

Informationen zum Mikrozensus 2001 und Hinweise zur Verarbeitung des Setups zum Einlesen der Daten

Inhalt

- Was ist der Mikrozensus?
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Regionale Angaben
- Zugang zum Mikrozensus
- Service-Angebot des German Microdata Lab bei ZUMA
- Hinweise zur Verarbeitung des Setups
- Recodierungen und Missing-Kategorien im Mikrozensus 2001 im Vergleich zum Schluesselverzeichnis
- Hochrechnung, Gewichtung, Vergleich mit Veroeffentlichungen des Statistischen Bundesamtes
- Pruefung des eingelesenen Materials Mikrozensus 2001
- Besonderheiten im Mikrozensus 2001
- Uebersicht der Merkmale des Mikrozensus 2001
- Hinweise auf Materialien zum Mikrozensus

Was ist der Mikrozensus?

Der Mikrozensus ist eine repraesentative 1-Prozent-Bevoelkerungsstichprobe, die bereits seit 1957 in Westdeutschland und seit 1991 auch in den neuen Bundeslaendern erhoben wird. Sie umfasst 2001 Informationen zu bevoelkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 720.000 Personen in 332.000 Haushalten (Fallzahl ohne Anpassung an die aktuelle Bevoelkerungsfortschreibung).

Datenerhebung

Die Daten des Mikrozensus 2001 wurden in der Berichtswoche vom 23. bis 29. April 2001 durch die Statistischen Landesaeemter erhoben. Die Befragung findet im Regelfall muendlich statt, ein kleiner Teil erfolgt mittels schriftlicher Auskunft der Befragten. Fuer alle Beteiligten besteht gemaess der gesetzlichen Regelungen Auskunftspflicht. Nur wenige Fragen sind freiwillig zu beantworten. Der Mikrozensus ist eine Haushaltsstichprobe, d.h. alle Personen in einem Haushalt werden entweder direkt befragt, bzw. eine erwachsene Person gibt Auskunft ueber die weiteren (alle minderjaehrigen) Haushaltsmitglieder. Der Mikrozensus umfasst auch auslaendische Haushalte. Zur Erhebung ihrer Angaben wird der gleiche Fragebogen verwendet wie bei deutschen Haushalten. Die Daten werden schlielich von den Statistischen Landesaeemtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengespielt.

Stichprobe

Fuer den Mikrozensus 2001 gilt das Stichprobendesign von 1990. Die Stichprobe wurde aus dem Datenmaterial der Volkszaehlung 1987 gezogen. Das stichprobenmethodische Grundkonzept des Mikrozensus ist die einstufige Klumpenstichprobe. Die Klumpen sind Auswahlbezirke, die ganze Gebaeude oder Gebaeudeteile umfassen. Zur Verbesserung der Regionalisierbarkeit wurde die durchschnittliche Groee der Auswahlbezirke 1990 verkleinert. Ein Auswahlbezirk umfasst durchschnittlich 9 Wohnungen (vor 1990: 23 Wohnungen). Jaehrlich wird die Stichprobe durch eine Ergaenzungsauswahl der Bautaetigkeitsstatistik aktualisiert, um z.B. Neubauten zu erfassen.

Waehrend fuer das fruehere Bundesgebiet die Grundausswahl auf der Basis der Volkszaehlung 1987 vorgenommen werden konnte, war fuer die neuen Bundeslaender die Volkszaehlung 1981 als Auswahlgrundlage zu alt. Aus diesem Grund wurde in den neuen Bundeslaendern und Ost-Berlin das "Bevoelkerungsregister Statistik" in einem weitgehend analogen Verfahren zur Abgrenzung der Auswahlheiten herangezogen. Im Mikrozensus 2001 sind Informationen zum Auswahlbezirk, der Gebaeudeschicht, Bogenart und der Zahl der Haushalte enthalten.

