

Einführung in SPSS

Empirisches Forschungsprojekt

Tarek Carls

21. Oktober 2023

2. Überblick über SPSS

- Modular aufgebautes Programmpaket.
- Basismodul für Datenmanagement und statistische Analysen.
- Ergänzende Module für spezielle statistische Verfahren.
- Ursprünglicher Name: Statistical Package for the Social Sciences.
- Aktueller Name: IBM SPSS Statistics (SPSS).

2. Einsatzgebiete und Beliebtheit

- Einsatz in verschiedenen Gesellschaftsbereichen.
- Standard in empirischen Wissenschaften.
- Vorteile:
 - Vielfältige Formate und Funktionen.
 - Diverse Möglichkeiten zur Anwendung.
 - Komplexe statistische Analysen möglich.
- Nachteile:
 - Etwas veraltet
 - Preis
 - nicht open source
 - kommt bei komplexeren Modellen an seine Grenzen

2.1 Starten und Beenden von SPSS

- Nach dem Starten von SPSS erscheint eine Dialogbox.
- Die Dialogbox fasst häufig benötigte Aktionen zusammen, z.B. *Neues Dataset* und *Zuletzt verwendete Dateien*.
- Es gibt Möglichkeiten, sich über Neuerungen zu informieren, die SPSS Community zu besuchen, und erste Schritte mit Hilfe des SPSS-Lernprogramms zu tätigen.
- Das Dialogfeld kann optional deaktiviert werden.
- Für die Beendigung von SPSS gibt es mehrere Optionen.
- Man kann über das Menü die Aktion Datei → Beenden auswählen.
- Alternativ kann man die in SPSS geöffneten Fenster einzeln schließen.
- Beim Schließen einer Datei, in der Änderungen vorgenommen wurden, fragt SPSS, ob die Datei gespeichert werden soll.

2.2 SPSS-Grundstruktur

- Drei Hauptfenster: Daten-Editor, Viewer und Syntaxfenster.
- Gemeinsamkeiten: Titelzeile, Menüleiste, Statusleiste und Symbolleiste.
- Inhalte können separat gespeichert werden.
- Parallelöffnung von Fenstern möglich.

2.2.1 Daten-Editor

- Startet automatisch mit SPSS.
- Hier sieht man, wie der Name schon nahelegt, die Daten
- Zwei Ansichten: Datenansicht und Variablenansicht.
- Spreadsheet-ähnlicher Aufbau.
- Speichern führt zu Datendatei mit Endung '.sav'.

2.2.2 Viewer

- Separates Fenster, wird nicht automatisch mit SPSS gestartet. Wird dann automatisch geöffnet, wenn ihr etwas ausführt.
- Dokumentiert Datenaufbereitungs- und Analyseschritte.
- Zeigt Ergebnisse von statistischen Prozeduren.

2.2.3 Syntax-Editor

- Erlaubt Steuerung aller SPSS-Prozeduren über Syntax.
- Zugang: Datei → Neu → Syntax.
- Zweiteiliger Aufbau:
 - Links: Gliederung (auskommentierter Text, Syntaxbefehle).
 - Rechts: Eingabefenster für Syntax.
- Speichern als .sps-Datei: Datei → Speichern.

2.3 Prinzipielles zur Verwendung der Syntax

- Arbeiten mit SPSS Syntax bietet Vorteile.
- Aber: Erlernen kann anfangs herausfordernd sein.
- Empfehlung: Anfänger sollten über Menü starten.
- Erfahrung mit anderen Statistikprogrammen (z.B., SAS, R) kann helfen.

2.3.1 Vor- und Nachteile der Syntax

Vorteile:

- Erhöhte Effizienz.
- Wiederverwendbarkeit und Modifizierung.
- Protokollierung von Auswertungsschritten.
- Vorbereitung auf andere Software.
- Zugang zu speziellen Prozeduren.

Nachteile:

- Erlernensaufwand.
- Anwenderfreundlichkeit.
- Aufwendige Fehleranalyse.

2.3.2 Grundlagen zur Syntax und Anwendung

- Syntax kann auch indirekt über das Modul generiert werden. Dazu klickt ihr einfach auf die Analyse eurer Wahl und klickt anschliessend auf *Einfügen* statt auf *OK*.
- Dokumentation und Reproduktion von Auswertungsschritten.
- Veröffentlichung von Syntax und Daten in Forschung.

2.3.2 Bausteine der Syntax

- Anweisungen steuern SPSS.
- Unteranweisungen sind Zusätze zu Anweisungen.
- Spezifikationen für Details in Anweisungen/Unteranweisungen.
- Schlüsselwörter sind vordefinierte Wörter.

Regeln:

- Anweisung in neuer Zeile, Ende mit Punkt.
- Unteranweisungen durch Schrägstrich getrennt.
- Maximale Zeilenlänge: 80 Zeichen.

