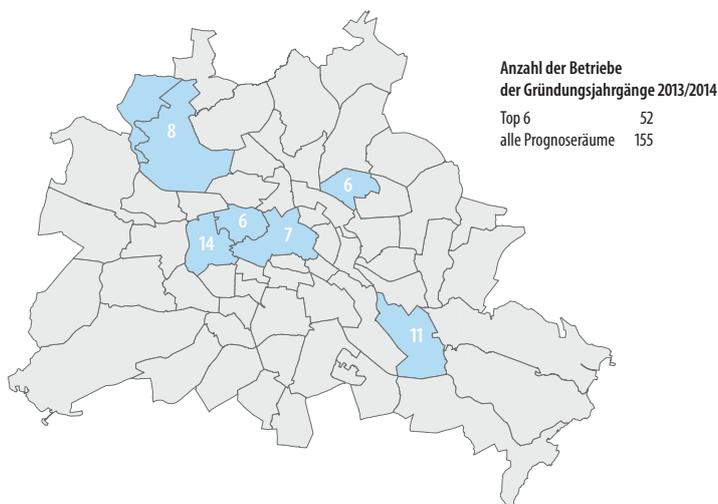
Ebenfalls stark auf wenige Standorte konzentriert ist das Gründungsgeschehen bei Lowtech-Industri-

d | Betriebsgründungen im Verarbeitenden Gewerbe nach Technologieklassen – Durchschnitt 2012 bis 2016 je 100 000 Beschäftigten



Quellen: Gewerbeanzeigenstatistik, AFID-Panel Industriebetriebe und Unternehmensregister der Statistischen Ämter des Bundes und der Länder, eigene Berechnungen

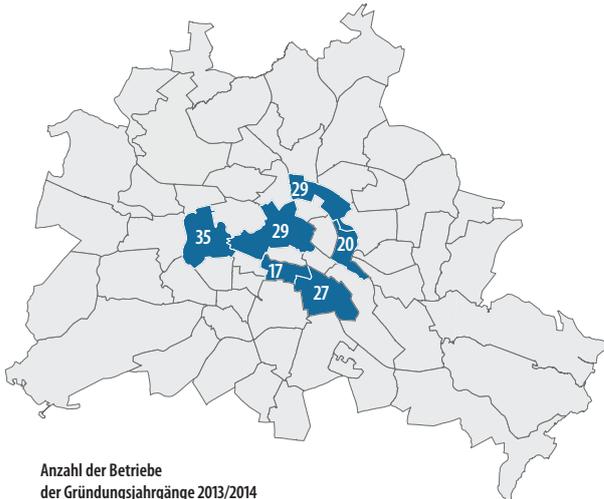
e | Betriebsgründungen in Berlin: Hightech-Betriebe



Quellen: Amt für Statistik Berlin-Brandenburg, Unternehmensregister, eigene Berechnungen

en. Hier entfallen sogar deutlich mehr als ein Drittel der Gründungen auf sechs Standorte (Abbildung f). Diese bevorzugten Gründungsstandorte liegen allesamt im hochverdichteten Innenstadtbereich Berlins. Die Konzentration spricht für eine hohe Bedeutung der räumlichen Kundennähe der im Bereich Lowtech dominierenden konsumnahen Produktion.

f | Betriebsgründungen in Berlin: Lowtech-Betriebe



Anzahl der Betriebe
der Gründungsjahrgänge 2013/2014

Top 6 157
alle Prognoseräume 448

Quellen: Amt für Statistik Berlin-Brandenburg, Unternehmensregister,
eigene Berechnungen

Fazit

Die Digitalisierung besitzt das Potenzial, auch die räumlichen Standortmuster der Industrie neu zu justieren. Dies gilt nicht nur für die internationale Arbeitsteilung, sondern gerade auch für die regionalen Strukturen in Deutschland. Eine Analyse des industriellen Gründungsgeschehens hierzulande weist darauf hin, dass Agglomerationen von den neuen (digitalen) Produktionen profitieren können. Verantwortlich für die neue Attraktivität der Agglomerationen für die (digitale) Industrie dürfte der Bedarf an räumlicher Nähe zu Forschungseinrichtungen und vor allem zu Konsumenten sein.

Als ein zentraler Engpass könnte sich allerdings die zunehmende Flächenknappheit in den Agglomerationen erweisen. Der fehlende bezahlbare Wohnraum in den Innenstädten setzt die Politik unter Druck, gerade in den Großstädten zusätzliche Wohnbauflächen zu aktivieren. Gleichzeitig bieten die Innenstädte durch ihre Nähe zu exzellenter Forschung und zahlungskräftigen Kunden aber auch für die neue (digitale) Industrie entscheidende Wachstumsvorteile (vgl. TU Berlin 2017).

Eine zentrale Aufgabe der Politik wird es daher sein, die bestehenden Nutzungskonflikte zwischen Wohnen und Gewerbe aufzulösen. Die offensive Anwendung von planungsrechtlichen Instrumenten der Nutzungsmischung wie im Fall der „urbanen Gebiete“ könnte wesentlich dazu beitragen. Gleichzeitig müssten tradierte Formen fester Planungsvorgaben wie bei Sicherungssatzungen für Gewerbeflächen flexibilisiert werden. In vielen Städten fehlt es derzeit zudem an einem leistungsfähigen Flächenmonitoring.



Foto: AFS Berlin-Brandenburg

Ramona Voshage leitet die Abteilung
Gesamtwirtschaft des Amtes für
Statistik Berlin-Brandenburg.



Foto: DIW Berlin

Prof. Dr. Martin Gornig ist
Forschungsdirektor *Industriepolitik*
am Deutschen Institut für Wirtschafts-
forschung und Honorarprofessor für
Stadt- und Regionalökonomie an der
Technischen Universität Berlin.

Danksagung:

Der Beitrag beruht auf Arbeiten im von der Hans-Böckler-Stiftung geförderten Forschungsprojekt „Industrie in der Stadt“ (vgl. Gornig et al. 2018).

Literatur

- Bertelsmann Stiftung, SNV – Stiftung neue Verantwortung, 2016: Auf dem Weg zum Arbeitsmarkt 4.0? Mögliche Auswirkungen der Digitalisierung auf Arbeitsmarkt und Beschäftigung in Deutschland bis 2030, Gütersloh/Berlin.
- BBSR – Bundesinstitut für Bau-, Stadt- und Raumforschung, 2019: Laufende Raumbearbeitung – Raumabgrenzungen. Zugriff: <https://www.bbsr.bund.de/BBSR/DE/Raumbearbeitung/Raumabgrenzungen/deutschland/regionen/Raumordnungsregionen/raumordnungsregionen-node.html> [abgerufen am 20.11.2019].
- Croon, Helmuth, 1963: Zur Entwicklung deutscher Städte im 19. und 20. Jahrhundert. In: *Studium Generale – Zeitschrift für die Einheit der Wissenschaften im Zusammenhang ihrer Begriffsbildungen und Forschungsmethoden* Bd. 9: 565–575.
- Davies, Steven J.; Haltiwanger, John; Schuh, Scott, 1996: Small business and job creation: Dissecting the myth and reassessing the facts. In: *Small Business Economics*. Bd. 8: 297–315.
- Eurostat, 2017: Glossar: Klassifikation des Verarbeitenden Gewerbes nach der Technologieintensität. Zugriff: http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Hightech_classification_of_manufacturing_industries/de [abgerufen am 10.11.2017].
- FDZ – Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder, 2012: AFiD-Panel Unternehmensregister 2002–2012. On-Site, Version 0. DOI: 10.21242/52111.2012.00.01.1.1.0
- Geoport Berlin, 2010: Flächennutzung der Planungsräume (LOR). FIS-Broker der Senatsverwaltung für Stadtentwicklung und Wohnen. Zugriff: <http://fbinter.stadt-berlin.de/fb/index.jsp> [abgerufen am 10.11.2017].
- Gerlach, Knut; Wagner, Joachim, 1995: Die Heterogenität der Arbeitsplatzdynamik innerhalb der Industrie. In: Frick, B.; Semlinger, K. (Hrsg.): *Betriebliche Modernisierung in personeller Erneuerung*. Berlin: 39–57.
- Gornig, Martin; Belitz, Heike; Geppert, Kurt; Löckener, Ralf; Schiersch, Alexander; Werwatz, Axel, 2018: *Industrie in der Stadt*. DIW Wochenbericht Bd. 85 (47): 1003–1027.
- Gornig, Martin; Schiersch, Alexander, 2016: Weak Investment Poses a Threat to Industry in Europe. *Intereconomics* Bd. 51 (5): 272–277.
- Gornig, Martin; Werwatz, Axel, 2018: Anzeichen für eine Reurbanisierung der Industrie. *DIW Wochenbericht* Bd. 85 (47): 1005–1011.
- Handrich, Lars; Pavel, Ferdinand; Proske, Sandra, 2008: Standort Berlin-Adlershof: kräftige Impulse für die Stadt. *DIW Wochenbericht* Bd. 75 (4): 41–46.
- Hüther, Michael, 2016: Digitalisierung: Systematisierung der Trends im Strukturwandel – Gestaltungsaufgabe für die Politik. *IW policy paper* 15.
- IAB – Institut für Arbeitsmarkt- und Berufsforschung, 2015: *Industrie 4.0 und die Folgen für Arbeitsmarkt und Wirtschaft*. IAB Forschungsbericht 8. Nürnberg.
- Klare, Jan, 2017: *Industriestandort Berlin*. Ergebnisse einer Auswertung des Unternehmensregisters. *Zeitschrift für amtliche Statistik Berlin-Brandenburg*. Bd. 11 (4): 58–63.
- Läßle, Dieter, 2016: *Produktion zurück in die Stadt*. *Bauwelt* 35.
- Malchin, Anja; Pohl, Ramona, 2007: *Firmendaten der amtlichen Statistik – Datenzugang und neue Entwicklungen im Forschungsdatenzentrum*. *Vierteljahreshefte zur Wirtschaftsforschung des DIW Berlin* Bd. 76 (3): 8–16.
- Malchin, Anja; Voshage, Ramona, 2009: *Official Firm Data for Germany*. *Schmollers Jahrbuch*. Bd. 129 (3): 501–513.
- OECD – Organization for Economic Cooperation and Development, 1987: *The Process of Job Creation and Destruction*. In: *Employment Outlook*.
- PwC – PricewaterhouseCoopers, 2014: *Industrie 4.0: Chancen und Herausforderungen der vierten industriellen Revolution*.
- Siebel, Walter, 2005: *Suburbanisierung*. In: *ARL – Akademie für Raumforschung und Landesplanung* (Hrsg.): *Handwörterbuch der Raumordnung*. Hannover: 1135–1140.
- Strotmann, Harald, 1999: *Zur Schaffung und zum Abbau von Arbeitsplätzen in der baden-württembergischen Industrie 1995 bis 1997*. *Baden-Württemberg in Wort und Zahl*. Nr. 2: 63–69.
- Strotmann, Harald, 2002: *Arbeitsplatzdynamik in der baden-württembergischen Industrie – Eine Analyse mit amtlichen Betriebspaneldaten*. Frankfurt am Main: 23.
- Technische Universität Berlin, 2017: *Aufgespürt: Raumprofile schnell wachsender Industrieunternehmen*. Projektbericht am Institut für Stadt und Regionalplanung, Berlin.
- Voshage, Ramona, 2019: *Zur Entwicklung der Berliner Industriebetriebe der Jahre 2003 bis 2012 im Vergleich deutscher Agglomerationsräume unter besonderer Berücksichtigung der Arbeitsplatzdynamik und der Modernisierung des industriellen Produktportfolios*. Dissertation. Technische Universität Berlin.

Zensus

Pretest zur Personenerhebung des Zensus 2021

von Mark Hoferichter, Ralf Mädler, Natalie Lück, Laura Tomm-Demski, Ulrike Warnke und Jonas-Otto Werner

Am 16. Mai 2021 findet der nächste Zensus in Deutschland statt.¹ Um die konzipierten Verfahrensabläufe und Erhebungsinstrumente für die Haupterhebung zu testen, wurde zwischen November 2019 und März 2020 in Berlin und Brandenburg ein Pretest durchgeführt. Der vorliegende Beitrag erläutert dessen Vorbereitung, Durchführung sowie Nachbereitung und gibt einen kurzen Überblick über die Ergebnisse.

Einleitung und Gesamtüberblick

In einer vom Statistischen Bundesamt geleiteten Arbeitsgruppe erarbeitete das Amt für Statistik Berlin-Brandenburg (AfS) zusammen mit anderen Statistischen Ämtern der Länder ein auf Rechtssicherheit, Qualität und Aktualität geprüfetes Verfahren der Personenerhebung im Rahmen des Zensus 2021.

