

Autorin: Dr. Heike Püttmann
Tabellen und Grafik: Dr. Heike Püttmann

Umstellung und Erweiterung der Haushaltsstatistiken

Neue Datengrundlage und Umstellung des Haushaltsgenerierungsverfahrens

Aus Anlass der Umstellung verschiedener IT-Verfahren der Landeshauptstadt München wurden auch die Haushaltsstatistiken überarbeitet und bieten in der neuen Fassung mehr Inhalte und Analysemöglichkeiten.

Begriffsklärungen

Haushalt - Personenverband

Haushalte an sich werden melderechtlich nicht erfasst. Unter Haushalte werden Personen verstanden, die aufgrund von Indizien, die einem Abzug des Melderegisters zu einem bestimmten Stichtag entstammen, zusammengeführt werden können.

Im Einwohnermelderegister gibt es lediglich Verzeigerungen zwischen familienrechtlich zusammengehörenden Personen eines Personen- bzw. Familienverbandes: z.B. Ehe- und Lebenspartner, Eltern bzw. Elternteile und deren Kinder. Letztere Verzeigerung wird dabei nur bis zur Vollendung des 18. Lebensjahres im Melderegister geführt und danach gelöscht. Personen einer solch verzeigerten Gruppe, die zudem noch an der gleichen Adresse wohnen, werden als Kernhaushalt bezeichnet. Im Gegensatz dazu steht der steuerrechtliche Personenverband, der diese Bedingung nicht erfüllen muss.

Auf Basis des Einwohnermelderegisterabzugs können somit nicht familienrechtlich verbundene Paare oder Eltern mit einem erwachsenen Kind die an der gleichen Adresse wohnen nicht als gemeinsamer Haushalt identifiziert werden. Um diese Personenverbände dennoch als Haushalte identifizieren zu können, bedarf es eines Haushaltsgenerierungsverfahrens. Auf Basis verschiedener Indizien der Zusammengehörigkeit werden darin schrittweise Personenverbände abgeleitet. Ein Personenverband ist somit der Oberbegriff für alle abgeleiteten Personenzusammenschlüsse. Die zuvor genannten Kernhaushalte sind eine Untermenge davon.

Kennzahlen der Bevölkerungsstatistik

Die Bevölkerungsstatistik unterscheidet in verschiedene Kennzahlen bezogen auf einer räumliche Einheit, die auch für die Haushaltsstatistik relevant sind. Die wichtigste Kennzahl ist die Hauptwohnsitzbevölkerung. Diese zählt alle Personen mit Hauptwohnung. Personen mit einer oder mehreren Nebenwohnungen werden als Nebenwohnsitzbevölkerung gezählt. Die Summe von Haupt- und Nebenwohnsitzbevölkerung sind die Wohnberechtigten. Daneben werden noch alle Wohnsitze bzw. nur Nebenwohnsitze gezählt.

Auswertungssichtweisen: Haushalte oder Personen mit Haushaltsmerkmalen

Fragestellungen zu Haushalten können sich unterscheiden in Fragen zu Gruppenmerkmalen von Haushalten bzw. Fragen zu Individualmerkmalen von Personen in Haushalten. Beispiele: die Frage „Anzahl der Haushalte nach Anzahl der Kinder im Haushalt“ bezieht sich auf das Gruppenmerkmal „Anzahl der Kinder im Haushalt“, die Frage „Anzahl Kinder unter sechs Jahren in Haushalten mit alleinig ausländischen Erwachsenen im Haushalt“ auf die Kombination des Individualmerkmals „Alter“ und des Gruppenmerkmals „Nationengruppen im Haushalt“.

Tabelle 1

Haushalte				
Haushalt-Nr.	Anzahl Personen	Alter jüngstes Kind	Geschlechter im Haushalt	Anzahl Generationen
1	4	2	m, w	2
2	1		w	1
3	2		m	1