2.4 Arbeiten mit dem Viewer

- Effizientes Suchen, Löschen, Editieren, Umstellen von Inhalten.
- Sie können Zahlen ändern, die Darstellung ändern, Tabellen exportieren....
- Markieren durch Klicken in der Gliederung oder auf Parts.
- Löschen: Taste [Entf] oder Bearbeiten → Löschen.
- Umstellen: Teile ziehen oder Bearbeiten → Ausschneiden und Einfügen.
- Kopieren: Teile mit gedrückter Kontrolltaste ziehen oder Bearbeiten → Kopieren und Einfügen.

2.4 Bearbeiten von Inhalten im Viewer

- Titel, Texte, Tabellen und Grafiken können durch Doppelklick aktiviert und editiert werden.
- Tabellen über den Pivot-Tabellen-Editor.
- Grafiken über den Diagramm-Editor.
- Editieren von Titeln
 - Klicken auf das Objekt, wo der Text/Titel eingefügt werden soll.
 - Einfügen → Neuer Titel oder Neuer Text.

2.5 Überblick über die allgemeine Menüleiste

- Zugriff aus jedem Fenster des Programms.
- Übliche Menüpunkte eines Microsoft-Office-Programms.
- SPSS-spezifische Menüpunkte: Daten, Transformieren, Analysieren.

2.6 Allgemeine Voreinstellungen

- Empfehlung für Einsteiger: Voreinstellungen unverändert lassen.
- Voreinstellungen anpassen über Bearbeiten → Optionen.
- Registerkarten: Allgemein, Viewer, Ausgabe.

2.6 Mögliche Optionen

- **Allgemein:** Anzeige der Variablenliste, Sprache der Benutzeroberfläche.
- **Viewer:** Schriftart und Schriftgröße für Blöcke.
- **Ausgabe:** Informationen zu den Variablen in der Gliederung oder (Pivot-)Tabellen.

Aufbereitung der Daten

- Meistens sind die Daten nicht in einem Zustand, in dem sie analysiert werden können, z.B. liegen Ihnen analoge Fragebögen vor, die Daten müssen erst noch rekodiert werden, etc...
- Daher: Notwendigkeit der Datenaufbereitung **vor** der Analyse.
- Beispiele:
 - Transfer von Papierdaten zu SPSS
 - Zusammenführen unterschiedlicher Dateien
 - Rekodieren/Berechnen von Variablen
 - ...

3.1 Kodierung und Kodeplan

- Das Codebook (Codeplan...) ist **das wichtigste (!)** Instrument, was Sie haben, um Ihre Datenanalyse zu planen, zu strukturieren, zu gestalten.
- Codebook:
 - Tabelle über **alle** Variablen, die Sie in Ihrem Survey erheben
 - Enthält genauen Wortlaut der Items
 - Enthält die Antwortmöglichkeiten
 - Enthält die Quellen
 - Ordnet Fragen des Fragebogens Variablennamen zu.
 - Ordnet Merkmalsausprägungen Nummern zu.
- Codebook und Codierung gewährleistet Eindeutigkeit und statistische Verarbeitung.
- Wenn Sie nach 6 Monaten noch wissen wollen, wie genau eine Variable erhoben wurde, werden Sie dankbar für ein sauber angelegtes Codebook sein...

3.3 Anlegen von Variablen in SPSS

- Variablenansicht in SPSS nutzen.
 - Einfach im Feld **Name** einen neuen Variablennamen eintragen
 - Beginn mit einem Buchstaben.
 - Keine Leerzeichen oder spezifische Sonderzeichen.
 - Maximale Länge: 64 Zeichen.
 - Groß- und Kleinschreibung ist variabel.
- Beachtung von Regeln für Variablennamen.
- Beispiel: Variablen v2_zuf, v11_alter und v18_chron_Erkrankung.

3.3.1 Vergabe von Wertelabels

- Klicken Sie mit der linken Maustaste auf die Schaltfläche mit den drei Punkten.
- Eine Dialogbox öffnet sich (siehe Abb. 3.3).
- Bei *Wert* können Sie den Wert der Merkmalsausprägung eingeben.
- Bei *Beschriftung* können Sie eine nähere Beschreibung eingeben.
- Eingaben mit OK bestätigen.

Beispielvariablen:

- v2 1 = ausgezeichnet, 2 = gut, 3 = weniger gut, 4 = schlecht,
9 = keine Angabe
- v11 999 = keine Angabe

3.3.3 Fehlende Werte

- Jede empirische Erhebung kann fehlende Werte aufweisen.
- Gründe: Übersehen von Fragen, Verweigerung von Antworten, Übernahme kompletter Datensätze.
- Bei SPSS gibt es zwei Möglichkeiten:
 - Benutzerdefinierte fehlende Werte
 - Benutzer legt einen Wert für fehlende Werte fest (zB 999)
 - Dieser Wert muss in SPSS definiert werden.
 - Systemdefinierte fehlende Werte
 - Lassen Sie die Zelle in der Datenansicht leer.
 - SPSS erkennt dies als fehlenden Wert.
 - In der Datenansicht erscheint ein Punkt (.).

