ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Dreschmitt, Kai; Pfahl, Miriam

Article

Die Personenerhebung im Zensus 2022

WISTA - Wirtschaft und Statistik

Provided in Cooperation with:
Statistisches Bundesamt (Destatis), Wiesbaden

Suggested Citation: Dreschmitt, Kai; Pfahl, Miriam (2024) : Die Personenerhebung im Zensus 2022, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 76, Iss. 6, pp. 29-40

This Version is available at:
https://hdl.handle.net/10419/309497

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DIE PERSONENERHEBUNG IM ZENSUS 2022

Kai Dreschmitt, Miriam Pfahl

Schlüsselwörter: Primärerhebung – Haushaltsstichprobe – Konzeption – Erhebungsinstrumente – Softwareentwicklung – Schulung

ZUSAMMENFASSUNG

Das registergestützte Zensusverfahren sieht vor, Verwaltungsregister als Datenquellen zu nutzen und diese durch zusätzlich primärstatistisch erhobene Angaben zu korrigieren und zu ergänzen. Die Personenerhebung ist eine tragende Säule des Zensus 2022 und zählt zu den ergänzenden Primärerhebungen. Dieser Beitrag gibt einen Überblick über die zentralen Vorbereitungen: Er beschreibt das Erhebungskonzept und stellt die Erhebungsinstrumente vor. Nach einem Überblick über den Aufbau der zur Personenerhebung genutzten IT-Systeme geht er auf das fachliche und IT-seitige Schulungskonzept ein und zeigt spezielle Herausforderungen auf. Auch die Datenaufbereitung und Plausibilisierung werden näher beleuchtet.

Keywords: *primary survey – household sample – conceptual design – survey instruments – software development – training*

ABSTRACT

The register-assisted census procedure involves using administrative registers as data sources and correcting and supplementing the registers with additional primary statistical data. The survey of individuals is a key pillar of the 2022 Census and one of the supplementary primary surveys. This article provides an overview of the central preparations, explains the survey design and presents the survey instruments. After providing an overview of the structure of the IT systems used in the survey of individuals, the article proceeds to describe the approach for statistical and IT training and highlights specific challenges. It also examines data processing and data editing in greater detail.

Kai Dreschmitt

studierte Soziologie und Philosophie und ist wissenschaftlicher Mitarbeiter im Referat „Räumliche Bevölkerungsbewegung, Gebietsgliederungen“ des Statistischen Bundesamtes. Für den Zensus 2022 war er im Rahmen der Personenerhebung für die Betreuung des Erhebungsunterstützungssystems und die Schulung der Nutzenden zuständig.

Miriam Pfahl

studierte Politikwissenschaft, Soziologie und Öffentliches Recht und ist Referentin im Referat „Gebäude und Wohnungen“ des Statistischen Bundesamtes. Für den Zensus 2022 war sie unter anderem für die Konzeption der Online-First-Strategie und die Entwicklung der Online-Fragebogen der Haushaltsstichprobe zuständig.

1

Einleitung

Das in Deutschland sowohl 2011 als auch 2022 angewandte registergestützte Zensusverfahren sieht vor, vornehmlich vorhandene Verwaltungsregister als Datenquellen zu nutzen und diese durch zusätzlich primärstatistisch erhobene Angaben zu korrigieren und zu ergänzen. Dies ist immer dann erforderlich, wenn die geforderten Informationen nicht oder in nicht ausreichender Qualität in den Registern vorliegen. Neben der Gebäude- und Wohnungszählung und der Befragung an Gemeinschaftsunterkünften zählt die Personenerhebung zu den ergänzenden Primärerhebungen des Zensus.

Die Personenerhebung umfasst:

- die Haushaltebefragung (Stichprobe),
- die Befragung an Anschriften mit Wohnheimen (Vollerhebung) und
- die Wiederholungsbefragung (Unterstichprobe der Haushaltebefragung).

Während die Gebäude- und Wohnungszählung sowie die Befragungen an Anschriften mit Wohnheimen und an Gemeinschaftsunterkünften als Vollerhebungen durchgeführt wurden, handelt es sich bei der Haushaltebefragung sowie der Wiederholungsbefragung[1] jeweils um eine Stichprobenbefragung. Um Informationen für ganz Deutschland zu erhalten, wurden die Befragungsergebnisse mithilfe mathematischer Verfahren auf die gesamte Bevölkerung hochgerechnet.

Für die Haushaltebefragung wurde eine zufällige Stichprobe aus deutschlandweit allen Anschriften mit Wohnraum gezogen. Alle Personen, die zum Zensusstichtag am 15. Mai 2022 an den gezogenen Anschriften wohnhaft waren, wurden für den Zensus befragt. Dies entspricht etwa 12 % der Einwohnerinnen und Einwohner Deutschlands, rund 10,2 Millionen Menschen.

In Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern wurden die soziodemografischen Zusatzmerkmale zur Bildung und Erwerbstätigkeit nur bei einer weiteren Unterstichprobe von etwa 40 % der Stichprobenanschriften erhoben.

Die zur Personenerhebung zählende Wiederholungsbefragung fand ebenfalls als Unterstichprobe an lediglich 4 % der Stichprobenanschriften statt. Im Zuge der Wiederholungsbefragung wurde somit ein kleiner Anteil der Auskunftspflichtigen aus der Haupterhebung erneut zu einzelnen Merkmalen befragt.

Die Personenerhebung verfolgte zwei Ziele:

1. Durch die sogenannte Existenzfeststellung leistete sie einen wichtigen Beitrag bei der statistischen Korrektur der Melderegister zur Ermittlung der amtlichen Einwohnerzahlen von Bund, Ländern und Gemeinden. Sie schuf damit die Grundlage für die Fortschreibung der Einwohnerzahlen für die Zeit zwischen zwei Volkszählungen (Ziel 1).
2. Zudem ermittelte die Befragung weitere zentrale Strukturmerkmale der Bevölkerung zu den Themen Wohnsituation, Staatsangehörigkeit und Zuwanderung, Bildung und Ausbildung, Erwerbstätigkeit, Arbeitsort, Wirtschaftszweig und Beruf (Ziel 2).