Erhebungsprogramm

Das Fragenprogramm des Mikrozensus besteht aus einem gleichbleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht. Im Grundprogramm, das sich an alle Haushalte richtet, sind im Mikrozensus 2001 folgende Bereiche enthalten:

- Fragen zur Wohnung/zum Haushalt
- Fragen zu den Personen im Haushalt
- Schulbesuch
- Erwerbsbeteiligung
- Erste gegenwärtige Erwerbstätigkeit
- Zweite Erwerbstätigkeit
- Arbeitsuche/Arbeitsplatzwechsel
- Altersvorsorge
- Ausbildung
- Unterhalt, Einkommen

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms (Auswahlsatz 1%) werden in jährlichen Ergänzungsprogrammen mit einem Auswahlsatz von 0,5% weitere Themen erfragt. Im Ergänzungsprogramm 2001 sind folgende zusätzliche Bereiche enthalten:

- Berufliche und allgemeine Aus- und Fortbildung
- Ergänzende Angaben zur Erwerbstätigkeit
- Frühere Erwerbstätigkeit
- Situation ein Jahr vor der Erhebung

Darüber hinaus gibt es ein Zusatzprogramm mit Merkmalen, die nur im Abstand von 4 Jahren erhoben werden (Auswahlsatz bis zu 0,5%). Der bislang existierende zwei- bzw. dreijährliche Turnus entfällt mit dem Mikrozensusgesetz vom 17.1.1996 ab 1996.

Im Rahmen der vierjährigen Zusatzprogramme enthält der Mikrozensus 2001 folgende Angaben:

- Zusatzangaben zur Erwerbstätigkeit
- Private und betriebliche Altersvorsorge

Bei der Variable EF98 (geringfügige Beschäftigung) hat sich im Vergleich zu dem Scientific Use File des Mikrozensus 1998 infolge des Gesetzes zur Neuregelung der geringfügigen Beschäftigungsverhältnisse (Inkrafttreten am 1.4.1999) die Geringfügigkeitsgrenze auf 630 DM erhöht. Es wird nicht mehr zwischen alten und neuen Bundesländern unterschieden.

Regionale Angaben

Im faktisch-anonymisierten Grundfile sind die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die nicht für die kleinen Bundesländer Bremen und Saarland ausgewiesen ist, enthalten. Mit Hilfe eines speziellen Codes der Variablen Gemeindegrößenklasse kann zwischen West- und Ost-Berlin (EF708=9) unterschieden werden. Die Ausprägungen der Variable Nationalität werden so aggregiert, dass eine Nationalitätengruppe in der Grundgesamtheit wenigstens 50.000 Einwohner umfasst. Bei den übrigen Variablen werden Merkmalsvergrößerungen vorgenommen, damit in den univariaten Randverteilungen jede Ausprägung mindestens 5.000 Fälle in der Grundgesamtheit enthält.

Zugang zum Mikrozensus

Der Mikrozensus 2001 kann von unabhängigen inländischen Forschungseinrichtungen als faktisch anonymisierte 70-Prozent-Substichprobe (Scientific Use file) im Rahmen der Vereinbarung zwischen dem Statistischen Bundesamt und dem BMBF (Bundesministerium für Bildung und Forschung) zu Sonderkonditionen (Bereitstellungsgebühr von jeweils 65 EURO) beim Statistischen Bundesamt erworben werden. Die Mikrozensus 1989, 1991, 1993, 1995 bis 2002 sind zu gleichen Konditionen zugänglich.

Service-Angebot des German Microdata Lab bei ZUMA

Das vorliegende Setup ist ein zentrales Serviceangebot des German Microdata Lab (GML) von ZUMA zum Einlesen der faktisch-anonymisierten Daten des Mikrozensus 2001. Für Interessenten und Nutzer des Mikrozensus besteht die Möglichkeit sich an die Abteilung Mikrodaten zu wenden und folgende Serviceleistungen in Anspruch zu nehmen:

- Beratung, inwieweit der Forschungsplan mit Daten aus dem Mikrozensus zu realisieren ist,
- Beratung und evtl. Unterstützung beim Verfassen des Antrags (Forschungsplans),
- Beratung zu technischen Fragen (EDV-Voraussetzungen, Datensicherheit etc.) und
- Beratung zu Auswertungs- und Analysefragen.

