

Informationen zum Mikrozensus 1989 und Hinweise zur Verarbeitung des Setups zum Einlesen der Daten

Inhalt

- Was ist der Mikrozensus?
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Besonderheiten
- Zugang zum Mikrozensus
- Service-Angebot des German Microdata Lab bei ZUMA
- Hinweise zur Verarbeitung des Setups Mikrozensus 1989
- Recodierungen und Missing-Kategorien im Setup Mikrozensus 1989 im Vergleich zum Schlüsselverzeichnis
- Hochrechnung, Gewichtung, Vergleich mit Veröffentlichungen des Statistischen Bundesamtes
- Prüfung des eingelesenen Materials Mikrozensus 1989
- Übersicht der Eigenschaften des Mikrozensus 1989
- Hinweise auf Materialien zum Mikrozensus

Was ist der Mikrozensus?

Der Mikrozensus ist eine repräsentative 1-Prozent-Bevölkerungsstichprobe, die bereits seit 1957 in Westdeutschland und seit 1991 auch in den neuen Bundesländern erhoben wird. Sie umfasst 1989 Informationen zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 600.000 Personen in 250.000 Haushalten.

Datenerhebung

Die Daten des Mikrozensus 1989 wurden in der Berichtswoche vom 24. bis 30. April 1989 durch die Statistischen Landesämter erhoben. Die Befragung findet im Regelfall mündlich statt, ein kleiner Teil erfolgt mittels schriftlicher Auskunft der Befragten. Für alle Beteiligten besteht gemäß 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 über die weiteren (alle minderjährigen) Haushaltsmitglieder. Der Mikrozensus umfasst auch ausländische Haushalte. Zur Erhebung ihrer Angaben wird der gleiche Fragebogen verwendet wie bei deutschen Haushalten. Die Daten werden schließlich von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengespielt.

Stichprobe

Der Mikrozensus 1989 basiert auf dem Stichprobendesign von 1972. Das stichprobenmethodische Grundkonzept des Mikrozensus ist die einstufige Klumpenstichprobe. Die Klumpen sind Auswahlbezirke, die ganze Gebäude oder Gebäudeteile umfassen. Innerhalb dieser Auswahlbezirke werden alle Haushalte befragt. Die durchschnittliche Größe der Klumpen beträgt in Gemeinden mit 20.000 und mehr Einwohnern 30 Haushalte (mit 75 Personen) und in kleineren Gemeinden 20 Haushalte (mit 50 Personen). Für Großgebäude beträgt die durchschnittliche Größe der Auswahlseinheiten 30 Haushalte (mit 75 Personen), in großen Anstalten 60 Personen. Die Auswahlseinheiten wurden auf der Basis der Volks- und Berufszählung 1970 ermittelt, wobei die Neubautätigkeit seit dem Stichtag der Volkszählung berücksichtigt wurden.

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 1989 folgende Bereiche enthalten:

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

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms (Auswahlsatz 1%) gibt es eine Reihe weiterer Merkmale, die mit einem Auswahlsatz bis zu 0,5% in einem Ergänzungsprogramm erhoben werden. Im Ergänzungsprogramm 1989 sind folgende zusätzliche Bereiche enthalten:

- Berufliche Verhältnisse, Tätigkeitsmerkmale
- Aus- und Weiterbildung
- Zusatzangaben für Ausländer
- Private und betriebliche Altersversorgung
- Behinderten Eigenschaften
- Fragen zur Gesundheit
- EG - Arbeitskräfte Stichprobe

Besonderheiten

Im Grundfile sind die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die aus Anonymisierungsgründen z.T. zusammengefasst wurden, enthalten. Bei den übrigen Variablen werden Merkmalsvergrößerungen vorgenommen, damit in den univariaten Randverteilungen jede Ausprägung mindestens 5.000 Fälle in der Grundgesamtheit enthält. Im Gegensatz zu älteren Mikrozensusdaten stehen im Mikrozensus 1989 erstmals auch Haushalts- und Familienbezüge zur Verfügung. Bisher war das Konzept "Wohnberechtigte Bevölkerung" vom Statistischen Bundesamt vorgegeben; seit dem Mikrozensus 1989 kann zwischen unterschiedlichen Grundkonzepten gewählt werden, welches die Analysemöglichkeit erheblich erhöht. Weiterhin kann zwischen Anstaltsbevölkerung und Bevölkerung in Privathaushalten unterschieden werden.

