

<b>Wichtige Informationen zur Nutzung der Setups für das Mikrozensus Scientific Use File 1987</b>
---

Diese Informationen sollen dabei helfen, das Scientific Use File sachgerecht auszuwerten. Dazu werden vom German Microdata Lab (GML) der GESIS–Leibniz-Institut für Sozialwissenschaften die Setups dreier Statistikprogramme (SPSS, SAS, STATA) bereitgestellt.

Ausführliche Informationen über den Mikrozensus und die Mikrozensus Scientific Use Files (SUF) enthalten die [Web-Seiten der GESIS](#). Weitere Informationen sind auf den [Web-Seiten der Forschungsdatenzentren](#) der Statistischen Ämter des Bundes und der Länder abrufbar.

Die vom GML bereitgestellten Setups für das Mikrozensus Scientific Use File 1987 dienen zum Einlesen des Rohdatenmaterials und zum Erstellen von Systemfiles der Statistikprogramme SPSS, SAS und Stata im dokumentierten Format.

*Spezifika der Statistikprogramme:* In den Originaldaten werden fehlende Werte i. d. R. nicht differenziert und nicht gesondert mit einem numerischen Wert ausgewiesen, sondern als „leer“ (blank) codiert und somit von den Auswertungsprogrammen als sogenannte System-Missings behandelt. In den Scientific Use Files wurde dies weitgehend übernommen. Jedoch wurden in diesen Fällen numerische Werte zugewiesen (z. B. „9“ oder „99“), die in SPSS als benutzerdefinierte fehlende Kategorie deklariert werden konnten. Die Definition benutzerdefinierter fehlender Werte ist in SAS und STATA ohne entsprechende Recodierung allerdings nicht möglich. Die SAS- und STATA-Setups enthalten stattdessen die Recodierung in System-Missings („.“), die nach Entfernen der diesen Block einschließenden Kommentarzeichen („/\*“ und „\*/“) aktiviert werden können.

Um eine reibungsfreie Aufbereitung der Daten zu gewährleisten, empfiehlt es sich, die in den Setups vorgesehenen Voreinstellungen zu berücksichtigen. Ist das entsprechende Systemfile erstellt, kann die einfache Fallzahl  $n = 438.823$  (ohne Gewichtung, ohne Selektion) zur Kontrolle, ob der Rohdatensatz fehlerfrei eingelesen wurde, mit der des erstellten Datensatzes verglichen werden. Unterscheiden sich die Fallzahlen, weist dies meist auf einen Fehler beim Einlesen hin.

Des Weiteren können die Verteilungen ausgewählter Variablen des Mikrozensus SUF 1987 zur Prüfung des fehlerfreien Ablaufs des Setups herangezogen werden (vgl.: <http://www.gesis.org/missy/studie/daten/randverteilungen/randverteilung-1987/>).

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 1987 sind dies die Reihe 3 (Haushalte und Familien),

die Reihe 4.1.1 (Stand und Entwicklung der Erwerbstätigkeit) und die Reihe 4.1.2 (Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen) der Fachserie 1.

Die Plausibilitätsprüfung des Mikrozensus SUF 1987 erfolgte durch den Vergleich des aufbereiteten Datensatzes mit den vom Statistischen Bundesamt veröffentlichten Tabellen in den entsprechenden Reihen der Fachserie 1. Dabei wurde nach den vom Statistischen Bundesamt verwendeten Bevölkerungs- und Erwerbskonzepten gewichtet und selektiert (vgl. „Zur Abgrenzung und Hochrechnung der Bevölkerungsbegriffe und zur Hochrechnung im Mikrozensus SUF 1987“ auf der Seite <http://www.gesis.org/missy/missy-home/auswahl-datensatz/mikrozensus-1987/>).

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 % 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 1987 auf ihre Plausibilität hin überprüft. Dabei zeigten sich keine Inkonsistenzen.

Bei der Aufbereitung des Mikrozensus Scientific Use Files 1987 wurden die Variablenlabels nach der neuen, erstmals für das Mikrozensus Scientific Use File 2003 verwendeten Systematik erstellt. Die Variablenlabels enthalten die Nummer der entsprechenden Frage im Selbstausfüllerbogen, den grundsätzlichen inhaltlichen Bezug der Variable und eventuelle Spezifizierungen; außerdem sind Freiwilligkeit und Zugehörigkeit zur Unterstichprobe gekennzeichnet. Ein im Vergleich mit früheren Files abweichendes Label bedeutet also i. d. R. nicht, dass sich am Inhalt der entsprechenden Variable etwas geändert hat.

Geschlechtsspezifische Berufsbezeichnungen sind in den Setups zum Mikrozensus 1987 - im Unterschied zu den Setups früherer Erhebungen - ausschließlich aus Platzgründen und zur besseren Lesbarkeit in der männlichen Form genannt.