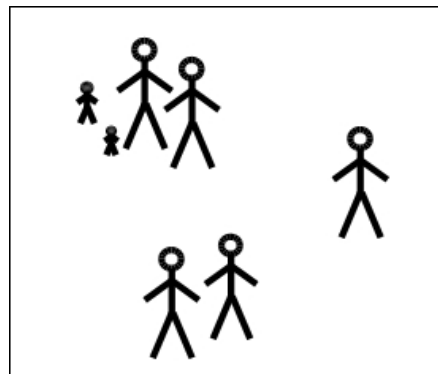
© Statistisches Amt München

Tabelle 2

Personen				
Haushalt-Nr.	Person-Nr.	Alter	Geschlecht	Anzahl Generationen
1	1	34	m	2
1	2	32	w	2
1	3	2	m	2
1	4	6	m	2
2	1	68	w	1
3	1	22	m	1
3	2	23	m	1

© Statistisches Amt München

Abbildung 1



Einfach dargestellt bedeutet die Sichtweise Haushalte somit ein Datensatz je Haushalt (=Haushaltsstatistiken), dagegen die Sichtweise Personen ein Datensatz je Person (=Personenstatistiken) – siehe Tabelle 1 und 2, dazu Abbildung 1. Es können nicht alle Kennzahlen der Bevölkerungsstatistik in beiden Sichtweisen abgebildet werden. So werden bei der Haushaltgenerierung nur der einzige bzw. der kommunalstatistisch wichtigste Wohnsitz einer Person berücksichtigt. Dadurch fehlen weitere Nebenwohnsitze in der Haushaltsstatistik.

Privathaushalte – Haushalte in Heimen und Anstalten

Haushalte in Heimen und Anstalten haben keine eigene Haushaltsführung, wie dies bei Privathaushalten der Fall ist. In München wie in den meisten anderen Kommunen gibt es kein Zentralregister für Adressen mit Heimen und Anstalten. Schwierigkeiten bereiten hier die unterschiedlichen Träger-schaften (privat, öffentlich, kirchlich) und erfassungstechnische bzw. meldepflichtige Zuständigkeiten (privat, kommunal, staatlich).

Die Bestimmung, ob Haushalte an einer Adresse als Privathaushalte bzw. Heim-/Anstalts-Haushalte zu zählen sind, erfolgt deshalb nach der Haushaltgenerierung auf Adressenebene aus den generierten Haushalten selber, durch Analyse der demographischen Daten

- Anzahl Erwachsene (Alter > 17)
- Anzahl Verwitwete
- Anzahl Ledige
- Anzahl Senioren (Alter > 65)
- Anzahl Kinder (Alter < 18)

- Anzahl junge Erwachsene (Alter ≥ 16 und Alter ≤ 32)
- Anzahl erwachsene Männer
- Anzahl Kinder unter 16
- Anzahl erwachsene Frauen
- Anzahl Personen
- Anzahl 1-Personen-Haushalte
- Anzahl Nicht-EU-Bürger (1. Staatsangehörigkeit ist nicht deutsch und ausländische ist nicht EU-Mitglied zum Stichtag)

und etlichen Verhältniszahlen zueinander. Sämtliche Haushalte an einer Adresse mit Heim-/Anstaltscharakter werden der Adressenkategorie Heim-/Anstalt zugerechnet. Dazu zählen z.B. auch die Hausmeisterwohnungen in Heimen.

Die Privathaushalte sind demnach statistisch abgeleitet, mit der Adresse als kleinste räumliche Einheit. Eine Verknüpfung mit Wohnungen ist u.a. auch deshalb nicht möglich.

Die selben statistischen Ableitungsregeln werden sowohl im Altverfahren als auch im neuen Verfahren angewendet.

Datengrundlagen und Haushaltgenerierungsverfahren

Altverfahren bis 12/2011

Im Altverfahren beruhte das Haushaltgenerierungsverfahren auf einer individuellen Statistikdatei des Bevölkerungsbestandes. Dieser wurde halbjährlich zum 30.6. und 31.12. erzeugt, mit je einem Datensatz je Wohnsitz einer Person. Aus dieser Datei wurden mit einem eigenentwickelten Haushaltgenerierungsprogramm der Landeshauptstadt München¹⁾ die Haushalte bestimmt. Die daraus erstellte neue Ergebnisdatei hatte je Haushalt einen Datensatz mit acht Individualmerkmalen je Person und bis zu 20 Personen je Haushalt. Die Zeitreihe dieser Daten reicht von 12/1995 bis 12/2011.

Die Bestimmung der Heim- / Anstaltsadressen bzw. Privatadressen erfolgte mit der zuvor dargestellten Methode.

