**Migration und Flucht im SOEP: Übungsaufgaben**

Das SOEP stellt mit seinen verschiedenen Migrations- und Fluchtstichproben ein breites Spektrum an Informationen zu Personen mit Flucht- und Migrationshintergrund bereit.

Im BIOIMMIG Datensatz finden Sie relevante Informationen zur Flucht- und Migrationsgeschichte, wie etwa Flucht- und Migrationsmotive, die Umstände nach der Ankunft in Deutschland aber auch Informationen zu Verwandten im Ursprungsland und dem Wunsch ins Herkunftsland zurückzukehren in aufbereiteter Form. Weitere Informationen zu diesem Datensatz, sowie eine Liste der darin enthaltenen Variablen finden Sie in der Dokumentation der generierten Variablen. (<https://www.diw.de/de/diw_01.c.552837.de/soep_dokumentation_generierte_variablen_v32_1.html>)

Im Folgenden werden wir mithilfe dieses Datensatzes und weiteren Informationen aus dem SOEP eine Statusvariable erstellen, mithilfe welcher Sie unterscheiden können, ob Personen mit Migrationshintergrund auch einen Fluchthintergrund haben oder nicht.

**Schritt 1: Beginnen Sie mit Basisvariablen aus BIOIMMIG (Do-File Zeilen 16-97)**

1. In welcher Variable finden Sie Informationen darüber, welchen Status die jeweilige Person bei Ihrer Einwanderung in Deutschland hatte?
2. Identifizieren Sie diese Variable im BIOIMMIG Datensatz und laden Sie sie aus dem Datensatz, zusammen mit der Personennummer und dem Befragungsjahr
3. Welche Ausprägungen hat diese Variable?
4. Generieren sie sich auf Basis dieser Variable die Variable „Flucht“ die nur noch drei Gruppen unterscheidet und zwar:

0 = Fälle zu denen keine Informationen vorliegen

1. = Alle Personen ohne Fluchthintergrund
2. = Asylsuchende / Geflüchtete
3. Es kommt vor, dass zunächst keine Informationen über den Status bei Einwanderung vorliegt, dies sich jedoch in einem späteren Jahr ändert. Beschränken Sie den Datensatz auf die letzte Beobachtung die für die jeweilige Person zur Verfügung steht, da so die Ausprägung mit dem meisten Informationsgehalt genutzt wird.
4. Speichern Sie den generierten Datensatz auf Ihrem persönlichen Laufwerk zwischen.

**Schritt 2: Basisvariablen aus PPFAD und Gewichte anfügen (Do-File Zeilen 98-158)**

Einige Variablen aus PPFAD sollten Sie standardmäßig immer an Ihren Datensatz anfügen.

* 1. Laden Sie sich aus PPFAD die folgenden Informationen:
     + Unveränderliche Personennummer „persnr“
     + Haushaltsnummer „hhnr“ und die aktuelle Haushaltsnummer „bfhhnr“
     + die Netto-Variable mit Informationen zum Interviewtyp „bfnetto“
     + das Geschlecht der Person „sex“
     + das Geburtsjahr „gebjahr“
     + sowie die Variablen zum Migrationshintergrund „migback“, „germborn“ „corigin“ „immiyear“
     + Informationen über den Befragungstatus: „bfnetto“ und „psample“
  2. Mergen Sie den zuvor generierten Datensatz über die Migrationsgruppe an.
  3. Fügen sie dem Datensatz die entsprechenden Personenhochrechnungsfaktoren hinzu (Tipp: Datensatz „phrf.dta“).
  4. Behalten Sie nur Individuen für die in 2015 ein Jugend- oder Personenfragebogen realisiert wurde.

**Schritt 3: Generieren Sie eine Statusvariable (Do-File Zeilen 159-174)**

Generieren Sie nun mit den vorhandenen Daten eine Statusvariable die folgende fünf Ausprägungen aufweist und benennen Sie die Gruppen entsprechend.

|  |  |  |
| --- | --- | --- |
| 0 | nicht genügend Informationen  zum Migrationshintergrund vorhanden | keine Infos |
| 1 | in Deutschland geborene ohne Migrationshintergrund | kein migback |
| 2 | in Deutschland geborene mit Migrationshintergrund | 2. Generation |
| 3 | im Ausland geborene ohne Informationen zum Status | Einwanderer ohne Info |
| 4 | im Ausland geborene ohne Flucht/Asyl Titel | Einwanderer, keine Flucht |
| 5 | im Ausland geborene mit Flucht/Asyl Titel | Einwanderer, Flucht |

**Schritt 4: Generieren Sie eine Statusvariable (Do-File Zeilen 175-269)**

1. Wie viele Geflüchtete (im Ausland geborene mit Flucht/Asyl Titel) sind nun in Ihrem Datensatz?
2. Wie viele sind es wenn man die Personenhochrechnungsfaktoren beachtet? Deuten Sie die Ergebnisse.
3. Wie viele Personen werden von der Stichprobe unter Beachtung der Hochrechnungsfaktoren repräsentiert?
4. Wie groß ist der Anteil der über-40-Jährigen unter den Geflüchteten?