Die nachgelagerte Wiederholungsbefragung diente dazu, die Qualität der ermittelten Einwohnerzahl zu bewerten.

Der Beitrag zeichnet die Stationen der konzeptionellen und technischen Vorbereitungen für die Erhebung 2022 nach. Auf die Durchführungsphase geht er dabei nur am Rande ein, da diese maßgeblich durch die Statistischen Ämter der Länder betreut wurde. Kapitel 2 beleuchtet die Vorbereitungen sowie die Neuerungen im Vergleich zum Zensus 2011. Die Kapitel 3 und 4 beschreiben das Erhebungskonzept und die Erhebungsinstrumente. Nach einem Überblick über die qualitativen und quantitativen Pretests in Kapitel 5 geht Kapitel 6 näher auf den Aufbau der zur Personenerhebung genutzten IT-Systeme ein. Darüber hinaus stellt Kapitel 7 das fachliche und IT-seitige Schulungskonzept vor und beschreibt spezielle Herausforderungen. Der auf die Erhebungsphase folgenden Datenaufbereitung und Plausibilisierung widmet sich das vorletzte Kapitel, bevor ein Fazit hinsichtlich der zentralen Erfolgsfaktoren gezogen wird.

1 Auf die Wiederholungsbefragung geht dieser Beitrag nicht weiter ein. Der Fokus liegt auf der Haushaltebefragung auf Stichprobenbasis.

2

Vorbereitungen und Neuerungen

2.1 Nach dem Zensus ist vor dem Zensus

Um die Personenerhebung 2022 vorzubereiten, wurden zunächst die Verfahren und Prozesse des Zensus 2011 umfassend geprüft und bewertet. Die Erfahrungsberichte der Statistischen Ämter der Länder, eine externe Evaluierung im Auftrag des Bundesministeriums des Innern und für Heimat, Empfehlungen des Rats für Sozial- und Wirtschaftsdaten sowie die Auswertung des Fragebogens 2011 lieferten allesamt wertvolle Erkenntnisse und Hinweise für die Neukonzeption der folgenden Zensusrunde. Das Grundmodell für die Personenerhebung 2022 sollte im Wesentlichen dem des Zensus 2011 entsprechen, jedoch mit Optimierungen unter Genauigkeits-, Zeit-, Effizienz- und Kostenaspekten. Neue Datenquellen sowie informationstechnische und methodische Neuerungen waren zu berücksichtigen, sofern diese einen zusätzlichen Nutzen stiften würden und für den Zensus eingesetzt werden könnten.

Zur konzeptionellen Vorbereitung der Personenerhebung 2022 wurde das Teilprojekt Personenerhebung (Konzeption & Aufbereitung) eingerichtet, welches sich aus Mitarbeitenden des Statistischen Bundesamtes und der Statistischen Landesämter zusammensetzte. Die Arbeitsgruppe Haushaltsstichprobe, die neben den Mitgliedern des Teilprojekts aus weiteren Vertreterinnen und Vertretern der Statistischen Ämter der Länder und zusätzlich der Kommunen bestand, nahm die im Teilprojekt erarbeiteten Konzepte ab.

Die Aufgaben des Teilprojekts Personenerhebung waren im Einzelnen:

- Entwickeln und Bereitstellen eines bundeseinheitlichen Konzepts, um die Personenerhebung durchzuführen
- Entwickeln und Bereitstellen bundeseinheitlicher Erhebungsinstrumente für die Existenzfeststellung, für die Erhebung der Ziel-1-Merkmale sowie für die Erhebung der Zusatzmerkmale (Ziel 2)
- Formulieren und Bereitstellen von fachlichen Anforderungen an die IT-Entwicklung, um die Personenerhebung durchzuführen
- Bereitstellen der betriebsbereiten IT-Anwendungen, um die Personenerhebung durchzuführen und aufzubereiten
- Erstellen eines Schulungskonzepts und Bereitstellen von Schulungsunterlagen für die Mitarbeitenden der Statistischen Landesämter, der Erhebungsstellen und für die Erhebungsbeauftragten.

2.2 Neuerungen

Zu den wesentlichen Neuerungen im Zensus 2022 zählt unter anderem die zentrale Ansiedlung der IT-Systeme im Statistischen Bundesamt. Hier folgte man den Handlungsempfehlungen der Evaluationsergebnisse des Zensus 2011 und verzichtete auf eine dezentrale Aufteilung der IT-Systeme im Statistischen Verbund[2] zugunsten einer zentralen Lösung (Dittrich, 2019). So wurden insbesondere die Kommunikation zwischen den verschiedenen IT-Schnittstellen optimiert und Effizienzsteigerungen erreicht. Da alle Daten an einer zentralen Stelle gesammelt und weiterverarbeitet wurden, konnten sie zeitnah validiert und auf ihre Richtigkeit überprüft werden – beispielsweise in Bezug auf die sogenannte erhebungsteilübergreifende Plausibilisierung. Mit dieser erfolgte bereits während der Erhebung ein Abgleich der Erhebungsdaten mit den Angaben aus den Melderegistern (Bretnütz und andere, 2024).

Um die Erhebung zu optimieren und die Erhebungskonzepte zeitgemäß anzupassen, wurde unter anderem der Einsatz mobiler Endgeräte bei den persönlichen Befragungen im Zensus 2022 geprüft. Gleichzeitig legten die Projektbeteiligten eine Online-First-Strategie für die primärstatistischen Erhebungen fest. Wenngleich die Auskunftgebenden auch im Zensus 2022 zwischen verschiedenen Optionen der Meldungsabgabe wählen konnten, stellte die Selbstauskunft über den Online-Fragebogen den präferierten Erhebungsweg[3] dar. Das

2 Den Statistischen Verbund bilden die Statistischen Ämter des Bundes und der Länder.

3 Bei der Personenerhebung wurde die Online-First-Strategie für die Erhebung der Zusatzmerkmale (Bildung und Erwerbstätigkeit) verfolgt. Für die vorangehende Existenzfeststellung war der Einsatz von Erhebungsbeauftragten vorgesehen (siehe hierzu auch das folgende Kapitel 3).

sollte durch eine Reihe gezielter Maßnahmen erreicht werden (Gaedke und andere, 2024; Freier/Mosel, 2019; Neutze und andere, 2017). Ziel war, bei der Personenerhebung eine Online-Selbstausfüller-Quote von mehr als 50 % zu erreichen. (Hentschke und andere, 2024, hier: Übersicht 3) Einige Bundesländer schufen zudem die Voraussetzungen, die interviewergestützten persönlichen Befragungen mithilfe von Tablets durchzuführen. Papierfragebogen sollten nur ersatzweise eingesetzt werden.

