

# B12 BAYERN IN ZAHLEN

Statistik-Magazin

Mehr als  
**3,1** Mio.  
Haushalte\*

## ZENSUS 2022

Trennungen und  
Zusammenlegungen  
in der Haushalte-  
generierung

## Die Cell-Key-Methode

in den Forschungs-  
datenzentren der  
Statistischen Ämter  
des Bundes und  
der Länder



modern,  
amtlich –  
und kostenlos

312

Kostenloses Abonnement  
der Online- und Print-Ausgabe  
unter [vertrieb@statistik.bayern.de](mailto:vertrieb@statistik.bayern.de)  
oder Telefon 0911 98208-6311  
[www.statistik.bayern.de](http://www.statistik.bayern.de)



## IMPRESSUM

**Bayern in Zahlen – Statistik-Magazin**  
Jahrgang 155. (78.)  
Bestell-Nr. Z10001 202408  
ISSN 0005-7215  
Erscheinungsweise monatlich

**Herausgeber, Druck und Vertrieb**  
Bayerisches Landesamt für Statistik  
Nürnberg Straße 95 | 90762 Fürth

**Bildnachweis**  
Titel: © fottoo – stock.adobe.com  
Innen: © Bayerisches Landesamt für Statistik  
(wenn nicht anders vermerkt)  
Wir danken der IG Fotografie des Landesamts  
für ihre Unterstützung.

**Papier**  
Gedruckt auf umweltfreundlichem Papier,  
chlorfrei gebleicht

**Vertrieb**  
E-Mail [vertrieb@statistik.bayern.de](mailto:vertrieb@statistik.bayern.de)  
Telefon 0911 98208-6311

**Auskunftsdienst**  
E-Mail [info@statistik.bayern.de](mailto:info@statistik.bayern.de)  
Telefon 0911 98208-6563

**Hinweis** Diese Druckschrift wird im Rahmen der Öffentlichkeitsarbeit der Bayerischen Staatsregierung herausgegeben. Sie darf weder von Parteien noch von Wahlwerbern oder Wahlhelfern im Zeitraum von fünf Monaten vor einer Wahl zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für Landtags-, Bundestags-, Kommunal- und Europawahlen. Missbräuchlich ist während dieser Zeit insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken und Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist gleichfalls die Weitergabe an Dritte zum Zwecke der Wahlwerbung. Auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl darf die Druckschrift nicht in einer Weise verwendet werden, die als Parteinahme der Staatsregierung zugunsten einzelner politischer Gruppen verstanden werden könnte. Den Parteien ist es gestattet, die Druckschrift zur Unterrichtung ihrer eigenen Mitglieder zu verwenden.

© Bayerisches Landesamt für Statistik,  
Fürth 2024

Vervielfältigung und Verbreitung, auch  
auszugsweise, mit Quellenangabe gestattet.

# Bayerische Industrie im Mai 2024 deutlich im Minus

## Umsätze um 7,3 Prozent und Auftragseingänge um 11,7 Prozent unter Vorjahresniveau

Im Mai 2024 verzeichnet das Verarbeitende Gewerbe Bayerns ein Umsatzminus von nominal 7,3 Prozent gegenüber dem entsprechenden Vorjahreszeitraum. Das Verarbeitende Gewerbe, bezogen auf den Berichtskreis „Betriebe mit 50 oder mehr Beschäftigten“, umfasst hierbei auch immer den Bergbau und die Gewinnung von Steinen und Erden.

Von den 37,9 Milliarden Euro Gesamtumsatz entfallen 23,2 Milliarden Euro auf Umsätze mit dem Ausland (– 6,1 Prozent) und darunter rund 7,3 Milliarden Euro auf Umsätze mit den Ländern der Eurozone (– 5,5 Prozent). Die Anteile der Auslandsumsätze und der Umsätze mit den Eurozonenländern am Gesamtumsatz machen 61,2 bzw. 19,3 Prozent aus.

Die Nachfrage nach Gütern des Verarbeitenden Gewerbes verringert sich im Mai 2024 gegenüber dem Vorjahresmonat preisbereinigt um 11,7 Prozent. Die Bestelleingänge aus dem Inland (– 17,0 Prozent) geben dabei im Vorjahresvergleich wesentlich stärker nach als die aus dem Ausland (– 8,3 Prozent).

Der Personalstand des Verarbeitenden Gewerbes liegt Ende Mai 2024 bei 1,196 Millionen Beschäftigten und damit um gut 200 Personen über dem Vorjahresmonat. ■

Ausführliche Ergebnisse enthalten die Statistischen Berichte „Verarbeitendes Gewerbe in Bayern im Mai 2024“ und „Index des Auftragseingangs für das Verarbeitende Gewerbe in Bayern im Mai 2024“. Kostenlos abrufbar unter [www.statistik.bayern.de/statistik/wirtschaft\\_handel/verarbeitendes\\_gewerbe](http://www.statistik.bayern.de/statistik/wirtschaft_handel/verarbeitendes_gewerbe)

Weitere kostenlos abrufbare Ergebnisse zum Auftragseingangsindex finden Sie online in unserer GENESIS Datenbank unter: [www.statistikdaten.bayern.de/genesis/online](http://www.statistikdaten.bayern.de/genesis/online)

**Mai  
2023**



**Mai  
2024**



Nachfrage nach Gütern des Verarbeitenden Gewerbes



Schloss Neuschwanstein



# INHALT

08 | 2024



2 IMPRESSUM

3 NACHRICHT DES MONATS

6 STATISTIK KOMMUNAL  
Passau

8 KURZ MITGETEILT

ÖFFENTLICH PRÄSENTIERT

18 **Workshop** „Zwischen Narrativen und Fakten:  
Wie können Statistiken die Öffentlichkeit erreichen?“

22 **75 Jahre Bayerischer Rundfunk**  
Das Landesamt als Teil der Kärwa mit dabei

24 DIE CELL-KEY-METHODE  
in den Forschungsdatenzentren der Statistischen  
Ämter des Bundes und der Länder (Teil 1)  
*Stefanie Setzer, Johannes Rohde,  
Volker Güttgemanns, Patrick Rothe*

40 TRENNUNGEN UND ZUSAMMENLEGUNGEN  
IN DER HAUSHALTEGENERIERUNG  
IM ZENSUS 2022  
Mathematische Modellierung  
und algorithmische Verfahren  
*Dr. Dirk Meierling, Dipl.-Math.*

62 NACHGEFRAGT BEI Dr. Dirk Meierling

70 ZURÜCKGEBLICKT  
ZENSUS 2011: Ablauf der Haushaltegenerierung

82 VERBRAUCHERPREISINDEX – Juni 2024

84 BAYERISCHER ZAHLENSPIEGEL

104 STATISTIK BEWEGT  
Tourismus in Bayern im Jahr 2023 –  
Burgen und Schlösser

105 NEU ERSCIENEN / ZEICHENERKLÄRUNG





Landwirtschaftliche  
Betriebe

48

2020\*

Einwohner

53 907

31.12.2022\*

Siedlungs- und  
Verkehrsfläche

32,7%

31.12.2022\*

## Mehr Daten zu Passau

sowie den anderen 24 kreisfreien Städten und den 2031 kreisangehörigen Gemeinden Bayerns in:

\* Statistik kommunal: [https://s.bayern.de/statistik\\_kommunal](https://s.bayern.de/statistik_kommunal)

\*\* Stadt.Land.Zahl: [www.statistikportal.de/de/stadt-land-zahl](http://www.statistikportal.de/de/stadt-land-zahl)

\*\*\* Statistisches Jahrbuch für Bayern: <https://s.bayern.de/jahrbuch>

© Geobasisdaten: Bayerische Vermessungsverwaltung 2024

Alle Daten abgerufen am 15.07.2024.





# PASSAU

## NIEDERBAYERN

Durchschnittsalter

43,6

31.12.2022\*\*

Gewässer

7,3%

31.12.2022\*

Bruttoinlands-  
produkt (BIP)  
je Einwohner

70 466

2021\*\*

Rang 22 von 400 kreisfreien  
Städten und Kreisen in  
Deutschland



# BAYERN WEITERHIN MAGNET FÜR ZUZÜGE: Wanderungsgewinn im Jahr 2023 bei über 99 000 Personen

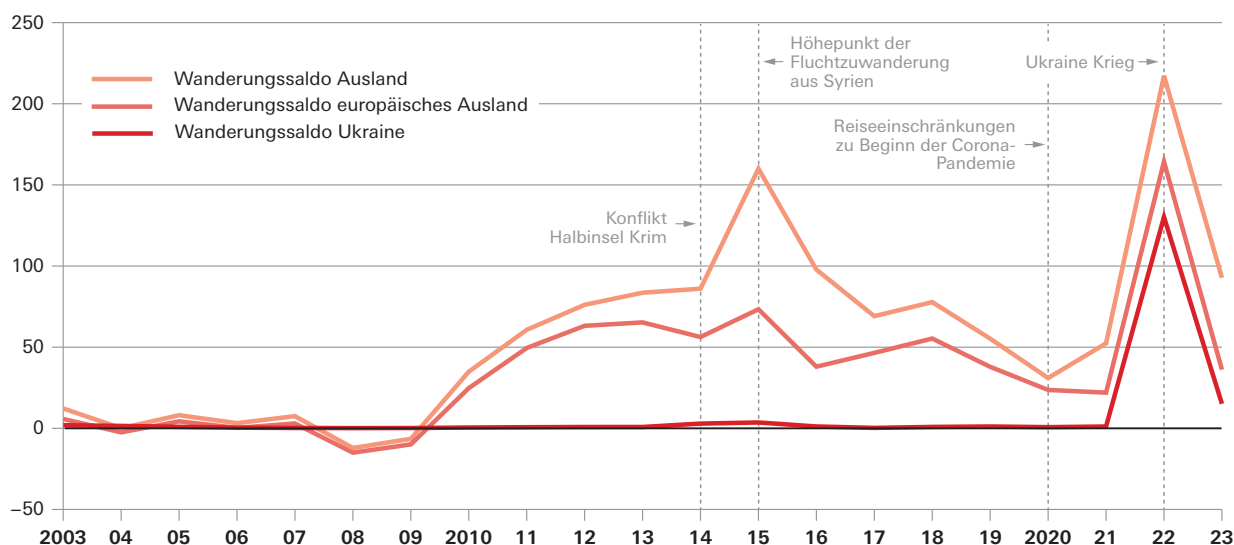
*Saldo unter dem Wert des Vorjahres, aber fast doppelt so hoch wie 2021*

Die Wanderungsbilanz Bayerns wies im Jahr 2023 einen Überschuss in Höhe von 99 124 Personen auf. Es ergab sich der Saldo aus insgesamt 455 877 Zuzügen und 356 753 Fortzügen über die Grenzen des Freistaats.

Aus dem Ausland zogen im abgelaufenen Berichtsjahr 335 582 Personen nach Bayern, aus dem restlichen Bundesgebiet 120 295 Personen. Von den Fortziehenden verließen 242 842 Personen den Freistaat in Richtung Ausland, während 113 911 ins restliche Bundesgebiet zogen. Gegenüber dem Ausland erzielte

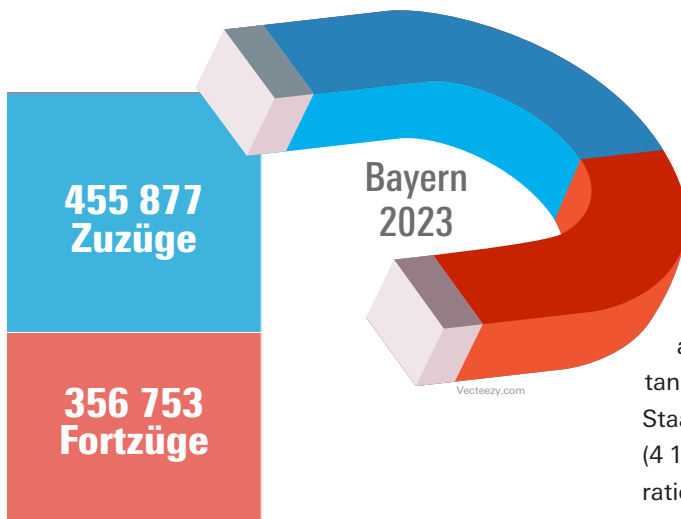
Bayern im Jahr 2023 somit einen Wanderungsgewinn in Höhe von 92 740 Personen, gegenüber dem restlichen Bundesgebiet in Höhe von 6 384 Personen. Insgesamt lag der Wanderungsgewinn im Jahr 2023 mit 99 124 Personen circa 123 000 Personen unter dem des Vorjahres, als 222 568 Personen (Ausland: 217 255, restliches Bundesgebiet: 5 313) mehr zu- als fortgezogen sind. Dieser hohe Wert war stark beeinflusst durch die verstärkte Zuwanderung aufgrund des Angriffskrieges in der Ukraine. Im Vergleich zum Jahr 2021 hat sich der Saldo von 52 772 auf 99 124 Personen fast verdoppelt.

**Wanderungssaldo Bayerns über die Bundesgrenze im Zeitraum 2003 bis 2023**  
in Tausend



Stand: 15.06.2024





Betrachtet man einzelne Staatsangehörigkeiten, so gab es im Jahr 2023 die höchsten positiven Auslandswanderungssalden bei Personen aus der Ukraine (15 113), gefolgt von Personen aus Syrien (13 628), der Türkei (12 696), Afghanistan (7 741), Indien (6 939), dem Kosovo (5 062) sowie Staatsangehörigen aus Bosnien und Herzegowina (4 178), Rumänien (3 296) und der Russischen Föderation (3 117).

Weiter entfielen 46 Prozent des Wanderungsüberschusses gegenüber dem Ausland auf Personen mit einer asiatischen Staatsangehörigkeit (42 870) und 39 Prozent auf Personen mit einer europäischen Staatsangehörigkeit (36 165). Zu dem hohen Wanderungsüberschuss aus Europa hat insbesondere die weiterhin verstärkte Zuwanderung von ukrainischen Staatsangehörigen (15 113) aufgrund des Ukraine-Krieges geführt. Zehn Prozent des positiven Wanderungssaldos mit dem Ausland sind auf Personen mit einer afrikanischen Staatsangehörigkeit (9 277) zurückzuführen. Etwas über vier Prozent des Wanderungsgewinns gegenüber dem Ausland wurde durch Personen mit einer nord- oder südamerikanischen Staatsangehörigkeit (3 829) erzielt.

Innerhalb Bayerns kam es im Jahr 2023 zu 609 280 übergemeindlichen Wanderungsbewegungen, was im Vergleich zum Vorjahr (2022: 589 626) ein Plus von 19 654 innerbayerischen Umzügen darstellt.

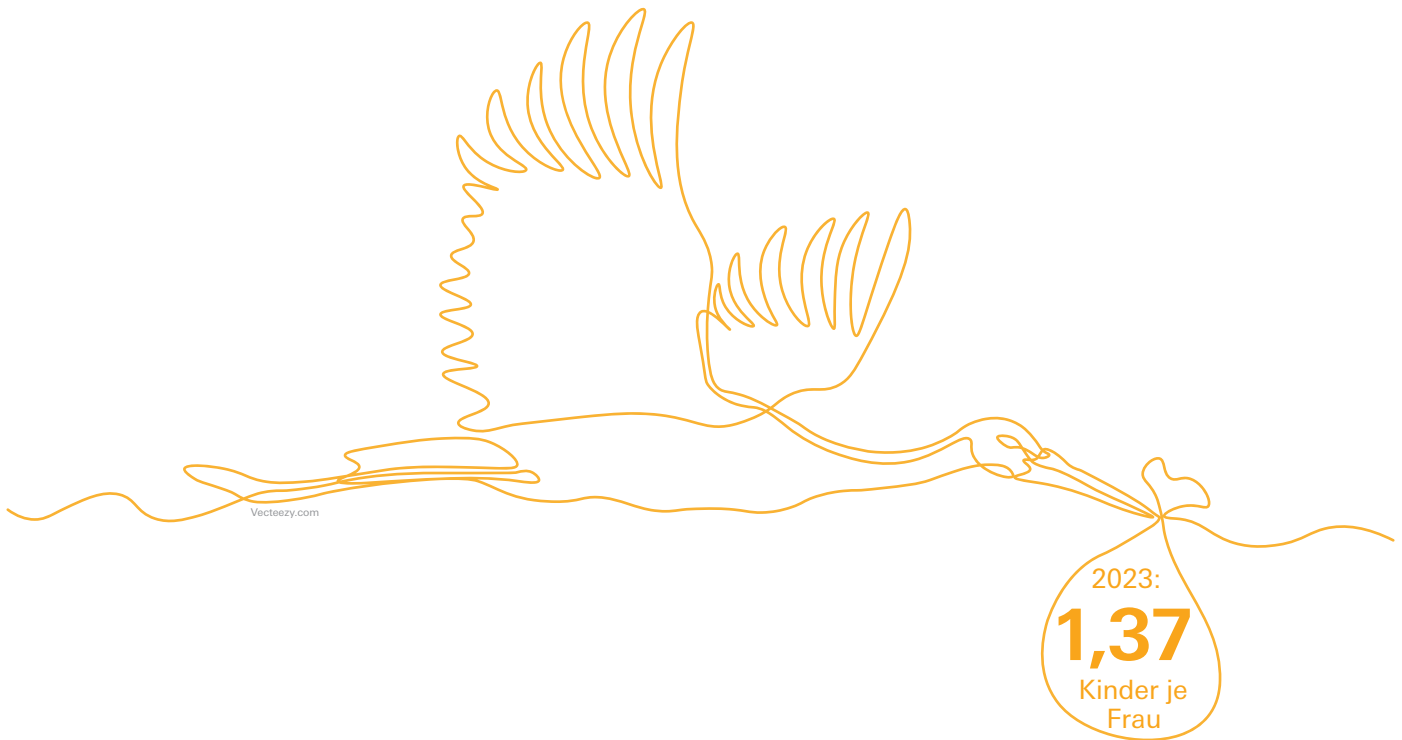
Auf Regierungsbezirksebene konnten im Jahr 2023 alle Regierungsbezirke ein deutliches Wanderungsplus verbuchen (Oberbayern: 23 265, Niederbayern: 11 219, Oberpfalz: 11 756, Oberfranken: 9 835, Mittelfranken: 13 681, Unterfranken: 9 434, Schwaben: 19 934). ■

Ausführliche Informationen erhalten Sie kostenlos unter [www.statistik.bayern.de/statistik/gebiet\\_bevoelkerung/wanderungen/index.html#link\\_1](http://www.statistik.bayern.de/statistik/gebiet_bevoelkerung/wanderungen/index.html#link_1).

Daten zu den Wanderungen in Bayern erhalten Sie über folgenden Link: [www.statistikdaten.bayern.de/genesis/online?language=de&sequenz=tabellen&selectionname=12711](http://www.statistikdaten.bayern.de/genesis/online?language=de&sequenz=tabellen&selectionname=12711)

### Wanderungssaldo über die Bundesgrenzen im Zeitraum 2018 bis 2023

Kontinente aus Staatsangehörigkeiten	Wanderungssaldo über Bundesgrenzen											
	2018		2019		2020		2021		2022		2023	
	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %
Europa .....	55 319	71	37 892	69	23 647	76	21 991	42	164 309	76	36 165	39
Asien.....	13 797	18	11 706	21	5 587	18	25 288	48	42 740	20	42 870	46
Amerika .....	3 043	4	3 154	6	803	3	3 028	6	3 374	2	3 829	4
Afrika .....	5 361	7	2 300	4	946	3	1 940	4	6 380	3	9 277	10
Australien und Ozeanien .....	111	0	122	0	- 63	0	24	0	18	0	99	0
Staatenlos, unbekannt, ungeklärt, ohne Angabe .....	152	0	71	0	73	0	91	0	434	0	500	1
<b>Insgesamt .....</b>	<b>77 783</b>	<b>100</b>	<b>55 245</b>	<b>100</b>	<b>30 993</b>	<b>100</b>	<b>52 362</b>	<b>100</b>	<b>217 255</b>	<b>100</b>	<b>92 740</b>	<b>100</b>



## ZUSAMMENGEFASSTE GEBURTENZIFFER IN BAYERN

im Jahr 2023 bei 1,37 Kindern je Frau

*Nach hoher zusammengefasster Geburtenziffer im Jahr 2021  
erneut deutlicher Rückgang*

Im Jahr 2023 nimmt im Freistaat neben der Zahl der Lebendgeborenen auch die zusammengefasste Geburtenziffer ab und erreicht einen Wert von 1,37 Kindern je Frau. Zum Vergleich: Im Zeitraum von 2000 bis 2015 lag die TFR in Bayern bei durchschnittlich 1,38 Kindern pro Frau (niedrigster Wert 2006: 1,32; höchster Wert 2015: 1,48) und pendelte sich in den Jahren 2016 bis 2020 bei 1,55 bzw. 1,56 Kindern pro Frau ein. Im Jahr 2021 – dem ersten Jahr nach Beginn der Corona-Pandemie – erreichte die TFR mit 1,61 Kindern je Frau den höchsten Wert seit 1972. Ob die zuletzt niedrige TFR eine Schwankung darstellt oder sich weiter verstetigt, wird sich in den nächsten Jahren zeigen.

Die TFR ähnelt im langjährigen Vergleich in allen sieben Regierungsbezirken Bayerns dem Niveau von vor 10 Jahren:

- Schwaben: 1,48 Kindern je Frau (2013: 1,45)
- Niederbayern: 1,41 (2013: 1,40)
- Oberpfalz: 1,41 (2013: 1,35)
- Unterfranken: 1,39 (2013: 1,38)
- Mittelfranken: 1,35 (2013: 1,41)
- Oberbayern: 1,34 (2013: 1,43)
- Oberfranken: 1,33 (2013: 1,33)



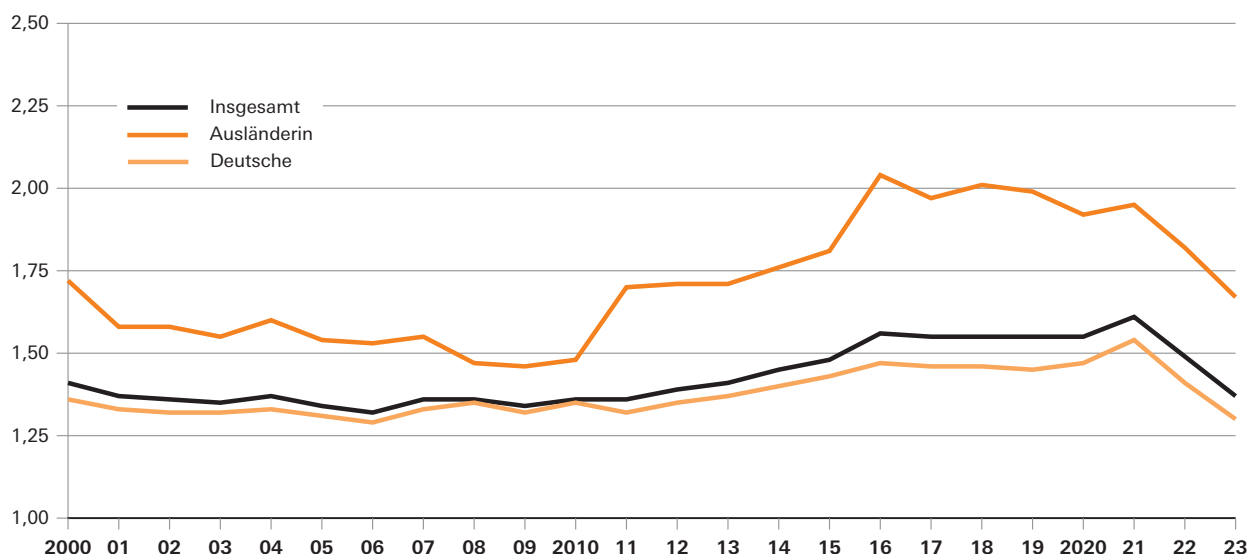
Auf Kreisebene weist auch im Jahr 2023 der Landkreis Neustadt a.d.Waldnaab mit 1,77 Kindern je Frau (2013: 1,29) bayernweit die höchste zusammengefasste Geburtenziffer auf, gefolgt von der kreisfreien Stadt Kaufbeuren mit 1,71 Kindern je Frau (2013: 1,55) und den Landkreisen Amberg-Weizsach mit 1,67 Kindern je Frau (2013: 1,34), Unterallgäu mit 1,67 Kindern je Frau (2013: 1,58) und Donau-Ries mit 1,61 Kindern je Frau (2013: 1,61). Die niedrigsten Werte verzeichnen die Universitätsstädte Bayreuth mit 0,95 Kindern je Frau (2013: 1,19), Passau mit 0,98 Kindern je Frau (2013: 1,12), Regensburg mit 1,07 Kindern je Frau (2013: 1,23), Bamberg mit 1,09 Kindern je Frau (2013: 1,30), Würzburg mit 1,11 Kindern je Frau (2013: 1,19) und Erlangen mit 1,18 Kindern je Frau (2013: 1,39). Städte mit Universitäten haben zwar eine vergleichsweise junge Bevölkerung mit vielen Frauen im fertilen Alter, bei den meisten dieser Frauen liegt der Fokus aber auf der Ausbildung oder dem Berufseinstieg. Entsprechend fällt die Anzahl der Lebendgeborenen auf eine große Gesamtzahl an Frauen zwischen 15 und 49 Jahren, wodurch die TFR in diesen Städten relativ niedrig ist.

Wird die zusammengefasste Geburtenziffer im Freistaat getrennt für Frauen mit deutscher und ausländischer Staatsangehörigkeit betrachtet, ergibt sich für das Jahr 2023 folgendes Bild: Bei Frauen mit deutscher Staatsangehörigkeit betrug die zusammengefasste Geburtenziffer 1,30 Kinder je Frau (2013: 1,37); bei Frauen mit ausländischer Staatsangehörigkeit 1,67 Kinder je Frau (2013: 1,71).

Die zusammengefasste Geburtenziffer ist ein Maß zur Beschreibung des aktuellen Geburtenverhaltens und gibt an, wie viele Kinder eine Frau durchschnittlich in ihrem Leben bekommen würde, wenn ihr Geburtenverhalten so wäre, wie das aller Frauen zwischen 15 und 49 Jahren im jeweils betrachteten Kalenderjahr. Sie wird auch totale Fertilitätsrate oder Gesamtfertilitätsrate (engl. Total Fertility Rate; TFR) genannt. ■

Die hohe zusammengefasste Geburtenziffer in Landkreisen wie z. B. Neustadt a.d.Waldnaab oder Amberg-Weizsach ist auch auf die dort stationierten US-Streitkräfte zurückzuführen. Seit dem Jahr 2018 werden in der amtlichen Statistik teilweise auch Geburten ausländischer Streitkräfte erfasst und können nicht mehr von den übrigen Geburten unterschieden werden. Dementsprechend hat sich die Gesamtzahl der – in der Geburtenstatistik ausgewiesenen – Geburten in Landkreisen mit US-Streitkräften erhöht. Die ausländischen Streitkräfte selbst sind jedoch von der Meldepflicht befreit. Infolgedessen werden bei der Berechnung der zusammengefassten Geburtenziffer in diesen Landkreisen nicht alle zu den Geburten beitragenden Mütter berücksichtigt, was zu einer Überschätzung der TFR führt.

**Zusammengefasste Geburtenziffer in Bayern nach der Nationalität der Mutter von 2000 bis 2023**  
Kinder je Frau



# ZENSUS 2022:

## 13,04 Millionen Einwohner und 6,41 Millionen Wohnungen in Bayern

*Gebäude- und Wohnungszählung erfasst erstmals Energieträger und Nettokaltmiete*

### Bevölkerungszahl in Bayern niedriger als angenommen

Nach den jetzt vorliegenden Ergebnissen des Zensus 2022 lebten am 15. Mai 2022 13 039 684 Einwohner in Bayern. Im Durchschnitt hat der Freistaat zum Stichtag des Zensus, dem 15. Mai 2022, insgesamt 2,2 Prozent weniger Einwohner als nach der bisher gültigen amtlichen Einwohnerzahl aus der Bevölkerungsfortschreibung auf der Basis des Zensus 2011 angenommen. Zum 15. Mai 2022 wohnten 1 860 563 Ausländerinnen und Ausländer in Bayern, das sind 161 610 Personen (8,0 Prozent) weniger als in der amtlichen Bevölkerungsfortschreibung nachgewiesen. Zu der gesamten Korrektur der Bevölkerungszahlen um 291 435 Einwohner tragen damit Personen mit nicht-deutscher Staatsangehörigkeit 55,3 Prozent bei.

### Bevölkerung in den Städten

Unter den drei größten Städten Bayerns sieht man für Nürnberg keine Änderung zur Bevölkerungsfortschreibung. Augsburg hat eine Abweichung von –1,7 Prozent und München verzeichnet eine Abweichung von –2,0 Prozent Abweichung der Bevölkerungszahl nach unten gegenüber der Bevölkerungsfortschreibung auf Basis des Zensus 2011.



Quelle: Statistische Ämter des Bundes und der Länder

Bei den Städten zwischen 100 000 und 250 000 Einwohnern reicht die Bandbreite von –6,7 Prozent in Regensburg bis zu einem Plus von 2,7 Prozent in Würzburg. In Fürth sind die Abweichungen marginal, die Werte daher stabil. In den Städten zwischen 50 000 und 100 000 Einwohnern zeigen Bamberg (–7,9 Prozent), Landshut (–8,9 Prozent) und Kempten (–7,0 Prozent) höhere Abweichungen von der Bevölkerungsfortschreibung. Bei den anderen Städten in dieser Größenklasse halten sich die Werte gerade im Blick auf die durchschnittliche Abweichung von –2,2 Prozent in ganz Bayern im Rahmen. Neu-Ulm weist einen positiven Saldo mit 0,5 Prozent auf.



### Bevölkerungszahlen in den Gemeinden

Der Zensus 2022 ermittelt die Bevölkerungszahlen aller 10 786 Gemeinden Deutschlands. In 3 163 Gemeinden (29 Prozent) gab es kaum Unterschiede zur bisherigen Bevölkerungsfortschreibung (Abweichung nach oben oder unten weniger als ein Prozent). Darunter 503 bayerische Gemeinden. Bundesweit haben 5 989 Gemeinden (56 Prozent) zum 15. Mai 2022 mindestens ein Prozent weniger Einwohner als bisher ausgewiesen. In den bayerischen Gemeinden haben 1 432 mindestens ein Prozent an Einwohnern verloren.

### Abweichungen bei den kleinen Gemeinden etwas über dem Durchschnitt

Die Korrekturen der Bevölkerungszahlen liegen in den kleineren Gemeinden unter 10 000 Einwohnern auch in Bayern mit –2,5 Prozent über dem Landesdurchschnitt.

„Beim Zensus 2011 waren die kleinen Gemeinden noch von der Korrektur durch die Personenerhebung ausgenommen. Das Bundesverfassungsgericht hat 2018 die Methodik des registergestützten Zensus und den Zensus 2011 bestätigt, aber empfohlen, in Zukunft die Korrekturstichprobe auch auf die kleinen Gemeinden auszuweiten. Dieser Empfehlung ist der Bundesgesetzgeber gefolgt, sodass im Zensus 2022 erstmalig ein stichprobenbasiertes Korrekturverfahren auch bei Gemeinden unter 10 000 Einwohnern durchgeführt wurde“, so Dr. Thomas Gößl, Präsident des Bayerischen Landesamts für Statistik.

Die Gemeinden unter 10 000 Einwohnern sind für Bayern besonders wichtig, weil insgesamt 1 829 Gemeinden (89 Prozent) in diese Größenklasse fallen. Dabei gibt es in 15 Gemeinden Abweichungen über –10 Prozent, die bis zu –18,5 Prozent in Reit im Winkl und –38,0 Prozent in Balderschwang reichen. Ähnlich hohe Abweichungen wie in diesen klassischen Ferienorten in Bayern zeigen sich in den Feriengebieten an Nord- und Ostsee. Die amtliche Statistik wird zusammen mit den Gemeinden analysieren, inwieweit ein hoher Einsatz von Saisonarbeitskräften oder nur nominelle Erstwohnsitzanmeldungen zu den Abweichungen beiträgt.

### Zensus gibt Auskunft über Demographie in Bayern

Das durchschnittliche Alter der Bevölkerung in Bayern liegt bei 43,7 Jahren und damit knapp unter dem bundesweiten Durchschnitt (44,3 Jahre). Im Vergleich zum Zensus 2011, ist das Durchschnittsalter in Bayern um mehr als ein Jahr gestiegen (42,5 Jahre). Regionale Unterschiede werden in der Altersstruktur erkennbar. So haben Schwaben und Oberbayern den größten Anteil von Kindern- und Jugendlichen, während in den übrigen Regierungsbezirken deutlich mehr ältere Menschen leben.

### Gebäude- und Wohnungszählung

Am Zensus-Stichtag gab es in Bayern 3 135 414 Wohngebäude und darin insgesamt 6 412 657 Wohnungen. Damit entspricht die Zahl der Wohnungen aus der Gebäude- und Wohnungszählung im Wesentlichen der Zahl der Wohnungen aus der Bestandsfortschreibung vom 31. Dezember 2021 mit 6 434 162 Wohnungen in Wohngebäuden. Seit dem vorherigen Zensus 2011 ist die Zahl der Wohnungen in Wohngebäuden bayernweit um rund 648 214 Wohnungen gestiegen.

### Wohneigentumsquote

In Bayern lebten zum Stichtag 2022 insgesamt 48,6 Prozent der Haushalte in ihrer eigenen Wohnung – das sind etwa 5 Prozentpunkte mehr als im deutschen Mittel (43,7 Prozent). Die höchste Wohneigentumsquote hat weiterhin Niederbayern mit 58,6 Prozent. Der Regierungsbezirk mit dem geringsten Anteil ist Oberbayern mit 40,5 Prozent.

### Zweithöchster Mietpreis in Bayern nach Hamburg

Erstmals wurde beim Zensus 2022 die Nettokaltmiete für alle Wohnungen erhoben. Die durchschnittliche Nettokaltmiete lag in Bayern bei 8,74 Euro pro m<sup>2</sup>, der zweithöchste Mietpreis im Bundesvergleich nach Hamburg mit 9,16 Euro pro m<sup>2</sup>.

Regional lassen sich hier deutliche Unterschiede feststellen: Der Regierungsbezirk Oberbayern weist mit 10,95 Euro pro m<sup>2</sup> die höchsten, der Regierungsbezirk Oberfranken mit 5,99 Euro die niedrigsten durchschnittlichen Mietpreise auf. München ist bayernweiter Spitzenreiter unter den Landkreisen und kreisfreien Städten mit einer Nettokaltmiete von 12,89 Euro pro m<sup>2</sup>. Die teuerste Gemeinde aber ist Neubiberg mit einer durchschnittlichen Nettokaltmiete von 13,84 Euro pro m<sup>2</sup>. Überdurchschnittliche Mietpreise sind erwartungsgemäß für neuere Wohngebäude, die ab 2010 fertig gestellt wurden, sowie in und um die Großstädte erkennbar.

### Kontinuierlicher Ausbau bei der Nutzung regenerativer Energien

Nach den vorliegenden Ergebnissen werden fast drei Viertel (71,5 Prozent) aller Wohngebäude in Bayern mit Gas (34,4 Prozent) oder Öl (37,0 Prozent) beheizt und weitere 5,3 Prozent mit Fernwärme. Eine bayerische Besonderheit ist der hohe Einsatz von Holz und Holzpellets. Im bundesweiten Vergleich liegt der Freistaat mit 12,8 Prozent deutlich über dem gesamtdeutschen Einsatz von nur 5,8 Prozent. Erneuerbare Energiequellen zum Heizen von Wohngebäuden spielen im Gesamtbestand bislang eine untergeordnete Rolle.

Seit der Jahrtausendwende ist ein kontinuierlicher Ausbau bei der Nutzung regenerativen Energien sichtbar. In Wohngebäuden ab 2010 wurden verstärkt solar- und geothermische Anlagen eingebaut. In dieser Altersgruppe werden 35,2 Prozent der Wohngebäude mit Solar-/Geothermie oder Wärmepumpen geheizt. Hier lässt sich ein deutlicher Trend erkennen, denn in der Altersgruppe 1990 bis 2009 lag der durchschnittliche Wert noch bei 5,4 Prozent und den Altersgruppen davor bei unter zwei Prozent.

### Wie entsteht die amtliche Einwohnerzahl?

In allen Ländern ist es Aufgabe des statistischen Landesamtes, die amtliche Einwohnerzahl in den Gemeinden festzustellen, nach den Ergebnissen des Zensus und durch Verwaltungsakt. Die Verfahren zur Feststellung der amtlichen Einwohnerzahl beginnen nach der Sommerpause im September 2024. Dabei werden die Gemeinden in einer Anhörung umfassend informiert und beteiligt. Erst danach wird durch Verwaltungsakt die Einwohnerzahl festgestellt.

### Regionalisierte Veröffentlichungen zu den Ergebnissen online verfügbar

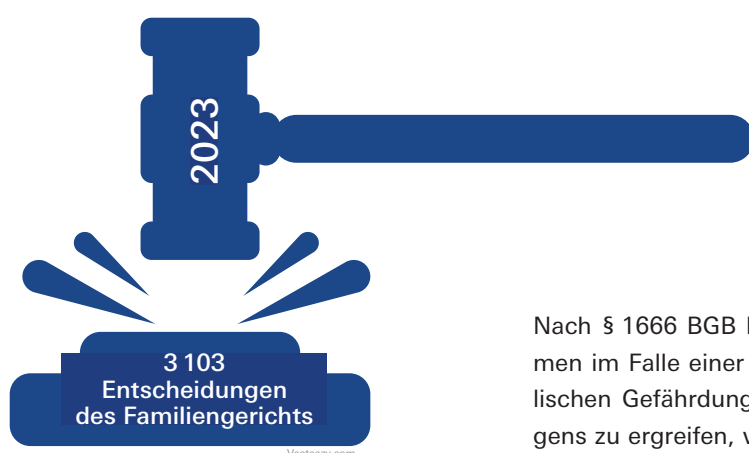
Das bayerische Landesamt für Statistik stellt für jede regionale Einheit ein Zensus-Dossier mit den wichtigsten Zahlen und Eckdaten aus dem Zensus 2022 kostenfrei zum Download zur Verfügung. Unter [www.zensus2022.bayern.de](http://www.zensus2022.bayern.de) sind die Dossiers als PDF-Datei oder Excel-Format abrufbar. ■

Weitere Informationen zum Zensus 2022 entnehmen Sie der Seite im Zensus Sonderheft unter <https://statistik.bayern.de/zensus2022>



# GERICHTLICHE MASSNAHMEN aufgrund einer Gefährdung des Kindeswohls im Jahr 2023 auf Vorjahresniveau

*Übertragung der elterlichen Sorge ganz oder teilweise  
auf das Jugendamt oder einen Dritten in 1 967 Fällen*



Nach § 1666 BGB hat das Familiengericht Maßnahmen im Falle einer körperlichen, geistigen oder seelischen Gefährdung des Kindes oder seines Vermögens zu ergreifen, wenn die Eltern nicht gewillt oder in der Lage sind, diese Gefahr abzuwenden.

Im Laufe des Jahres 2023 werden in Bayern insgesamt 3 103 Fällen Entscheidungen des Familiengerichts getroffen, um für Kinder und Jugendliche Maßnahmen einzuleiten. Im Vergleich zum Vorjahr ist die Fallzahl der Entscheidungen nahezu unverändert (2022: 3 107). In 1 585 Fällen sind Jungen und in 1 518 Fällen Mädchen betroffen.

In 1 967 Fällen wird der Entzug der elterlichen Sorge entschieden und diese somit auf das Jugendamt oder einen Dritten als Vormund oder Pfleger (gem. § 1666 Abs. 3 Nr. 6 BGB) übertragen – in 696 Fällen vollständig und in 1 271 Fällen teilweise.

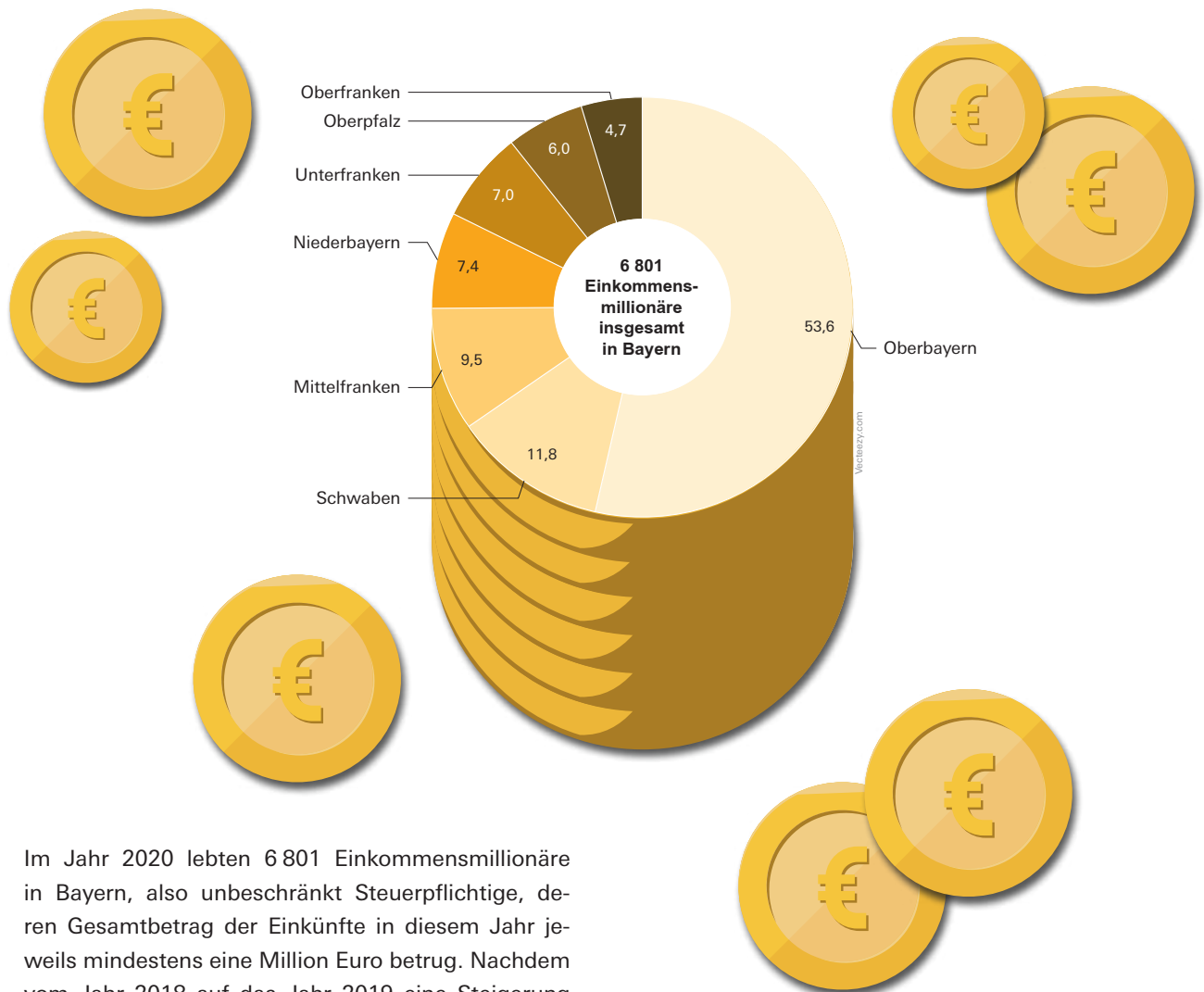
Des Weiteren wird in 646 Fällen die Inanspruchnahme von Leistungen der Kinder- und Jugendhilfe aufgelegt, in 107 Fällen werden die Erklärungen von Personensorgeberechtigten ersetzt und in 383 Fällen werden Ge- oder Verbote gegenüber Personensorgeberechtigten oder Dritten ausgesprochen.

Die Anrufung des Familiengerichts durch das Jugendamt ist immer dann erforderlich und verpflichtend, wenn sich die Gefährdung des Kindeswohls nur durch eine gerichtliche Entscheidung abwenden lässt. Das Familiengericht entscheidet dann im Einzelfall, ob und welche gerichtlichen Maßnahmen einzuleiten sind. Zu den möglichen Maßnahmen zählen Weisungen, Gebote und Verbote oder Auflagen zur Wahrnehmung der elterlichen Sorge. ■

Ausführliche Ergebnisse enthält der Statistische Bericht „Kinder- und Jugendhilfe Ergebnisse zu Teil I“, kostenlos abrufbar unter: [www.statistik.bayern.de/statistik/bildung\\_soiales/kinder\\_jugend\\_hilfe/index.html#link\\_4](http://www.statistik.bayern.de/statistik/bildung_soiales/kinder_jugend_hilfe/index.html#link_4)

# 6 801 EINKOMMENS MILLIONÄRE IN BAYERN IM JAHR 2020

*Höchste Millionärsdichte im Landkreis Starnberg*



Im Jahr 2020 lebten 6 801 Einkommensmillionäre in Bayern, also unbeschränkt Steuerpflichtige, deren Gesamtbetrag der Einkünfte in diesem Jahr jeweils mindestens eine Million Euro betrug. Nachdem vom Jahr 2018 auf das Jahr 2019 eine Steigerung um 4,5 Prozent zu verzeichnen war, erhöhte sich die Anzahl der Einkommensmillionäre abermals innerhalb eines Jahres (von 2019 auf 2020) um 436 Steuerpflichtige bzw. 6,8 Prozent. Im gleichen Zeitraum verzeichnete der Gesamtbetrag der Einkünfte der Einkommensmillionäre einen Anstieg von 2,3 Prozent und belief sich im Jahr 2020 auf insgesamt 18,2 Milliarden Euro. Im Durchschnitt kamen die Einkommensmillionäre im Jahr 2020 auf einen Gesamtbetrag der Einkünfte pro Steuerpflichtigen von 2,7 Millionen Euro.

Die Einkünfte aus Gewerbebetrieb, die mit 11,6 Milliarden Euro den Großteil (63,6 Prozent) der Einkünfte der bayerischen Einkommensmillionäre 2020 ausmachten, erhöhten sich um 0,3 Prozent im Vergleich zu 2019. Die Einkünfte aus nichtselbständiger Arbeit nahmen von 2019 auf 2020 um 4,8 Prozent zu und beliefen sich 2020 auf 3,3 Milliarden Euro. Dies entsprach einem Anteil von 18,0 Prozent am Gesamtbetrag der Einkünfte.

Die Einkommensmillionäre im Freistaat machten nur 0,1 Prozent aller unbeschränkt Steuerpflichtigen im Jahr 2020 aus, sie erzielten jedoch mit 5,3 Prozent einen überproportional hohen Anteil am Gesamtbetrag der Einkünfte. Gleichzeitig entrichteten sie mit 6,7 Milliarden Euro einen Anteil von 10,3 Prozent der gesamten festzusetzenden Einkommensteuer. Der Anteil an der festgesetzten Einkommensteuer fiel aufgrund des progressiven Steuersatzes höher aus als der Anteil am Gesamtbetrag der Einkünfte.

53,6 Prozent der bayerischen Einkommensmillionäre bzw. 3 644 unbeschränkt Steuerpflichtige hatten im Jahr 2020 ihren Wohnsitz in Oberbayern, wobei 2 042 in der Stadt oder im Landkreis München lebten. 803 Spitzenverdiener (11,8 Prozent) beherbergte der Regierungsbezirk Schwaben und 643 (9,5 Prozent) der Regierungsbezirk Mittelfranken. In Niederbayern waren 506 (7,4 Prozent), in Unterfranken 477 (7,0 Prozent), in der Oberpfalz 406 (6,0 Prozent) und in Oberfranken 322 (4,7 Prozent) Einkommensmillionäre ansässig.

