

## Informationen zum Mikrozensus 2000 und Hinweise zur Verarbeitung des SAS-Setups zum Einlesen der Daten

### Inhalt

- Was ist der Mikrozensus?
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Besonderheiten
- Zugang zum Mikrozensus
- Service-Angebot des German Microdata Labs (GML) bei ZUMA
- Hinweise zur Verarbeitung des Setups
- Hochrechnung, Gewichtung, Vergleich mit Veröffentlichungen des Statistischen Bundesamtes
- Prüfung des eingelesenen Materials MZ2000
- Hinweis auf Probleme im MZ2000**
- Übersicht der Merkmale des Mikrozensus
- 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 2000 Informationen zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 719.000 Personen in 329.000 Haushalten (Fallzahl ohne Anpassung an die aktuelle Bevölkerungsfortschreibung).

---

### Datenerhebung

Die Daten des Mikrozensus 2000 wurden in der Berichtswoche vom 8. bis 14. Mai 2000 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 zusammengestellt.

---

### Stichprobe

Für den MZ98 gilt das Stichprobendesign von 1990. Die Stichprobe wurde aus dem Datenmaterial der Volkszählung 1987 gezogen. Das stichprobenmethodische Grundkonzept des Mikrozensus ist die einstufige Klumpenstichprobe. Die Klumpen sind Auswahlbezirke, die ganze Gebäude oder Gebäudeteile umfassen. Zur Verbesserung der Regionalisierbarkeit wurde die durchschnittliche Größe der Auswahlbezirke 1990 verkleinert. Ein Auswahlbezirk umfasst durchschnittlich 9 Wohnungen (vor 1990: 23 Wohnungen). Jährlich wird die Stichprobe durch eine Ergänzungsauswahl der Bautätigkeitsstatistik aktualisiert, um z.B. Neubauten zu erfassen.

Während für das frühere Bundesgebiet die Grundausswahl auf der Basis der Volkszählung 1987 vorgenommen werden konnte, war für die neuen Bundesländer die Volkszählung 1981 als Auswahlgrundlage zu alt. Aus diesem Grund wurde in den neuen Bundesländern und Ost-Berlin das „Bevölkerungsregister Statistik“ in einem weitgehend analogen Verfahren zur Abgrenzung der Auswahlseinheiten herangezogen. Im MZ2000 sind Informationen zum Auswahlbezirk, der Gebäudeschicht, 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 2000 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 2000 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 2000 folgende Angaben:

- Zusatzangaben zur beruflichen Ausbildung
- Pendlereigenschaft, -merkmale
- Zusatzangaben für Ausländer
- Zusatzangaben zur Erwerbstätigkeit

**Nutzer der Scientific Use Files der Mikrozensus 1996 - 1998 werden darauf hingewiesen, dass sich aufgrund einer Änderung im Erhebungsprogramm seit dem MZ1999 vor allem in Bezug auf die Variablen zu (Aus- und Fort-)Bildungsabschlüssen und Umschulungen (EF288-EF302, EF568, EF570, EF603, EF606, EF619, EF622, EF667 und EF670) deutliche Änderungen in den Variablenamen und den value labels im MZ2000 ergeben haben. Eine direkte Übertragung von Analysesyntaxen, die auf der Basis der Scientific Use Files der Mikrozensus 1996 - 1998 entwickelt wurden, ist daher nicht möglich. Informationen zu den Veränderungen der Variablen finden sich im [ZUMA-Methodenbericht 2003/08](#).**

**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 2000 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€) beim Statistischen Bundesamt erworben werden. Die Mikrozensus 1989, 1991, 1993, 1999-2002 sind zu gleichen Konditionen zugänglich.

---

## Service-Angebot des German Microdata Lab (GML) bei ZUMA

Das vorliegende Setup ist ein zentrales Serviceangebot des German Microdata Labs (GML) von ZUMA zum Einlesen der faktisch-anonymisierten Daten des Mikrozensus 2000. Für Interessenten und Nutzer des Mikrozensus besteht die Möglichkeit sich an das German Microdata Lab (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.

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

---

## Hinweise zur Verarbeitung des Setups

Das SAS-Setup **setup\_00.sas** erfüllt folgende Aufgaben:

- Einlesen der Rohdaten, wie sie vom StBA bereitgestellt werden
- Umwandeln von alphanumerischen Variablen in numerische Variablen
- Ersetzen von fehlenden Werten (Blanks, @)
- Hinzufügen der Variable Labels
- Hinzufügen der Value Labels

Das Rohdatenfile des StBA enthält alle Variablen des Mikrozensus 2000. Damit das obige Setup 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 eine SAS-Datei zu erzeugen, der nur die von Ihnen bestellten Variablen enthält, können Sie in einem weiteren Datenschnitt Ihre spezifische Variablenauswahl sichern.