Der Ablauf dieses Verfahrens sollte im Vorfeld durch die Statistischen Ämter der Länder im Rahmen eines Pretests umgesetzt und getestet werden. Die Feldphase erfolgte in Berlin und Brandenburg zwischen November 2019 und März 2020 – unter möglichst realistischen Bedingungen der Haupterhebung des Zensus 2021. Für die Umsetzung wurde im Vorfeld ein grober einheitlicher Arbeits- und Zeitplan konzipiert, der länderindividuell angepasst wurde (Abbildung a). Das Statistische Bundesamt gab einen Nettostichprobenumfang von 9052 deutschlandweit erfolgreich zu befragenden Personen vor. Dies entspricht circa 0,1% des Stichprobenumfangs der Haupterhebung und war notwendig, um valide Rückschlüsse für das zukünftige Verfahren treffen zu können. Bis auf Hessen und Mecklenburg-Vorpommern führten alle Statistischen Ämter der Länder den Pretest in ihrem Bundesland durch. Ziel war es, Aussagen über die Praktikabilität, mögliche Probleme und Optimierungspotenziale

Die Personenerhebung im Zensus 2021 hat zwei inhaltliche Ziele: die Einwohnerzahl festzustellen (Ziel 1) sowie zusätzliche Strukturmerkmale der Bevölkerung zu ermitteln (Ziel 2). Dies soll zum einen durch eine persönliche Befragung vor Ort zu Merkmalen der Person erfolgen. Die weiteren Strukturmerkmale zu den Themen Wohnsituation, Staatsangehörigkeit und Zuwanderung, Bildung und Ausbildung, Erwerbstätigkeit, Arbeitsort, Wirtschaftszweig und Beruf sollen zum anderen durch eine zusätzliche Online-Befragung ermittelt werden.

a | Übersicht Zeitplan Pretest

2019						2020					
vor August	August	September	Oktober	November	Dezember	Januar	Februar	März	April	Mai	
Erhebungskonzeption											
		Durchführungsvorbereitung									
		Erstellung Arbeitsbezirke									
		Unterlagendruck und Materialbeschaffung									
		Rekrutierung der Erhebungsbeauftragten									
		Schulung der Erhebungsbeauftragten									
		Erhebungsdurchführung									
		Öffentlichkeitsarbeit									
		Feldphase der Erhebungsbeauftragten									
		Erinnerungsgeschäft									
							Erhebungsnachbereitung				

Quelle: eigene Darstellung

¹ Aufgrund der Corona-Pandemie soll der für 2021 vorgesehene Zensus verschoben werden. Organisatorische und rechtliche Fragen zu einer Änderung des Zensusgesetzes und des EU-Rechts im Hinblick auf das weitere Verfahren bzw. eine Stichtagsverschiebung werden derzeit geklärt (Stand: Juni 2020). Unabhängig davon laufen die Vorbereitungen weiter.

des Verfahrens hinsichtlich der Durchführung des Zensus 2021 zu treffen. Dabei wurden die Erhebungsorganisation, die Erhebungsdurchführung und die Erhebungsinstrumente erprobt. Abweichend zur Haupterhebung ist die Befragung im Pretest keine Stichprobe, die eine Hochrechnung der Ergebnisse möglich machen soll. Es bestand keine Auskunftspflicht und es wurden keinerlei Abgleiche mit dem Melderegister durchgeführt. Die Aufbereitungsprozesse der Daten erfolgten ohne eine spezielle Fachanwendung.

Für die Durchführung der Erhebung (Abbildung b) wurden Interviewende, sogenannte Erhebungsbeauftragte, rekrutiert und geschult. Direkt im Anschluss an die Schulung erhielten diese ihre Erhebungsunterlagen. Ihre erste Aufgabe bestand darin, die ihnen zugewiesenen Anschriften zu sichten und die Anschriftenbefunde im dafür vorgesehenen Mantelbogen² zu erfassen. Während der Begehung warfen die Erhebungsbeauftragten Terminankündigungsschreiben mitsamt Informationsflyer in die jeweiligen Briefkästen der auskunftgebenden Personen ein. Zum vereinbarten Termin wurden die angetroffenen Personen in den Haushalten zu persönlichen Merkmalen befragt und diese Angaben im Haushaltsbogen³ erfasst. Die meisten Haushalte (rund 90 %) sollten darüber hinaus zu den weiteren soziodemografischen Merkmalen befragt werden. Hierfür erhielten die jeweiligen Personen Online-Zugangsdaten für den Online-Fragebogen zum Selbstausfüllen (CAWI). In Ausnahmefällen konnten sich die Auskunftgebenden auch einen Papierfragebogen (PAP) geben oder die weitere Befragung persönlich-mündlich (PAPI) durch den Interviewenden durchführen lassen. Als zusätzlichen Erhebungsweg gab es die Möglichkeit eines computergestützten Interviews (CATI) über eine Hotline, die für den Pretest beim AfS eingerichtet wurde.

Durchführungsvorbereitung Erstellen der Arbeitsbezirke

Das Statistische Bundesamt hatte die erfolgreiche Befragung von deutschlandweit 9 052 Personen vorgegeben. Auf Berlin fielen 433 und auf Brandenburg 302 Personen, die am Pretest teilnehmen sollten. Die dafür notwendigen Stichprobenanschriften wurden vom Statistischen Bundesamt im August 2019 bereitgestellt. Dafür gaben die Statistischen Ämter der Länder bereits im Vorfeld mindestens zwei Gemeinden oder Stadtteile von Großstädten je Bundesland vor, die für sie erhebungsorganisatorisch günstig lagen.

Das AfS meldete für Brandenburg die Stadt Potsdam sowie die Gemeinde Hoppegarten, für Berlin die Ortsteile Friedrichsfelde und Schöneberg. In diesen vier Raumeinheiten wurde jeweils eine Zufallsauswahl von Anschriften mit Wohnraum und mindestens einer gemeldeten Person getroffen.

Da für den Pretest keine Auskunftspflicht bestand, wurde mit einer Bereitschaft in der Bevölkerung zu einer freiwilligen Teilnahme von lediglich 10 % gerechnet. Um die Teilnahme zu steigern, wurden Auskunftgebende mit 7,50 EUR vergütet. Auf dieser Grundlage wurde eine Responserate von 24 % erwartet.

Anhand dieser Annahme plante das AfS die Einteilung der Arbeitsbezirke⁴. Die Arbeitsbezirke stellten die Arbeitspakete für die Interviewenden dar und beinhalteten zwischen vier und 20 Anschriften. Somit enthielt ein Arbeitsbezirk anhand des Melderegisters in Summe circa 150 Personen.

Für das „Clustern“ der Stichprobenanschriften zu Arbeitsbezirken wurden die Anschriften mit der jeweiligen Anzahl der gemeldeten Personen mittels der Geoinformationssoftware QGIS auf einer Karte dargestellt. Daraufhin wurden die Anschriften manuell auf Arbeitsbezirke verteilt, bis in allen Bezirken etwa 150 gemeldete Personen akkumuliert waren. Ziel dabei war es, zum einen so viele Arbeitsbezirke zu erstellen, dass die Anzahl der zu erwartenden Personen insgesamt ausreichend groß war, um die vorgegebene Nettostichprobe für Berlin und Brandenburg zu erreichen. Zum anderen sollten möglichst kleinräumige Cluster gebildet werden, um den Erhebungsbeauftragten lange Wege zwischen den Anschriften zu ersparen.

Zunächst wurden für Berlin zwölf und für Brandenburg zehn Arbeitsbezirke erstellt. In Berlin wurden damit 1921 gemeldete Personen an 70 Anschriften und in Brandenburg 1476 gemeldete Personen an 126 Anschriften erwartet. In einer ersten Einteilung waren nur rund 100 zu erwartende Personen je Arbeitsbezirk vorgesehen. Da sich aber in der Rekrutierungsphase abzeichnete, dass nicht ausreichend Interviewende zur Verfügung stehen werden, wurde die Anzahl der Personen je Arbeitsbezirk auf 150 erhöht.

Als sich während der Feldphase des Pretests zeigte, dass die Responserate mit rund 18 % teilnehmenden Personen zu gering war, um die geplante Nettostichprobe in Berlin und Brandenburg zu erreichen, wurden weitere Arbeitsbezirke eingeteilt: für Berlin sieben Bezirke mit 1129 gemeldeten Personen an 34 Anschriften und für Brandenburg

b | Ablauf der Erhebung im Pretest



² https://www.statistik-berlin-brandenburg.de/zensus2021/pdf/Mantelbogen_Pretest.pdf

³ https://www.statistik-berlin-brandenburg.de/zensus2021/pdf/Haushaltsbogen_Pretest.pdf

⁴ Ein Arbeitsbezirk definiert das Einsatzgebiet mit einer Anzahl an Anschriften, an denen ein Erhebungsbeauftragter in Summe circa 150 Personen anhand des Melderegisters erwarten konnte.

fünf Arbeitsbezirke mit 747 gemeldeten Personen an 44 Anschriften. Hierfür wurden Anschriften aus einem erweiterten Stichprobenanschriftenbestand genutzt, der vom Statistischen Bundesamt für eine Vorgehensweise ohne monetäre Anreize übermittelt wurde. Insgesamt umfasste der Pretest damit über 5 000 Personen an mehr als 250 Anschriften (Tabelle 1).

Erstellen der Erhebungsunterlagen

Im Zuge der Erhebung kamen viele Unterlagen zum Einsatz. Hierzu gehörte ein Mantelbogen je Anschrift, mindestens ein Haushaltsbogen, Infolyer und Terminankündigungsschreiben je Haushalt sowie ein Anschreiben für die Online-Befragung (IDEV) pro Person und als Alternative dazu auch ein Papierfragebogen.

Die kompletten Unterlagen mussten bereits zur Schulung der Erhebungsbeauftragten bereitgestellt werden, um einen zusätzlichen Anfahrtsweg zu vermeiden. Dies hatte zur Folge, dass die Druckauflage für die Arbeitsbezirke überschlagen werden musste, da die genaue Anzahl der Haushalte und Personen im Voraus nicht bekannt war (Tabelle 2). Für die Schätzung wurde die Anzahl der an den Anschriften gemeldeten Personen und darauf aufbauend die erwartete Anzahl der Haushalte herangezogen und um einen Sicherheitsaufschlag ergänzt.

Während das Statistische Bundesamt die Infolyer für die Statistischen Ämter der Länder bereitstellte, wurden die Haushaltsbogen und Terminankündigungsschreiben des AfS von einem externen Dienstleister gedruckt. Der Druck aller übrigen Unterlagen erfolgte hausintern. Hierfür mussten die Musterunterlagen länderspezifisch für Berlin und Brandenburg angepasst werden.

Gewinnung von Erhebungsbeauftragten

Für die Durchführung des Pretests wurde eine ausreichende Anzahl an Erhebungsbeauftragten benötigt. Um die kalkulierten 30 Freiwilligen zu gewinnen, wurde zunächst hausintern die Bereitschaft dazu abgefragt. Aufgrund der schlechten Resonanz wurden auch externe Strategien zur Gewinnung entwickelt. Zunächst wurde erörtert, welche Zielgruppen primär angesprochen werden sollten. Aufgrund ihrer hohen Flexibilität wurden Studierende als potenzielle Interviewende adressiert. Ein im AfS erstelltes Plakat wurde an ausgewählten Berliner und Branden-

burger Universitäten und Fachhochschulen sowie in Studierendenwohnheimen in Potsdam ausgehängt.

Da auch durch diese Maßnahme der Rücklauf nicht ausreichend war, wurden weitere mögliche Orte für Aushänge recherchiert. Als geeignet wurden öffentliche Einrichtungen erachtet, die im Umkreis des Erhebungsgebiets lagen, in denen der Bedarf an Interviewenden am größten war. In Absprache mit den Zuständigen der Einrichtungen vor Ort wurden im Verlauf des Pretests im Berliner Bezirk Lichtenberg (Rathaus, Bezirksamt, Finanzamt) und in Potsdam (Rathaus, Jobcenter) weitere Aushänge angebracht. Um die Suche noch zusätzlich auszuweiten, wurden in verschiedenen Potsdamer Kultur- und Bürgerzentren Aushänge platziert.