Zugang zum Mikrozensus

Der Mikrozensus 1989 kann von unabhängigen inländischen Forschungseinrichtungen als faktisch anonymisierte 70-Prozent-Substichprobe (Scientific Use File) im Rahmen der Vereinbarung zwischen dem Statistischen Bundesamt und der GESIS zu Sonderkonditionen (Bereitstellungsgebühr von jeweils 65 Euro) beim Statistischen Bundesamt erworben werden. Die Mikrozensus 1991, 1993, 1995 bis 2004 sind zu gleichen Konditionen zugänglich.

Service-Angebot des German Microdata Lab bei ZUMA

Das vorliegende Setup zum Einlesen der faktisch-anonymisierten Daten des Mikrozensus 1989 ist ein zentrales Serviceangebot des German Microdata Lab (GML) von ZUMA. Für Interessenten und Nutzer des Mikrozensus besteht die Möglichkeit sich an das GML 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.

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

Hinweise zur Verarbeitung des Setups Mikrozensus 1989

Das Stata-Setup beruht auf dem SPSS-Setup und wurde mit Hilfe des Programms 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.

Das Stata-Setup setup_89.do 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 1989. Damit das obige Stata-Setup für jede mögliche Variablen-Auswahl verwendbar ist, wurden alle Variablen erfasst. Variablen, die Sie nicht bestellt haben, sind in Ihrem Rohdatenfile auf 0 oder BLANK gesetzt. Um einen Systemfile zu erzeugen, der nur die von Ihnen bestellten Variablen enthält, fügen Sie im Stata-Setup (setup_89.do) vor der save-Zeile den Ausdruck keep + Variablenamen hinzu und nennen dort Ihre bestellten Variablen.

Beispiel: keep ef1 ef2
save Mz89.dta

Stata benötigt für temporäre Operationen ca. das Dreifache des Plattenplatzes des Rohdatenfiles. D.h. es sollten ca. 500 MB Plattenplatz zur Verfügung stehen.

Recodierungen und Missing-Kategorien im Setup Mikrozensus 1989 im Vergleich zum Schlüsselverzeichnis

Im Setup des Mikrozensus 1989 wurden von ZUMA zum Teil Recodierungen vorgenommen, die zu Abweichungen in der numerischen Bezeichnung der Value Labels im Vergleich zum Schlüsselverzeichnis führen. Im Folgenden werden die wichtigsten Regeln, die hinter diesen Recodierungen im Setup stehen, benannt:

1. Die meisten Recodierungen im Setup wurden vorgenommen, um eine einheitliche Kennzeichnung der Missing-Kategorien zu erreichen. Diese sind im Falle der Ausprägung 'Entfällt' mit 9; 99 oder 999 und im Falle der Ausprägung 'Angabe fehlt/keine Angabe' mit 8; 98 und 998 gekennzeichnet. Das hat zur Folge, dass der Ausprägung 'Nein' - im Schlüsselverzeichnis in der Regel mit der Zahl 9 versehen - eine Zahl zugewiesen wurde, die sich an die anderen Ausprägungen anschließt. Die 0 für 'Angabe fehlt/keine Angabe' wurde durch die 8, 98 oder 998 ersetzt.
2. Weitere Recodierungen gibt es, wenn sich innerhalb der Zahlenfolgen Lücken ergaben (z.B. 1, 2, 4, 5, 6 und 9 wurde zu 1, 2, 3, 4, 5 und 9).
3. Zum Teil wurden Recodierungen vorgenommen um die Merkmalsausprägungen in eine "logische Folge" zu bringen. Das betrifft die Variablen EF108-EF110 und EF220 und EF221. So wurde zum Beispiel bei der Variable "Stellung im Beruf" (EF110) die Ausprägung 'Direktor, Amtsleiter, Geschäftsführer oder Betriebsleiter/Werksleiter' von 0 (Schlüsselverzeichnis) auf 10 (Setup) recodiert, so dass sie sich an die Ausprägung 'Abteilungsleiter, Prokurist' (9) anschließt und nicht der Ausprägung 'Auszubildender, Praktikant, Volontär' (1) vorangestellt ist.
4. In einigen Variablen wurde die Gruppe der Personen in Gemeinschaftsunterkünften, die als Missing geführt wird, von 0 auf einen anderen numerischen Wert gesetzt, damit sie sich am Ende der Häufigkeitsauszählung wiederfindet. Dies ist bei den Variablen, welche die Haushalte und die Familien betreffen, der Fall.
5. Zahlenfolgen in der Form von 0, 1, 2, 3... wurden in die Form von 1, 2, 3, 4... umgewandelt. Dies betrifft die Variable "Schulbesuch" (EF56), alle Variablen, die auf der Frage nach der "Stellung im Beruf" basieren (EF94, EF115, EF194, EF217, EF227), und die Variablen zum "sonstigen öffentlichen und privaten

Einkommen" (EF144 und EF145). Die Regeln gelten nicht für alle im Setup aufgeführten Variablen. Als Grundlage für das vorliegende Setup diente ein Setup, das zunächst für das ZUMA-File (d.h. für eine bestimmte Merkmalsauswahl aus dem Mikrozensus Grundfile) erstellt wurde. Zu einem späteren Zeitpunkt wurde das Setup um die restlichen Variablen des Mikrozensus Grundfiles ergänzt. Dieses enthält nun alle im Mikrozensus Grundfile verfügbaren Variablen und wird von ZUMA für alle Nutzer bereitgestellt. Bei den ergänzten Variablen wurde möglichst auf Recodierungen verzichtet. Fehlende Werte wurden bei diesen Variablen wie folgt codiert: Sofern die 0 nicht besetzt war, wurden Missings mit 0 codiert. War die 0 vergeben, so wurde auf die 9; 99 oder 999 zurückgegriffen. Falls diese Zahlen ebenfalls besetzt waren, wurden die Missings mit 8 oder 10 gekennzeichnet. Betroffen sind die Variablen zur Kranken- und Rentenversicherung (EF78, EF82, EF83, EF86), zum Thema Verwendung von Jodsalz (EF87, EF89) zur Schichtarbeit und den geleisteten Arbeitsstunden in der Berichtswoche (EF101, EF111, EF112, EF117, EF118, EF128), Fragen zur Krankheit, Unfall und Behinderung (EF166U2 - EF166U8, EF167U1, EF167U2) sowie zum Raucherverhalten (EF137, EF166U9, EF166U11).

Hochrechnung, Gewichtung und Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes

Sollen die Ergebnisse von Häufigkeitsauszählungen der 70%-Stichprobe mit Veröffentlichungen des Statistischen Bundesamtes zum Mikrozensus 1989 verglichen werden, muss der Hochrechnungsfaktor (Gewicht) des Mikrozensus (EF257U4: "Hochrechnungsfaktor der Bezugsperson (ggf. der Person in Gemeinschaftsunterkunft)" verwendet werden. Um auf 1% der Bevölkerung hochzurechnen wird der Hochrechnungsfaktor mit 100/70 multipliziert:

```
gen gewicht=EF257U4/0.7; tabulate <varname> [iw=gewicht]
```

Bei Auszählungen von Fragen zur Gesundheit und Behinderteneigenschaft, die bei einer 0,5% Stichprobe erhoben wurden, sowie den Fragen zur privaten und betrieblichen Altersvorsorge (0,25% Stichprobe) ist bei der Hochrechnung zusätzlich der Auswahlsatz der Unterstichprobe zu berücksichtigen:

```
gen gewicht=EF257U4 /0.7/0.5; tabulate <varname> [iw=gewicht]   bzw.  
gen gewicht=EF257U4/0.7/0.25; tabulate <varname> [iw=gewicht]
```

