



DIE HAUSHALTEGENERIERUNG IM ZENSUS 2022

Digitales Fachgespräch zum Zensus 2022



Gliederung

- (1) Ausgangslage und Ziele der Haushaltegenerierung
- (2) Überblick über die Haushaltegenerierung
- (3) Bestandteile der Haushaltegenerierung
 - Haushaltebildung nach harten Kriterien
 - Namensabgleich der Wohnungsnutzer aus der GWZ
 - Haushaltebildung nach statistischen Kriterien
 - Korrekturverfahren: Bereinigung des Personenbestands um Karteileichen und Fehlbestände



AUSGANGSLAGE UND ZIELE DER HAUSHALTEGENERIERUNG

1



Ausgangslage

1

- ▶ Keine amtlichen Register zu Haushalts-, Familien- und Wohnverhältnissen in Deutschland
- ▶ Unionsrechtliche Verpflichtung zur Bereitstellung von Daten im Rahmen des Zensus
- ▶ Lösung: Gewinnung der Daten durch Verknüpfung der Erhebungsteile des Zensus mithilfe statistischer Verfahren



Ziele der Haushaltegenerierung

1

- ▶ Verknüpfung der Erhebungsteile des Zensus zu einem zensustypischen Datensatz, der fachlich und regional kleinteilige Auswertungen auf Einzeldatensatzebene ermöglicht
- ▶ Korrektur des Personenbestands um hochgerechnete Karteileichen und Fehlbestände auf Einzeldatensatzebene (aber keine Veränderung der amtlichen Einwohnerzahl!)
- ▶ Bereitstellung von Daten zu Zahl und Struktur der Wohnhaushalte in Deutschland



ÜBERBLICK ÜBER DIE HAUSHALTEGENERIERUNG

2



Grundlegendes zur Haushaltegenerierung

2

- ▶ Automatisiertes, zentrales Verfahren ohne Benutzer-
eingriffe durch Bearbeiter
- ▶ Bedingung für den Start: Abschluss aller Erhebungsteile
- ▶ Haushaltsdefinition: Wohnhaushalt
- ▶ Grundprinzip: Nutzung von Register- und
Befragungsangaben soweit möglich, erst danach
Verwendung statistischer Verfahren



Verwendete Datenquellen

2

▶ Referenzdatenbestand:

- Anschriftenbestand
- Personenbestand (Melderegister und Ergebnisse der Personenerhebungen)

▶ Gebäude- und Wohnungszählung:

- Gebäude- und Wohnungsbestand aus der Befragung der Eigentümer

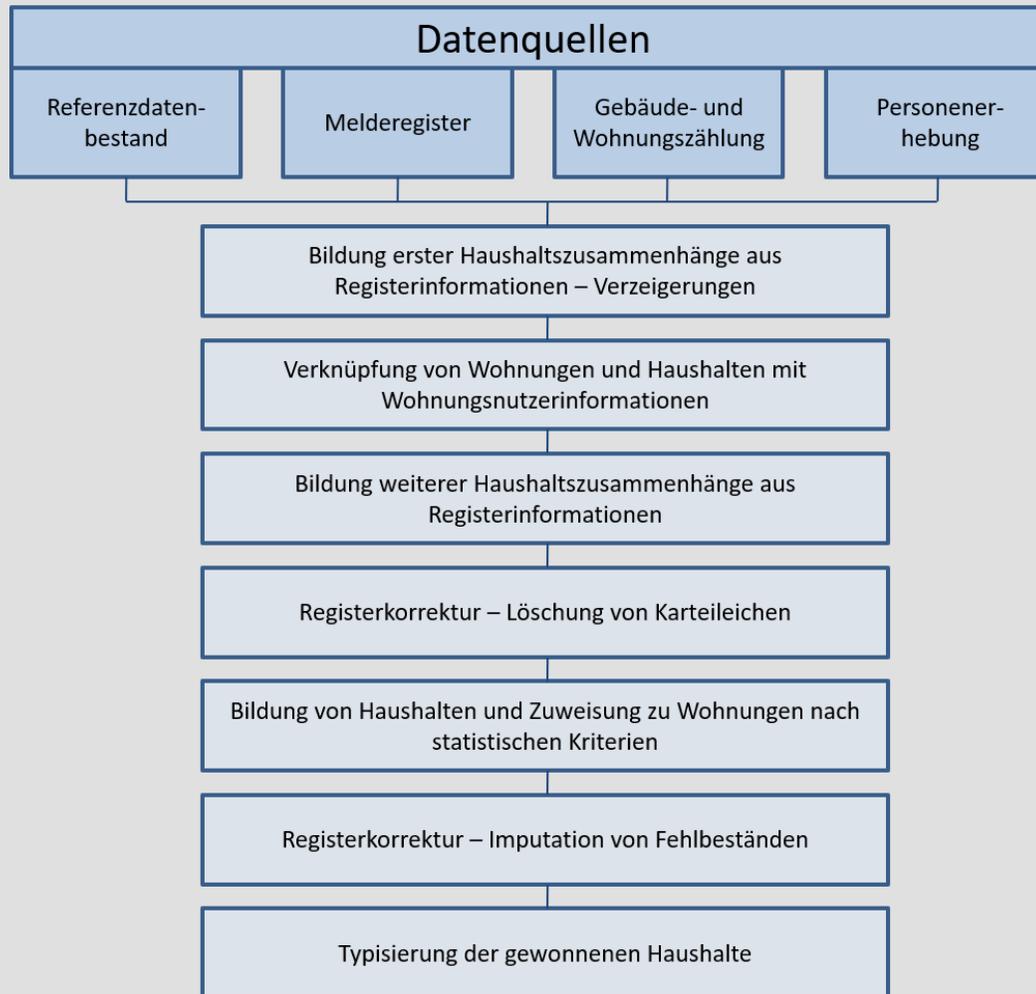
▶ Hochrechnungen aus der Haushaltsstichprobe:

- Einwohnerzahl insgesamt
- Demografische Eckwerte für Karteteileichen und Fehlbestände
- Hochgerechnete Kohorten für die Bevölkerungsfortschreibung



Ablauf der Haushaltegenerierung

2





BESTANDTEILE DER HAUSHALTEGENERIERUNG

3



Haushaltebildung nach harten Kriterien

3

- ▶ An Normalanschriften: Rekonstruktion der Haushaltszusammenhänge durch Nachbildung von Beziehungen
 - Prioritäre Auswertung von gesetzlichen Beziehungen in den Melderegistern (Verzeigerungen)
 - Erkennung von Haushaltszusammenhängen auf Basis weiterer Melderegisterinformationen (z.B. durch Namensvergleich)

- ▶ An Befragungsanschriften: Übernahme des Haushaltszusammenhangs aus der Befragung, Nachbildung von Beziehungen aus weiteren Angaben und Melderegisterinformationen



Namensabgleich der Wohnungsnutzer aus der GWZ 3

- ▶ Angabe von (bis zu) 2 Wohnungsnutzern in der GWZ durch den Eigentümer
- ▶ Probabilistischer Abgleich der Wohnungsnutzernamen mit allen Personendatensätzen einer Anschrift
- ▶ Bei hinreichender Übereinstimmung Verknüpfung von Personen und Wohnungen
- ▶ Aufdeckung weiterer Haushaltszusammenhänge (z.B. unverheiratete Paare)



Haushaltbildung nach statistischen Kriterien

3

- ▶ Anwendung statistischer Kriterien erst nach Ausschöpfung aller Möglichkeiten aus Melderegister- und Befragungsangaben
- ▶ Verknüpfung von Haushalten und Wohnungen auf Basis von Wohnflächenbedarf und der Nutzungsart von Wohnungen
- ▶ Bildung von Haushalten und Verknüpfung mit Wohnungen nach der geschätzten Häufigkeit von 3 Haushaltsgrößenklassen



Das Korrekturverfahren

3

- ▶ Korrektur des Einzeldatenbestands um Karteteileichen und Fehlbestände
 - Befragungsanschriften: Ergebnis der Existenzfeststellung
 - Normalanschriften: Hochrechnung auf Gemeindeebene

- ▶ Problem: Hochrechnung nur bzgl. der wichtigsten demografischen Merkmale möglich

- ▶ Ziel des Korrekturverfahrens ist eine möglichst verzerrungsfreie Löschung bzw. Imputation von Karteteileichen bzw. Fehlbeständen



Das Korrekturverfahren: Karteileichen

3

- ▶ Zweistufiges Verfahren
 - Optimierungsverfahren zur Bestimmung der multivariaten Verteilung der Karteileichen auf verschiedene demografische Klassen
 - Auswahl von konkreten Personen zur Löschung innerhalb der demografischen Klassen auf Basis von Karteileichenwahrscheinlichkeiten

- ▶ Wahrscheinlichkeitsprädiktion erfolgt auf Basis der Ergebnisse der Haushaltsstichprobe (Machine Learning)

- ▶ Vereinfachtes Verfahren für Nebenwohnsitzpersonen



Das Korrekturverfahren: Fehlbestände

3

- ▶ Anzahl der Fehlbestände ergibt sich aus den Karteileichen in der jeweiligen demografischen Klasse
- ▶ Imputation durch Kopie von konkreten Personen im Datenbestand, Verteilung auf Anschriften
- ▶ Allokationsproblem: Verteilung der imputierten Personen auf Anschriften innerhalb einer Gemeinde nach Haushaltsmerkmalen



**VIELEN DANK FÜR IHRE
AUFMERKSAMKEIT!**