Für weitere Informationen zum Mikrozensus siehe <http://www.gesis.org/Dauerbeobachtung/GML/index.htm>

Hinweise zur Verarbeitung des Setups

Das Stata-Setup beruht auf dem SPSS-Setup und wurde mit Hilfe des Programmes StatTrans automatisch erstellt. Dadurch sind die Variable und Value Labels gekürzt bzw. abgeschnitten. Vollständige Labellisten finden sich z.B. in der Datendokumentation, im Internet und im SPSS-Setup.

WICHTIGER HINWEIS:

Bevor das Stata-Setup korrekt durchgeführt werden kann, ist es nötig, zunächst den Rohdatensatz über einen Texteditor zu öffnen und dort die Dezimalstellenkommata durch Dezimalpunkte automatisch zu ersetzen.

Das Stata-Setup setup_01.do/setup_01.dct erfüllt folgende Aufgaben:

- Einlesen der Rohdaten, wie sie vom StBA bereitgestellt werden
- Umwandeln von alphanumerischen Variablen in numerische Variablen
- Ersetzen von fehlenden Werten (Blanks, @)
- Recodieren der Missing Values
- Umformatierung des Hochrechnungsfaktors
- Hinzufügen der Variable Labels
- Hinzufügen der Value Labels
- Speichern des endgültigen Stata-Systemfile

Das Rohdatenfile des StBA enthält alle Variablen des Mikrozensus 2001. Damit die obigen Stata-Setups für jede mögliche Variablen-Auswahl verwendbar sind, wurden alle Variablen erfasst. Variablen, die Sie nicht bestellt haben, sind in Ihrem Rohdatenfile auf 0 oder BLANK gesetzt. Um ein Systemfile zu erzeugen, das nur die von Ihnen bestellten Variablen enthält, fügen Sie im Stata-Setup (setup_01.do) vor der save-Zeile den Ausdruck `keep + Variablennamen` hinzu und nennen dort Ihre bestellten Variablen.

Beispiel: `keep ef1 ef2
save Mz01.dta`

Recodierungen und Missing-Kategorien im Mikrozensus 2001 im Vergleich zum Schlüsselerzeichnis

Im Setup des Mikrozensus 2001 wurde groeßtenteils auf Recodierungen verzichtet. Recodierungen wurden nur vorgenommen, um den im Rohdatensatz nicht gekennzeichneten Missings einen numerischen Wert zuzuordnen. Die Missing-Kategorien im Schlüsselerzeichnis werden durch das Stata-Setup nicht in Systemmissings umgewandelt. Den Missing-Kategorien werden stattdessen numerische Werte zugeordnet.

Im Folgenden wird die Systematik dieser Recodierungen kurz vorgestellt:

1. Soweit die 0 in den einzelnen Variablen nicht vergeben ist, wird die Missing-Kategorie mit der 0 gekennzeichnet.
2. Ist die 0 besetzt, so wird die Missing-Kategorie mit der 9 oder der 99 gekennzeichnet.
3. Ist auch die 9 oder die 99 in der Variable schon vergeben, so wird die 8 oder 98 gewählt.

Hochrechnung, Gewichtung und Vergleich mit den Veroeffentlichungen des Statistischen Bundesamtes

Sollen die Ergebnisse von Haeufigkeitsauszaehlungen der 70%-Stichprobe mit Veroeffentlichungen des Statistischen Bundesamtes zum Mikrozensus 2001 verglichen werden, muss der entsprechende Hochrechnungsfaktor (Gewicht) des Mikrozensus (EF750, EF751, EF755, EF756) verwendet werden. Vgl. hierzu :

http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_2001/hochrechnungsfaktoren_mz01.htm