Insgesamt kamen im Jahr 2020 in Bayern im Durchschnitt 5,2 Einkommensmillionäre auf 10 000 Einwohner, im Jahr 2019 waren es 4,8. Die höchste Millionärsdichte gab es im Landkreis Starnberg mit 22,8 Einkommensmillionären je 10 000 Einwohner. Es folgten der Landkreis Miesbach mit 14,2, der Landkreis München mit 14,0, die Stadt München mit 10,4 und die kreisfreie Stadt Landshut mit 9,7 Einkommensmillionären je 10 000 Einwohner. ■

Bei der Interpretation der Ergebnisse der Statistik ist zu berücksichtigen, dass diese sich auf unbeschränkt Steuerpflichtige beziehen und zusammenveranlagte Ehepaare und Personen in eingetragenen Lebenspartnerschaften als ein Steuerpflichtiger gezählt werden. Datengrundlage der Lohn- und Einkommensteuerstatistik sind die anonymisierten Ergebnisse des Steuerfestsetzungsverfahrens durch die Finanzverwaltung.

Der mehrjährige Abstand zwischen dem Statistikjahr und der Erstellung der Ergebnisse ist weitgehend auf die steuerrechtlich festgelegten Antrags- und Erklärungsfristen sowie auf den Zeitbedarf in den Finanzverwaltungen für die Bearbeitung der Steuererklärungen zurückzuführen.

Die hier dargestellten Ergebnisse für 2020 sind damit die aktuellsten verfügbaren Daten. Die Einwohnerzahlen basieren auf den Ergebnissen der Bevölkerungsfortschreibung auf Grundlage des Zensus 2011.

Ausführliche Ergebnisse enthält der Bericht „Einkommen der natürlichen Personen in Bayern 2020“. Kostenlos unter [www.statistik.bayern.de/statistik/haushalte\\_steuern/steuern/#link\\_1](http://www.statistik.bayern.de/statistik/haushalte_steuern/steuern/#link_1)







## Workshop „Zwischen Narrativen und Fakten: Wie können Statistiken die Öffentlichkeit erreichen?“

*am 3. Juni 2024 in Fürth*

„Traue keiner Statistik ...“ – jeder kennt diese Redewendung. Amtliche und wissenschaftliche Statistiken helfen uns, wirtschaftspolitische und gesellschaftliche Zusammenhänge besser zu verstehen. Wer sie richtig liest, tut sich leichter, zwischen Fakten und „Fake News“ zu unterscheiden. Aber Statistiken können fehlinterpretiert oder falsch dargestellt werden. Wie steht es um die Kompetenz, statistische Informationen zu verstehen und zu interpretieren? Wie können Nutzerinnen und Nutzer Statistiken selbst richtig lesen und korrekt in den Gesamtzusammenhang einordnen?

Diesen Fragen widmete sich der Workshop „Zwischen Narrativen und Fakten: Wie können Statistiken die Öffentlichkeit erreichen?“ am 3. Juni 2024 in Fürth. Der Workshop wurde vom Bayerischen Landesamt für Statistik gemeinsam mit dem Ludwig Erhard ifo Forschungszentrum für Soziale Marktwirtschaft und Institutionenökonomik in Fürth im Rahmen des Statistik Netzwerk Bayern organisiert. Ziel des Workshops war es, durch acht Vorträge eine Vielzahl von Expertinnen und Experten aus Forschung, Bildung und Journalismus zusammenzubringen und gemeinsam mit den etwa 100 Teilnehmerinnen und Teilnehmern über dieses Thema zu diskutieren.



Oben von links nach rechts: Prof. Dr. Joachim Weimann, Katharina Schüller, Prof. Dr. Ursula Münch.  
Unten von links nach rechts: Prof. Dr. Henrik Müller und die Moderatoren Nina Storfinger und Eberhard Beck.

Der Workshop, der von Herrn Eberhard Beck, ifo Institut, und Frau Nina Storfinger, Bayerisches Landesamt für Statistik, organisiert und moderiert wurde, gliederte sich in zwei Sessions, um das Thema aus unterschiedlichen Blickwinkeln betrachten zu können. Nach den Begrüßungsworten der beiden Initiatoren des Workshops, Dr. Thomas Gößl, Präsident des Bayerischen Landesamts für Statistik, und Prof. Dr. Sarah Necker, Leiterin Ludwig Erhard ifo Zentrum für Soziale Marktwirtschaft und Institutionenökonomik, beschäftigte sich die erste Session mit der Frage „Wie funktionieren Daten in den Medien“. Prof. Dr. Ursula Münch, Direktorin der Akademie für Politische Bildung in Tutzing, referierte in ihrer Keynote darüber, dass gerade in Zeiten von Kriegen und Krisen Menschen besonders empfänglich für gezielt verbreitete Desinformationen sind. Prof. Dr. Münch erläuterte, inwiefern Verschwörungsglauben und Wissenschaftsleugnung Gefahren für unsere freiheitlich demokratische Grundordnung darstellen und wie dieser Entwicklung begegnet werden kann. Der Keynote von Prof. Dr. Münch folgten drei weitere Vorträge,

die aufzeigten, wie Daten in den Medien verfälscht dargestellt werden, welche Fehlschlüsse aus den Darstellungen gezogen werden und wie Journalistinnen und Journalisten beim Aufbau von Statistikenkenntnissen unterstützt werden könnten.

### Veranstaltungshinweis

Am 19. September 2024 wird im Bayerischen Landesamt für Statistik in Fürth bereits der nächste Workshop, dieses Mal zum Thema statistische Geheimhaltung, stattfinden. Im Rahmen der Veranstaltung wird ein detaillierter Überblick zum aktuellen Stand der statistischen Geheimhaltung und neuen Entwicklungen im Bereich der Anonymisierung statistischer Daten innerhalb und außerhalb der statistischen Ämter gegeben. Interessierte sind herzlich zur Teilnahme eingeladen. Weitere Infos dazu sind unter [www.statistiknetzwerk.bayern.de/themen/workshops/index.html](http://www.statistiknetzwerk.bayern.de/themen/workshops/index.html) zu finden.

Die zweite Session stellte die Frage „Wie wirken Daten auf den Menschen?“ in den Mittelpunkt. Prof. Dr. Florian Zimmermann, Professor an der Universität Bonn und Forschungsdirektor am Institut zur Zukunft der Arbeit GmbH, eröffnete die Session mit einer Keynote. Seine Vorstellung einer Forschungsstudie war ein maßgeblicher Impulsgeber für den Workshop. In der referierten Studie wurde der Einfluss von Geschichten auf die Überzeugungen im Gegensatz zum Einfluss von Statistiken auf die Überzeugungen untersucht. Es zeigte sich, dass der durchschnittliche Einfluss von Geschichten im Lauf des Tages um 33 Prozent abnahm, bei Statistiken dagegen um 73 Prozent. Die sich in dieser Session angeschlossenen Vorträge zeigten ebenfalls eindrücklich, wie Daten von Menschen, das heißt den datenkonsumierenden Personen, aufgegriffen und oftmals fehlinterpretiert werden und welche Folgen dies mit sich bringt beziehungsweise mit sich bringen könnte.

Den Abschluss des Workshops bildete eine Podiumsdiskussion, an der Prof. Dr. Ursula Münch, Prof. Dr. Florian Zimmermann sowie Tassilo Forchheimer, Leiter des BR Franken, teilnahmen. Moderiert wurde die Diskussion von Dr. Thomas Gößl und Prof. Dr. Sarah Necker. Kernaussagen der Diskussion waren unter anderem die Wichtigkeit eines vertrau-

ensvollen Miteinanders von Medien und Datenproduzenten, sowie die Möglichkeit, sich als Medienschaffender intensiv in Datensätze einzuarbeiten. Darüber hinaus wurde betont, dass die amtliche Statistik mit ihrer gesetzlich verankerten Neutralität einen wichtigen Beitrag zur Verringerung von Fehlinterpretationen leistet. Zudem sollte die amtliche Statistik auch darauf hinweisen, welche Informationen aus den Zahlen herausgelesen werden können und welche nicht. Gleichzeitig wurde herausgehoben, dass die Wissenschaft auch als Ausbildungsstätte von künftigen Statistikerinnen und Statistikern sowie Datenjournalistinnen und Datenjournalisten wesentlich zur Steigerung von Datenkompetenzen beiträgt. Weiterer zentraler Punkt in der Diskussion war, dass verlässliche und möglichst vollständige Daten benötigt werden, um gute Narrative bilden zu können.

Die Teilnehmerinnen und Teilnehmer tauschten sich im Anschluss an die Veranstaltung mit den Referentinnen und Referenten im Rahmen eines gemeinsamen Abendessens im Stadtparkcafé Fürth aus. ■

*Dipl-Soz. Nina Storfinger*

Weitere Informationen zum Programm und zu den Referentinnen und Referenten finden Sie unter: [www.statistiknetzwerk.bayern.de/themen/workshops/narrative\\_und\\_fakten/index.html](http://www.statistiknetzwerk.bayern.de/themen/workshops/narrative_und_fakten/index.html)



Von links nach rechts: Prof. Dr. Sarah Necker, Tassilo Forchheimer, Prof. Dr. Ursula Münch, Prof. Dr. Florian Zimmermann, Dr. Thomas Gößl.



# TEAMPLAYER GESUCHT!



Aktuelle Stellenangebote  
sowie weitere Informationen  
zum Bewerbungsverfahren  
finden Sie unter:  
**[www.statistik.bayern.de/karriere](http://www.statistik.bayern.de/karriere)**



Bayerisches Landesamt für  
Statistik





*„Unser Infostand auf der BR-Kärwa war aus meiner Sicht ein voller Erfolg – viele Besucher, interessante Gäste und eine gelungene Atmosphäre. Die Zusammenarbeit mit den Kolleginnen und Kollegen vor Ort war sehr angenehm und es war toll, unser Landesamt auf dieser Veranstaltung zu repräsentieren.“*

*– Ulrike Schaumburg –*

*„Da ist ja richtig was los!“*

*– Margit Lott –*

## 75 Jahre Bayerischer Rundfunk

### Das Landesamt als Teil der Kärwa mit dabei

*Im Jahr 1949 eröffnete der erste BR-Intendant Rudolf Scholtz das Gelände des Bayerischen Rundfunks in der Wallensteinstraße in Nürnberg. 75 Jahre später feiert der BR im Rahmen einer „Kärwa im Park“ sein Jubiläum drei Tage lang mit zahlreichen Fahrgeschäften, Mitmachaktionen und Blicken hinter die Kulissen. Das Bayerische Landesamt für Statistik folgte der Einladung und nutzte die Gelegenheit, sich den zahlreichen Besucherinnen und Besuchern mit einem Stand am Marktplatz zu präsentieren.*

*„Ich empfand die Stimmung als gut. Mein Eindruck war, dass die Besucher recht interessiert waren an den Statistiken und diese auch als realitätsnah dann empfanden.“*

*– Jasmin Gründel –*

*„Unsere Präsenz war sinnvoll: Viele Besucher erlebten einen Aha-Effekt, da sie uns bisher nicht kannten. Viel Interesse an unseren Plakaten und etliche kluge Fragen.“*

*– Dr. Christina Wübbecke –*



*„Super Veranstaltung mit jeder Menge netten und wissbegierigen Leuten, denen man die amtliche Statistik näherbringen konnte. Toller Auftritt vom Landesamt.“*

*– Pascal Forster –*

### Statistik und Tradition verbinden

Traditionell wurde die Kärwa am Freitagnachmittag auf dem Gelände des BR mit einem Fassanstich eröffnet. Passend zum Thema einer Kirchweih präsentierte das Bayerische Landesamt für Statistik aus seinem Statistikportfolio die Entwicklung der Bierpreise der letzten 18 Jahre und stellte eine Übersicht der Reallohne vor, die dem Motto „Wie viel Geld kann ich auf einer Kärwa ausgeben?“, folgte. Besonders beliebt bei den vielen Gästen war das Schätzspiel zur Anzahl der Lebendgeburten und dem Durchschnittsalter der Mutter bei der Geburt des ersten Kindes. Auf diese Weise wurde unterhaltsam der Nutzen der amtlichen Statistik im lockeren Rahmen vermittelt. Statistik zum Anfassen. Nicht nur das Herzstück der amtlichen Statistik, wie die Datenbank Genesis oder die Laufenden Wirtschaftsrechnungen, waren Teil der Präsentation. Mit anschaulichen Plakaten, Flyern und Tablets zeigte man das breite und vielfältige Angebot der statistischen Erhebungen und Veröffentlichungen. Viele Besucherinnen und Besucher erinnerten sich an die alte Quelleverwaltung und waren überrascht, dass dieses Haus nun die amtliche Statistik beheimatet.



### Durchweg hoher Andrang am Stand des Landesamts

Drei Tage lang kamen viele Gäste zum Stand des Landesamts. Die Besucherinnen und Besucher interessierten sich nicht nur für die einzelnen Statistiken, sondern auch generell für das Amt an sich, wie viele Beschäftigte hier arbeiten und welche Aufgaben die amtliche Statistik hat. Die durchweg positive Resonanz bestätigt: Solche Veranstaltungen sind wichtig. Das Amt sieht darin einen großen Nutzen, den gesetzlich verankerten Informationsauftrag in verständlicher und interessanter Form der Bevölkerung zu vermitteln. Man hat gemerkt, dass das Landesamt in Fürth trotz der räumlichen Nähe noch nicht so bekannt ist. Und auf die Frage, was das Landesamt mit 75 Jahren BR zu tun habe, folgt die Antwort, dass das Landesamt für Statistik genau die qualitativ hochwertigen Daten zur Verfügung stellt, die die Medienlandschaft für ihre seriöse Berichterstattung benötigt.

Ein großer Dank geht an die mehr als 20 Mitarbeiterinnen und Mitarbeiter, die den Stand über die drei Tage hinweg begleitet und betreut haben sowie an die vielen Kolleginnen und Kollegen im Haus, die im Voraus an der Planung und Umsetzung beteiligt waren und auch auf einen Besuch vorbeigekommen sind. ■

*„Die Leute in Bayern lieben Ratespiele und gutes Merchandise.“*

*– Lena Erhardt –*

*Franziska Strauch  
Tobias Schwär*





# DIE CELL-KEY- METHODE

**in den Forschungsdatenzentren  
der Statistischen Ämter des Bundes  
und der Länder**

**Teil 1:  
Vorstellung des neuen  
Geheimhaltungsverfahrens**

Stefanie Setzer, Johannes Rohde, Volker Güttgemanns, Patrick Rothe

Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder führen für ausgewählte Statistiken die Cell-Key-Methode als neues Verfahren zur Ergebnisgeheimhaltung ein. Dieses Verfahren schützt die Befragten vor der Reidentifikation, indem es durch die Überlagerung der Fallzahlen eine Unsicherheit über die Anzahl der tatsächlich zum Ergebnis beitragenden Fälle schafft. Der Artikel stellt die Funktionsweise der Cell-Key-Methode vor und bietet dabei sowohl eine einfach zu verstehende Einführung in die Thematik als auch detaillierte methodische Informationen.



## 1 Einleitung

Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ) stellen Mikrodaten für die wissenschaftliche Nutzung bereit. Um den Datenschutz hierbei zu gewährleisten, gibt es zwei Möglichkeiten: entweder die Anonymisierung, bei der die Daten vor der Bereitstellung an die Wissenschaft so verändert werden, dass bei der Auswertung keine Geheimhaltungsrisiken entstehen können, oder die Geheimhaltung, bei welcher der Schutz der Daten durch eine Veränderung der Ergebnisse erzeugt wird. Die Wahl der Vorgehensweise hängt dabei vom gewählten Zugangsweg ab: Bei den sogenannten Off-Site-Zugangswegen, bei denen die Daten an die Nutzenden übermittelt werden, erfolgt eine Anonymisierung der Daten, was jedoch immer mit einem Informationsverlust einhergeht. Bei Nutzung der On-Site-Zugangswege, also eines Gastwissenschaftsarbeitsplatzes oder der kontrollierten Datenfernverarbeitung, verbleiben die Daten in den geschützten Räumen der amtlichen Statistik. Dort kann in der Regel das volle Informationspotenzial der Daten erhalten bleiben, die erzeugten Ergebnisse werden dafür aber einer Geheimhaltungsprüfung unterzogen.

Die Auswirkungen dieser Geheimhaltungsprüfung kennt jede Person, die schon einmal einen der On-Site-Zugangswege der Forschungsdatenzentren genutzt hat: Nach Bereitstellung der Ergebnisse springen häufig drei große X ins Auge. Dieses Sperrmuster verwenden die Forschungsdatenzentren üblicherweise, wenn die Veröffentlichung eines Ergebnisses ein Geheimhaltungsrisiko darstellt. Doch warum nehmen die Forschungsdatenzentren die Geheimhaltung überhaupt so ernst? Gäbe es Alternativen zu den drei großen X?

Kapitel 2 beantwortet zunächst die erste Frage, warum die Geheimhaltung den Forschungsdatenzentren so wichtig ist. Wie die Cell-Key-Methode als alternatives Verfahren zur Sperrung mit den drei großen X funktioniert, erläutert Kapitel 3. Danach erläutert Kapitel 4 die Methodik der Cell-Key-Methode formell. Ein kurzes Fazit mit dem Hinweis auf den zweiten Aufsatzteil beschließt den Beitrag.

### Stefanie Setzer

*ist Diplom-Soziologin und Referentin im Referat „Forschungsdatenzentrum, Methoden der Datenanalyse“ des Statistischen Bundesamtes. Schwerpunkt ihrer Arbeit ist die fachliche und methodische Weiterentwicklung des Arbeitsbereichs.*

### Dr. Johannes Rohde

*hat Wirtschaftswissenschaften an der Leibniz Universität Hannover studiert und dort 2015 seine Promotion im Bereich Statistik abgeschlossen. Bei IT.NRW leitet er den Service „Mathematisch-statistische Methoden und experimentelle Statistik“.*

### Volker Güttgemanns

*hat einen Master of Science in Wirtschaftswissenschaften und war von 2017 bis 2023 stellvertretende Leitung der Geschäftsstelle des Forschungsdatenzentrums der Statistischen Ämter der Länder.*

### Patrick Rothe

*hat Sozialwissenschaften an der Universität Mannheim studiert und ist seit 2011 im Bayerischen Landesamt für Statistik tätig. Seit 2018 leitet er dort das Sachgebiet „Grundsatzfragen der amtlichen Statistik, Digitalisierung, Forschungsdatenzentrum, Kompetenzzentrum Analyse“. Inhaltlich beschäftigt er sich schwerpunktmäßig unter anderem mit der statistischen Geheimhaltung.*

*Der vorliegende Beitrag ist in der Zeitschrift WISTA Wirtschaft und Statistik, Ausgabe 3/2024 erschienen und wird hier im Originalwortlaut mit Originalabbildungen abgedruckt. Das Bayerische Landesamt für Statistik dankt den Autoren und dem Statistischen Bundesamt (Destatis) für die freundliche Nachdruckgenehmigung.*





*Der Schutz der anvertrauten Daten hat daher für die amtliche Statistik – und damit auch für die Forschungsdatenzentren – stets die oberste Priorität.*

## 2 Geheimhaltung in den Forschungsdatenzentren

### 2.1 Warum nehmen die Forschungsdatenzentren Geheimhaltung so ernst?

Diese Frage lässt sich einfach beantworten: weil sie gesetzlich dazu verpflichtet sind. Die Pflicht zur Geheimhaltung ist in § 16 Bundesstatistikgesetz (BStatG) geregelt. Danach sind „Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik gemacht werden, [...] geheim zu halten“ (§ 16 Absatz 1 BStatG). Dieser Absatz regelt aber auch Ausnahmen, die es den statistischen Ämtern und der Wissenschaft ermöglichen, Daten und Ergebnisse unter bestimmten Voraussetzungen zu veröffentlichen. Eine Veröffentlichung ist beispielsweise möglich, wenn die Einzelangaben mit den Ergebnissen anderer Befragter zusammengefasst wurden oder wenn die Einzelangaben den Betroffenen nicht zuzuordnen sind. Diese beiden Ausnahmen ermöglichen die Bereitstellung von Daten und Ergebnissen und begründen gleichzeitig die Pflicht zur Geheimhaltung. Ergebnisveröffentlichungen sind erlaubt, solange aus den Ergebnissen keine Rückschlüsse auf Einzelne gezogen werden können.

Der Grund für diese gesetzliche Regelung und den dadurch festgelegten hohen Stellenwert der Geheimhaltung ist gut nachvollziehbar: Für die Erhebungen der amtlichen Statistik besteht oft Auskunftspflicht. Die Befragten – seien es Personen, Unternehmen, Betriebe oder Sonstige – können demnach häufig nicht selbst entscheiden, welche Informationen sie von sich preisgeben wollen. Um diesen Eingriff in die informationelle Selbstbestimmung auszugleichen, garantiert der Gesetzgeber den Befragten, dass ihnen ihre Angaben nicht zugeordnet werden können. Gleiches gilt für Erhebungen mit freiwilliger Teilnahme.

Das Vertrauen der Befragten in die Nicht-Zuordenbarkeit ihrer Angaben ist die Grundlage dafür, dass Fragen ohne Sorge vor Enthüllung wahrheitsgemäß beantwortet werden, und trägt somit maßgeblich zur hohen Qualität der Daten bei. Der Schutz der anvertrauten Daten hat daher für die amtliche Statistik – und damit auch für die Forschungsdatenzentren – stets die oberste Priorität.

## 2.2 Der bisherige Standard: die Zellspernung

Bisher stellen die statistischen Ämter die Geheimhaltung in der Regel mithilfe der Zellspernung sicher.<sup>1</sup> Bei dieser Form der Geheimhaltung werden alle Angaben, die ein Geheimhaltungsrisiko darstellen, durch ein Sperrmuster („XXX“) ersetzt. Dieses Verfahren hat sich in der amtlichen Statistik bewährt, weist jedoch einige gravierende Nachteile auf:

- **Informationsverlust:** Bei der Zellspernung werden nicht nur die kritischen Angaben selbst gesperrt (Primärspernung). Um eine Rückrechnung dieser Werte zu verhindern, müssen sie mit an sich unkritischen Angaben gegengespart werden (Sekundärspernung). Wenn gesperrte Angaben über andere Tabellen rückrechenbar sind, müssen auch hier Sperrungen umgesetzt werden (tabellenübergreifende Sperrung). So kann ein einzelner zu sperrender Wert schnell eine Vielzahl weiterer Sperrungen an sich unkritischer Werte nach sich ziehen.
- **Hoher Aufwand:** Da die Geheimhaltungsprüfung bei der Zellspernung in der Regel nicht vollständig automatisiert erfolgen kann, ist die Geheimhaltung sehr zeit- und ressourcenintensiv und Nutzende müssen teils lange auf ihre Ergebnisse warten.
- **Unzufriedenheit:** Durch dieses Verfahren können nicht alle interessierenden Werte veröffentlicht werden, teilweise müssen sogar ganze Tabellen gesperrt werden. Daher führt die Zellspernung häufig zu unzufriedenen Datennutzenden.

<sup>1</sup> Beim Zensus 2011 wurde außerdem das Verfahren SAFE – Sichere Anonymisierung Für Einzelangaben genutzt (Höhne, 2015).

<sup>2</sup> Die Cell-Key-Methode wurde ursprünglich vom australischen Statistikamt (Australian Bureau of Statistics) entwickelt.

## 2.3 Das neue Geheimhaltungsverfahren: die Cell-Key-Methode

Aufgrund der beschriebenen Nachteile der Zellspernung haben die statistischen Ämter für erste Statistiken die Einführung der Cell-Key-Methode (CKM) beschlossen.<sup>2</sup> Diese Entscheidung wirkt sich unmittelbar auf die Datenbereitstellung in den Forschungsdatenzentren aus, da die Sicherstellung der Geheimhaltung stets in Einklang mit den fachseitig festgelegten Geheimhaltungsregeln erfolgt. Bei der Cell-Key-Methode handelt es sich um ein datenveränderndes Verfahren für die Geheimhaltung von Fallzahltabellen. Die Vorteile des Verfahrens sind:

- Mit der Cell-Key-Methode gibt es keine Sperrungen.
- Die Ergebnisse weisen eine hohe Datenqualität auf und sind tabellenübergreifend konsistent.
- Der Aufwand für die Geheimhaltungsprüfung von Tabellen ist deutlich geringer.
- Aufdeckungsrisiken durch Fehler bei der tabellenübergreifenden Geheimhaltung können ausgeschlossen werden.

Diesen Vorteilen stehen aber auch Nachteile gegenüber:

- Tabellen sind nach der Anwendung der Cell-Key-Methode nicht mehr additiv.
- Gerade bei kleinen Fallzahlen kann die Veränderung der Werte relativ stark ausfallen.
- Da es sich bei der Cell-Key-Methode originär um ein Verfahren für die Geheimhaltung von Fallzahltabellen handelt, können zusätzliche Aufwände bei der Prüfung der Ergebnisse multivariater Analysemethoden entstehen.
- Die Cell-Key-Methode ist weniger eingängig als andere Geheimhaltungsverfahren und bedarf daher umfangreicher Erläuterung.

Das Verfahren der Cell-Key-Methode wird im Folgenden für die Grundform der Geheimhaltung von Fallzahltabellen vorgestellt.

### 3 Funktionsweise der Cell-Key-Methode

Bei der Cell-Key-Methode handelt es sich um ein post-tabulares datenveränderndes Geheimhaltungsverfahren. Das bedeutet, dass das Verfahren erst bei der Ergebniserstellung ansetzt und dass die Geheimhaltung durch eine Veränderung von Fallzahlen erfolgt. Die Schutzwirkung wird dadurch erzielt, dass Unsicherheit bezüglich der Originalfallzahl geschaffen wird, indem Tabellenwerte mit einem Fehlerterm überlagert werden. Das Verfahren stellt dabei sicher, dass die Überlagerung konsistent ist, dass also logisch identische Fallzahlen über alle Tabellen hinweg identisch bleiben. So nehmen Randsummen einer Kreuztabelle, beispielsweise „Bundesland x Alter“, immer die gleichen Werte an, die auch in den entsprechenden Fallzahltabellen der beiden Merkmale ausgegeben werden. Das gilt unabhängig davon, ob der Wert als Innen- oder als Randfeld einer Tabelle auftritt.

Die folgenden Abschnitte erläutern die Funktionsweise der Cell-Key-Methode Schritt für Schritt. Hierbei ist zu beachten, dass ein großer Vorteil der Cell-Key-Methode gerade darin besteht, dass die eigentliche Geheimhaltung von Fallzahltabellen weitgehend automatisiert erfolgt.

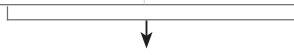
#### 3.1 Record Keys

Für die Anwendung der Cell-Key-Methode wird an den Ausgangsdatensatz zunächst ein zusätzliches Merkmal angespielt, das den sogenannten Record Key enthält. Dieser besteht aus einer Zufallszahl zwischen 0 und 1, die jeder Beobachtungseinheit fest zugeordnet wird.

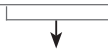
#### Übersicht 1

Anfügen der Record Keys an den Ausgangsdatensatz

ID	Alter	Einkommen	Record Key
1	jung	mittel	0,54
2	jung	hoch	0,68
3	alt	niedrig	0,14
4	alt	mittel	0,93
5	jung	mittel	0,51
6	alt	mittel	0,37
7	alt	niedrig	0,84
8	alt	hoch	0,19
9	alt	mittel	0,26
10	jung	hoch	0,43
11	alt	mittel	0,99
12	jung	mittel	0,74
13	jung	mittel	0,65
14	alt	niedrig	0,79
15	alt	mittel	0,25



Ausgangsdatensatz



Angespielter  
Record Key

Zur Veranschaulichung zeigt Übersicht 1 Daten für eine fiktive Gemeinde, in der das Alter und das Einkommen aller erwachsenen Bewohner klassiert erfasst ist. Aus Gründen der Übersichtlichkeit wird ein Record Key mit zwei Nachkommastellen vergeben, in echten Anwendungsfällen ist die Anzahl der Nachkommastellen in der Regel höher.

#### 3.2 Übergangsmatrix

In der Übergangsmatrix wird festgelegt, mit welchem Wert eine Fallzahl überlagert wird (Kleber/Gießing, 2018). Dafür wird für jede Originalfallzahl beschlossen, mit welcher Wahrscheinlichkeit diese zu einem bestimmten anderen Wert verändert wird. In Tabelle 1 bliebe die Originalfallzahl 10 zum Beispiel mit einer Wahrscheinlichkeit von 0,7 eine 10, würde mit einer Wahrscheinlichkeit von je 0,1 zu einer 9 oder 11 und mit einer Wahrscheinlichkeit von je 0,05 zu einer 8 oder 12. So lassen sich verschiedene Rahmenbedingungen festlegen, beispielsweise die maximale Abweichung, die Wahrscheinlichkeit für den Erhalt einer Originalfallzahl, der Ausschluss von 1 und 2 in der überlagerten Tabelle oder eine höhere Bleibewahrscheinlichkeit für höhere Fallzahlen.



Tabelle 1  
Fiktives Beispiel einer Übergangsmatrix

Original- häufigkeit	Zielhäufigkeit																	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0,7	0	0	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0,3	0	0	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0,5	0,35	0,15	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0,25	0,5	0,2	0,05	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0,05	0,2	0,5	0,2	0,05	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0,05	0,2	0,5	0,2	0,05	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0,05	0,2	0,5	0,2	0,05	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0,05	0,2	0,5	0,2	0,05	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0,05	0,2	0,5	0,2	0,05	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0,05	0,1	0,7	0,1	0,05

■ Nullen bleiben unverändert  
■ Keine 1 oder 2 in der überlagerten Tabelle  
■ Maximalabweichung +/-2  
■ Höhere Bleibewahrscheinlichkeit bei höheren Fallzahlen

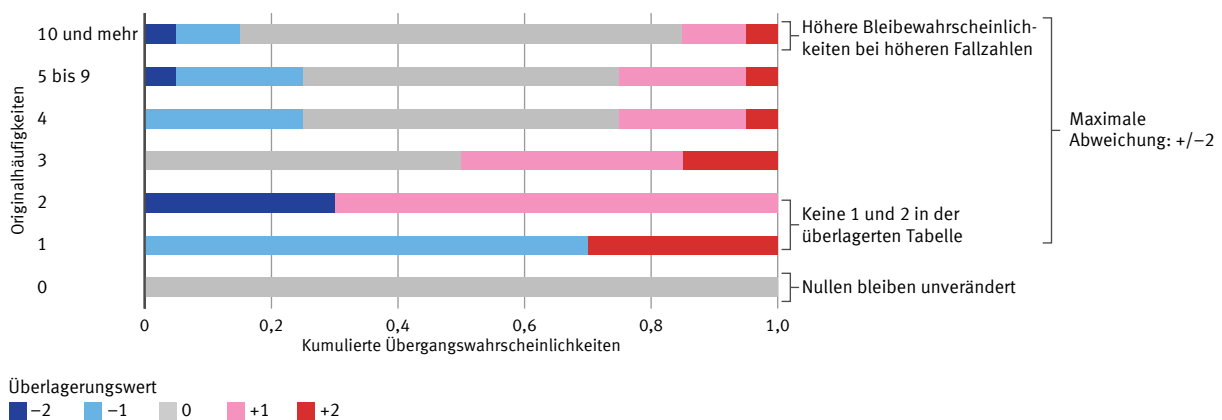
Dabei ist zu beachten, dass es sich hierbei lediglich um ein fiktives Beispiel einer Übergangsmatrix handelt, um deren mögliche Ausgestaltung vereinfacht darzustellen. Denkbar wäre beispielsweise auch die Festlegung, dass alle ausgewiesenen Nullen echte Nullen sind, dass kein Wert unverändert bleibt oder dass starke Überlagerungen wahrscheinlicher sind als geringe. Da die Übergangsmatrix somit steuert,

wie (un-)ähnlich sich die originale und die überlagerte Tabelle sind, stellt sie das Kernstück der Cell-Key-Methode dar, das mit viel Aufwand ausgestaltet wird. Diese tatsächlich verwendeten Übergangsmatrizen unterliegen ebenso wie die festgelegten Rahmenbedingungen der strengen Geheimhaltung und werden an die jeweiligen Bedarfe der Statistiken spezifisch angepasst.



Grafik 1

Fiktives Beispiel eines Überlagerungstableaus



### 3.3 Überlagerungstableau

Auf Basis der Übergangsmatrix wird ein Überlagerungstableau erstellt (Kleber/Gießing, 2018). Zur Veranschaulichung stellt Grafik 1 das Überlagerungstableau auf Basis der Übergangsmatrix aus Tabelle 1 als Raster dar. Hierfür werden die Wahrscheinlichkeiten der einzelnen Zeilen der Übergangsmatrix kumuliert. Daraus lässt sich im sogenannten Lookup-Schritt ablesen, welcher Überlagerungswert für die jeweilige Originalfallzahl aus der zugehörigen kumulierten Übergangswahrscheinlichkeit resultiert.

Zum besseren Verständnis stellt Übersicht 2 die Veränderungen der Originalfallzahl „4“ exemplarisch dar.

#### Übersicht 2

Kumulierte Übergangswahrscheinlichkeiten für die Originalfallzahl „4“

Veränderung zu	Entspricht Überlagerung mit	Wahrscheinlichkeit	Kumulierte Wahrscheinlichkeit
3	-1	0,25	0,25
4	0	0,5	0,75
5	1	0,2	0,95
6	2	0,05	1

### 3.4 Tabellenerstellung

Bei der Erstellung der überlagerten Tabellen kommen schließlich die Originalfallzahl, die kumulierte Übergangswahrscheinlichkeit für diese Originalfallzahl und die eingangs erzeugten Record Keys zusammen, um den Überlagerungswert zu bestimmen:

Im Zuge der Erstellung von Fallzahltabellen mit der Cell-Key-Methode werden nicht nur die Fallzahlen ermittelt. Für jede Tabellenzelle werden darüber hinaus die Record Keys aller Beobachtungseinheiten addiert, die zur entsprechenden Tabellenzelle beitragen. Relevant sind von der Summe der Record Keys allerdings nur die Nachkommastellen, der Wert vor dem Komma wird daher auf 0 gesetzt. So entsteht ein Wert zwischen 0 und kleiner 1, der sogenannte Cell Key.

Übersicht 3 veranschaulicht diesen Mechanismus für die Merkmalskombination „junge Befragte mit mittlerem Einkommen“ sowie für die Summe der Personen mit niedrigem Einkommen.

Die für alle Originalfallzahlen berechneten Cell Keys werden jetzt an das Übergangstableau zurückgespielt. Hier erfolgt der Abgleich, in welchem Raster die kumulierte Wahrscheinlichkeit dem ermittelten Cell Key entspricht. Aus dieser Spalte wird dann der Überlagerungswert für die jeweilige Originalfallzahl abgelesen. Die farbige Markierung des betreffenden Rasters verdeutlicht, mit welchem Wert die Originalfallzahl überlagert wird.

Im Beispiel ergibt sich für die vier jungen Personen mit mittlerem Einkommen aus dem ermittelten Cell Key von 0,44 ein Überlagerungswert von 0. Diese Fallzahl bleibt also unverändert. Für die drei Personen mit niedrigem Einkommen ergibt sich ein Cell Key von 0,77, was laut Überlagerungstableau einer Überlagerung von +1 entspricht. Diese Fallzahl wird also von 3 auf 4 verändert. Siehe Grafik 2

Wird das Verfahren auf alle Felder der Tabelle angewandt, verändert sich die originale Fallzahltableau anhand des in Übersicht 4 dargestellten Mechanismus.

Der überlagerten Tabelle sieht man zunächst nicht an, dass sie nicht die Originalfallzahlen enthält. Auf den zweiten Blick wird aber schnell deutlich, dass sich die Innenfelder in der Regel nicht zu den Randsummen summieren. Diese und andere Auswirkungen der Cell-Key-Methode stellt in dieser Ausgabe ein weiterer Beitrag vor (Rothe und andere, 2024).

## 4 Formelle Erläuterung der Cell-Key-Methode

Die Cell-Key-Methode eignet sich in ihrer Grundform für die Geheimhaltung von Fallzahltableaus. Mittlere steht auch eine Erweiterung der Cell-Key-Methode auf Wertetabellen zur Verfügung (Gießing/Tent, 2019), die in der deutschen amtlichen Statistik bislang allerdings lediglich für einige Wertmerkmale des Zensus 2022 angewendet wird. Im Folgenden beschränkt sich die formelle Erläuterung der Methodik der Cell-Key-Methode daher auf deren Grundform für Fallzahltableaus in Anlehnung an die Originalveröffentlichung von Fraser/Wooton (2005).

Die Grundidee der Cell-Key-Methode besteht darin, bei den Nutzenden Unsicherheit dahingehend zu erzeugen, ob eine veröffentlichte Fallzahl der Originalfallzahl entspricht oder diese leicht verändert wurde. Dazu erfolgt eine Perturbation aller Originalfallzahlen mittels eines Überlagerungswertes. Bezeichnet man mit  $d_i \in \mathbb{Z}$  den dem  $i$ -ten Tabellenfeld mit Originalfallzahl  $j_i \in \mathbb{N}_0$  zugeordneten Überlagerungswert, so ergibt sich die veränderte Fallzahl  $k_i$  gemäß

$$k_i = j_i + d_i.$$

### Übersicht 3

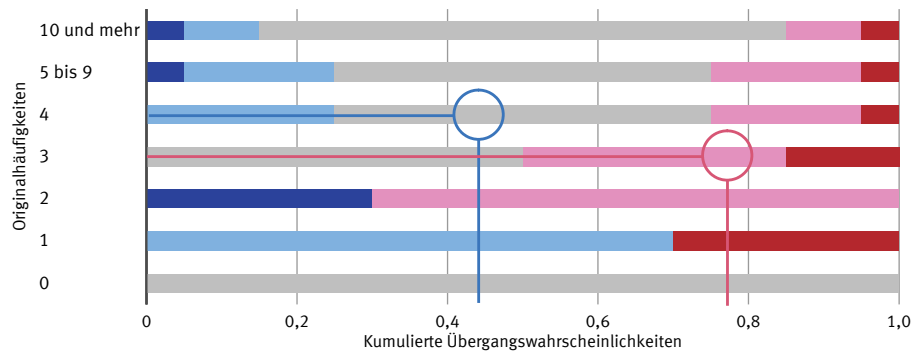
Erstellen der Cell Keys für zwei Beispiele durch die Addition der zugehörigen Record Keys (RK) und das Auf-null-Setzen der Zahl vor dem Komma

ID	Alter	Einkommen	Record Key
1	jung	mittel	0,54
2	jung	hoch	0,68
3	alt	niedrig	0,14
4	alt	mittel	0,93
5	jung	mittel	0,51
6	alt	mittel	0,37
7	alt	niedrig	0,84
8	alt	hoch	0,19
9	alt	mittel	0,26
10	jung	hoch	0,43
11	alt	mittel	0,99
12	jung	mittel	0,74
13	jung	mittel	0,65
14	alt	niedrig	0,79
15	alt	mittel	0,25

		Alter		
		jung	alt	Summe
Einkommen	niedrig	0	3	3 $\Sigma RK = 0,14 + 0,84 + 0,79 = 1,77$ → Cell Key = 0,77
	mittel	4 $\Sigma RK = 0,54 + 0,51 + 0,74 + 0,65 = 2,44$ → Cell Key = 0,44	5	9
	hoch	2	1	3
	Summe	6	9	15

## Grafik 2

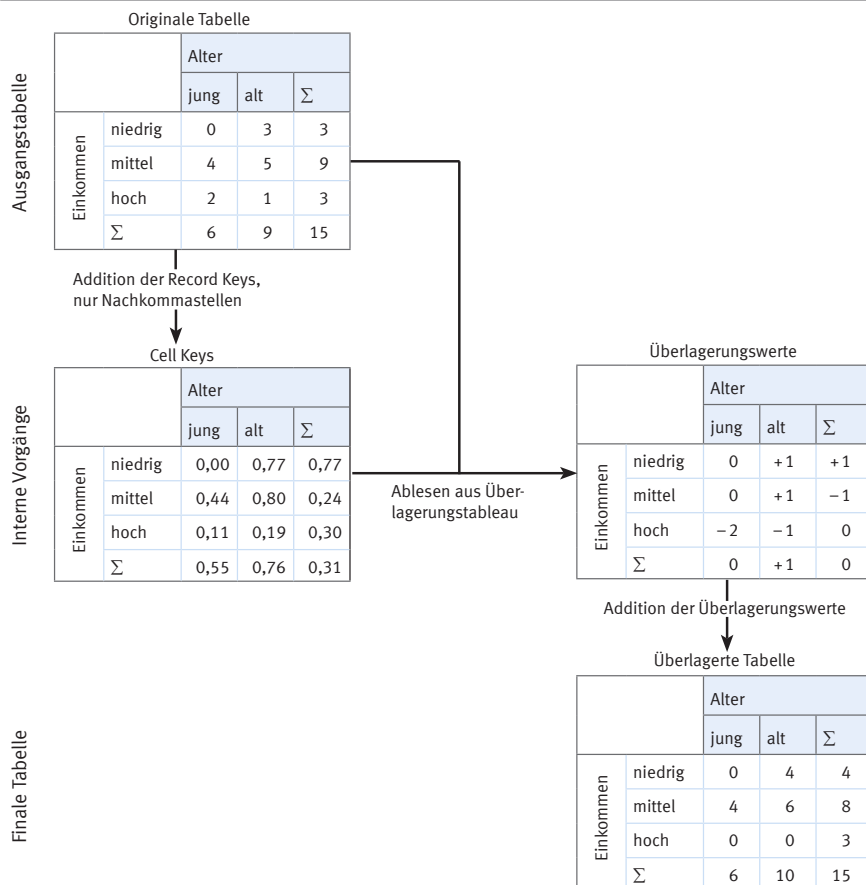
Ablesen ("Lookup") der Überlagerungswerte aus dem Überlagerungstableau für die 4 jungen Personen mit mittlerem Einkommen und einem Cell Key von 0,44 (blau) und den 3 Personen mit niedrigem Einkommen und einem Cell Key von 0,77 (rot)



Überlagerungswert  
 -2 -1 0 +1 +2

## Übersicht 4

Darstellung der Arbeitsschritte der Cell-Key-Methode





Wie bei allen datenverändernden Verfahren gehört zu den wesentlichen Eigenschaften der Cell-Key-Methode, dass alle Originalfallzahlen unabhängig ihrer Kritikalität beziehungsweise ihres Aufdeckungsrisikos verändert werden können.

#### 4.1 Anforderungen an die Überlagerungswerte

Für die Überlagerungswerte  $d_i$  gelten für alle  $i$  folgende Eigenschaften:

- Unverzerrtheit:**  $E[d_i] = 0$ , das heißt aus der Überlagerung der Tabellenfelder ergibt sich keine systematische Verzerrung des Gesamtergebnisses.
- Nicht-Negativität:**  $d_i \geq -j_i$ , das heißt für jedes Tabellenfeld wird sichergestellt, dass sich durch die Überlagerung keine negative Fallzahl ergibt ( $\forall i: k_i \geq 0$ ).
- Ganzzahligkeit:**  $d_i \in \mathbb{Z}$ , das heißt die Überlagerungswerte müssen ganzzahlig sein, damit auch für die veränderten Fallzahlen Ganzzahligkeit gewährleistet ist. Aus den Bedingungen (b) und (c) ergibt sich somit  $k_i \in \mathbb{N}_0$ .

#### 4.2 Parameter der Cell-Key-Methode

Die Anwendung der Cell-Key-Methode setzt die Festlegung von Parametern und Bedingungen voraus, durch welche das für die Zuweisung der Überlagerungswerte maßgebende stochastische Modell determiniert wird. Dabei wird zwischen zwingend festzulegenden Parametern sowie zusätzlichen optional zu setzenden Restriktionen unterschieden. Die konkrete Ausgestaltung der Parametrisierung erfolgt dabei stets durch die für die betreffende Statistik zuständige Fachseite.

Konkret ist die Festlegung zweier Parameter obligatorisch:

- **Maximale Abweichung**  $D \in \mathbb{N}$  zwischen originalen und veränderten Fallzahlen, sodass  $\forall i: |d_i| \leq D$ .
- **Varianz**  $V$  der Abweichungen von den Originalfallzahlen: Die Varianz entspricht

$$V := \text{Var}[d_i] = E[d_i^2] = \sum_{i=-D}^D (p_i \cdot d_i^2).$$

Durch die Festlegung von  $D$  ergibt sich als Wertebereich für  $V$  implizit  $(0; D^2]$ .

Die Festlegung der maximalen Abweichung zwischen originalen und veränderten Fallzahlen dient insbesondere der Steuerung des durch die Datenveränderung induzierten Informationsverlustes und damit der Qualität der geheim gehaltenen Ergebnisse. Mit der Varianz der Abweichungen von den Originalfallzahlen wird die Unsicherheit kalibriert, die bei den Nutzenden durch die Datenveränderung erzeugt werden soll. Eine geringe Varianz impliziert dabei, dass ein hoher Anteil der Originalfallzahlen nur geringfügig oder gar nicht (also  $d_i = 0$ ) verändert wird. Je größer  $V$  gewählt wird, umso größer ist die Wahrscheinlichkeit, dass (unter Berücksichtigung der festgelegten Maximalabweichung  $D$ ) hohe Abweichungen zwischen originalen und veränderten Fallzahlen auftreten.

### 4.3 Übergangsmatrix

Die maximale Abweichung von der Originalfallzahl sowie die Varianz der Überlagerungswerte bestimmen maßgeblich die Wahrscheinlichkeiten, mit der ein bestimmter Überlagerungswert einer Originalfallzahl zugeordnet wird. Im Folgenden bezeichnet  $p_{jk} \in [0;1]$  die Wahrscheinlichkeit, dass die Originalfallzahl  $j$  zur Fallzahl  $k$  verändert wird (beziehungsweise den Überlagerungswert  $d = |j-k|$  erhält). Da diese Wahrscheinlichkeit nur von der Höhe der Originalfallzahl abhängt, jedoch nicht von einem konkreten Tabellenfeld  $i$ , kann hier auf den Index  $i$  verzichtet werden. Die (bedingten) Übergangswahrscheinlichkeiten werden für alle Kombinationen möglicher Originalfallzahlen und veränderter Fallzahlen in der sogenannten Übergangsmatrix  $\mathbf{T}$  gesammelt:

$$\mathbf{T} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \dots & p_{0k} & \dots \\ p_{10} & p_{11} & p_{12} & \dots & p_{1k} & \dots \\ p_{20} & p_{21} & p_{22} & \dots & p_{2k} & \dots \\ \vdots & \vdots & \vdots & \ddots & & \\ p_{j0} & p_{j1} & p_{j2} & & p_{jk} & \\ \vdots & \vdots & \vdots & & & \ddots \end{pmatrix}$$

$\mathbf{T}$  ist eine quadratische Matrix, die folgende Eigenschaften aufweist:

- $\mathbf{T}$  enthält auf der Hauptdiagonalen die sogenannten Bleibewahrscheinlichkeiten  $p_{jj}$  mit  $j = k$ , das heißt die Wahrscheinlichkeiten, dass eine Originalfallzahl  $j$  unverändert bleibt.
- Zeilenweise enthält  $\mathbf{T}$  die bedingten Wahrscheinlichkeitsverteilungen für die Überlagerungen der Originalfallzahlen  $j$ , das heißt  $\forall j: \sum_k p_{jk} = 1$ .
- Durch die zwingende Festlegung einer maximalen Abweichung  $D$  von den jeweiligen Originalfallzahlen ist  $p_{jk} = 0$  für alle  $k \notin \{j-D, \dots, j+D\}$ .
- Für kleine Fallzahlen  $j < D$  kann das eigentlich durch  $D$  festgelegte Intervall für die möglichen Fallzahlen nach Datenveränderung  $[j-D, \dots, j+D]$  nicht vollständig ausgeschöpft werden, um der Bedingung der Nicht-Negativität der veränderten Fallzahlen zu genügen. Das tatsächliche Intervall für die möglichen Fallzahlen nach Veränderung der Originalfallzahl  $j$  lautet somit  $\Pi = [\max\{j-D; 0\}, \dots, j+D]$ . Negative Überlagerungswerte sind bei kleinen Originalfallzahlen (bei gleichzeitiger Gewährleistung der Unverzerrtheit der veränderten Fallzahlen) somit nur eingeschränkt möglich (Höhne / Höninger, 2018).