Bis zum Ende des Pretests erwies sich die Rekrutierung der Erhebungsbeauftragten vor allem im Land Brandenburg als herausfordernd. Im Hinblick auf die Haupterhebung ist es daher erforderlich, die Suche über viele verschiedene Kanäle zu streuen. Auch die Nutzung von Social-Media-Kanälen sollte ernsthaft in Betracht gezogen werden, um die benötigte große Anzahl an Erhebungsbeauftragten gewinnen zu können. Die Schwierigkeiten bei der Akquirierung für den Pretest haben verdeutlicht, dass das Anwerben über die klassischen Wege für die Haupterhebung 2021 nicht ausreichend sein wird.

Packen der Taschen

Um die Erhebung schnellstmöglich zu starten und den Aufwand für alle Beteiligten gering zu halten, wurden die Erhebungsunterlagen direkt nach der Schulung an die Interviewenden übergeben. Bereits im Vorfeld der Schulung wurden die Taschen mit den entsprechenden Erhebungsunterlagen gepackt. Die dafür beschafften Notebook-Taschen wurden mit den benötigten Erhebungsunterlagen sowie Zubehör ausgestattet und zur Gewährleistung des Datenschutzes um ein Zahlenschloss ergänzt.

Jede Tasche sollte genügend Erhebungsunterlagen bereitstellen, gleichzeitig aber nicht zu viele, da dies sowohl die Übersichtlichkeit als auch den Tragekomfort beeinträchtigte. Insgesamt waren beispielsweise circa 90 bis 120 Haushaltsankündigungsschreiben in einer Erhebungstasche. Um die Übersichtlichkeit über die verschiedenen Dokumente zu gewährleisten, wurden diese in unterschiedlich farbige Folien platziert. Dieses Vorgehen, beziehungsweise die Anordnung der Erhebungsunterlagen, wurde von

1 | Übersicht der Arbeitsbezirke in Berlin und Brandenburg

Bundesland	Arbeitsbezirke	Anzahl Anschriften	Anzahl gemeldeter Personen
Berlin.....	19	104	3 050
Brandenburg..	15	170	2 223
Insgesamt	34	274	5 273

Quelle: eigene Berechnung

2 | Mengenangabe zu den gedruckten Unterlagen

Art der Unterlage	Anzahl Berlin	Anzahl Brandenburg	Insgesamt
Mantelbogen.....	104	170	274
Haushaltsbogen.....	2 072	1 608	3 680
Papierfragebogen.....	100	80	180
Infolyer.....	3 000	2 200	5 200
Anschreiben Online-Befragung.....	1 700	1 400	3 100
Erstterminankündigung.....	1 450	1 580	3 030
Zweitterminankündigung..	1 040	860	1 900

Quelle: eigene Berechnung

den meisten Interviewenden als vorteilhaft beurteilt und erwies sich in der Erhebungsdurchführung als äußerst nützlich.

Schulung

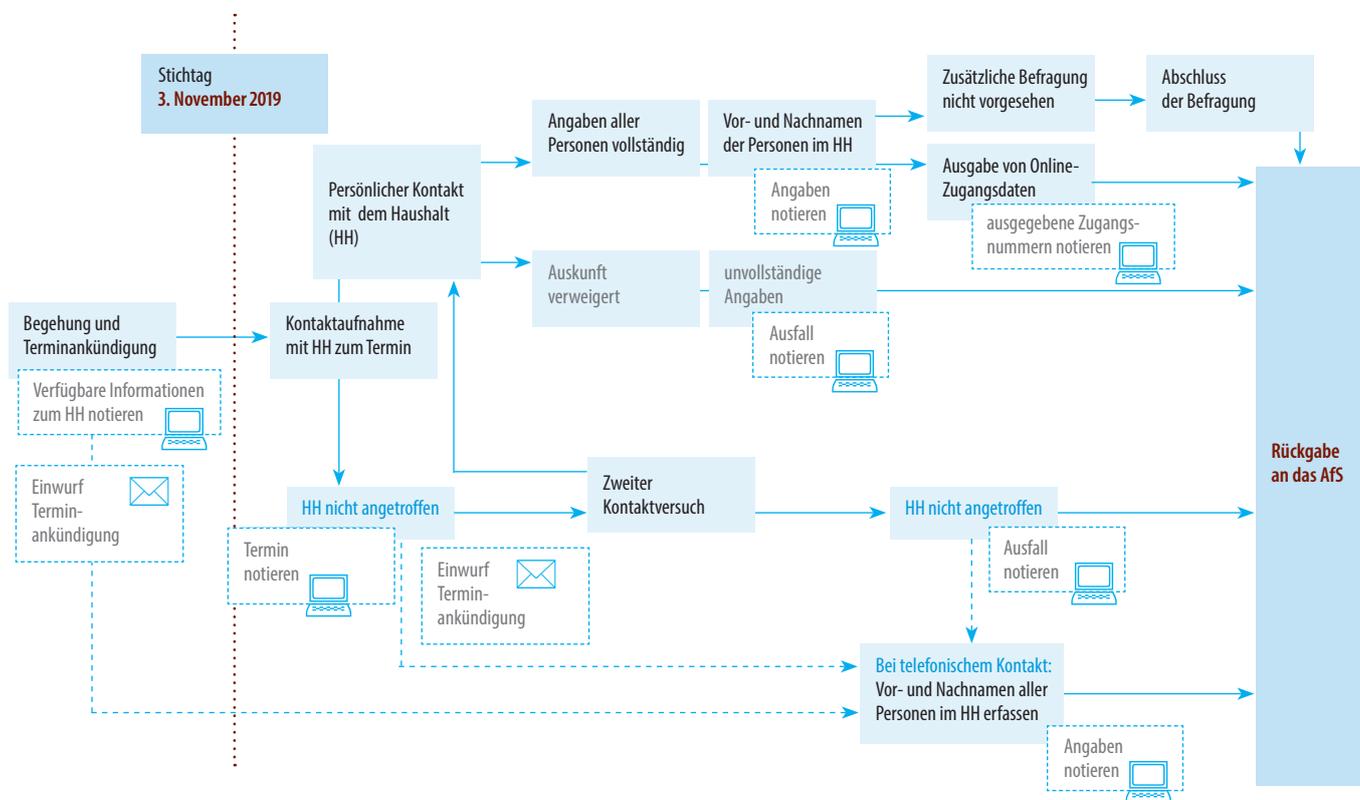
Grundlegendes Ziel der Schulung war es, die rekrutierten Erhebungsbeauftragten auf ihre Tätigkeit vorzubereiten. Zentraler Aspekt der Schulungsvorbereitung war die Terminfindung. Dabei sollten möglichst wenige Schulungstermine stattfinden, um ressourcensparend agieren zu können. Insgesamt fanden zwischen November 2019 und Februar 2020 fünf Schulungstermine mit insgesamt 20 Teilnehmenden statt.

Neben der organisatorischen Vorbereitung wurden die Schulungen auch inhaltlich vorbereitet. Hierzu wurde den Statistischen Ämtern der Länder eine durch die UAG Schulung⁵ erstellte Power-Point-Präsentation im Oktober 2019 zur Verfügung gestellt. Die Vorlage enthielt alle zu schulenden Aspekte, wurde jedoch punktuell zusammengefasst, umgestellt und durch länderspezifische Gegebenheiten ergänzt. Der Einführungsteil stellte den Zensus allgemein vor. Zugrundeliegende Fragestellungen waren, warum und wie der Zensus durchgeführt wird und welche Ziele dabei verfolgt werden. Der zweite

Abschnitt behandelte die rechtlichen Rahmenbedingungen und Datenschutzbestimmungen. Dabei wurde insbesondere auf die einzuhaltende Geheimhaltung hingewiesen. Um die Erhebung praxisorientiert und vor allem in der Gesamtheit darzustellen, wurde diese ab dem zweiten Schulungstermin in einem kurzen Rollenspiel dargestellt. Dieser Part wurde von den Teilnehmenden als sehr hilfreich erachtet. Im Anschluss wurde auf die Einzelheiten der Erhebung eingegangen, zunächst mit Blick auf die Erhebungsunterlagen. Dort wurden schwerpunktmäßig Mantel- und Haushaltsbogen besprochen. Danach wurde der Ablauf der Erhebung geschult. Wie Abbildung c zeigt, gehörte dazu die eigentliche Begehung der Adressen und das Ausfüllen der Erhebungsinstrumente. Des Weiteren wurde die Kontaktaufnahme zu den auskunftgebenden Personen, die persönliche Befragung sowie die Ausgabe von Zugangsdaten zur Online-Befragung erklärt. Zuletzt wurden organisatorische Aspekte, wie die Rückgabe der Unterlagen und Kommunikationskanäle, besprochen.

Im Nachgang der Schulung wurden die Teilnehmenden per Bestellsurkunde zu ehrenamtlichen Erhebungsbeauftragten und damit zur Einhaltung des Statistik- und Datengeheimnisses laut §§ 14 und 16 BstatG⁶ verpflichtet. In diesem Zuge wurde ihnen

c | In der Schulung genutzte Darstellung zum Ablauf der Erhebung



Quelle: Statistisches Bundesamt

⁵ Die UAG Schulung ist eine Unterarbeitsgruppe des Teilprojektes „Personenerhebung: Konzeption und Aufbereitung für den Zensus 2021“. Mitglieder

sind das Statistische Bundesamt, die Statistischen Ämter der Länder Berlin und Brandenburg sowie Baden-Württemberg (Stand Oktober 2019).

⁶ Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I

S. 2394), das zuletzt durch Artikel 10 Absatz 5 des Gesetzes vom 30. Oktober 2017 (BGBl. I S. 3618) geändert worden ist.

der Ausweis für ihre Tätigkeit ausgegeben, mit dem Hinweis, dass dieser ausschließlich für die Arbeit als Erhebungsbeauftragte verwendet werden darf. Mit der erfolgten Verpflichtung konnten anschließend die Erhebungsunterlagen überreicht werden. Da die Anzahl der Teilnehmenden in den Schulungen recht klein war, konnten die Unterlagen teilweise einzeln besprochen werden. Im Zuge dessen wurden auch Verständnisfragen oder Unklarheiten geklärt.

Erhebungsdurchführung Öffentlichkeitsarbeit

Es war von hoher Relevanz, die Bürgermeisterinnen und Bürgermeister bzw. Landräte der betreffenden Ortsteile in Berlin (Friedrichsfelde, Schöneberg) und Brandenburg (Potsdam, Hoppegarten) über den Pretest in Kenntnis zu setzen. Diese wurden im Namen des Vorstands des AfS über die rechtlichen und inhaltlichen Rahmenbedingungen des Pretests informiert, um auch auf eventuelle Nachfragen der Bürgerinnen und Bürger entsprechend reagieren zu können. Im Zuge der Öffentlichkeitsarbeit wurde auch die zuständige Polizei informiert, für den Fall, dass Personen ihr Misstrauen bezüglich der Erhebungstätigkeit melden.

Auftakt der Öffentlichkeitsarbeit im Rahmen des Pretests war die Landespressekonferenz im Dezember 2019. Diese diente vor allem der Vorstellung der Statistischen Jahrbücher 2019 für die Länder Berlin und Brandenburg. In der Pressemappe für die anwesenden Medienvertreterinnen und -vertreter befand sich auch der Flyer zur Interviewenden-Suche. Dieser wurde in einem Bericht der Potsdamer Neueste Nachrichten aufgegriffen. Die Zeitung berichtete neben den Erkenntnissen aus den Jahrbüchern 2019 auch darüber, dass Interviewende für die Monate Januar und Februar 2020 in Potsdam, Hoppegarten und Berlin-Schöneberg gesucht wurden. Diese Veröffentlichung hatte durchschlagenden Erfolg: Viele Personen bekundeten ihr Interesse an der Tätigkeit, von denen sechs als Erhebungsbeauftragte verpflichtet werden konnten.

Da Anfang Februar 2020 abzusehen war, dass die Nettostichprobe für Brandenburg mit den bereits rekrutierten Erhebungsbeauftragten voraussichtlich nicht zu erreichen war, wurde nochmals auf unterschiedlichen Wegen um diese geworben. Zum einen wurde eine Pressemitteilung (Nr. 27 vom 07.02.2020) erstellt, in der sowohl der Zensus 2021 als auch der Pretest dargestellt wurden. Gleichzeitig wurden Personen, die an einer Tätigkeit als Interviewende in Potsdam interessiert waren, dazu aufgerufen sich zu melden. Diese Interviewenden-Suche wurde auch im Intranet der Verwaltungen des Landes Brandenburg veröffentlicht.

