

Erläuterungen zum Datensatz „BAMF-Flüchtlingsstudie 2014“

Stand: 07.12.2017

1. Aufbau des Datensatzes und Arten von Variablen

Der Datensatz enthält insgesamt 211 Variablen. Davon sind:

191 **Kernvariablen**, die den Fragebogeninhalt im engeren Sinne widerspiegeln und an den Anfangsbuchstaben **e** bzw. **v** zu erkennen sind.

20 **neu gebildete Variablen** (am Ende des Datensatzes), die daran zu erkennen sind, dass sie „inhaltliche“ Namen und keine e- oder v-Bezeichnung haben. Sie sollen die Analysen des Forschungsberichtes zum Projekt (Worbs/Bund/Böhm 2016) leichter nachvollziehbar machen. Darunter fallen auch zwei GewichtungsvARIABLEN ganz am Ende des Datensatzes sowie vier so genannte Plausibilisierungsvariablen (Endung **_plausi**). Letztere sollen bei komplexeren, über mehrere Items abgefragten Sachverhalten – schulische Bildung, Besuch von Sprachkursen, Besuch von Beratungsstellen sowie Aufenthaltsdauer/Aufenthaltstitel – Befragte mit insgesamt plausiblen Angaben von solchen Befragten unterscheiden, bei denen das nicht der Fall ist.

Bei den Kernvariablen gibt es noch folgende Untergruppen:

Statusvariablen sind an der Endung **s** erkennbar. Sie beziehen sich auf Fragen, bei denen Mehrfachantworten zulässig waren, und unterscheiden die gegebenen Antworten in konsistente (Code 55) und inkonsistente (Code 66) sowie die Restkategorie keine Angabe (99). Nähere Erläuterungen zu diesen Codes befinden sich im nächsten Abschnitt.

Gegenüber dem Original-Fragebogeninhalt **veränderte Variablen** sind an die Endung **_n** erkennbar. Dabei handelt es sich um kategorisierte Antworten auf offene Fragen oder Vergrößerungen von Antwortkategorien, um die Identität einzelner Befragter zu schützen.

Insgesamt zehn Variablen, die ursprünglich im Fragebogen enthalten waren, sind aus dem Datensatz ganz entfernt worden. Dies ist im Codebuch vermerkt. Dabei handelt es sich einerseits um offene Nachfragen bei an sich geschlossenen Fragen, („Sonstiges, und zwar ...“), bei denen keine Kategorisierung möglich oder sinnvoll war und bei denen durch die Nennungen Rückschlüsse auf die Identität von Befragten möglich gewesen wären. Zum anderen wurden drei Variablen wegen offensichtlicher Verständnisprobleme der Befragten und dementsprechend nicht validen Antwortmustern entfernt. Nähere Erläuterungen dazu finden sich im Forschungsbericht zum Projekt bzw. im Codebuch.

2. Art der Vercodung

Neben den inhaltlich belegten Variablenausprägungen (Ziffern 0,1,2 usw.) wurden standardmäßig folgende Codes vergeben (vgl. auch Worbs/Bund/Böhm 2016: 51f.):

55 - Konsistent (bei Fragen mit zulässigen Mehrfachantworten in der jeweiligen "Statusvariable"). Nur Fälle mit dieser Kennzeichnung sollten bei der Auswertung der jeweiligen Variablen berücksichtigt werden.

66 - Inkonsistent/falsch verstanden für Antworten, die mit einiger Wahrscheinlichkeit oder sicher falsch sind oder wo die Kombination der Antworten bei Fragen mit zulässigen Mehrfachantworten keinen Sinn ergibt. Ein Beispiel wäre die Zustimmung sowohl bei „ich habe (noch) keinen Beruf erlernt und auch nicht studiert“ **und** bei „Ich habe ein Studium abgeschlossen und besitze ein Abschlusszeugnis“ bei Frage 34.

77 - Unzulässige Mehrfachantwort bei Fragen, bei denen nur eine Antwort möglich war.

88 - Weiß nicht.

98 - nicht übersetzte offene Angabe (aus fremdsprachigem Fragebogen). Bei offenen Fragen sind solche Antworten als "NUE" gekennzeichnet. Hintergrund dieses Codes ist, dass aus Kapazitätsgründen nicht alle offenen Angaben in fremdsprachig ausgefüllten Fragebögen übersetzt werden konnten (vgl. Worbs/Bund/Böhm 2016: 52f.).

99 - Keine Angabe (wobei keine Filter vorliegt, d.h. der/die Befragte hätte eine Antwort geben sollen)

. - Missing (Frage trifft für diese Person nicht zu, beispielsweise Frage 14 nach der Situation im Haushalt lebender Kindern bei Personen, die entweder keine Kinder haben oder bei denen diese nicht mit ihm Haushalt leben).

Der Verweis „Siehe V71“ bei einigen offenen Angaben bezieht sich auf außerhalb des Datensatzes erfasste Anmerkungen der Codierer/-innen zu dem jeweiligen Fall. Ebenfalls außerhalb des Datensatzes wurden auch die z.T. umfangreichen „Originalantworten“ bei den Variablen v65 und v70 erfasst. Sie können bei Bedarf bei den Datengebern angefordert werden.