#### 4.4 Weitere optionale Restriktionen zur Kalibrierung der Übergangsmatrix

Neben den oben genannten zwingend festzulegenden Parametern  $D$  und  $V$  können optional weitere Nebenbedingungen formuliert werden, welche die Gestalt der Übergangsmatrix beeinflussen. Mögliche Restriktionen lauten:

- **Festlegung einer Bleibewahrscheinlichkeit:** Es kann festgelegt werden, dass ein bestimmter Anteil  $P_V$  aller Originalfallzahlen unverändert bleiben soll. In diesem Fall entsprechen  $P_V \cdot 100\%$  aller Fallzahlen in den geheim gehaltenen Tabellen ihrem jeweiligen Originalwert.
- **Original-Nullen sollen unverändert bleiben:**  $p_{00} = 1$  beziehungsweise  $\forall k > 0: p_{0k} = 0$ . In der Realität nicht existierende Ausprägungen von Merkmalskombinationen bleiben auch nach Anwendung der Cell-Key-Methode ausgeschlossen.
- **Ausschluss kleiner Fallzahlen nach Datenveränderung:** Die Ausgabe sehr kleiner veränderter Fallzahlen kann optional bis einschließlich eines Schwellenwerts  $m > 0$  ausgeschlossen werden. Für die entsprechenden veränderten Fallzahlen  $k = 1, \dots, m$  gilt dann  $\forall j: p_{jk} = 0$ . Die entsprechenden Spaltenvektoren für die ausgeschlossenen veränderten Fallzahlen in  $\mathbf{T}$  entsprechen dann dem Nullvektor. Häufig wird fachseitig die 1 als veränderte Fallzahl ausgeschlossen ( $\forall j: p_{j1} = 0$ ).

- **Gewährleistung einer symmetrischen Verteilung:**  $\forall d \in \{-D, \dots, D\}: p_{j(j-d)} = p_{j(j+d)}$ , das heißt Veränderungen einer Originalfallzahl um die Überlagerungswerte  $-d$  und  $d$  besitzen die identische Wahrscheinlichkeit. Hieraus ergibt sich eine symmetrische Verteilung der Übergangswahrscheinlichkeiten für jede Fallzahl  $j$  um die entsprechende Bleibewahrscheinlichkeit  $p_{jj}$ . Dabei ist zu beachten, dass die Symmetrieeigenschaft der Überlagerungsverteilung nur für solche Originalfallzahlen  $j$  gewährleistet werden kann, für welche  $j \geq D$  gilt. Werden zusätzlich kleine veränderte Fallzahlen bis einschließlich des Schwellenwerts  $m$  ausgeschlossen, erweitert sich diese Bedingung zu  $j > D + m$ .

#### 4.5 Berechnung der Übergangsmatrix

Auf Basis der obligatorisch festzulegenden Parameter sowie unter Berücksichtigung der weiteren optionalen Restriktionen an die Ausgestaltung der Übergangsmatrix ist für jede Originalfallzahl die konkrete (bedingte) Verteilung der Überlagerungswerte zu berechnen. Marley/Leaver (2011) schlagen hierzu die Maximierung der Entropie der (bedingten) Überlagerungsverteilungen vor. Die Entropie stellt ein Streuungsmaß einer Wahrscheinlichkeitsverteilung dar, dessen Maximierung in diesem Zusammenhang als Minimierung des Aufdeckungsrisikos interpretiert werden kann (Marley/Leaver, 2011). Bezeichnet  $\Pi$  die Menge der für die betreffende Originalfallzahl möglichen Fallzahlen nach Datenveränderung, so lautet das Maximierungsproblem zur Berechnung der Überlagerungsverteilung für Originalfallzahl  $j$

$$\max_{\mathbf{p}_{jk}} \left\{ - \sum_{k \in \Pi} (p_{jk} \cdot \log_2 p_{jk}) \right\}.$$



Einschließlich aller formulierten Nebenbedingungen ergibt sich somit ein nicht-lineares Gleichungssystem. Gießing (2016) stellt einen Ansatz zur Lösung des Optimierungsproblems mittels eines Lagrange-Ansatzes vor, welcher im R-Paket *ptable* (Enderle, 2023) zur Erstellung von CKM-Übergangsmatrizen angewendet wird. Weitere Ausführungen zur Maximierung der Entropie sind Enderle/Vollmar (2019) zu entnehmen.

Für alle Originalfallzahlen  $j > m + D$  ist die Überlagerungsverteilung strukturell identisch, das heißt  $\forall l \geq 0: p_{jk} = p_{(j+l)(k+l)}$ . Die resultierende Überlagerungsverteilung für  $j = m + D + 1$  gilt somit – jeweils um  $a \in \mathbb{N}$  Spalten in  $\mathbf{T}$  nach rechts verschoben – auch für alle größeren Originalfallzahlen  $j + a$ .

#### 4.6 Record Keys und Cell Keys

Nach der erfolgten Spezifizierung der Übergangsmatrix  $\mathbf{T}$  werden den Originalfallzahlen konkrete Überlagerungswerte zugeordnet. Dies erfolgt anhand des vorliegenden Datenmaterials.

Dazu wird zunächst jedem Merkmalsträger  $s = 1, \dots, S$  im Mikrodatsatz eine feste Zufallszahl  $r_s \sim U[0; 1]$  zugewiesen. Die Zufallszahlen werden aus einer stetigen Gleichverteilung gezogen und als **Record Keys** bezeichnet. Der einem Merkmalsträger zugewiesene Record Key bleibt für alle Auswertungen, die für die betroffene Statistik erstellt werden, mindestens für die laufende Berichtsperiode identisch.

Auf Ebene der Tabellenfelder wird anhand der Record Keys eine „Kennziffer“ für jedes einzelne Tabellenfeld gebildet, welche zur Zuweisung des Überlagerungswertes für das betreffende Tabellenfeld verwendet wird. Dieser sogenannte **Cell Key**  $c_i$  wird für Tabellenfeld  $i$  gemäß

$$c_i = \sum_{s \in I} r_s - \left\lfloor \sum_{s \in I} r_s \right\rfloor \sim U[0; 1]$$

berechnet, wobei die Menge  $I$  alle Merkmalsträger  $s$  enthält, die zu Tabellenfeld  $i$  beitragen. Zur Berechnung des Cell Keys für Tabellenfeld  $i$  wird die Summe

der Record Keys aller zu  $i$  beitragenden Merkmalsträger um deren nächst kleineren ganzzahligen Betrag reduziert (hinterer Term mit unterer Gauß-Klammer). Diese Rechenoperation ist notwendig, da gleichverteilte Zufallsvariablen ihre Verteilungseigenschaft bei Summierung (vorderer Term) verlieren und durch die Korrektur die Eigenschaft einer stetigen Gleichverteilung für die Cell Keys wiederhergestellt wird. Im Ergebnis weist nach diesem Schritt jedes zu überlagernde Tabellenfeld einen spezifischen Cell Key mit einem Wert  $c_i \in [0; 1]$  auf, der von den konkreten Merkmalsträgern beziehungsweise deren Record Keys abhängt, die zum Tabellenfeld beitragen. Durch die hier dargestellte Vorgehensweise wird tabellenübergreifende Konsistenz der veränderten Fallzahlen sichergestellt, da logisch identische Tabellenfelder (das heißt mit identischen beitragenden Merkmals-trägern) stets den gleichen Cell Key erhalten.

#### 4.7 Zuweisung der Überlagerungswerte

In einem letzten Schritt werden die generierten Cell Keys genutzt, um anhand der spezifizierten Übergangsmatrix zu entscheiden, welcher konkrete Überlagerungswert einem Tabellenfeld zugewiesen wird.

Die Übergangsmatrix wird dazu in ein sogenanntes Überlagerungstableau überführt, indem die Überlagerungsverteilung für jede Originalfallzahl  $j$  schrittweise aggregiert wird. Dazu sei die Verteilungsfunktion  $F_j$  gemäß

$$F_j(d) := \sum_{b \leq d; d \in A} p_{b|j}$$

definiert, wobei die Menge  $\Lambda_j = \{d_1, d_2, \dots, d_{n-1}, d_n\}$  alle für  $j$  infrage kommenden Überlagerungswerte enthält und  $p_{d|j}$  die Wahrscheinlichkeit einer Überlagerung

der Originalfallzahl  $j$  mit Überlagerungswert  $d$  bezeichnet. Die Höhe der einzelnen Übergangswahrscheinlichkeiten wird dabei über die Breite der Intervalle  $[0; F_j(d_1)]$ ,  $(F_j(d_1); F_j(d_2)]$ , ...,  $(F_j(d_{n-1}); F_j(d_n)]$  abgebildet, wobei  $F_j(d_n) = 1$  ist. Die Vereinigungsmenge aller Intervalle deckt das Intervall  $[0; 1]$  somit vollständig und überlappungsfrei ab.

Da die Cell Keys auf dem Intervall  $[0; 1)$  gleichverteilt sind, kann die Zuweisung des Überlagerungswerts auf Basis eines einfachen Abgleichs zwischen dem Cell Key eines Tabellenfeldes und der Verteilungsfunktion der Überlagerungswerte vorgenommen werden. Dabei wird der Überlagerungswert  $d_i$  zum Tabellenfeld  $i$  mit Cell Key  $c_i$  anhand des folgenden Mechanismus zugeordnet:

$$d_i = d_r | d_r \in A: c_i \in (F_j(d_{r-1}); F_j(d_r)]$$

Als Überlagerungswert für das Tabellenfeld  $i$  wird somit der Überlagerungswert  $d_r$  ausgewählt, falls der Cell Key des Tabellenfeldes in das Teilintervall fällt, welches durch die Verteilungsfunktionswerte von  $d_{r-1}$  und  $d_r$  aufgespannt wird (und dessen Breite der Wahrscheinlichkeit einer Überlagerung mit  $d_r$  entspricht). Dieser Zuweisungsmechanismus wird auf alle zu überlagernden Tabellenfelder angewendet. Durch die Gleichverteilungseigenschaft der Cell Keys ist gewährleistet, dass über alle Tabellen hinweg der Anteil an mit einem bestimmten Überlagerungswert überlagerten Tabellenfeldern der jeweiligen Übergangswahrscheinlichkeit  $p_{dij} \cdot 100\%$  entspricht.

## 5. Fazit

Mit der Cell-Key-Methode hält ein neues Geheimhaltungsverfahren Einzug in die amtliche Statistik und damit auch in die Forschungsdatenzentren. Die Cell-Key-Methode ist ein Verfahren, das auf einer post-tabularen stochastischen Überlagerung basiert. Die feste Zuordnung eines Record Keys zu jeder Beobachtungseinheit stellt sicher, dass Veränderungen der originalen Fallzahlen konsistent und über verschiedene Ergebnisläufe hinweg replizierbar erfolgen. Dies geht jedoch zulasten der Additivität der Ergebnistabellen. Welche Auswirkungen die Anwendung der Cell-Key-Methode auf Tabellenergebnisse und darauf basierende Kennzahlen darüber hinaus hat, beschreibt im Detail der Artikel „Die Cell-Key-Methode in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder – Teil 2: Auswirkungen des neuen Geheimhaltungsverfahrens“ in Ausgabe 9/2024 dieser Zeitschrift (Rothe und andere, 2024). ■

## Literatur

Enderle, Tobias/Vollmar, Meike. Geheimhaltung in der Hochschulstatistik. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2019, Seite 87 ff.

Verfügbar unter: [www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2019/06/geheimhaltung-hochschulstatistik-062019.pdf?\\_\\_blob=publicationFile](http://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2019/06/geheimhaltung-hochschulstatistik-062019.pdf?__blob=publicationFile)

Enderle, Tobias. ptable: Generation of Perturbation Tables for the Cell-Key Method. R package version 1.0.0. 2023. [Zugriff am 30. April 2024]. Verfügbar unter: <https://cran.r-project.org/web/packages/ptable/index.html>

Fraser, Bruce/Wooton, Janice. A proposed method for confidentialising tabular output to protect against differencing. Work session on statistical data confidentiality. Supporting paper. Genf 2005. [Zugriff am 30. April 2024]. Verfügbar unter: <https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2005/wp.35.e.pdf>

Giessing, Sarah/Tent, Reinhard. Concepts for generalising tools implementing the cell key method to the case of continuous variables. In: Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality. Den Haag 2019. [Zugriff am 30. April 2024]. Verfügbar unter: [https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2019/mtg1/SDC2019\\_S2\\_Germany\\_Giessing\\_Tent\\_AD.pdf](https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2019/mtg1/SDC2019_S2_Germany_Giessing_Tent_AD.pdf)

Giessing, Sarah. Computational Issues in the Design of Transition Probabilities and Disclosure Risk Estimation for Additive Noise. In: Domingo-Ferrer, Josep/Peji-Bach, Mirjana (Herausgeber). Privacy in Statistical Databases. LNCS (Lecture Notes in Computer Science). 2016. Ausgabe 9867, Seite 237 ff. DOI: 10.1007/978-3-319-45381-1\_18

Höhne, Jörg. Das Geheimhaltungsverfahren SAFE. In: Zeitschrift für amtliche Statistik Berlin Brandenburg. Ausgabe 2/2015, Seite 16 ff. [Zugriff am 30. April 2024]. Verfügbar unter: [www.statistischebibliothek.de/mir/receive/BBHeft\\_mods\\_00016121](http://www.statistischebibliothek.de/mir/receive/BBHeft_mods_00016121)

Höhne, Jörg/Höninger, Julia. Die Cell-Key-Methode – ein Geheimhaltungsverfahren. In: Zeitschrift für amtliche Statistik Berlin Brandenburg. Ausgabe 3+4/2018, Seite 14 ff. [Zugriff am 30. April 2024]. Verfügbar unter: [www.statistischebibliothek.de/mir/receive/BBHeft\\_mods\\_00036268](http://www.statistischebibliothek.de/mir/receive/BBHeft_mods_00036268)

Kleber, Birgit/Gießing, Sarah. Geheimhaltung beim Zensus 2021. In: Methoden – Verfahren – Entwicklungen. Nachrichten aus dem Statistischen Bundesamt. Ausgabe 2/2018, Seite 3 ff. [Zugriff am 30. April 2024]. Verfügbar unter: [www.destatis.de/DE/Methoden/Qualitaet/methoden-verfahren-entwicklung-02\\_2018.pdf?\\_\\_blob=publicationFile&v=1](http://www.destatis.de/DE/Methoden/Qualitaet/methoden-verfahren-entwicklung-02_2018.pdf?__blob=publicationFile&v=1)

Marley, Jennifer K./Leaver, Victoria L. A Method for Confidentialising User-Defined Tables: Statistical Properties and a Risk-Utility Analysis. In: Proceedings of 58th World Statistical Congress. 2011. [Zugriff am 30. April 2024]. Verfügbar unter: <https://2011.isiproceedings.org/papers/450007.pdf>

Rohde, Johannes/Seifert, Christiane/Gießing, Sarah/Setzer, Stefanie (unter Mitarbeit von Breitenfeld, Jörg/Brings, Stefan/Höhne, Jörg/Höninger, Julia/Rothe, Patrick/Schedding-Kleis, Ulrike). Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens. Version 1.1 vom 23.04.2021. Internes Dokument des Statistischen Verbunds (Statistische Ämter des Bundes und der Länder).

Rothe, Patrick/Güttgemanns, Volker/Rohde, Johannes/Setzer, Stefanie. Die Cell-Key-Methode in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. – Teil 2: Auswirkungen des neuen Geheimhaltungsverfahrens. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2024, Seite 45 ff. Verfügbar unter: [www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2024/03/cell-key-methode-teil2-032024.pdf?\\_\\_blob=publicationFile](http://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2024/03/cell-key-methode-teil2-032024.pdf?__blob=publicationFile)

## Rechtsgrundlagen

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.



# **TRENNUNGEN UND ZUSAMMENLEGUNGEN IN DER HAUSHALTEGENERIERUNG IM ZENSUS 2022**

**Mathematische Modellierung  
und algorithmische Verfahren**

Dr. Dirk Meierling, Dipl.-Math.



Ziel des Zensus 2022 ist neben der Ermittlung der aktuellen Bevölkerungszahlen die Bereitstellung von Daten zur Anzahl und Struktur von Haushalten und Familien und deren Wohnsituation. Das Teilprojekt Haushaltegenerierung ist dabei von zentraler Bedeutung: Durch Integration von Register- und Erhebungsdaten wird ein bundesweiter Gesamtdatenbestand erstellt, der verschiedenste inhaltliche und regionale Auswertungen bezüglich der Haushalts-, Familien- und Wohnverhältnisse in Deutschland ermöglicht. Nach der Bildung von ersten Haushaltszusammenhängen aus Melderegisterdaten ist das Ziel des zweiten Schritts der Haushaltegenerierung, eindeutige Zuordnungen zwischen Haushalten und Wohnungen zu ermitteln. Ist keine eindeutige Zuordnung zwischen Haushalten und Wohnungen möglich, so erfolgen Korrekturen in Form von Haushaltstrennungen und -zusammenlegungen sowie Wohnungszusammenlegungen. Aufgrund fachlicher Vorgaben ist die Trennung von Haushalten ein nicht triviales kombinatorisches Optimierungsproblem. Dieser Beitrag gibt einen Überblick über die mathematische Modellierung und skizziert die zur Lösung verwendeten algorithmischen Verfahren.

## Die Haushaltegenerierung im Zensus 2022

### Überblick

Die Haushaltegenerierung ist als eines der Teilprojekte ein wichtiger Baustein des Zensus 2022. Zentrale Aufgabe des Verfahrens ist die Gewinnung von Informationen über die Zahl und Struktur von Wohnhaushalten, einer statistischen Einheit, die im Zensus nicht direkt erfasst wird. Insbesondere werden dabei Relationen zwischen Haushaltsmitgliedern geknüpft, beispielsweise Nachkommenbeziehungen zwischen Eltern und Kindern oder Ehen.

Eine weitere Aufgabe der Haushaltegenerierung ist die statistische Korrektur von Über- und Untererfassungen – sogenannten Karteileichen und Fehlbeständen. Diese werden durch Vergleich der Bevölkerungszahlen laut Melderegister mit der Anzahl der bei der Haushalteinbefragung festgestellten Personen durch das Teilprojekt Hochrechnung ermittelt.

Nach Abschluss der Haushaltegenerierung liegt ein Ergebnisdatenbestand vor, der auf verschiedenen regionalen Ebenen unterschiedlichste Auswertungen hinsichtlich der Haushalts-, Familien- und Wohnverhältnisse in Deutschland ermöglicht. Die Ergebnisse der Haushaltegenerierung werden an interne und externe Auswertungsdatenbanken geliefert (siehe Abb. 1).

### Dr. Dirk Meierling



*Dirk Meierling promovierte im Fach Mathematik an der Rheinisch-Westfälischen Technischen Hochschule Aachen. Nach mehrjähriger Mitarbeit in Forschungsprojekten an der RWTH Aachen und der Universität Ulm*

*ist er seit Juni 2019 als Referent im Sachgebiet „Verfahrensentwicklung“ des Bayerischen Landesamts für Statistik tätig. Er war im Teilprojekt „Haushaltegenerierung“ des Zensus 2022 für mathematische Modellierung, Algorithmik und Tests verantwortlich. Seit Januar 2024 ist er mit der Anforderungsanalyse für Softwareprojekte betraut.*

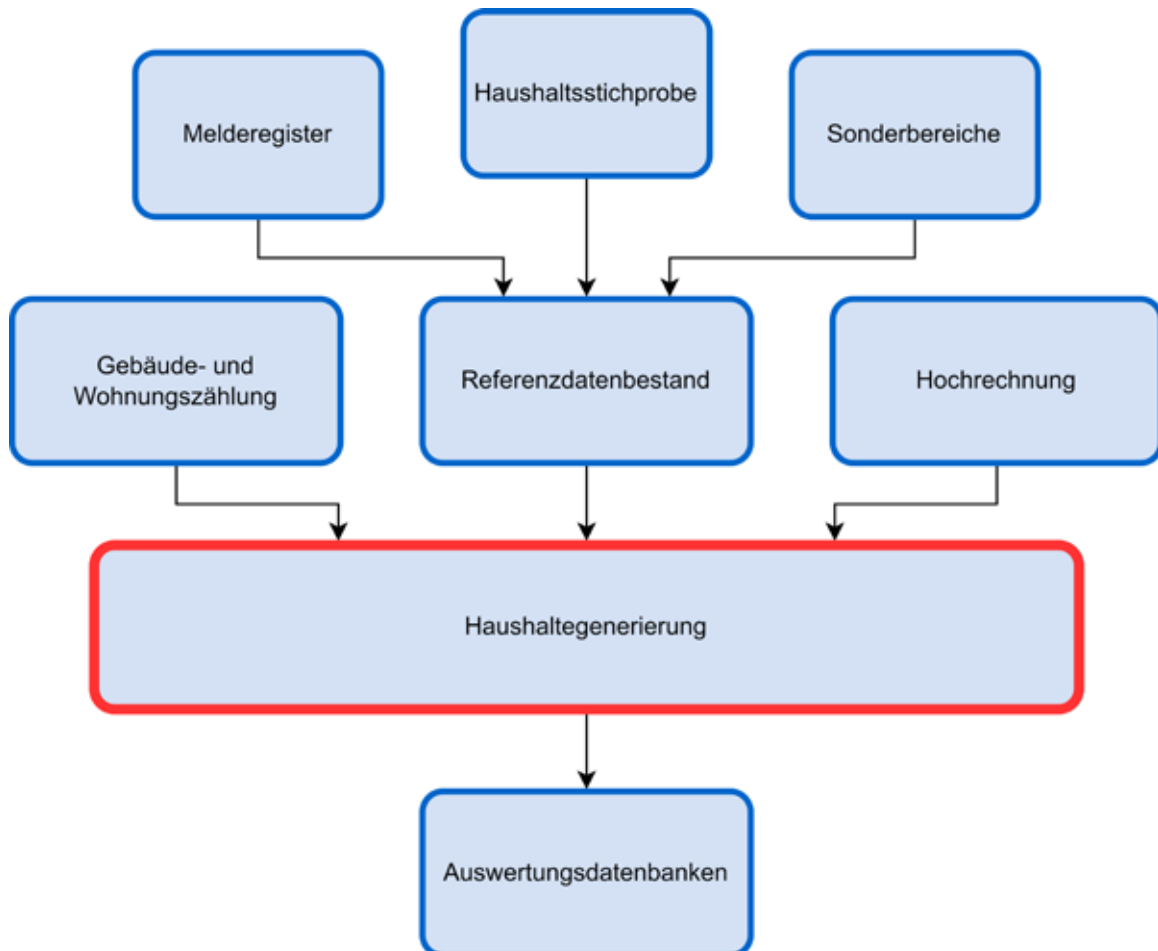
### Datenquellen

Die Haushaltegenerierung führt Daten aus verschiedenen Quellen des Zensus 2022 zusammen, verknüpft und bereinigt sie. Zu den Datenquellen zählen

- die Register der Einwohnermeldeämter der Kommunen,
- die Gebäude- und Wohnungszählung,
- die Haushaltsstichprobe,
- die Sonderbereichserhebung,
- die Ergebnisse der Hochrechnung.

Die der Haushaltegenerierung zugeliferten Daten beziehen sich demnach auf Gemeinden, Anschriften, Gebäude, Wohnungen und Personen. Sie stammen einerseits aus amtlichen Quellen, etwa den Registern der Einwohnermeldeämter, und andererseits aus primärstatistischen Erhebungen, wie zum Beispiel der Gebäude- und Wohnungszählung, der Haushaltsstichprobe und Erhebungen an Sonderbereichsanschriften.

Abb. 1

**Die Haushaltegenerierung im Kontext des Zensus 2022**

Die Melderegisterdaten der kommunalen Einwohnermeldeämter ermöglichen die Zuordnung von Personen zu Anschriften und Gemeinden und enthalten Informationen über gesetzliche Beziehungen zwischen Personen, wie Ehen, Lebenspartnerschaften<sup>1</sup> oder Eltern-Kind-Beziehungen.

Die Gebäude- und Wohnungszählung liefert Gebäude- und Wohnungsdaten auf Anschriftenebene. Diese erlauben insbesondere Zuordnungen von Personen zu Wohnungen über einen Abgleich von Nutzernamen.

Personenerhebungen an Anschriften der Haushaltsstichprobe und an Sonderbereichsanschriften liefern für einen Teil der Anschriften primärstatistische Daten über tatsächlich dort wohnende Personen und Haushaltszusammenhänge.

Auf Grundlage der Personenerhebungen werden zudem Hochrechnungen erstellt, die der Bereinigung der amtlichen Registerdaten um Über- und Unterfassungen dienen.

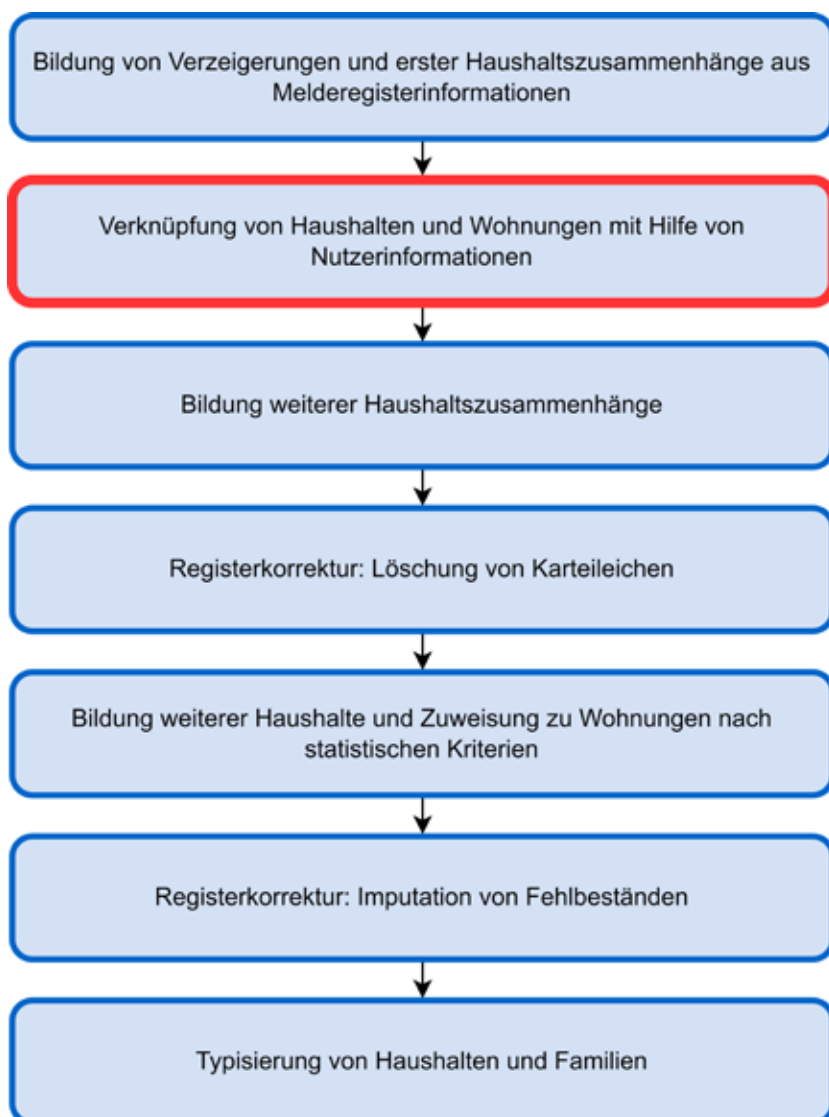
<sup>1</sup> Das Lebenspartnerschaftsgesetz (LPartG) ermöglichte von August 2001 bis September 2017 zwei Menschen gleichen Geschlechts in der Bundesrepublik Deutschland die Begründung einer Lebenspartnerschaft.

### Ablauf

Im Verfahren der Haushaltegenerierung werden die Eingangsdaten durch sequenzielle Ausführung von Arbeitsschritten, sogenannte Module, zusammengeführt (siehe Abb. 2). Dabei werden Haushaltszusammenhänge konstruiert und typisiert, Registerdaten bereinigt und Datensätze zur weiteren Auswertung erstellt. Mit Ausnahme der Registerkorrektur, die auf Gemeindeebene durchgeführt wird, operieren alle Module des Verfahrens auf einzelnen Anschriften.

Abb. 2

#### Ablauf der Haushaltegenerierung





Eine wesentliche Aufgabe der Haushaltegenerierung ist die Identifikation und Erzeugung von Relationen zwischen statistischen Einheiten (Personen, Wohnungen und Haushalten). Diese Relationen können hierarchisch sein, das heißt über- mit untergeordneten Einheiten verknüpfen, etwa Personen mit Haushalten oder Haushalte mit Wohnungen. Ebenso sind nicht-hierarchische Relationen möglich, die gleichgeordnete Einheiten miteinander verknüpfen, wie bei Relationen zwischen Personen, die Beziehungsstrukturen innerhalb von Haushalten abbilden, etwa Eltern-Kind-Beziehungen oder Ehen. Hierarchische Relationen sind dabei immer eindeutig – beispielsweise ist jede Person eindeutig einem Haushalt und jeder Haushalt eindeutig einer Wohnung zugeordnet.

Im ersten Schritt der Haushaltegenerierung werden Haushaltszusammenhänge mit Hilfe von Verzeigerungen<sup>2</sup> aus den Melderegistern gebildet. Eine Beziehung zwischen zwei Personen einer Anschrift wird genau dann gebildet, wenn im Melderegister Verweise zwischen den Personen vorhanden sind. Anschließend werden über identifizierte Beziehungen verbundene Personenmengen zu Haushalten zusammengefasst (siehe [19] für eine detaillierte Darstellung dieses Schritts).

Hauptthema dieses Beitrages ist der zweite Schritt der Haushaltegenerierung, die „Verknüpfung von Haushalten und Wohnungen mit Hilfe von Nutzerinformationen“, dessen Ziel die Bildung eindeutiger Zuordnungen zwischen Haushalten und Wohnungen ist.

Da in den Melderegistern keine Zuordnungen zwischen Personen und Wohnungen vermerkt sind, wird zunächst anschriftenweise ein maschineller Datenabgleich zwischen den Namen der gemeldeten Personen und den Wohnungsnutzerinformationen der Gebäude- und Wohnungszählung<sup>3</sup> durchgeführt. Dabei identifizierte Wohnungsnutzer und ihre Haushalte werden mit den entsprechenden Wohnungen verknüpft. Ist keine eindeutige Zuordnung zwischen Haushalten und Wohnungen möglich, so können Haushalte getrennt und zusammengelegt sowie Wohnungen zusammengelegt werden. Haushaltszusammenlegungen dienen etwa der Rekonstruktion der Haushaltszugehörigkeit von Partnern nichtehelicher Lebensgemeinschaften, über die keine Angaben in den Melderegistern vorliegen. Eine Haushaltstrennung kann beispielsweise dann nötig sein, wenn verwandte Personen unterschiedliche Wohnungen an derselben Anschrift bewohnen.

<sup>2</sup> Verzeigerungen sind in den Personendaten-sätzen der Melderegister enthaltene Verweise auf andere Personendatensätze. Sie enthalten Informationen darüber, ob eine Person in gesetzlichen Beziehungen zu anderen Personen steht. Verwiesen wird auf Partner einer Ehe oder Lebenspartnerschaft, auf minderjährige Kinder sowie auf gesetzliche Vertreter, das heißt Eltern oder Vormunde.

<sup>3</sup> Im Rahmen der Gebäude- und Wohnungszählung wurden für jede Wohnung die Namen von bis zu zwei Bewohnerinnen oder Bewohnern erfragt.

Im weiteren Verlauf der Haushaltegenerierung werden Haushaltszusammenhänge in mehreren Arbeitsschritten verfeinert. Zunächst werden die Daten aus den Melderegistern auf weitere Hinweise auf das Zusammenleben von Personen in einem gemeinsamen Haushalt analysiert, beispielsweise das Alter von Personen, Familien- oder Geburtsnamen sowie Zuzugsdaten. Anschließend werden Haushalte, die in den vorherigen Schritten noch keiner Wohnung zugeordnet werden konnten, anhand von statistischen Kriterien auf Wohnungen verteilt. Nach Abschluss der Haushaltegenerierung ist jede Person einem Haushalt und jeder Haushalt einer Wohnung zugeordnet.

Zudem werden in zwei Schritten Registerkorrekturen durchgeführt und abschließend Haushalte und Familien typisiert.

Verknüpfung von Haushalten mit Wohnungen mithilfe von Nutzerinformationen

Ausgangssituation

Nach der Bildung erster Haushaltszusammenhänge und des Abgleichs von Wohnungsnutzerinformationen ist die Menge der Personen jeder zu bearbeitenden Anschrift in Haushalte unterteilt. Zwischen Personen desselben Haushalts können Beziehungen existieren; haushaltsübergreifende Beziehungen sind nicht möglich. Darüber hinaus können Personen (auch verschiedener Haushalte) als Nutzer von Wohnungen identifiziert worden sein. Sogenannte Mehrfachnutzungen wurden in einem vorbereitenden Arbeitsschritt bereinigt, sodass jede Person Nutzer höchstens einer Wohnung ist. In Abb. 3 ist eine Beispielsanschrift graphisch dargestellt. Für die graphische Darstellung von Anschriften wird der folgende Symbolkatalog verwendet.



Tab. 1 Symbolkatalog bei der Anschriftendarstellung





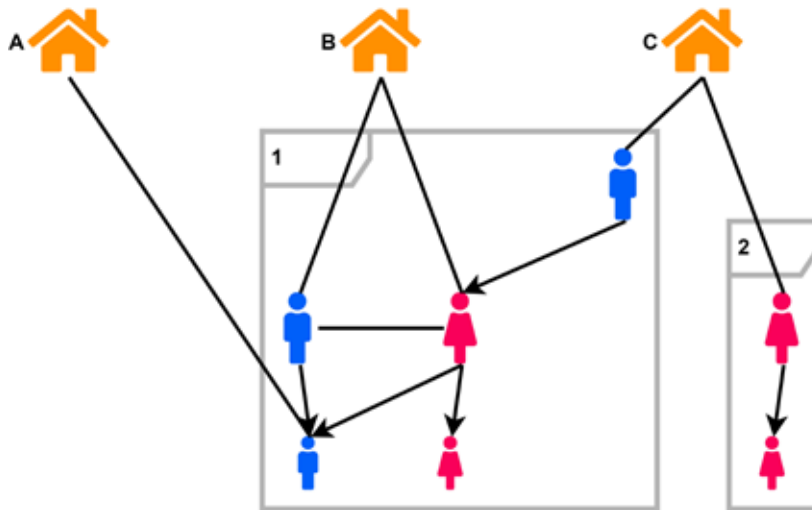
Symbol	Bedeutung
	Wohnung/zusammengelegte Wohnung
	Haushalt
	Personen (weiblich: volljährig/minderjährig; männlich: volljährig/minderjährig)
	Relationen

Abb. 3

**Situation an einer Beispielschrift vor Ausführung des zweiten Schritts der Haushaltegenerierung**



### Zielsetzung

Ziel dieses Schritts ist die eindeutige Zuordnung von Haushalten zu Wohnungen unter Berücksichtigung möglichst vieler bereits ermittelter Informationen (Beziehungen, Haushaltszusammenhänge und Nutzerinformationen). Haushalte und Wohnungen ohne identifizierte Nutzer werden dabei nicht bearbeitet. Jeder verarbeitete Haushalt soll nach Abschluss der Modulausführung genau einer Wohnung und jede verarbeitete Wohnung genau einem Haushalt zugeordnet sein.

### Vorgehen

Um die Zielvorgabe der eindeutigen Zuordnung zwischen Haushalten und Wohnungen zu erfüllen, müssen folgende Konstellationen aufgelöst werden:

- Haushalte mit Nutzern verschiedener Wohnungen sowie
- Wohnungen mit Nutzern in verschiedenen Haushalten.

Diese Konstellationen können, wie in der Beispielschrift in Abb. 3, auch in Kombination auftreten. (Für Haushalt 1 wurden vier Nutzer dreier Wohnungen identifiziert, für Wohnung 3 wurde sowohl ein Nutzer in Haushalt 1 als auch in Haushalt 2 ermittelt. Haushalt 1 ist somit an Konstellationen beider Typen beteiligt.)



Marko Novkov / stock.adobe.com

Abb. 4  
Beispielanschrift nach Haushaltstrennung

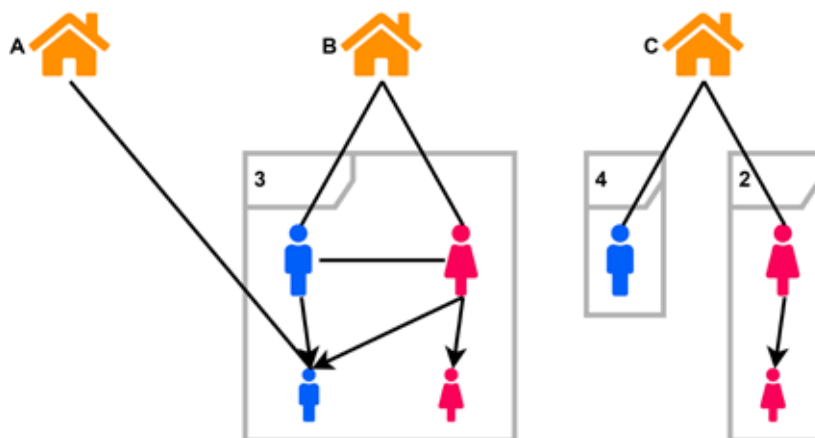
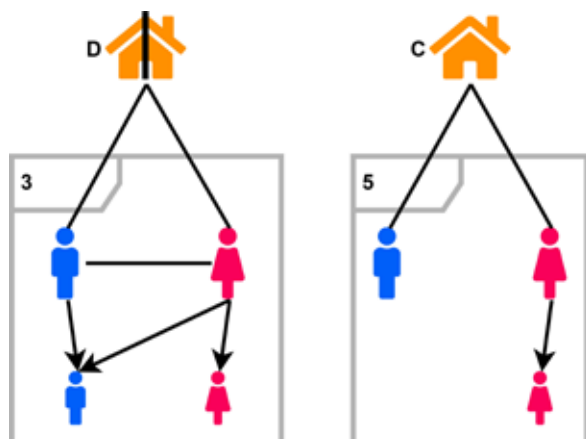


Abb. 5  
Beispielanschrift nach Wohnungs- und Haushaltszusammenlegung





Zur Auflösung problematischer Konstellationen werden zunächst sogenannte Haushaltstrennungen durchgeführt. Jeder Haushalt mit zwei (oder mehr) Personen, die als Nutzer verschiedener Wohnungen identifiziert wurden, wird so in Teilhaushalte getrennt, dass die Nutzer jedes (neu erzeugten) Teilhaushalts möglichst wenige Wohnungen nutzen (im Idealfall eine eindeutige). Da dabei verschiedene fachliche Vorgaben einzuhalten sind (siehe folgender Abschnitt „Fachliche Vorgaben bei Haushaltstrennungen“), ist eine ideale Trennung nicht immer möglich. Daher können die erwähnten Konstellationen auch nach Durchführung der Trennungen auftreten, allerdings nicht mehr in Kombination (siehe Abb. 4).

Verbleibende nicht-eindeutige Konstellationen werden durch Wohnungs- und Haushaltszusammenlegungen aufgelöst. Zunächst werden für jeden Haushalt, der Nutzer verschiedener Wohnungen enthält, alle solchen Wohnungen zu einer einzigen Wohnung zusammengelegt. Abschließend werden für jede Wohnung, für die Nutzer in verschiedenen Haushalten identifiziert wurden, alle entsprechenden Haushalte zusammengelegt (siehe Abb. 5).

Haushaltstrennungen sind der algorithmisch anspruchsvollste Teil dieses Moduls und werden nachfolgend genauer betrachtet.

### Fachliche Vorgaben bei Haushaltstrennungen

Da im Rahmen der Haushaltegenerierung haushaltsübergreifende Beziehungen nicht möglich sind, müssen bei der Trennung eines Haushalts auch Beziehungen „gelöscht“ werden<sup>4</sup>. Die Deaktivierung von Partnerbeziehungen (Ehen und eingetragenen Lebenspartnerschaften) ist dabei im Wesentlichen ausgeschlossen. Andere Beziehungstypen (Eltern-Kind-Beziehungen oder Vormundschaften) können gelöscht werden, sollen aber nach Alter und Geschlecht der beteiligten Personen sowie Typ der Beziehung priorisiert werden. Ziel der Priorisierungen ist, dass eher Kanten vom Typ Vormund-Mündel gelöscht werden als Kanten vom Typ Elternteil-Kind und eher Kanten von Vätern zu Kindern als Kanten von Müttern zu Kindern. Für Kanten vom Typ Elternteil-Kind werden eher Kanten von älteren Eltern beziehungsweise zu älteren Kindern gelöscht als solche von jüngeren Eltern beziehungsweise zu jüngeren Kindern.

Generell ausgeschlossen ist die Deaktivierung von Nutzerrelationen zwischen Personen und Wohnungen.

Eine weitere Einschränkung betrifft die Struktur der Teilhaushalte nach den Trennungen: Durch Haushaltstrennungen dürfen keine Haushalte entstehen, die ausschließlich minderjährige Personen<sup>5</sup> enthalten.

<sup>4</sup> Relationen werden in der Haushaltegenerierung nicht gelöscht, sondern deaktiviert und historisiert, sodass Zugriffe und Auswertungen auch nach Abschluss möglich sind.

<sup>5</sup> In der Haushaltegenerierung gelten Personen als minderjährig, wenn sie 14 Jahre alt oder jünger sind.

*In der Haushaltegenerierung wurden insgesamt 25 901 Haushalte getrennt, davon 25 749 in zwei und 152 Haushalte in drei Teilhaushalte.*



### **Modellierung als kombinatorisches Optimierungsproblem**

Die im vorherigen Abschnitt geschilderte Ausgangssituation an jeder Anschrift lässt sich in natürlicher Weise als ein mathematisches Objekt, einen sogenannten Graphen, modellieren (siehe Infokasten „Graphen, Knoten, Kanten, Färbungen und Gewichte“).

#### **Knoten und Kanten**

Die Knotenmenge des Graphen besteht aus den Personen und Wohnungen der Anschrift. Partnerschaften zwischen Personen und Nutzerrelationen werden durch ungerichtete Kanten, Eltern-Kind- und Vormund-Mündel-Beziehungen durch gerichtete Kanten repräsentiert.

### **Graphen, Knoten, Kanten, Färbungen und Gewichte**

In der diskreten Mathematik, genauer in der Graphentheorie, bezeichnet der Begriff Graph eine abstrakte Struktur, die aus einer Menge von Objekten besteht, die paarweise in Beziehung zueinander stehen können. Die Objekte nennt man Knoten, die Beziehungen Kanten.

Graphisch werden die Knoten typischerweise als Punkte oder Kreise, die Kanten als Geraden oder Linien dargestellt.

Eine Kante kann gerichtet oder ungerichtet sein. In einem klassischen Stammbaum wird beispielsweise jede Person durch einen Knoten repräsentiert. Verheiratete Personen sind durch eine ungerichtete Kante verbunden, Personen durch gerichtete Kanten mit ihren direkten Nachkommen.

Weitere Beispiele für Strukturen, die durch Graphen modelliert werden können, sind etwa Eisenbahnnetze (die Knoten sind Bahnhöfe, die Kanten Zugverbindungen), chemische Moleküle (die Knoten sind Atome, die Kanten chemische Bindungen) oder soziale Netzwerke (die Knoten sind Personen, die Kanten Bekanntschaftsbeziehungen).

Für graphentheoretische Modellierungen existieren eine Fülle von Erweiterungen zur Erhöhung des Informationsgrads. Ist jeder Kante eine Zahl (ihr Gewicht) zugeordnet, so spricht man von gewichteten Graphen oder Netzwerken. Die Gewichte repräsentieren beispielsweise Kosten, Entfernungen oder Kapazitäten. Ist jedem Knoten eine Beschriftung (seine Farbe) zugeordnet, so spricht man von (knoten-)gefärbten Graphen. Die Farben können beispielsweise der Klassifikation der Knoten dienen, etwa in Start- und Zielpunkte oder Erwachsene und Minderjährige.