Am 10. Februar erreichte das AfS eine Anfrage der Märkischen Allgemeinen (MAZ), mit der Bitte um Hintergrundinformationen zum Pretest und zum Zensus 2021. Ein MAZ-Redakteur wurde auf das Thema aufmerksam, da er selbst in Potsdam an seiner privaten Anschrift befragt wurde. Infolgedessen befragte er den Erhebungsbeauftragten zu seiner Tätigkeit, welcher an das AfS verwies. Nachdem

die inhaltlichen Fragen geklärt waren, erschien am 12. Februar 2020 in der MAZ ein Artikel mit dem Titel „Testlauf für die Volkszählung“.

Nur wenige Tage später stellte der Radiosender Antenne Brandenburg eine O-Ton-Anfrage zur Vorbereitung des Zensus 2021. Diesbezüglich offene Punkte zu Zeitpunkt, Zweck und Auswahl der auskunftgebenden Personen wurden in einem Telefonat mit dem AfS geklärt.

Kommunikation mit Auskunftgebenden – Hotline und CATI

Während des Pretests wurde für die Länder Berlin und Brandenburg jeweils eine telefonische Hotline geschaltet. Diese fungierte als Informationsstelle für auskunftgebende Personen. Die Hotlines für Berlin und Brandenburg wurden ausgiebig genutzt. Die Berliner Hotline verzeichnete während der Erhebung insgesamt 40 Anrufe, die Brandenburger Nummer wurde 42-mal gewählt. Der häufigste Grund für den Anruf der Auskunftgebenden war die Verweigerung der Teilnahme am Pretest. Ebenfalls häufig erfolgte eine Kontaktaufnahme, um darüber zu informieren, dass Personen, die an Stichprobenanschriften lebten, zum vorgeschlagenen Termin nicht zu Hause waren. Einige Anrufe bezogen sich auf den Vorgang der Online-Meldung. Auch wurde um neue Zugangsdaten gebeten oder Personen berichteten in seltenen Fällen von nicht funktionierenden Online-Kennungen. Ein kleiner Teil der Anrufe bezog sich auf die allgemeine Seriosität der Anstreben und der Erhebung. Weiterhin wurden den Erhebungsablauf betreffende Informationen, wie beispielsweise Sterbefälle an gezogenen Anschriften, übermittelt.

Neben der Option der telefonischen Kontaktaufnahme wurden auch zwei E-Mail-Adressen während der Erhebung kommuniziert. Diese wurden ebenfalls intensiv genutzt. Häufige Themen waren hier die nicht gewünschte Teilnahme oder die Vereinbarung eines neuen Termins. Nach erfolgtem Einverständnis der Auskunftgebenden wurde die betreffende E-Mail an den zuständigen Erhebungsbeauftragten weitergeleitet, welcher die Kommunikation übernahm. Auch Online-Zugangsdaten wurden per E-Mail angefragt. Insgesamt verzeichnete die Berliner E-Mail-Adresse 35 E-Mails, die Brandenburger E-Mail-Adresse erhielt 33 Anfragen.

Für Auskunftgebende gab es zusätzlich über die Hotline die Möglichkeit einer telefonischen Auskunftserteilung (CATI). Nur auf explizite Nachfrage wurde die Möglichkeit eines Telefoninterviews mit dem AfS angeboten. Während der gesamten Dauer des Pretests wurde jedoch kein CATI wahrgenommen.

Die Kommunikation mit den Auskunftgebenden war in den überwiegenden Fällen positiv und konstruktiv. Auch bei angekündigten Verweigerungen der Teilnahme war die Stimmung freundlich und aufgeschlossen. Einigen Anrufenden war jedoch eine allgemeine Ablehnungshaltung anzumerken, die nach Erläuterungen zum Zensus, dem Pretest und seinen Zwecken größtenteils abgemildert werden konnte. Zudem gab es Personen, die sich mit einem

Anruf vom amtlichen Charakter der Erhebung und der Seriosität überzeugen wollten. Diese Telefonate wurden ebenfalls für allgemeine Erklärungen zum Zensus und zum Pretest genutzt.

Betreuung der Erhebungsbeauftragten

Während der jeweiligen Feldphase bestand kein regelmäßiger Kontakt zwischen dem AfS und den Erhebungsbeauftragten, insbesondere um das selbstständige Arbeiten zu ermöglichen. Die insgesamt 21 rekrutierten Interviewenden, von denen später 17 im Einsatz waren, wurden gleichwohl umfassend betreut. Bei eventuellen Fragen oder Problemen gab es die Möglichkeit, jederzeit Rücksprache mit dem AfS zu halten. Eine entsprechende Hotline wurde für diesen Zweck bereitgestellt.

Nach der Abarbeitung ihrer Arbeitsbezirke wurde die vollständige und unverzügliche Rückgabe der Erhebungsunterlagen organisiert. Alle Unterlagen wurden fristgerecht abgegeben, sodass keine Erhebungsbeauftragten diesbezüglich gemahnt werden mussten. Ein Sonderfall dazu wird im Abschnitt „Problemfälle“ näher erläutert. Um alle Datenschutzbestimmungen einzuhalten, erfolgte die Rückgabe ausschließlich persönlich; die Unterlagen wurden dabei auf Vollständigkeit, Vollzähligkeit und Wertbarkeit überprüft.

Während der Rückgabe wurden die im Haushaltsbogen verzeichneten erfolgreich befragten Personen gezählt, vermerkt und den Interviewenden mitgeteilt. So konnten diese bereits einschätzen, wie hoch ihre Aufwandsentschädigung ausfallen würde. Abschließend füllten sie einen Evaluationsfragebogen aus.

Mit Beginn der Rekrutierung von Erhebungsbeauftragten ab Oktober 2019 wurde eine durchgängige Erreichbarkeit für diese über eine Hotline sowie E-Mail-Funktionspostfächer gewährleistet. Die drei primären Kontaktgründe waren:

- Fragen, die während der Pretest-Durchführung aufkamen, beispielsweise zu Anschriften, Sonderfällen oder Rückfragen der auskunftgebenden Personen
- Organisatorisches, wie beispielsweise der Ausfall des Erhebungsbeauftragten (kurz- oder langfristig)
- die Nachforderung von Unterlagen und Materialien für die Erhebung.

Einige Interviewende meldeten sich während ihrer Tätigkeit gar nicht, da sie zum einen keine Fragen und zum anderen ihre Kontaktdaten an die Auskunftgebenden weitergegeben hatten und die Korrespondenz mit diesen selbstständig führten. Andere nahmen einen größeren Betreuungsumfang in Anspruch und meldeten sich des Öfteren sowohl telefonisch als auch per E-Mail. Fehlende Unterlagen konnten per Post übersandt oder persönlich nach Absprache im AfS abgeholt werden.

Kontakt mit den Erhebungsbeauftragten erfolgte meist aus zwei Gründen:

- Übermittlung der Kontaktdaten derjenigen auskunftgebenden Personen, deren Absage weitergeleitet werden musste, sowie
- zur individuellen Terminvereinbarung für die Rückgabe der Erhebungsunterlagen.

Problemfälle

Von 21 rekrutierten Erhebungsbeauftragten zogen vier aus unterschiedlichen Gründen ihre Teilnahme an der Tätigkeit zurück. Diese waren unter anderem:

- keine Erlaubnis zur Nebenbeschäftigung durch den Arbeitgeber
- Überlastung aufgrund der hauptberuflichen Tätigkeit
- Enttäuschung über die Teilnahmebereitschaft der Auskunftgebenden
- lange Fahrtwege.

Mit einer Person brach der Kontakt nach erfolgreich absolvierter Schulung und Verpflichtung vollständig ab. Sämtliche Kontaktversuche via Telefon und E-Mail wurden ignoriert, sodass die Rechtsabteilung des AfS schließlich in den Fall involviert wurde. In enger Zusammenarbeit wurde ein Schreiben mit Aufzählung der gescheiterten Kontaktversuche sowie der Bitte um Herausgabe der Erhebungsunterlagen postalisch zugestellt. Als auch dieses ignoriert wurde, wurden zwei Postzustellungsurkunden (PZU) zugestellt; zunächst mit der Bitte um Rückgabe der Tasche mit allen Erhebungsunterlagen samt Erhebungsbeauftragten-Ausweis mit Fristsetzung, da sonst rechtliche Schritte eingeleitet werden würden. Persönlich meldete sich die Person nicht, kurz vor Fristende wurde die Tasche dann schließlich vor die Tür der Poststelle am Standort Alt-Friedrichsfelde abgestellt. Der Ausweis fehlte jedoch, sodass die zweite PZU folgte, mit der dringlichen Bitte um Herausgabe. Der Ausweis wurde anschließend postalisch übermittelt. Mit der Erhebungstätigkeit wurde nie angefangen, sodass der Arbeitsbezirk weiter verteilt werden konnte.

Ausgehend vom Pretest kann davon ausgegangen werden, dass der Betreuungsbedarf für die Erhebungsbeauftragten während der Haupterhebung sehr arbeits- und zeitaufwendig sein wird. Empfehlenswert ist es, ein einheitliches Verfahren für den Umgang mit dem Abbruch der Erhebungstätigkeit zu entwickeln. Dies gilt insbesondere, wenn nicht bekannt ist, ob bereits mit der Erhebungstätigkeit angefangen wurde.

Vergütungsmodell/Abrechnung

Die Erhebungsbeauftragten erhielten für ihre ehrenamtliche Tätigkeit im Rahmen des Pretests eine Aufwandsentschädigung. Diese setzte sich aus einem pauschalen einmaligen Sockelbetrag in Höhe von 150,00 EUR für Schulung, Fahrtkosten und Begehung der Anschriften sowie aus zusätzlichen 5,00 EUR pro erfolgreich befragter Person zusammen. Bei Bearbeitung eines weiteren Arbeitsbezirks wurde der Sockelbetrag nochmals ausgezahlt. Antwortausfälle wurden nicht berücksichtigt. Das Modell, welches möglichst einfach und übersichtlich gehalten werden sollte, wurde im Vorfeld mit den Rechtsaufsichten des AfS abgesprochen. Die Abrechnung erfolgte im wöchentlichen Rhythmus.

Erinnerungsgeschäft

Mit Rückgabe der Unterlagen durch die Erhebungsbeauftragten wurde ersichtlich, welche Personen nicht an der Befragung teilgenommen hatten. Da-

raufhin wurden die auskunftgebenden Personen mithilfe von Erinnerungsschreiben an die Teilnahme der Online-Befragung erinnert. Zum einen wurden jene adressiert, die an der Ziel-1-Befragung teilgenommen hatten und IDEV-Zugänge für die Ziel-2-Befragung entgegennahmen, aber diese nach mehr als 14 Tagen immer noch nicht durchgeführt hatten (Fall 1). Zum anderen wurden die Haushalte angeschrieben, die von den Interviewenden auch beim zweiten Kontaktversuch nicht angetroffen wurden (Fall 2).

Im ersten Fall konnten die Auskunftgebenden personenscharf angesprochen und mit einem persönlichen IDEV-Zugang beschickt werden. Da für den zweiten Fall weder Vornamen noch Personenanzahl bekannt waren, wurde festgelegt, den gesamten Haushalt zu adressieren und vier IDEV-Kennungen beizulegen.

Da die Unterlagen zu den fertig bearbeiteten Arbeitsbezirken von den Erhebungsbeauftragten zwischen Dezember 2019 und März 2020 eintrafen, wurden die Erinnerungen in mehreren Versandwellen verschickt. Dadurch lagen die Befragung durch die Erhebungsbeauftragten sowie die Erinnerung zur Teilnahme an der Online-Befragung zeitlich nicht zu weit auseinander. Insgesamt wurden 433 Einzelpersonen und 300 Haushalte mit Erinnerungsschreiben beschickt.

3 | Erinnerungsgeschäft in Berlin und Brandenburg

Rücklauf Erinnerungsschreiben	Erinnerte Personen	Teilnahmen	%
Fall 1	433	88	20
Fall 2	300	24	8

Quelle: eigene Berechnung

Die Erinnerungsschreiben im Fall 1 wurden von den Auskunftgebenden gut angenommen. Wie Tabelle 3 zeigt, beteiligte sich jede/r Fünfte daraufhin doch noch an der Online-Befragung. Im zweiten Fall war der Rücklauf erwartungsgemäß geringer. Aus den 300 Haushalten nahmen 24 Personen

an der Befragung teil. Der Aufwand für das Erinnerungsverfahren lohnte sich dennoch, da durch dieses insgesamt 112 weitere Teilnehmende gewonnen werden konnten.

