

Informationen zur Nutzung des Setups des Mikrozensus Scientific Use Files 1973

Der vorliegende Text soll Ihnen dabei helfen, den durch das Forschungsdatenzentrum der Statistischen Landesämter (FDZ) bereit gestellten Rohdatensatz, den Sie im ASCII-Format zusammen mit dem SPSS-Setup (*setup73.sps*) erhalten haben, korrekt einzulesen. Zudem wird auf „Besonderheiten“ des Mikrozensus Scientific Use File (SUF) 1973 hingewiesen.

Ausführliche Informationen über den Mikrozensus und die Mikrozensus Scientific Use Files werden auf den Websites des Zentrums für Umfragen, Methoden und Analysen (ZUMA) bereitgestellt:

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

Weitere Informationen sind auf den Seiten des Statistischen Bundesamtes abrufbar:

http://www.destatis.de/themen/d/thm_mikrozen.php

Das German Microdata Lab (GML) im ZUMA bietet für den Mikrozensus 1973 zahlreiche Informationen an: Erhebungsbogen, Schlüsselverzeichnis, thematische Übersicht, relevante Klassifikationen u.m. Auf der folgenden Website finden Sie zudem einen ZUMA-Methodenbericht zur Dokumentation und Aufbereitung des Mikrozensus SUF 1973:

http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/MZ_1973/index.htm.

Das vom GML bereitgestellte SPSS-Setup für das Mikrozensus Scientific Use File 1973 (*setup73.sps*) dient zum Einlesen des Rohdatenmaterials und zum Erstellen eines SPSS-Systemfiles. Es beinhaltet Programmanweisungen zum Umwandeln von nicht-numerischen Variablen in numerische Variablen, zur Definition von fehlenden Werten sowie zum Versehen der Variablen und ihrer Ausprägungen mit entsprechenden Labeln.

Um eine reibungsfreie Aufbereitung der Daten zu gewährleisten, empfiehlt es sich, die in dem Setup vorgesehenen Voreinstellungen zu berücksichtigen und je nach Leistung des verarbeitenden Systems das Setup schrittweise durchführen zu lassen. Sind die Daten eingelesen und ist das entsprechende Systemfile erstellt, kann die folgende Fallzahl (ohne Gewichtung, ohne Selektion) zur Kontrolle, ob der Rohdatensatz fehlerfrei eingelesen wurde, mit der des eingelesenen Datensatzes verglichen werden. Unterscheiden sich die Fallzahlen, weist dies auf einen Fehler beim Einlesen hin.

Fallzahl Mikrozensus 1973 (ohne Gewichtung, ohne Selektion) = 448.366

Des Weiteren können die Verteilungen ausgewählter Variablen des Mikrozensus SUF 1973 zur Prüfung des fehlerfreien Ablaufs des Setups herangezogen werden (vgl.:

http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_1973/randverteilungen_mz73.htm).

Als Referenz zur Plausibilitätsprüfung der Mikrozensus Scientific Use Files dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 1973 ist dies die Fachserie A, Bevölkerung und Kultur, Reihe 6: Entwicklung der Erwerbstätigkeit 1973.

Die Plausibilitätsprüfung des Mikrozensus SUF 1973 erfolgte durch den Vergleich des aufbereiteten Datensatzes mit den vom Statistischen Bundesamt veröffentlichten Tabellen in den entsprechenden Reihen der Fachserie. Dabei wurde nach den vom Statistischen Bundesamt verwendeten Bevölkerungs- und Erwerbskonzepten gewichtet und selektiert. (Zur Abgrenzung und Hochrechnung der Bevölkerungsbegriffe im Mikrozensus 1973 siehe:

http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_1973/hochrechnungsfaktoren_mz73.htm)

Bedingt durch die Ziehung der 70%-Substichprobe ergeben sich Abweichungen zwischen den Häufigkeiten der Variablen des Mikrozensus Scientific Use Files und den in den Fachserien des Statistischen Bundesamtes veröffentlichten Zahlen (beruhend auf den Original-Mikrozensusdaten). Die meisten Variablen weichen nur in geringem Maße (meist 0% bis 1%, max. 5 %) von den veröffentlichten Daten ab. Abweichungen über 5 Prozent wurden bei Variablen festgestellt, deren Merkmalsausprägungen mit sehr geringen Fallzahlen besetzt sind (z.B. Staatsangehörigkeit und Wirtschaftsbereiche).