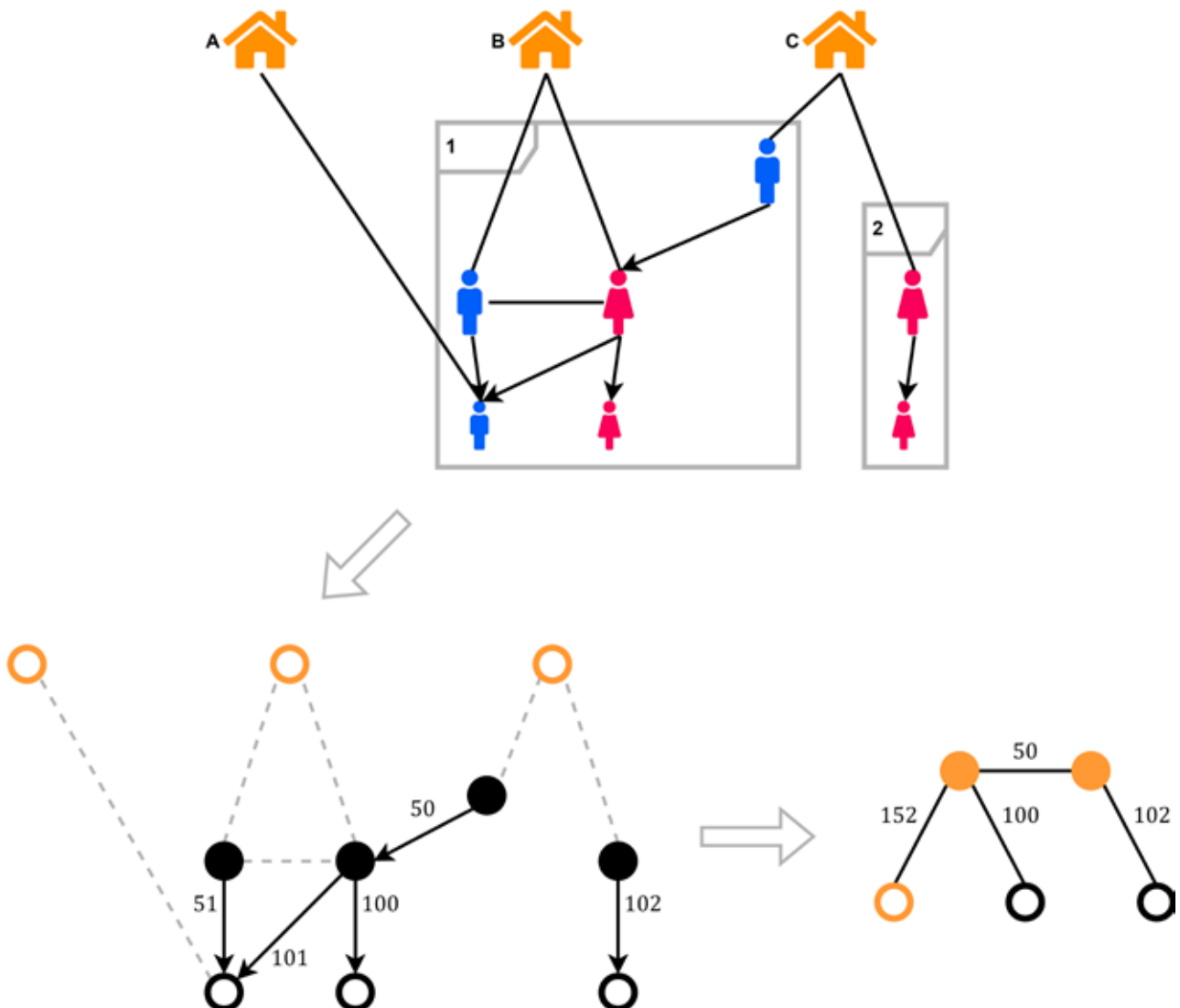
## Knotenfärbung

Zur Unterscheidung von Personen und Wohnungen werden die Knoten des Graphen mit zwei unterschiedlichen Farben gefärbt (in den folgenden Abbildungen schwarz für Personen und orange für Wohnungen). Zusätzlich werden die Knoten des Graphen zur Unterscheidung der minder- und volljährigen Personen der Anschrift markiert (in den folgenden Abbildungen sind Knoten, die volljährige Personen repräsentieren, ausgefüllt, alle anderen Knoten sind leer<sup>6</sup>). Die Knoten eines Anschriftengraphen unterscheiden sich also in zwei binären Merkmalen: minderjährig / volljährig und Wohnung / keine Wohnung. Sie sind also mit vier „Farben“ gefärbt, die die entsprechenden Klassen repräsentieren (Abb. 6 zeigt beispielsweise den Graphen der Beispielanschrift aus Abb. 3).

<sup>6</sup> Wohnungen sind dabei aus technischen Gründen „minderjährig“, da sie, wie minderjährige Personen, nicht isoliert werden dürfen.

Abb. 6

### Graph der Beispielanschrift und Vereinfachungen



## Gewichte

Die fachliche Vorgabe der Priorisierung von Beziehungstypen wird durch Kantengewichte modelliert. Die Gewichte wurden dabei nach fachlichen Kriterien so festgelegt, dass Kanten vom Typ Elternteil-Kind etwa zehnmal so schwer sind wie Kanten vom Typ Vormund-Mündel. Kanten von Müttern zu Kindern sind etwa doppelt so schwer wie Kanten von Vätern zu Kindern. Kanten vom Typ Elternteil-Kind von jüngeren Eltern beziehungsweise zu jüngeren Kindern sind etwas schwerer als von älteren Eltern beziehungsweise zu älteren Kindern.

Da die Orientierungen der Kanten und das Geschlecht der Personen nach Berechnung der Gewichte keine Rolle mehr spielen, können sie danach vernachlässigt werden.

## Kantenkontraktionen

Gemäß der fachlichen Problemstellung sind Kanten gewisser Typen untrennbar. Diese Vorgabe wird durch sogenannte Kantenkontraktionen im Graphen modelliert: Ist eine Kante untrennbar, so identifizieren wir ihre Endknoten (siehe Abb. 6, die Endknoten der grauen gestrichelten Kanten werden schrittweise identifiziert). Bei der Identifikation von Knoten werden die Färbungen (Eigenschaften) der Ursprungsknoten auf den neuen Knoten übertragen: Da gemäß fachlicher Vorgaben minderjährige Personen durch Trennungen nicht isoliert werden dürfen, ist ein neuer Knoten genau dann „minderjährig“, wenn alle Ursprungsknoten „minderjährig“ waren; ein neuer Knoten repräsentiert genau dann (mindestens) eine Wohnung, wenn mindestens einer der Ursprungsknoten eine Wohnung repräsentiert hat.

Bei der Kontraktion von Kanten können parallele Kanten (Mehrfachkanten) zwischen Knoten auftreten. Unter Addition ihrer Gewichte werden diese zu einer neuen Kante verschmolzen.

## Formulierung des kombinatorischen Optimierungsproblems

In diesem Abschnitt formalisieren wir das Problem der Haushaltstrennung als sogenanntes kombinatorisches Optimierungsproblem auf Graphen (siehe [18] für weiterführende Informationen).

### HAUSHALTSTRENNUNG

*Instanz:* Ein Graph  $G$  mit Knotenmenge  $V$  und Kantenmenge  $E$ , eine Teilmenge  $X \subseteq V$  nicht isolierbarer Knoten, eine Teilmenge  $T \subseteq V$  von  $k$  Terminalen, eine Funktion  $c: E \rightarrow \mathbb{N}$ , die jeder Kante eine natürliche Zahl als Kosten zuweist.

*Aufgabe:* Finde eine Kantenmenge  $S \subseteq E$  minimaler Kosten  $c(S) = \sum_{e \in S} c(e)$ , deren Löschung  $G$  in eine maximale Anzahl von Komponenten trennt, sodass jede Komponente von  $G - S$  mindestens einen Knoten aus  $T$  und einen Knoten aus  $V \setminus X$  enthält.

Im Falle der Haushaltstrennung bestehen die Menge  $T$  aus allen Knoten des vereinfachten Anschriftengraphen, die eine Wohnung repräsentieren, und die Menge  $X$  aus allen Knoten, die eine minderjährige Person repräsentieren. Durch die Formulierung der Aufgabenstellung wird sichergestellt, dass durch Haushaltstrennungen keine Haushalte entstehen, die ausschließlich minderjährige Personen enthalten.

<sup>7</sup> Genauer: Das zugehörige Entscheidungsproblem ist NP-schwer. Daher existieren vermutlich auch keine effizienten Algorithmen. Diese Vermutung ist ein Spezialfall eines der wichtigsten ungelösten Probleme der Komplexitätstheorie der theoretischen Informatik – die Frage, ob P gleich NP ist. Es ist eines der sieben Millenium Prize Problems des Clay Mathematics Institute, deren Lösung seit dem Jahr 2000 mit jeweils einer Million US-Dollar dotiert ist [16].



Den Spezialfall  $X = \emptyset$  dieses Problems findet man in der Literatur unter dem Namen MINIMUM MULTITERMINAL CUT [11], [12] beziehungsweise MINIMUM MULTIWAY CUT [9]. Selbst für MINIMUM MULTITERMINAL CUT ist kein effizienter Algorithmus bekannt, der das Problem exakt löst<sup>7</sup>. Verschiedene Spezialfälle sind aber in Polynomialzeit lösbar (siehe Infokasten „Algorithmen und Laufzeiten“). Im Folgenden geben wir einen kurzen Überblick über relevante Spezialfälle:

- Instanzen von MINIMUM MULTITERMINAL CUT mit genau zwei Terminalen sind Netzwerk-Fluss-Probleme, für die diverse Algorithmen mit polynomieller Laufzeit existieren [13].
- Für Instanzen mit mehr als zwei Terminalen existieren exakte Algorithmen für MINIMUM MULTITERMINAL CUT mit exponentieller Laufzeit [7], [20] und Approximationsalgorithmen mit polynomieller Laufzeit [5], [6], [8], [17].
- Für spezielle Klassen von Graphen, etwa Bäume und gewisse Verallgemeinerungen (siehe auch Infokasten „Bäume und verallgemeinerte Bäume“) ist MINIMUM MULTITERMINAL CUT lösbar in Polynomialzeit lösbar [3], [9], [14].

## Algorithmische Implementierung in der Haushaltegenerierung

### Anforderungen und Herausforderungen

Neben Vollständigkeit und Korrektheit der Algorithmen war die Laufzeit der im Rahmen des Teilprojekts erstellten Fachanwendung ein wichtiges Kriterium. Nach Lieferung aktualisierter Daten der zuliefernden Teilprojekte oder Veränderung von Parametern sollten möglichst schnell aktuelle Ergebnisse zur Verfügung stehen.

Organisatorische Herausforderungen ergaben sich vor allem durch gesetzliche Regelungen zur statistischen Geheimhaltung. Zum Schutz der Zensusdaten wurde eine eigene Betriebsumgebung mit stark eingeschränktem Zugriff aufgebaut. So war im Vorfeld der Entwicklung keine Analyse der Datenqualität möglich, insbesondere waren Größe und Struktur der Probleminstanzen a priori nicht genau bekannt. Darüber hinaus konnte die entwickelte Fachanwendung weder mit Echtdaten getestet, noch konnten Fehlerfälle detailliert analysiert werden.

### Algorithmen und Laufzeiten

Ein Algorithmus besteht aus einer Folge von Anweisungen, die in elementare Schritte – Variablenzuweisungen, Vergleiche, arithmetische Operationen, Fallunterscheidungen – zerlegt werden können. Für jede gültige Eingabe ist die Berechnung des Algorithmus eine eindeutige, endliche Folge von elementaren Schritten, die zu einer bestimmten Ausgabe führt.

Existieren für eine Funktion  $f$  Konstanten  $\alpha, \beta > 0$ , sodass ein Algorithmus  $A$  seine Berechnung für alle Eingaben der Größe  $x$  nach höchstens  $\alpha f(\beta x)$  elementaren Schritten beendet, dann ist die Laufzeit von  $A$  gleich  $O(f)$ .

Ist  $f$  ein Polynom in  $x$ , so spricht man von polynomieller Laufzeit; ist  $f$  eine Exponentialfunktion in  $x$ , so spricht man von exponentieller Laufzeit. Algorithmen mit polynomieller Laufzeit werden häufig auch als effizient bezeichnet.

Als inhaltliche Herausforderung erwies sich das Volumen der zu verarbeitenden Datensätze und erzeugten Relationen. Einen Überblick über die für das hier diskutierte Modul relevanten Kennzahlen des finalen Laufs der Haushaltegenerierung liefert Tabelle 2 (siehe [4] für weitere Informationen).

**Tab. 2 Anzahl der verarbeiteten aktiven Datensätze im Modul „Verknüpfung von Haushalten und Wohnungen mit Hilfe von Nutzerinformationen“ der Haushaltegenerierung**

Entität	Anzahl Eingangsdatensätze	Anzahl Ausgangsdatensätze
Anschrift .....	23 249 567	23 249 567
Person .....	89 569 565	89 569 565
Wohnung .....	44 111 023	43 988 280
Haushalt .....	52 555 031	49 162 791
Kante .....	176 607 295	207 721 765

In der Haushaltegenerierung wurden insgesamt 25 901 Haushalte getrennt, davon 25 749 in zwei und 152 Haushalte in drei Teilhaushalte. Bei den Trennungen wurden 47 130 Kanten deaktiviert (42 278 vom Typ Elternteil-Kind, 3 195 vom Typ Vormund-Mündel, 1 647 vom Typ Ehe und 10 vom Typ eingetragene Lebenspartnerschaft). Durch Haushaltstrennungen entstanden keine Teilhaushalte, die ausschließlich minderjährige Personen enthalten. Neben den Haushaltstrennungen wurden 6 223 360 Teilhaushalte zu 3 109 517 Haushalten und 234 714 Teilwohnungen zu 111 971 Wohnungen zusammengelegt.

Struktur der Anschriftengraphen

Wie im vorherigen Abschnitt geschildert, waren die tatsächlichen Größen der Problem instanzen (Anschriftengraphen) und insbesondere ihre Struktur zum Zeitpunkt der Programmierung unklar. Plausibel erschien jedoch die Annahme, dass ein Großteil der auftretenden Haushaltsstrukturen auf „natürliche Weise“ gewachsen ist – im Wesentlichen durch Schließung von Partnerschaften, Vergrößerung durch Nachkommen, Verkleinerungen durch Auszüge oder Sterbefälle. Dieses natürliche Wachstum resultiert graphentheoretisch in einer besonderen Struktur (siehe Infokasten „Bäume und verallgemeinerte Bäume“).

Abweichungen von der Struktur können in besonderen Haushaltskonstellationen auftreten. Als mögliche Auslöser für Abweichungen von dieser Struktur wurden in der Haushaltegenerierung in erster Linie Vormund-Mündel-Beziehungen identifiziert (eine Vormundschaft etwa zwischen Geschwistern folgt nicht den im Infokasten beschriebenen Konstruktionsvorschriften).

### Bäume und verallgemeinerte Bäume

Einige Klassen von Graphen können durch einfache Konstruktionsvorschriften beschrieben werden.

Ein Graph heißt Baum (englisch „tree“), wenn er aus einem einzelnen Knoten schrittweise durch folgende Regel konstruiert werden kann: füge einen neuen Knoten zum Graphen hinzu und verbinde ihn durch eine neue Kante mit einem beliebigen Knoten des Graphen. (Diese Operation tritt beispielsweise beim Zuzug eines Ehepartners auf.)

Analog wird ein Graph als 2-tree bezeichnet, wenn er aus zwei Knoten, die durch eine Kante verbunden sind, schrittweise durch folgende Regel konstruiert werden kann: Füge einen neuen Knoten zum Graphen hinzu und verbinde ihn durch zwei neue Kanten mit zwei beliebigen Knoten des Graphen, die durch eine Kante verbunden sind. Diese Operation tritt beispielsweise bei der Geburt eines gemeinsamen Kindes eines Ehepaares auf.

Graphen, die aus 2-trees durch Entfernen von Knoten beziehungsweise Kanten oder Kontraktion von Kanten konstruiert werden können, nennt man partial 2-trees. Das Entfernen eines Knotens tritt beispielsweise bei Auszügen oder Sterbefällen auf.

Die Erwartung, dass diese Konstellationen selten auftreten, wurden im Nachhinein in den Echtdaten bestätigt. Im finalen Lauf der Haushaltegenerierung wurden 217 476 Kanten vom Typ Vormund-Mündel gebildet und lediglich 3 195 dieser Kanten bei Haushaltstrennungen deaktiviert. Darüber hinaus waren alle Haushalte, für die Trennungen durchgeführt wurden, partial 2-trees. Bei der Verarbeitung von vorläufigen, unbereinigten Echtdaten in Evaluierungsläufen wurden jedoch vereinzelt Abweichungen von der Struktur festgestellt.

Eine genauere Analyse der in [9] beschriebenen Algorithmen, führte zu dem Ergebnis, dass sie auch im Fall  $X \neq \emptyset$  anwendbar sind, wenn gleichzeitig  $X \cap T = \emptyset$  gilt (das heißt bei Auftreten minderjähriger Personen, wenn gleichzeitig keine minderjährigen Wohnungsnutzer vorliegen). Das Auftreten minderjähriger Wohnungsnutzer wurde a priori als selten eingeschätzt. Auch diese Einschätzung wurde nach Lieferung der Echtdaten bestätigt: Von insgesamt 94 086 Personen in getrennten Haushalten wurden 62 886 Personen als volljährige und nur 10 599 als minderjährige Wohnungsnutzer identifiziert.

### Verfahren

Im Folgenden skizzieren wir ausschnittsweise die Struktur der implementierten Algorithmen. Die im vorherigen Abschnitt erläuterten Annahmen über die vorliegenden Haushaltsstrukturen und die Charakterisierung von (verallgemeinerten) Bäumen durch einfache Konstruktionsvorschriften legten einen Ansatz nahe, der dem Prinzip der dynamischen Programmierung folgt (siehe Infokasten „Dynamische Programmierung“).

### Dynamische Programmierung

Dynamische Programmierung ist eine Methode der mathematischen Optimierung, die in den 1950er-Jahren vom amerikanischen Mathematiker Richard Bellman entwickelt wurde [1], [2]. Sie kann verwendet werden, wenn ein Optimierungsproblem so in Teilprobleme aufgeteilt werden kann, dass sich eine optimale Lösung des Gesamtproblems aus den optimalen Lösungen der Teilprobleme zusammensetzt. In der dynamischen Programmierung werden zuerst die kleinsten Teilprobleme optimal gelöst, die Lösungen werden dann zu einer Lösung eines nächstgrößeren Teilproblems zusammengesetzt. Dieses Verfahren wird bis zur Lösung des ursprünglichen Problems fortgesetzt.

Wir beschränken unsere Betrachtung auf die kleinere beschriebene Graphenklasse der Bäume und die Problemstellung, eine optimale Trennung aller Terminalen zu finden (oder festzustellen, dass eine solche Trennung nicht existiert).

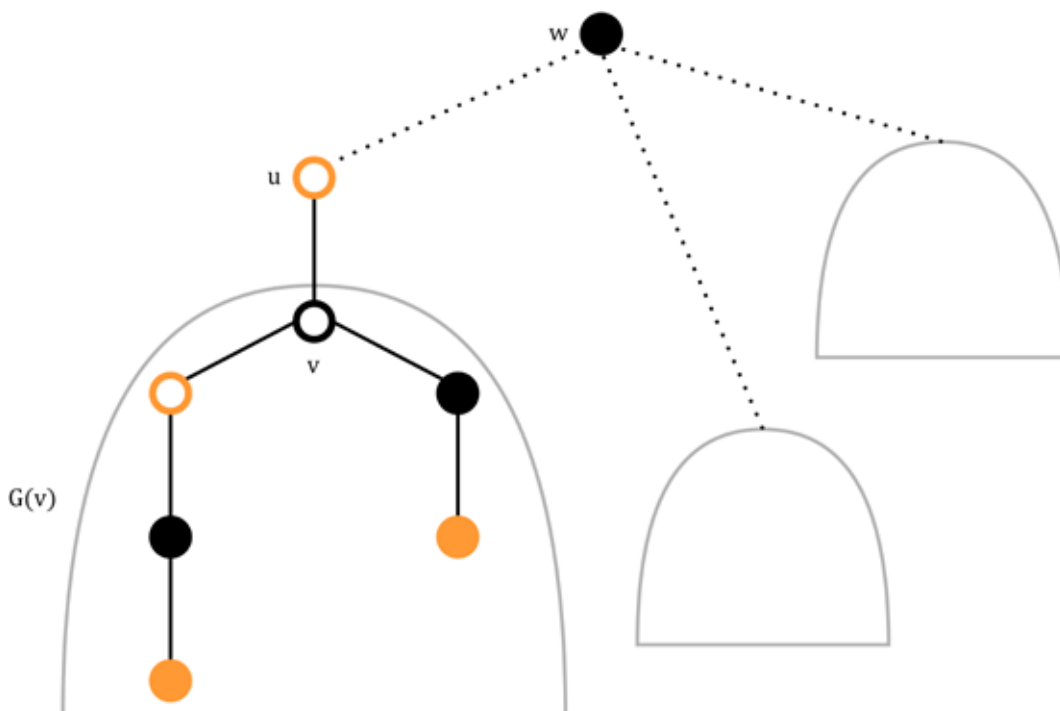
Sei also  $G = (V, E)$  ein Baum mit Kostenfunktion  $c: E \rightarrow \mathbb{N}$ ,  $T \subseteq V$  eine Menge von Terminalen und  $X \subseteq V$ . Wir markieren einen beliebigen Knoten  $w$  des Baumes  $G$  als sogenannte Wurzel. Den Ansatz der dynamischen Programmierung wenden wir folgendermaßen an: Auf kleinen Teilen des Baumes lösen wir zunächst vier Teilprobleme, deren optimale Lösungen anschließend zu Lösungen der vier Teilprobleme auf größeren Teilbäumen zusammengesetzt werden.<sup>8</sup>

Für einen beliebigen Knoten  $v$  bezeichnen wir mit  $G(v)$  den Teil des Baumes, der aus dem Knoten  $v$  selbst und allen Knoten besteht, die von  $w$  aus gesehen hinter  $v$  liegen (siehe Abb. 7). In jedem so definierten Teilbaum  $G(v)$  lösen wir nun folgende vier Trennungsprobleme.

<sup>8</sup> Das Lösen von vier Teilproblemen statt eines Problems erscheint zunächst unintuitiv und überkompliziert. Wir werden jedoch sehen, dass alle vier Teile zur Lösung des Ausgangsproblems nötig sind.

Abb. 7

#### Aufteilung des Optimierungsproblems in Teilprobleme





## TEILPROBLEME

*Instanz:* Ein Teilbaum  $G(v)$  von  $G$ .

*Aufgabe:* Finde eine Kantenmenge  $S$  in  $G(v)$  minimaler Kosten  $c(S) = \sum_{e \in S} c(e)$ , deren Löschung  $G(v)$  so in Komponenten trennt, dass jede Komponente von  $G(v) - S$  mit Ausnahme der Komponente  $C(v)$ , in der  $v$  liegt, genau einen Knoten aus  $T$  und mindestens einen Knoten aus  $V \setminus X$  enthält. Für  $C(v)$  gelte:

- P1:  $C(v)$  enthält genau einen Knoten aus  $T$  und mindestens einen aus  $V \setminus X$ ;
- P2:  $C(v)$  enthält genau einen Knoten aus  $T$  und keinen aus  $V \setminus X$ ;
- P3:  $C(v)$  enthält keinen Knoten aus  $T$  und mindestens einen aus  $V \setminus X$ ;
- P4:  $C(v)$  enthält weder einen Knoten aus  $T$  noch aus  $V \setminus X$ .

Man beachte, dass eine Lösung von TEILPROBLEM 1 in  $G(v)$  eine Lösung der eigentlichen Problemstellung ist.

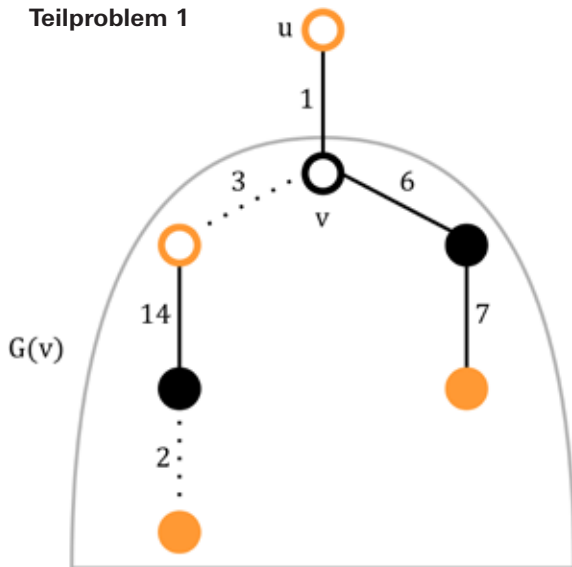
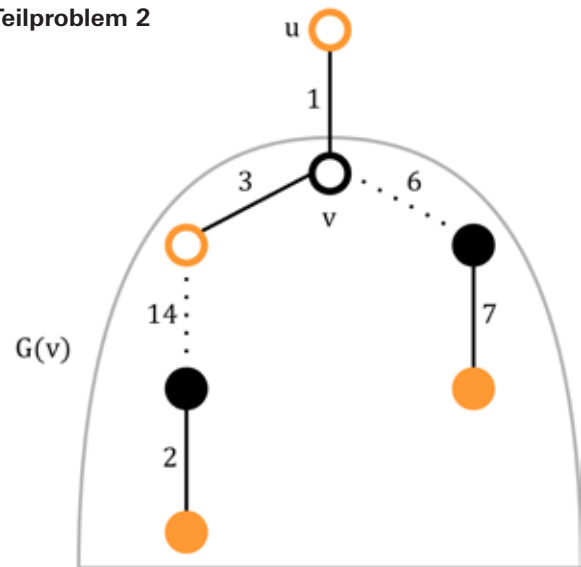
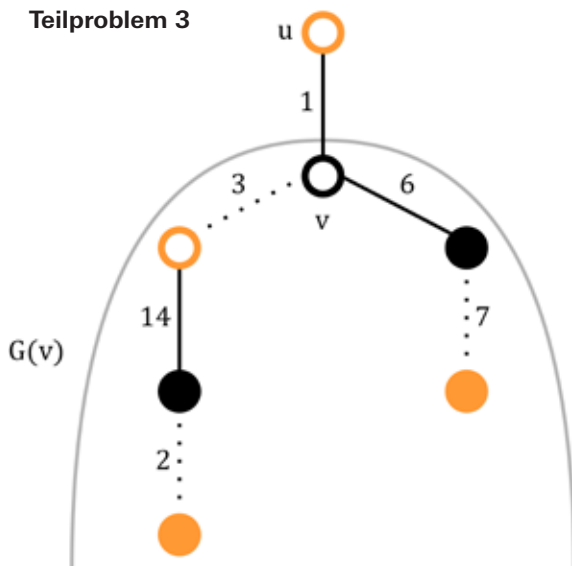
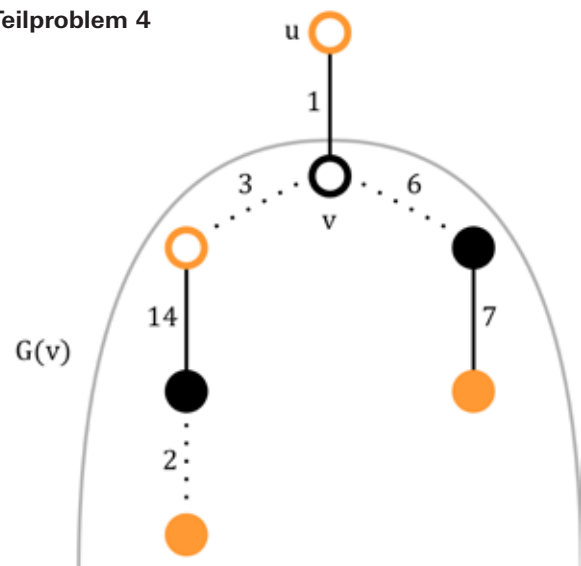
Im Folgenden skizzieren wir an einem Beispiel, wie Lösungen der Teilprobleme in  $G(v)$  zu Lösungen von größeren Teilproblemen zusammengesetzt werden können. Abb. 8 zeigt die optimalen Lösungen der vier Teilprobleme in einem Teilgraphen  $G(v)$ .

Da  $G$  ein Baum ist, ist der Knoten  $v$ , wie bereits in Abb. 7 zu sehen, mit genau einem Knoten  $u$  verbunden, der nicht in  $G(v)$  liegt. Die Lösungen von  $G(v)$  setzen wir nun zu Lösungen von  $G(u)$  zusammen (siehe Abb. 9).

Da  $u$  ein Terminal ist, existieren in diesem Beispiel in  $G(u)$  keine Lösungen von TEILPROBLEM 3 bzw. 4.

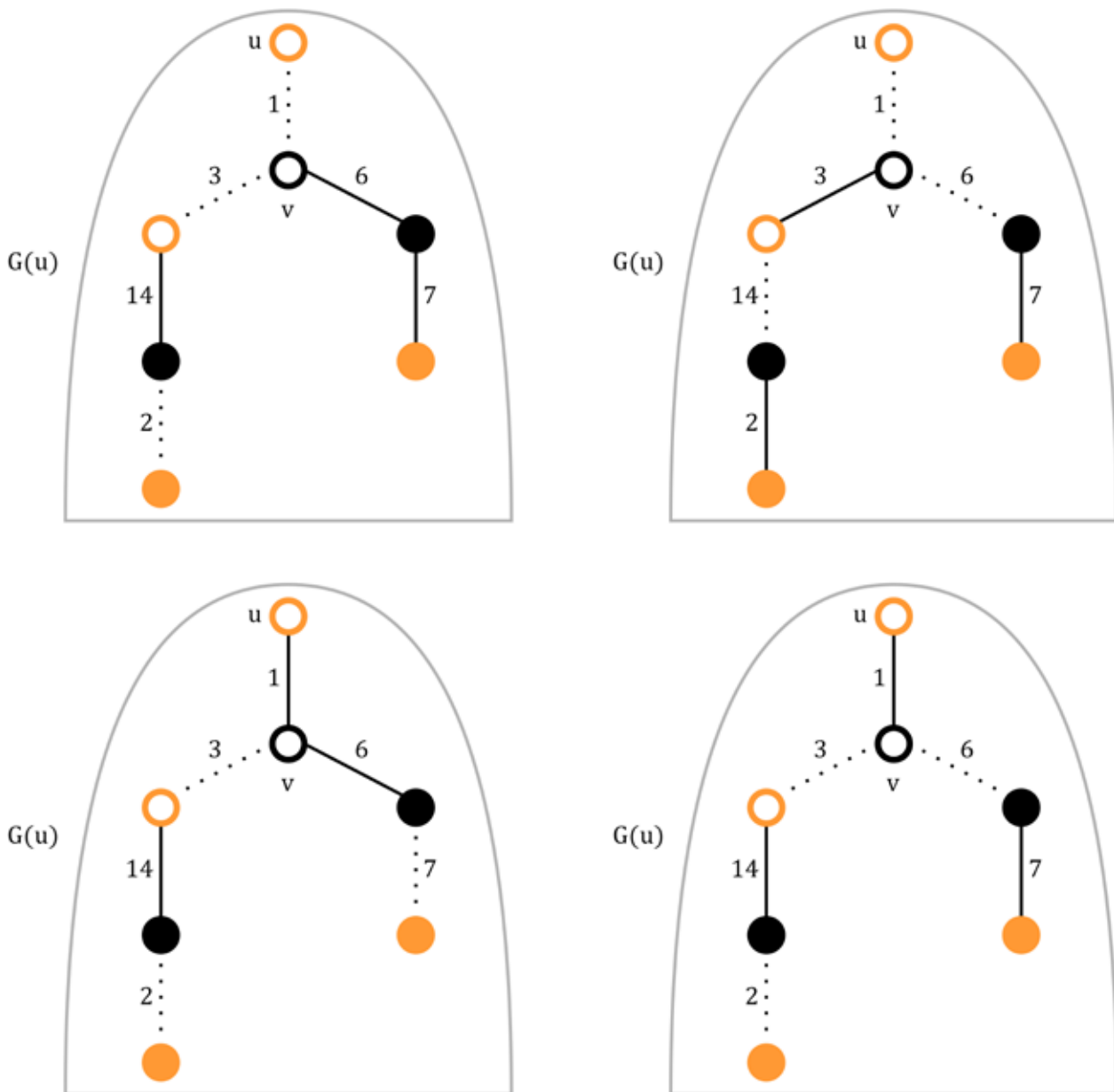
Die Lösungen von TEILPROBLEM 1 und 2 in  $G(v)$  können nur durch Löschung der Kante  $uv$  zu zulässigen Lösungen in  $G(u)$  erweitert werden (bei Erhalt von  $uv$  lägen zwei Terminale in einer gemeinsamen Komponente vor). Die Lösung von TEILPROBLEM 4 in  $G(v)$  kann nur durch Erhalt der Kante  $uv$  zu einer zulässigen Lösung in  $G(u)$  erweitert werden (bei Löschung von  $uv$  wäre  $v$  isoliert). Alle drei Erweiterungen sind Lösungen von TEILPROBLEM 2 in  $G(u)$ . Die erste Erweiterung hat das geringste Gewicht; die anderen Erweiterungen können daher verworfen werden.

Abb. 8

**Lösungen der Teilprobleme 1 bis 4****Teilproblem 1****Teilproblem 2****Teilproblem 3****Teilproblem 4**

Auch die Lösung von TEILPROBLEM 3 in  $G(v)$  kann nur durch Erhalt der Kante  $uv$  zu einer zulässigen Lösung in  $G(u)$  erweitert werden (bei Löschung von  $uv$  entstünde eine Komponente ohne Terminal, die  $u$  nicht enthält). Die Erweiterung ist dann eine Lösung von TEILPROBLEM 1 in  $G(u)$ .

Abb. 9

**Zusammensetzen von Lösungen von Teilproblemen**

Durch ähnliche Überlegungen können in allen Fällen bereits berechnete Lösungen zu Lösungen immer größerer Teilprobleme kombiniert werden bis schließlich Lösungen aller vier Teilprobleme in  $G(w)$  vorliegen.

Jeder Schritt des Algorithmus benötigt eine konstante Anzahl von Operationen (Additionen und Vergleiche von Gewichten). Da in jedem Schritt ein neuer Knoten verarbeitet wird, ist die Anzahl der Schritte und damit auch die Anzahl der Operationen linear in der Anzahl der Knoten des Baumes.

## Zusammenfassung der Ergebnisse und Fazit

Das Problem der Trennung von Haushalten im Rahmen des Schritts „Verknüpfung von Haushalten und Wohnungen mit Hilfe von Nutzerinformationen“ der Haushaltegenerierung des Zensus 2022 erwies sich als „gutartiges“ kombinatorisches Optimierungsproblem. Eine einfachere Variante des Problems war in der Literatur bereits bekannt, sodass auch vorhandene Lösungsansätze modifiziert und erweitert werden konnten. Als Verallgemeinerung dieses bekannten Problems ist die abstrahierte Form der Haushaltstrennung auch von allgemeinem wissenschaftlichem Interesse. In der Haushaltegenerierung standen allerdings nicht die theoretische Lösbarkeit, sondern insbesondere die Aspekte Laufzeit sowie Implementierungs- und Testaufwand im Vordergrund. Durch ausführliche vorausgehende mathematische Analyse, sorgfältige Modellierung und hohen Testaufwand konnten die fachlichen Vorgaben sowie Ziele bezüglich Laufzeit und Fehlerverhalten eingehalten werden.

## Literatur

- [1] Bellman, Richard: The theory of dynamic programming, in: Bull. Amer. Math. Soc., Bd. 60, 1954, doi:10.1090/S0002-9904-1954-09848-8, S. 503 – 515.
- [2] Bellman, Richard: Dynamic Programming, Princeton, NJ, USA: Princeton University Press, 1957.
- [3] Bertsimas, Dimitris / Teo, Chung-Piaw / Vohra, Rakesh: Analysis of LP relaxations for multiway and multicut problems, in: Networks: An International Journal, Bd. 34, Nr. 2, 1999, doi:10.1002/(SICI)1097-0037(199909)34:2<102::AID-NET3>3.0.CO;2-X, S. 102 – 114.
- [4] Bienk, Stefan: Softwareentwicklung im Zensusanteilprojekt, in: Bayern in Zahlen, Ausgabe 7/2024. [www.statistik.bayern.de/mam/produkte/biz/z1000g\\_202407.pdf](http://www.statistik.bayern.de/mam/produkte/biz/z1000g_202407.pdf), S. 48 – 55.
- [5] Buchbinder, Niv / Naor, Joseph / Schwartz, Roy: Simplex Partitioning via Exponential Clocks and the Multiway Cut Problem, in: Proceedings of the forty-fifth Annual ACM Symposium on Theory of Computing, 2013, doi:10.1145/2488608.2488675, S. 535 – 544.
- [6] Călinescu, Grigoriu / Karloff, Howard J. / Rabani, Yuval: An improved approximation algorithm for MULTIWAY CUT, in: Journal of Computer and System Sciences, Bd. 60, Nr. 3, 2000, doi:10.1006/jcss.1999.1687, S. 564 – 574.



- [7] Cao, Yixin / Chen, Jianer / Fan, Jia-Hao: An  $O^*(1.84^k)$  parameterized algorithm for the multiterminal cut problem, in: Information Processing Letters, Bd. 114, Nr. 4, 2014, doi:10.1016/j.ipl.2013.12.001, S. 167 – 173.
- [8] Cheung, Kevin K. H. / Cunningham, William H. / Tang, Lawrence: Optimal 3-terminal cuts and linear programming, in: Mathematical Programming, Bd. 106, Nr. 1, 2006, doi:10.1007/s10107-005-0668-2, S. 1 – 23.
- [9] Chopra, Sunil / Rao, M. R.: On the multiway cut polyhedron, in: Networks, Bd. 21, Nr. 1, 1991, doi:10.1002/net.3230210106, S. 51 – 89.
- [10] Clay Mathematics Institute: The Millenium Prize Problems: P vs NP: [www.claymath.org/millennium/p-vs-np/](http://www.claymath.org/millennium/p-vs-np/) (abgerufen am 16.04.2024).
- [11] Cunningham, William H.: The optimal multiterminal cut problem, in: DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Bd. 5, 1991, doi:10.1090/dimacs/005/07, S. 105 – 120.
- [12] Dahlhaus, Elias / Johnson, David S. / Papadimitriou, Christos H. / Seymour, Paul D. / Yannakakis, Mihalis: The Complexity of Multiterminal Cuts, in: SIAM J. Comput., Bd. 23, Nr. 4, 1994, doi:10.1137/S0097539792225297, S. 864 – 894.
- [13] Dinic, Efim A.: Algorithm for solution of a problem of maximum flow in networks with power estimation, in: Soviet Math. Dokl., Bd. 11, Nr. 5, 1970, S. 1277 – 1280.
- [14] Erdős, Peter L. / Székely, László A.: On weighted multiway cuts in trees, in: Mathematical programming, Bd. 65, Nr. 1, 1994, doi:10.1007/BF01581691, S. 93 – 105.
- [15] Fujishige, Satoru: A maximum flow algorithm using MA ordering, in: Operations Research Letters, Bd. 31, Nr. 3, 2003, doi:10.1016/S0167-6377(02)00237-7, S. 176 – 178.
- [16] Goldberg, Andrew V. / Tarjan, Robert E.: A new approach to the maximum-flow problem, in: Journal of the ACM, Bd. 35, Nr. 4, 1988, doi:10.1145/48014.61051, S. 921 – 940.
- [17] Karger, David R. / Klein, Philip N. / Stein, Clifford / Thorup, Mikkel / Young, Neal E.: Rounding algorithms for a geometric embedding of minimum multiway cut, in: Mathematics of Operations Research, Bd. 29, Nr. 3, 2004, doi:10.1287/moor.1030.0086, S. 436 – 461.
- [18] Korte, Bernhard / Vygen, Jens: Combinatorial Optimization: Theory and Algorithms, 6. Auflage, Berlin, Heidelberg, Deutschland: Springer, 2018, doi:10.1007/978-3-662-56039-6.
- [19] Seifert, Jakob: Haushaltegenerierung: Melderegister und Graphentheorie, erscheint in: Bayern in Zahlen, Ausgabe 12/2024.
- [20] Xiao, Mingyu: Simple and improved parameterized algorithms for multiterminal cuts, in: Theory of Computing Systems, Bd. 46, Nr. 4, 2010, doi:10.1007/s00224-009-9215-5, S. 723 – 736.

# NACHGEFRAGT

## BEI



### DR. DIRK MEIERLING

*Referent im Sachgebiet  
„Verfahrensentwicklung“*

*Dirk Meierling promovierte im Fach Mathematik an der Rheinisch-Westfälischen Technischen Hochschule Aachen. Nach mehrjähriger Mitarbeit in Forschungsprojekten an der RWTH Aachen und der Universität Ulm ist er seit Juni 2019 als Referent im Sachgebiet „Verfahrensentwicklung“ des Bayerischen Landesamts für Statistik tätig. Er war im Teilprojekt „Haushaltegenerierung“ des Zensus 2022 für mathematische Modellierung, Algorithmik und Tests verantwortlich. Seit Januar 2024 ist er mit der Anforderungsanalyse für Softwareprojekte betraut.*

**Die Haushaltegenerierung ist ein Teilprojekt des Zensus 2022. Können Sie in kurzen Worten beschreiben, was die Haushaltegenerierung ist und wie sie abläuft? Wie viele Personen waren im Landesamt beteiligt?**

**DR. MEIERLING:** In der Haushaltegenerierung werden Daten aus verschiedenen Quellen des Zensus zusammengeführt. Ziel ist die Ermittlung von Haushaltszusammenhängen. Durch Auswertung der Merkmale von Gebäuden, Wohnungen und einzelnen Personen werden in mehreren Schritten Wohnhaushalte gebildet und Wohnungen zugeordnet. Am Ende des Verfahrens ist jede Person genau einem Haushalt und jeder Haushalt genau einer Wohnung zugeordnet.

Zudem wird in der Haushaltegenerierung geprüft, ob die Bevölkerungszahl laut Melderegister mit der Anzahl der bei der Haushaltsstichprobe ermittelten Personen übereinstimmt. Über- und Untererfassungen – sogenannte Karteileichen und Fehlbestände – werden für die Bevölkerungszahlen des Zensus statistisch korrigiert.

Auf Seite der Anwendungsentwicklung bestand das Team aus bis zu 15 Personen mit entsprechenden Spezialisierungen, einschließlich externem Personal; im Team der Fachseite hier im Landesamt waren bis zu fünf Personen tätig.

”  
Die Verwendung von Verfahren  
des maschinellen Lernens, einem  
Teilgebiet der künstlichen Intelligenz,  
war eine wichtige Neuerung im  
Vergleich zum Zensus 2011.

**Welche Bedeutung haben die Ergebnisse der Haushaltegenerierung und für wen sind sie eine wichtige Informationsgrundlage?**

**DR. MEIERLING:** Die Haushaltegenerierung stellt Informationen über Anzahl, Größe und Struktur von Haushalten und deren Wohnsituation bereit. Nach ihrer Durchführung liegt ein Datenbestand für das gesamte Bundesgebiet vor, der auch auf regionalen Ebenen, wie beispielsweise in bayerischen Gemeinden, hinsichtlich Haushalts-, Familien- und Wohnverhältnissen ausgewertet werden kann. Für die Politik sind diese Daten eine wichtige Grundlage für Planungsentscheidungen.

**Welche Aufgaben hatten Sie in der Haushaltegenerierung?**

**DR. MEIERLING:** Gemeinsam mit meinem ehemaligen Kollegen Jakob Seifert war ich vor allem für drei Bereiche verantwortlich. Erstens für die vollständige und widerspruchsfreie Spezifikation der fachlichen Anforderungen. Dies geschah in enger Zusammenarbeit mit den Kollegen des Fachbereichs hier im Landesamt. Zweitens für eine wissenschaftlich fundierte Analyse und mathematische Modellierung der Problemstellungen, die sich aus den fachlichen Anforderungen ergaben. Dies betraf vor allem den Bereich der „Computational Complexity“, der den Ressourcenverbrauch der Anwendung hinsichtlich Laufzeit und Speicherbedarf umfasst. Schließlich war ich verantwortlich für den Entwurf und die Implementierung von Lösungsalgorithmen für die modellierten Probleme unter Berücksichtigung der Ergebnisse der vorangegangenen Analyse.

### Worin lagen dabei die größten Herausforderungen?

**DR. MEIERLING:** Zum einen im Bereich der mathematischen Analyse und Modellierung, da Ergebnisse und Methoden aus verschiedensten Bereichen der Mathematik und Informatik zur Anwendung kamen. Darüber hinaus vor allem in der Qualität und Quantität der Eingangsdaten. Diese stammen aus verschiedensten Quellen, wie Primärerhebungen oder Registerauszügen, und unterscheiden sich daher stark in ihrer Güte. Beispielsweise wurden schon beim letzten Zensus 2011 viele Fehler in den Melderegistern festgestellt, die aber aufgrund des Rückspielverbots nicht korrigiert wurden. Außerdem standen in der Entwicklungsphase keinerlei Echtdateien für Tests zur Verfügung. Zur Durchführung von Lasttests haben wir daher beispielsweise erfolgreich einen Testdatengenerator implementiert, der zur Erzeugung großer Mengen randomisierter, synthetischer, aber dennoch realitätsnaher Eingabedaten diente.

Herausfordernd war auch der Bereich der Spezifikation der fachlichen Anforderungen. Einige Konkretisierungen und Änderungen erfolgten von Fachseite erst in späten Entwicklungsphasen nach Vorliegen von Erhebungsdaten. Zum Beispiel wurden bei der Hochrechnung der Bevölkerungszahlen durch das Statistische Bundesamt laut eigenem Pressestatement nicht in allen Gemeinden die gesetzten Präzisionsziele erreicht, sodass wir fachliche und technische Anpassungen in einer späten Projektphase umsetzen mussten.

### Welche Methoden und Technologien kamen in der Haushaltegenerierung zum Einsatz? Gab es Neuerungen im Vergleich zur Haushaltegenerierung im Zensus 2011?

**DR. MEIERLING:** Zur Lösung der mathematischen Problemstellungen wurden eigens zu diesem Zweck entworfene und implementierte Spezialalgorithmen verwendet. Zusätzlich kamen auch spezielle einschlägige Softwarepakete zum Einsatz, die über entsprechende Schnittstellen in die Anwendung eingebunden wurden.

Die Verwendung von Verfahren des maschinellen Lernens, einem Teilgebiet der künstlichen Intelligenz, war eine wichtige Neuerung im Vergleich zum Zensus 2011. Dabei wurden auf Trainingsdaten, die aus den Erhebungsdaten von Personen, Wohnungen und Anschriften bestanden, zunächst statistische Modelle (sogenannte „Random Forests“) trainiert. Diese Prädiktionsmodelle ermöglichten in der Haushaltegenerierung dann Schlussfolgerungen über Eigenschaften von Personen und Wohnungen. Vereinfacht ausgedrückt lieferten die Modelle Antworten auf Fragen wie „Mit welcher Wahrscheinlichkeit ist diese Person eine Karteileiche?“ oder „Wie viele Personen bewohnen diese Wohnung?“.

Die Prädiktionen kamen beispielsweise in folgendem Kontext zur Anwendung: War eine Zuordnung von Haushalten zu Wohnungen aufgrund fehlender Daten nicht möglich, so wurde die Anzahl der Bewohner jeder Wohnung mithilfe eines Random Forests prädiziert; in einem weiteren Schritt erfolgte dann die Verteilung der Haushalte auf die Wohnungen möglichst passend zur Schätzung unter Verwendung von Algorithmen der kombinatorischen Optimierung.



**Die hochkomplexen und anspruchsvollen Arbeiten zur Haushaltegenerierung scheinen Ihnen viel Freude bereitet zu haben. Welche weiteren Projekte sind für Sie vorgesehen und worauf freuen Sie sich besonders?**

**DR. MEIERLING:** Bereits seit einem halben Jahr arbeite ich an einem neuen Projekt, das nicht weniger spannend als die Haushaltegenerierung ist: Die bisherige Fachanwendung zum Verbraucherpreisindex (siehe auch nebenstehende Informationen) wird neu aufgestellt. Mit diesem Software-System wird sowohl die monatliche Berechnung des Verbraucherpreisindex als auch – in fünfjährigem Turnus – dessen Revision durchgeführt. Neben der Aktualisierung des Wägungsschemas, dient die Revision der Umstellung des Verbraucherpreisindex auf ein neues Basisjahr und der Einführung methodischer Verbesserungen. Derzeit läuft hierzu ein Vorprojekt, das auf insgesamt zwei Jahre ausgelegt ist. In enger Zusammenarbeit mit dem Fachbereich übernehme ich die Konsolidierung der fachlichen Anforderungen. Die Ergebnisse werden dann in ein neues Fachkonzept und erneuerte technische Spezifikationen überführt. Ich freue mich, dass ich beim Bayerischen Landesamt für Statistik mit vielen spannenden Herausforderungen konfrontiert werde und an den Lösungen mitwirken darf.

