

Informationen zum Mikrozensus 1993 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 1993 im Vergleich zum Schlüsselverzeichnis
- Besonderheiten des Setups Mikrozensus 1993
- Hochrechnung, Gewichtung, Vergleich mit Veröffentlichungen des Statistischen Bundesamtes
- Prüfung des eingelesenen Materials
- Übersicht der Eigenschaften des Mikrozensus 1993
- Hinweise auf Materialien zum Mikrozensus

Was ist der Mikrozensus?

Der Mikrozensus ist eine repräsentative 1-Prozent-Bevölkerungsstichprobe, die bereits seit 1957 in Westdeutschland. Im Jahre 1991 wurden erstmals auch die neuen Bundesländer mit einbezogen. Die Stichprobe umfasst 1993 Informationen zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 800.000 Personen in 350.000 Haushalten.

Datenerhebung

Die Daten des Mikrozensus 1993 wurden in der Berichtswoche vom 19. bis 25. April 1993 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

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 wurde dabei nach der Zahl ihrer Wohnung in drei Größenklassen bzw. Schichten eingeteilt, zusätzlich wurden Gebäude mit vermuteten Gemeinschaftsunterkünften einer Sonderschicht zugeordnet. Im Gegensatz zum früheren Stichprobenplan ergebe sich für die einzelnen Schichten eine geringer 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 "Bevölkerungsregister 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 1993 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 1993 sind folgende zusätzliche Bereiche enthalten:

- Aus- und Weiterbildung
- Zusatzangaben für Ausländer
- Wohnsituation der Haushalte
- Pendlereigenschaften/-merkmale
- 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, 1995 - 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 1993 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 setup93.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 1993. 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 (setup93.sps) in der letzten save outfile - Anweisung die Option /keep = (Variablennamen) 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 1993 im Vergleich zum Schlüsselverzeichnis

Im Setup des Mikrozensus 1993 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, die 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) und Variablen, die auf der Frage nach der "Stellung im Beruf" basieren (EF94, EF115, EF194, EF217, EF227) sowie 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 (EF76, EF78, EF82, EF83), zur Schichtarbeit und den geleisteten Arbeitsstunden in der Berichtswoche (EF101, EF111, EF112, EF128).

Besonderheiten des Setups Mikrozensus 1993

Der Mikrozensus 1993 weist zwei Besonderheiten auf:

1. Der vom Statistischen Bundesamt gelieferte Rohdatensatz enthielt 886 Sätze von Personen, die nicht tatsächlich befragt wurden und zu denen keine Angaben vorliegen; die Variablen enthalten nur fehlende Werte. Diese Ausfallsätze wurden bei der Erstellung des SPSS-Systemfiles ausgeschlossen (SELECT IF NOT MISSING(EF26))
2. Durch einen Programmfehler im Statistischen Bundesamt wurde die Angabe zum Eheschließungsjahr (EF39, EF214) '1990' gelöscht und statt dessen der Wert 'ohne Angabe' vergeben.