- EF750 Personenbezogener Hochrechnungsfaktor
- EF751 Haushaltsfaktor
- EF755 Personenbezogener Hochrechnungsfaktor fuer die disproportionale 0,45%-Unterstichprobe/Ergaenzungsprogramm (Der Auswahlatz der Unterstichprobe liegt auf Regierungsbezirksebene bei 0,4%, 0,6%, 0,8% oder 1% und betraegt im Bundesdurchschnitt 0,45%).
- EF756 EP-Faktor (Haushalts- und Familienfaktor) im Regierungsbezirk

Um auf 1% der Bevölkerung hochzurechnen, wird der Hochrechnungsfaktor mit 100/70 multipliziert. z.B. gen gewicht = ef750/0.7; tabulate <varname> [iw=gewicht]

In den Fachserien des Statistischen Bundesamtes werden die Ergebnisse meist auf die gesamte Bevoelkerung hochgerechnet. Zur Herstellung der Vergleichbarkeit muessen in diesem Fall Ergebnisse eigener Berechnungen noch zusaetzlich zur Hochrechnung und Gewichtung mit 100 multipliziert werden.

Weiterhin sind zum Vergleich mit den vom Statistischen Bundesamt veroeffentlichten Zahlen die jeweiligen Bevoelkerungs-, und Erwerbskonzepte zu beruecksichtigen:

Bevoelkerungs-/ Erwerbskonzept mit Filterung / Selektion:

'Bevoelkerung in Privathaushalten':

(EF506==1)

'Bevoelkerung in Privathaushalten am Ort der Hauptwohnung':

(EF505==1 | EF505==2)

'Erwerbstaetige':

(EF504==1 & (EF505==1 | EF505==2))

'Erwerbslose' nach MZ (ohne Verfuegbarkeitskriterium):

((EF504==2 | EF504==3) & (EF505==1 | EF505==2))

'Erwerbslose' nach ILO (mit Verfügbarkeitskriterium):

(EF504==2 & (EF505==1 | EF505==2))

'Nichterwerbspersonen':

(EF504==4 & (EF505==1 | EF505==2))

'Erwerbstätige (EU-Arbeitskräfteerhebung, 0,45% Substichprobe)':

(EF504==1 & (EF505==1 | EF505==2) & EF738==1)

Pruefung des eingelesenen Materials Mikrozensus 2001

Zur Pruefung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 2001 fehlerfrei eingelesen wurde, kann die folgende Fallzahl (ohne Gewichtung, ohne Selektion) mit der des eingelesenen Datensatzes verglichen werden. Unterscheiden sich die Fallzahlen, weist dies auf einen Fehler beim Einlesen hin. (Anmerkung: Handelt es sich um unterschiedliche Stichproben, koennen sich die Fallzahlen ebenfalls unterscheiden.)

Fallzahl Mikrozensus 2001 (ohne Gewichtung, ohne Selektion)= 503961

Besonderheiten im Mikrozensus 2001

Im Mikrozensus Scientific Use File 2001 wird in der EF512 eine falsche Fallzahl für die Missing-Kategorie "Person in Gemeinschafts-/Anstaltsunterkunft" angegeben. Es werden drei Fälle zuviel ausgewiesen (vgl. EF506). Für die Bandsatzerweiterungen Haushaltsbezugsperson, Bezugsperson in der Familie, Ehefrau der Bezugsperson in der Familie und Lebenspartner der Bezugsperson im Haushalt gibt es jeweils zwei Variablen zum höchsten beruflichen Ausbildungs- oder Hochschul-/Fachhochschulabschluss, die in den Values variieren. Die Ausprägungen der Variablen EF568, EF603, EF619 und EF667 stimmen mit denen der Mikrozensus Scientific Use Files bis zum Jahr 1998 überein. Die Values der Variablen EF570, EF606, EF622 und EF670 entsprechen den Antwortmöglichkeiten im Erhebungsbogen seit dem Jahr 1999. Bei Zeitvergleichen der angesprochenen Variablen mit Mikrozensen bis zum Erhebungszeitpunkt 1998 bieten sich die Variablen EF568, EF603, EF619 und EF667 an, bei Vergleichen mit Mikrozensen seit dem Jahr 1999 dagegen die Variablen EF570, EF606, EF622 und EF670. Die letztgenannten Variablen können zudem so recodiert werden, dass sie mit den entsprechenden Variablen der Mikrozensen bis zum Jahr 1998 korrespondieren.