*Das Gespräch führte Bettina Mösch.*

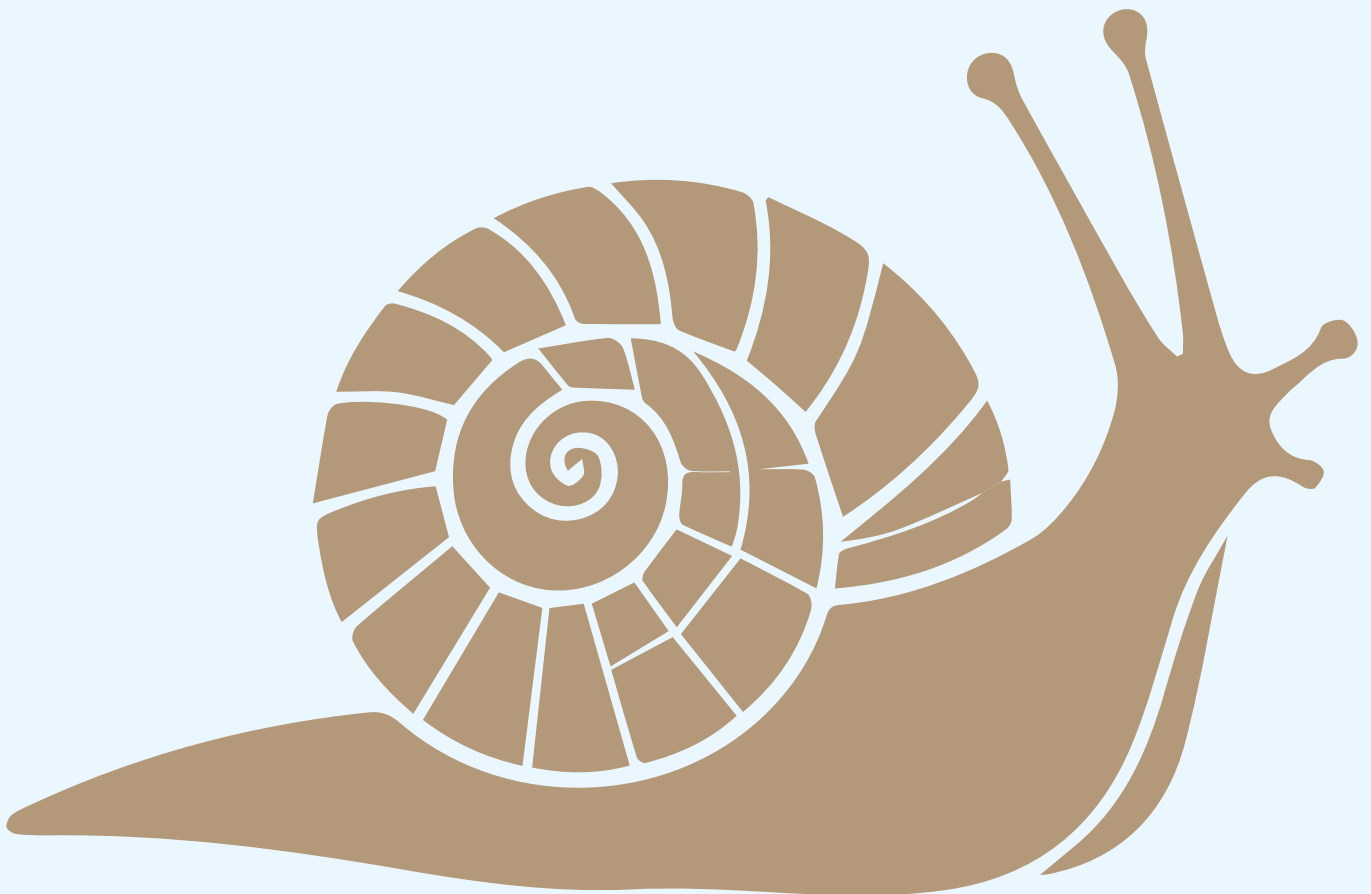
### Verbraucherpreisindex

Der Verbraucherpreisindex misst die durchschnittliche Preisentwicklung aller Waren und Dienstleistungen, die private Haushalte für Konsumzwecke kaufen, wie beispielsweise Nahrungsmittel, Bekleidung, Elektroartikel, Kraftfahrzeuge oder Mieten. Die Veränderung des Verbraucherpreisindex im Vergleich zum Vorjahresmonat bzw. zum Vorjahr wird als Inflationsrate bezeichnet. Zur Berechnung des Verbraucherpreisindex wird ein „Warenkorb“ verwendet, der rund 700 Güter und Dienstleistungen umfasst und eine repräsentative Auswahl sämtlicher von privaten Haushalten in Deutschland gekauften Waren und Dienstleistungen bildet. Die Preiserhebung erfolgt durch die Statistischen Landesämter und das Statistische Bundesamt. Mit welchen Gewichten die Güterarten des Warenkorbs in den Gesamtindex einfließen, ist im Wägungsschema festgehalten, das für den Verbraucherpreisindex üblicherweise alle fünf Jahre aktualisiert wird.

Der Verbraucherpreisindex ist für die Geldpolitik der europäischen Zentralbank von zentraler Bedeutung. Er spielt zudem eine große Rolle bei Tarifverhandlungen und wird sehr häufig auch in privaten Verträgen (zum Beispiel Miet- oder Pachtverträgen) in sogenannten Wertsicherungsklauseln als Wertsicherungsmaßstab verwendet.

# 3 012 660

Gebäude mit Wohnraum  
und Wohngebäude  
in Bayern  
am 9. Mai 2011



\* Quelle: Bayerisches Landesamt für Statistik:  
Zensus 2011 – Gemeindedaten Gebäude und Wohnungen  
Ergebnisse für Bayern (Mai 2014)

# ZENSUS 2011: Ablauf der Haushaltegenerierung\*

Dipl.-Stat. Ingrid Kreuzmair, Dipl.-Math. Marco Reisch

Ein wichtiges Ziel des Zensus 2011 ist es – neben der Berechnung der amtlichen Einwohnerzahl – Daten zu Zahl und Struktur von Haushalten und deren Wohnsituation zu gewinnen. Da diese Informationen jedoch nicht in den Melderegistern vorhanden sind, wurde mit der Haushaltegenerierung ein neuartiges regelbasiertes Verfahren entwickelt, um diese Haushaltszusammenhänge zu ermitteln, weitere kombinierte Auswertungen zu ermöglichen und einen zur amtlichen Einwohnerzahl passenden und entsprechenden Datensatz zu erhalten. Dabei werden – in einer Vielzahl von Prozessschritten – die für eine Anschrift vorliegenden Personeninformationen zu Haushalten zusammengefasst und den in der Gebäude- und Wohnungszählung (GWZ) erfassten Wohnungen an der Anschrift zugeordnet. Die Haushaltegenerierung bildet also die Wohnhaushalte an einer Anschrift ab, für die einerseits reine Haushaltsergebnisse (z. B. Anzahl der Ein- und Mehrpersonenhaushalte), andererseits kombinierte Ergebnisse (z. B. durchschnittliche Wohnungsgröße von Haushalten mit Kindern) ausgewiesen werden können.

## Datenbasis der Haushaltegenerierung

Auf Grund des für einen Zensus erstmalig in Deutschland durchgeführten registergestützten Verfahrens bedient sich auch die Haushaltegenerierung einer Vielzahl unterschiedlicher Datenquellen. Basis aller Erhebungsteile bildet das Anschriften- und Gebäuderegister, welches alle Anschriften mit Wohnraum der Bundesrepublik Deutschland enthält und eigens für diesen Zensus aufgebaut wurde. Dieses Register ermöglicht es, in allen anderen Erhebungs- und Aufbereitungsschritten – speziell auch im Verfahren der Haushaltegenerierung – die Anschrift selbst als kleinste regionale Einheit zu erfassen.

Weiterhin bilden die Registerdaten der Einwohnermeldeämter die Grundlage für einen Einzeldatensatz auf Personenebene, welcher bereits vor der Haushaltegenerierung verschiedenen Aufbereitungs- und Bereinigungsschritten zugeführt wurde, aber auch noch im Haushaltegenerierungsverfahren selbst weiteren Bereinigungsschritten unterliegt. So liegen die Melderegisterdaten bereits zu Beginn der Haushaltegenerierung vor, bereinigt um die Ergebnisse der Befragung in Wohnheimen und Ge-

meinschaftsunterkünften (gemäß § 8 ZensG 2011) und der Befragung zur Klärung des Wohnsitzes (gemäß § 15 ZensG 2011). Als weitere ergänzende Befragungen zur Ermittlung der amtlichen Einwohnerzahlen dienten die Haushaltebefragung (gemäß § 7 ZensG 2011) in Gemeinden ab 10 000 Einwohnern und die Befragung zur Klärung von Unstimmigkeiten (gemäß § 16 ZensG 2011) in Gemeinden mit weniger als 10 000 Einwohnern. Die Ergebnisse dieser beiden primärstatistischen Erhebungen werden im



Foto: Statistisches Bundesamt.

\* Quelle: Bayern in Zahlen 09/2012

Haushaltegenerierungsverfahren korrigierend in den Einzelpersonenbestand eingearbeitet.

Für die Ermittlung von Zahl und Struktur von Haushalten und deren Wohnsituation werden Daten zu Gebäuden und Wohnungen benötigt. Da in Deutschland kein Register mit dieser Information existiert, wurde in einer begleitenden Vollerhebung, der Gebäude- und Wohnungszählung, der Bestand aller Eigentümer und Verwalter von Wohnraum erfragt. Die Ergebnisse der Befragung werden von der Haushaltegenerierung genutzt, um unter anderem Haushalte besser abgrenzen zu können.

Man kann also zusammenfassen: Im Rahmen der Haushaltegenerierung werden sämtliche Datenquellen des Zensus 2011 zusammengeführt und zu einem zensustypischen Datensatz verknüpft, der fachlich und regional tief gegliederte Auswertungen zulässt.

Dieser wird im Rahmen des in der Haushaltegenerierung eingebetteten Korrekturverfahrens noch mit den berechneten amtlichen Einwohnerzahlen synchronisiert. Hierzu fließen hochgerechnete Ergebnisse aus der Haushaltebefragung in das Verfahren ein.

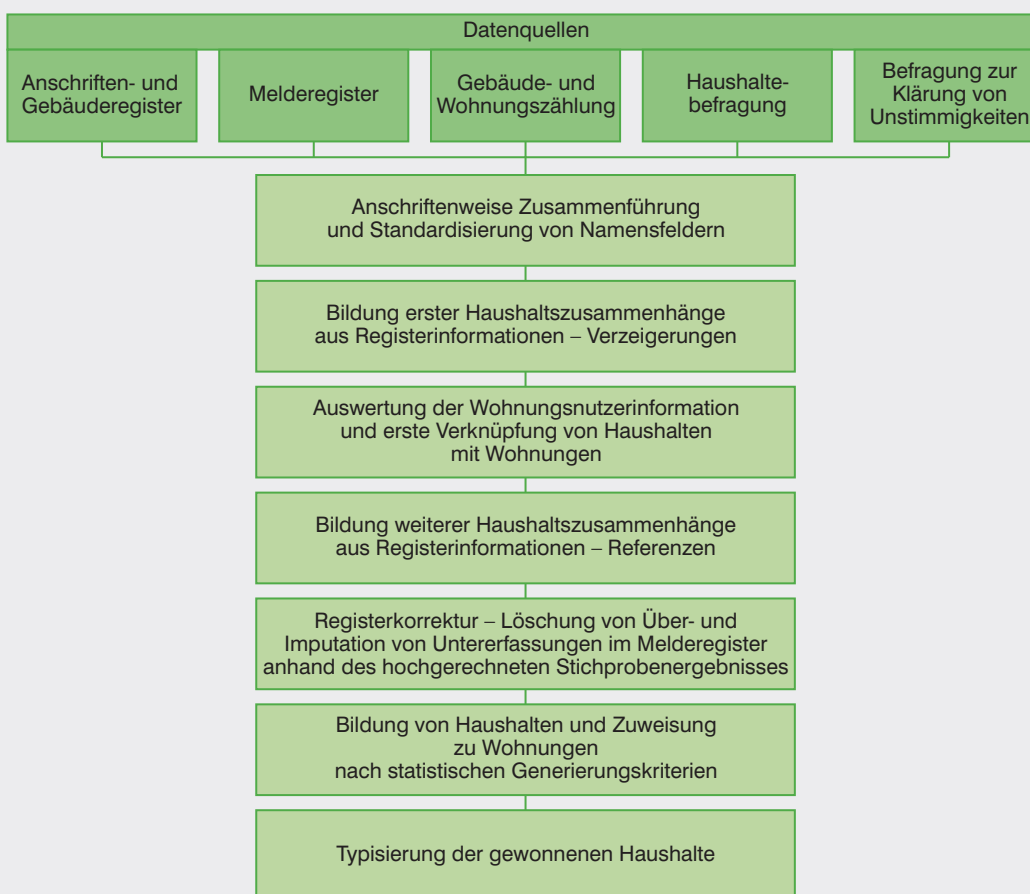
### Grundstruktur des Verfahrens der Haushaltegenerierung (vgl. Abbildung 1)

Die Bezeichnung Haushaltegenerierung lässt den Trugschluss zu, dass dieses Verfahren einzig und allein dazu dient, Haushaltszusammenhänge zwischen im Melderegister vorhandenen Personen zu bilden. Zusätzlich erfüllt die Haushaltegenerierung aber noch die beiden wichtigen Ziele der Korrektur des Registerpersonenbestands und der Erstellung eines Zensustypischen Datensatzes.

Die namensgebende Bildung von Haushalten selbst ist ein regelbasiertes Verfahren, welches Merkmale

Komponenten der Zusammenführung und Haushaltegenerierung

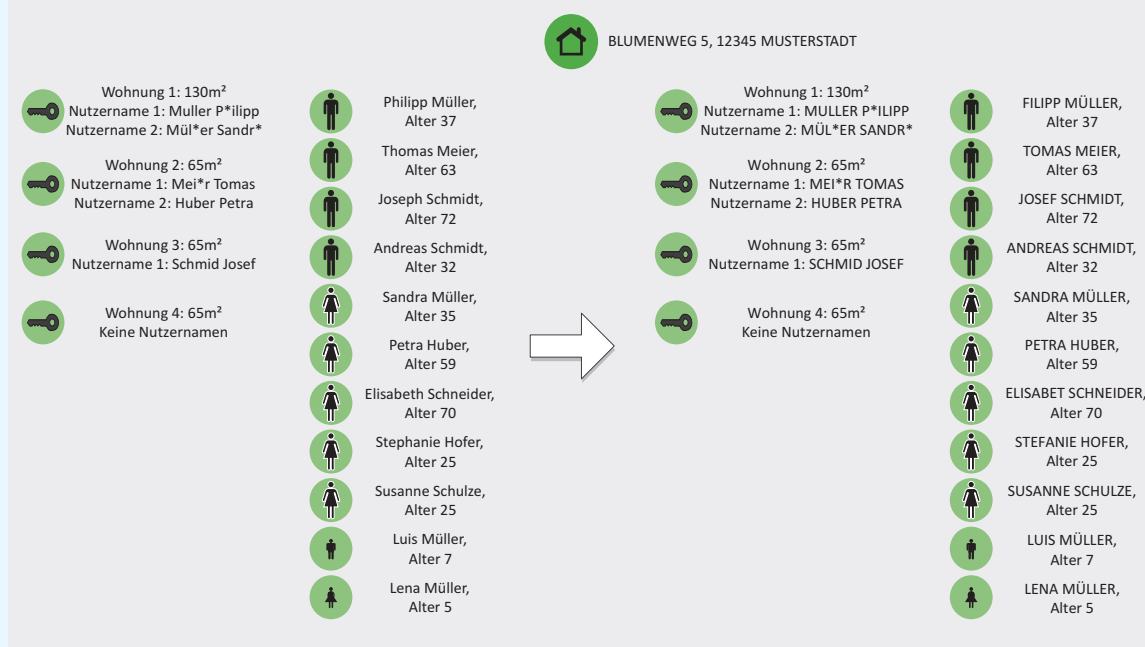
Abb. 1





## Standardisierung der Namensfelder

Abb. 2



aus den unterschiedlichen Datenquellen kombiniert, analysiert und zu Haushalten in Wohnungen zusammengeführt. Dabei ist es möglich, regional tief gegliederte Ergebnisse über Haushaltszusammenhänge zu erhalten.

Die Korrektur wird auf dem Registerpersonenbestand mit Hilfe der Erkenntnisse aus der Haushaltebefragung durchgeführt. Dabei ist zu beachten, dass die mittels Primärstatistik festgestellten Über- bzw. Untererfassungen der Melderegister in zwei separaten Schritten korrigiert werden. Dies ist notwendig, um Einzeldaten für flexible Auswertungen bereitzustellen.

Somit werden insgesamt drei große Ziele mit dem Verfahren Haushaltegenerierung verfolgt. Diese werden allerdings nicht in einzelnen Verfahrensteilen umgesetzt, da keine strikte Trennung möglich ist. Zum Beispiel hat die Registerkorrektur Einfluss auf Haushaltszusammenhänge und umgekehrt, und auch der Zensustypische Datensatz wird sukzessive optimiert.

#### Anschriftenweise Zusammenführung und Standardisierung von Namensfeldern

(vgl. Abbildung 2)

Die gemeinsame Datenbasis aller Erhebungsteile, das Anschriften- und Gebäuderegister, ermöglicht

bereits zum Start der Haushaltegenerierung eine Vorgehensweise auf Anschriftenebene. So ist zum Beispiel bekannt, welche Personen gemäß Melderegister und welche Gebäude und Wohnungen gemäß der Gebäude- und Wohnungszählung einer bestimmten Anschrift zuzuordnen sind. Um nun die an einer Anschrift gemeldeten Personen sinnvoll auf Wohnhaushalte zu verteilen und die resultierenden Haushalte noch sinnvoll den vorhandenen Wohnungen zuzuordnen, bedarf die Haushaltegenerierung unterschiedlicher maschineller Abgleichverfahren. Die Gefahr bei rein maschinellen Verfahren besteht darin, dass auf Grund möglicher Fallkonstellationen Falschzuordnungen erfolgen. Um diese Gefahr zu minimieren, ist es notwendig, dass die dem Verfahren zugeführten Informationen in einer einheitlichen Form und Struktur vorliegen. Dies betrifft zum einen die personenbezogenen Informationen zu Namen aus dem Melderegister, welche verwendet werden, um zwischenmenschliche Beziehungen zwischen Personen herzustellen, und den Informationen zu Wohnungsnutzern aus der Gebäude- und Wohnungszählung, welche zur Verknüpfung von Personen mit Wohnungen herangezogen werden. Die Wohnungsnutzernamen resultieren aus der für die Auskunftspflichtigen der Gebäude- und Wohnungszählung gegebene Möglichkeit, bis zu zwei

Personen namentlich zu benennen, welche eine Wohnung zum Stichtag des Zensus 2011 bewohnten.

Die Standardisierung der Namensfelder umfasst:

- eine Bereinigung um sämtliche Sonderzeichen und überflüssige Leerzeichen
- das gezielte Ersetzen bestimmter Buchstaben/ Buchstabenfolgen, welche bedingt durch die verschiedenen Datenquellen des Öfteren variieren (z. B. th -> t, ph -> f, ß -> ss etc.)
- das Umsetzen aller Klein- in Großbuchstaben.

Damit ist nach diesem Bereinigungsschritt in allen Datenquellen eine vergleichbare Datenbasis enthalten und das Ziel, den Abgleich einzelner Merkmale zu erleichtern, erfüllt.

### Bildung erster Haushaltszusammenhänge aus Registerinformationen – Verzeigerungen (vgl. Abbildung 3)

Das Melderegister enthält Verweise zwischen Personen, welche Verzeigerungen genannt werden. Diese Verzeigerungen geben eindeutige Hinweise auf zwischenmenschliche Beziehungen zwischen zwei gemeldeten Personen. Dabei kann zwischen zwei Verzeigerungsarten unterschieden werden. Zum einen werden mit Hilfe von Verzeigerungen eingetragene Partnerschaften zwischen zwei Personen gekennzeichnet. Dies sind im Einzelnen die klassische Ehe zwischen zwei nicht gleichgeschlechtlichen Personen und die eingetragene Lebenspartnerschaft zwischen gleichgeschlechtlichen Personen. Zum anderen werden mit Hilfe von Verzeigerungen gesetzliche Vertretungen gekennzeichnet. Dabei kann

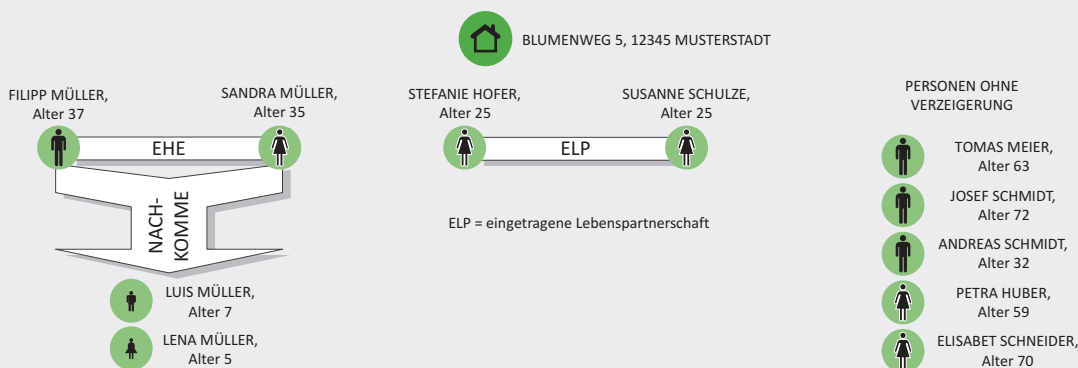
es sich um Eltern-Kind-Beziehungen, aber auch um sonstige gesetzliche Vertretungen handeln, wobei Kinder üblicherweise lediglich bis zum Abschluss des achtzehnten Lebensjahres verzeigert sind.

In diesem Prozessabschnitt werden die Verzeigerungen aus dem Melderegister genutzt, um erste Mehrpersonenhaushalte zu generieren. Zu beachten ist hierbei, dass eine Person nur an höchstens einer Partnerschaftsbeziehung teilnehmen und auch höchstens zwei Elternteile haben kann. Nach diesem Schritt ist jede Registerperson einem vorläufigen Haushalt zugeordnet, welcher sich im weiteren Verlauf der Haushaltgenerierung immer wieder verändern kann. Diese Haushalte bestehen entweder aus Einzelpersonen oder aus einfachen Familienstrukturen resultierend aus den Verzeigerungen.

Der Erfolg dieses Generierungsschritts hängt unmittelbar von der Qualität der im Melderegister geführten Informationen ab. Im Idealfall kann die im Melderegister geführte Ordnungsnummer jeder Person bereits zur Verzeigerung zweier Personen verwendet werden, wenn diese im entsprechenden Feld einer zweiten Person angegeben ist. Ist diese Ordnungsnummer nicht angegeben, so müssen über personenbezogene Merkmale wie Familienname, Vorname und Geburtsdatum – für die jeweils ein separates Feld für die einzelnen Verzeigerungsarten vorhanden ist – Beziehungen zwischen Personen erstellt werden. Um möglichst alle im Melderegister vorhandenen Verzeigerungen erfolgreich nutzen zu können, ist an dieser Stelle ein mehrstufiges Abgleichverfahren notwendig. Hierbei werden

### Auswertung der Verzeigerungen

Abb. 3



zunächst bidirektional angegebene Verzeigerungen ausgewertet, das heißt Verzeigerungen, die bei beiden Beziehungspartnern angegeben sind.

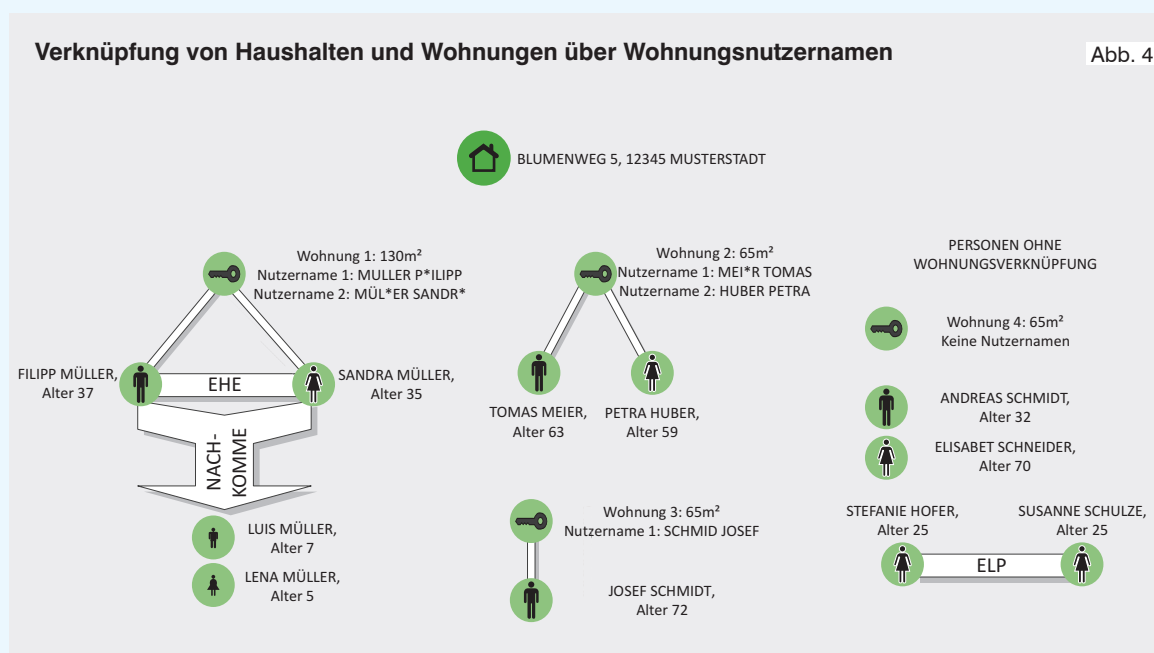
An primärstatistisch erhobenen Anschriften gilt hier als zusätzliche Einschränkung nicht die Anschriftenebene als alleinige Generierungsgrundlage, sondern zusätzlich der Haushalt selbst. Dies ist dadurch begründet, dass an primärstatistisch erhobenen Anschriften bereits aus der Erhebung hervorgeht, welche Personen gemeinsam einen Wohnhaushalt bilden. Dieser primärstatistische Haushaltszusammenhang darf in keinem Fall verletzt werden. Auch ist zu beachten, dass Verzeigerungen immer anhand von Melderegisterangaben generiert werden. Damit haben primärstatistisch ermittelte Fehlbestände zu diesem Zeitpunkt lediglich die Möglichkeit, mit einer weiteren Person verzeigert zu werden, wenn diese zweite Person auch schon im Melderegister gemeldet war, also kein Fehlbestand ist. Somit kann in diesem Schritt kein Haushalt generiert werden, der rein aus primärstatistisch erhobenen Fehlbestandspersonen besteht.

#### Auswertung der Wohnungsnutzerinformation und erste Verknüpfung von Haushalten mit Wohnungen (vgl. Abbildung 4)

Um erste Verknüpfungen zwischen den aus der Gebäude- und Wohnungszählung gelieferten Woh-

nungen und den im Melderegister gemeldeten Personen herzustellen, findet ein maschinelles Abgleichverfahren von Namensfeldern der beiden Erhebungsteile statt, der sogenannte maschinelle Namensabgleich. Dies ist ein mehrstufiges Verfahren, welches sich dem Prinzip der sukzessiven Massenreduktion bedient. Dabei werden nach jeder erfolgreichen Zuordnung die beiden zu vergleichenden Datenmengen um die an der Zuordnung beteiligten Elemente verringert. Die verbleibenden Elemente durchlaufen erneut das komplette Abgleichverfahren. Weiterhin werden zunächst sehr strikte Regelungen eingesetzt, um exakte bzw. nahezu vollständige Übereinstimmung zwischen den zu vergleichenden Zeichenketten zu finden. Im Laufe des Verfahrens werden diese Regelungen immer mehr aufgeweicht, um auch weniger gute aber dennoch ausreichende Übereinstimmung im Abgleich zuzulassen. Sofern in einem Durchlauf des Verfahrens mindestens zwei Elemente der beiden Datenmengen als identisch eingestuft werden können, existiert auch ein Paar mit dem höchsten Maß an Übereinstimmung. Für dieses Paar findet eine Verlinkung statt. Dieser Verfahrensschritt führt also zu Verknüpfungen zwischen Personen und Wohnungen bzw. zwischen bereits generierten Haushalten und Wohnungen.

Da die Auskunftspflichtigen der GWZ die Möglichkeit hatten, bis zu zwei Namen von Wohnungsnut-



zern anzugeben, können auch in diesem Schritt bestehende Haushalte durch Zusammenfassung oder Trennung verändert werden, sofern der Haushaltszusammenhang nicht bereits aus einer primärstatistischen Erhebung bekannt ist. Eine Zusammenfassung von Haushalten findet immer dann statt, wenn zwei Namen von Wohnungsnutzern angegeben wurden und diese auf zwei Registerpersonen verweisen, die sich nach dem ersten Schritt noch in unterschiedlichen Haushalten befinden. Somit ist dies nahezu die einzige Möglichkeit, eine klassische Wohngemeinschaft abzubilden. Aber auch Personen in nicht eingetragenen Partnerschaften, welche in der Regel anhand von Registerinformation nur sehr eingeschränkt identifizierbar sind, können in diesem Schritt zueinander finden.

Es besteht auch die Möglichkeit, dass die Wohnungsnutzerangaben unterschiedlicher Wohnungen auf ein und denselben bereits im ersten Schritt generierten Haushalt verweisen. Für die weitere Vorgehensweise ist in diesem Fall entscheidend, welche Personen des Haushalts als Wohnungsnutzer der Wohnungen identifiziert wurden. In den meisten Fällen werden dadurch die unterschiedlichen Wohnungen als Teilwohnungen zu einer Hauptwohnung zusammengefasst. Es existiert jedoch auch eine Konstellation, in der die im ersten Schritt gefundene Haushaltsstruktur wieder aufgelöst werden muss. Immer dann, wenn in einem Haushalt eine Vorfahre-Nachfahre-Beziehung besteht (z.B. Mutter und Kind), der Nachfahre bereits eine definierte Altershürde überschritten hat und zusätzlich als Wohnungsnutzer einer anderen Wohnung als der Vorfahre identifiziert wird, so wird der im ersten Schritt generierte Haushalt wieder aufgetrennt. Somit würden Vorfahre und Nachfahre jeweils einer anderen Wohnung zugeordnet.

Eine große Hürde stellen hierbei die oftmals belegten Wohnungsnutzernamen dar, da beim maschinellen Verfahren der Beleglesung einzelne Zeichen zum Teil falsch beziehungsweise nicht erkannt wurden. Aus diesem Grund wurde der maschinelle Namensabgleich bereits ein erstes Mal im Prozess der Datenaufbereitung ausgeführt, so dass alle Wohnungsnutzerinformationen, die nicht zu einer maschinellen Verknüpfung von Personen und Wohnungen geführt haben, einem manuellen Zu-

ordnungsschritt zugeführt werden konnten. Dieser manuelle Zuordnungsschritt wurde durch Bearbeiter der Statistischen Ämter der Länder durchgeführt und seine Ergebnisse im Lieferdatensatz der Gebäude- und Wohnungszählung an die Haushaltgenerierung hinterlegt. Damit sind auch manuell erzielte Verknüpfungen in diesem Verfahrensschritt der Haushaltgenerierung neben den Ergebnissen eines zweiten Laufs des maschinellen Namensabgleichs nutzbar, um Verknüpfungen zwischen Personen und Wohnungen herzustellen.

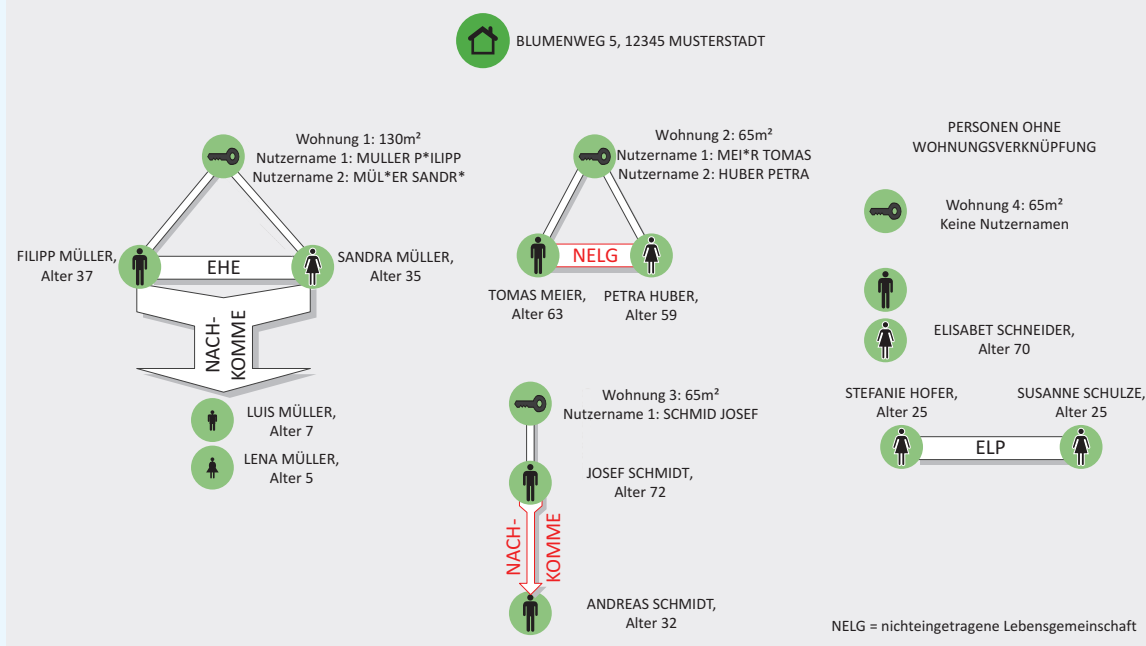
Eine weitere Hürde stellen erhebungsteilübergreifende Unplausibilitäten in den Daten dar. So besteht zum Beispiel die Möglichkeit, dass bereits durch Primärstatistik erhobene Haushaltszusammenhänge nicht mit den Angaben zu den Wohnungsnutzern der GWZ zusammenpassen. Wenn also zwei Personen, welche gemäß Befragung in unterschiedlichen Haushalten leben, als Wohnungsnutzer einer Wohnung identifiziert werden können, so kann im Einzelfall nicht entschieden werden, welcher der Haushalte mit der Wohnung verknüpft werden soll. Dies führt dazu, dass keiner der beiden Haushalte in diesem Schritt mit dieser Wohnung verknüpft werden kann, da die Datenlage nicht eindeutig ist. Die Zuordnung erfolgt somit in einem späteren Verfahrensschritt.

#### **Bildung weiterer Haushaltszusammenhänge aus Registerinformationen – Referenzen (vgl. Abbildung 5)**

Da die im ersten Generierungsschritt ausgewerteten Verzeigerungen lediglich klassische Haushaltsstrukturen abbilden – und diese auch nur, wenn sie explizit im Melderegister angegeben sind – ist es notwendig, dass personenbezogene Registerinformationen genutzt werden, um Indizien für weitere Haushaltszusammengehörigkeiten zu erhalten. Dabei werden neben den klassischen Haushaltsstrukturen, wie sie auch anhand der Verzeigerungen bereits gebildet werden konnten, auch nichteingetragene Partnerschaften oder Großeltern-Beziehungen erstellt. Hierzu kommt wiederum ein mehrstufiges regelbasiertes Verfahren zur Anwendung. Je nachdem, welcher Beziehungstyp gesucht wird, werden Kombinationen unterschiedlicher Registerinformationen verwendet. Beispielhaft werden nicht eingetragene

## Referenzenbildung anhand weiterer Registerinformationen

Abb. 5



Partnerschaften anhand der Informationen zu Familienstand, Zuzugsdatum und Zuzugsanschrift genutzt, wobei als Ausschlusskriterium die Namenskonstellation dient.

Zunächst wird zusätzlich zur Anschriftenebene die bereits generierte Haushaltsebene berücksichtigt. Dies liegt darin begründet, dass über die Wohnungsnutzerangaben bis zu zwei eigenständige Haushalte zu einem neuen Haushalt zusammengefasst werden konnten, ohne dass bereits Beziehungen zwischen den ehemals eigenständigen Haushalten bestehen. Da die Information des gemeinsamen Nutzens einer Wohnung den Verdacht nahe legt, dass auch ein ‚familiäres‘ Verhältnis zwischen mindestens zwei Personen existiert, beschränkt sich die Suche zunächst auf Haushaltsebene. Dabei gilt zum Beispiel der Grundsatz, dass zwei Personen mit gemeinsamen Kindern mindestens in einer nichteingetragenen Partnerschaft leben.

Erst wenn innerhalb der vorhandenen Haushalte keine Beziehungen mehr gefunden werden konnten, geht das Verfahren dazu über, auch haushaltsübergreifende Beziehungen zwischen Personen zu suchen. Falls zwischen zwei Personen Indizien für

das Führen eines gemeinsamen Haushalts gefunden werden können, so ist das Abbilden dieser Beziehung immer mit dem Zusammenfassen zweier bisher eigenständiger Haushalte verbunden. Grundsätzlich ist somit erforderlich, dass die Haushalte nicht bereits mit unterschiedlichen Wohnungen verknüpft wurden. In diesem Fall liefert die primärstatistische Erhebung der Gebäude- und Wohnungszählung die Information, dass zwei Haushalte nicht zusammengefasst werden und keine haushaltsübergreifenden Beziehungen gebildet werden dürfen.

Weiterhin gilt, dass primärstatistisch erhobene Haushaltszusammenhänge durch die Haushaltsgenerierung nicht verändert werden dürfen. Somit entfällt für primärstatistisch erhobene Anschriften der Schritt der haushaltsübergreifenden Beziehungssuche.

### Registerkorrektur – Löschung von Über- und Imputation von Untererfassungen im Melderegister anhand des hochgerechneten Stichprobenergebnisses

Beim Zensus 2011 wird in Gemeinden mit 10 000 oder mehr Einwohnern eine Haushaltsstichprobe durchgeführt. Zweck dieser Stichprobe ist neben



der Erhebung von nicht in Registern verfügbaren Daten primär die gemeindeweise Gewinnung von demografischen und haushaltsstatistischen Informationen zu Über- und Untererfassungen (Karteileichen und Fehlbestände) in den Melderegistern. Um einen qualitativ hochwertigen, fachlich und regional flexibel auswertbaren Zensuseinzeldatensatz zu erhalten, muss eine Bereinigung der Karteileichen und Fehlbestände auf der Basis der hochgerechneten Stichprobenergebnisse vorgenommen werden. Zu diesem Zweck wurde ein Verfahren entwickelt, welches die gemeindeweise aggregierten Vorgaben aus der Haushaltsstichprobe umsetzt.

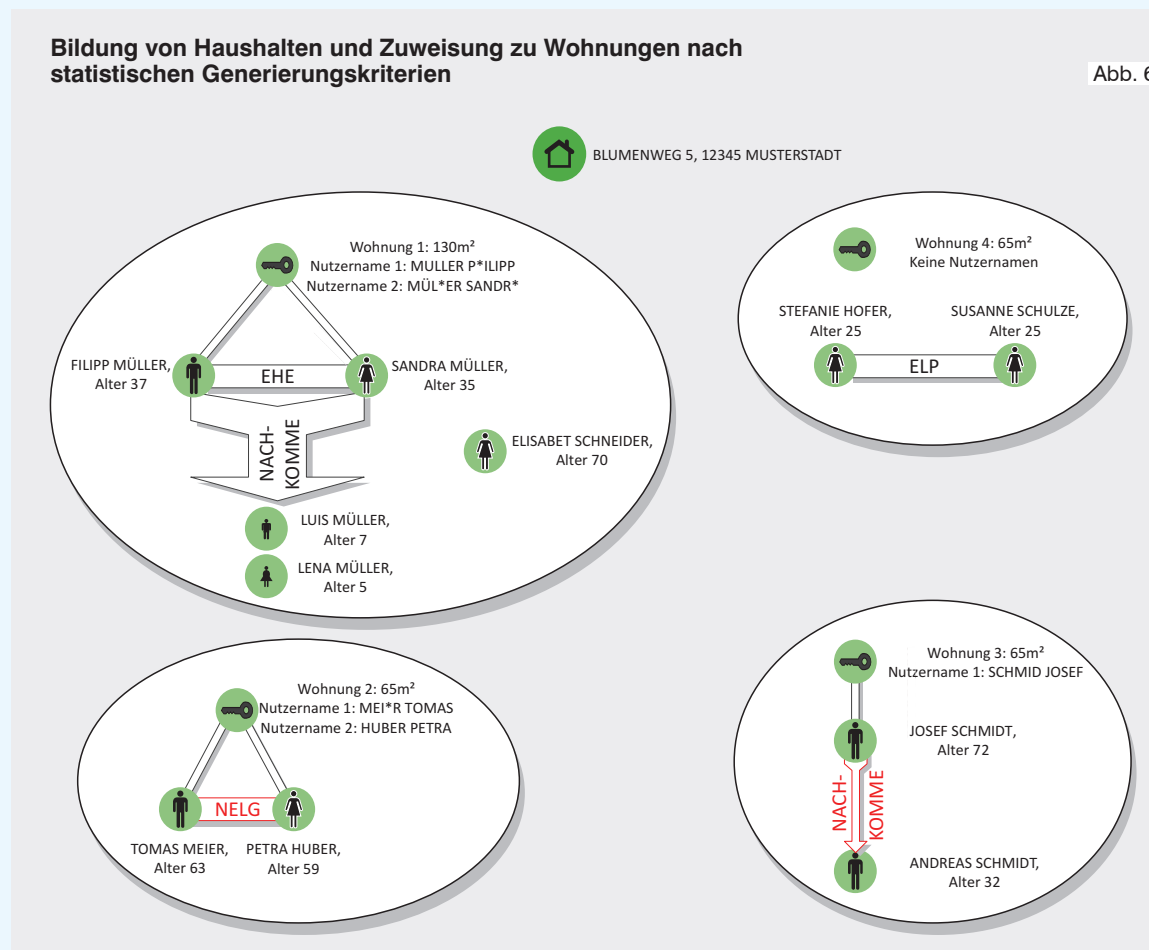
Für detailliertere Informationen zur Registerkorrektur sei auf den Artikel „ZENSUS 2011: Korrekturverfahren im Rahmen der Haushaltegenerierung“ von Dr. Michael Fürnrohr und Katrin Hofmeister verwiesen, der für eine kommende Ausgabe von „Bayern in Zahlen“ geplant ist.

### Bildung von Haushalten und Zuweisung zu Wohnungen nach statistischen Generierungskriterien (vgl. Abbildung 6)

Bisher wurden an einer Anschrift – abgesehen von den im Korrekturverfahren mit reinen Fehlbestandshaushalten belegten Wohnungen – nur die Wohnungen mit den erzeugten Haushalten zusammengeführt, in denen mindestens ein Wohnungsnutzername vorhanden und auch mit den an dieser Anschrift vorhandenen Personen über den Namensabgleich zusammenzuführen war. Die ggf. noch vorhandene Restmenge aller anderen, an dieser Anschrift noch vorhandenen, unverknüpften bewohnten Wohnungen, für die entweder keine Wohnungsnutzernamen angegeben wurden oder keine zu der Angabe der Wohnungsnutzer passende Person an der Anschrift gefunden werden konnte, steht noch für eine Zusammenführung mit den ggf. noch an dieser Anschrift vorhandenen und noch nicht mit einer Wohnung verknüpften Haushalten zur Verfü-

#### Bildung von Haushalten und Zuweisung zu Wohnungen nach statistischen Generierungskriterien

Abb. 6



gung. Es können also noch bewohnte Wohnungen, die noch mit keinem Haushalt verknüpft sind, und/oder Haushalte, die noch mit keiner Wohnung verknüpft sind, vorliegen. Dabei sind an einer Anschrift folgende Szenarien möglich.

**1) An einer Anschrift ist die Menge der noch unverknüpften Haushalte und die Menge der noch unverknüpften bewohnten Wohnungen exakt gleich.**

- a) *sowohl noch unverknüpfte Haushalte als auch unverknüpfte bewohnte Wohnungen, nämlich genau so viele unverknüpfte Haushalte wie unverknüpfte bewohnte Wohnungen*

Die noch unverknüpften Haushalte werden sukzessive mit den noch unverknüpften bewohnten Wohnungen zusammengeführt.

Bei diesem Verfahren wird darauf geachtet, dass die Haushaltstypen nach Möglichkeit den Wohnungen des entsprechenden Typs zugeordnet werden, d.h. ein Haushalt mit ausschließlich Nebenwohnsitzpersonen erhält bevorzugt eine Ferien- und Freizeitwohnung, ein Haushalt mit mindestens einer nicht-meldepflichtigen Person erhält bevorzugt eine Streitkräfte- und Diplomatenwohnung, ein Haushalt mit mindestens einer Hauptwohnsitzperson und ausschließlich meldepflichtigen Personen erhält bevorzugt eine normale Wohnung. Innerhalb dieser bevorzugten Abstufungen werden Haushalte bevorzugt mit den Wohnungen zusammengeführt, für die eine Bewohnerzahl angegeben wurde, die der des Haushalts entspricht. Stehen dann immer noch mehrere Haushalte mehreren Wohnungen für eine potentielle Zusammenführung gegenüber, so werden für die Haushalte Prioritätswerte vergeben, die sich auf das Alter und die Nationalität der Haushaltsmitglieder sowie auf das Bundesland, in dem sich der Haushalt befindet, beziehen. Für die Prioritätswerte der Wohnungen sind Wohnfläche und Raumzahl maßgeblich. Anhand dieser Prioritätswerte werden die Haushalte und Wohnungen schließlich zusammengeführt.

Anschließend stehen keine unverknüpften Haushalte und keine unverknüpften bewohnten Wohnungen mehr zur Verfügung.

- b) *weder unverknüpfte Haushalte noch unverknüpfte bewohnte Wohnungen*

Da weder unverknüpfte Haushalte noch unverknüpfte bewohnte Wohnungen vorhanden sind, die noch verarbeitet werden können, ist die Haushaltegenerierung an dieser Stelle beendet.

**2) An einer Anschrift ist die Menge der noch unverknüpften Haushalte größer als die Menge der noch unverknüpften bewohnten Wohnungen.**

- a) *sowohl noch unverknüpfte Haushalte als auch unverknüpfte bewohnte Wohnungen, jedoch mehr unverknüpfte Haushalte als unverknüpfte bewohnte Wohnungen*

Die noch unverknüpften Haushalte werden – wie oben beschrieben – so lange sukzessive mit den noch unverknüpften bewohnten Wohnungen zusammengeführt, wie noch solche Wohnungen vorhanden sind.

Für die anschließend immer noch unverknüpften Haushalte erfolgt – je nach Art der Anschrift – ein weiteres Verfahren, um auch diese Haushalte noch mit einer Wohnung zusammenführen zu können.

- b) *noch unverknüpfte Haushalte, jedoch keine unverknüpften bewohnten Wohnungen mehr*

An nicht primärstatistisch erhobenen Anschriften wird anhand einer für diesen Zweck aus den Ergebnissen der Haushaltsstichprobe erzeugten gemeinde- und auch kreisweiten Haushaltshochrechnung versucht, die noch unverknüpften Haushalte mit den bereits verknüpften Haushalten zusammenzulegen und damit mit den Wohnungen der bereits verknüpften Haushalte zusammenzuführen. Mit den dabei neu entstehenden Haushaltsstrukturen wird versucht, die hochgerechneten Haushaltszahlen zu erreichen. Die Zusammenlegung der Haushalte erfolgt weiterhin auf Anschriftenebene, die Hochrechnung der Haushalte muss jedoch auf Gemeinde- bzw. Kreisebene erfüllt werden.