3 Erhebungskonzept

Die Personenerhebung war eine Kombination aus einer Vor-Ort-Befragung durch Interviewerinnen und Interviewer (Erhebungsbeauftragte) sowie – für weitere Fragen – einer sich anschließenden Selbstauskunft durch die Auskunftspflichtigen. Hierzu wurden kleinräumig (in der Regel in den Kommunalverwaltungen auf Kreisebene) Erhebungsstellen eingerichtet, die die Erhebung koordinierten und leiteten. Anders als bei der Gebäude- und Wohnungszählung war bei der Personenerhebung Interviewpersonal erforderlich, da nur die Anschriften festgelegt waren, an denen die Existenzfeststellung und die weiteren Befragungen durchgeführt werden sollten.

In einem ersten Schritt mussten die Interviewerinnen und Interviewer die ausgewählten Anschriften begehen und zunächst erhebungsorganisatorische Angaben erfassen (zum Beispiel Prüfung des Anschriftentyps, Anmerkungen zum Auffinden der Anschrift und zur Anzahl der Haushalte). Dies sollte zum einen sicherstellen, dass die korrekten Anschriften aufgesucht und beispielsweise keine Hinterhäuser, Seiteneingänge oder Ähnliches bei der späteren Befragung übersehen werden. Zum anderen hinterließen die Erhebungsbeauftragten Ankündigungsschreiben mit Informationen zum Zensus sowie eine Terminankündigung für die Interviews. Zum vereinbarten Termin ermittelten sie dann in einer persönlichen Befragung, welche Personen zum Zensusstichtag 15. Mai 2022 in den Haushalten wohnhaft und damit zählungsrelevant waren. Außerdem erhoben sie die für die Existenzfeststellung geforderten Merkmale[4] sowie weitere Angaben[5]. Für einen Großteil der Auskunftspflichtigen endete die Befragung an dieser Stelle. Lediglich bei einer Untermenge wurden darüber hinaus Informationen zur Bildung und Erwerbstätigkeit erfragt. Diese Angaben konnten die Auskunftspflichtigen selbstständig tätigen, bevorzugt online über den Fragebogen im Internet (Online First), auf Wunsch auch in Papierform oder bei Bedarf durch Interviewerunterstützung – in einigen Bundesländern tabletbasiert, gegebenenfalls auch telefonisch.

Bei Nichterreichbarkeit griff ein detailliertes Verfahren bestehend aus erneutem Kontaktversuch und Erinnerungs- und Mahnverfahren auf postalischem Weg, um alle Personen einer Anschrift zu erreichen.

Dieser kaskadenartig aufgebaute Katalog von Schritten war im Zuge der Existenzfeststellung zwingend und einheitlich zu durchlaufen. So wurde sichergestellt, dass für jede Person an einer Anschrift ihre Existenz beziehungsweise Nicht-Existenz im Hinblick auf die Einwohnerzahlermittlung eindeutig festgestellt werden konnte.

4 Erhebungsinstrumente

Bei der Personenerhebung kamen unterschiedliche standardisierte Erhebungsinstrumente zum Einsatz. Sie unterschieden sich insbesondere hinsichtlich der jeweiligen Nutzenden (Erhebungsbeauftragte oder Auskunftspflichtige), der Zielgruppe (Haushaltebefragung oder Erhebung an Wohnheimen) sowie des Merkmalskranzes, der jeweils zu erheben war. ↘ Übersicht 1

4.1 Erhebungsliste für Erhebungsbeauftragte

Bei der Begehung der Erhebungsbezirke[6] nutzten die Erhebungsbeauftragten eine Erhebungsliste, die sich aus einem Mantelbogen (je Erhebungsbezirk) sowie

4 Familienname, Vorname, Geburtsdatum, Geschlecht sowie an Anschriften mit Wohnheimen zusätzlich Geburtsname, Geburtsort und Geburtsstaat.

5 Staatsangehörigkeit, Familienstand, Wohnungsstatus, gegebenenfalls Hauptwohnung.

6 Ein Erhebungsbezirk entsprach im Regelfall einer bestimmten (Stichproben-)Anschrift, wobei Anschriften mit vielen Haushalten gegebenenfalls in mehrere Erhebungsbezirke aufgeteilt wurden.

Übersicht 1
Erhebungsinstrumente der Personenerhebung

Erhebungsinstrumente für Erhebungsbeauftragte	
Anschriftenebene	Mantelbogen: je Erhebungsbezirk
Haushaltsebene	Haushaltsbogen: je Haushalt
	Haushaltsbogen Wiederholungsbefragung: je Haushalt
Erhebungsinstrumente für befragte Personen (Fragebogen)	
Personenebene	Fragebogen zur Befragung an Ziel-1-Anschriften
	Fragebogen zur Befragung an Ziel-1- und Ziel-2-Anschriften
	Fragebogen zur Befragung an Ziel-1-Anschriften mit Wohnheimen
	Fragebogen zur Befragung an Ziel-1- und Ziel-2-Anschriften mit Wohnheimen
	Fragebogen zur Wiederholungsbefragung

einem oder mehreren Haushaltsbogen (je Haushalt) zusammensetzte.

Der Mantelbogen diente dazu, Ergebnisse und (Anschriften-)Befunde[7] der vor Ort an einem Erhebungsbezirk vorgefundenen Situation zu erfassen. Zudem enthielt er die Information, ob an der betreffenden Anschrift eine Ziel-1- oder eine Ziel-2-Erhebung vorgesehen war.