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

Weiterhin sind zum Vergleich mit den vom Statistischen Bundesamtes veröffentlichten Zahlen die jeweiligen Bevölkerungs- und Erwerbskonzepte (z.B. tabulate <varname> if (...)) zu berücksichtigen:

'Bevölkerung in Privathaushalten':

```
ef27==1
```

'Bevölkerung am Ort der Hauptwohnung':

```
Ef26==1 | ef26==2
```

'Privathaushalte':

```
ef27==1 & ef33==1
```

'Bevölkerung am Familienwohnsitz':

```
ef28==1
```

'Familien':

```
ef28==1 & ef30==1
```

'Erwerbstätige':

```
(ef34>=1 & ef34<=4 | ef34>=10) & (ef26==1 | ef26==2)
```

'Erwerbslose' nach MZ (ohne Verfügbarkeitskriterium):

```
ef34>=5 & ef34<=7 & (ef26==1 | ef26==2)
```

'Erwerbslose' nach ILO (mit Verfügbarkeitskriterium):

```
(ef34>=5 | ef34<=7) & (ef26==1 | ef26==2) & (ef68==1)
```

'Nichterwerbspersonen':

```
(ef34==8 | ef34==9) & (ef26==1 | ef26==2)
```

Prüfung des eingelesenen Materials Mikrozensus 1989

Zur Prüfung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 1989 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, können sich die Fallzahlen ebenfalls unterscheiden.)

Fallzahl Mikrozensus 1989 (ohne Gewichtung, ohne Selektion)= **385831**

Übersicht der Eigenschaften des Mikrozensus 1989

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, in der Regel mündlich
Aufbereitungseinheiten:	Person, Haushalt
Auswahlverfahren:	Einstufig geschichtete Stichprobe
Schichtungsmerkmale:	nach Bundesland, Gemeindegrößenklasse, Straßenart, (Anstalten, Großgebäude)
Auswahleinheiten:	Segment (Straßenabschnitt, räumlich benachbarte Wohnungen oder Häuser, nach Anfangsbuchstaben gruppierte Familiennamen); ein Auswahlbezirk verbleibt vier Jahre in der Stichprobe, in jedem Jahr scheidet 1/4 der Auswahlbezirke aus (rotierendes Panel)
Stichprobenumfang:	ca. 250.000 Haushalte, ca. 600.000 Personen
Auswahlgrundlage:	Volkszählung 1970, unter Berücksichtigung der Neubautätigkeit
Auswahlsatz:	1 Prozent
Auswahltechnik:	Zufallsauswahl
Hochrechnungsverfahren:	Freie Hochrechnung, Multiplikation mit 100. Die Anpassung an die Eckwerte der Bevölkerungsfortschreibung erfolgt durch Doppeln oder Streichen von Personen getrennt nach den Merkmalen Geschlecht, Nationalität (Deutsche/Ausländer) und Soldat/Wehrpflichtiger auf der Ebene der Regierungsbezirke.
Scientific Use File:	Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70%-Substichprobe der Haushalte der Mikrozensus-Daten 1989, bei der alle Personen eines ausgewählten Haushaltes in das Scientific Use File aufgenommen werden. Gezogen wird die Stichprobe durch a) Anordnen nach Land, Regierungsbezirk, Gemeindegrößenklasse und Haushaltsgröße, b) Durchnummerieren der Haushalte, sowie c) Ziehen aller Sätze, die nicht die Endziffern 2, 5 und 9 einer ganzzahligen Zufallszahl haben.
Gewicht in der Stichprobe:	Im Grundfile liegt zur Hochrechnung ein Hochrechnungsfaktoren (Gewicht) vor: EF257U4.
Regionalinformationen:	Die Datei enthält an Regionalinformationen Angaben zum Bundesland und zur Gemeindegrößenklassen.