Erhebungsnachbereitung Kosten

Da die Teilnahme am Pretest der Personenerhebung freiwillig war, wurde jeder Person, die sich an der Online-Befragung beteiligte, ein Incentive in Höhe von 7,50 EUR ausbezahlt. Am Ende des IDEV-Formulars wurde daher nach der Kontoverbindung gefragt. Die beim Statistischen Bundesamt eingegangenen Informationen wurden an die entsprechenden Statistischen Ämter der Länder weitergeleitet. Diese Datensätze wurden zunächst auf die zur Auszahlung wesentlichen Informationen reduziert und dann auf mögliche Doppelungen und fehlende Angaben überprüft. Die bereinigten Datensätze wurden im Weiteren an das Finanzmanagement des AfS übergeben, welches die Auszahlung veranlasste.

Nicht wenige Auskunftgebende verzichteten auf die Auszahlung. Sie gaben keine Kontoverbindung

an. Dies berichteten auch einige Erhebungsbeauftragte, welche die Ziel-2-Befragung persönlich vor Ort vorgenommen hatten. Um auch hier eine Auszahlung gewährleisten zu können, wurden an die Papierfragebogen Dokumentationsblätter zur Angabe der Kontoverbindung angehängt. Viele der persönlich Befragten nutzten diese jedoch nicht.

In einigen Fällen wurden falsche oder nicht vollständige Kontoverbindungen übermittelt. Da die Anschriften dieser Personen bekannt waren, wurden diese postalisch angeschrieben, mit der Bitte die korrekte IBAN mitzuteilen, sofern eine Auszahlung des Betrages gewünscht wird. Insgesamt wurden 86 Personen bzw. Familien angeschrieben, in 46 Fällen konnte aufgrund einer Rückmeldung die Überweisung vorgenommen werden.

Auswertung Anzahl der Teilnahmen

Parallel zur Erhebungsphase fand die Fortschrittsverfolgung und Auswertung der digitalen und analogen Rückläufe statt. Anfang April konnte der folgende finale Stand festgehalten werden:

Von den 34 erstellten Arbeitsbezirken wurden letztendlich 31 erfolgreich von den Erhebungsbeauftragten bearbeitet, wobei sie 252 Anschriften mit einer nach dem Melderegister erwarteten Personenanzahl von 4825 Personen begingen. Dabei konnten 2528 Haushalte dokumentiert werden, von denen 776 am Pretest teilnahmen. In diesen Haushalten nahmen wiederum 1554 Personen an Teil 1 der Befragung teil.

38 von den erfolgreich Befragten wohnten an einer Ziel-1-Anschrift und hatten mit Teil 1 ihre Befragung direkt gegenüber den Interviewenden vollständig abgeschlossen. 50 Personen beantworteten Teil 2 mit einem Papierfragebogen. Die Mehrheit der Teilnehmenden entschied sich dafür, den zweiten Teil selbstständig online auszufüllen. Zu diesem Zweck wurden 1449 individuelle IDEV-Kennungen für die Online-Befragung herausgegeben. Allerdings wurden von diesen nur 814 genutzt, was einer Aktivierungsquote von 57,2% entspricht.

Addiert ergibt sich aus allen drei Teilnahmemodi (38 nur Ziel 1+ 50 PAPI/PAP + 814 CAWI) eine Gesamtteilnahmezahl von 902 Personen, womit in Berlin als auch in Brandenburg die Zielvorgabe des Statistischen Bundesamtes erreicht werden konnte (Tabelle 4).

Die Vorgabe konnte jedoch nur eingehalten werden, weil an erheblich mehr Anschriften erhoben wurde als im Voraus geplant. Da sich statt eines ge-

4 | Erhebungserfolge in Berlin und Brandenburg

Bundesland	Vorgabe des Statistischen Bundesamtes	Vollständige Rückläufer	%	potenziell erreichbare Personen ¹	Rücklaufquote in %
Berlin.....	433	526	121,0	2 722	19,3
Brandenburg...	302	376	125,0	2 103	17,9
Insgesamt	735	902	123,0	4 825	18,7

1 nach Melderegister

Quelle: eigene Berechnung

planten Rücklaufs von 24 % nur eine geringe Teilnahmebereitschaft von zwischenzeitlich 15 % einstellte, musste der schwache Rücklauf durch eine Korrektur kompensiert werden. Dafür wurde im Laufe des Pretests die Erhebung von 22 auf 31 Arbeitsbezirke ausgeweitet.

In Folge dieser Aufstockung zeigte sich, dass bei wiederholt eingesetzten Interviewenden eine Professionalisierung erfolgte. Zum einen arbeiteten sie schneller, zum anderen war die Teilnahmebereitschaft der auskunftgebenden Personen etwas höher. Daraus ergab sich, dass die Vorgabe des Statistischen Bundesamtes um mehr als 20 % übertroffen wurde.

Evaluation

Während der Rückgabe der Erhebungsunterlagen durch die Erhebungsbeauftragten wurden diese um eine mündliche und schriftliche Evaluation gebeten. Dadurch konnte Feedback eingeholt werden, das Probleme im Ablauf der Erhebung und Einschätzungen zu den Rahmenbedingungen der Tätigkeit umfasste.

Beispielsweise wurde berichtet, dass im Erhebungsablauf bei einigen auskunftgebenden Personen eine erhöhte Skepsis gegenüber einer Befragung durch den Staat zu erkennen war. Zudem bestand ein hoher Erklärungsbedarf darüber, was der Zensus 2021 sei. Hier zeigte sich, dass es im Vorlauf der Haupterhebung bereits frühzeitig Aufklärung mittels einer breit angelegten Öffentlichkeitsarbeit geben muss.

Die meisten Interviewenden waren mit der Anzahl und Lage ihrer zugewiesenen Anschriften zufrieden, dennoch schätzten etwas mehr als die Hälfte den zeitlichen Aufwand gegenüber der Aufwandsentschädigung als zu hoch ein.

Ausblick und Anpassungen für die Haupterhebung

Insgesamt verlief der Pretest durch den engagierten Einsatz der Mitarbeiterinnen und Mitarbeiter des AfS und der Erhebungsbeauftragten sehr erfolgreich. Zum einen wurden die Zielvorgaben für Berlin und Brandenburg vollständig erfüllt, zum anderen wurden wertvolle Erkenntnisse im Hinblick auf die Haupterhebung gesammelt.

Durch die Teilnahme am bundesweit durchgeführten Pretest war es unter anderem möglich, die Arbeitsutensilien der Erhebungsbeauftragten nicht nur unter „Laborbedingungen“ in den Statistischen Ämtern der Länder zu erproben, sondern auch die Tauglichkeit der verwendeten Taschen, Mappen oder Schreibgeräte unter Echtbedingungen im Feld zu testen. Nichtsdestotrotz werden die großen Mengen an zu beschaffenden Unterlagen, die Schulungsvorbereitungen im Allgemeinen und eine klar strukturierte Über- und Rückgabe der Dokumente die Mitarbeiterinnen und Mitarbeiter der Erhebungsstellen vor immense Herausforderungen stellen. Hier gilt es, die Erfahrungen des Pretests zu nutzen und die Prozesse detailliert und strukturiert vorzubereiten.

Eine verständliche und auf die Bedürfnisse der Erhebungsbeauftragten ausgerichtete Schulung ist essenziell für eine erfolgreiche Durchführung des

Zensus 2021. Wichtiger Erkenntnisgewinn konnte an dieser Stelle durch das qualitative Feedback der Schulungsteilnehmenden gewonnen werden. Insgesamt wurde die 3,5 Stunden andauernde Schulung durchgängig gelobt und als gute Vorbereitung auf die Erhebungstätigkeit angesehen. Die anschließende etwa 20-minütige Übergabe der Unterlagen wurde zur Klärung aufgetretener Verständnisfragen genutzt. Zudem konnte durch den persönlichen Kontakt die Distanz zwischen den Mitarbeitenden des AfS und den Erhebungsbeauftragten minimiert werden, um eine Vertrauensbasis für etwaige Rückfragen zu schaffen. Zum gegenwärtigen Zeitpunkt wird während der Haupterhebung von einer Schulungsdauer von vier bis fünf Stunden unter anderem wegen der größeren Anzahl an Schulungsteilnehmenden gerechnet. Hier gilt es die Inhalte weiter auf ein absolut notwendiges Maß zu reduzieren, um keine negativen Effekte aufgrund sinkender Aufmerksamkeitspannen zu bewirken.

Als wichtiges Kernelement des Pretests sollte das Online-First-Prinzip auf seine Wirksamkeit getestet werden. Hierbei stellte es eine Herausforderung dar, die Vorteile des Verfahrens (Zeit- und Kostenersparnis, geringer Aufwand, diskrete Durchführung, modernes Medium mitsamt Filterführung, Barrierefreiheit sowie die daraus resultierende freie Entscheidung über Zeitpunkt und Ort der Beantwortung der Fragen) zu den Auskunftgebenden zu transportieren. Im Zuge des Testlaufs haben 90 % der Teilnehmenden zum Ende der ersten Befragung Online-Zugangsdaten entgegengenommen. Inwieweit dieser Wert als tatsächliche Auskunftsbereitschaft oder als Vermeidungsstrategie zu interpretieren ist, um einer weiteren Befragung zu entgehen, kann an dieser Stelle nur gemutmaßt werden. In dieser Form nicht zu erwarten war das Verhalten der interviewten Personen, die die Existenzfeststellung mit dem Erhebungsbeauftragten durchführten (Merkmale zur Person und des Wohnstatus werden festgestellt), die Erhebungsunterlagen auch entgegennahmen, diese dann aber nicht ausfüllten. Die hier beobachteten Rücklaufquoten betragen nur etwa 50 %. Auch das Erinnerungsverfahren konnte den Rücklauf nur marginal erhöhen (um etwa 7%). Sicherlich muss hier die Freiwilligkeit der Erhebung berücksichtigt werden, jedoch lassen sich zwei Kernthesen ableiten:

1. Die Inaugenscheinnahme durch die Erhebungsbeauftragten stellt ein unabdingbares Kernelement des Zensus 2021 dar, um Ergebnisse in hinreichender Qualität zu erzielen. In der derzeitigen Ausgestaltung des Verfahrens ist von stark erhöhten Margen im Erinnerungs- und Mahnverfahren bezüglich der Ziel-2-Merkmale auszugehen. Die grundsätzliche Wirksamkeit der Online-First-Strategie für die Gewinnung der sozioökonomischen Merkmale kann jedoch bestätigt werden. Insbesondere Punkt 2 muss in der Arbeitsplanung der Arbeitsprozesse in den Erhebungsstellen Berücksichtigung finden. Weiterhin sollte die Erhebung von einer Informationskampagne begleitet werden, welche unterstützend die Online-Ausrichtung des Verfahrens flankiert.

2. Im Zuge der Vorüberlegungen zu den Projektrisiken des Zensus 2021 wurde die Gewinnung von Erhebungsbeauftragten als ein großes Risiko in der Durchführungsphase ausgemacht. Diese Überlegungen wurden durch den Pretest bestätigt und noch einmal verstärkt. Es benötigte mehrere Anläufe, um die erforderliche (geringe) Anzahl von Interviewenden zu erreichen. Hier zeigte sich, dass eher konservative Methoden der Bekanntmachung wie Aushänge oder Flyer in entsprechenden Institutionen (Hochschulen, Behörden etc.) nur geringe Resonanz zur Folge haben. Am erfolgreichsten war der Zeitungsartikel im Zuge der Landespressekonferenz, aus dem ein hohes Maß an Interesse, Teilnahmebereitschaft und Korrespondenz resultierte. Grundsätzlich sollten die Veröffentlichungskanäle breit und modern gedacht werden. Es wird maßgeblich sein, dass es eine sehr starke Öffentlichkeitsarbeit und breit angelegte Kampagnen, sowohl online als auch offline, geben wird und dies vor allem frühzeitig erfolgt. So kann im Vorfeld bereits Interesse geweckt und auf etwaige Rekrutierungsprobleme frühzeitig reagiert werden.

Wichtige Rückschlüsse konnten zum Aufbau und zur Handhabung vieler Erhebungsinstrumente gewonnen werden. Um einige umgesetzte Anpassungen zu nennen, wird nun beispielsweise die Jahreszahl der Befragungstermine auf den Erhebungsunterlagen vorgedruckt sein. Eine Fragestellung zu den meldepflichtigen Personen wird in der Haupterhebung anders formuliert. Weiterhin wurden Checkboxen auf den Erhebungsunterlagen entfernt oder versetzt, um Arbeitserleichterungen für die Interviewenden zu bewirken. Auch sind Veränderungen im Layout und der Schriftart sowie Formatänderungen vorgenommen worden. Die Akzeptanz der Anschreiben mit Infoflyern war eher gering, viele wurden im Vorfeld entsorgt und vielen auskunftgebenden Per-

sonen war der Begriff „Zensus“ nur wenig oder gar nicht geläufig. Für die Haupterhebung ist es daher essenziell, die Wirkung der Ankündigungsschreiben für eine höhere Wahrnehmung zu verbessern und dadurch die sofortige Teilnahme zu erhöhen (und damit Erinnerungsverfahren zu reduzieren).