An primärstatistisch erhobenen Anschriften, an denen die Haushaltsstrukturen bereits vorgegeben sind, darf diese Zusammenlegung von Haushalten nicht erfolgen. Stattdessen

werden für die noch unverknüpften Haushalte zusätzliche bewohnte Wohnungen künstlich erzeugt. Dazu wird eine Wohnung aus demselben Gebäude kopiert, wobei nach Möglichkeit eine benötigte Mindestwohnfläche berücksichtigt wird.

**3) An einer Anschrift ist die Menge der noch unverknüpften Haushalte kleiner als die Menge der noch unverknüpften bewohnten Wohnungen.**

- a) *sowohl noch unverknüpfte Haushalte als auch unverknüpfte bewohnte Wohnungen, jedoch weniger unverknüpfte Haushalte als unverknüpfte bewohnte Wohnungen*

Die noch unverknüpften Haushalte werden – wie oben beschrieben – sukzessive mit den noch unverknüpften bewohnten Wohnungen zusammengeführt.

Für die anschließend immer noch unverknüpften bewohnten Wohnungen erfolgt – je nach Art der Anschrift – eine Plausibilisierung mit dem Ziel, dass letztendlich die Zahl der Haushalte mit der Zahl der bewohnten Wohnungen übereinstimmt.

- b) *keine unverknüpften Haushalte mehr, jedoch noch unverknüpfte bewohnte Wohnungen*

An nicht primärstatistisch erhobenen Anschriften wird die Art der Wohnungsnutzung der noch unverknüpften bewohnten Wohnungen von bewohnt auf (laut Melderegister) leer stehend korrigiert.

An primärstatistisch erhobenen Anschriften, werden – als Gegenstück zur künstlichen Erzeugung von bewohnten Wohnungen – die noch unverknüpften bewohnten Wohnungen gelöscht.

Es gibt eine EU-Typisierung, in der nur Hauptwohnsitzpersonen berücksichtigt werden, eine sog. nationale EU-Typisierung, die der EU-Typisierung entspricht, zusätzlich jedoch auch Nebenwohnsitzpersonen berücksichtigt, sowie eine nationale Typisierung, die für Haupt- und Nebenwohnsitzpersonen erfolgt und an die Kriterien des Mikrozensus angelehnt ist.

Es werden jeweils Ehen, Lebenspartnerschaften, Nichteheliche Lebensgemeinschaften mit und ohne Kinder, Alleinerziehende mit Kindern und alle dieser Gruppen jeweils mit und ohne sonstige Personen ausgewiesen. Sonstige Ein- und Mehrpersonenhaushalte werden ebenfalls typisiert. Zusätzlich zu diesen Haushaltsmerkmalen (Haushaltstyp, Haushaltsgröße, Stellung einer Person im Haushalt) werden auch noch sogenannte Kernfamilienmerkmale (Familientyp, Familiengröße, Stellung einer Person in der Familie) ausgewiesen. Kernfamilien bestehen aus mindestens zwei miteinander verbundenen Personen und können, müssen aber nicht, mit dem sie umfassenden Haushalt übereinstimmen, sie bilden jedoch immer eine Teilmenge dieses Haushalts. Größter Unterschied der Typisierungen ist, dass die EU-Typisierung mehrere Kernfamilien in einem Haushalt zulässt (z.B. in verschiedenen Generationen), die nationale Typisierung den Haushalt ausgehend von einer Bezugsperson mit maximal einer Kernfamilie identifiziert.

#### Fazit

Die Haushaltegenerierung ist die Synthese der verschiedenen Erhebungsteile des Zensus 2011. Dadurch ermöglicht sie fachlich und regional tiefgegliederte, erhebungsteilübergreifende Kombinationsauswertungen.

#### Typisierung der gewonnenen Haushalte

Sämtliche von der Haushaltegenerierung erzeugten Haushalte erhalten drei verschiedene Typisierungen.

# STATISTIK KOMMUNAL FÜR BAYERN

Die Ergebnisse in „Statistik kommunal“ stammen direkt aus der GENESIS-Datenbank des Landesamts.

Jährlich werden für den gesamten Freistaat, jeden der 7 Regierungsbezirke und 71 Landkreise sowie jede der 25 kreisfreien Städte und 2031 kreisangehörigen Gemeinden fundierte Regionalprofile erstellt – bestehend aus jeweils rund 2 200 spezifischen Daten in 33 Tabellen und 21 Graphiken. Zusätzlich lassen Zeitreihen über mehrere Jahre bzw. Jahrzehnte Entwicklungen erkennen.

Kostenloser Download unter:

[https://s.bayern.de/statistik\\_kommunal](https://s.bayern.de/statistik_kommunal)



bayerischer/stock.adobe.com



Bayerisches Landesamt für  
Statistik





# INFLATION IN BAYERN IM JUNI 2024 BEI 2,7%

*Gesamtindex ohne Nahrungsmittel und Energie im Juni bei 3,3%; Preise für Energie (Haushaltsenergie und Kraftstoffe) fallen um 1,9%*

Im Freistaat steigen die Verbraucherpreise im Juni gegenüber dem Vorjahresmonat um 2,7%. Dabei kosten Nahrungsmittel 1,4% mehr als im Vorjahr. Energie kann hingegen 1,9% günstiger bezogen werden. Ohne Nahrungsmittel und Energie verzeichnet das Bayerische Landesamt für Statistik eine Preissteigerung von 3,3%.

Im Vergleich zum Vormonat steigen die Verbraucherpreise im Juni 2024 um 0,2%. Nahrungsmittel werden im Vergleich zum Mai 2024 (0,4%) teurer. Die Preise für Kraftstoffe (–2,7%) und Heizöl (–1,3%) liegen deutlich unter dem Vormonatswert.

## Inflationsrate im Juni

Die Inflationsrate, gemessen als %uale Veränderung des Verbraucherpreisindex gegenüber dem entsprechenden Vorjahresmonat, liegt im Juni 2024 bei 2,7%. Im Vergleich zum Vormonat steigen die Verbraucherpreise im Juni 2024 um 0,2%.

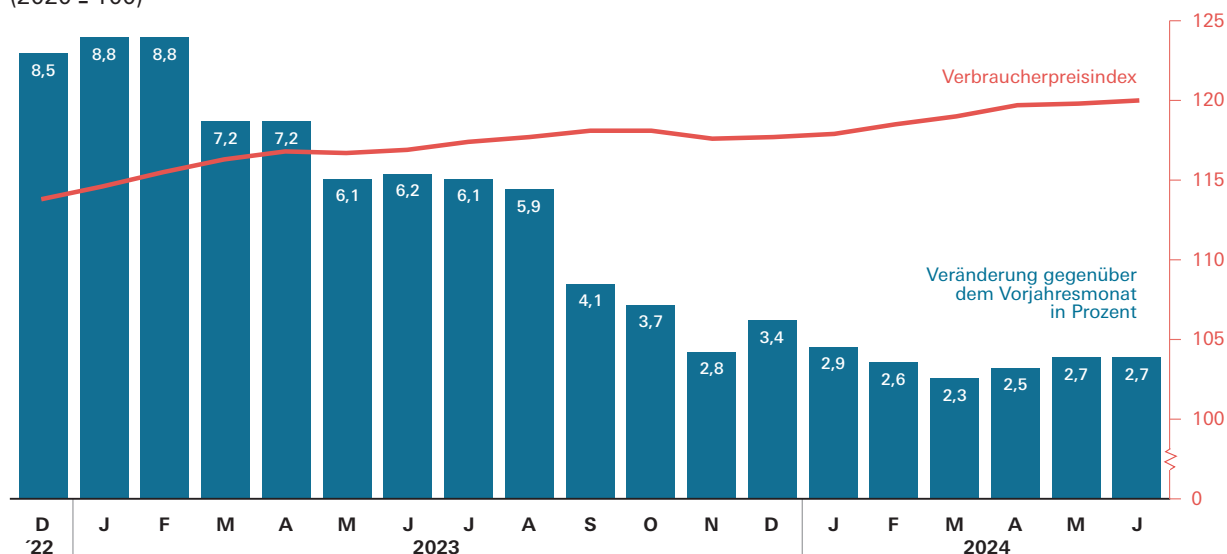
Die Teuerungsrate des Gesamtindex ohne Nahrungsmittel und Energie, in der öffentlichen Diskussion oft als Kerninflationsrate bezeichnet, beläuft sich in diesem Monat auf 3,3% im Vergleich zum Vorjahresmonat.

## Nahrungsmittel werden geringfügig teurer

Im Jahresvergleich müssen die Verbraucherinnen und Verbraucher im Juni 2024 leichte Preissteigerungen bei Nahrungsmitteln hinnehmen (+1,4%). Im Vergleich zum Vormonat steigen die Preise um 0,4%. Obst wird mit einem Plus von 1,0% und Gemüse von 0,8% gegenüber dem Vorjahr teurer. Beim Gemüse sind vor allem Kartoffeln mit 14,4% deutlich im Preis gestiegen. Karotten hingegen sind 20,8% günstiger als im Vorjahresmonat. Bei den beliebten Tomaten verzeichnen die Expertinnen und Experten des Bayerischen Landesamts für Statistik immerhin noch eine Preissenkung von 3,2%.



**Verbraucherpreisindex für Bayern von Dezember 2022 bis Juni 2024**  
(2020 = 100)



### Kraftstoffe werden zum Start in die Feriensaison günstiger

Die Preise für Energie liegen im Juni 2024 1,9% unter dem Vorjahreswert. Im Vergleich zum Vorjahresmonat wird Erdgas mit 9,6% deutlich günstiger. Brennholz, Holzpellets oder andere feste Brennstoffe (–20,1%) sowie Strom (–4,0%) können ebenfalls deutlich günstiger bezogen werden, während Heizöl (+8,9%) im Vergleich zum Vorjahr teurer ist. Kraftstoffe liegen leicht unter dem Wert des Vorjahres (–0,8%).

Tanken wird zum Start in die Feriensaison im Vergleich zum Vormonat Mai deutlich günstiger (–2,7%). Die Preise für Brennholz, Holzpellets oder andere feste Brennstoffe und beim Heizöl fallen im gleichen Zeitraum um jeweils 1,3%. Gas (+0,2%) und Strom (+0,6%) werden hingegen teurer.

### Wohnungsmieten – moderate Veränderungen

Die Preisentwicklung bei Wohnungsmieten ohne Nebenkosten verläuft leicht unterdurchschnittlich im Vergleich zum Gesamtindex. Gegenüber dem Vorjahresmonat erhöhen sie sich im Juni 2024 um 2,6%.

Die Presseinformation zum Monatsbericht Juni 2024 enthält vorläufige Ergebnisse.

Regionalisierte Zahlen stehen nicht zur Verfügung.

Ausführliche Ergebnisse enthält der Statistische Bericht „Verbraucherpreisindex für Bayern. Monatliche Indexwerte von Mai 2020 bis Juni 2024 mit Gliederung nach Haupt- und Sondergruppen“, kostenlos abrufbar unter:  
[www.statistik.bayern.de/statistik/preise\\_verdienste/preise](http://www.statistik.bayern.de/statistik/preise_verdienste/preise)

# BAYERISCHER ZAHLENSPIEGEL

Einheit	Vorjahres- monat	2023		2024					
		November	Dezember	Januar	Februar	März	April	Mai	Juni

## Preise

### Verbraucherpreisindex (2020 = 100)

Gesamtindex .....	%	116,9	117,6	117,7	117,9	118,5	119,0	119,7	119,8	120,0
Nahrungsmittel und alkoholfreie Getränke .....	%	128,8	129,9	130,1	130,6	130,2	130,3	130,8	131,0	131,3
Alkoholische Getränke und Tabakwaren .....	%	117,8	119,5	119,1	120,4	121,0	121,1	121,8	121,7	121,3
Bekleidung und Schuhe .....	%	108,0	110,3	110,6	103,7	107,3	110,4	112,1	112,3	111,1
Wohnung, Wasser, Strom, Gas und andere Brennstoffe .....	%	115,0	115,3	115,2	116,3	116,4	116,3	117,4	117,4	117,5
Möbel, Leuchten, Geräte u. a. Haushaltszubehör .....	%	119,9	120,6	120,3	120,7	120,6	120,6	120,6	120,0	119,8
Gesundheit .....	%	104,2	105,3	105,6	107,3	107,4	107,6	107,6	106,9	107,7
Verkehr .....	%	123,6	124,1	124,2	123,1	124,1	125,8	126,2	126,4	125,7
Post und Telekommunikation .....	%	99,7	99,8	100,0	99,9	99,8	99,6	99,1	99,2	99,0
Freizeit, Unterhaltung und Kultur .....	%	114,9	113,6	114,6	112,5	114,1	114,7	114,8	115,3	116,4
Bildungswesen .....	%	115,8	120,7	120,8	124,4	124,4	126,5	127,0	127,0	126,9
Gaststätten- und Beherbergungsdienstleistungen .....	%	119,0	120,4	119,7	122,3	123,1	123,5	124,4	125,6	127,3
Andere Waren und Dienstleistungen .....	%	112,1	113,8	114,0	116,3	116,8	117,6	118,7	118,9	119,4
Dienstleistungen ohne Nettokaltmiete .....	%	112,7	113,1	114,0	114,2	115,1	116,5	117,0	117,8	118,6
Nettokaltmiete .....	%	107,0	108,1	108,3	108,8	108,9	109,1	109,4	109,7	109,8

### Preisindex für Bauwerke<sup>1</sup> (2021 = 100)

Wohngebäude insgesamt (reine Baukosten) .....	%	125,8	126,1	.	.	127,2	.	.	127,7	.
davon Rohbauarbeiten .....	%	121,6	120,3	.	.	121,1	.	.	121,3	.
Ausbauarbeiten .....	%	129,1	130,7	.	.	132,0	.	.	132,8	.
Schönheitsreparaturen in einer Wohnung .....	%	124,0	125,2	.	.	125,2	.	.	127,0	.
Bürogebäude .....	%	126,8	127,5	.	.	128,8	.	.	129,5	.
Gewerbliche Betriebsgebäude .....	%	126,8	127,3	.	.	128,5	.	.	129,2	.
Straßenbau .....	%	124,2	125,9	.	.	127,7	.	.	128,6	.

### Nachrichtlich: Ergebnisse für Deutschland

### Verbraucherpreisindex (2020 = 100)

Gesamtindex .....	%	116,8	117,3	117,4	117,6	118,1	118,6	119,2	119,3	119,4
Nahrungsmittel und alkoholfreie Getränke .....	%	130,4	131,2	131,2	132,3	132,0	131,7	132,2	132,1	132,3
Alkoholische Getränke und Tabakwaren .....	%	117,4	119,3	118,9	120,0	120,6	121,2	122,0	121,5	121,9
Bekleidung und Schuhe .....	%	107,0	109,4	109,3	103,6	106,4	109,8	111,1	111,2	110,3
Wohnung, Wasser, Strom, Gas und andere Brennstoffe .....	%	114,5	114,5	114,4	115,2	115,3	115,3	115,9	115,9	116,0
Möbel, Leuchten, Geräte u. a. Haushaltszubehör .....	%	117,7	118,5	118,5	118,5	118,4	118,4	118,6	118,2	118,1
Gesundheitspflege .....	%	104,8	105,4	105,5	106,8	107,3	107,4	107,6	107,6	107,8
Verkehr .....	%	122,9	123,4	123,8	122,6	123,7	125,1	125,7	125,7	124,9
Post und Telekommunikation .....	%	99,9	100,0	100,0	99,9	99,9	99,6	99,2	99,3	99,2
Freizeit, Unterhaltung und Kultur .....	%	114,9	113,7	114,5	112,3	113,8	114,7	114,8	115,7	116,5
Bildungswesen .....	%	108,4	110,6	110,7	112,8	112,9	113,4	113,6	113,7	113,7
Gaststätten- und Beherbergungsdienstleistungen .....	%	119,9	121,0	121,1	123,4	124,2	125,1	125,9	126,7	127,7
Andere Waren und Dienstleistungen .....	%	113,1	114,5	114,6	116,8	117,2	118,1	118,9	119,3	120,0

1 Einschließlich Mehrwertsteuer.

noch: Preise	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni
noch: Nachrichtlich: Ergebnisse für Deutschland										
Außenhandels-, Erzeuger- und Großhandelspreise in Deutschland										
Index der Einfuhrpreise <sup>1</sup> (2021 ± 100) .....	%	113,2	112,8	111,7	111,7	111,5	111,9	112,7	112,7	...
Ausfuhrpreise <sup>2</sup> (2021 ± 100) .....	%	114,2	113,8	113,5	113,6	113,8	113,9	114,4	114,4	...
Index der Erzeugerpreise gew. Produkte <sup>2</sup> (Inlandsabsatz); (2021 ± 100) .....	%	130,4	128,3	127,3	127,6	127,1	127,3	127,5	127,5	...
Vorleistungsgüterproduzenten .....	%	119,2	115,6	115,9	116,1	116,1	116,2	116,6	117,0	...
Investitionsgüterproduzenten .....	%	113,0	114,0	114,1	114,9	115,1	115,3	115,5	115,7	...
Konsumgüterproduzenten zusammen .....	%	124,3	123,3	123,4	123,5	123,6	124,3	124,7	124,9	...
Gebrauchsgüterproduzenten .....	%	116,9	117,2	117,1	117,5	117,6	117,7	117,8	117,7	...
Verbrauchsgüterproduzenten .....	%	125,5	124,3	124,4	124,5	124,6	125,4	125,9	126,0	...
Energie .....	%	163,4	159,9	155,8	155,7	153,8	153,8	153,7	152,9	...
Erzeugung landwirtschaftlicher Produkte <sup>2</sup> (2020 ± 100) .....	%	139,4p	135,7p	137,6p	138,5p	140,1p	141,3p	142,4p	144,3p	...
Pflanzliche Erzeugung .....	%	138,1p	139,1p	141,9p	144,3p	145,9p	145,6p	147,7p	151,4p	...
Tierische Erzeugung .....	%	140,2	133,5	134,8	134,8p	136,3p	138,6p	138,9p	139,7p	...
Großhandelsverkaufspreise <sup>2</sup> (2015 ± 100) .....	%	117,4	116,9	116,2	116,3	116,2	116,4	116,9	117,0	116,7
darunter Großhandel mit										
Nahrungs- und Genussmitteln, Getränken, Tabakwaren ...	%	124,9	123,2	123,4	124,0	123,6	124,1	124,5	125,5	126,2
festen Brennstoffen, Mineralölerzeugnissen .....	%	123,2	132,8	126,4	125,4	125,7	126,4	129,4	126,4	122,9
Einzelhandel und Kraftfahrzeughandel zusammen (2020 ± 100) .....	%	119,9	121,4	121,4	121,1	121,6	121,7	121,8	121,6	121,5
darunter Einzelhandel mit Waren verschiedener Art .....	%	125,6	126,6	126,4	127,0	127,0	127,1	127,6	127,4	127,6
Einzelhandel mit Nahrungs- und Genussmitteln, Getränken und Tabakwaren .....	%	126,6	127,5	127,4	128,1	128,0	128,1	128,6	128,4	128,7
Kraftfahrzeughandel .....	%	122,5	124,9	125,5	125,1	125,6	125,3	124,4	124,2	124,1

### Gewerbeanzeigen<sup>3</sup>

Gewerbeanmeldungen .....	1 000	9,8	10,0	8,6	12,3	11,5	10,4	9,9	9,4	...
Gewerbeabmeldungen .....	1 000	7,4	9,7	12,2	11,5	9,2	8,1	7,7	6,8	...

### Produzierendes Gewerbe

#### Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden<sup>4</sup>

Betriebe mit 50 oder mehr Beschäftigten .....	Anzahl	4 012	3 990	3 990	3 893	3 949	3 979	3 975	3 974	...
Beschäftigte .....	1 000	1 195	1 207	1 205	1 197	1 197	1 198	1 197	1 196	...
davon Vorleistungsgüterproduzenten .....	1 000	409	409	408	401	402	402	401	400	...
Investitionsgüterproduzenten .....	1 000	583	593	593	593	593	594	593	592	...
Gebrauchsgüterproduzenten .....	1 000	34	33	33	32	32	32	32	32	...
Verbrauchsgüterproduzenten .....	1 000	168	170	169	169	169	169	169	169	...
Energie .....	1 000	2	2	2	2	2	2	2	2	...
Geleistete Arbeitsstunden .....	1 000	145 382	153 900	122 838	148 418	152 138	147 138	152 162	134 498	...
Bruttoentgelte .....	Mill. Euro	6 198	7 527	6 118	6 423	6 019	5 887	6 130	6 527	...
Umsatz (ohne Mehrwertsteuer) .....	Mill. Euro	40 830	44 526	38 205	34 773	38 612	39 287	39 755	37 865	...
davon Vorleistungsgüterproduzenten .....	Mill. Euro	10 087	9 669	7 810	8 632	8 932	9 246	9 126	8 710	...
Investitionsgüterproduzenten .....	Mill. Euro	25 117	29 266	25 316	20 971	24 332	24 536	24 982	23 666	...
Gebrauchsgüterproduzenten .....	Mill. Euro	.	.	.	.	.	.	.	.	...
Verbrauchsgüterproduzenten .....	Mill. Euro	4 215	4 254	3 885	4 040	4 045	4 180	4 231	4 172	...
Energie .....	Mill. Euro	.	.	.	.	.	.	.	.	...
darunter Auslandsumsatz .....	Mill. Euro	24 670	26 417	22 727	20 602	23 537	23 417	23 680	23 166	...

#### Index der Produktion für das Verarbeitende Gewerbe, Bergbau und Gewinnung von Steinen und Erden (preisbereinigt) (2015 ± 100)<sup>4</sup>

Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden .....	%	103,0	107,4	89,6	91,3	100,1	99,7	101,8	88,9	...
Bergbau und Gewinnung von Steinen und Erden .....	%	100,7	94,4	55,2	56,0	68,3	88,0	100,2	92,3	...
Verarbeitendes Gewerbe .....	%	103,0	107,5	89,8	91,5	100,2	99,8	101,8	88,9	...
Vorleistungsgüterproduzenten .....	%	97,2	97,0	74,6	85,2	89,4	90,8	92,5	84,5	...
Investitionsgüterproduzenten .....	%	108,8	117,7	101,8	94,5	108,7	107,2	109,5	90,8	...
Gebrauchsgüterproduzenten .....	%	.	.	.	.	.	.	.	.	...
Verbrauchsgüterproduzenten .....	%	98,2	99,1	87,7	96,9	98,7	97,9	100,5	96,4	...
Energie .....	%	.	.	.	.	.	.	.	.	...

1 Ohne Zölle, Abschöpfungen, Währungsausgleichsbeträge und Einfuhrumsatzsteuer.

2 Ohne Mehrwertsteuer.

3 Ohne Reisegewerbe.

4 Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

noch: Produzierendes Gewerbe	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni
Index des Auftragseingangs im Verarbeitenden Gewerbe (preisbereinigt) (2021 = 100) <sup>1</sup>										
Verarbeitendes Gewerbe <sup>2</sup> insgesamt .....	%	102,8	98,9	119,5	95,7	94,2	101,9	103,0	91,6	...
Inland .....	%	108,9	112,7	142,9	93,6	98,9	101,4	96,6	91,1	...
Ausland .....	%	99,2	90,8	105,7	96,9	91,4	102,2	106,7	91,9	...
Vorleistungsgüterproduzenten .....	%	109,4	102,3	119,4	95,2	89,5	100,8	95,6	84,9	...
Investitionsgüterproduzenten .....	%	99,2	95,8	120,9	94,7	94,2	101,4	105,5	93,4	...
Gebrauchsgüterproduzenten .....	%	94,2	107,2	86,8	83,6	92,7	101,2	95,3	86,7	...
Verbrauchsgüterproduzenten .....	%	121,6	131,6	107,2	131,4	147,7	125,8	128,6	125,2	...

**Baugewerbe****Bauhauptgewerbe/Vorbereitende Baustellenarbeiten, Hoch- und Tiefbau<sup>3</sup>**

Tätige Personen (einschließlich tätiger Inhaber) im Bauhauptgewerbe .....	1 000	109	109	106	104	105	107	108	108	...
Geleistete Arbeitsstunden .....	1 000	11 169	11 125	6 325	5 305	7 842	9 626	11 355	9 846	...
davon Wohnungsbau .....	1 000	3 575	3 370	2 011	1 734	2 541	2 881	3 334	2 980	...
gewerblicher und industrieller Bau .....	1 000	3 686	3 865	2 317	2 209	3 076	3 554	3 945	3 399	...
öffentlicher und Verkehrsbau .....	1 000	3 909	3 890	1 997	1 362	2 226	3 191	4 076	3 467	...
Entgelte .....	Mill. Euro	440,4	537,0	387,8	344,0	342,4	370,3	435,3	433,8	...
Baugewerblicher Umsatz (ohne Umsatzsteuer) .....	Mill. Euro	1 936,3	2 283,4	2 030,4	926,6	1 110,7	1 551,8	1 772,9	1 868,7	...
davon Wohnungsbau .....	Mill. Euro	533,2	564,0	551,1	253,1	331,3	421,2	458,3	470,5	...
gewerblicher und industrieller Bau .....	Mill. Euro	777,4	898,7	824,1	412,7	466,3	657,0	721,7	768,3	...
öffentlicher und Verkehrsbau .....	Mill. Euro	625,7	820,6	655,3	260,8	313,0	473,6	592,8	629,8	...

Messzahlen (2015 = 100)

Index des Auftragseingangs im Bauhauptgewerbe insg. ....	Messzahl	157,4	122,8	175,0	129,4	147,7	216,3	189,9	167,7	...
davon Wohnungsbau .....	Messzahl	138,3	110,5	132,2	89,0	134,8	159,3	138,8	143,0	...
gewerblicher und industrieller Bau .....	Messzahl	151,5	150,5	174,5	145,1	155,5	244,0	239,5	178,5	...
öffentlicher und Verkehrsbau .....	Messzahl	179,9	101,4	210,9	144,8	149,3	231,6	175,4	175,7	...
darunter Straßenbau .....	Messzahl	193,4	77,6	79,8	121,3	129,7	250,5	177,2	196,0	...

**Ausbaugewerbe/Bauinstallation u. sonst. Ausbaugewerbe<sup>4</sup>**

Tätige Personen (einschließlich tätiger Inhaber) im Ausbaugewerbe .....	1 000	86	.	87	.	.	88	.	.	...
Geleistete Arbeitsstunden .....	1 000	24 552	.	24 539	.	.	24 433	.	.	...
Entgelte .....	Mill. Euro	827,0	.	956,5	.	.	881,7	.	.	...
Ausbaugewerblicher Umsatz (ohne Umsatzsteuer) .....	Mill. Euro	2 819,9	.	4 732,0	.	.	2 885,9	.	.	...

**Energie- und Wasserversorgung**

Betriebe .....	Anzahl	392	393	393	399	398	398	397	396	...
Beschäftigte .....	Anzahl	36 862	38 533	38 494	38 586	38 655	38 768	38 993	39 065	...
Geleistete Arbeitsstunden .....	1 000	4 337	4 825	4 014	4 883	4 827	4 582	4 877	4 142	...
Bruttolohn- und -gehaltssumme .....	Mill. Euro	182	317	187	183	184	201	247	210	...
Bruttostromerzeugung der Kraftwerke der allg. Versorgung...	Mill. kWh	1 844,5	2 159,9	2 351,8	2 492,1	1 902,9	1 925,0	1 778,0	1 946,0	...
Nettostromerzeugung der Kraftwerke der allg. Versorgung...	Mill. kWh	1 765,3	2 067,9	2 260,1	2 392,6	1 820,4	1 840,3	1 707,5	1 881,3	...
darunter in Kraft-Wärme-Kopplung .....	Mill. kWh	302,9	587,3	684,1	774,0	548,7	495,6	343,1	308,9	...
Nettowärmeerzeugung der Kraftwerke der allg. Versorgung ...	Mill. kWh	737,8	1 243,2	1 384,7	1 616,4	1 131,7	1 002,4	791,3	599,4	...

**Handwerk (Messzahlen)<sup>5</sup>**

Beschäftigte (Index) <sup>6</sup> (30.09.2020 = 100) .....	Messzahl	96,3	.	95,7	.	.	95,3p	.	.	...
Umsatz <sup>7</sup> (VjD 2020 = 100) (ohne Umsatzsteuer) .....	Messzahl	99,3	.	134,1	.	.	98,4p	.	.	...

**Bautätigkeit und Wohnungswesen****Baugenehmigungen<sup>8</sup>**

Wohngebäude <sup>9</sup> (nur Neu- und Wiederaufbau) .....	Anzahl	1 562	1 218	1 062	1 020	1 171	...	...	...	...
darunter mit 1 oder 2 Wohnungen .....	Anzahl	1 341	870	840	839	972	...	...	...	...
Umbauter Raum .....	1 000 m <sup>3</sup>	2 245	2 440	1 829	1 742	1 941	...	...	...	...
Veranschlagte Baukosten .....	Mill. Euro	1 064	1 303	952	964	1 039	...	...	...	...
Wohnfläche .....	1 000 m <sup>2</sup>	391	437	315	308	337	...	...	...	...
Nichtwohngebäude (nur Neu- und Wiederaufbau) .....	Anzahl	475	543	432	450	464	...	...	...	...
Umbauter Raum .....	1 000 m <sup>3</sup>	3 218	3 748	3 695	2 444	4 547	...	...	...	...
Veranschlagte Baukosten .....	Mill. Euro	699	924	1 137	710	949	...	...	...	...
Nutzfläche .....	1 000 m <sup>2</sup>	434	525	497	378	5 832	...	...	...	...
Wohnungen insgesamt (alle Baumaßnahmen) .....	Anzahl	4 235	5 823	4 118	3 770	4 127	...	...	...	...
Wohnräume <sup>10</sup> insgesamt (alle Baumaßnahmen) .....	Anzahl	16 431	18 406	14 324	13 704	14 781	...	...	...	...

<sup>1</sup> Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). Volumenindex.<sup>2</sup> Nur auftragseingangsmeldepflichtige Wirtschaftsklassen.<sup>3</sup> Bau von Gebäuden, Tiefbau, Abbrucharbeiten und vorbereitende Baustellenarbeiten u. a.; Betriebe von rechtlichen Einheiten mit 20 oder mehr tätigen Personen.<sup>4</sup> Bauinstallation und sonstiger Ausbau. Ab Berichtsjahr 2021: Betriebe von rechtlichen Einheiten mit 20 oder mehr Personen

(Von Berichtsjahr 2018 bis einschließlich Berichtsjahr 2020: Betriebe von rechtlichen Einheiten mit 23 oder mehr tätigen Personen). Vierteljahresergebnisse (März=1, Juni=2, September=3, Dezember=4).

<sup>5</sup> Zulassungspflichtiges Handwerk laut Anlage A der Handwerksordnung.<sup>6</sup> Am Ende des Kalendervierteljahres.<sup>7</sup> Vierteljahresergebnisse (März=1, Juni=2, September=3, Dezember=4).<sup>8</sup> Die Monatsergebnisse sind vorläufig, da diese keine Tekturen (nachträgliche Baugenehmigungsänderungen) enthalten.<sup>9</sup> Einschließlich Wohnheime.<sup>10</sup> Wohnräume mit jeweils mindestens 6 m<sup>2</sup> Wohnfläche sowie abgeschlossene Küchen.

	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni

## Handel und Gastgewerbe

### Außenhandel

<b>Einfuhr insgesamt (Generalhandel)<sup>1,2</sup></b> .....	Mill. Euro	20 922,0	21 350,8	17 379,6	19 016,9	19 392,2	19 715,3	20 235,0	18 300,6	...
darunter Güter der Ernährungswirtschaft .....	Mill. Euro	1 128,5	1 128,1	1 023,9	1 130,7	1 074,4	1 096,3	1 148,2	1 110,4	...
Güter der gewerblichen Wirtschaft .....	Mill. Euro	18 139,0	18 495,1	14 925,5	15 986,1	16 529,1	16 646,2	16 969,0	15 651,5	...
davon Rohstoffe .....	Mill. Euro	941,2	640,6	755,7	625,0	773,2	902,7	803,4	907,5	...
Halbwaren .....	Mill. Euro	614,9	594,3	450,4	514,9	559,9	540,3	566,5	496,3	...
Fertigwaren .....	Mill. Euro	16 582,9	17 260,2	13 719,4	14 846,3	15 196,0	15 203,2	15 599,2	14 247,7	...
davon Vorerzeugnisse .....	Mill. Euro	1 183,4	1 038,7	760,3	1 033,5	1 072,4	1 033,1	1 096,7	1 023,2	...
Enderzeugnisse .....	Mill. Euro	15 399,4	16 221,6	12 959,1	13 812,7	14 123,7	14 170,2	14 502,5	13 224,5	...
darunter aus <sup>3</sup>										
Europa .....	Mill. Euro	13 508,2	13 619,5	10 766,6	12 200,6	12 605,6	12 594,9	13 092,5	11 286,1	...
darunter aus EU-Ländern <sup>4</sup> insgesamt .....	Mill. Euro	11 569,9	11 787,1	9 312,0	10 549,0	10 846,8	10 984,1	11 364,0	9 828,7	...
darunter aus Belgien .....	Mill. Euro	453,9	479,5	361,4	409,2	399,0	433,6	397,1	341,2	...
Bulgarien .....	Mill. Euro	112,0	123,6	92,7	115,7	119,5	114,0	119,6	106,6	...
Dänemark .....	Mill. Euro	90,3	91,4	73,0	82,4	82,1	86,1	90,4	71,6	...
Finnland .....	Mill. Euro	64,1	62,4	44,2	69,5	63,0	63,9	73,0	59,2	...
Frankreich .....	Mill. Euro	710,1	707,8	692,5	687,1	683,7	667,8	744,7	608,1	...
Griechenland .....	Mill. Euro	65,5	49,1	56,8	57,6	60,2	73,1	80,8	62,2	...
Irland .....	Mill. Euro	205,5	163,9	154,7	150,5	209,2	241,1	190,3	227,3	...
Italien .....	Mill. Euro	1 180,7	1 281,9	1 033,9	1 189,9	1 206,1	1 243,0	1 233,1	1 139,5	...
Luxemburg .....	Mill. Euro	22,2	20,6	21,0	20,5	20,9	21,6	23,1	18,8	...
Niederlande .....	Mill. Euro	1 154,4	979,7	838,0	855,8	791,3	782,9	867,0	773,7	...
Österreich .....	Mill. Euro	1 677,0	1 711,6	1 261,6	1 478,8	1 581,8	1 606,6	1 715,3	1 439,8	...
Polen .....	Mill. Euro	1 217,4	1 314,7	1 077,2	1 168,2	1 143,8	1 216,3	1 194,4	1 078,9	...
Portugal .....	Mill. Euro	168,0	156,3	114,2	151,3	153,7	155,2	147,7	142,4	...
Rumänien .....	Mill. Euro	420,5	444,3	320,3	380,7	416,3	420,9	433,1	338,0	...
Schweden .....	Mill. Euro	164,1	175,5	129,5	154,0	157,9	145,1	147,6	137,0	...
Slowakei .....	Mill. Euro	584,1	547,0	409,3	449,6	474,9	457,8	458,3	426,0	...
Slowenien .....	Mill. Euro	142,3	154,5	112,6	144,1	144,4	146,0	154,1	125,7	...
Spanien .....	Mill. Euro	412,5	407,6	323,2	433,9	452,1	458,3	476,1	432,0	...
Tschechien .....	Mill. Euro	1 448,8	1 628,5	1 184,1	1 453,0	1 470,3	1 468,2	1 528,7	1 289,9	...
Ungarn .....	Mill. Euro	1 145,4	1 141,0	897,0	953,6	1 079,3	1 042,3	1 141,3	874,5	...
Vereinigtes Königreich .....	Mill. Euro	520,2	508,8	419,4	447,6	491,5	328,8	417,5	356,7	...
Russische Föderation .....	Mill. Euro	33,8	33,7	26,4	31,9	29,1	23,6	10,3	8,9	...
Afrika .....	Mill. Euro	529,2	476,2	486,3	515,4	400,9	521,8	394,6	562,3	...
darunter aus Südafrika .....	Mill. Euro	113,6	144,3	148,9	110,0	102,5	123,0	68,5	152,6	...
Amerika .....	Mill. Euro	1 418,0	1 540,0	1 278,5	1 449,7	1 345,0	1 452,2	1 408,1	1 251,9	...
darunter aus den USA .....	Mill. Euro	1 086,7	1 119,1	985,6	1 088,2	1 050,4	1 076,5	1 078,8	976,2	...
Asien .....	Mill. Euro	5 426,2	5 676,3	4 814,5	4 813,8	5 007,7	5 112,0	5 300,6	5 166,0	...
darunter aus der Volksrepublik China .....	Mill. Euro	2 884,0	3 254,1	2 606,4	2 508,6	2 537,7	2 548,3	2 661,1	2 548,0	...
Japan .....	Mill. Euro	307,9	335,4	283,2	322,9	287,6	275,5	310,6	270,0	...
Australien, Ozeanien und übrige Gebiete .....	Mill. Euro	40,4	38,8	33,8	37,4	32,9	34,5	39,2	34,3	...
<b>Ausfuhr insgesamt (Spezialhandel)<sup>1,2</sup></b> .....	Mill. Euro	19 202,9	21 497,8	17 205,7	18 919,4	20 391,4	20 465,5	21 084,4	18 089,1	...
darunter Güter der Ernährungswirtschaft .....	Mill. Euro	1 057,7	991,4	880,5	1 081,2	994,8	1 019,9	1 054,6	1 033,0	...
Güter der gewerblichen Wirtschaft .....	Mill. Euro	17 459,7	19 782,9	15 754,4	16 604,4	18 102,8	18 237,4	18 769,4	15 988,7	...
davon Rohstoffe .....	Mill. Euro	144,1	126,7	110,0	117,9	112,1	123,4	132,4	119,4	...
Halbwaren .....	Mill. Euro	766,6	777,4	591,1	739,0	767,6	762,9	810,6	767,3	...
Fertigwaren .....	Mill. Euro	16 549,0	18 878,8	15 053,3	15 747,6	17 223,0	17 351,0	17 826,4	15 102,0	...
davon Vorerzeugnisse .....	Mill. Euro	1 236,3	1 154,1	885,0	1 213,1	1 211,8	1 216,8	1 266,8	1 205,5	...
Enderzeugnisse .....	Mill. Euro	15 312,7	17 724,7	14 168,3	14 534,4	16 011,2	16 134,2	16 559,6	13 896,4	...
davon nach										
Europa .....	Mill. Euro	12 368,9	13 970,3	10 899,3	12 577,7	12 747,8	13 079,1	13 344,8	12 007,6	...
darunter in EU-Länder <sup>4</sup> insgesamt .....	Mill. Euro	10 090,9	11 311,2	8 913,6	10 372,7	10 250,8	10 524,3	10 835,6	9 765,5	...
darunter nach Belgien .....	Mill. Euro	691,9	651,9	592,1	643,3	632,3	613,8	652,2	575,4	...
Bulgarien .....	Mill. Euro	73,0	83,5	66,4	74,8	80,9	85,9	86,3	70,4	...
Dänemark .....	Mill. Euro	182,6	230,9	165,3	194,5	154,1	179,0	199,0	164,6	...
Finnland .....	Mill. Euro	138,6	169,6	121,2	135,2	124,8	130,5	136,3	114,7	...
Frankreich .....	Mill. Euro	1 133,8	1 436,5	1 218,2	1 332,7	1 360,6	1 321,5	1 389,9	1 159,1	...
Griechenland .....	Mill. Euro	75,7	87,0	69,2	77,5	106,6	94,4	105,2	88,8	...

1 Vorläufige Ergebnisse.

2 Nachweis einschließlich „nicht aufgliederbares Intrahandelsresultat“.

3 Ohne Schiffs- und Luftfahrzeugbedarf, Polargebiete und nicht ermittelte Länder und Gebiete.

4 Januar 2020: EU 28. Ab Februar 2020 EU 27 (ohne Vereinigtes Königreich).



noch: Handel und Gastgewerbe	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni
Irland .....	Mill. Euro	93,6	91,7	116,1	111,0	130,3	98,9	102,4	104,4	...
Italien .....	Mill. Euro	1 317,6	1 361,1	1 058,6	1 255,0	1 313,3	1 359,7	1 337,1	1 209,0	...
Luxemburg .....	Mill. Euro	53,1	55,2	51,5	60,8	71,1	63,8	80,2	46,4	...
Niederlande .....	Mill. Euro	973,9	1 243,6	985,3	930,0	765,3	776,4	817,8	752,7	...
Österreich .....	Mill. Euro	1 476,5	1 603,5	1 253,5	1 497,4	1 543,2	1 586,4	1 626,8	1 474,9	...
Polen .....	Mill. Euro	869,4	1 038,8	761,3	1 010,2	986,5	1 112,4	1 061,5	979,2	...
Portugal .....	Mill. Euro	123,6	135,6	94,2	130,4	148,6	154,7	149,0	129,0	...
Rumänien .....	Mill. Euro	283,4	317,6	230,7	300,2	287,6	307,2	302,2	292,2	...
Schweden .....	Mill. Euro	316,5	338,6	252,8	303,3	292,3	289,7	296,7	271,5	...
Slowakei .....	Mill. Euro	247,1	238,2	162,5	251,0	242,1	246,6	262,3	239,8	...
Slowenien .....	Mill. Euro	92,4	102,4	82,7	98,3	97,0	118,3	104,4	98,7	...
Spanien .....	Mill. Euro	646,6	758,6	574,4	620,5	626,0	625,7	650,7	614,2	...
Tschechien .....	Mill. Euro	664,4	675,5	519,8	647,7	643,4	672,2	702,9	665,2	...
Ungarn .....	Mill. Euro	412,3	463,2	344,1	487,3	427,1	458,5	534,7	489,5	...
Vereinigtes Königreich .....	Mill. Euro	981,3	1 164,6	827,8	948,6	1 048,2	1 139,6	1 146,4	916,7	...
Russische Föderation .....	Mill. Euro	66,3	77,5	43,5	65,4	76,7	61,9	65,0	49,6	...
Afrika .....	Mill. Euro	217,5	269,6	252,0	220,5	236,2	224,2	251,3	240,5	...
darunter nach Südafrika .....	Mill. Euro	79,4	66,8	66,6	69,0	73,1	78,2	79,7	70,4	...
Amerika .....	Mill. Euro	2 934,8	3 510,6	2 647,5	2 765,9	3 550,6	3 113,5	3 631,7	2 522,8	...
darunter in die USA .....	Mill. Euro	2 290,5	2 852,3	2 035,5	2 172,7	2 894,0	2 408,7	2 824,4	1 904,3	...
Asien .....	Mill. Euro	3 461,2	3 511,2	3 189,1	3 149,2	3 536,6	3 832,0	3 636,2	3 131,3	...
darunter in die Volksrepublik China .....	Mill. Euro	1 597,5	1 513,7	1 287,8	1 379,6	1 538,3	1 667,8	1 561,3	1 324,3	...
nach Japan .....	Mill. Euro	244,7	261,2	261,8	241,6	286,3	260,1	243,7	216,2	...
Australien, Ozeanien und übrige Gebiete .....	Mill. Euro	220,5	236,1	217,7	206,0	320,3	216,7	220,4	186,9	...
<b>Großhandel (2015 = 100)<sup>1</sup></b>										
Index der Großhandelsumsätze nominal .....	Messzahl	157,2	171,2	164,0	155,9	160,2	...	...	...	...
Index der Großhandelsumsätze real .....	Messzahl	123,0	137,1	134,5	126,8	128,1	...	...	...	...
Index der Beschäftigten im Großhandel .....	Messzahl	106,5	107,4	107,0	105,8	106,3	...	...	...	...
<b>Einzelhandel (2015 = 100)<sup>2</sup></b>										
Index der Einzelhandelsumsätze nominal .....	Messzahl	152,8	171,5	173,9	140,7	138,2	152,0	...	...	...
Einzelhandel mit Waren verschiedener Art <sup>3</sup> .....	Messzahl	142,1	143,0	160,5	131,3	131,6	148,1	...	...	...
Facheinzelhandel mit Nahrungsmitteln, Getränken und Tabakwaren <sup>3</sup> .....	Messzahl	131,4	133,5	149,0	120,1	125,6	137,6	...	...	...
Apotheken; Facheinzelhandel mit medizinischen, orthopädischen und kosmetischen Artikeln <sup>3</sup> .....	Messzahl	157,7	159,5	167,8	152,5	153,6	156,1	...	...	...
Sonstiger Facheinzelhandel <sup>3</sup> .....	Messzahl	135,8	141,3	150,4	121,5	125,2	136,2	...	...	...
Einzelhandel (nicht in Verkaufsräumen) .....	Messzahl	225,7	306,0	278,7	225,3	203,8	216,8	...	...	...
Index der Einzelhandelsumsätze real .....	Messzahl	127,4	142,7	144,1	116,7	113,9	124,6	...	...	...
Index der Beschäftigten im Einzelhandel .....	Messzahl	106,8	107,1	106,9	105,8	105,5	105,2	...	...	...
<b>Kfz-Handel (2015 = 100)<sup>4</sup></b>										
Index der Umsätze im Kfz-Handel nominal .....	Messzahl	138,0	166,8	140,8	133,4	154,0	...	...	...	...
Index der Umsätze im Kfz-Handel real .....	Messzahl	109,4	126,7	105,9	100,8	115,9	...	...	...	...
Index der Beschäftigten im Kfz-Handel .....	Messzahl	103,8	107,3	106,8	106,8	106,9	...	...	...	...
<b>Gastgewerbe (2015 = 100)</b>										
Index der Gastgewerbeumsätze nominal .....	Messzahl	115,5	116,2	129,2	103,7	110,3	120,4	...	...	...
Hotels, Gasthöfe, Pensionen und Hotels garnis .....	Messzahl	113,3	119,7	128,6	103,5	114,5	120,8	...	...	...
Sonstiges Beherbergungsgewerbe .....	Messzahl	246,6	242,9	217,0	163,5	187,4	211,5	...	...	...
Restaurants, Cafés, Eisdielen und Imbisshallen .....	Messzahl	119,2	116,0	135,0	105,0	110,3	124,0	...	...	...
Sonstiges Gaststättengewerbe .....	Messzahl	116,8	113,2	132,9	102,7	107,6	121,4	...	...	...
Kantinen und Caterer .....	Messzahl	132,1	137,9	133,1	120,2	124,3	130,2	...	...	...
Index der Gastgewerbeumsätze real .....	Messzahl	87,8	85,8	95,6	77,7	82,3	89,0	...	...	...
Index der Beschäftigten im Gastgewerbe .....	Messzahl	97,2	103,2	103,2	99,6	100,6	102,4	...	...	...
<b>Tourismus<sup>5</sup></b>										
Gästeankünfte .....	1 000	3 590	2 461	2 591	2 091	2 444	2 824	2 946	3 758	...
darunter Auslands Gäste .....	1 000	742	533	678	465	590	583	660	792	...
Gästeübernachtungen .....	1 000	9 051	6 084	6 375	5 581	6 370	1 085	7 235	9 781	...
darunter Auslands Gäste .....	1 000	1 669	1 256	1 496	1 058	1 303	1 304	1 441	1 761	...