Im Haushaltsbogen wurden die erhebungsorganisatorischen Angaben zum Haushalt (zum Beispiel angekündigter Befragungstermin oder Auskunftsverweigerung), die allgemeinen Angaben zu den einzelnen Haushalten eines Erhebungsbezirks sowie die erfragten Erhebungsergebnisse zu den einzelnen Personen erfasst.

Zur Unterstützung der Interviewerinnen und Interviewer standen für den Haushaltsbogen Übersetzungshilfen in 15 Sprachen zur Verfügung (Arabisch, Bulgarisch, Englisch, Französisch, Griechisch, Italienisch, Kroatisch, Persisch, Polnisch, Portugiesisch, Rumänisch, Russisch, Spanisch, Türkisch, Ukrainisch).

4.2 Fragebogen je befragte Person

Insgesamt gab es fünf verschiedene Fragebogen (siehe Übersicht 1): Da alle Haushalte zu den Ziel-1-Merkmalen, aber nur eine Unterstichprobe auch zu den Ziel-2-Merkmalen befragt wurden, wurden an Normalanschriften jeweils zwei unterschiedliche Fragebogen verwendet. Der Kurzfragebogen deckte nur die Ziel-1-Merkmale ab, der längere Fragebogen sowohl die Ziel-1- als auch die Ziel-2-Merkmale. Der Ziel-1-Fragebogen kam nur zum Einsatz, falls die Erhebungsbeauftragten nicht alle Informationen zu einer Auskunftsperson über den Haushaltsbogen erfassen konnten. Da an Wohnheimanschriften zusätzliche Merkmale zu erheben waren, unterschieden sich die Fragebogen je nach Zielgruppe. Zudem gab es einen separaten Fragebogen für die Wiederholungsbefragung.

Im Gegensatz zu den Mantel- und Haushaltsbogen waren die fünf Fragebogen auch in elektronischer Form als Online-Formulare (IDEV[8]) zur Selbstauskunft hinsichtlich der weiterführenden Fragen (beispielsweise zur Erwerbstätigkeit) verfügbar. Alle Papierfragebogen enthielten zudem den Link zum Online-Formular und Zugangsdaten zu den Online-Fragebogen, um bei Interesse jederzeit eine Online-Teilnahme zu ermöglichen. Ein wesentlicher Vorteil der Online-Fragebogen war, dass diese außer in Deutsch in 15 weiteren Sprachen[9] bereitgestellt wurden und somit von Personen mit schlechten Deutschkenntnissen in ihren jeweiligen Muttersprachen aufgerufen und ausgefüllt werden konnten (Gaedke und andere, 2024).

7 Die sogenannten Befunde spiegeln das Ergebnis der Begehung wider. Hier konnten unter anderem Auffälligkeiten vermerkt werden, zum Beispiel wenn es sich um eine rein gewerblich genutzte Anschrift handelt, die Anschrift nicht vorgefunden wurde oder eine Baulücke ohne Bewohnerinnen und Bewohner besteht.

8 IDEV steht für Internet Datenerhebung im Verbund und ist ein einheitliches Online-Meldeverfahren, das seit 2005 von den Statistischen Ämtern des Bundes und der Länder angeboten wird.

9 Arabisch, Bulgarisch, Englisch, Französisch, Griechisch, Italienisch, Kroatisch, Persisch, Polnisch, Portugiesisch, Rumänisch, Russisch, Spanisch, Türkisch, Ukrainisch.

5

Pretests

Um die für den Zensus 2022 angepassten Erhebungsverfahren und -instrumente vor ihrem Einsatz zu erproben, wurden im Vorfeld mehrere qualitative sowie ein quantitativer Pretest[10] – jeweils mit unterschiedlichen Zielsetzungen – durchgeführt.

5.1 Qualitative Pretests

Zunächst wurden die für die Personenerhebung erstellten Erhebungsinstrumente im Pretestlabor des Statistischen Bundesamtes getestet. Hierfür hat eine kleine Gruppe Testpersonen (jeweils zwischen 20 und 25 Personen) sowohl die Papier- als auch die Online-Fragebogen unter Beobachtung ausgefüllt und anschließend in einem kognitiven Interview dazu Feedback gegeben. Von Interesse waren dabei vor allem die Verständlichkeit der Fragen, der Antworten und der Hinweistexte, der Filterführung (im Papierfragebogen) sowie der generellen Handhabbarkeit der Erhebungsinstrumente. Durch die Kombination von Beobachtung und Interview konnten potenzielle Stolperfallen erkannt und Nachbesserungen vorgenommen werden. Dies betraf sowohl gestalterische Aspekte als auch Formulierungen, um zum Beispiel relevante Hinweistexte besser wahrzunehmen und die Fragen und Antwortkategorien verständlicher zu machen.

Ebenfalls qualitativ und mit Fokus auf Verständlichkeit und Handhabbarkeit wurden diese Erhebungsinstrumente aus der Perspektive der Erhebungsbeauftragten getestet. Hierfür wurden zum einen Fokusgruppendiskussionen mit erfahrenen Interviewerinnen und Interviewern durchgeführt, zum anderen Interviewsituationen mit Testpersonen nachgestellt, beobachtet und im Anschluss ebenfalls kognitive Interviews durchgeführt. Wenngleich sich im Ergebnis zeigte, dass der Fragebogen grundsätzlich auch für ein Interview mit Erhebungsbeauftragten geeignet ist und die Konzeption als Selbstausfüller keine erkennbaren negativen Auswirkungen hat, konnten dennoch weitere Optimierungspotenziale identifiziert werden.