---

GESIS - Leibniz-Institut für Sozialwissenschaften  
German Microdata Lab  
B2, 1  
68159 Mannheim  
Tel.: 0621-12 46-265  
Fax: 0621-12 46-100  
<http://www.gesis.org/gml>  
Kontakt: Andreas Herwig, Bernhard Schimpl-Neimanns  
E-Mail: [gml@gesis.org](mailto:gml@gesis.org)

---

## Weitere Informationen zum Mikrozensus

### Inhalt

- Was ist der Mikrozensus?
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Besonderheiten
- Zugang zum Mikrozensus
- Service-Angebot des German Microdata Lab der GESIS
- Hinweise zur Verarbeitung des Setups
- Hochrechnung, Gewichtung, Vergleich mit amtlichen Veröffentlichungen
- Prüfung des eingelesenen Materials
- Übersicht der Eigenschaften des Mikrozensus 1987
- Hinweise auf Materialien zum Mikrozensus

### Was ist der Mikrozensus?

Der Mikrozensus ist eine repräsentative 1-Prozent-Bevölkerungsstichprobe, die seit 1957 erhoben wird. Die Erhebung von 1987 umfasst Informationen zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten von ca. 61,5 Mio. Personen in ca. 27 Mio. Haushalten.

---

### Datenerhebung

Wegen der Volkszählung 1987 wurde für den Mikrozensus 1987 die Berichtswoche auf den 23. bis 29. März 1987 vorverlegt. Die Berichtswoche liegt damit im Vergleich zu den vorangegangenen Mikrozensusen (letzte feiertagsfreie Woche im April) an einem früheren Zeitpunkt im Jahr. Dies muss bei Querschnittvergleichen mit anderen Mikrozensusen wegen eventuell auftretender saisonaler Schwankungen berücksichtigt werden. Der Mikrozensus 1987 wurde durch die Statistischen Landesämter erhoben. Die Befragung fand mündlich oder schriftlich statt. Für alle Beteiligten bestand gemäß der gesetzlichen Regelungen (Mikrozensusgesetz 1985) Auskunftspflicht. Der Mikrozensus 1987 ist eine Haushaltsstichprobe, d. h. alle Personen in einem Haushalt wurden entweder direkt befragt, oder eine volljährige Person gab Auskunft über die weiteren (alle minderjährigen) Haushaltsmitglieder. Die Daten wurden von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengespielt.

---

### Stichprobe

Für den Mikrozensus 1987 galt der Stichprobenplan von 1972 (Flächenstichprobe auf der Basis der Volkszählung von 1970). Das stichprobenmethodische Grundkonzept des Mikrozensus 1987 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 Auswahlbezir-

ke 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 gleich bleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht. Im Grundprogramm sind im Mikrozensus 1987 folgende Themenbereiche enthalten:

- Merkmale der Person, Familien-, Haushaltszusammenhang, Haupt- und Nebenwohnung, Aufenthaltsdauer, Staatsangehörigkeit (Wohn-/ Lebensgemeinschaft freiwillig)
- Erwerbstätigkeit, Arbeitsuche, Arbeitslosigkeit, Nichterwerbstätigkeit, Kind im Vorschulalter, Schüler, Student
- Quellen des Lebensunterhalts, Höhe des Einkommens
- Angaben zur gesetzlichen und privaten Krankenversicherung und zur gesetzlichen Rentenversicherung

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms (Auswahlsatz 1 %) werden in Ergänzungsprogrammen mit unterschiedlichen Auswahlsätzen weitere Angaben erhoben. Das Ergänzungsprogramm mit einem Auswahlsatz von 1 % umfasst 1987:

- Berufliche Verhältnisse, Tätigkeitsmerkmale
- Aus- und Weiterbildung
- Zusatzangaben für Ausländer
- Wohnsituation der Haushalte
- Zusatzfragen zur Migration (neu)

Das Ergänzungsprogramm mit einem Auswahlsatz von 0,4 % umfasst die EG-Arbeitskräftestichprobe.

Gesamtüberblick der Erhebungsprogramme

- [2005-2012](#)
- [1996-2004](#)
- [1985-1995](#)
- [1973-1985](#)

---

## Besonderheiten

Im Grundfile des Mikrozensus 1987 sind die Regionalangaben Bundesland und Gemeindegrößenklasse enthalten. Die Ausprägungen der Variable Nationalität werden so aggregiert, dass eine Nationalitätengruppe in der Grundgesamtheit wenigstens 50.000 Einwoh-

ner umfasst. Bei den übrigen Variablen werden Merkmalsvergrößerungen vorgenommen, damit in den univariaten Randverteilungen jede Ausprägung mindestens 5.000 Fälle in der Grundgesamtheit enthält. Im File des Mikrozensus 1987 sind nicht alle im Grundprogramm erhobenen Informationen enthalten. Es fehlen die Angaben zu Urlaubs- und Erholungszeiten und zur Wohnsituation des Haushalts.