1 Einschließlich Handelsvermittlung.

2 Einschließlich Tankstellen.

3 In Verkaufsräumen.

4 Sowie Instandhaltung und Reparatur von Kfz. Ohne Tankstellen.

5 Beherbergungsbetriebe mit zehn oder mehr Gästebetten (einschl. Campingplätze mit zehn oder mehr Stellplätzen).

	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni

## Verkehr

### Straßenverkehr

Zulassung fabrikneuer Kraftfahrzeuge insgesamt <sup>1</sup> .....	Anzahl	70 699	56 593	54 290	45 078	53 840	71 984	69 097	63 794	77 538
darunter Krafträder <sup>2</sup> .....	Anzahl	6 135	1 204	825	1 747	4 494	8 463	6 924	5 539	5 094
Personenkraftwagen und sonst. „M1“-Fahrzeuge ..	Anzahl	57 729	48 760	48 058	37 666	42 833	55 595	54 563	52 027	61 526
Lastkraftwagen .....	Anzahl	4 438	4 946	3 967	3 831	4 470	5 561	5 173	4 150	8 637
Zugmaschinen .....	Anzahl	1 874	1 283	1 057	1 487	1 634	1 999	1 984	1 718	1 785
sonstige Kraftfahrzeuge .....	Anzahl	425	289	237	209	288	298	337	263	386
Beförderte Personen im Schienennah- und gewerblichen Omnibuslinienverkehr insg. (Quartalsergebnisse) <sup>3</sup> .....	1 000	302 744	.	320 871	.	.	327 459	.	.	...
davon öffentliche und gemischtwirtschaftliche Unternehmen ..	1 000	250 248	.	264 372	.	.	266 961	.	.	...
private Unternehmen .....	1 000	52 495	.	56 500	.	.	60 498	.	.	...
Straßenverkehrsunfälle insgesamt <sup>4</sup> .....	Anzahl	35 432	33 755	33 710	29 667	25 883	29 447	33 384	30 585	...
davon Unfälle mit Personenschaden .....	Anzahl	4 954	3 553	2 854	2 812	2 406	3 362	4 175	3 998	...
mit nur Sachschaden .....	Anzahl	30 478	30 202	30 856	26 855	23 477	26 085	29 209	26 587	...
Getötete Personen <sup>5</sup> .....	Anzahl	39	32	38	29	22	31	47	36	...
Verletzte Personen .....	Anzahl	6 121	4 608	3 770	3 793	3 015	4 321	5 308	4 868	...

### Luftverkehr Fluggäste

Flughafen München   Ankunft .....	1 000	1 609	1 527	1 179	1 260	1 246	1 466	1 759	1 893	...
Abgang .....	1 000	1 662	1 428	1 259	1 180	1 269	1 527	1 692	1 920	...
Flughafen Nürnberg   Ankunft .....	1 000	169	124	103	104	93	113	153	185	...
Abgang .....	1 000	192	104	118	84	103	130	149	206	...
Flughafen Memmingen   Ankunft .....	1 000	125	111	103	111	93	106	134	146	...
Abgang .....	1 000	130	97	117	95	96	113	133	153	...

### Eisenbahnverkehr <sup>6</sup>

Güterempfang .....	1 000 t	2 239	2 643	2 008	2 039	2 250	2 344	2 429	...	...
Güterversand .....	1 000 t	1 616	2 090	1 678	1 803	2 129	2 185	2 348	...	...

### Binnenschifffahrt <sup>7</sup>

Güterempfang insgesamt .....	1 000 t	250	214	190	200	198	302	164	...	...
davon auf dem Main .....	1 000 t	144	103	80	56	79	122	77	...	...
auf der Donau .....	1 000 t	105	111	110	143	119	180	87	...	...
Güterversand insgesamt .....	1 000 t	187	234	132	131	210	282	117	...	...
davon auf dem Main .....	1 000 t	112	154	75	68	131	151	78	...	...
auf der Donau .....	1 000 t	74	81	57	63	80	132	39	...	...

## Geld und Kredit

### Kredite und Einlagen <sup>8,9</sup>

Kredite an Nichtbanken insgesamt .....	Mill. Euro	694 102	.	703 224	.	.	697 923	.	.	...
darunter Kredite an inländische Nichtbanken <sup>10</sup> .....	Mill. Euro	591 397	.	596 612	.	.	587 648	.	.	...
davon kurzfr. Kredite an Nichtbanken insgesamt .....	Mill. Euro	73 088	.	69 226	.	.	74 094	.	.	...
Unternehmen und Privatpersonen <sup>11</sup> ..	Mill. Euro	70 182	.	66 634	.	.	70 959	.	.	...
inländ. öffentliche Haushalte <sup>12</sup> ...	Mill. Euro	2 906	.	2 592	.	.	3 135	.	.	...
mittelfr. Kredite an Nichtbanken insgesamt <sup>13</sup> .....	Mill. Euro	89 678	.	96 455	.	.	96 119	.	.	...
Unternehmen u. Privatpersonen <sup>11</sup> ...	Mill. Euro	88 335	.	94 635	.	.	94 206	.	.	...
inländ. öffentliche Haushalte <sup>12</sup> ...	Mill. Euro	1 343	.	1 820	.	.	1 913	.	.	...
langfr. Kredite an Nichtbanken insgesamt <sup>14</sup> .....	Mill. Euro	531 336	.	537 543	.	.	527 710	.	.	...
Unternehmen u. Privatpersonen <sup>11</sup> ...	Mill. Euro	505 853	.	511 008	.	.	500 998	.	.	...
inländ. öffentliche Haushalte <sup>12</sup> ...	Mill. Euro	25 483	.	26 535	.	.	26 712	.	.	...

1 Daten des Kraftfahrt-Bundesamtes.

2 Einschließlich Leichtkrafträder, dreirädrige und leichte vierrädrige Kraftfahrzeugen.

3 Vorläufige Ergebnisse.

4 Soweit durch die Polizei erfasst, Vorläufige Ergebnisse.

5 Einschließlich der innerhalb 30 Tagen an den Unfallfolgen verstorbenen Personen.

6 Berücksichtigung der Nachkorrekturen erst zum Berichtsjahresende.

7 Schiffsgüterumschläge an den Häfen des Main-Donau-Kanals werden dem Donaugebiet zugeordnet.

8 Aus Veröffentlichungen der Deutschen Bundesbank Frankfurt am Main – Quartalsergebnisse der in Bayern tätigen Kreditinstitute (einschließlich Bausparkassen).

9 Stand am Monatsende.

10 Ohne Treuhandkredite.

11 Einschl. Kredite (Einlagen) an ausländische Nichtbanken.

12 Ohne Kredite (Einlagen) an ausländische öffentliche Haushalte.

13 Laufzeiten von über 1 Jahr bis 5 Jahre.

14 Laufzeiten über 5 Jahre.

noch: Geld und Kredit	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni
Einlagen von Nichtbanken insgesamt <sup>1</sup> (Monatsende) ....	Mill. Euro	780 537	.	767 910	.	.	761 370	.	.	...
davon Sicht- und Termineinlagen <sup>2</sup> .....	Mill. Euro	687 919	.	688 972	.	.	685 802	.	.	...
davon von Unternehmen und Privatpersonen .....	Mill. Euro	638 823	.	646 093	.	.	641 627	.	.	...
von öffentlichen Haushalten .....	Mill. Euro	49 096	.	42 879	.	.	44 175	.	.	...
Spareinlagen .....	Mill. Euro	92 618	.	78 938	.	.	75 568	.	.	...
darunter bei Sparkassen .....	Mill. Euro	32 305	.	27 099	.	.	25 734	.	.	...
bei Kreditbanken .....	Mill. Euro	19 724	.	16 775	.	.	16 313	.	.	...

<b>Zahlungsschwierigkeiten</b>										
Insolvenzen insgesamt .....	Anzahl	1 031	1 046	1 011	1 065	1 137	1 127	1 044	1 146	...
darunter mangels Masse abgelehnt .....	Anzahl	89	97	93	87	92	109	99	105	...
davon Unternehmen .....	Anzahl	245	213	252	215	256	248	260	283	...
darunter mangels Masse abgelehnt .....	Anzahl	72	67	61	59	67	81	76	73	...
Verbraucher .....	Anzahl	483	508	462	513	540	543	491	518	...
darunter mangels Masse abgelehnt .....	Anzahl	0	1	2	1	2	0	1	2	...
ehemals selbstständig Tätige .....	Anzahl	256	273	233	292	288	298	247	283	...
darunter mangels Masse abgelehnt .....	Anzahl	10	20	18	15	14	21	14	12	...
sonstige natürliche Personen, Nachlässe .....	Anzahl	47	52	64	45	53	38	46	62	...
darunter mangels Masse abgelehnt .....	Anzahl	7	9	12	12	9	7	8	18	...
Voraussichtliche Forderungen insgesamt .....	1 000 Euro	271 185	388 085	812 336	693 987	493 162	381 408	530 750	346 939	...
davon Unternehmen .....	1 000 Euro	202 534	309 407	741 301	606 322	413 980	290 472	394 715	238 712	...
Verbraucher .....	1 000 Euro	22 846	28 793	24 553	27 983	33 753	31 644	26 244	28 823	...
ehemals selbstständig Tätige .....	1 000 Euro	43 046	41 258	41 439	49 550	37 711	56 780	105 260	71 253	...
sonstige natürliche Personen, Nachlässe .....	1 000 Euro	2 759	8 626	5 042	10 132	7 717	2 512	4 530	8 151	...

## Verdienste

Bruttomonatsverdienste <sup>3</sup> der vollzeitbeschäftigten Arbeitnehmer in der Land- und Forstwirtschaft, Fischerei, im Produzierenden Gewerbe und im Dienstleistungsbereich ....	Euro	4 502	.	.	.	.	.	...	.	.
Männer .....	Euro	4 769	.	.	.	.	.	...	.	.
Frauen .....	Euro	3 905	.	.	.	.	.	...	.	.
Anforderungsniveau 1 <sup>4</sup> .....	Euro	2 825	.	.	.	.	.	...	.	.
Anforderungsniveau 2 <sup>4</sup> .....	Euro	3 676	.	.	.	.	.	...	.	.
Anforderungsniveau 3 <sup>4</sup> .....	Euro	5 174	.	.	.	.	.	...	.	.
Anforderungsniveau 4 <sup>4</sup> .....	Euro	6 974	.	.	.	.	.	...	.	.
Land- und Forstwirtschaft, Fischerei .....	Euro	3 013	.	.	.	.	.	...	.	.
Produzierendes Gewerbe .....	Euro	4 526	.	.	.	.	.	...	.	.
Bergbau und Gewinnung von Steinen und Erden .....	Euro	3 852	.	.	.	.	.	...	.	.
Verarbeitendes Gewerbe .....	Euro	4 722	.	.	.	.	.	...	.	.
Energieversorgung .....	Euro	4 983	.	.	.	.	.	...	.	.
Wasserversorgung; Abwasser- und Abfallentsorgung und Beseitigung von Umweltverschmutzungen .....	Euro	3 697	.	.	.	.	.	...	.	.
Baugewerbe .....	Euro	3 826	.	.	.	.	.	...	.	.
Dienstleistungsbereich .....	Euro	4 500	.	.	.	.	.	...	.	.
Handel; Instandhaltung u. Reparatur von Kraftfahrzeugen...	Euro	4 191	.	.	.	.	.	...	.	.
Verkehr und Lagerei .....	Euro	3 654	.	.	.	.	.	...	.	.
Gastgewerbe .....	Euro	2 908	.	.	.	.	.	...	.	.
Information und Kommunikation .....	Euro	6 282	.	.	.	.	.	...	.	.
Erbringung von Finanz- und Versicherungs- dienstleistungen .....	Euro	5 956	.	.	.	.	.	...	.	.
Grundstücks- und Wohnungswesen .....	Euro	(5 158)	.	.	.	.	.	...	.	.
Erbringung von freiberuflichen, wissenschaftlichen und technischen Dienstleistungen .....	Euro	5 733	.	.	.	.	.	...	.	.
Erbringung von sonstigen wirtschaftlichen Dienst- leistungen .....	Euro	3 653	.	.	.	.	.	...	.	.
Öffentliche Verwaltung, Verteidigung, Sozialversicherung ...	Euro	4 140	.	.	.	.	.	...	.	.
Erziehung und Unterricht .....	Euro	4 650	.	.	.	.	.	...	.	.
Gesundheits- und Sozialwesen .....	Euro	4 174	.	.	.	.	.	...	.	.
Kunst, Unterhaltung und Erholung .....	Euro	(4 256)	.	.	.	.	.	...	.	.
Erbringung von sonstigen Dienstleistungen .....	Euro	3 741	.	.	.	.	.	...	.	.

1 Ohne Verbindlichkeiten gegenüber Geldmarktfonds und ohne Einlagen aus Treuhandkrediten.

2 Einschließlich Sparbriefe.

3 Berichtsmonat April, ohne Sonderzahlungen.

4 Anforderungsniveau 1: Helfer; Anforderungsniveau 2: Fachkraft; Anforderungsniveau 3: Spezialist; Anforderungsniveau 4: Experte.

	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni

## Landwirtschaft

### Schlachtungen<sup>1</sup>

Gewerbl. Schlachtungen und Hausschl. (ohne Geflügel) .....	1 000	400,9	422,8	398,2	398,5	366,7	402,6	395,2	366,7	...
darunter Rinder .....	1 000	64,0	77,2	63,1	67,0	65,2	65,6	65,5	59,5	...
darunter Kälber <sup>2</sup> .....	1 000	1,0	1,2	1,5	0,9	0,8	1,3	0,8	0,8	...
Jungrinder <sup>3</sup> .....	1 000	0,3	0,4	0,3	0,2	0,2	0,2	0,2	0,2	...
Schweine .....	1 000	327,8	335,3	323,4	324,6	292,4	320,8	320,3	287,6	...
Schafe .....	1 000	8,2	9,4	10,9	6,4	8,5	15,2	8,6	9,4	...
darunter gewerbliche Schlachtungen (ohne Geflügel) ....	1 000	399,9	420,1	395,9	396,8	363,5	400,9	394,2	365,1	...
darunter Rinder .....	1 000	63,7	76,6	62,6	66,6	63,6	65,2	65,2	59,3	...
darunter Kälber <sup>2</sup> .....	1 000	0,9	1,1	1,4	0,9	0,8	1,2	0,7	0,7	...
Jungrinder <sup>3</sup> .....	1 000	0,3	0,3	0,2	0,2	0,2	0,2	0,2	0,2	...
Schweine .....	1 000	327,4	333,9	322,1	323,4	291,0	319,9	319,7	287,3	...
Schafe .....	1 000	7,9	8,7	10,4	6,2	8,3	14,8	8,4	9,1	...

### Durchschnittliches Schlachtgewicht<sup>4</sup>

Rinder .....	kg	365,8	357,5	362,2	315,8	308,7	315,8	307,1	302,0	...
darunter Kälber <sup>2</sup> .....	kg	45,6	124,4	64,8	155,0	155,0	154,5	156,6	159,2	...
Jungrinder <sup>3</sup> .....	kg	186,8	201,3	164,9	182,1	177,0	175,4	165,3	198,1	...
Schweine .....	kg	98,1	99,1	97,5	97,1	96,9	96,9	97,4	97,2	...

### Gesamtschlachtgewicht<sup>5</sup>

Gewerbl. Schlachtungen und Hausschl. (ohne Geflügel)	1 000 t	55,3	60,6	54,1	55,5	51,9	54,9	55,0	51,8	...
darunter Rinder .....	1 000 t	23,0	27,2	22,3	24,1	23,6	23,8	23,9	21,7	...
darunter Kälber <sup>2</sup> .....	1 000 t	0,0	0,1	0,1	0,1	0,1	0,2	0,1	0,1	...
Jungrinder <sup>3</sup> .....	1 000 t	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	...
Schweine .....	1 000 t	32,2	33,2	31,5	31,2	28,1	30,8	30,9	27,6	...
Schafe .....	1 000 t	0,2	0,2	0,2	0,1	0,2	0,3	0,2	0,2	...
darunter gewerbliche Schlachtungen (ohne Geflügel) ....	1 000 t	55,2	60,3	53,8	55,2	51,6	54,7	54,8	51,6	...
darunter Rinder .....	1 000 t	22,9	27,0	22,2	24,0	23,4	23,7	23,8	21,6	...
darunter Kälber <sup>2</sup> .....	1 000 t	0,0	0,1	0,1	0,1	0,1	0,2	0,1	0,1	...
Jungrinder <sup>3</sup> .....	1 000 t	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	...
Schweine .....	1 000 t	32,1	33,1	31,4	31,1	28,0	30,7	30,8	27,6	...
Schafe .....	1 000 t	0,1	0,2	0,2	0,1	0,2	0,3	0,2	0,2	...

### Geflügel

Hennenhaltungsplätze <sup>6</sup> .....	1 000	4 906	4 896	4 892	4 939	4 955	4 964	4 955	4 955	...
Legehennenbestand <sup>6</sup> .....	1 000	3 927	3 982	3 974	3 966	4 152	4 159	4 065	4 007	...
Konsumeier <sup>6</sup> .....	1 000	95 391	97 125	99 247	99 725	100 139	107 996	100 068	98 570	...
Geflügelfleisch <sup>7</sup> .....	1 000 t	16,1	15,6	15,9	15,8	14,6	15,6	16,6	15,7	...

### Getreideanlieferungen<sup>8,9</sup>

Roggen und Wintermenggetreide .....	1 000 t	1,2	1,5	1,1	0,9	1,6	1,2	1,4	...	...
Weizen .....	1 000 t	8,5	15,3	12,3	11,9	21,1	13,4	14,7	...	...
Gerste .....	1 000 t	5,9	9,3	5,5	3,9	8,0	6,2	5,9	...	...
Hafer und Sommermenggetreide .....	1 000 t	0,6	0,5	0,2	0,2	0,1	0,3	0,1	...	...

### Vermahlung von Getreide<sup>8,9</sup>

Getreide insgesamt .....	1 000 t	92,6	111,8	102,4	106,8	112,6	111,5	106,8	...	...
darunter Roggen und -gemenge .....	1 000 t	7,9	8,6	8,7	9,0	9,4	9,0	9,1	...	...
Weizen und -gemenge .....	1 000 t	84,7	103,3	93,7r	97,8	103,2	102,5	97,6	...	...

### Vorräte in zweiter Hand<sup>8,9</sup>

Roggen und Wintermenggetreide .....	1 000 t	33,0	46,1	47,4	43,8	42,3	35,9	30,4	...	...
Weizen .....	1 000 t	384,5	652,5	697,7	580,6	598,7	542,4	493,8	...	...
Gerste .....	1 000 t	237,2	343,4	325,8	289,2	278,7	256,6	233,2	...	...
Hafer und Sommermenggetreide .....	1 000 t	22,0	27,3	25,4	14,4	13,2	12,1	13,0	...	...
Mais .....	1 000 t	82,4	73,9	69,6	61,0	57,4	44,6	37,4	...	...

1 Gewerbliche Schlachtungen und Hausschlachtungen von Tieren inländischer und ausländischer Herkunft.

2 Höchstens 8 Monate alt.

3 Kälber über 8, aber höchstens 12 Monate alt.

4 Von gewerblich geschlachteten Tieren inländischer Herkunft.

5 Bzw. Schlachtmenge, einschließlich Schlachtfette, jedoch ohne Innereien.

6 In Betrieben mit einer Haltungskapazität von mindestens 3 000 Legehennen.

7 Alle Geflügelschlachtereien, die nach dem EG-Hygienericht im Besitz einer Zulassung sind.

8 Nach Angaben des Bundesinformationszentrums Landwirtschaft (BZL) in der Bundesanstalt für Landwirtschaft und Ernährung.

9 Anlieferung vom Erzeuger an Handel, Genossenschaften, Mühlen und sonstige Verarbeitungsbetriebe.

	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni

## Bierabsatz

Bierabsatz insgesamt .....	1 000 hl	2 249r	1 694	1 619	1 614	1 718	1 824	2 235	2 297	...
davon Bier der Steuerklassen bis 10 .....	1 000 hl	182	103	98	112	123	144	208	214	...
11 bis 13 .....	1 000 hl	2 039r	1 558	1 497	1 471	1 558	1 646	1 994	2 057	...
14 oder darüber .....	1 000 hl	28	32	24	31	37	34	33	26	...
darunter Ausfuhr zusammen .....	1 000 hl	565r	371	313	370	448	458	564	583	...
davon in EU-Länder .....	1 000 hl	337r	233	206	220	253	271	337	340	...
in Drittländer .....	1 000 hl	228	139	108	150	195	188	227	243	...

## Bevölkerung und Erwerbstätigkeit

Bevölkerungsstand <sup>1</sup> .....	1 000	13 369	13 442	13 435	...	...	...	...	...	...
Natürliche Bevölkerungsbewegung <sup>2</sup>										
Eheschließungen <sup>2</sup> .....	Anzahl	3 284	3 005	...	...	...	...	...	...	...
je 10 000 Einwohner .....	Anzahl	2,5	2,2	...	...	...	...	...	...	...
Lebendgeborene <sup>3</sup> .....	Anzahl	9 665	8 817	...	...	...	...	...	...	...
je 10 000 Einwohner .....	Anzahl	7,2	6,6	...	...	...	...	...	...	...
Gestorbene <sup>4</sup> .....	Anzahl	12 608	12 757	...	...	...	...	...	...	...
je 10 000 Einwohner .....	Anzahl	9,4	9,5	...	...	...	...	...	...	...
und zwar im 1. Lebensjahr Gestorbene .....	Anzahl	40	36	...	...	...	...	...	...	...
je 1 000 Lebendgeborene .....	Anzahl	4,1	4,1	...	...	...	...	...	...	...
in den ersten 7 Lebenstagen Gestorbene .....	Anzahl	19	18	...	...	...	...	...	...	...
je 1 000 Lebendgeborene .....	Anzahl	2,0	2,0	...	...	...	...	...	...	...
Überschuss										
der Geborenen bzw. der Gestorbenen (–) .....	Anzahl	– 2 943	– 3 940	...	...	...	...	...	...	...
je 10 000 Einwohner .....	Anzahl	– 2,2	– 2,9	...	...	...	...	...	...	...
Totgeborene <sup>3</sup> .....	Anzahl	30	42	...	...	...	...	...	...	...
Wanderungen <sup>2</sup>										
Zuzüge über die Landesgrenze .....	Anzahl	37 491	33 581	...	...	...	...	...	...	...
darunter aus dem Ausland .....	Anzahl	28 171	25 527	...	...	...	...	...	...	...
Fortzüge über die Landesgrenze .....	Anzahl	25 746	24 031	...	...	...	...	...	...	...
darunter in das Ausland .....	Anzahl	16 876	15 973	...	...	...	...	...	...	...
Zuzüge aus den anderen Bundesländern .....	Anzahl	9 320	8 054	...	...	...	...	...	...	...
Fortzüge in die anderen Bundesländer .....	Anzahl	8 870	8 058	...	...	...	...	...	...	...
Wanderungsgewinn bzw. -verlust (–) .....	Anzahl	11 745	9 550	...	...	...	...	...	...	...
Innerhalb des Landes Umgezogene <sup>5</sup> .....	Anzahl	50 444	48 908	...	...	...	...	...	...	...
Arbeitsmarkt <sup>6</sup>										
Sozialversicherungspflichtig Beschäftigte am Arbeitsort .....	1 000	5 899	.	5 947	.	.	...	.	.	...
Frauen .....	1 000	2 722	.	2 748	.	.	...	.	.	...
Ausländer <sup>7</sup> .....	1 000	1 037	.	1 094	.	.	...	.	.	...
Teilzeitbeschäftigte .....	1 000	1 701	.	1 743	.	.	...	.	.	...
darunter Frauen .....	1 000	1 351	.	1 375	.	.	...	.	.	...
nach zusammengefassten Wirtschaftsabschnitten (WZ 2008)										
A Land- und Forstwirtschaft, Fischerei .....	1 000	29	.	29	.	.	...	.	.	...
B–F Produzierendes Gewerbe .....	1 000	1 834	.	1 840	.	.	...	.	.	...
B–E Produzierendes Gewerbe ohne Baugewerbe .....	1 000	1 490	.	1 499	.	.	...	.	.	...
C Verarbeitendes Gewerbe .....	1 000	1 408	.	1 412	.	.	...	.	.	...
F Baugewerbe .....	1 000	344	.	341	.	.	...	.	.	...
G–U Dienstleistungsbereiche .....	1 000	4 036	.	4 078	.	.	...	.	.	...
G–I Handel, Verkehr und Gastgewerbe .....	1 000	1 260	.	1 268	.	.	...	.	.	...
J Information und Kommunikation .....	1 000	269	.	274	.	.	...	.	.	...
K Finanz- und Versicherungsdienstleister .....	1 000	180	.	180	.	.	...	.	.	...
L Grundstücks- und Wohnungswesen .....	1 000	41	.	41	.	.	...	.	.	...
M–N Freiberufliche, wissenschaftliche, technische Dienstleister; sonst. wirtschaftliche Dienstleister .....	1 000	794	.	800	.	.	...	.	.	...
O–Q Öffentliche Verwaltung, Verteidigung, Sozialversicherung; Erziehung und Unterricht; Gesundheit und Sozialwesen .....	1 000	1 311	.	1 333	.	.	...	.	.	...
R–U Kunst, Unterhaltung und Erholung; sonstige Dienstleister; Private Haushalte; Exterritoriale Organisationen und Körperschaften .....	1 000	182	.	183	.	.	...	.	.	...

1 Fortschreibung des Bevölkerungsstandes auf der Basis des Zensus 2011. Die Bevölkerungszahlen ab Mai 2022 werden – voraussichtlich ab Herbst 2024 – auf Basis des Zensus 2022 revidiert.

2 Die Zahlen der natürlichen Bevölkerungsbewegung und der Wanderungen geben den jeweils aktuellen Stand des Monats im noch nicht abgeschlossenen Berichtsjahr wieder. Bis zum Ende des Jahres können Nachmeldungen der Städte und Gemeinden für die einzelnen Monate erfolgen, so dass sich die endgültigen Monatsergebnisse noch ändern können.

3 Nach der Wohngemeinde der Mutter.

4 Ohne Totgeborene; nach der Wohngemeinde der Verstorbenen.

5 Ohne Umzüge innerhalb der Gemeinden.

6 Auswertungen aus der Beschäftigungsstatistik der Bundesagentur für Arbeit. Zahlenwerte vorläufig.

7 Ab März 2021: Einschl. Staatenlose sowie Personen ohne Angabe zur Staatsangehörigkeit.



noch: Bevölkerung und Erwerbstätigkeit	Einheit	Vorjahres- monat	2023		2024					
			November	Dezember	Januar	Februar	März	April	Mai	Juni
Arbeitslose .....	1 000	244,0	254,3	262,9	294,5	294,2	281,8	273,8	269,8	271,1
darunter Frauen .....	1 000	113,8	119,8	119,9	127,7	127,2	124,5	124,6	123,6	124,3
Arbeitslosenquote insgesamt <sup>1</sup> .....	%	3,2	3,3	3,4	3,9	3,8	3,7	3,6	3,5	3,5
Frauen .....	%	3,2	3,3	3,3	3,5	3,5	3,5	3,5	3,4	3,4
Männer .....	%	3,2	3,3	3,5	4,1	4,1	3,9	3,7	3,6	3,6
Ausländer <sup>2</sup> .....	%	7,8	8,2	8,5	9,4	9,4	9,1	8,8	8,1	8,2
Jugendliche .....	%	2,5	2,6	2,7	3,0	3,1	3,0	2,9	2,8	2,8
Kurzarbeiter .....	1 000	31,0	28,2	65,7	...	...	...	...	...	...
Gemeldete Stellen <sup>3</sup> .....	1 000	151,1	144,8	140,3	137,6	137,9	138,3	136,2	134,8	133,0

## Öffentliche Sozialleistungen

(Daten der Bundesagentur für Arbeit)

### Arbeitslosenversicherung (SGB III – Arbeitsförderung –)<sup>4</sup>

Anspruchsberechtigte von Arbeitslosengeld I .....	1 000	121,4	122,1	133,0	154,7	157,4	146,3	136,5	...	...
darunter Leistungsbeziehende von Arbeitslosengeld I ...	1 000	117,5	117,5	128,3	149,6	152,3	141,2	131,6	...	...
Ausgaben für Arbeitslosengeld I <sup>5</sup> .....	Mill. Euro	242,7	247,1	253,3	284,6	334,7	336,6	310,9	291,8	289,2

## Steuern

### Gemeinschaftsteuern

darunter Steuern vom Einkommen .....	Mill. Euro	5 054,0	...	...	...	...	...	...	...	...
davon Lohnsteuer .....	Mill. Euro	4 380,2	...	...	...	...	...	...	...	...
veranlagte Einkommensteuer .....	Mill. Euro	248,5	...	...	...	...	...	...	...	...
nicht veranlagte Steuern vom Ertrag .....	Mill. Euro	280,1	...	...	...	...	...	...	...	...
Abgeltungsteuer .....	Mill. Euro	60,9	...	...	...	...	...	...	...	...
Körperschaftsteuer .....	Mill. Euro	84,3	...	...	...	...	...	...	...	...
Umsatzsteuer (Mehrwertsteuer) .....	Mill. Euro	2 508,4	...	...	...	...	...	...	...	...
<b>Landessteuern</b> .....	Mill. Euro	325,8	...	...	...	...	...	...	...	...
darunter Erbschaftsteuer .....	Mill. Euro	125,9	...	...	...	...	...	...	...	...
Grunderwerbsteuer .....	Mill. Euro	162,2	...	...	...	...	...	...	...	...
Biersteuer .....	Mill. Euro	13,1	...	...	...	...	...	...	...	...
<b>Gemeindesteuern</b> <sup>6, 7, 8</sup> .....	Mill. Euro	.	.	...	.	.	...	.	.	...
darunter Grundsteuer A .....	Mill. Euro	.	.	...	.	.	...	.	.	...
Grundsteuer B .....	Mill. Euro	.	.	...	.	.	...	.	.	...
Gewerbesteuer (brutto) .....	Mill. Euro	.	.	...	.	.	...	.	.	...

### Steuereinnahmen des Bundes

darunter Anteil an den Steuern vom Einkommen <sup>9, 10</sup> .....	Mill. Euro	1 917,7	...	...	...	...	...	...	...	...
Anteil an der Gewerbesteuerumlage <sup>9, 11</sup> .....	Mill. Euro	118,7	...	...	...	...	...	...	...	...

### Steuereinnahmen des Landes

darunter Anteil an den Steuern vom Einkommen <sup>9, 10</sup> .....	Mill. Euro	1 917,7	...	...	...	...	...	...	...	...
Anteil an der Gewerbesteuerumlage <sup>9, 11, 12</sup> .....	Mill. Euro	162,0	...	...	...	...	...	...	...	...

### Steuereinnahmen der Gemeinden/Gv<sup>7, 8, 9</sup>

darunter Anteil an der Lohn- und veranlagter Einkommensteuer <sup>8, 13</sup> .....	Mill. Euro	612,4	...	...	...	...	...	...	...	...
Anteil an den Steuern vom Umsatz .....	Mill. Euro	.	.	...	.	.	...	.	.	...
Gewerbesteuer (netto) <sup>6, 14</sup> .....	Mill. Euro	.	.	...	.	.	...	.	.	...

1 Arbeitslose in Prozent aller zivilen Erwerbspersonen.

2 Ab September 2021: Einschl. Staatenlose sowie Personen ohne Angabe zur Staatsangehörigkeit.

3 Ohne geförderte Stellen.

4 Daten nach Revision.

5 Einschl. Arbeitslosengeld bei beruflicher Weiterbildung.

6 Vierteljährliche Kassenstatistik.

7 Quartalsbeträge (jeweils unter dem letzten Quartalsmonat nachgewiesen).

8 Einschließlich Steueraufkommen der Landkreise.

9 Quelle: Bundesministerium der Finanzen (BMF).

10 März, Juni, September und Dezember: Termin von Vierteljahreszahlungen.

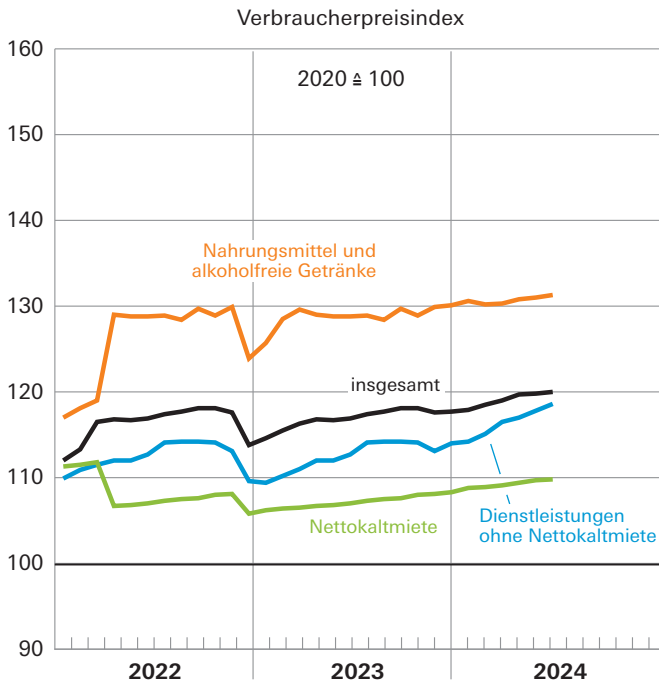
11 April, Juli, Oktober und Dezember: Termin von Vierteljahreszahlungen.

12 Einschließlich Erhöhungsbetrag.

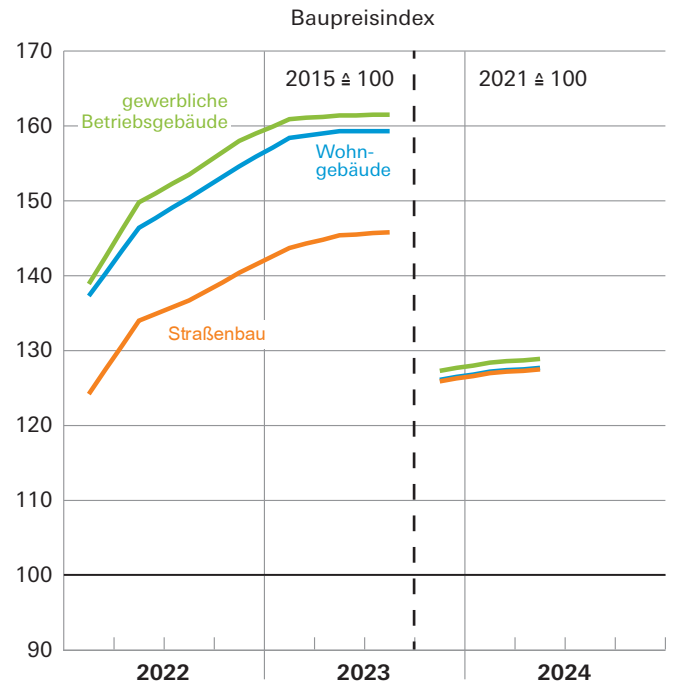
13 Einschließlich Zinsabschlag.

14 Nach Abzug der Gewerbesteuerumlage.

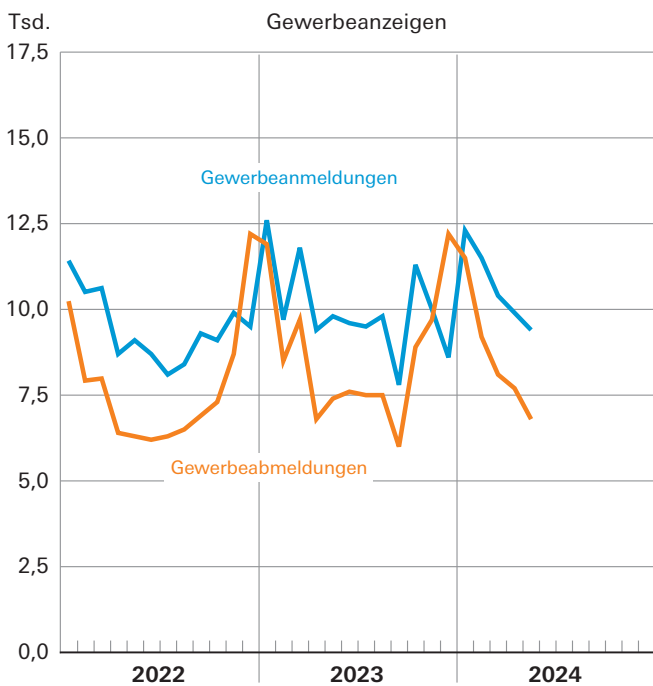
## Preise



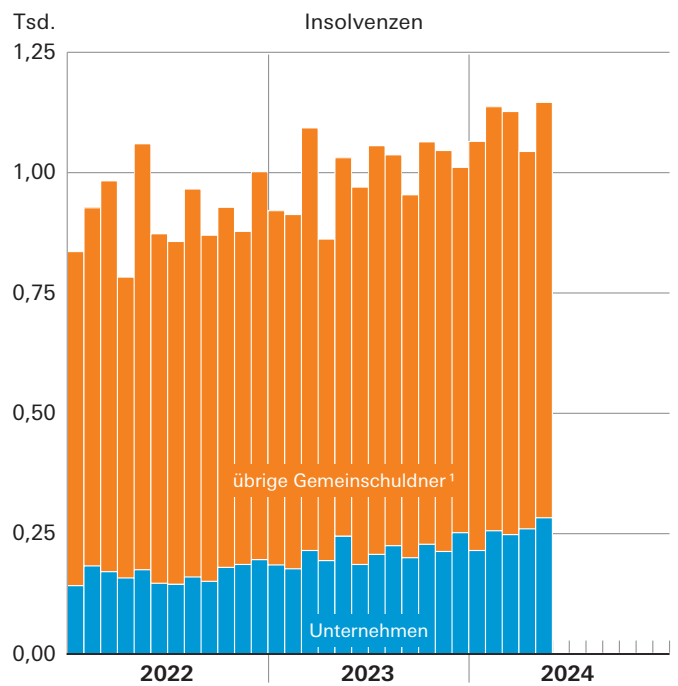
Weitere Informationen und Statistiken zum Thema Verbraucherpreisindex unter: <http://q.bayern.de/vpi>



Weitere Informationen und Statistiken zum Thema Baupreisindex unter: <http://q.bayern.de/bpi>

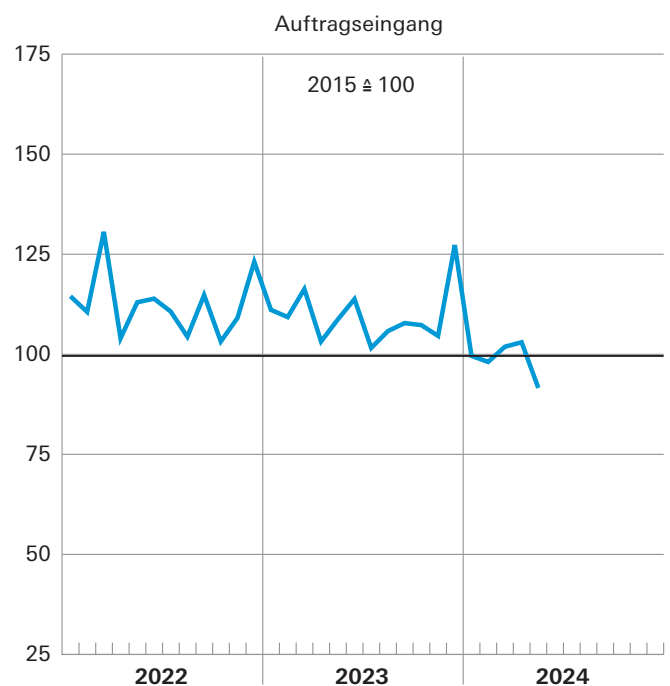
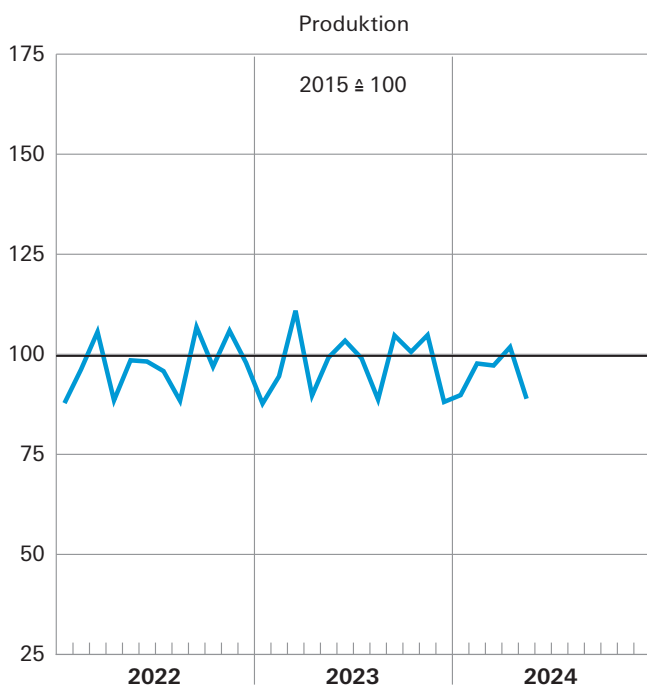
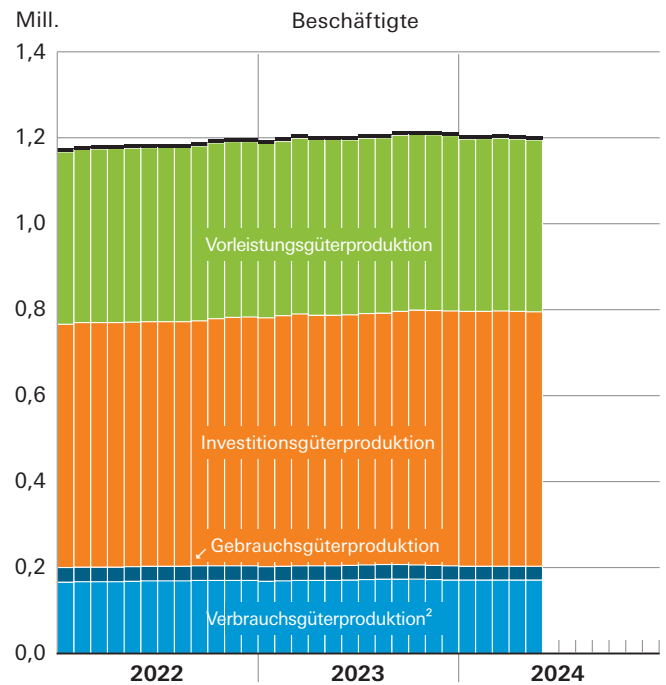
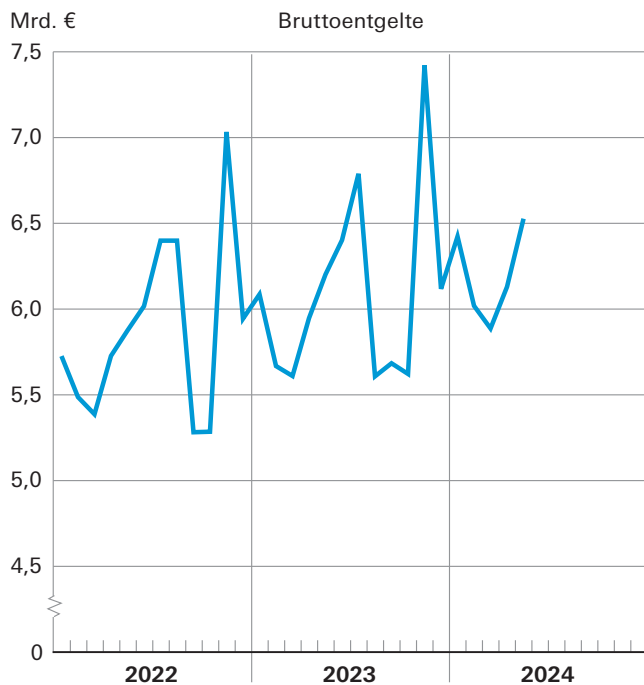


Weitere Informationen und Statistiken zum Thema Gewerbeanzeigen unter: <http://q.bayern.de/gewerbeanzeigen>



Weitere Informationen und Statistiken zum Thema Insolvenzen unter: <http://q.bayern.de/insolvenzen>

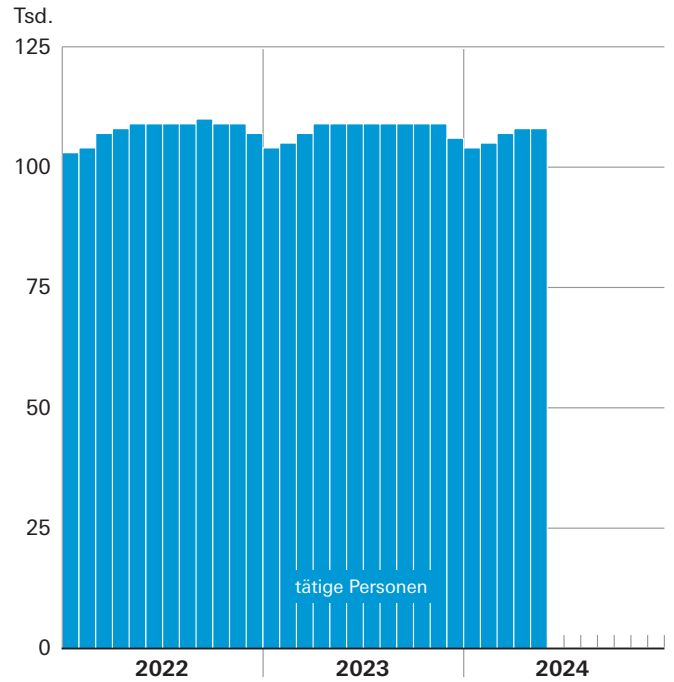
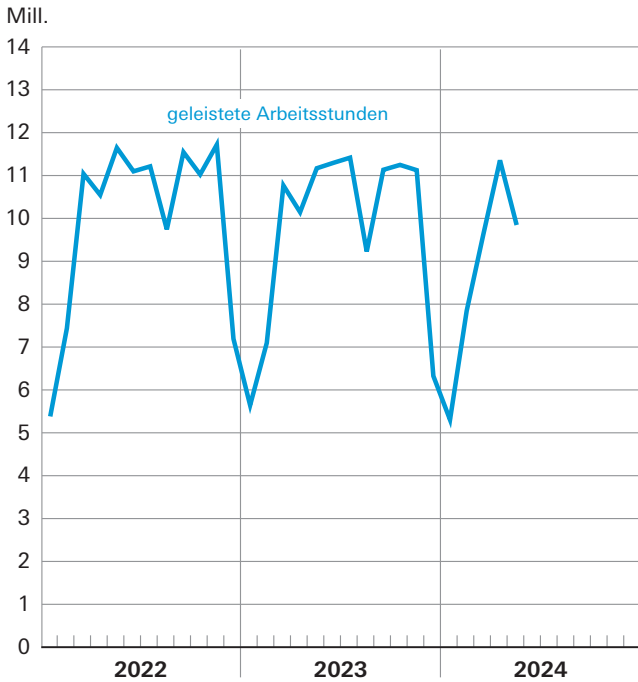
1 Einschließlich Verbraucherinsolvenzen.