5.2 Quantitativer Pretest

Neben den unter Laborbedingungen durchgeführten qualitativen Pretests fand im Zeitraum von November 2019 bis März 2020 auch eine praktische Pilotstudie unter realistischen Bedingungen der Haupterhebung im Feld statt. Befragt wurden insgesamt 8 632 Personen. Hierbei wurden nicht nur die Erhebungsinstrumente, sondern die komplette Erhebungsorganisation sowie die nachfolgende Aufbereitung erprobt. Zudem sollten Erkenntnisse zur Nutzung der verschiedenen Erhebungswege und insbesondere der Online-Fragebogen gewonnen werden (Estatico und andere, 2021). Methodik und Abläufe entsprachen dabei weitestmöglich denen der Haupterhebung: Es wurde eine Anschriftenstichprobe gezogen, um die Testpersonen zu gewinnen. Die Erhebungsbeauftragten erfassten zunächst erhebungsorganisatorische Angaben und stellten die Auskunftspflicht fest. Dann erhoben sie mindestens die für die Existenzfeststellung geforderten Informationen in einem persönlichen Interview und händigten denjenigen Personen, die auch für die Befragung der Zusatzmerkmale vorgesehen waren, Zugangsdaten für den Online-Fragebogen aus. Auf Wunsch erhoben sie diese Angaben im persönlichen Interview oder führten eine computergestützte Telefonbefragung durch.[11]

Nicht nur die Feldarbeit und die Datenanalyse lieferten wichtige Informationen, sondern auch die im Online-Fragebogen protokollierten Paradaten zur Nutzung, Handhabbarkeit und Verständlichkeit des elektronischen Fragebogens. Neben Informationen über die genutzte Geräteklasse (PC, Tablet, Smartphone) und das Betriebssystem wurden beispielsweise die Verweildauer je Seite, das Hin- und Herspringen im Fragebogen sowie die Dauer des Ausfüllvorgangs erfasst und ausgewertet. Die optionale Nachbefragung am Ende des Online-Formulars lieferte ebenfalls weitere Erkenntnisse zur Handhabung des Online-Fragebogens (Gaedke und andere, 2024).

Die Auswertungen des quantitativen Pretests haben bestätigt, dass sich sowohl die eingesetzten Erhebungsinstrumente als auch die erhebungsorganisatorischen Abläufe im Feld bewährt haben. Gleichwohl konnte auch hier weiterer Optimierungsbedarf identifiziert werden.

10 Ein Pretest ist ein Vorab-Testlauf mit wenigen Probandinnen und Probanden, um die Qualität der Erhebungsinstrumente und gegebenenfalls des Erhebungsablaufs zu prüfen.

11 Abweichend zur Haupterhebung wurden keine weiteren persönlichen oder schriftlichen Kontaktversuche vorgenommen, da die vollständige Erreichung aller Zielpersonen beim Pretest nicht nötig war.

Zudem wurden wichtige Hinweise für das erfolgreiche Gelingen der Online-First-Strategie, die anstehenden Schulungen des Erhebungspersonals sowie eine zielgerichtete Öffentlichkeitsarbeit für den Zensus 2022 gewonnen.

6

Aufbau der IT-Unterstützung

Für die Personenerhebung wurden zwei IT-Anwendungen (sogenannte Fachanwendungen) maßgeschneidert entwickelt, um die Erhebung zu organisieren und durchzuführen und die Erhebungsdaten aufzubereiten: das Erhebungsunterstützungssystem sowie die Kodierung der Berufsbezeichnungen.

Die gesetzlichen Vorgaben zum Datenschutz und zur IT-Sicherheit erhöhten die Anforderungen sowohl an die zur Ergebniserfassung genutzte Software als auch an die dahinterliegende IT-Infrastruktur (Hardware). Hard- und Software hatten die millionenfachen Datenmeldungen und Nutzereingaben zuverlässig und sicher zu verarbeiten. Um die Zensus-Daten zu schützen wurde eine eigens dafür vorgesehene Sicherheitsumgebung aufgebaut, welche den notwendigen IT-Standards entspricht.

Gleichzeitig war für die Softwareentwicklung die Benutzerfreundlichkeit (Usability) des Produkts und damit eine intuitive Bedienung durch die Mitarbeitenden in den Kommunen und den Statistischen Ämtern der Länder wichtig.

6.1 Erhebungsunterstützungssystem

Das Erhebungsunterstützungssystem wurde als ein bundesländerübergreifendes Verwaltungs- und Steuerungsinstrument konzipiert, um die Aufgaben der Personenerhebung, das heißt die Befragung an Wohnheimen, die Befragung an Gemeinschaftsunterkünften, die Haushaltebefragung auf Stichprobenbasis sowie die Wiederholungsbefragung, zu organisieren und zu bearbeiten. Zu den zentralen Funktionen zählten:

- Verwaltung von Erhebungsstellen und Erhebungsbeauftragten,
- Organisation der Schulungen der Erhebungsbeauftragten,
- automatisierte Einteilung von Anschriften in Erhebungsbezirke,
- Zuweisungen der Erhebungsbezirke zu Erhebungsbeauftragten,
- Druck von Erhebungsunterlagen,
- Protokollierung der Übergabe und Entgegennahme von Erhebungsunterlagen an Erhebungsbeauftragte,
- automatisierter Empfang und Verarbeitung von Online-Meldungen,
- manuelle Eingabe von Erhebungsergebnissen,
- Organisation des Mahnwesens,
- automatisierte Integration der Ergebnisse der Fachanwendung Kodierung der Berufe,
- verschiedene Möglichkeiten zur Arbeitsstandkontrolle im Erhebungsprozess,
- Ticketsystem für Kommunen und Statistische Landesämter, um Auffälligkeiten zu klären und Stichprobenanschriften zu pflegen,
- verschiedene Exportfunktionalitäten zu Qualitätskennzahlen, Arbeitsstandkontrollen, Unterlagen der Erhebungsbeauftragten, Mahnwesen, Beleglesung und vieles mehr,
- Validierung der Erhebungsergebnisse und automatisierte Übertragung an den Referenzdatenbestand.

Dabei bestimmte ein Rollen- und Rechtekonzept mit zwei Hierarchieebenen (Mitarbeitende der Statistischen Landesämter und Mitarbeitende der Erhebungsstellen), welche Funktionalitäten die jeweiligen Nutzenden sehen und bearbeiten konnten.