Mängel dieses Verfahrens bzw. der Ergebnisse waren u.a., dass

- es nicht mit anderen Kommunen wg. teils abweichender Methodik vergleichbar war und
- nur die Wohnberechtigte Bevölkerung bzw.
- nur Gruppenmerkmale auswertbar waren.

Neues Verfahren ab 1/2012

Für das neue Verfahren wird ein annähernd dem Statistikdatensatz Bevölkerungsbestand des Deutschen Städtetages entsprechendes Dateiformat verwendet. Je Wohnsitz einer Person ist ein Datensatz mit sämtlichen verfügbaren Individualmerkmalen enthalten. Dieser wird monatlich zum Monatsletzten erzeugt und liegt ab 6/2011 für die interne Qualitätsprüfung und ab 1/2012 für Auswertungen vor. Die Dateigrundlage für die Bevölkerungsbestandsstatistik ist derzeit eine andere.

Die Haushaltgenerierung wird mit dem Programm HHGen des KOSIS-Verbandes (Kommunales Statistisches Informationssystem) durchgeführt²⁾.

Für sämtliche HHGen-Steuerparameter werden die voreingestellten Werte verwendet, wobei nur der einzige bzw. der kommunalstatistisch wichtigste Wohnsitz einer Person für die Haushaltgenerierung berücksichtigt wird.

Für die Haushaltgenerierung werden acht Generierungsstufen durchlaufen. Eingangsmerkmale für die Haushaltgenerierung sind:

- aktuelle und letzte Wohnadresse
- Einzugsdatum
- Zuzugsdatum
- Zuzugsherkunftsgemeinde
- Geburtsdatum
- Geschlecht
- Familienstand
- erste Staatsangehörigkeit
- Namesübereinstimmungs-Nummer des aktuellen bzw. vorherigen Namens sowie des Geburtsnamens.

Die ermittelten Haushaltsinformationen werden im Datensatz an den dafür vorgesehenen Datensatzstellen geschrieben und das Ergebnis als neue Datei abgespeichert. Diese entspricht somit im Format der Ausgangsdatei, lediglich erweitert um inhaltliche Informationen der Haushalte. HHGen bietet auch die Möglichkeit Heim- und Anstaltsadressen zu ermitteln. Jedoch zeigte sich eine zu große Abweichung zu echten Heimadressen einer stichprobenartigen externen Heim-/Anstaltsliste. So wird auf eine Nutzung der mit HHGen bestimmten Heim- und Anstaltsadressen verzichtet. Stattdessen wird die Bestimmung der Adressen mit Privathaushalten bzw. Heimen / Anstalten wieder mit der zuvor dargestellten Methode durchgeführt.

Zwei Datenmodelle

Erweiterung der Datenverarbeitungsmethoden

Da sich die Inhalte und die Datenstruktur der Eingangsdaten vom alten und neuen Verfahren unterscheiden (aggregierte Haushalte vs. Personeneinzel-daten) bzw. die Auswertungsmöglichkeiten erweitert werden sollten, wurden für die neue Datengrundlage zwei neue Datenmodelle (Sternschematas) realisiert. Diese, die Auf- und Verarbeitung der Daten sowie die Bereitstellung von Auswertungen erfolgt im Data Warehouse ZIMAS (Zentrales Informationsmanagement und Analysesystem) des Statistischen Amtes.

Die Faktentabelle für die Personenstatistiken enthält wie die Ausgangsdatei einen Datensatz je Person, angereichert um weitere abgeleitete Haushaltsmerkmale des der jeweiligen Person zugehörigen Haushaltes. Die Faktentabelle der Haushaltsstatistiken enthält bis zu zwei Datensätzen je Haushalt: einen Datensatz für die Gruppenmerkmale der Hauptwohnsitzbevölkerung und / oder einen Datensatz für die Gruppenmerkmale der Nebenwohnsitzbevölkerung.