Weiterhin wurden verschiedene generierte Merkmale (sog. Bandsatzergänzungen) des Mikrozensus 1973 auf ihre Plausibilität hin überprüft.

Im Mikrozensus SUF 1973 sind Familienvorstände enthalten, die unter 15 Jahre alt sind. Diese sollten bei Analysen mit entsprechenden Bandsatzergänzungen gegebenenfalls ausgeschlossen werden.

Zentrum für Umfragen, Methoden und Analysen (ZUMA)
German Microdata Lab
B2,1
68159 Mannheim
Tel: 0621-1246-265
Fax: 0621-1246-100
<http://www.gesis.org/Dauerbeobachtung/GML/index.htm>

Kontakt: Paul Lüttinger, Yvonne Lechert (GML)
Email: mikrodaten@zuma-mannheim.de

Weitere Informationen zum Mikrozensus 1973

Inhalt

- Was ist der Mikrozensus?
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Der faktisch anonymisierte Mikrozensus - Das Scientific Use File -
- Zugang zum Mikrozensus
- Service-Angebot des German Microdata Lab bei ZUMA
- Hinweise zur Verarbeitung des Setups
- Recodierungen und Missing-Kategorien im Mikrozensus SUF 1973 im Vergleich zum Schlüsselverzeichnis
- Hochrechnung, Gewichtung, Vergleich mit Veröffentlichungen des Statistischen Bundesamtes
- Prüfung des eingeleseenen Materials Mikrozensus SUF 1973
- Übersicht der Merkmale des Mikrozensus 1973

Was ist der Mikrozensus?

Der Mikrozensus ist mit einem Auswahlsatz von 1% der Bevölkerung die größte jährliche Haushaltsbefragung in Europa. Sie wird seit 1957 in Westdeutschland und seit 1991 in den neuen Bundesländern durchgeführt. Grundlage des Mikrozensus ist die gesamte wohnberechtigte Bevölkerung in Deutschland. Der Mikrozensus ist durch eine hohe Kontinuität im Frageprogramm und aufgrund der Auskunftspflicht für die meisten Fragen durch einen niedrigen Unit-Nonresponse gekennzeichnet. Der Mikrozensus ermöglicht differenzierte Analysen auf Personen-, Familien- und Haushaltsebene sowohl im Quer- als auch im Längsschnitt und eignet sich für die Beobachtung von Strukturveränderungen sowie zu tiefgegliederten Analysen.

Datenerhebung

Die Angaben im Mikrozensus 1973 beziehen sich auf die Berichtswoche vom 6.-12. Mai 1973. Die Daten des Mikrozensus werden durch die Statistischen Landesämter erhoben. Die Befragungen finden i.d.R. mündlich statt, ein kleiner Teil erfolgt mittels schriftlicher Auskunft der Befragten. Für alle Beteiligten besteht gemäß der gesetzlichen Regelungen, Auskunftspflicht. Ausnahmen hierzu sind 1973 einige Fragen zum Zusatzprogramm, die jedoch im Scientific Use File nicht enthalten sind (Fragen zu Urlaubsreisen, Krankheiten und Unfälle). 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 (sog. Proxy Interviews). Der Mikrozensus umfasst auch ausländische Haushalte. Zur Erhebung ihrer Angaben wird der gleiche Fragebogen verwendet wie bei deutschen Haushalten. Die Daten werden von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengestellt.