Verarbeitendes Gewerbe<sup>1</sup>

Weitere Informationen und Statistiken  
zum Thema Verarbeitendes Gewerbe  
unter: <http://q.bayern.de/verarbeitendesgewerbe>

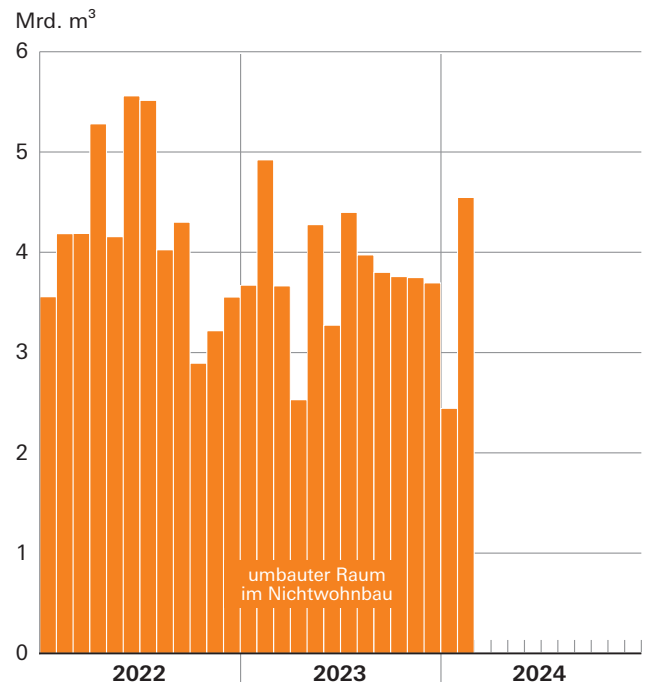
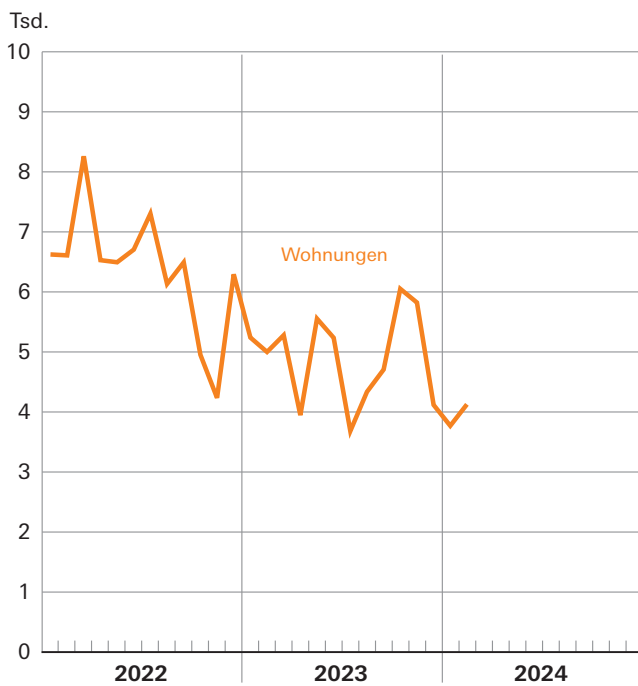
1 Sowie Bergbau und Gewinnung von Steinen und Erden; nur Betriebe mit 50 oder mehr Beschäftigten. 2 Einschließlich Energie.

### Bauhauptgewerbe



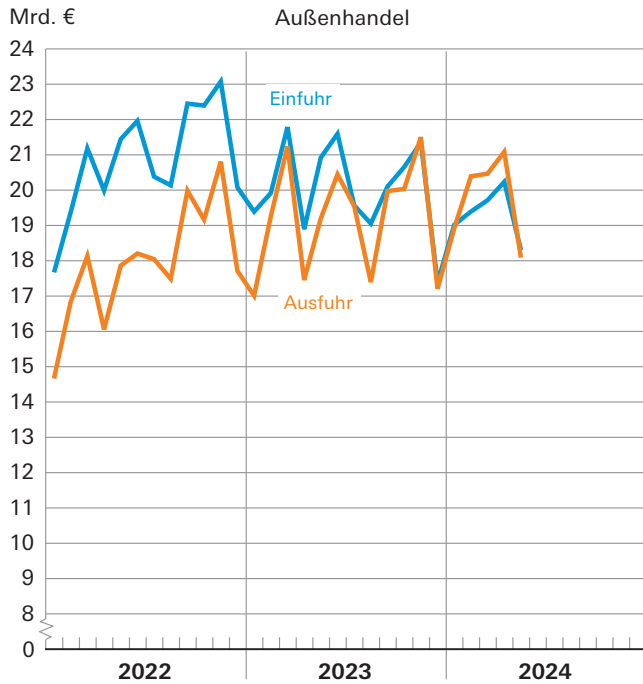
Weitere Informationen und Statistiken  
zum Thema Baugewerbe  
unter: <http://q.bayern.de/baugewerbe>

### Baugenehmigungen

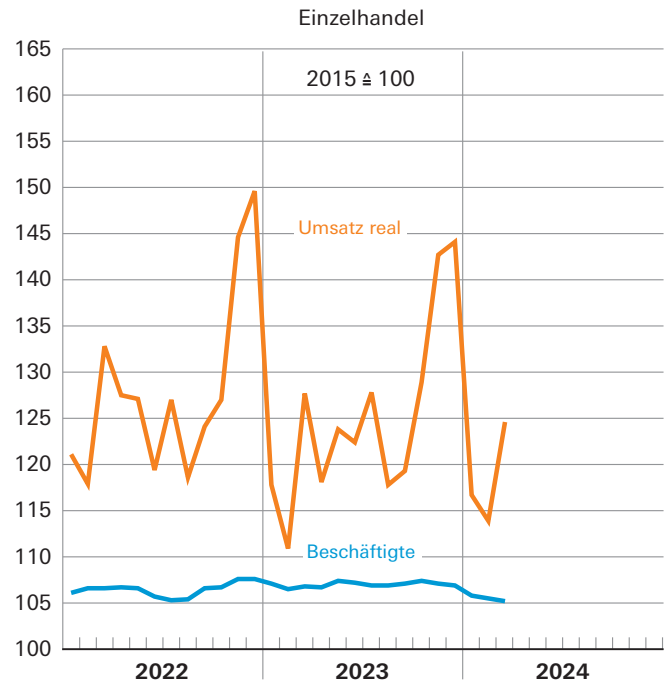


Weitere Informationen und Statistiken  
zum Thema Baugenehmigungen  
unter: <http://q.bayern.de/bautatigkeit>

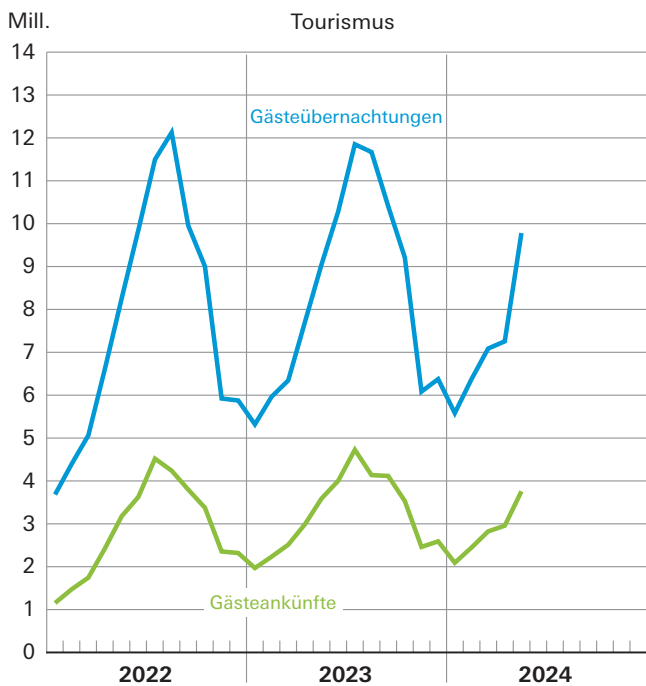
## Handel und Gastgewerbe



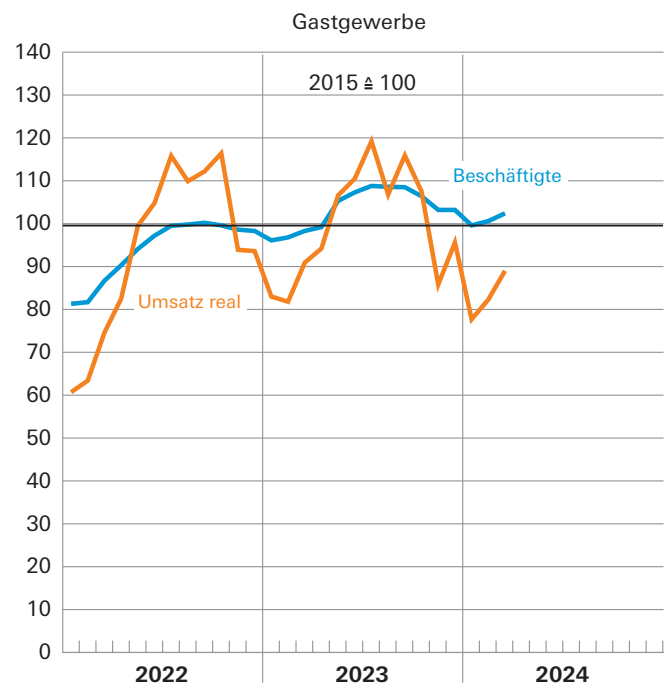
Weitere Informationen und Statistiken  
zum Thema Außenhandel unter:  
<http://q.bayern.de/aussenhandel>



Weitere Informationen und Statistiken  
zum Thema Einzelhandel unter:  
<http://q.bayern.de/binnenhandel>



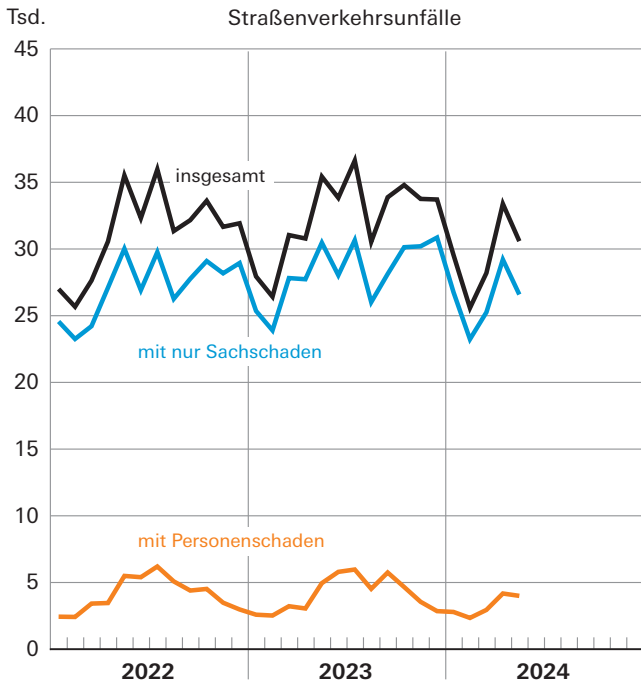
Weitere Informationen und Statistiken  
zum Thema Tourismus unter:  
<http://q.bayern.de/fremdenverkehr>



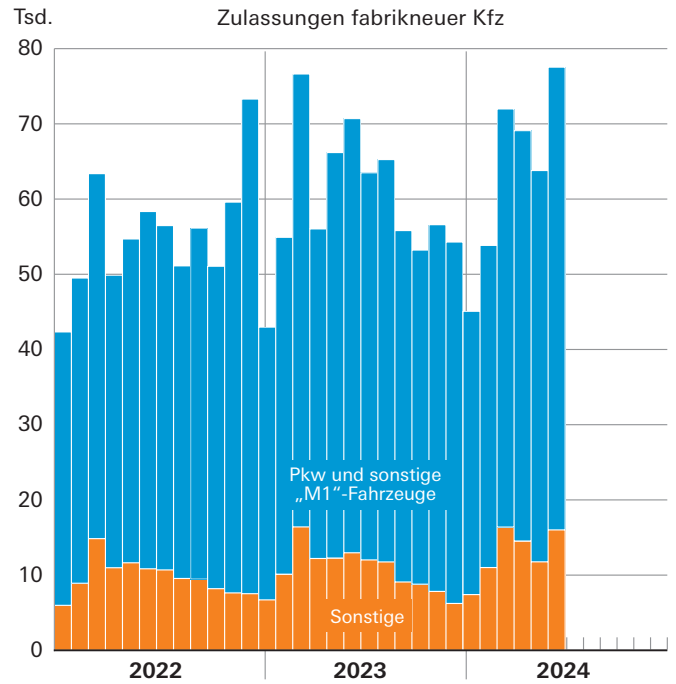
Weitere Informationen und Statistiken  
zum Thema Gastgewerbe unter:  
<http://q.bayern.de/gastgewerbe>



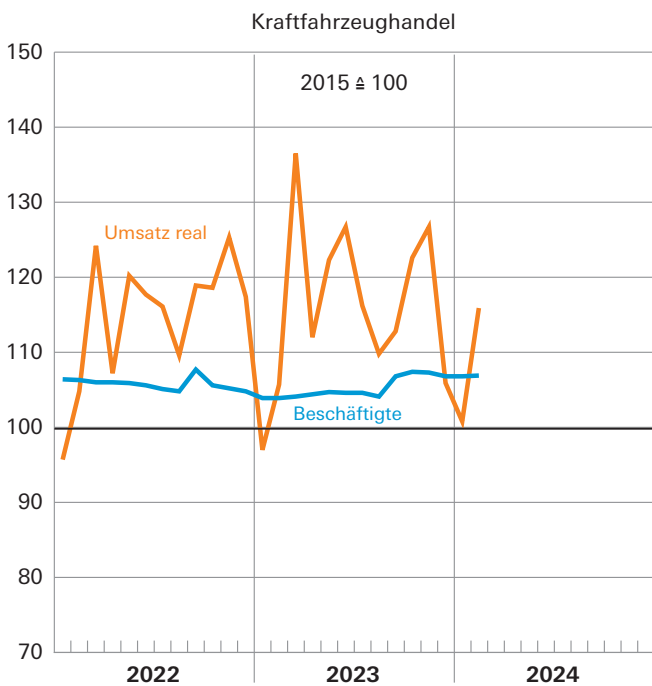
## Verkehr



Weitere Informationen und Statistiken  
zum Thema Straßenverkehrsunfälle unter:  
<http://q.bayern.de/unfaelle>

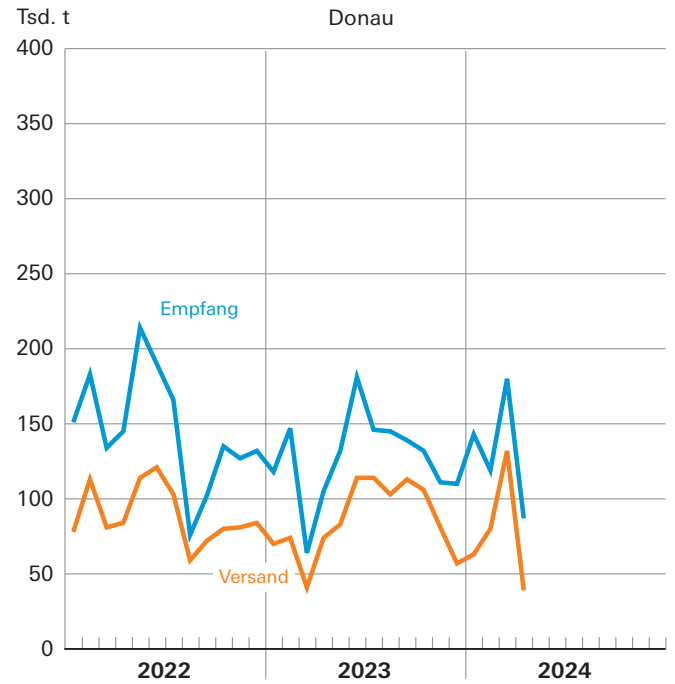
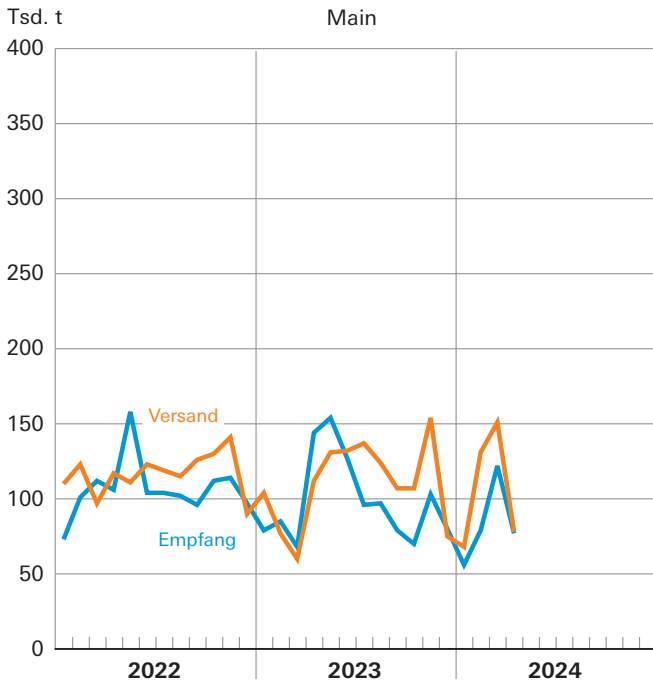


Weitere Informationen und Statistiken  
zum Thema Kfz-Zulassungen unter:  
<http://q.bayern.de/zulassungen>

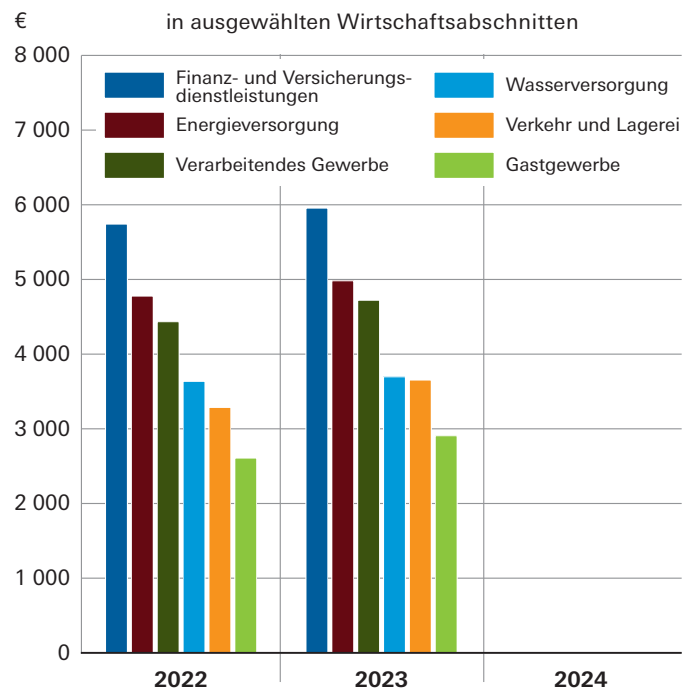
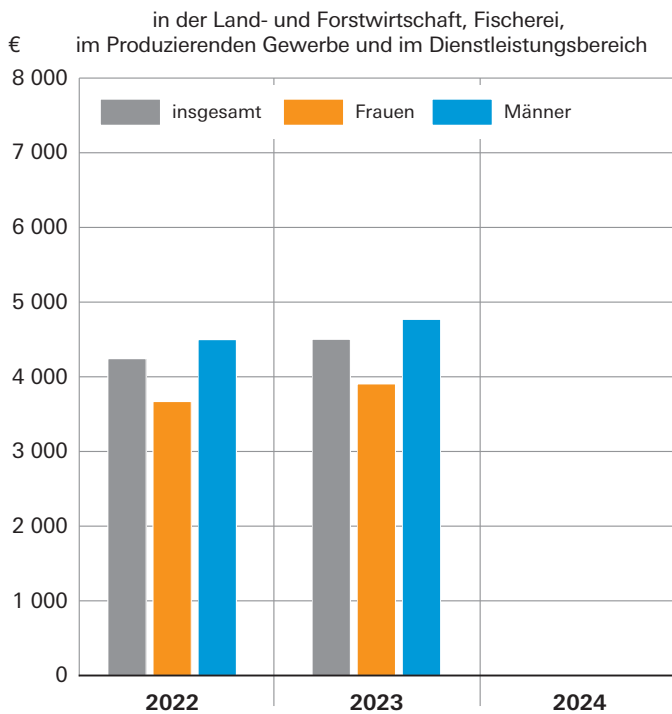


Weitere Informationen und Statistiken  
zum Thema Einzelhandel unter:  
<http://q.bayern.de/kfz-handel>

### Binnenschifffahrt



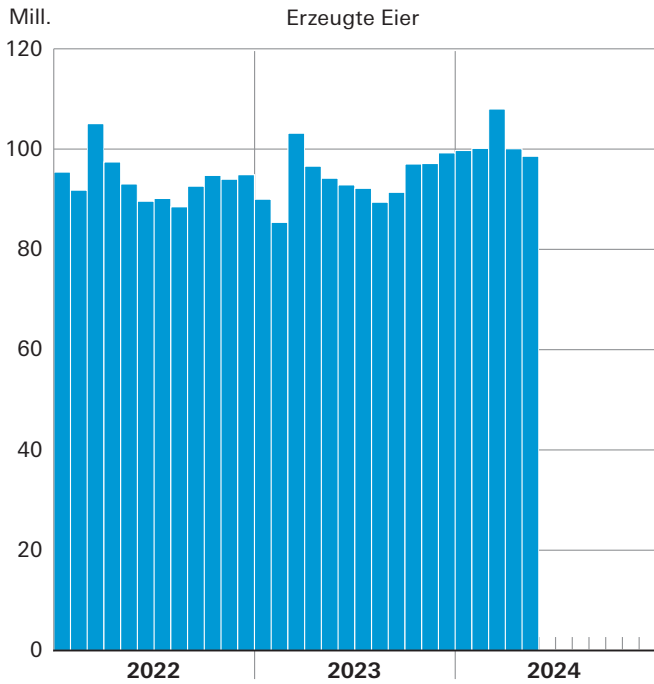
### Bruttomonatsverdienste<sup>1</sup> der vollzeitbeschäftigten Arbeitnehmer



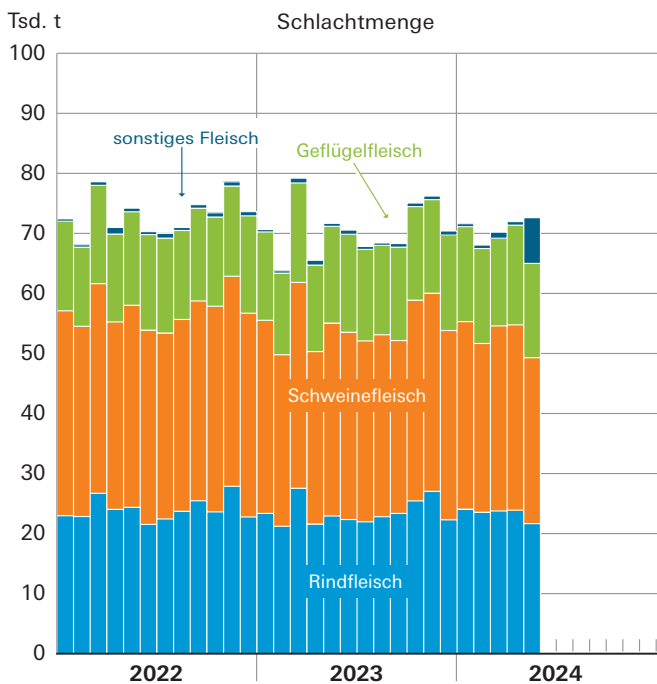
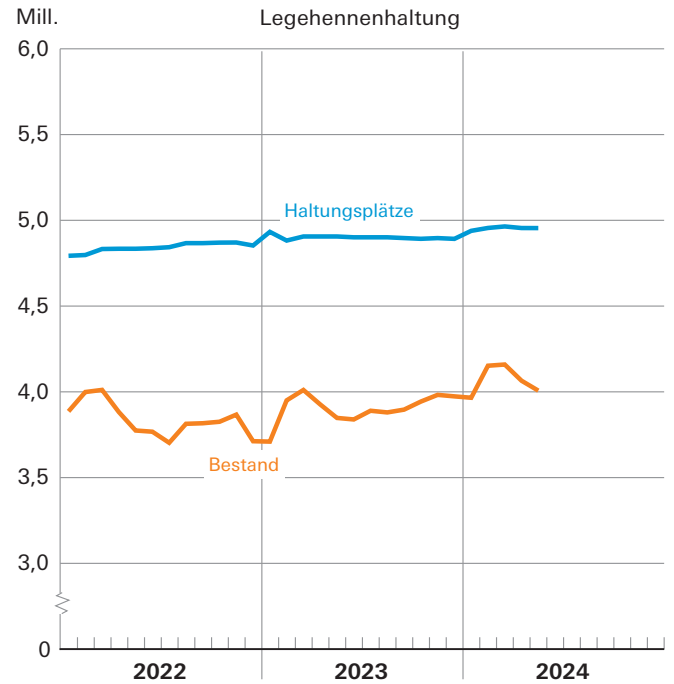
Weitere Informationen und Statistiken  
zum Thema Verdienste unter:  
<http://q.bayern.de/verdienste>

<sup>1</sup> Jeweils zum Berichtsmonat April ohne Sonderzahlungen.

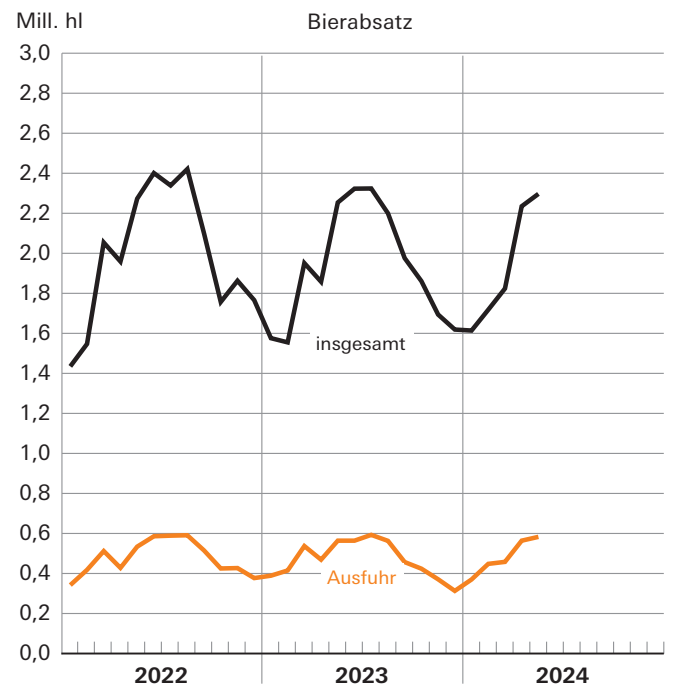
## Landwirtschaft



Weitere Informationen und Statistiken  
zum Thema Landwirtschaft unter:  
<http://q.bayern.de/tiererzeugnisse>

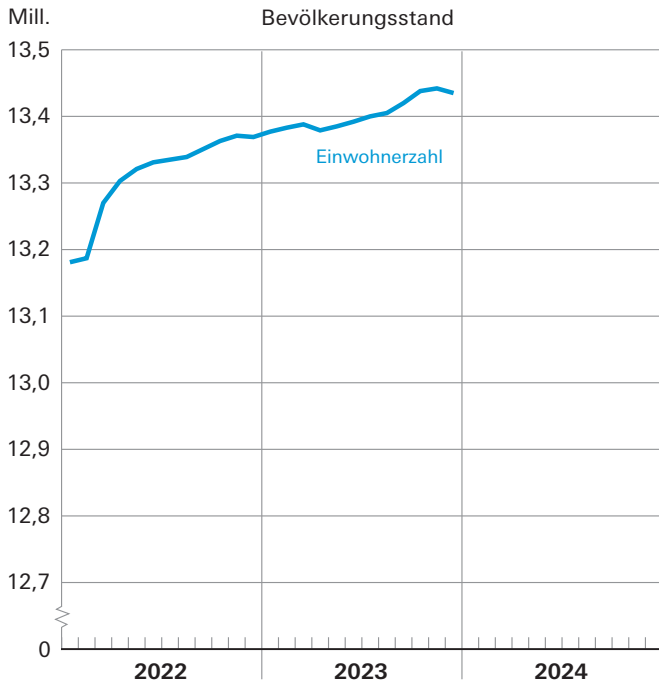


Weitere Informationen und Statistiken  
zum Thema Schlachtmengen unter:  
<http://q.bayern.de/tiererzeugnisse>

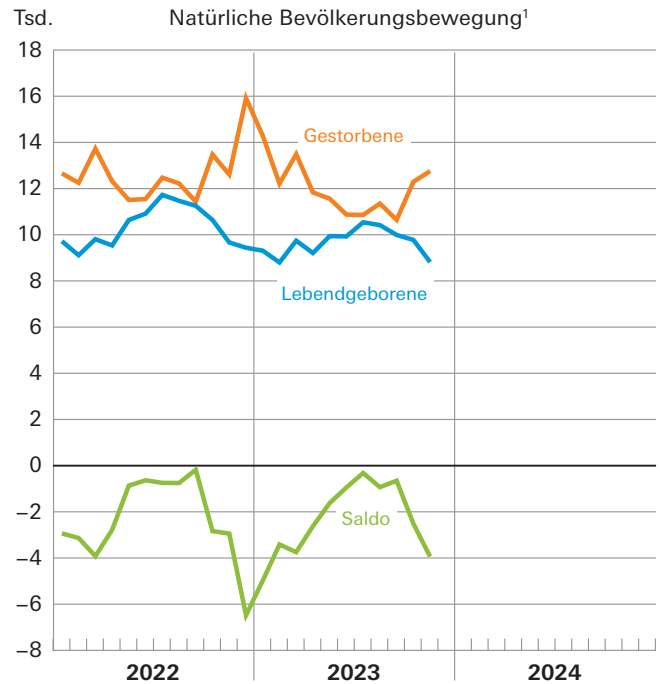


Aus: Statistisches Bundesamt, Fachserie 14, Reihe 9.2.1:  
Finanzen und Steuern, Absatz von Bier  
<http://q.bayern.de/bierabsatz>

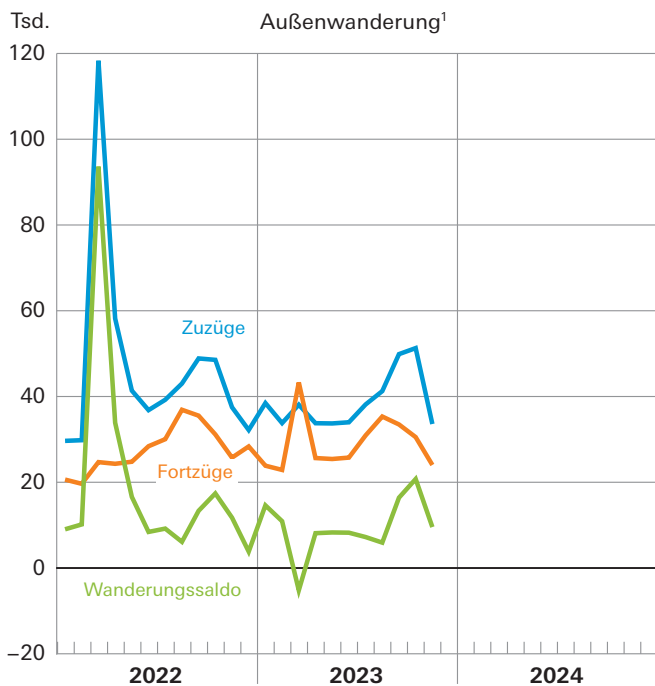
## Bevölkerung



Weitere Informationen und Statistiken  
zum Thema Bevölkerung unter:  
<http://q.bayern.de/bevoelkerung>



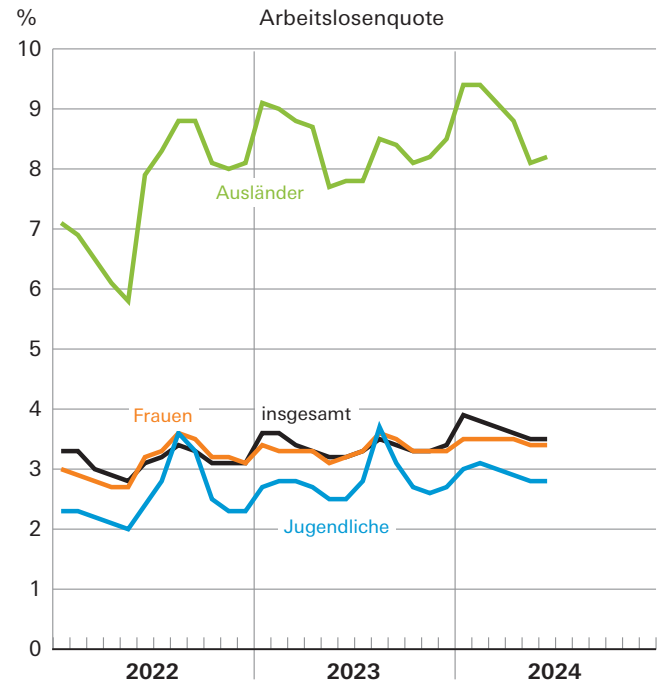
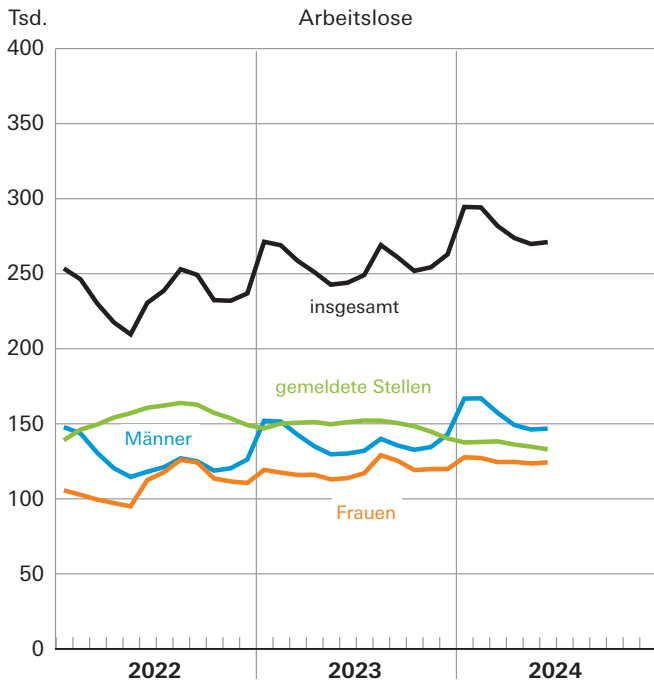
Weitere Informationen und Statistiken  
zum Thema natürliche Bevölkerungsbewegung unter:  
<http://q.bayern.de/bewegungen>



Weitere Informationen und Statistiken  
zum Thema Wanderungen unter:  
<http://q.bayern.de/wanderungen>

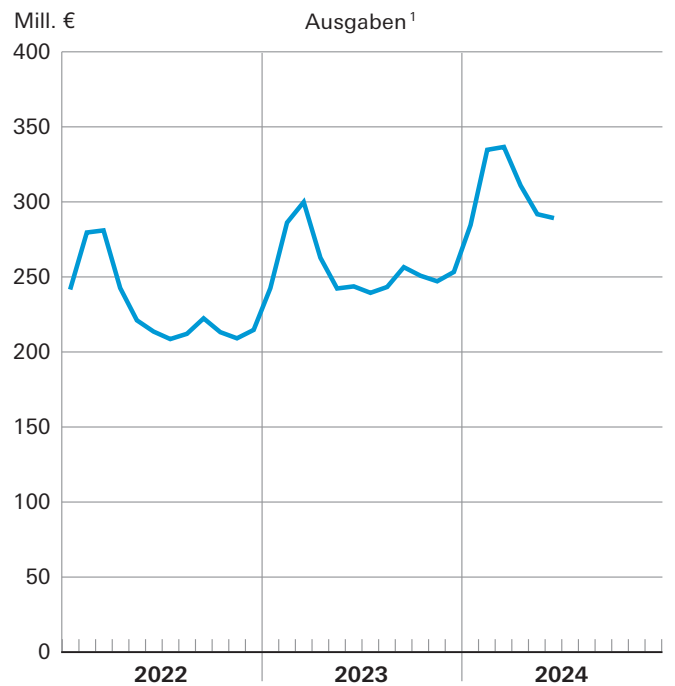
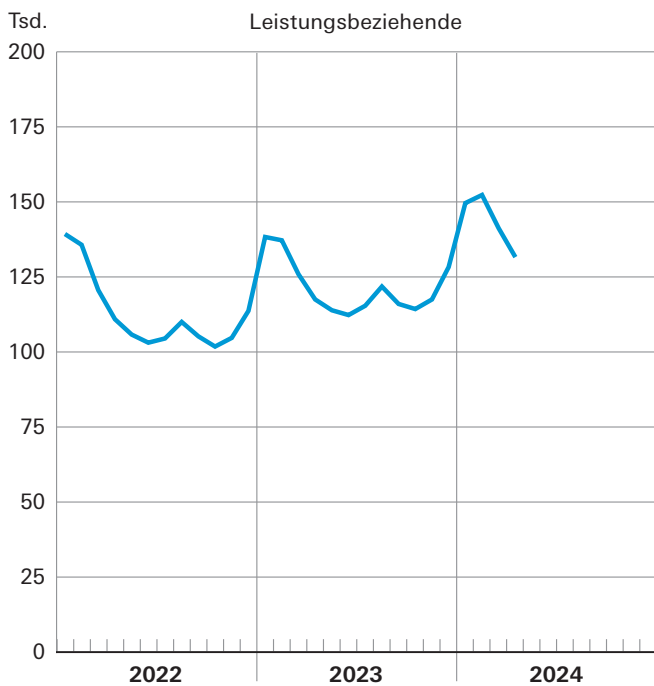
<sup>1</sup> Die Zahlen der natürlichen Bevölkerungsbewegung und der Wanderungen geben den jeweils aktuellen Stand des Monats im noch nicht abgeschlossenen Berichtsjahr wieder. Bis zum Ende des Jahres können Nachmeldungen der Städte und Gemeinden für die einzelnen Monate erfolgen, so dass sich die endgültigen Monatsergebnisse noch ändern können.

## Arbeitsmarkt



Weitere Informationen und Statistiken  
zum Thema Arbeitsmarkt unter:  
<http://q.bayern.de/erwerbstaetigkeit>

## Arbeitslosengeld I



Weitere Informationen und Statistiken  
zum Thema Leistungsbeziehende unter:  
<http://q.bayern.de/leistungsbeziehende>

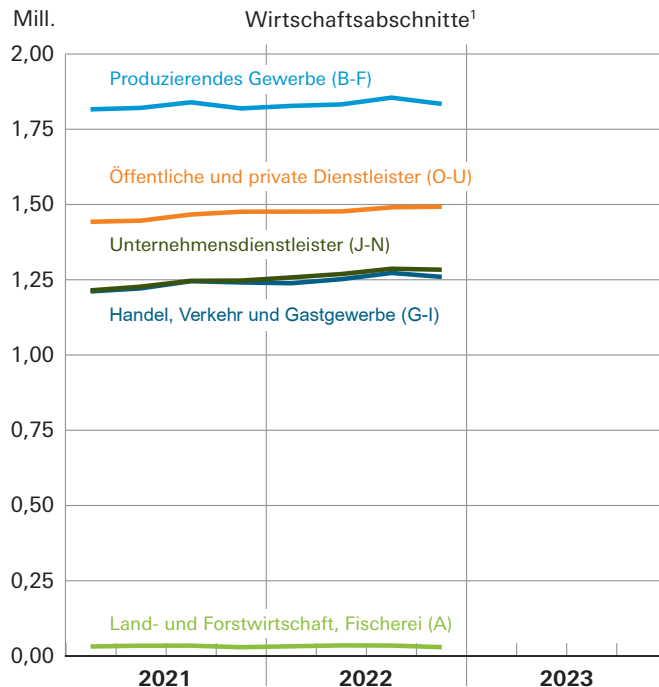
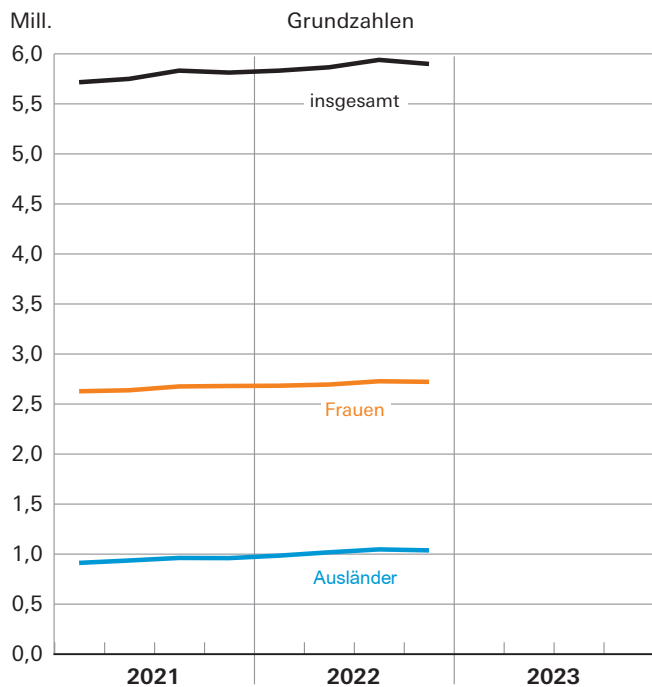


Weitere Informationen und Statistiken  
zum Thema Sozialausgaben unter:  
<http://q.bayern.de/sozialhilfeausgaben>

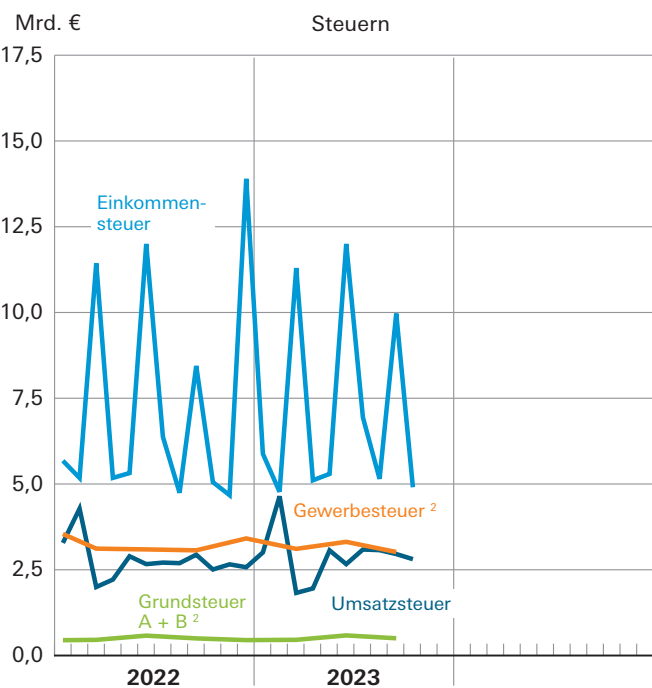
<sup>1</sup> Ab 2016 inklusive Arbeitslosengeld bei beruflicher Weiterbildung.



## Sozialversicherungspflichtig Beschäftigte am Arbeitsort



Weitere Informationen und Statistiken  
zum Thema Beschäftigte unter:  
<http://q.bayern.de/erwerbstaetigkeit>



Weitere Informationen und Statistiken  
zum Thema Steuern unter:  
<http://q.bayern.de/steuern>

1 Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008); in Klammern WZ-Code (vgl. Statistischer Bericht A6501C). 2 Quartalswerte.

# Tourismus in Bayern im Jahr 2023 – **Burgen und Schlösser**



Hier geht's zum Video:  
[www.statistik.bayern.de/presse/  
mitteilungen/2024/pm188](http://www.statistik.bayern.de/presse/mitteilungen/2024/pm188)



Alle bisher erschienenen Videoclips befinden sich hier:  
[www.statistik.bayern.de/presse/mediathek](http://www.statistik.bayern.de/presse/mediathek)

NEUERSCHEINUNGEN

STATISTISCHE BERICHTE

Europawahl in Bayern 2024

- Europawahl in Bayern am 9. Juni 2024  
Endgültiges Ergebnis

Gewerbeanzeigen

- Gewerbeanzeigen in Bayern im Juni 2024

Produzierendes Gewerbe

- Verarbeitendes Gewerbe in Bayern im Mai 2024  
(sowie Bergbau und Gewinnung von Steinen und Erden)
- Verarbeitendes Gewerbe in den kreisfreien Städten  
und Landkreisen Bayerns 2023  
(sowie Bergbau und Gewinnung von Steinen und Erden)
- Index der Produktion für das Verarbeitende Gewerbe  
in Bayern im Mai 2024 (sowie Bergbau und Gewinnung  
von Steinen und Erden)  
Basisjahr 2015
- Index des Auftragseingangs für das Verarbeitende  
Gewerbe in Bayern im Mai 2024  
Basisjahr 2015

Bauhauptgewerbe

- Bauhauptgewerbe in Bayern im Mai 2022

Energie- und Wasserversorgung

- Energiewirtschaft in Bayern – Teil I: Monatsergebnisse
- Energiewirtschaft in Bayern – Teil I: Endgültige Monats-  
ergebnisse

Bautätigkeit

- Baugenehmigungen in Bayern im Mai 2024

Handel

- Ausfuhr und Einfuhr Bayerns im Mai 2024

Tourismus und Gastgewerbe

- Tourismus in Bayern im Mai 2024

Straßenverkehr

- Straßenverkehrsunfälle in Bayern im April 2024

Gemeindefinanzen

- Kommunalfinanzen in Bayern Rechnungs-  
ergebnisse 2022

Steuern

- Einkommen der natürlichen Personen in Bayern 2020
- Ergebnisse der Lohn- und Einkommensteuerstatistik

Preise und Preisindizes

- Verbraucherpreisindex für Bayern  
Monatliche Indexwerte von Januar 2015 bis Juni 2024
- Verbraucherpreisindex für Bayern  
Monatliche Indexwerte von Januar 2015 bis Juni 2024



Alle Veröffentlichungen sind  
im Internet verfügbar unter  
[www.statistik.bayern.de/produkte](http://www.statistik.bayern.de/produkte)

ZEICHENERKLÄRUNG

- 0 mehr als nichts, aber weniger als  
die Hälfte der kleinsten in der Tabelle  
nachgewiesenen Einheit
- nichts vorhanden oder  
keine Veränderung
- / keine Angaben, da Zahlen nicht  
sicher genug
- Zahlenwert unbekannt,  
geheimzuhalten oder nicht rechenbar

- ... Angabe fällt später an
- x Tabellenfach gesperrt,  
da Aussage nicht sinnvoll
- ( ) Nachweis unter dem Vorbehalt,  
dass der Zahlenwert erhebliche  
Fehler aufweisen kann
- p vorläufiges Ergebnis
- r berichtigtes Ergebnis

- s geschätztes Ergebnis
- D Durchschnitt
- Δ entspricht
- 321 aktuellster Zahlenwert bzw.  
entsprechender vergleichbarer  
Vorjahreswert

AUF - UND  
ABRUNDEN

Im Allgemeinen ist ohne Rücksicht auf die  
Endsummen auf- bzw. abgerundet worden.  
Deshalb können sich bei der Summie-  
rung von Einzelangaben geringfügige  
Abweichungen zu den ausgewiesenen End-  
summen ergeben. Bei der Aufgliederung  
der Gesamtheit in Prozent kann die Summe  
der Einzelwerte wegen Rundens vom Wert  
100% abweichen. Eine Abstimmung auf  
100% erfolgt im Allgemeinen nicht.

✱ Für 3 109 517 Wohnungen wurde in der Haushaltegenerierung durch maschinellen Abgleich von Nutzerinformationen mehr als ein zugehöriger Haushalt identifiziert. Aufgrund dessen wurden 6 223 360 Teilhaushalte zu 3 109 517 Haushalten zusammengelegt.