---

### **Zugang zum Mikrozensus**

Die Mikrozensus Scientific Use Files können von unabhängigen inländischen Forschungseinrichtungen gegen ein Nutzungsentgelt von 250 EURO (Stand: April 2012) bei den [Forschungsdatenzentren](#) der Statistischen Ämter des Bundes und der Länder bestellt werden.

---

### **Service-Angebot des German Microdata Lab der GESIS**

In Kooperation mit dem Statistischen Bundesamt stellt das German Microdata Lab (GML) dem Nutzer der Mikrozensus Scientific Use Files ausführliche Informationen zur Verfügung und bietet Setups zum Einlesen des Rohdatenmaterials und zur Erstellung von Systemfiles in Datenformaten gängiger Softwarepakete an. Für Interessenten und Nutzer des Mikrozensus besteht weiterhin 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 Setup erfüllt folgende Aufgaben:

- Einlesen der Rohdaten
  - Recodieren und Deklarieren der Missing Values
  - Hinzufügen der Variable Labels
  - Hinzufügen der Value Labels
  - Speichern des endgültigen Systemfiles
- 

### **Hochrechnung, Gewichtung und Vergleich mit amtlichen Veröffentlichungen**

Sollen die Ergebnisse von Häufigkeitsauszählungen der 0,7 %-Substichprobe des Scientific Use File mit Veröffentlichungen des Statistischen Bundesamtes zum Mikrozensus 1987 verglichen werden, muss der Hochrechnungsfaktor des Mikrozensus verwendet werden.

Um auf 1 % der Bevölkerung hochzurechnen muss mit dem Kehrwert des Auswahlrates gewichtet werden (hier in SPSS-Syntax):

```
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 Gewichtung mit 100 multipliziert werden.

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

Weiterhin sind zum Vergleich mit den vom Statistischen Bundesamtes veröffentlichten Zahlen die jeweiligen Bevölkerungs- und Erwerbskonzepte zu berücksichtigen. Eine gesonderte Aufstellung ist unter dem Abschnitt „[Zur Abgrenzung der Bevölkerungs- und Erwerbskonzepte im Mikrozensus 1987](#)“ abzurufen.

## Prüfung des eingelesenen Materials

Zur Prüfung, ob der Rohdatensatz des faktisch anonymisierten Materials Mikrozensus 1987 fehlerfrei eingelesen wurde, kann die folgende Fallzahl  $n = 438.823$  (ohne Gewichtung, ohne Selektion) mit der des eingelesenen Datensatzes verglichen werden. Unterscheiden sich die Fallzahlen weist dies auf einen Fehler beim Einlesen hin.

## Übersicht der Eigenschaften des Mikrozensus 1987:

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. zusammen liegende Gebäude bzw. Gebäudeteile); ein Auswahlbezirk verbleibt vier Jahre in der Stichprobe, in jedem Jahr scheidet 1/4 der Auswahlbezirke aus (rotierendes Panel)
Stichprobenumfang:	27 Tsd. Haushalte, 611 Tsd. Personen
Auswahlgrundlage:	Volkszählung 1970
Auswahlrate:	1 Prozent

Auswahltechnik: Systematische Ziehung mit Zufallsstart

**Scientific Use File:** Das Scientific Use File (Grundfile) ist eine faktisch-anonymisierte 70 %-Substichprobe der Haushalte der Mikrozensus-Daten 1987, bei der alle Personen eines ausgewählten Haushalts in das Scientific Use File aufgenommen werden. Gezogen wird diese Stichprobe durch a) Anordnen nach Land, Regierungsbezirk, Auswahlbezirksnummer, b) Durchnummerieren der Haushalte, sowie c) Ziehen aller Sätze, die nicht die Endziffern 2, 5 und 9 einer ganzzahligen Zufallszahl haben.

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.

---

### Hinweise auf Materialien zum Mikrozensus:

Informationen zum Mikrozensus und seiner Konzeption:

- Statistisches Bundesamt (Hrsg.): Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957. Wiesbaden, Dezember 2004.
- Nourney, Martin: Stichprobenplan des Mikrozensus ab 1972. In: Wirtschaft und Statistik 11, 1973, S.631-638

Ergebnisse des Mikrozensus 1987: 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 3 umfasst zudem Ergebnisse über Privathaushalte, Kinder und Familien.

Darüber hinaus sind Publikationen in der Zeitschrift "Wirtschaft und Statistik" des Statistischen Bundesamtes zu finden.

Ein umfassendes Informationsangebot bietet das GML der GESIS an:

<http://www.gesis.org/gml>.

Publikationen, die auf den Scientific-Use-Files der Mikrozensusen basieren, finden sich auf der Webseite: <http://www.gesis.org/missy/studie/literatur/veroeffentlichungen-auf-basis-der-scientific-use-files-des-mikrozensus/>.