Stichprobe

Für den Mikrozensus 1973 galt der Stichprobenplan von 1972 (Flächenstichprobe auf der Basis der Volkszählung von 1970). Das stichprobenmethodische Grundkonzept des Mikrozensus 1973 ist eine einstufige Klumpenstichprobe. Die bebaute Fläche der Bundesrepublik wurde auf der Grundlage der Wohnungszählung von 1970, die im Rahmen der Volkszählung stattfand, in etwa gleich große Flächen (Auswahlbezirke bzw. Segmente) aufgeteilt. Die Neubautätigkeit wurde dabei berücksichtigt. 1 Prozent dieser Auswahlbezirke gelangen in die Stichprobe. Innerhalb eines Auswahlbezirkes wurden alle Haushalte befragt. Grundlage für die Auswahlbezirke sind ganze Gebäude oder bei größeren Gebäuden Gebäudeteile. Die Gebäude der Auswahlbezirke wurden dabei nach der Zahl ihrer Wohnung in drei Größenklassen bzw. Schichten eingeteilt, zusätzlich wurden Gebäude mit vermuteten Gemeinschaftsunterkünften einer Sonderschicht zugeordnet. Ein Auswahlbezirk umfaßte durchschnittlich 23 Wohnungen.

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 1973 folgende Bereiche enthalten:

- Fragen zum Haushalt/Wohnung
- Fragen zu den Personen im Haushalt
- Staatsangehörigkeit
- Vertriebenen/Flüchtlingsstatus
- Schulbesuch
- Erwerbsbeteiligung
- Jahr der Eheschließung
- Zahl und Alter der Kinder im HH und Familie
- Erste gegenwärtige Erwerbstätigkeit
- Zweite Erwerbstätigkeit
- Arbeitsuche/Arbeitsplatzwechsel
- Unterhalt, Einkommen
- Altersvorsorge
- Krankenversicherung

Als Zusatzprogramm wurden 1973 folgende Merkmale erfaßt:

mit einem Auswahlsatz von 1%:

- Urlaubs- und Erholungsreisen
- Vermögenswirksame Leistungen
- Behinderungen von Kindern
- Krankheiten und Unfälle

mit einem Auswahlsatz von 0,1%:

- Arbeitszeit am Wochenende
- Nacht- und Sonntagsarbeit
- Ausbildungsabsichten der Eltern für ihre Kinder
- Krankheiten und Unfälle
- berufliche Mobilität der Frauen
- Einkaufsmöglichkeiten und Einkaufsverhalten in Großstädten
- Ziel und Dauer von Kurzreisen

Im Scientific Use File des MZ73 sind die Fragen des Zusatzprogramms nicht enthalten.

Der faktisch anonymisierte Mikrozensus – Das Scientific Use File

Bei den vorliegenden Daten des Mikrozensus Scientific Use File (SUF) 1973 handelt es sich um eine faktisch anonymisierte 70-Prozent-Substichprobe des Mikrozensus 1973. Die faktisch anonymisierte Substichprobe wurde vom Forschungsdatenzentrum der Statistischen Landesämter als systematische Zufallsauswahl aus dem Originalmaterial gezogen. Das Material wurde dazu nach Bundesland, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen in Privathaushalten, Auswahlbezirksnummer und Nummer der Wohnung sortiert. Durch diese Anordnung ist gewährleistet, dass die Substichprobe hinsichtlich dieser Merkmale nur geringe zufallsbedingte Abweichungen aufweist. In diesem Prozess wurden Gemeinschaftsunterkünfte wie Wohnungen behandelt, das Merkmal „Zahl der Personen in Privathaushalten“ wurde nicht zur Ziehung der Substichprobe herangezogen. Bei Gemeinschaftsunterkünften erhielt jede Person eine eigene fortlaufende („Wohnungs“-)Nummer. Zur Stichprobenziehung diente die letzte (Einer-)Stelle der Wohnungsnummer. Alle Wohnungen, deren Einerstelle einer von drei zufällig gezogenen einstelligen Zahlen entsprach, gelangten nicht in die Stichprobe (30% der Wohnungen).

Weiterhin unterscheidet sich das Mikrozensus SUF vom Originalfile des Mikrozensus dadurch, dass bestimmte Variablen im Scientific Use File, bedingt durch die Anonymisierung, in klassierter und vergrößerter Form verfügbar sind.