---

## 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 2000 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\\_2000/hochrechnungsfaktoren\\_mz00.htm](http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_2000/hochrechnungsfaktoren_mz00.htm)

EF750 Personenbezogener Hochrechnungsfaktor

EF751 Haushaltsfaktor

EF755 Personenbezogener Hochrechnungsfaktor für die disproportionale 0,45%-Unterstichprobe/Ergänzungsprogramm (Der Auswahlatz der Unterstichprobe liegt auf Regierungsbezirksebene bei 0,4%, 0,6%, 0,8% oder 1% und beträgt 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: Beispiel bei personenbezogenen Auswertungen:

```
compute gewicht=EF750 * 100/70  
weight by 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 Bundesamt veröffentlichten Zahlen die jeweiligen Bevölkerungs-, und Erwerbskonzepte zu berücksichtigen:

'Bevölkerung in Privathaushalten':

select if (EF506=1)

'Bevölkerung in Privathaushalten am Ort der Hauptwohnung':

select if (EF505=1 or EF505=2)

'Erwerbstätige' :

select if (EF504=1 and (EF505=1 or EF505=2))

'Erwerbslose' nach MZ (ohne Verfügbarkeitskriterium):

select if (EF504=2 or EF504=3) and (EF505=1 or EF505=2)

'Erwerbslose' nach ILO (mit Verfügbarkeitskriterium):

select if (EF504=2 and (EF505=1 or EF505=2))

'Nichterwerbspersonen':

select if (EF504=4 and (EF505=1 or EF505=2))

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

select if (EF504=1 and (EF505=1 or EF505=2) and EF738=1)

---

### **Prüfung des eingelesenen Materials MZ2000**

Zur Prüfung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 2000 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 MZ2000 (ohne Gewichtung, ohne Selektion)= **503185**

---

### **Hinweis auf Problem im Mikrozensus 2000: EF512**

#### **EF53:**

Die Variable EF53 (Zuzugsjahr) umfasst im MZ2000 die ausländische Bevölkerung und die deutsche Population, letztere jedoch nur im Rahmen der Erhebung der EU-Arbeitskräfteerhebung. Alle Personen, die keine deutsche Staatsbürgerschaft besitzen, weisen gültige Werte bei der Variable EF53 auf. Personen, die entweder nur die deutsche Staatsbürgerschaft haben oder die deutsche Staatsbürgerschaft als 2. Staatsbürgerschaft besitzen, haben - insofern sie in der EU-Arbeitskräfteerhebung enthalten sind (ef738=1) - ebenfalls gültige Werte. Wenn Deutsche oder Ausländer mit 2. deutscher Staatsbürgerschaft nicht in der EU-Arbeitskräfteerhebung enthalten sind, besitzen sie keinen gültigen Wert bei der Variable EF53. Diese Unterscheidung resultiert aus einer unterschiedlichen Reihenfolge der entsprechenden Fragen in den verschiedenen Fragebögen (Erhebungsbogen 1 vs. Erhebungsbogen 1+E). Im Fragebogen der EU-Arbeitskräfteerhebung (Erhebungsbogen 1+E) wird zunächst das Einreisedatum erfragt und danach erst die Staatsbürgerschaft (daher liegen dort auch für Deutsche Informationen zum Einreisedatum vor). Bei Personen, die nicht in der Substichprobe enthalten sind (Erhebungsbogen 1), wird dagegen erst nach der Staatsbürgerschaft und dann nach dem Einreisedatum gefragt (daher liegt in dieser Fragebogenversion für Personen, die die deutsche Staatsbürgerschaft als 1. oder 2. Staatsbürgerschaft haben, keine Information über das Einreisedatum vor).

Folglich kann im MZ2000 für Personen mit deutscher Staatsbürgerschaft, die in der EU-Arbeitskräfteerhebung enthalten sind, festgestellt werden, ob es sich um Personen handelt, die im Laufe ihres Lebens nach Deutschland eingewandert sind (in diesem Fall ist das Einreisedatum erfasst), vormals also eine andere Staatsbürgerschaft oder eine zusätzliche Staatsbürgerschaft hatten, oder ob es sich um Personen handelt, die von Geburt an ausschließlich die deutsche Staatsbürgerschaft besaßen.

**EF512:**

Bei einer Typisierung im Mikrozensus 2000 wurde eine Person fälschlicherweise als Person in Gemeinschafts-/Anstaltsankünften verkodet (EF512). Diese Fehlzuordnung kann anhand der Variablen EF4, EF712, EF506 und EF512 identifiziert werden. Der problematische Fall weist die folgende Kombination von Charakteristika auf:

Fehlverkodung bei EF512, wenn die folgende Merkmalskombination zutrifft (1 Fall):  
(EF506=1) & (EF4=27) & (EF712=6) & (EF512=0).

---

**Übersicht der Merkmale des Mikrozensus 2000:**

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 (Selbstaussfüller)
Aufbereitungseinheiten:	Person, Haushalt
Auswahlverfahren:	Einstufig geschichtete Stichprobe
Schichtungsmerkmale:	nach Bundesland, Regierungsbezirk, Anpassungsschicht, Regionalschicht, Gebäudeschicht
Auswahleinheiten:	Zählbezirke (i.d.R. zusammenliegende Gebäude bzw. Gebäudeteile); ein Auswahlbezirk verbleibt vier Jahre in der Stichprobe, in jedem Jahr scheidet 1/4 der Auswahlbezirke aus (rotierendes Panel)
Stichprobenumfang:	ca. 330.000 Haushalte, ca. 730.000 Personen
Auswahlgrundlage:	Volkszählung 1987, unter Berücksichtigung der Neubautätigkeit (alte Bundesländer), Bevölkerungsregister Statistik (neue Bundesländer/Ost-Berlin)
Auswahlsatz:	1 Prozent
Auswahltechnik:	Systematische Ziehung mit Zufallsstart
<b>Scientific Use File</b>	Der Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70%-Substichprobe der Mikrozensus-Daten 2000.
Gewicht in der Stichprobe:	Im Grundfile liegen zur Hochrechnung personen-, haushalts- bzw. familienspezifische Hochrechnungsfaktoren (Gewichte) vor: EF750, EF751, EF755, EF756. (s.o.)
Regionalinformationen:	Die Datei enthält an Regionalinformationen Angaben zum Bundesland und zur Gemeindegrößenklassen sowie zum Auswahlbezirk.

---

**Hinweise auf Materialien zum Mikrozensus:**Informationen zum Mikrozensus und seiner Konzeption:

Emmerling, Dieter/Riede, Thomas (1997): 40 Jahre Mikrozensus. Wirtschaft und Statistik (3):160-174.

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

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

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

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

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

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

Lüttinger, P./Riede, T.: Der Mikrozensus: amtliche Daten für die Sozialforschung, ZUMA-Nachrichten Nr. 41, November 1997: 19-43. [PDF-Datei / zipped PDF-Datei]

Zur Weitergabe faktisch anonymisierter Einzeldaten:

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

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

Zur Fehlerrechnung im Mikrozensus:

Statistisches Bundesamt (2001): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstätigkeit 2000 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Krug, W./Nourney, M./Schmidt, J.: (1994): Wirtschafts- und Sozialstatistik. Gewinnung von Daten. München: Oldenbourg (3. völlig Neubearb. Auflage) [s. insbes. Seite 242-246]

Ergebnisse des Mikrozensus 2000:

Statistisches Bundesamt (2001): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 3, Haushalte und Familien 2000 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2001): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstätigkeit 2000 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (2001): Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.2, Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen 2000 (Ergebnisse des Mikrozensus), Stuttgart: Metzler-Poeschel.

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.

Darüber hinaus sind Veröffentlichungen in der Zeitschrift 'Wirtschaft und Statistik' des Statistischen Bundesamtes zu finden. Ein umfassendes Informationsangebot bietet das German Microdata Lab (GML), ZUMA an: <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](http://www.gesis.org/Dauerbeobachtung/GML/Service/MZ-SUF-Literatur/MZ_Literatur.htm)

-----

Autorin: Dr. Annette Kohlmann  
Zentrum für Umfragen, Methoden und Analysen (ZUMA)  
German Microdata Lab (GML)  
B2,1  
68159 Mannheim  
Tel: 0621-1246-371  
Fax: 0621-1246-100  
Email: [mikrodaten@zuma-mannheim.de](mailto:mikrodaten@zuma-mannheim.de)

---