Um möglichst detaillierte Informationen der Haushaltsmitglieder in den Haushaltsstatistiken anzubieten, wurden ausgewählte Individualmerkmale per Listenaggregation ohne Doppelnennungen zusammengefasst (Vergleiche Tabelle 1 und 2, Attribut Geschlechter im Haushalt bzw. Geschlecht im Haushalt-Nr. 1). Dies erfolgte z.T. für alle bzw. nur für die erwachsenen Haushaltsmitglieder. Möglichkeiten der technischen Realisierung von Listenaggregationen in einer Oracle-Datenbank gibt z.B.: <http://www.oracle-base.com/articles/misc/string-aggregation-techniques.php>. So können aggregierte Listen mit und ohne Wiederholungen der Attributausprägungen, sortiert bzw. unsortiert oder mit benutzerdefinierten Trennzeichen erstellt werden. Auswertungen auf Listenaggregate erfolgen dann mit Ähnlichkeitssuche oder regulären Ausdrücken.

Verbesserung der Qualitätssicherung

Im Altverfahren wurde die Datenqualität erst am Ende des Bearbeitungsprozesses mit einem Testskript geprüft. Ein Rückverfolgung zum Fehler verursachenden Attribut war damit nicht möglich. In den neuen Datenmodellen wurde nun ein Fehler-Logging während der Datenverarbeitung eingeführt. Im Fehlerfall stehen nun detaillierte Informationen zur Verfügung in welchem Datensatz welches Attribut zu einem Fehler geführt hat.

Jeder Lauf von HHGen produziert ein ASCII-Logdatei. Darin enthalten sind Statistiken über die Korrekturen im Rahmen der Plausibilisierung, die verwendeten Steuerparametereinstellungen, die durchlaufenen Haushaltgenerierungsstufen, Statistiken der gelesenen, ausgeschlossenen und verarbeitenden Datensätze sowie der generierten Haushalte. Da eventuelle Änderungen im Zeitverlauf mit solchen Einzelprotokollen nicht erkennbar sind, wird das monatliche Protokoll aufbereitet und in eine eigene Datenbanktabelle importiert. Durch die Betrachtung der Protokoll-Zeitreihen kann so schnell auf noch nicht bekannte Änderungen im Quellsystem des Einwohnerwesens oder der Auswirkung einer neuen HHGen-Programmversion geschlossen werden.

Die neuen Auswertungsmöglichkeiten

Statt bisher Halbjahresstatistiken stehen nun Monatsstatistiken zur Verfügung. Für die Haushaltsstatistiken werden folgende weitere Informationen bereit gestellt:

- Alter [jüngstes|ältestes] & [weibliches|männliches] Haushaltsmitglied
- Haupt- bzw. Nebenwohnsitzhaushalte
- Haupt- und Nebenwohnsitzbevölkerung bzw. wohnberechtigte Bevölkerung
- Listenaggregate der Religionsgruppen Erwachsene
- Listenaggregate der Geschlechter der Erwachsenen bzw. Kinder
- Anzahl 1. Staatsangehörigkeiten der Erwachsenen
- Anzahl Religionskombinationen der Erwachsenen
- Anzahl Familienstände der Erwachsenen

Ferner können auch die namentlichen Merkmalskombinationen der drei letztgenannten Merkmale (z.B. „Anzahl Haushalte mit mindestens einer Person der Nationalität x?“ oder „Anzahl Haushalte mit mindestens einer Person der Religion y?“) vom Statistischen Amt ausgewertet werden (siehe Beispiele weiter unten).

Personen in Haushalten können nun nach folgenden Haushaltsmerkmalen ausgewertet werden

- Haupt- und Nebenwohnsitzbevölkerung bzw. wohnberechtigte Bevölkerung
- Hauptwohnsitze, Nebenwohnsitze, Wohnsitze
- frei wählbare Alter bzw. Altersklassen
- Anzahl Wohn- und Nebenwohnsitze in München
- Wohndauer in München bzw. Dauer des Wohnverhältnisses
- Migrationshintergrund

Für beide Sichten stehen

- die bisherige Haushaltstypisierung sowie die Typisierung nach HHSTAT (KOSIS Gemeinschaft Haushalts- und Bevölkerungsstatistiken) bzw. BfLR (Bundesforschungsanstalt für Landeskunde und Raumordnung)
- Kombination der Nationengruppen mit Drill auf Nationengruppen EU (nur Erwachsene bzw. aller Haushaltsmitglieder)
- Anzahl Personen im Haushalt / Haushaltsgröße
- Anzahl Kinder
- Anzahl Generationen
- zur Verfügung.

