

Informationen zum Mikrozensus 1995 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
- Recodierungen und Missing-Kategorien im Mikrozensus 1995 im Vergleich zum Schlüsselverzeichnis
- Besonderheiten des Setups Mikrozensus 1995
- Hochrechnung, Gewichtung, Vergleich mit Veröffentlichungen des Statistischen Bundesamtes
- Prüfung des eingelesenen Materials
- Übersicht der Eigenschaften des Mikrozensus
- Hinweise auf Materialien zum Mikrozensus

Was ist der Mikrozensus?

Der Mikrozensus ist eine repräsentative 1-Prozent-Bevölkerungstichprobe, die bereits seit 1957 in Westdeutschland erhoben wird. Im Jahre 1991 wurden erstmals auch die neuen Bundesländer mit einbezogen. Die Stichprobe umfasst 1995 Informationen zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 800.000 Personen in 350.000 Haushalten.

Datenerhebung

Die Daten des Mikrozensus 1995 wurden in der Berichtswoche vom 24. bis 30. April 1995 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 zusammengespült.

Stichprobe

Mit dem Mikrozensus 1990 wurde ein neuer Stichprobenplan eingeführt. Das stichprobenmethodische Grundkonzept des Mikrozensus ist nach wie vor die einstufige Klumpenstichprobe. Innerhalb eines Auswahlbezirkes werden alle Haushalte befragt. Grundlage für die Auswahlbezirke sind ganze Gebäude oder bei größeren Gebäuden Gebäudeteile. Die Gebäude der Auswahlgrundlage wurden dabei nach der Zahl ihrer Wohnungen in drei Größenklassen bzw. Schichten eingeteilt, zusätzlich wurden Gebäude mit vermuteten Gemeinschaftsunterkünften einer Sonderschicht zugeordnet. Im Gegensatz zum früheren Stichprobenplan ergeben sich für die einzelnen Schichten eine geringere Durchschnittsgröße, im Jahre 1989 umfasste eine Schicht noch durchschnittlich 23 Wohnungen, im Auswahlplan 1990 nur noch durchschnittlich 9 Wohnungen. Zusätzlich wurde eine stärkere regionale Schichtung eingeführt. Als regionale Schichten wurden Raumeinheiten von durchschnittlich ca. 350.000 Einwohnern herangezogen. In den alten Bundesländern ergaben sich 170, in den neuen Bundesländern 31 regionale Schichten. Auf Grund dieser Umstellung auf stärkere regionale Schichtung und die Reduzierung der Klumpengröße sind die Angaben nun stärker regionalisierbar und auch unterhalb der Regierungsbezirksebenen aussagekräftig. In den alten Bundesländern basiert die Grundausswahl auf den Daten

der Volkszählung 1987; in den neuen Bundesländern und Ostberlin wurden hingegen die Daten des "Bevölkerungsregisters Statistik" herangezogen. Die Neubautätigkeit wird entsprechend berücksichtigt.

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 1995 folgende Bereiche enthalten:

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

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 1995 sind folgende zusätzliche Bereiche enthalten:

- Aus- und Weiterbildung
- Zusatzangaben für Ausländer
- Wohnsituation der Haushalte
- Fragen zur Gesundheit
- Behinderteneigenschaften
- Private und betriebliche Altersvorsorge
- EG - Arbeitskräfte Stichprobe

Besonderheiten

Im Grundfile sind die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die aus Anonymisierungsgründen z.T. zusammengefasst wurden, enthalten. 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 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 1989, 1991, 1993, 1996 - 2002 sind zu gleichen Konditionen zugänglich.

Service-Angebot des German Microdata Lab bei ZUMA

Das vorliegende Setup zum Einlesen der faktisch-anonymisierten Daten des Mikrozensus 1995 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

Das SPSS-Setup setup95.sps 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 und Deklarieren der Missing Values
- Umformatierung des Hochrechnungsfaktors
- Hinzufügen der Variable Labels
- Hinzufügen der Value Labels
- Speichern des endgültigen SPSS-Systemfile

Das Rohdatenfile des StBA enthält alle Variablen des Mikrozensus 1995. Damit das obige SPSS-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 SPSS-Setup (setup95.sps) in der letzten save outfile - Anweisung die Option /keep = (Variablenamen) hinzu und nennen dort Ihre bestellten Variablen.