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 1993 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
```

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 1993 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 1993 (ohne Gewichtung, ohne Selektion)= 513830

Übersicht der Eigenschaften des Mikrozensus 1993

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 (neü Bundesländer/Ost-Berlin), Aktualisierung unter Berücksichtigung der Neubautätigkeit
Auswahlsatz:	1 Prozent
Auswahltechnik:	Systematische Ziehung mit Zufallsstart
Hochrechnungsverfahren:	Gebundene Hochrechnung (Anpassung) mit Eckzahlen aus der laufenden Bevölkerungsfortschreibung pro regionaler Anpassungsschicht nach den Merkmalen Deutscher / Ausländer und Geschlecht; für Zeit- und Berufssoldaten sowie Wehrpflichtige auf Regierungsbezirksebene
Scientific Use File:	Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70%-Substichprobe der Haushalte der Mikrozensus-Daten 1993, bei der alle Personen eines ausgewählten Haushalts 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 Hochrechnungsfaktoren (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:

- Cornelsen, Carsten (1994): Entwicklung der Erwerbstätigkeit nach Wirtschaftsbereichen und Berufen. In: *Wirtschaft und Statistik*, 12/94. S. 991-997
- Emmerling, Dieter/Riede, Thomas (1994): Zur Freiwilligkeit in der Auskunftserteilung im Mikrozensus. In: *Wirtschaft und Statistik*, 6/94. S. 435-449
- Gruber, Stefan/Riede, Thomas (1994): Erste Ergebnisse der Mikrozensus 1993. In: *Wirtschaft und Statistik*, 7/94. S. 518-533
- Hartmann, P.: Der Mikrozensus als Datenquelle für die Sozialwissenschaften. In: *ZUMA-Nachrichten* Nr. 24, Mai 1989, S.6-25
- 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
- Lüttinger, Paul/Riede, Thomas (1997): Der Mikrozensus. Amtliche Daten für die Sozialforschung. *ZUMA-Nachrichten* Nr.41, S. 19-43
- 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
- Niemeyer, Frank/Voit, Hermann (1995): Lebensformen der Bevölkerung 1993. In: *Wirtschaft und Statistik*, 6/95. S. 437-445
- Riede, Thomas/Emmerling, Dieter (1994): Analysen zur Freiwilligkeit der Auskunftserteilung im Mikrozensus. In: *Wirtschaft und Statistik*, 9/94. S. 733-742
- Schimpl-Neimanns, Bernhard (1998): Analysemöglichkeiten des Mikrozensus. *ZUMA-Nachrichten* Nr.42, S. 91-119
- Schimpl-Neimanns, Bernhard/Schmidt, Karen (1997): Konstruktion von Bandsatz-Erweiterungen zum Haushalt und zur Familie im Mikrozensus mittels satzübergreifender Operationen in SPSS ZUMA - Technischer Bericht 97/08
- Schmidt, Karen/Schimpl-Neimanns, Bernhard (1998): Vergleichbarkeit der faktisch anonymisierten Mikrozensusen 1989, 1991, 1993 und 1995ZUMA - Technischer Bericht 98/09
- Statistisches Bundesamt (Hrsg.): *Arbeitsunterlagen zum Mikrozensus*, Wiesbaden, September 1997
- Statistisches Bundesamt (Hrsg.): *Inhaltliche Fragen bevölkerungsstatistischer Stichproben am Beispiel des Mikrozensus*. Bericht zur Konferenz vom 21. und 22. Oktober 1988. Heft 10 der Schriftenreihe *Ausgewählte Arbeitsunterlagen zur Bundesstatistik*. Wiesbaden 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* In: *ZUMA-Nachrichten* Nr. 30, Mai 1992, S.7-65

Zur Fehlerrechnung im Mikrozensus:

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

Ergebnisse des Mikrozensus 1993:

- Fachserie 1, Statistisches Bundesamt (Hrsg.), 1993: *Bevölkerung und Erwerbstätigkeit*
Daraus: Reihe 3: Haushalte und Familien
Reihe 4.1.1: Stand und Entwicklung der Erwerbstätigkeit
Reihe 4.1.2: Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen

Yvonne Lechert, Karen Schmidt, Bernhard Schimpl-Neimanns; April 1996

Bernhard Schimpl-Neimanns (Ergänzungen; Juni 1999)

Überarbeitung:

- Änderungen der VARIABLE und VALUE LABELS

- Ergänzung der VARIABLE LABELS um die Fragebogen-Nummern. Die Fragennummern werden in Klammern angegeben. Wenn nicht anders bezeichnet, beziehen sich die Fragen-Nr. auf den Bogen '1+E'; Bogen '2' enthält die Fragen ohne Auskunftspflicht; 'E' bezieht sich auf die spezifischen Fragen der EG-Arbeitskräftestichprobe

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