Der Statistische Verbund arbeitete die Anforderungen an die Software aus und formulierte sie in einem Lastenheft. Die nachfolgende Softwareentwicklung durch einen IT-Dienstleister erfolgte inkrementell.[12] Dieses Vorgehen hatte den Vorteil, dass Fehler frühzeitig im Entwicklungsprozess erkannt und korrigiert werden konnten. Jeder Auslieferung eines Inkrements folgte eine mehrwöchige Qualitätssicherungsphase durch einen

12 Inkrementelle Softwareentwicklung ist ein iterativer Ansatz, bei dem Software in überschaubaren Teilmodulen entwickelt wird und regelmäßig funktionsfähige Zwischenversionen erstellt und an den Auftraggeber ausgeliefert werden. Jedes Inkrement bringt neue Funktionen oder Verbesserungen vorangegangener Inkremente mit sich, die sofort getestet und verwendet werden können.

externen Dienstleister mit anschließender Abnahme durch das Teilprojekt Personenerhebung.

In insgesamt elf Austauschterminen mit den künftigen Anwenderinnen und Anwendern wurden die aktuellen Inkremente präsentiert, um Feedback aus der Endnutzendenperspektive zu erhalten und gegebenenfalls Anpassungen vorzunehmen. Diese Termine waren mit Vertreterinnen und Vertretern der Statistischen Landesämter sowie der Kommunen besetzt, um deren Anforderungen und Erfahrungen bei der Programmierung berücksichtigen zu können. In einer mehrwöchigen Testphase vor Inbetriebnahme der Anwendung konnten darüber hinaus wertvolle Hinweise der Anwenderinnen und Anwender der Statistischen Landesämter umgesetzt werden.

6.2 Kodierung der Berufe

Die Fachanwendung Kodierung der Berufe diente der Nachkodierung von Berufsinformationen aus der Personenerhebung. Zentrale Funktionen der Fachanwendung waren die automatisierte und die manuelle Kodierung von Berufsangaben sowie die nachträgliche automatisierte Differenzierung von Meisterabschlüssen.[13] Analog zur Fachanwendung Erhebungsunterstützungssystem wurde auch die Fachanwendung Kodierung der Berufe iterativ und inkrementell entwickelt.

7

Konzeption und Durchführung der Schulungen

Alle an der Personenerhebung im Zensus 2022 Beteiligten waren für die jeweils anstehenden Aufgaben zu schulen. Dies waren im Wesentlichen drei Gruppen:

- Mitarbeitende der Statistischen Ämter der Länder in Fach- und Querschnittsbereichen (zum Beispiel Auskunftsdienst),
- die kommunalen Erhebungsstellen, die sich zusammensetzen aus Mitarbeitenden der Kommunalverwaltungen, aber auch eigens für den Zensus neu eingestelltem Personal,
- die Erhebungsbeauftragten, die zumeist von den Kommunen speziell für die Durchführung der Interviews eingestellt wurden.

Insgesamt wirkten mehr als 100 000 Personen an der Durchführung des Zensus 2022 mit. Diese verfügten über ein sehr heterogenes Vorwissen und sollten die unterschiedlichsten Rollen im Gesamtprozess wahrnehmen. Sie wurden zum einen fachlich (Konzepte, Erhebungsinstrumente, Erhebungsablauf, Rechtsgrundlagen und Datenschutz) und zum anderen technisch (in der Bedienung der IT-Anwendungen) auf die anstehenden Aufgaben vorbereitet.

7.1 Kaskadenprinzip

Um die unterschiedlichen Zielgruppen mit ihren jeweiligen Aufgabenfeldern angemessen zu schulen, war ein kaskadenartiger Wissenstransfer über zum Teil mehrtägige Multiplikatorenschulungen vorgesehen. Die Zuständigkeit für die Durchführung wurde dabei jeweils von einer Stufe zur nächsten weitergegeben und lag nicht bei einer zentralen Stelle. Die Schulungen waren ursprünglich meist in Präsenz geplant, wurden jedoch pandemiebedingt überwiegend online durchgeführt, sodass es möglich war, größere Schulungsgruppen zu bilden.

Zum Start der Schulungskaskade hat das Statistische Bundesamt Multiplikatoren der Statistischen Landesämter in den IT-Fachanwendungen in Form von Online-Schulungen sowie durch das Bereitstellen umfangreicher Materialien geschult.

Bei den IT-Schulungen handelte es sich um praktische Anwendungsschulungen in einer eigenen, geschützten Schulungsumgebung mit Schulungsdaten, in der die fachlichen Abläufe und der Umgang mit der Software erprobt wurden.[14] Die Präsentation der verschiedenen Funktionen wurde dabei jeweils durch thematisch passende Übungen begleitet.

13 Die Differenzierung von Meisterabschlüssen war notwendig, um die spätere Typisierung der Berufsabschlüsse im internationalen Klassifikationssystem ISCED 2011 (Internationale Standardklassifikation des Bildungswesens) vorzunehmen. Sie erfolgte anhand eines Abgleichs der Angaben aus dem Fragebogen mit einer in der Fachanwendung hinterlegten Liste.

14 Die Schulungsumgebungen standen den Anwenderinnen und Anwendern auch im Nachgang der Schulungen weiterhin zur Verfügung, um beispielsweise neue Mitarbeitende einzuarbeiten.

Die Multiplikatoren der Statistischen Landesämter gaben die Inhalte sowohl an die Kolleginnen und Kollegen im jeweiligen Statistischen Landesamt als auch an die Leitungen der Erhebungsstellen weiter.[15] Letztere wiederum schulten alle weiteren Mitarbeitenden der Erhebungsstellen sowie die Erhebungsbeauftragten in den jeweils für sie relevanten Themen. So war es möglich, innerhalb des vorgegebenen Zeitraums die große Anzahl aller in den Zensus involvierten Personen in den relevanten Bereichen zu schulen.

