Zurück

Nutzerinformationen zur 25%-Stichprobe der Sozialhilfeempfängerstatistik 1999

1. Beschreibung des Datensatzes

Die Sozialhilfeempfängerstatistik ist eine Bestandsstatistik, die im vorliegenden Datensatz die Anzahl der Empfänger von laufender Hilfe zum Lebensunterhalt (HLU) außerhalb von Einrichtungen zum Stichtag 31.12.1999 ausweist. Diese werden auch als "Sozialhilfeempfänger im engeren Sinne" bezeichnet. Gemäß § 132 Abs. 2 des Bundessozialhilfegesetzes stellen die Statistischen Ämter der Länder dem Statistischen Bundesamt jedes Jahr u.a. Einzelangaben aus einer Zufallsstichprobe mit einem Auswahlsatz von 25 vom Hundert der Empfänger von laufender Hilfe zum Lebensunterhalt zur Verfügung. Der Datensatz für das Jahr 1999 umfasst insgesamt 698.037 Fälle (Hochrechnungsfaktor = 4). In der Datei sind sowohl personen- als auch haushaltsbezogene Daten enthalten. Insgesamt 61 Variablen, wie z.B. Geburtsjahr, Geschlecht, Staatsangehörigkeit, Erwerbsstatus, höchster Schul- bzw. Berufsausbildungsabschluss, Haushaltstyp, Anspruch und Bruttobedarf je Monat, vorhandene Einkommensarten etc. ermöglichen eine Vielzahl von Auswertungsmöglichkeiten. Die vollständige Variablenliste kann der Datensatzbeschreibung entnommen werden, die als gesonderte Datei auf der CD-ROM enthalten ist.

Methodisch gesehen ist die 25%-Stichprobe eine interessante Datenquelle, da sie dem Nutzer einerseits eine sehr große Anzahl von Einzeldatensätzen bietet und somit hohe Ansprüche hinsichtlich der Repräsentativität erfüllt. Darüber hinaus kann der Anwender seine errechneten Stichprobenergebnisse mit den aggregierten Gesamtzahlen der Totalerhebung abgleichen, welche beispielsweise in den Fachserien oder Arbeitsunterlagen des Statistischen Bundesamtes veröffentlicht sind.

Aus datenschutzrechtlichen Gründen musste eine Anonymisierung der Daten vorgenommen werden, wodurch zum Teil eine Modifizierung der Variablen erforderlich wurde. Die Auswirkungen dieser Anpassungen werden im Folgenden näher erläutert:

Im anonymisierten Material sind lediglich drei **Gebietsstände** ausgewiesen. Das *"frühere Bundesgebiet Nord inkl. Berlin-Ost"* umfasst die drei Stadtstaaten Berlin, Bremen und Hamburg so-

wie die Länder Schleswig-Holstein, Niedersachsen und Nordrhein-Westfalen. Das *"frühere Bundesgebiet Süd"* ist definiert durch die fünf verbleibenden Bundesländer Hessen, Rheinland-Pfalz, Baden-Württemberg, Bayern und das Saarland. Die *"neuen Länder ohne Berlin-Ost"* werden durch die fünf verbleibenden ostdeutschen Bundesländer abgedeckt.

Die Variable **"Stellung zum Haushaltsvorstand"** liefert wichtige Informationen zur Bestimmung des Haushaltstyps. Sie hat fünf Ausprägungen: Haushaltsvorstand (1), Ehegatte(in) (2), Kind (3), Verwandte(r) oder Verschwägerte(r) (4) und sonstige Person (5). Der Haushaltsvorstand ist die Person, welche die Generalkosten des Haushalts trägt. Aus diesem Grund steht dem Haushaltsvorstand ein erhöhter Eckregelsatz zu. Eine nicht ganz eindeutige Abgrenzung verbirgt sich hinter der Klassifizierung "Kind". Der Oberbegriff "Stellung zum Haushaltsvorstand" lässt vermuten, dass es sich hierbei um die leiblichen Kinder des Haushaltsvorstandes handelt. Der Begriff ist aber in dem hier verwendeten Zusammenhang weiter gefasst. Tatsächlich werden als Kinder neben ehelichen und nichtehelichen auch Stief-, Adoptiv- und Pflegekinder mit einbezogen, die zusammen mit ihren Eltern oder einem Elternteil in einem Haushalt wohnen.