So sind im Mikrozensus SUF die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse enthalten. Auch bei den weiteren Variablen des Scientific Use Files werden je nach Bedarf Merkmalsvergrößerungen vorgenommen, so dass jede Ausprägung in den univariaten Verteilungen des Originalmaterials mindestens 5.000 Fälle in der Grundgesamtheit aufweist. Die Ausprägungen der Variable Nationalität werden so aggregiert, dass jede Nationalitätengruppe in der Grundgesamtheit wenigstens 50.000 Einwohner umfasst.

Zugang zum Mikrozensus

Der Mikrozensus 1973 kann von unabhängigen inländischen Forschungseinrichtungen als faktisch anonymisierte 70-Prozent-Substichprobe (Scientific Use File) im Rahmen der Vereinbarung zwischen den Statistischen Ämtern und dem BMBF (Bundesministerium für Bildung und Forschung) zu Sonderkonditionen (Bereitstellungsgebühr von jeweils 65 EURO) bei den beiden Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder für konkrete Forschungsvorhaben bezogen werden.

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 1973. 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 des Mikrozensus zu realisieren ist
- Beratung und evtl. Unterstützung beim Verfassen des Antrags (Forschungsplans)
- Beratung zu technischen Fragen (EDV-Voraussetzungen, Datensicherheit etc.)
- Beratung zu Auswertungs- und Analysefragen

Hinweise zur Verarbeitung des Setups

Das SPSS-Setup *setup73.sps* erfüllt folgende Aufgaben:

- Einlesen der Rohdaten, wie sie vom StBA bereitgestellt werden
- Umwandeln von alphanumerischen Variablen in numerische Variablen
- Recodieren und Deklarieren der Missing Values
- Hinzufügen der Variable Labels
- Hinzufügen der Value Labels
- Speichern des endgültigen SPSS-Systemfiles

Das Rohdatenfile enthält alle Variablen des Mikrozensus SUF 1973. 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 (*setup73.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`

Recodierungen und Missing-Kategorien im Mikrozensus SUF 1973 im Vergleich zum Schlüsselverzeichnis

Im Setup des Mikrozensus 1973 wurde größtenteils auf Recodierungen verzichtet. Recodierungen wurden nur vorgenommen, um den im Rohdatensatz nicht gekennzeichneten Missings einen numerischen Wert zuzuordnen.

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 Veröffentlichungen des Statistischen Bundesamtes

Um von der 70% Substichprobe auf 1% der Bevölkerung hochzurechnen muss mit folgendem Faktor gewichtet werden:

```
compute gewicht=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:

http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_1973/hochrechnungsfaktoren_mz73.htm

Prüfung des eingelesenen Materials Mikrozensus SUF 1973

Zur Prüfung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 1973 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 1973 (ohne Gewichtung, ohne Selektion)= **448.366**

Überblick: Mikrozensus 1973

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, mündlich, schriftlich
Aufbereitungseinheiten	Person, Haushalt
Auswahlverfahren	Einstufig geschichtete Stichprobe
Schichtungsmerkmale	nach Bundesland, Regierungsbezirk, Gebäudeschicht, Gemeindegrößenklasse
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	232 Tsd. Haushalte, 619 Tsd. Personen
Auswahlgrundlage	Volkszählung 1970
Auswahlsatz	1 Prozent
Auswahltechnik	Systematische Ziehung mit Zufallsstart
Scientific Use File	Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70%-Substichprobe der Wohnungen der Mikrozensus-Daten 1973, bei der alle Personen einer ausgewählten Wohnung in das Scientific Use File aufgenommen werden. Gezogen wird die Stichprobe durch a) Anordnen nach Land, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen in Privathaushalten und Auswahlbezirksnummer, b) Durchnummerieren der Wohnungen, sowie c) Ziehen aller Sätze, die nicht die Endziffern 2, 5 und 9 einer ganzzahligen Zufallszahl haben (0,1,3,4,6,7,8).
Regionalinformationen	Die Datei enthält an Regionalinformationen Angaben zum Bundesland und zu Gemeindegrößenklassen. Die Gemeindegrößenklassen wurden entsprechend der Anonymisierungsregeln so zusammengefaßt, dass in jedem Bundesland jede Größenklasse mindestens 400.000 Einwohner enthält.