Uebersicht der Merkmale des Mikrozensus 2001

Erhebungseinheiten:	Haushalte/Personen
Zuordnung Person/Haushalt:	Personen mit mehreren Haushalten werden dem Haushalt zugerechnet, der ihre Hauptwohnung ist. Hauptwohnung eines verheirateten Einwohners, der nicht dauernd von seiner Familie getrennt lebt, ist die vorwiegend benutzte Wohnung seiner Familie.
Erhebungsverfahren:	Interview, zu ca. 80% mündlich und zu ca. 20% schriftlich (Selbstaussueller)
Aufbereitungseinheiten:	Person, Haushalt
Auswahlverfahren:	einstufig geschichtete Stichprobe
Schichtungsmerkmale:	nach Bundesland, Regierungsbezirk, Anpassungsschicht, Regionalschicht, Gebaedeschicht
Auswahleinheiten:	Zaehlbezirke (i.d.R. zusammenliegende Gebaeude bzw. Gebaedeteile); ein Auswahlbezirk verbleibt vier Jahre in

	der Stichprobe, in jedem Jahr scheidet 1/4 der Auswahlbezirke aus (rotierendes Panel)
Stichprobenumfang:	ca. 332.000 Haushalte, ca. 720.000 Personen
Auswahlgrundlage:	Volkszaehlung 1987, unter Berueck- sichtigung der Neubautaetigkeit (alte Bundeslaender), Bevoelkerungsregister Statistik (neue Bundeslaender/Ost-Berlin)
Auswahlsatz:	1 Prozent
Auswahltechnik:	Systematische Ziehung mit Zufallsstart
Scientific Use File:	Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70%-Sub- stichprobe der Haushalte der Mikrozensus- Daten 2001, bei der alle Personen eines ausgewaehlten Haushaltes in das Scientific Use File aufgenommen werden. Gezogen wird die Stichprobe durch a) Anordnen nach Land, Regierungsbezirk, Gemeindegroessenklasse und Haushaltsgroesse, b) Durchnummerieren der Haushalte, sowie c) Ziehen aller Saetze, die nicht die Endziffern 2, 5 und 9 einer ganzzahligen Zufallszahl haben.
Gewicht in der Stichprobe:	Im Grundfile liegen zur Hochrechnung personen-, haushalts- bzw. familien- spezifische Hochrechnungsfaktoren (Gewichte) vor: EF750, EF751, EF755, EF756.(s.o.)
Regionalinformationen:	Die Datei enthaelt an Regional- informationen Angaben zum Bundesland und zur Gemeindegroessenklassen sowie zum Auswahlbezirk.

Hinweise auf Materialien zum Mikrozensus

Informationen zum Mikrozensus und seiner Konzeption:

Emmerling, D.; Riede, T. (1997): 40 Jahre Mikrozensus. *Wirtschaft und Statistik* (3): 160-174.

Heidenreich, H.-J.(1994): Hochrechnung des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis* (S. 112-123). Opladen: Westdeutscher Verlag.

Meyer, K. (1994): Zum Auswahlplan des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis* (S. 106-111). Opladen: Westdeutscher Verlag.

Statistisches Bundesamt (1997): *Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957*. Wiesbaden.

Statistisches Bundesamt (1989): Inhaltliche Fragen bevoelkerungsstatistischer Stichproben am Beispiel des Mikrozensus. Bericht zur Konferenz vom 21. und 22. Oktober 1988. Heft 10 der Schriftenreihe Ausgewaehlte Arbeitsunterlagen zur Bundesstatistik. Wiesbaden.