Die meisten Auswertungsanforderungen können mit ZIMAS und dem im städtischen Intranet zugänglichen Business Intelligence Werkzeug Oracle Discoverer abgedeckt werden. In speziellen aber seltenen Ausnahmen müssen die Auswertungen wegen der Komplexität der Verknüpfungen und Aggregationen weiterhin in SQL-Datenbankabfragen formuliert werden.

Ein Beispiel für die neuen Auswertungsmöglichkeiten der Personenstatistiken mit Haushaltsmerkmalen stellt Tabelle 3, siehe Seite 36, dar.

Die Auswertung Alter und Migrationshintergrund von Kindern in Alleinerziehenden-Haushalten ist die Kombination der Individualmerkmale Alter und Migrationshintergrund mit dem Gruppenmerkmal Alleinerziehenden-Haushalt. Tabelle 4; Seite 36, ist ein Beispiel für eine tiefer gehende Analyse von Privathaushalten vom Typ „Ehepaar“ mit mindestens einer erwachsenen Person römisch-katholischer Konfession und mindestens einer erwachsenen deutschen Person.

Tabelle 3

**Das Alter und der Migrationshintergrund
von Kindern in Alleinerziehenden-Haushalten**
(Stand 31.12.2013)

Alter	Kinder			zusammen
	Deutsche ohne Migrationshintergrund	Deutsche mit Migrationshintergrund	Ausländer/innen	
Zusammen	18 382	13 155	6 429	37 966
darunter				
bis unter 3 Jahre	1 898	1 657	657	4 212
3 bis unter 6 Jahre	2 168	2 208	748	5 124
6 bis unter 9 Jahre	2 665	2 429	835	5 929
9 bis unter 12 Jahre	3 172	2 371	948	6 491
12 bis unter 15 Jahre	3 849	2 418	1 348	7 615
15 bis unter 18 Jahre	4 630	2 072	1 893	8 595

© Statistisches Amt München

**Die Privathaushalte vom Typ „Ehepaar“ 1)
mit mindestens einer erwachsenen Person römisch-katholischer Konfession und
mindestens einer erwachsenen deutschen Person**
(Stand: 31.12.2013)

Tabelle 4

Nationengruppen Erwachsene im Haushalt	Haushalte mit Kindern		Haushalte ohne Kinder	
	alle Erwachsenen im Haushalt sind römisch-katholischer Konfession	Erwachsene haben unterschiedliche Konfessionen 2)	alle Erwachsenen im Haushalt sind römisch-katholischer Konfession	Erwachsene haben unterschiedliche Konfessionen 2)
Insgesamt	12 275	22 176	34 818	36 803
darunter				
gemischt	2 343	5 741	3 109	6 749
darunter				
Ausländer/innen EU, Ausländer/innen nicht EU, Deutsche	10	275	13	263
Ausländer/innen EU, Deutsche	1 876	3 258	2 642	4 050
Ausländer/innen nicht EU, Deutsche	457	2 208	454	2 436
Ausländer/innen, Deutsche				
Deutsche	9 932	16 435	31 709	30 054

Hinweis: Mehr als zwei Nationengruppen treten auf in Haushalten mit mindestens einer weiteren erwachsenen Personen neben den Ehe- / Lebenspartnern.

1) Haushaltstyp HHSTAT (Ehepaare und Lebenspartnerschaften).- 2) Einschl. „ohne Angabe“.

© Statistisches Amt München

Vergleichbarkeit

Die Kennzahlen des neuen Verfahrens sind nicht mit den Kennzahlen aus Altverfahren vergleichbar. Zum einen basieren die Kennzahlen auf verschiedenen Datengrundlagen, zum anderen werden unterschiedliche Methoden und Eingangsmerkmale für die Haushaltegenerierung verwendet. Ferner berücksichtigt HHGen im Gegensatz zum Altverfahren nur noch die Staatsangehörigkeiten der Erwachsenen Haushaltsmitglieder. Dies hat zur Folge, dass Haushalte mit ausländischen Eltern und Optionskind(ern) als ausländische und nicht wie zuvor als gemischte Haushalte in die Ergebnisse eingehen.

Die Quantität der Änderungen zeigt ein Vergleich beider Verfahrensergebnisse für 12/2011: die Zahl der Privathaushalte steigt mit dem neuen Verfahren um 1,9% an. Dabei erhöht sich insbesondere der Anteil sonstiger Mehrpersonenhaushalte um 4,7% bzw. der Anteil nichtehelicher Paare verringert sich um 5,7%.

