

Aufgabe 1 - Datensatz in SPSS einlesen und formatieren

Es wurde folgende Studie durchgeführt:

- ❖ In zwei Kliniken wurden Daten von Patienten erhoben.
 - ❖ In jeder Klinik wurde unabhängig von der anderen Klinik eine Patienten-ID vergeben.
 - ❖ Die Daten wurden bereits in der jeweiligen Klinik eingegeben, jeweils von einer Person.
 - ❖ Jeder Patient wurde genau einmal befragt.

 - ❖ Patienten durften nur in die Studie eingeschlossen werden, wenn sie 18 Jahre oder älter waren.

 - ❖ Es wurde das Geschlecht erfasst.
 - ❖ Es wurde das Alter erfasst.
 - ❖ Es wurde nach der Anzahl eigener Kinder gefragt.

 - ❖ Es wurde erfasst, ob die Patienten über ausreichend Deutschkenntnisse verfügten, um selbst zwei deutsche Fragebögen (Fragebogen A und Fragebogen B) auszufüllen.
 - ❖ Bei zu geringen Deutschkenntnissen erhielten sie die Fragebögen auf Englisch.

 - ❖ In Fragebogen A wurde nach der allgemeinen Gesundheit gefragt. Mögliche Werte liegen zwischen 0=schlechteste Gesundheit und 10=beste Gesundheit.
 - ❖ In Fragebogen B wurde ebenfalls nach der allgemeinen Gesundheit gefragt. Mögliche Werte liegen zwischen 0=schlechteste Gesundheit und 100=beste Gesundheit.
- a) Laden Sie das **Codebuch zum Übungsblatt** von der Homepage herunter. Schauen Sie sich das Codebuch an.
 - b) Was wurde im Codebuch evtl. ungünstig festgelegt?
 - c) Ergänzen Sie im Codebuch zu allen Variablen das Skalenniveau.
 - d) Laden Sie den **Datensatz zum Übungsblatt** von der Homepage herunter.
 - e) Lesen Sie den Datensatz in SPSS ein.
 - f) Speichern Sie den Datensatz als SPSS-Datensatz (Dateiendung .sav).
 - g) Füllen Sie die Variablenansicht gemäß der Vorgaben im Codebuch aus. Speichern Sie den formatierten Datensatz erneut.
 - h) Klicken Sie in der Datenansicht auf „Wertbeschriftungen“ (Symbol mit rosa Pfeil mit rotem A oben und roter 1 links). Schauen Sie sich die formatierten Daten kurz an.

Aufgabe 2 – Plausibilitätsprüfung der formatierten Daten

Verwenden Sie den formatierten Datensatz aus Aufgabe 1.

- a) Prüfen Sie die Plausibilität für alle kategorialen Variablen, indem Sie sich die Häufigkeitsverteilungen aller kategorialen Variablen ausgeben lassen. Welche Eingaben können nicht stimmen?
- b) Prüfen Sie die Plausibilität für alle metrischen Variablen, indem Sie sich geeignete Maßzahlen für alle metrischen Variablen ausgeben lassen. Welche Eingaben können nicht stimmen?
- c) Verwenden Sie geeignete Analysen, um weitere, nicht so offensichtliche Fehler (oder auch eventuelle Ursachen für die Fehler) in den Daten zu finden?

Hier ein paar Hinweise (**bitte möglichst erst nach eigenen Versuchen lesen**):

- Gibt es ein Problem mit den ID-Variablen? Sind diese für jede Klinik eindeutig?
- Sind alle Patienten alt genug, um in die Studie eingeschlossen zu werden?
- Hängt die fehlerhafte Kodierung des Geschlechts mit der Person zusammen, die die Daten eingegeben hat? Bei wem müsste man nachfragen?
- Ist der Zusammenhang für die allgemeine Gesundheit aus Fragebogen A und die allgemeine Gesundheit aus Fragebogen B plausibel? Gab es evtl. einen Übersetzungsfehler?

Aufgabe 3 – Selbständiges Erstellen eines Datensatzes

- a) Überlegen Sie sich 3-10 Variablen, die Sie gerne für eine Studie erfassen wollen (gerne auch mehr).
- b) Erstellen Sie ein **Codebuch** für diese Variablen. (Speichern Sie dieses.)
- c) Erfinden Sie einen **Datensatz** entsprechend Ihres Codebuches (mit mindestens 8 Patienten, besser deutlich mehr). (Speichern Sie diesen.)
- d) Lesen Sie diesen Datensatz in SPSS ein.
- e) Formatieren Sie den Datensatz entsprechend Ihres Codebuches, d.h. füllen Sie die Variablenansicht vollständig aus. (Speichern Sie diesen Datensatz als SPSS-Datensatz.)

Aufgabe 4 - Fehlersuche

- a) Speichern Sie den (formatierten) Datensatz aus Aufgabe 3 zusätzlich unter einem neuen Dateinamen (z.B. Daten_mit_Fehlern.sav)
- b) Bauen Sie in diesen Datensatz absichtlich Fehler ein, z.B.
 - In der Datenansicht:
 - Tippfehler, d.h. Werte, die offensichtlich nicht existieren können
 - In der Variablenansicht:
 - Fehlende oder fehlerhafte Spezifizierung für fehlende Werte
 - Fehlerhafte Wertelabel, d.h. fehlende oder falsche Angaben (Spalte „Werte“ in der Variablenansicht)
 - Ein falscher Typ bzw. ein falsches Messniveau (Skalenniveau)
 - Sonstige Fehler, die Ihnen einfallen
- c) Suchen Sie sich einen Partner. Stellen Sie diesem Ihr Codebuch und Ihre formatierten, mit Fehlern versehenen Daten zur Verfügung. Lassen Sie Ihren Partner nach den Fehlern suchen.