Beispiel: SAVE OUTFILE = 'name.sav'
/keep=EF1 EF2

SPSS 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 Mikrozensus 1995 im Vergleich zum Schlüsselverzeichnis

Im Setup des Mikrozensus 1995 wurde größtenteils auf Recodierungen verzichtet. Recodierungen wurden nur vorgenommen, um den im Rohdatensatz nicht gekennzeichneten Missings einen numerischen Wert zu zuordnen. Im Folgenden wird die Systematik dieser Recodierungen kurz vorgestellt:

1. Soweit die 9 bzw. 99 oder 999 in den einzelnen Variablen nicht vergeben ist, wird die Missing-Kategorie mit diesen Zahlen bezeichnet.
2. Ist die 9 bzw. 99 oder 999 besetzt, so wird die Missing-Kategorie mit der 0 gekennzeichnet.
3. Ist auch die 0 in der Variable schon vergeben, so wird die 8 bzw. die 98 oder 998 gewählt.

Lediglich in den Variablen EF221 und EF231 (Arbeitszeittyp) wurden weitere Recodierungen vorgenommen. Hier wird die Lücke innerhalb der Zahlenfolge geschlossen (1,2,3,4,5,7,8,9,0 und "leer" wurde zu 1,2,3,4,5,6,7,8,0 und 9 als Missing-Kategorie)

Besonderheiten des Setups Mikrozensus 1995

Der Mikrozensus 1995 weist drei Besonderheiten auf:

1. Im Gegensatz zu Scientific Use Files (SUF) aus anderen Jahrgängen sind im SUF 1995 oftmals die Kategorien "ohne Angabe" und "entfällt" in eine Missing-Kategorie zusammengefasst.
2. Es finden sich im Datensatz des SUF 1995 bei einigen Variablen einzelne Fälle, die nicht einer der vorhandenen Kategorien zugeordnet sind. Diese werden, soweit möglich, im Setup direkt mit der Missing-Kategorie "Entfällt" zusammengefasst oder aber extra als Missing ausgewiesen.
3. In der Bandsatzerweiterung EF228 "überwiegender Lebensunterhalt, Ehefrau der Familienbezugsperson" befinden sich im Rohdatenfile die zusätzlichen Ausprägungen "8" und "9", die laut Schlüsselverzeichnis mit der Kategorie "1 = Erwerbstätigkeit, Berufstätigkeit" zusammengefasst sein sollten. In den Kategorien "8" und "9" befinden sich entsprechend auch nur erwerbstätige / berufstätige Personen. Für die Analyse empfiehlt es sich daher die drei Kategorien "1", "8" und "9" zusammenzufassen.

In der Bandsatzerweiterung EF226 "Wirtschaftszweig der Ehefrau der Familienbezugsperson" gibt es geringe Abweichungen zur Variable EF92 "Wirtschaftszweig". Einige Personen, die in der Variablen EF92 eine gültige Angabe haben, sind in der EF226 der Kategorie "entfällt" zugeordnet. Zur Behebung dieses Problems vergleiche: Schimpl-Neimanns, B./Schmidt, K., 1997: Konstruktion von Bandsatz-Erweiterungen zum Haushalt und zur Familie im Mikrozensus mittels satzübergreifenden Operationen in SPSS. ZUMA-Technischer Bericht Nr. 97/08.

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 1995 verglichen werden, muss der Hochrechnungsfaktor (Gewicht) des Mikrozensus (EF254: "Hochrechnungsfaktor (Haushalts-, Familienfaktor) in der regionalen Anpassungsschicht") verwendet werden. Um auf 1% der Bevölkerung hochzurechnen, wird der Hochrechnungsfaktor mit 100/70 multipliziert:

```
compute gewicht=EF254 * 100/70
weight by 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:

```
compute gewicht=EF257U4 * (100/70 * 100/50)
weight by gewicht
```

bzw.