Aus der Variable **"Geburtsjahr"** lässt sich das Alter der HLU-Empfänger berechnen. Im Rahmen der Anonymisierung wurde das Höchstalter auf 90 Jahre festgesetzt. Den Personen, die vor 1908 geboren wurden, ist das Geburtsjahr 1908 zugewiesen worden. Alle Personen, die im Berichtsjahr 1999 im Alter von 91 Jahren und mehr waren, erscheinen demnach bei Bildung der Variable "Alter" unter der Restkategorie der 91-jährigen.

Als ein Teil der laufenden Hilfe zum Lebensunterhalt werden bestimmten Personengruppen, beispielsweise behinderten Menschen und Pflegebedürftigen, zusätzlich **"Mehrbedarfszuschläge"** gewährt, welche im einzelnen der Datensatzbeschreibung entnommen werden können. Ein Sozialhilfeempfänger kann von insgesamt sieben unterschiedlichen Mehrbedarfszuschlägen maximal vier gleichzeitig erhalten. Dadurch sind für dieses Merkmal im Datensatz vier Positionen vorgesehen, welche jeweils die Ausprägungen 1-7 annehmen können.

Ähnlich wie bei der Höchstaltersbegrenzung auf 90 Jahre (siehe oben) war es in dieser Stichprobe aus datenschutzrechtlichen Gründen erforderlich, bei bestimmten Variablen eine Abschneidegrenze festzulegen. So musste bei den Merkmalen "**Nummer der Person**", **"Anzahl der Personen in der Bedarfsgemeinschaft"**, "**Zahl der Haushaltsmitglieder"** und **"Zahl aller HLU-Empfänger im Haushalt"** eine Beschränkung auf 10 Personen vorgenommen werden. Das bedeutet, dass bei-

spielsweise eine Bedarfsgemeinschaft, die real aus 12 Personen besteht, im vorliegenden Datensatz lediglich mit 10 Personen erfasst wird. Da dies nur für relativ wenige Fälle zutrifft, ist die zu erwartende Untererfassung als gering zu beurteilen.

Die in der vorliegenden Stichprobe enthaltenen **Zahlbeträge** werden aus Gründen der Vergleichbarkeit mit denen der Vollerhebung (veröffentlicht u.a. in den Fachserien und Arbeitsunterlagen des Statistischen Bundesamts) noch in Deutschen Mark ausgewiesen. Falls Umrechnungen von DM in EURO erfolgen, gilt der **Umrechnungsfaktor 1 Euro = 1,95583 DM.**

Die Beträge **"Bruttobedarf"**, "**Anerkannte Bruttokaltmiete"**, **"Anspruch der Bedarfsgemeinschaft"** und die aus diesen Angaben ermittelte Größe "**Angerechnetes Einkommen"** wurden auf Zehnerstellen gerundet (z.B.: 1452,-DM ➔ 1450,-DM). Um Rückschlüsse auf einzelne Empfänger zu vermeiden, wurden bei diesen Beträgen Höchstgrenzen eingeführt. Der maximal ausgewiesene *Bruttobedarf* liegt bei 5.000,-DM. Der größte ausgewiesene Wert für die *anerkannte Bruttokaltmiete* beläuft sich auf 1.800,-DM; beim *Nettoanspruch* beträgt er 3.500,-DM. Dies ist auch der Höchstwert für das *angerechnete Einkommen*.

Die **"Längste bisherige Dauer der Hilfegewährung an die Bedarfsgemeinschaft in der aktuellen Zusammensetzung"** und die **"Bisherige Dauer der ununterbrochenen Hilfegewährung für mindestens ein Mitglied der Bedarfsgemeinschaft"** ist auf 120 Monate beschränkt. Einer Bezugsdauer, die länger als 120 Monate anhält, ist der Wert "120" zugeordnet. Diese Kappungsgrenzen führen dazu, dass die durchschnittliche Bezugsdauer gegenüber dem nicht-anonymisierten Material und gegenüber den Daten der Totalerhebung geringer ausfällt.

Die Variable "**Typ der Bedarfsgemeinschaft"** hat insgesamt 24 Ausprägungen und ist aus den Variablen **"Stellung zum Haushaltsvorstand"**, "**Geburtsjahr"**, **"Geschlecht"** und "**Anzahl der Personen in der Bedarfsgemeinschaft"** konstruiert. Die Haushaltstypisierung wurde in der vorliegenden Stichprobe weitgehend an die der Standardtabellen und Veröffentlichungen des Statistischen Bundesamts angepasst. Dementsprechend wurden auch hier die gleichgeschlechtlichen Paare in der Kategorie **"nichteheliche Lebensgemeinschaft"** erfasst.