7.2 Schulungsmaterialien

Ergänzend zu den Schulungen sicherten umfangreiche, zentral erstellte Schulungsmaterialien die Qualität sowie ein bundesweit einheitliches Vorgehen. So konnten die Inhalte auch nach den Schulungen nachgeschlagen werden. Die unterschiedlichen Schulungsmaterialien waren rollenadäquat aufbereitet. Zu ihnen zählen:

- klassische Handbücher – Nutzungs- und Schulungshandbücher, die als Nachschlagewerk dienten und alle Aufgaben, die vor, während und nach der Erhebung durch die jeweiligen Akteure zu erledigen waren, sowie die Funktionen der Fachanwendungen beschrieben,
- Präsentationen – anschauliche Darstellung der Schulungsinhalte zur visuellen Unterstützung der Schulungen,
- Kurzanleitungen für Erhebungsbeauftragte – komprimierte Form der Handbücher; kurze und primär grafische Anleitungen, maximal eine Doppelseite; Überblick über die wichtigsten Punkte/Abläufe,
- kurze Video-Tutorials zum Erhebungsunterstützungssystem für die Erhebungsstellen – zeigten zentrale Funktionen/Abläufe in der Fachanwendung,
- Übungsaufgaben zum Erproben der IT-Anwendungen.

8

Datenaufbereitung

Bevor die im Zuge der Befragungen gewonnenen Erhebungsdaten ausgewertet und veröffentlicht werden konnten, waren verschiedene Aufbereitungsschritte zu durchlaufen.

8.1 Beleglesung

Die Rohdaten wurden zunächst in ein maschinell weiterverarbeitbares Format gebracht: Während die Antworten aus den elektronischen Erhebungsinstrumenten bereits in digitaler Form vorlagen, wurden die auf Papier getätigten Angaben zunächst digitalisiert. Im Rahmen der Beleglesung wurden die Fragebogen eingescannt, die Angaben mithilfe einer speziellen Software ausgelesen und gemäß zuvor definierten Vorgaben kodiert.[16] Die Freitextangaben zum Beruf und zur Tätigkeitsbeschreibung wurden zusätzlich als digitales Bild gespeichert und sofern erforderlich während der manuellen Kodierung der Berufsangaben gesichtet und klassifiziert.

8.2 Kodierung der Berufsangaben

Beim Thema Erwerbstätigkeit wurden die Auskunftspersonen unter anderem nach ihrer Berufsbezeichnung gefragt. Hierbei handelte es sich um eine offen formulierte Frage.

Um die Angaben der Auskunftspflichtigen sowohl in der nationalen als auch in der internationalen Klassifikation der Berufe (Bundesagentur für Arbeit, 2021; International Labour Organization, 2012) berichten zu können, erfolgte eine Kodierung über drei verschiedene Vorgehensweisen:

1. Bereits durch die Auskunftspflichtigen selbst beim Ausfüllen des Onlinefragebogens: Nach Eingabe eines Stichwortes konnte in dem eingebundenen Klassifikationsserver nach der Berufsbezeichnung gesucht und eine zutreffende Angabe aus einer vorgegebenen Liste ausgewählt werden.

15 Die Erhebungsstellen wurden nicht in der Fachanwendung Kodierung der Berufe geschult, da diese ausschließlich von den Statistischen Landesämtern genutzt wurde.

16 Für die Beleglesung waren ausschließlich die Fragebogen relevant, die für die Erhebung von Zusatzmerkmalen (Fragebogen HH2) eingesetzt wurden.

2. Durch eine automatische Kodierung, die anhand eines Abgleichs der eingegebenen Stichwörter mit der auf dem Klassifikationsserver hinterlegten Liste mit Berufen erfolgte.
3. Manuell, durch geschulte Mitarbeitende in den Statistischen Landesämtern, ebenfalls anhand der Klassifikationsserver-Liste.[17]

Ziel der elektronischen Unterstützung während des Erhebungsprozesses war, die Anzahl der manuell zu kodierenden Fälle im nachgelagerten Prozess der Aufbereitung so gering wie möglich zu halten. Die automatische und die manuelle Kodierung erfolgten in der Fachanwendung Kodierung der Berufe. Insgesamt konnten so 63 % der Fragebogen mit Berufsangaben (etwa 2,05 Millionen) entweder über den elektronischen IDEV-Fragebogen oder über die Fachanwendung Kodierung der Berufe automatisiert kodiert werden. Noch manuell kodiert wurden 37 % der Fragebogen mit Berufsangaben (etwa 1,2 Millionen).[18]

8.3 Integration, Plausibilisierung und Imputation

Demografische Angaben zu Personen lagen aus bis zu drei Quellen vor: aus dem Melderegister, der Erhebungsliste und dem Fragebogen. Während der Aufbereitung wurden diese Daten zusammengeführt und auf jeweils eine Angabe reduziert. Hierbei galt die Vorfahrtsregel: Melderegisterangaben vor Angaben aus der Erhebungsliste (Angaben aus den Interviews) vor Antworten aus dem Fragebogen. In einem zweiten Schritt wurden die Angaben zu den Zusatzmerkmalen Bildung und Erwerbstätigkeit auf Vollzähligkeit, Vollständigkeit, korrekte Befüllung und Widersprüchlichkeit geprüft und im Bedarfsfall maschinell bereinigt. Hierbei kam unter anderem die Imputations-Software CANCEIS[19] zum Einsatz, die nach dem Nearest-Neighbour-Prinzip arbeitet. Dabei werden fehlende oder fehlerhafte Werte eines Datensatzes (Empfänger) mit vollständigen und plausiblen Angaben eines anderen Datensatzes (Spender) derselben Erhebung ersetzt, sofern Spender- und Empfängerdatensätze sich in den übrigen vorhandenen Angaben überwiegend ähnlich sind (Bretnütz und andere, 2024).