```
compute gewicht=EF257U4 * (100/70 * 100/25)
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 Bundesamtes veröffentlichten Zahlen die jeweiligen Bevölkerungs- und Erwerbskonzepte zu berücksichtigen:

'Bevölkerung in Privathaushalten':

```
select if (EF27=1)
```

'Bevölkerung am Ort der Hauptwohnung':

```
select if (EF26=1 or EF26=2)
```

'Privathaushalte':

```
select if (EF27=1) and (EF33=1)
```

'Bevölkerung am Familienwohnsitz':

```
select if (EF28=1)
```

'Familien':

```
select if (EF28=1) and (EF30=1)
```

'Erwerbstätige':

```
select if (EF34>=1 and EF34<=4) and (EF26=1 or EF26=2)
```

'Erwerbslose' nach MZ (ohne Verfügbarkeitskriterium):

```
select if (EF34>=5 and EF34<=7) and (EF26=1 or EF26=2)
```

'Erwerbslose' nach ILO (mit Verfügbarkeitskriterium):

```
select if (EF34>=5 and EF34<=7) and (EF26=1 or EF26=2) and (EF68=1)
```

'Nichterwerbspersonen':

```
select if (EF34=8 or EF34=9) and (EF26=1 or EF26=2))
```

Prüfung des eingelesenen Materials

Zur Prüfung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 1995 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 1995 (ohne Gewichtung, ohne Selektion)= 512509

Übersicht der Eigenschaften des Mikrozensus 1995

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, 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. 350.000 Haushalte, ca. 800.000 Personen
Auswahlgrundlage:	Volkszählung 1987 (alte Bundesländer), Bevölkerungsregister Statistik (neue Bundesländer/Ost-Berlin), Aktualisierung unter Berücksichtigung der Neubautätigkeit
Auswahlsatz:	1 Prozent
Auswahltechnik:	Systematische Ziehung mit Zufallsstart
Hochrechnungsverfahren:	In einem zweistufigen Verfahren werden zunächst die bekannten Ausfälle kompensiert und dann in einem zweiten Schritt die Stichprobenergebnisse gebunden an Eckzahlen aus der laufenden Bevölkerungsfortschreibung hochgerechnet. Diese Hochrechnung erfolgt dabei grundsätzlich auf der Ebene von bundesweit 123 sogenannten regionalen Anpassungsschichten. Das sind regionale Einheiten mit einer derzeitigen Durchschnittsgröße von etwa 660.000 Einwohnern. Die Anpassungsklassen werden dabei gebildet durch die Angaben über die Zahl von Deutschen und Ausländern in der Gliederung nach Geschlecht. Die Anpassung für Soldaten erfolgt dagegen getrennt auf Regierungsbezirksebene.
Scientific Use File:	Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70% Substichprobe der Haushalte der Mikrozensus-Daten 1995, bei der alle Personen eines ausgewählten Haushaltes in das Scientific Use File aufgenommen werden. Gezogen wird diese 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 Hochrechnungsfaktor (Gewicht) vor: EF254
Regionalinformationen:	Die Datei enthält an Regionalinformationen Angaben zum Bundesland und zur Gemeindegrößenklassen.

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.

Greiner, Ulrich (1996): Erste Ergebnisse des Mikrozensus April 1995. *Wirtschaft und Statistik* (5): 304-312.

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*, Wiesbaden, Mai 1990.

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.

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 (1995): *Fachserie 1, Reihe 4.1.1. Stand und Entwicklung der Erwerbstätigkeit*, S. 18-22.

Ergebnisse des Mikrozensus 1995:

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

Statistisches Bundesamt (1996): *Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1, Stand und Entwicklung der Erwerbstätigkeit 1995 (Ergebnisse des Mikrozensus)*, Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt (1996): *Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.2, Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen (Ergebnisse des Mikrozensus)*, Stuttgart: Metzler-Poeschel.

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

Matthias Engelhard, Karen Schmidt, Ramona Winkler, April 1997

Vollständige Überarbeitung: Helga Christians, Oktober 2003

Ergänzungen: Julia H. Schroedter, August 2004

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