Hinweise auf Materialien zum Mikrozensus:

Informationen zum Mikrozensus und seiner Konzeption:

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

Hartmann, Peter: Der Mikrozensus als Datenquelle für die Sozialwissenschaften. In: ZUMA-Nachrichten Nr. 24, Mai 1989, S. 6-25

Lüttinger, Paul; Riede, Thomas: Der Mikrozensus: amtliche Daten für die Sozialforschung. In: ZUMA-Nachrichten Nr. 41, November 1997, S. 19-43

Beschreibung der Erstellung der faktisch anonymisierten Einzeldaten (Grundfile):

Müller, Walter; Blien, Uwe; Knoche, Peter; Wirth, Heike u.a.: Die faktische Anonymität von Mikrodaten, Band 19 der Schriftenreihe Forum der Bundesstatistik, hrsg. v. Statistischem Bundesamt, Metzler-Poeschel, Stuttgart, 1991

Weitergabe faktisch anonymisierter Einzelangaben aus dem Mikrozensus 1989. In: Statistisches Bundesamt (Hrsg.): Methoden, Verfahren, Entwicklungen, 2/1993

Wirth, Heike: Die faktische Anonymität von Mikrodaten. Ergebnisse und Konsequenzen eines Forschungsprojektes. In: ZUMA-Nachrichten Nr.30, Mai 1992, S. 7-65

Beschreibung der Stichprobenziehung:

Krug, Walter; Nourney, Martin; Schmidt, Jürgen: Wirtschafts- und Sozialstatistik. München: Oldenbourg, 1994 (3., völlig neu bearbeitete Auflage)

Nourney, Martin: Stichprobenplan des Mikrozensus ab 1972. In: Wirtschaft und Statistik 11, 1973, S.631-638.

Ergebnisse des Mikrozensus 1989:

Fachserie 1: Bevölkerung und Erwerbstätigkeit

Reihe 4.1.1: Stand und Entwicklung der Erwerbstätigkeit

Reihe 4.1.2: Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen

Reihe 2 enthält Informationen über Ausländer

Reihe 3 umfasst zudem Ergebnisse über Privathaushalte, Kinder und Familien

Darüber hinaus sind Veröffentlichungen in der Zeitschrift 'Wirtschaft und Statistik' des Statistischen Bundesamtes zu finden:

Erwerbstätigkeit im April 1989. Ergebnis des Mikrozensus (Heidenreich); Heft 7, 1990, S. 451-465.

Dienstleistungsberufe im Produzierenden Gewerbe. Ergebnis des Mikrozensus (Riede), Heft 9, 1990, S. 623-631.

Entwicklung der Erwerbstätigkeit nach Berufen. Ergebnis des Mikrozensus 1989 (Cornelsen), Heft 11, 1990, S. 768-774 (430*-431*).

Gesundheitszustand der Bevölkerung. Ergebnis des Mikrozensus 1989 (Gräß), Heft 2, 1991, S. 108-115.

Erwerbstätige mit Hochschulabschluss nach Hauptfachrichtungen (Cornelsen), Heft 5, 1991, S. 320-328 (194*-197*).

Gesundheitsrisiko "Rauchen". Ergebnis des Mikrozensus 1989 (Brückner), Heft 5, 1991, S. 341-352.

Beruf und Tätigkeitsmerkmale der Erwerbstätigen. Ergebnis des Mikrozensus 1989 (Cornelsen), Heft 8, 1991, S. 531-537 (306*-308*).

Autoren:

Hansjörg Frenzel, Paul Lüttinger, Anke Nau (Februar 1994)

Karen Schmidt (Überarbeitung; Februar 1999)

Bernhard Schimpl-Neimanns (Ergänzungen und Tests; Juni 1999)

Helga Christians (Überarbeitung; September 2003).

Tobias Hubert (Ergänzungen zu STATA, August 2004)

Yvonne Lechert (Überarbeitung, Juli 2006)

Zentrum für Umfragen, Methoden und Analysen (ZUMA)

German Microdata Lab

Postfach 12 21 55

68072 Mannheim

Tel.: 0621 / 1246-265 Fax: 0621 / 1246-100

E-Mail: mikrodaten@zuma-mannheim.de
