



## Übung Einführung Statistik und EDV – WS 2023/24

### Übung 3 (Namberger) – Datenmanagement III

- Fälle auswählen (Wdh.)
- Daten umcodieren
- Daten aggregieren
- Doppelte Fälle ermitteln
- Datumsextraktion
- Variablen berechnen
- Daten umstrukturieren
- Daten in einem anderen Format speichern (Excel)

#### Aufgabe 1 (freiwillige Hausaufgabe):

Die Datei „BY2011HSO2\_MUC.sav“ beinhaltet Messwerte für Luftschadstoffkonzentrationen ( $\mu\text{g}/\text{m}^3$ ) in München. Enthalten sind die Werte für den Schadstoff Schwefeldioxid ( $\text{SO}_2$ ) an zwei verschiedenen Messstationen in München (DEBY002 und DEBY037). Für jede der beiden Stationen gibt es stündliche Messwerte für jeden Tag im Jahr 2011. Die Variable „Datum“ bezeichnet den entsprechenden Tag im Jahr (z.B. 2011/07/26 für den 26.07.2011). Die Variablen „Wert01“ bis „Wert24“ stehen für die Stunde des Tages, an dem die Messung vorgenommen wurde (z.B. Wert07 für 07.00 Uhr morgens). Entsprechend hat die Datei 730 Zeilen (2 Stationen x 365 Tage).

Dieses Datenformat (Struktur und Anordnung von Spalten und Zeilen) haben Sie so vom Umweltbundesamt erhalten. Für ihr Forschungsprojekt benötigen Sie allerdings ein anderes Datenformat, um die Schadstoffkonzentrationen mit einer speziellen Software weiterverarbeiten zu können. Die Datei „Example\_dry deposition analysis\_SO2.xls“ zeigt Ihnen das Zielformat, in welches Sie die Daten umwandeln müssen. Für das Forschungsprojekt benötigen Sie für jede Stunde den durchschnittlichen Messwert der zwei Stationen.

Verwenden Sie Ihre Kenntnisse des Datenmanagements, um die Daten entsprechend des Zielformats umzustrukturieren.

Vollziehen Sie dabei die folgenden Arbeitsschritte:

- Wert „-999“ umcodieren in „systemdefiniert fehlend“ (für fehlende Werte)
- Aggregieren der Messwerte der beiden Stationen DEBY002 und DEBY037 (Ziel: Durchschnittswerte)
- Doppelte Fälle löschen
- Daten umstrukturieren (transponieren der Stundenwerte)
- aus Variable [Datum] drei neue Variablen erstellen ([YEAR], [MONTH] und [DAY])
- übrige Variablennamen ändern, unnötige Variablen löschen und/oder neue Variablen erstellen
- Exportieren in eine Excel-Datei (\*.xlsx)

**Tipp:** Die Ergebnisdatei muss 8.760 Zeilen (plus Titelzeile) haben.