8.4 Typisierung

Abschließend wurden aus dem nunmehr plausiblen Datenmaterial zusätzliche Merkmale abgeleitet, die nicht explizit durch die Datengewinnung bereitgestellt wurden. So errechnet sich beispielsweise das Merkmal Erwerbsstatus aus einer Vielzahl einzelner Erhebungsmerkmale rund um das Thema Erwerbstätigkeit. Diese sogenannten Auswertungsmerkmale wurden anhand der vorliegenden Erhebungsmerkmale (= Basismerkmale) im sogenannten Internen Auswertungssystem zum Zensus 2022[20] erstellt (typisiert) und schließlich über die Zensusdatenbank der interessierten Öffentlichkeit zur Verfügung gestellt.[21]

9

Fazit

Die Personenerhebung bildet eine der zentralen Säulen des registergestützten Zensusverfahrens in Deutschland. Dieser Beitrag zeichnet die Stationen der konzeptionellen und technischen Vorbereitungen für die Erhebung 2022 nach. Obwohl im Zensus 2011 bereits einmal eine Personenerhebung durchgeführt wurde, brachten die Vorbereitungen für den Zensus 2022 große organisatorische und technische Herausforderungen mit sich.

17 Die manuellen Kodierungen wurden darüber hinaus nach einem Vieraugenprinzip qualitätsgesichert.

18 Gründe für den verhältnismäßig hohen Anteil an manuellen Kodierungen liegen in der Qualitätssicherung. Die automatische Kodierung wurde nur bei sehr hoher Übereinstimmung verwendet.

19 Die Software CANCEIS (Canadian Census Edit and Imputation System) wurde 1992 vom kanadischen Statistikamt Statistics Canada/Statistique Canada für die Aufbereitung und Imputation von Zensusdaten entwickelt.

20 Das Interne Auswertungssystem zum Zensus 2022 ist eine zensusspezifische Anwendung, um die statistischen Einzeldaten aus den verschiedenen Erhebungsteilen sowie weitere Metadaten des Zensus 2022 langfristig sicher zu halten. Zugleich bietet die Anwendung die Möglichkeit flexibler und individueller Auswertungen einschließlich statistischer Geheimhaltung der Ergebnisse.

21 Bei der Zensusdatenbank handelt es sich um ein öffentlich zugängliches Recherchetool, mit dem die Ergebnisse des Zensus 2022 im Detail und in Tabellenform für die Regionalebenen Bund, Länder, Kreise, Gemeinden und Bezirke (Berlin, Hamburg) abgerufen und mit anderen Merkmalen kombiniert werden können. Zu beachten ist, dass die aus der Stichprobenerhebung hochgerechneten Ergebnisse zu Bildung und Erwerbstätigkeit nur für Gemeinden und Gemeindeverbände mit einer Bevölkerungszahl von mindestens 10 000 sowie flächendeckend ab Kreisebene ausgewiesen werden.

Hierbei konnten drei maßgebliche Erfolgsfaktoren identifiziert werden:

1. Ein zentraler Erfolgsfaktor bei der Bewältigung der vielfältigen Aufgaben der Personenerhebung waren die kaskadenartigen Schulungen. Sie haben die bundeseinheitlichen Erhebungsstandards sowie die Nutzungshinweise für die IT-Anwendungen innerhalb kürzester Zeit an die immens große Zahl der Erhebungsbeteiligten weitergegeben.
2. Um die hohen Qualitätsstandards für amtliche Daten zu erreichen, war – neben den bundesweit einheitlichen Erhebungsinstrumenten – vor allem das einheitliche Vorgehen beim Aufsuchen der Auskunftspflichtigen durch die Interviewerinnen und Interviewer entscheidend. Hierzu trugen auch maßgeblich die rollenadäquat aufbereiteten Schulungsunterlagen bei.
3. Nicht zuletzt brachte die zentrale Ansiedlung der IT-Systeme erhebliche Vorteile bei der Datenverarbeitung mit sich. Ebenfalls bewährte sich die inkrementelle Softwareentwicklung, die gewährleistete, dass Fehler frühzeitig erkannt und Anregungen der Nutzenden – auch noch zu einem fortgeschrittenen Zeitpunkt der Entwicklung – umgesetzt wurden.

LITERATURVERZEICHNIS

Bretnütz, Hella/Estatico, Sabrina/Walther, Sabrina/Hüsgen-Brodhäcker, Fabienne/Lange, Kerstin/Kleber, Birgit/Haberzettl, Benjamin. *Aufbereitung und Plausibilisierung der primärstatistischen Erhebungsteile im Zensus 2022*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 51 ff.

Bundesagentur für Arbeit. *Klassifikation der Berufe – überarbeitete Fassung 2020*. Nürnberg, 2021. [Zugriff am 7. November 2024]. Verfügbar unter: statistik.arbeitsagentur.de

Dittrich, Stefan. *Der registergestützte Zensus 2021. Weiterentwicklung des Zensus 2011 und weitere Schritte hin zu einem registerbasierten Zensus*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 5 ff.

Estatico, Sabrina/Tornow, Tobias/Walther, Sabrina. *Generalprobe für den Zensus 2022: Erkenntnisse aus der Pilotstudie zur Haushaltebefragung*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2021, Seite 126 ff.

Freier, Benjamin/Mosel, Juliane. *Online First als Leitgedanke für effiziente Primärerhebungen beim Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 46 ff.

Gaedke, Annika/Pfahl, Miriam/Strohalm, Anna. *Die Online-First-Strategie und die Online-Fragebogen im Zensus 2022*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 92 ff.

Hentschke, Janine/Tobies, Cara-Aileen/Weber, Susanne/Claus, Dennis. *Kern-Qualitätskennzahlen und Zielwerte im Zensus 2022*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 106 ff.

International Labour Organization. *International Standard Classification of Occupations 2008 (ISCO-08)*. Genf, 2012. [Zugriff am 7. November 2024]. Verfügbar unter: www.ilo.org

Neutze, Michael/Pfahl, Miriam/Schweinert-Albinus, Stefan. *Das Innovationspotenzial des Zensus 2021*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 37 ff.

Herausgeber
Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung
Dr. Daniel Vorgrimler
Redaktion: Ellen Römer

Ihr Kontakt zu uns
www.destatis.de/kontakt

Erscheinungsfolge
zweimonatlich, erschienen im Dezember 2024
Ältere Ausgaben finden Sie unter www.destatis.de sowie in der Statistischen Bibliothek.

Artikelnummer: 1010200-24006-4, ISSN 1619-2907

© Statistisches Bundesamt (Destatis), 2024
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.