Über die Erfahrungen aus den Pretests sollte ein Austausch zwischen den teilnehmenden Statistischen Ämtern der Länder auf einem vom Statistischen Bundesamt für den 31. März und 1. April 2020 geplanten Treffen stattfinden. Aufgrund von COVID-19 wurde entschieden, auf ein persönliches Treffen zu verzichten und alternativ eine leitfadengestützte, schriftliche Abfrage durchzuführen. Darin konnten dem Statistischen Bundesamt die hier geschilderten Vorgehensweisen, Herausforderungen, Erfolge, Probleme und Verbesserungsvorschläge kommuniziert werden.

Als äußerst positiven Nebeneffekt eines vorgelagerten Pretests für ein Großprojekt wie den Zensus 2021 muss die Chance einer zusätzlichen Einarbeitung der neuen Kolleginnen und Kollegen betrachtet werden. Hier konnten Erfahrungen gewonnen werden, die sich voraussichtlich als sehr hilfreich für die Bewältigung der umfangreichen und komplexen zukünftigen Arbeitsprozesse erweisen werden. Hierdurch kann das Erhebungsgeschäft effizienter gestaltet und die Ergebnisqualität erhöht werden. In der Gesamtschau konnte der Pretest einige Verbesserungspotenziale offenbaren, zeigte aber auch, dass die Statistischen Ämter des Bundes und der Länder gut vorbereitet sind, um ein Mammutprojekt wie den Zensus 2021 mit hoher Aktualität und Genauigkeit solide durchzuführen.

Mark Hoferichter ist Teilprojektleiter im Bereich Personenbefragungen, Erhebungsstelle Berlin des Projektes *Zensus 2021* im Amt für Statistik Berlin-Brandenburg.

Ralf Mädler ist Teilprojektleiter im Bereich Personenbefragungen, Erhebungsstellen Brandenburg des Projektes *Zensus 2021* im Amt für Statistik Berlin-Brandenburg.

Natalie Lück und

Laura Tomm-Demski sind Sachbearbeiterinnen im Bereich Personenbefragungen, Erhebungsstellen Brandenburg des Projektes *Zensus 2021* im Amt für Statistik Berlin-Brandenburg.

Ulrike Warnke und **Jonas-Otto Werner** sind Sachbearbeiterin und Sachbearbeiter im Bereich Personenbefragungen, Erhebungsstelle Berlin des Projektes *Zensus 2021* im Amt für Statistik Berlin-Brandenburg.

Neuerscheinung

▣ „Bildung in Deutschland 2020“ mit Schwerpunktthema zur Digitalisierung

Mit „Bildung in Deutschland 2020“ wurde der inzwischen achte vom Bundesministerium für Bildung und Forschung (BMBF) und der Kultusministerkonferenz (KMK) in Auftrag gegebene Bildungsbericht veröffentlicht. Auf Basis der amtlichen Statistiken sowie sozialwissenschaftlicher Daten und Studien wurde das gesamte Bildungsgeschehen von der vorschulischen Bildung bis zur Weiterbildung im Erwachsenenalter dargestellt und durch eine Analyse der Rahmenbedingungen für den Bildungserwerb vervollständigt. Um dabei der Heterogenität der Bildungslandschaft in Deutschland gerecht zu werden, wurde eine Vielzahl an Ergebnissen auch auf regionaler Ebene präsentiert. Der Bericht wurde von einer unabhängigen Gruppe von Wissenschaftlerinnen und Wissenschaftlern erstellt, die verschiedenen deutschen Forschungseinrichtungen und den Statistischen Ämtern des Bundes und der Länder angehören. Das Amt für Statistik Berlin-Brandenburg vertritt die Statistischen Ämter der Länder in diesem Konsortium.

Der langjährige Trend zu größerer Bildungsbeteiligung und höheren Abschlüssen schlägt sich immer stärker im Bildungsstand der Gesamtbevölkerung nieder: Im Jahr 2018 besaß ein Drittel der Bevölkerung die Hochschulreife, 2008 waren es noch 24 %. Gleichzeitig führte die immer frühere Bildungsbeteiligung im vorschulischen Bereich sowie die steigenden Studierendenzahlen zu einem Anstieg der Bildungsteilnehmenden auf 17,2 Mill. Diese Entwicklung scheint nun an ihre Grenzen zu stoßen: Neben stagnierenden, in einigen Bundesländern sogar rückläufigen Übergangsquoten ans Gymnasium deuten auch sinkende Absolventenquoten mit (Fach-)Hochschulreife (2014: 53 %; 2018: 50 %) darauf hin. Zudem sind die Anzahl und der Anteil an Jugendlichen, die ohne einen Schulabschluss von der Schule abgehen, weiter gestiegen (2014: 5,8 %; 2018: 6,8 %).

Positiv ist, dass sich die Durchlässigkeit im Bildungssystem weiter erhöht hat. Abschlüsse sind weniger an spezifische Schularten gekoppelt und früh zu treffende Bildungsentscheidungen, wozu etwa die Wahl der weiterführenden Schule gehört, können leichter korrigiert werden. Auch die berufliche Ausbildung eröffnet mehr Möglichkeiten der höheren Qualifizierung. Ein Beispiel für diesen Trend insgesamt ist die gestiegene Anzahl der Studienanfängerinnen und -anfänger ohne schulische Studienberechtigung (2000: 0,4 %; 2018: 3,5 %).

Die aktuelle Corona-Pandemie verdeutlicht, welchen hohen Stellenwert Kitas und Schulen für die Sozialisierung haben und wie wichtig der Bereich der Familie jenseits der formalen Bildungsangebote ist. Insbesondere die sozioökonomische Situation der Familie steht im deutlichen Zusammenhang mit den Bildungserfolgen der Kinder. Im Bericht wird diese durch die drei Risikolagen eines niedrigen Bildungsstands der Eltern (Risikolage formal gering qualifizierter Eltern), der elterlichen Erwerbslosigkeit (soziale Risikolage) und der Armutsgefährdung des Haushalts (finanzielle Risikolage) analysiert. 2018 war fast jede bzw. jeder dritte Minderjährige von mindestens einer dieser Risikolagen betroffen.

Das Schwerpunktthema des diesjährigen Berichts ist „Bildung in einer digitalisierten Welt“: Die zunehmende Digitalisierung in allen Bildungsbereichen offenbart die Potenziale, verweist aber zugleich auf noch bestehende Herausforderungen, die die Corona-Pandemie besonders deutlich macht. So ist die technische Infrastruktur an deutschen Schulen auch im internationalen Vergleich weiterhin unterdurchschnittlich. Doch die Technik allein genügt nicht – es fehlt zudem an Konzepten, digitale Technologien an Bildungsinstitutionen didaktisch sinnvoll einzusetzen.



Der Bericht „Bildung in Deutschland 2020“ ist unter www.bildungsbericht.de kostenfrei als PDF-Dokument verfügbar. Über den QR-Code gelangen Sie direkt dahin.

Historisches

Die Spanische Grippe im Deutschen Reich

von Iris Hoßmann-Büttner

Das Jahr 1918 ist von bewegenden Ereignissen geprägt: Nach vier Jahren ging der 1. Weltkrieg mit der Kapitulation der deutschen Armee zu Ende. Kaiser Wilhelm II. wurde am 9. November 1918 zur Abdankung gezwungen; am selben Tag rufen sowohl der Sozialdemokrat Philipp Scheidemann als auch der Kommunist Karl Liebknecht die Republik aus. Der Krieg ist zwei Tage später vorbei, es folgen Straßenschlachten und bürgerkriegsähnliche Zustände in der Hauptstadt. Diese grundlegenden Ereignisse überschatten die öffentliche Wahrnehmung einer Grippewelle, die in die Geschichte als „Spanische Grippe“ einging und sich von 1918 bis 1920 als Pandemie auf der ganzen Welt ausbreitete.¹ Im Deutschen Reich infizierten sich Schätzungen zufolge innerhalb weniger Monate rund 4 Mill. Menschen – jeder Zehnte überlebte die Grippe nicht.²

Belegen lassen sich die Zahlen nur bedingt. Aufgrund der Kriegswirren konnten die Erhebungen der amtlichen Statistik nur in Teilen und nicht in allen Gebieten durchgeführt werden. Dazu kam der Umstand, dass die Krankheit nicht meldepflichtig war. Nur ein Bruchteil der Erkrankten suchte einen Arzt auf – eine Vielzahl war nicht krankenversichert und konnte die Kosten für eine Behandlung nicht aufbringen. Dazu traf der Virus im Jahr 1918 auf eine Bevölkerung in schlechtem Gesundheitszustand: Im vierten Jahr des Krieges waren die Menschen körperlich geschwächt und hatten dem aggressiven Influenza-Virus nichts entgegensetzen.³ Die klassischen Grippesymptome gingen oftmals mit einer Lungenentzündung einher, sodass bei einem tödlichen Ausgang oftmals Lungenentzündung als Todes-

a | Todesfälle 1914 bis 1923 nach Todesursachen und Geschlecht
hier: männlich

Länder und Landesteile	Jahr	II. Bewegung der Bevölkerung																	
		II.1. Todesursachen der in den Jahren 1921 ¹⁾ , 1922 ²⁾																	
		a. Männliches																	
		Es starben																	
		III. Infektionskrankheiten																	
		IV. Urtliche Krankheiten																	
		V. durch gewalttätige Einwirkungen																	
		VI. aus ander. u. unbek. Ursachen																	
		VII. unbek. Ursachen																	

Länder und Landesteile	Jahr	II. Bewegung der Bevölkerung																	
		II.1. Todesursachen der in den Jahren 1921 ¹⁾ , 1922 ²⁾																	
		a. Männliches																	
		Es starben																	
		III. Infektionskrankheiten																	
		IV. Urtliche Krankheiten																	
		V. durch gewalttätige Einwirkungen																	
		VI. aus ander. u. unbek. Ursachen																	
		VII. unbek. Ursachen																	

¹⁾ Ohne die an Polen (ohne Abtinnung), an die Tschechoslowakei, Dänemark, Belgien abgetretenen Gebiete und ohne Mangelgebiet von 20.10.1921 an Polen abgetretenen Teil von Oberschlesien. – ²⁾ Ausschließlich Mecklenburg-Strelitz. – ³⁾ Siehe Ann. 1, ausschließlich Mecklenburg. – ⁴⁾ Ausschließlich bei der Mecklenburg- und dem an Polen abgetretenen Teil der Provinz Posen, ausschließlich der beiden Jahre 1922 eingegangenen nachträglichen ständesmittlichen Beurteilungen. – ⁵⁾ Ausschließlich der beiden Mecklenburg.

¹⁾ Ohne die an Polen (ohne Abtinnung), an die Tschechoslowakei, Dänemark, Belgien abgetretenen Gebiete und ohne Mangelgebiet von 20.10.1921 an Polen abgetretenen Teil von Oberschlesien. – ²⁾ Ausschließlich Mecklenburg-Strelitz. – ³⁾ Siehe Ann. 1, ausschließlich Mecklenburg. – ⁴⁾ Ausschließlich bei der Mecklenburg- und dem an Polen abgetretenen Teil der Provinz Posen, ausschließlich der beiden Jahre 1922 eingegangenen nachträglichen ständesmittlichen Beurteilungen. – ⁵⁾ Ausschließlich der beiden Mecklenburg.

- 1 Den Namen „Spanische Grippe“ erhielt diese schwere Krankheit, weil erste Nachrichten darüber aus dem im 1. Weltkrieg neutralen Spanien kamen. Dieses Land hatte eine liberale Zensur.
- 2 Schneider, Sebastian (2020): Das große Sterben. In: RBB 24 Corona-Blog, https://www.rbb24.de/panorama/thema/2020/coronavirus/beitrag_neu/2020/05/spanische-grippe-gross-berlin-preussen-deutsches-reich-pandemie-rueckblick.html [abgerufen am 17.06.2020].
- 3 Dass es sich um einen Influenza-Virus handelte, wurde erst 1933 entdeckt.

ursache angegeben wurde. Zusätzlich war die Lungenerkrankung Tuberkulose in der Bevölkerung weit verbreitet. Für beide Todesursachen sind neben der Influenza deutlich gestiegene Todesfallzahlen für das Jahr 1918 festzustellen. Ein Impfstoff gegen die Spanische Grippe wurde zu dieser Zeit nicht gefunden und eine Behandlung mit Antibiotika sollte erst zehn Jahre später möglich sein.