2. Qualität der Stichprobe

Größere Datenbanken mit Individualdatensätzen enthalten immer eine Reihe inkonsistenter und schwer interpretierbarer Einzeldatensätze. Obwohl die 25%-Stichprobe verschiedenen Kontrollen unterzogen wurde, ist das Datenmaterial nicht vollkommen frei von solchen widersprüchlichen Einzeldatensätzen. Die vorgenommene Qualitätskontrolle basiert auf zwei Säulen: Auf der Prüfung der Einzeldatensätze und auf dem Vergleich der Stichprobenergebnisse mit denen der Totalerhebung.

Die Prüfung der Einzeldatensätze erfolgt bereits durch die Statistischen Landesämter. Diese prüfen die Plausibilität der Originaldaten. Dabei handelt es sich einerseits um eine *Signierkontrolle* (haben die Variablen auch wirklich nur die vorgesehenen Ausprägungen?) und andererseits um eine *Kombinationskontrolle* (Beispiel: Wenn a gilt, dann muss auch b vorkommen).

Zur Beurteilung der Qualität der Stichprobe wurden für einige ausgewählte Variablen die Abweichungen der Stichprobenergebnisse von denen der Totalerhebung ermittelt. Nach dem Gesetz der großen Zahl nimmt der Stichprobenfehler mit kleiner werdender Fallzahl zu. Je detaillierter die Untergliederung wird und je kleiner die Fallzahlen werden, desto wahrscheinlicher wird die Abweichung der Stichprobenergebnisse von denen der Totalerhebung. Durch die Anonymisierung kommt ein weiterer Aspekt hinzu: Bestimmte Informationen liegen nicht ganz so exakt wie im Originalmaterial vor. Trotz dieser Einschränkungen lässt sich sagen, dass der vorliegende Datensatz unserer Ansicht nach von guter Qualität ist.

Im Folgenden sind die Ergebnisse des Datenvergleichs zwischen Stichprobe und Totalerhebung zusammengefasst.

– Die Variable **"Typ der Bedarfsgemeinschaft"** zeigt in ihren einzelnen Ausprägungen im Großen und Ganzen keine stärkeren Abweichungen von der Totalerhebung. Lediglich bei dem Typ 18 (Sonstige Bedarfsgemeinschaften mit zwei volljährigen und mindestens einem minderjährigen Haushaltsangehörigen) weichen die Ergebnisse der Stichprobe um 23% von denen der Totalerhebung ab. Hier dürfte insbesondere zum Tragen kommen, dass dieser Bedarfsgemeinschaftstyp in der Stichprobenziehung nicht explizit berücksichtigt wird und somit die Fallzahlen in der Erhebung sehr gering sind.

- Die **Bezugsdauer** ist wie oben schon angesprochen in der Stichprobe auf den Höchstwert von 120 Monaten beschränkt. Das führt dazu, dass dieser Wert für Bedarfsgemeinschaftstypen, die überdurchschnittlich lange Hilfe zum Lebensunterhalt beziehen, unterschätzt wird. So liegt die Bezugsdauer von weiblichen Alleinstehenden in der anonymisierten Stichprobe um ca. sechs Monate unter dem Wert, der mit den Daten der Totalerhebung ermittelt wurde.
- Für die quantitativ bedeutenden Bedarfsgemeinschaftstypen lässt sich zusammenfassend sagen, dass die Stichprobenergebnisse für die **Zahlbeträge (Bruttobedarf, anerkannte Bruttokaltmiete, angerechnetes Einkommen und Nettoanspruch)** nur sehr gering von den Ergebnissen der Totalerhebung abweichen.
- Die Stichprobenergebnisse für die **Altersgruppen nach Geschlecht** weichen lediglich für Männer ab 75 Jahren um mehr als 2% von der Totalerhebung ab.
- Die Differenzierung nach den unterschiedlichen **Gruppen von Nichtdeutschen** und die weitere Untergliederung nach Alter zeigt in der Stichprobe erwartungsgemäß größere Abweichungen von den Daten der Totalerhebung. Für die große Gruppe der "Sonstigen Ausländer/innen" gilt dies in weit geringerem Ausmaß als für die vergleichsweise schwach besetzten Kategorien "EU-Ausländer/innen", "Asylberechtigte(r)" und "Bürgerkriegsflüchtling".
- Die **Mehrbedarfszuschläge** nach Altersgruppen spiegeln erneut das Problem geringer Fallzahlen in den jeweiligen Kombinationen wider. Bei geringer Zellenbesetzung besteht das Risiko, dass die Stichprobenergebnisse von denen der Totalerhebung stärker abweichen. In einigen Konstellationen mit einer Zellenbesetzung von unter 100 Fällen, liegt die prozentuelle Abweichung im zweistelligen Bereich.

