

## Übung Multivariate Analyseverfahren – SoSe 2016

Dr. Volker Ludwig

### Übungsblatt 1 – Abgabe bis 11.05.2016

Wir stellen einen Paneldatensatz für die beiden Jahre 1992 und 2010 aus den Originaldaten des SOEP zusammen, mit dem wir die Zeitverwendung für Hausarbeit analysieren können.

1. Erstellen Sie ein Do-file „03\_Übungsblatt1“. Legen Sie Ihre Arbeitsverzeichnisse fest, in dem Sie globale Macros `pfad1` und `pfad2` definieren (Befehl `global`). (`pfad1` ist der Ordner, in dem die SOEP-Originaldaten liegen; `pfad2` der Ordner, in dem Sie Ihre rekodierten Daten speichern wollen.)
2. Finden Sie die Variablen zum Zeitaufwand für Hausarbeit (in Std.) an einem typischen Werktag für die Wellen 1992 und 2010 in den Personendaten `$p.dta` des SOEP. Matchen Sie die Information zur Hausarbeit aus beiden Jahren und speichern Sie die Info als neuen Datensatz `hw1.dta` in `pfad2`.

(Hinweis: Erzeugen Sie zunächst für jedes der beiden Jahre einen Querschnittsdatsatz, der nur die Personnummer, das Befragungsjahr und die Variable zur Hausarbeit enthält. Nennen Sie die Variablen: `id`, `year`, `hw`. Matchen Sie dann die Infos aus beiden Jahren in einem Paneldatensatz.)

3. Erstellen Sie einen Datensatz `hw2.dta`, der für die beiden Jahre nur die folgenden Variablen enthält:
  - Personnummer (neuer Variablenname: `id`)
  - Befragungsjahr (`year`)
  - Aktuelle Haushaltsnummer (`hhnrakt`)
  - Personnummer PartnerIn (`pid`)
  - Familienstand (`famstd`)
  - Bruttomonatsverdienst in Euro (`labgro`)

Verwenden Sie generierte Variablen. Speichern Sie den neuen Datensatz (in `pfad2`).

4. Erstellen Sie einen Datensatz `hw3.dta` (in `pfad2`), der für jede Person nur die folgenden zeitkonstanten Variablen enthält:
  - Personnummer (`id`)
  - Geschlecht (`sex`)
  - Geburtsjahr (`gebjahr`)

Verwenden Sie die Infos aus dem Datensatz `ppfad.dta`.

5. Spielen Sie jetzt die Daten aus `hw1.dta`, ..., `hw3.dta` der Reihe nach zusammen (Befehl `merge`). Betrachten Sie dazu in jedem Schritt zuerst die

Datenstruktur der beiden Datensätze, die sie matchen wollen (`master` und `using`). Verwenden Sie den Befehl `list` (oder `browse`), um sich immer die ersten 20 Beobachtungen ausgeben zu lassen. Überlegen Sie sich dann, wie der `merge` Befehl lauten muss.

6. Speichern Sie den resultierenden Datensatz als `housework_92_10.dta` (in `pfad2`).
7. (\*) Erstellen Sie einen Paneldatensatz `hw4.dta`, der für die beiden Jahre und für jeden Haushalt nur die folgenden Variablen enthält:
  - Aktuelle Haushaltsnummer (`hhnrakt`)
  - Befragungsjahr (`year`)
  - Wohnort Ost-/Westdeutschland (`east`)
  - (Korrigiertes) Netto-Haushaltseinkommen (`hhinc`)

Verwenden Sie generierte Variablen (Originaldaten `$hgen.dta`). Matchen Sie die Daten `hw4.dta` mit dem Datensatz `housework_92_10.dta`.

8. (\*) Lassen Sie sich deskriptive Statistiken (Mittelwerte bzw. Anteile) für die Variablen im Datensatz ausgeben, getrennt nach Geschlecht und Erhebungsjahr. (Dazu müssen Sie zuerst Missings definieren.) Wie hat sich der Zeitaufwand von Frauen und Männern entwickelt?
9. Schicken Sie Ihr Do-file an [vludwig@soziologie.uni-muenchen.de](mailto:vludwig@soziologie.uni-muenchen.de). Geben Sie im Betreff der email das Kürzel „MVSO16IhrNachname“ an.

**Aufgaben mit (\*) sind Zusatzaufgaben.**