Weltweite Ausbreitung des Grippevirus durch 1. Weltkrieg begünstigt

Medizinhistoriker gehen davon aus, dass die Pandemie erstmals im Januar und Februar 1918 unter amerikanischen Soldaten in Kansas ausbrach. Wenige Wochen später traf das Virus mit den amerikanischen Truppentransporten in Frankreich ein und verbreitete sich in alle Himmelsrichtungen. Im März

1918 hatten deutsche Truppen in ihrer Frühjahrsoffensive erhebliche Geländegewinne an der Westfront gemacht. Dabei waren viele britische, französische und amerikanische Soldaten, die bereits infiziert waren, in deutsche Gefangenschaft geraten. Bald grassierte die Grippe auch unter den deutschen Frontsoldaten. Schätzungen gehen von bis zu 500 000 Infizierten aus. Das Zentralnachweisamt für Kriegerverluste und Kriegergräber belegte für das Jahr 1918 rund 378 000 Gefallene. Darunter finden sich gegliedert nach Todesursache 13 000 Influenza-Tote, 16 000 Todesfälle aufgrund von Lungentzündung und 6 000 Militärpersonen, die an Tuberkulose gestorben sind. Ein Jahr zuvor gab es 70 Influenza-Tote und jeweils 5 000 Tote mit der offiziellen Todesursache Lungentzündung bzw. Tuberkulose.⁴ Als die kranken Soldaten

b | Todesfälle 1914 bis 1923 nach Todesursachen und Geschlecht
hier: weiblich

18 II. Bewegung der Bevölkerung
II. Todesursachen der in den Jahren 1921¹⁾, 1922²⁾
b. Weibliches

Länder und Landesteile	Zahl der Gestorbenen ohne Vorgeborene	Es starben														
		I. Entwicklungs-Krankheiten und Folgen der Geburt			II. Infektionskrankheiten								11a. Tuberkulose anderer Organe	11b. Tuberkulose anderer Organe		
		1. 1. Entwicklungs-Krankheiten und Folgen der Geburt	2. 2. Allergische	3a. 3a. Keimheftiger	3b. 3b. Keimheftiger	4. 4. Scharlach	5. 5. Masern und Röteln	6. 6. Diphtherie und Kruppe	7. 7. Keuchhusten	8. 8. Typhus	9. 9. übertragene Tier-Krankheiten	10a. 10a. Rose (Erysipel)			10b. 10b. andere Virusinfektionskrankheiten	11. 11. Tuberkulose d. Lungen (Lungenentzündung)
Deutsches Reich ³⁾	1921 426 716	27 660	55 647	4 373	3 217	818	2 010	2 867	2 516	1 441	8	1 005	3 405	37 156	5 867	540
	1922 436 979	24 205	60 871	3 476	3 578	567	1 304	2 115	3 928	929	10	988	3 054	37 817	5 880	622
	1923 426 263	22 244	61 458	3 646	2 999	475	3 336	2 105	3 486	1 060	15	777	3 071	39 756	6 707	691

¹⁾ bis ³⁾ und ⁴⁾ siehe Anmerkung 1 bis 6 und 8 auf Seite 46/47.

19 II. Bewegung der Bevölkerung
II. Todesursachen der in den Jahren 1921¹⁾, 1922²⁾ und 1923³⁾ Gestorbenen nach Ländern und Landesteilen
Geschlecht
an

Länder und Landesteile	Zahl der Gestorbenen ohne Vorgeborene	Es starben														
		I. Entwicklungs-Krankheiten und Folgen der Geburt			II. Infektionskrankheiten								11a. Tuberkulose anderer Organe	11b. Tuberkulose anderer Organe		
		1. 1. Entwicklungs-Krankheiten und Folgen der Geburt	2. 2. Allergische	3a. 3a. Keimheftiger	3b. 3b. Keimheftiger	4. 4. Scharlach	5. 5. Masern und Röteln	6. 6. Diphtherie und Kruppe	7. 7. Keuchhusten	8. 8. Typhus	9. 9. übertragene Tier-Krankheiten	10a. 10a. Rose (Erysipel)			10b. 10b. andere Virusinfektionskrankheiten	11. 11. Tuberkulose d. Lungen (Lungenentzündung)
Deutsches Reich ³⁾	1921 426 716	27 660	55 647	4 373	3 217	818	2 010	2 867	2 516	1 441	8	1 005	3 405	37 156	5 867	540
	1922 436 979	24 205	60 871	3 476	3 578	567	1 304	2 115	3 928	929	10	988	3 054	37 817	5 880	622
	1923 426 263	22 244	61 458	3 646	2 999	475	3 336	2 105	3 486	1 060	15	777	3 071	39 756	6 707	691

¹⁾ bis ³⁾ und ⁴⁾ siehe Anmerkung 1 bis 6 und 8 auf Seite 46/47.

4 Statistisches Reichsamt (1925): Statistisches Jahrbuch für das Deutsche Reich 1924/25, Berlin.

ab Anfang Juni 1918 zur Erholung in die Heimat verlegt wurden, brachten sie die Erreger mit. Im Deutschen Reich durfte aus strategischen Gründen nicht über Erkrankungen an der Front berichtet werden. Doch als der weitaus größere Teil der Militärpersonen nach dem Kriegsende in die Heimat zurückkehrte und sich der Virus in der Zivilbevölkerung verbreitete, berichteten deutsche Zeitungen über Grippefälle im zivilen Bereich. Allerdings fiel die Berichterstattung sehr zurückhaltend aus und wurde nur als Grippe betitelt – zu groß war das Interesse an den politischen Umwälzungen im Herbst 1918.⁵

Bis 1920 hohe Zahl an Grippetoten

Rund 1,6 Mill. Menschen starben im Jahr 1918 – es sollte die höchste Sterbefallzahl in der Geschichte des Deutschen Reiches sein. Im letzten Jahr des 1. Weltkrieges gab es nochmals viele Gefallene an der Front. Doch die rapide steigenden Todesfälle ab Herbst 1918 erklären die Frontverluste nur bedingt. Besonders im Monat Oktober wurden mit 191 000 überdurchschnittlich viele Verstorbene registriert (Oktober 1917: 81 000). Dies fällt auch zeitlich mit der 2. Grippewelle zusammen. Verließ die erste Welle im Frühsommer 1918 noch relativ harmlos und führte nur selten zum Tode, brachte die 2. Grippewelle im September/Oktober 1918 einen mutierten und weitaus aggressiveren Virus unter die Bevölkerung. Jeder achte

Todesfall (12 % bzw. 188 000) ging auf Influenza zurück – unter den weiblichen Verstorbenen sogar jeder sechste (16 %). Weitere 10 % starben an Lungenentzündung und 8 % an Tuberkulose. Ein Jahr zuvor spielte Influenza mit 7 000 Todesfällen und einem Anteil von 1 % an allen Todesursachen keine Rolle in der Bevölkerung. Auch nach 1918 wurden weitere zwei Jahre erhöhte Zahlen von Grippetoten registriert: Im Jahr 1919 starben 43 000 Menschen an Influenza, 1920 waren es 58 000.⁶ Aufgrund des langen Zeitraums der Grippe-Pandemie und der Vielzahl an Influenza-Todesfällen, die nicht dem Grippevirus zugeordnet wurden, erscheint die Schätzung von rund 400 000 Todesfällen im Deutschen Reich nicht unrealistisch.⁷

Grippe-Pandemie legte weite Teile der Berliner Infrastruktur lahm

Die Stadt Berlin war ebenso wie der Rest des Landes vom Grippevirus überrollt worden – mehr als 3 000 Menschen sollte er den Tod bringen. Weitere 5 000 Menschen starben jeweils an Lungenentzündung und Tuberkulose. Die Allgemeinen Ortskrankenkassen meldeten ab Oktober 1918 fast täglich 2 000 Neuinfektionen, wobei die Erkrankten ohne Krankenversicherung nicht mitgezählt wurden. Am 24. Oktober 1918 wurde an 217 Schulen der Unterricht eingestellt. Ein Drittel der Schülerinnen und Schüler war erkrankt, sodass eine Schulschließung angeordnet wurde.

⁵ Micke, Tristan (2018): Vor 100 Jahren erreichte die Spanische Grippe Deutschland. In: Projekte und Netzwerke Marzahn-Hellersdorf, <https://www.berlin.de/projekte-mh/netzwerke/spaetlese/themen/gesundheits/artikel.716988.php> [abgerufen am 17.06.2020].

⁶ Statistisches Reichsamt (1925): Statistisches Jahrbuch für das Deutsche Reich 1924/25, Berlin.

⁷ Maybaum, Thorsten (2018): Spanische Grippe: Ein Virus – Millionen Tote. In: *Ärztblatt* 1/2018, <https://www.aerzteblatt.de/archiv/197155/Spanische-Grippe-Ein-Virus-Millionen-Tote> [abgerufen am 17.06.2020].

Der Begriff „Grippeferien“ findet sich in damaligen Zeitungsartikeln mit dem Hinweis, welche Schulen wie lange geschlossen bleiben. Eine einheitliche Regelung gab es diesbezüglich nicht: In Schöneberg dauerten die Grippeferien fünf Tage, in Wilmersdorf 14 Tage. Davor fand schon wochenlang verkürzter Unterricht statt, da eine große Zahl der Lehrkräfte krankheitsbedingt ausfiel. Auch der öffentliche Nahverkehr musste mit Ausfällen kämpfen: Bei der Berliner Straßenbahn hatten sich innerhalb weniger Wochen 1050 der insgesamt 9 000 Angestellten krankgemeldet. Auch die Post stellte die Zulieferung für einzelne Tage gänzlich ein, was etwa mit dem Ausfall des Internets heutzutage vergleichbar wäre. Ein generelles Verbot von Gottesdiensten und Versammlungen oder die Schließung von Restaurants und Gaststätten, wie anderenorts verordnet, wurde in Berlin nicht verhängt. Die Stadtverordnetenversammlung konnte sich nicht auf ein entschlossenes Vorgehen einigen, da in den Tagen vor der Novemberrevolution jeden Tag Tausende bei Kundgebungen auf Straßen und Plätze drängten. Aufgrund einer Grippe-Pandemie ein Versammlungsverbot oder die Beschränkung von Freizeitaktivitäten zu verkünden, wagten sich die Abgeordneten der Berliner Stadtverordnetenversammlung nicht.⁸

Iris Hoßmann-Büttner
ist Referentin in der Stabsstelle *Querschnittsanalysen und Digitale Transformation* des Amtes für Statistik Berlin-Brandenburg.

⁸ Vossische Zeitung, Morgen-Ausgabe vom 24.10.1918, <http://zefys.staatsbibliothek-berlin.de/kalender/auswahl/date/1918-10-24/27112366/> [abgerufen am 17.06.2020].

Save the date

▮ 40. Kongress der Deutschen Gesellschaft für Soziologie

14. bis 25. September 2020

Der 40. Kongress der Deutschen Gesellschaft für Soziologie (DGS) 2020 mit der Überschrift „Gesellschaft unter Spannung“ findet digital statt. Er zielt auf die Beschreibung, Analyse und Erklärung gesellschaftlicher Spannungen in den unterschiedlichsten Verhältnissen, die gegenwärtige Gesellschaften ausmachen: zwischen Gruppierungen verschiedenster Art, Kulturen und Milieus, Organisationen, Professionen und Institutionen, individuellen Akteurinnen und Akteuren sowie in Interaktionen, ihren räumlichen, zeitlichen und kommunikativen Aspekten, in Fragen von Macht und Ungleichheiten sowie der Rolle von Politik, Technik, Medien, Wissen, Sprache, Kunst und anderen Bereichen soziologischen Forschens.