Hartmann, P.(1989): Der Mikrozensus als Datenquelle fuer die Sozial-wissenschaften. ZUMA-Nachrichten Nr. 24: 6-25.

Schmidt, S. (2000): Erwerbstaetigkeit im Mikrozensus. Konzept, Definition, Umsetzung. ZUMA-Arbeitsbericht 2000/01.

Luettinger, P.; Riede, T. (1997): Der Mikrozensus: amtliche Daten fuer die Sozialforschung, ZUMA-Nachrichten Nr. 41: 19-43.

Zur Weitergabe faktisch anonymisierter Einzeldaten:

Mueller, W./Blien, U./Knoche, P./Wirth, H. (1991): Die Faktische Anonymitaet von Mikrodaten, (Band 19 der Schriftenreihe Forum der Bundesstatistik). Stuttgart: Metzler-Poeschel.

Wirth, H. (1992): Die faktische Anonymitaet von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes. ZUMA-Nachrichten Nr. 30

Zur Fehlerrechnung im Mikrozensus:

Krug, W.; Nourney, M.;Schmidt, J. (1994): Wirtschafts- und Sozialstatistik. Gewinnung von Daten. Muenchen: Oldenbourg (3. voellig neubearb. Auflage) [s. insbes. Seite 242-246].

Ergebnisse des Mikrozensus 2001:

Statistisches Bundesamt (2002): Bevoelkerung und Erwerbstaetigkeit, Fachserie 1, Reihe 3, Haushalte und Familien 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2002): Bevoelkerung und Erwerbstaetigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstaetigkeit 2001 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2002): Bevoelkerung und Erwerbstaetigkeit, Fachserie 1, Reihe 4.1.2, Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstaetigen 2001 (Ergebnisse des Mikrozensus), Stuttgart:Metzler-Poeschel.

Statistisches Bundesamt (2002): Leben und Arbeiten in Deutschland. Ergebnisse des Mikrozensus 2001.

Download unter:

http://www.destatis.de/presse/deutsch/pk/2002/mikrozensus_2001.pdf

Vergleichbarkeit der Mikrozensen:

Kollinger, N.; Kohlmann, A., (2003): Vergleich der Mikrozensen 1998 und 2000. ZUMA-Technischer Bericht Nr. 2003/08.

Lehnert, D.; Kohlmann, A., (2003): Vergleich der Mikrozensen 1997 und 1998. ZUMA-Technischer Bericht Nr. 2003/02.

Schimpl-Neimanns, B.; Schmidt, K., (1998): Vergleichbarkeit der faktisch anonymisierten Mikrozensen 1989, 1991, 1993, 1995. ZUMA-Technischer Bericht Nr. 98/09.

Staudenmaier B., Granato, N.; Schmidt, S. (2000): Vergleich der Mikrozensen 1995 und 1996. ZUMA-Technischer Bericht Nr.2000/08.

Staudenmaier B.; Schmidt, S., (2000): Vergleich der Mikrozensen 1996 und 1997. ZUMA-Technischer Bericht Nr. 2000/09.

Darueber hinaus sind Veroeffentlichungen in der Zeitschrift 'Wirtschaft und Statistik' des Statistischen Bundesamtes zu finden. Ein umfassendes Informationsangebot bietet das GML, ZUMA an unter:

<http://www.gesis.org/Dauerbeobachtung/GML/index.htm>

Publikationen, die auf den Scientific-Use-Files der Mikrozensen basieren, finden sich auf der Webseite:

http://www.gesis.org/Dauerbeobachtung/GML/Service/MZ-SUF-Literatur/MZ_Literatur.htm

Julia H. Schroedter
Zentrum fuer Umfragen, Methoden und Analysen (ZUMA)
German Microdata Lab (GML)
B2,1
68159 Mannheim
Tel: 0621-1246-264
Fax: 0621-1246-100
Email: mikrodaten@zuma-mannheim.de
April 2005
Überarbeitung für Stata: Tobias Hubert, Oktober 2004