3. Hinweise zur Berechnung von Zahlbeträgen

Eine Besonderheit des vorliegenden Datenmaterials ist die **Kopplung von personen- und haushaltsbezogenen Informationen**. In der Sozialhilfestatistik existieren zwei parallele Systeme: Es werden zum einen Daten über die einzelnen Sozialhilfeempfänger erhoben (personenbezogene Daten wie z.B. Geburtsjahr, Geschlecht, Bildungsabschluss etc.), zum anderen werden aber auch Informationen erfasst, welche sich auf den gesamten Haushalt beziehen, in dem der/die Sozialhilfeempfänger(in) lebt. Dies sind unter anderem die Zahlbeträge "Bruttobedarf der Bedarfsgemeinschaft", "Anerkannte Bruttokaltmiete", "Nettoanspruch der Bedarfsgemeinschaft" und "Angerechnetes Einkommen der Bedarfsgemeinschaft".

Die personenbezogenen Daten werden zunächst in einem anderen Datensatz erfasst als die, welche die haushaltsbezogenen Informationen aufweisen. Um möglichst viele Informationen über die Empfänger und die Haushalte der Empfänger zu erhalten, wurden die separaten Datensätze in der vorliegenden 25%-Stichprobe zusammengeführt. Damit in dem Datensatz komfortabel gearbeitet werden kann (beispielsweise durch den Einsatz von Filtern), sind die Zahlbeträge, welche sich auf den gesamten Haushalt beziehen, jeder Person in dem jeweiligen Haushalt zugeordnet worden (siehe Abbildung).

	lfd_p_nr	lfd_hhnr	personen	bedarf	miete	anspruch	angerech
1	1	1	4	2630	720	1750	880
2	2	1	4	2630	720	1750	880
3	3	1	4	2630	720	1750	880
4	4	1	4	2630	720	1750	880

Der Haushalt in diesem Beispiel besteht aus vier Personen. Die Zahlenwerte "bedarf", "miete", "anspruch" und "angerech" wurden allen Personen des Haushalts zugeordnet, obwohl sich diese Werte auf den Gesamthaushalt beziehen: Nicht jede der vier Personen hat einen Bedarf von 2630,- DM, sondern alle vier Personen zusammen. Gleiches gilt für die Miete, den Nettoanspruch und das angerechnete Einkommen.

Wichtig: Falls mit Hilfe der Stichprobe Mittelwerte berechnet werden (beispielsweise der Mittelwert des Nettoanspruchs), so ist darauf zu achten, dass sich die Zahlbeträge auf die Bedarfsgemeinschaften und nicht auf die einzelnen Personen beziehen. Um einen korrekten Wert zu erhal-

ten, darf der Wert 1750,- DM aus dem oben skizzierten Beispiel nur einmal und nicht viermal in die Berechnung einfließen! Bevor Berechnungen solcher Art durchgeführt werden, muss somit zunächst nach Bedarfsgemeinschaften gefiltert werden.

Wir schlagen vor, den Filter **lfd_p_nr = 1** zu verwenden.

Da die Personen eines Haushalts durchnummeriert sind, ist gewährleistet, dass der Zahlenwert jeder Bedarfsgemeinschaft bei der anschließenden Berechnung des Mittelwerts nur einmal berücksichtigt wird.

4. Lieferumfang

Die Daten stehen dem Nutzer als ASCII-File, sowie als SPSS- und SAS-Datei zur Verfügung. Ergänzend werden Dateien mitgeliefert, die das Konvertieren des ASCII-Files in eine SPSS- oder SAS-Datei ermöglichen. Die Verwendung der Konvertierungsdateien ("Einlesehilfe SPSS" bzw. "Einlesehilfe SAS") empfiehlt sich für alle Anwender, die abweichende Variablennamen oder Kommentierungen vergeben möchten bzw. die nur eine Auswahl der Variablen einlesen wollen. In diesen Fällen sind die beabsichtigten Änderungen in den Konvertierungsdateien vorzunehmen. Die gelieferten Variablen und ihre Spezifizierungen sind in der Datensatzbeschreibung dokumentiert.

Zurück