Weitere Informationen unter:

<https://kongress2020.sozioogie.de/aktuelles>

Save the date

▮ Vfs-Jahrestagung 2020

27. bis 30. September 2020

Unter dem Titel „Gender Economics. Geschlechterforschung in den Wirtschaftswissenschaften“ widmet sich die diesjährige Jahrestagung des Vereins für Socialpolitik (Vfs) virtuell der Darstellung von Geschlechterunterschieden in verschiedenen wirtschaftlich relevanten Bereichen, etwa hinsichtlich Bildungs- und Arbeitsmarkterfolgen, der Diskussion von Ursachen und der Entwicklung innovativer Maßnahmen zum Abbau bestehender Ungleichgewichte zwischen den Geschlechtern. Drei Keynotes führender internationaler Wissenschaftler geben Einblicke in diese Fragen und einen Überblick über die aktuelle Forschung. Raquel Fernandez (New York University) wird einen makroökonomischen Blick auf Geschlechterfragen werfen, unter Berücksichtigung kultureller Faktoren und Geschlechternormen. Muriel Niederle (Stanford University) wird analysieren, wie institutionelle Rahmenbedingungen diese Verhaltensunterschiede beeinflussen. Manuel Bagues (Warwick) wird eine Reihe von Arbeiten zu den Konsequenzen von Frauenquoten für wirtschaftliche Ergebnisse diskutieren. Im offenen Tagungsteil wird sich die gesamte Bandbreite aktueller wirtschaftswissenschaftlicher Forschung widerspiegeln.

Weitere Informationen unter:

<https://vfs2020.uni-koeln.de/de/>

Save the date

▮ Smart Country Convention

27. und 28. Oktober 2020

Die Digitalisierung von Städten, Gemeinden und Regionen ist das zentrale Thema der Smart Country Convention. Die Veranstaltung wird für die dritte Auflage in den digitalen Raum ziehen und virtuell stattfinden. Am 27. Oktober 2020 geht es ganztägig um das Thema E-Government, am 28. Oktober 2020 liegt der Fokus auf Smart City. Das Programm wird live aus dem Studio im hub27 auf dem Berliner Messegelände übertragen. Die Teilnehmenden werden zahlreiche Möglichkeiten zum virtuellen Austausch erhalten. Die Smart Country Convention richtet sich an Vertreterinnen und Vertreter von Bund, Ländern, Landkreisen, Städten und Gemeinden sowie kommunalen Unternehmen.

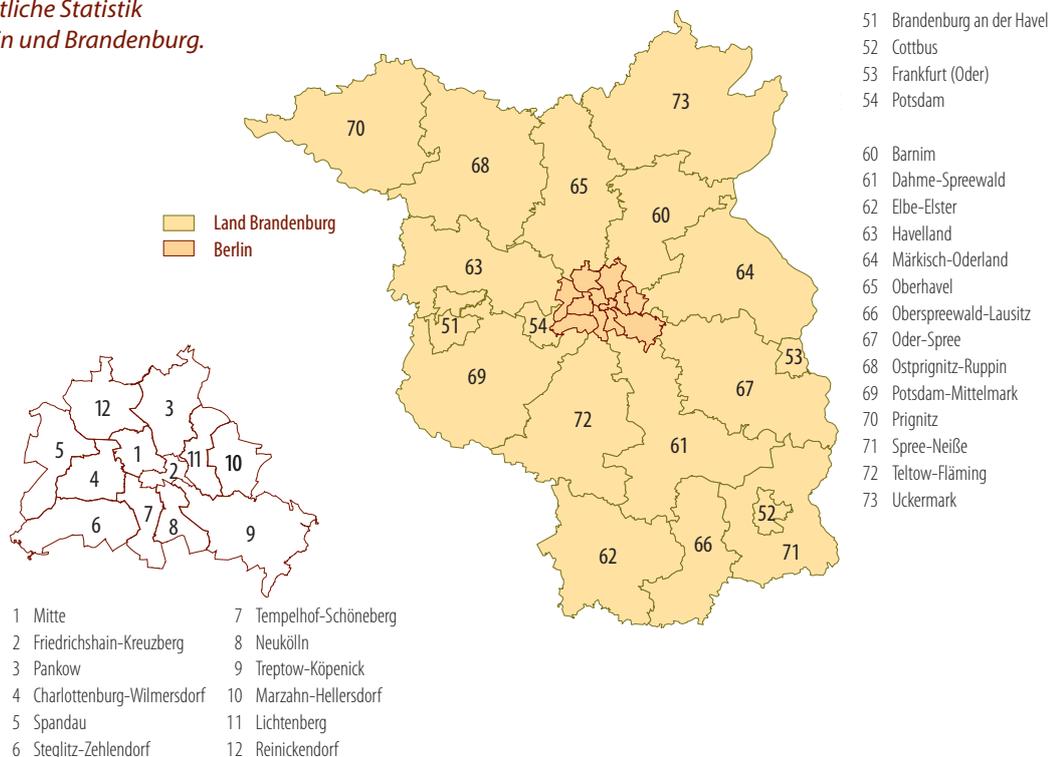
Weitere Informationen unter:

<https://www.smartcountry.berlin/>

statistik Berlin Brandenburg

- | Wir berichten fachlich unabhängig, neutral und objektiv über die Ergebnisse der amtlichen Statistik.
- | Wir haben den gesetzlichen Auftrag zur Datenerhebung mit der Möglichkeit zur Auskunftspflichtung.
- | Wir garantieren die Einhaltung des Datenschutzes.
- | Wir wenden adäquate statistische Methoden und Verfahren an und erhöhen kontinuierlich das erreichte Qualitätsniveau.
- | Wir gewährleisten regionale und zeitliche Vergleichbarkeit unserer Statistiken durch überregionale Kooperation.
- | Wir ermöglichen jedermann Zugang zu statistischen Ergebnissen.

Wir sind der führende Informationsdienstleister für amtliche Statistik in Berlin und Brandenburg.



Unter

www.statistik-berlin-brandenburg.de

finden Sie einen Überblick über das gesamte Leistungsspektrum des Amtes mit aktuellen Daten, Pressemitteilungen, Statistischen Berichten, regionalstatistischen Informationen, Wahlstatistiken und -analysen.

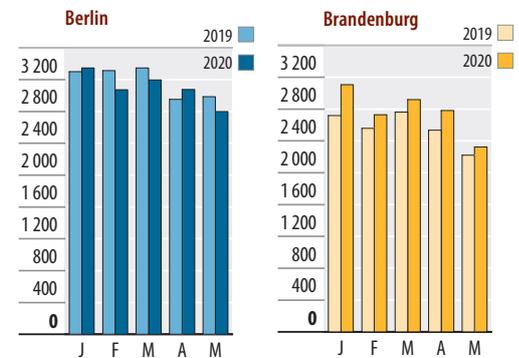
Sterbefallgeschehen in der Hauptstadtregion

In den ersten 22 Kalenderwochen des Jahres 2020 starben im Land Brandenburg 1 027 Personen bzw. 7,9 % mehr als im entsprechenden Vorjahreszeitraum. Dabei ist fast das gesamte bisherige Jahr 2020 von höheren Sterbefallzahlen als 2019 geprägt. So fiel etwa die Hälfte der Differenz zum Vorjahr in die Monate Januar und Februar. Von Ende März bis Mitte April kam es durch

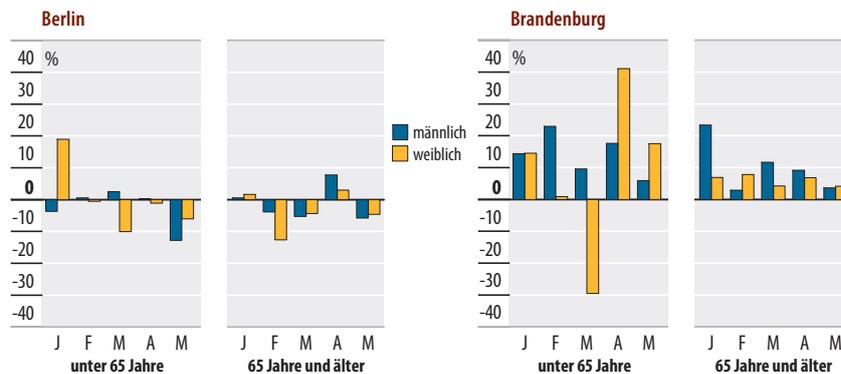
die Corona-Pandemie zu Ausschlägen im Sterbefallgeschehen. Vom 23. März bis zum 12. April (Kalenderwochen 13 bis 15) starben im Vergleich zum Vorjahreszeitraum 309 Personen (+17,6 %) mehr. Dabei ist zu berücksichtigen, dass in Brandenburg im Jahr 2020 bisher generell mehr Menschen starben als im Vorjahr. Nach Bereinigung um diesen Effekt ergibt sich dort eine Zunahme

der Sterblichkeit in den Kalenderwochen 13 bis 15 um 11,2 %. In Berlin starben in den ersten 22 Kalenderwochen dieses Jahres 519 Personen weniger als im entsprechenden Vorjahreszeitraum. Das entspricht einem Rückgang von 3,2 %. In den Kalenderwochen 13 bis 15 übertrafen die Sterbefallzahlen das Vorjahresniveau um 132 Personen (+6,3 %). Da in Berlin im Jahr 2020 bisher weniger Men-

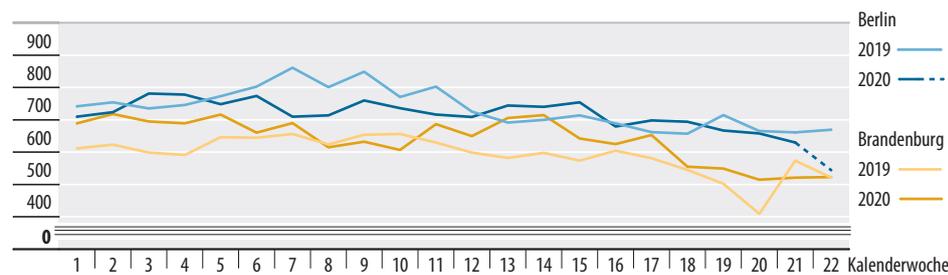
Eingegangene Sterbefallmeldungen am Sterbeort 2019 und 2020 in Berlin und Brandenburg nach ausgewählten Monaten



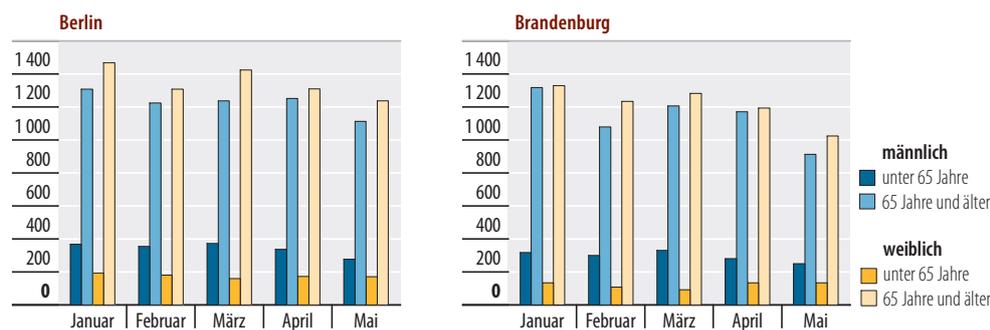
Prozentuale Veränderung der eingegangenen Sterbefallmeldungen am Sterbeort 2020 zum Vorjahresmonat in Berlin und Brandenburg nach ausgewählten Monaten und Geschlecht



Eingegangene Sterbefallmeldungen am Sterbeort 2019 und 2020 in Berlin und Brandenburg nach ausgewählten Kalenderwochen



Eingegangene Sterbefallmeldungen am Sterbeort 2020 in Berlin und Brandenburg nach ausgewählten Monaten, Altersgruppen und Geschlecht – Rohdaten



schen starben als im Vorjahr, dürfte die Übersterblichkeit in den Kalenderwochen 13 bis 15 unterzeichnet sein. Wird dieser Trend herausgerechnet, ergibt sich eine Zunahme der Sterblichkeit von 10,4 %.

In beiden Ländern ging in den darauffolgenden Wochen die Sterblichkeit zurück, sodass nach aktueller Datenlage von einer Normalisierung des Sterbgeschehens ausgegangen werden kann.

Die Sterbefälle werden betrachtet, um eine mögliche Übersterblichkeit durch COVID-19 und die Maßnahmen zur Eindämmung der Pandemie zu identifizieren. Eine erhöhte Sterblichkeit in Zusammenhang mit COVID-19 ist nach aktueller Datenlage auf einen kurzen Zeitraum beschränkt und schwach ausgeprägt. Bei den dargestellten Sterblichkeitsdaten handelt es sich um eine reine Fallauszählung ohne weitere statistische Aufbereitung der eingegangenen Sterbefallmeldungen. Vom Tod einer Person bis zur Meldung an die amtliche Statistik vergeht Zeit. In Brandenburg betrug im Jahr 2019 in 95 % der Fälle diese Zeitspanne zwei Wochen oder weniger. In Berlin waren 2019 nach acht Wochen 95 % aller Sterbefälle registriert. Daher sind die Sterbefallzahlen, insbesondere in Berlin, am aktuellen Rand unterschätzt.