Die Bestimmung der Heim- / Anstaltsadressen erfolgt, wie zuvor dargestellt, in beiden Verfahren gleich.

*Vergleichbarkeit der
Ergebnisse aus altem und
neuem Verfahren*

Die Familientypen des Altverfahrens können durch Umgruppierung des Merkmals „Haushaltstyp HHSTAT“ zwar annähernd für das neue Verfahren abgeleitet werden, jedoch gibt es hier methodisch bedingt größere Haushalte.

*Vergleichbarkeit mit der
Einwohnerbestandsstatistik*

Die Basisdateien für die Einwohnerbestands- und die Haushaltsstatistik sind nicht identisch und weisen geringe Abweichungen in den Kennzahlen Wohnberechtigte sowie Haupt- und Nebenwohnsitzbevölkerung auf. Da ferner nur die einzige bzw. kommunalstatistisch wichtigsten Wohnsitze bei der Haushaltgenerierung verwendet werden, ist die Zahl der Wohnberechtigten geringer als in der Einwohnerbestandsstatistik.

*Vergleichbarkeit mit anderen
Kommunen*

Statistiken aller Art können nur dann zu Vergleichen herangezogen werden, wenn die Kennzahldefinitionen und die angewendeten Methoden gleich sind. Etliche deutsche Kommunen sind deshalb Mitglied im KOSIS-Verbund, der zentral Datenverarbeitungsinstrumente für die Städtestatistik entwickelt und pflegt, u.a. auch HHGen. Bis Ende 2012 hat die Landeshauptstadt München zunächst einen eigenen Ansatz für die Ableitung der Haushaltsstatistiken verfolgt. Mit der Einführung von HHGen reiht sich München ein in den immer größer werdenden Nutzerkreis von HHGen und stellt nun mit einer Vielzahl von anderen deutschen Kommunen vergleichbare Haushaltsstatistiken zur Verfügung.

Die Bestimmung des Migrationshintergrundes wird seit der Einführung 12/2006 mit einer im Statistischen Amt der Landeshauptstadt München selbst entwickelten und innerstädtisch abgestimmten Methode durchgeführt. Die für die Bestimmung verwendeten Variablen weichen in Teilen von denen des ebenfalls vom KOSIS-Verbund angebotenen Datenverarbeitungsverfahrens MigraPro ab. Der Migrationshintergrund ist deshalb nur bedingt vergleichbar.

Ausblick

Derzeit wird das Einwohnermeldewesen in München mit dem Datenverarbeitungsverfahren EWO-PAMELA (Anbieter ekom21) verwaltet. Zum 1.5.2015 erfolgt eine Umstellung auf OK-EWO (Anbieter AKDB). Mit dieser Umstellung ändert sich auch die Schnittstellendatei Bevölkerungsbestand. Dies bietet die Möglichkeit die Bevölkerungsbestands- und Haushaltstatistik zukünftig auf die gleiche Dateigrundlage umzustellen. Eine Verbesserung der interkommunalen Vergleichbarkeit wird durch Einführung der Nutzung von MigraPro zur Bestimmung des Migrationshintergrundes erreicht werden.

Fazit

Mit der Umstellung der Datengrundlage und des Haushaltgenerierungsverfahrens kommt es bei einer Zeitreihenbetrachtung zu einem Bruch. Jedoch konnte mit der neuen Datengrundlage und den dafür neu entwickelten Datenmodellen die inhaltlichen Auswertungsmöglichkeiten der Haushaltsstatistiken wesentlich erweitert und damit ein echter Mehrwert geschaffen werden. Die Nutzung des in der Kommunalstatistik weit verbreiteten HHGen-Programms bietet jetzt zudem eine bessere interkommunale Vergleichbarkeit. Schließlich konnte durch den neu entwickelten Verarbeitungsablauf eine optimierte Qualitätsverfolgung und -sicherung erzielt werden.

1) Jung, Egon: Beschreibung des Verfahrens der Münchener Haushaltsstatistik. Münchner Statistik, Jahrgang 2006, 4. Quartalshft, S. 11ff.- 2) <http://www.staedtestatistik.de/hhstat.html> .