3.3.4 Messniveau

- Sie können für jede Variable in SPSS ein Messniveau angeben (Nominal, ordinal, metrisch)
- Bei neuen Variablen ist das Messniveau standardmäßig *Unbekannt*.

3.3.6 Verschieben und Löschen von Variablen

- Variablen können in der Variablenansicht verschoben werden.
- Um eine Variable zu löschen: Markieren Sie die Zeile und klicken Sie mit der rechten Maustaste, dann wählen Sie "Löschen".

Manuelle Dateneingabe in den Daten-Editor

- Wenn Sie Daten z.B. in Papierform erheben, müssen Sie die Daten manuell in den Dateneditor eingeben
- Zunächst legen Sie alle Variablen an, dann geben Sie die Werte ein - wie in Excel, selbsterklärend...
- Wechsel in die Datenansicht zur Nutzung des Daten-Editors
- Namen der Variablen im Kopf der Tabelle sichtbar
- Zeilenweise Dateneingabe empfohlen (ein Fragebogen = eine Zeile)

3.5 Einlesen von Datendateien

- Einlesen von SPSS-Formaten
 - Dateiendung: .sav
 - Einlesen über Menü: Datei → Öffnen → Daten
 - Beispiel: Einlesen der Datei Zentrum1.sav
- Einlesen von anderen Formaten
 - SPSS kann viele externe Datenformate einlesen
 - Wichtig: Variablenbezeichnungen sollten mit SPSS kompatibel sein
 - Beispiel: Einlesen der Datei Routinedaten.xls

Zusammenfügen von Datendateien

- Möglichkeit: „Fallweise zusammenfügen“: Für gleiche Variablen, aber unterschiedliche Fälle
- Möglichkeit: „Variablenweise zusammenfügen“: Für gleiche Fälle, aber unterschiedliche Variablen

3.7.2 Häufigkeitsauszählungen

Erstellung von Häufigkeitstabellen für kategoriale Variablen (nominal oder ordinal).

- Menü: Analysieren → Deskriptive Statistiken → Häufigkeiten
- Dialogbox "Häufigkeiten"
- Auswahl und Anzeigoptionen setzen
- Interpretation
 - Analyse der Häufigkeitstabelle
 - Erkennen von Eingabefehlern

3.7.3 Berechnung von Kennwerten

Berechnung von Kennwerten für Variablen mit Intervall- oder Verhältnisniveau.

- Menü: Analysieren → Deskriptive Statistiken → Häufigkeiten
- Dialogbox "Häufigkeiten: Statistik"
- Auswahl und Anzeigoptionen setzen
- Interpretation
 - Vergleich von Modus, Median und Mittelwert
 - Erkennen von Ausreißern und Fehlern

4 Modifikation von Variablen

Nach erfolgreicher Datenaufbereitung sind oft weitere Datenmodifikationen nötig:

- Umcodieren von Variablen
- Berechnung neuer Variablen

4.1 Umcodieren von Variablen

- Umcodieren: Bestehende Variablenausprägungen nach definierten Regeln ändern.
- Beispiel: Altersangaben (in Jahren) zu Altersklassen.
- Ergebnis: Informationsreduktion und Veränderung des Messniveaus (z.B., Intervall zu Ordinal).

Gründe für das Umcodieren:

- Nicht benötigter Detaillierungsgrad
- Bildung von Kategorien (z.B., Altersklassen)
- Umcodieren von Einzelitems für Summenskalen
- Uneinheitliche Kodierung

SPSS bietet zwei Umcodierungsoptionen:

- ① In dieselbe Variable
- ② In eine neue Variable

4.1.1 Umcodieren in dieselbe Variable

- Beispiel: Umcodieren der Items v2_zuf, v4_zuf, v7_zuf, v8_zuf für Patientenzufriedenheit.
- Ziel: Aussagekräftige Summenskala (hohe Werte = hohe Zufriedenheit).
- Umcodierregel: $1 \rightarrow 4$, $2 \rightarrow 3$, $3 \rightarrow 2$ und $4 \rightarrow 1$.
- Prozess in SPSS: Menü \rightarrow Transformieren \rightarrow Umcodieren \rightarrow In dieselben Variablen.

4.1.2 Umcodieren in andere Variablen

- Wählen Sie im Menü: *Transformieren* → *Umcodieren* → *Umcodieren in andere Variablen*.
- Dialogbox öffnet: Umcodieren in andere Variablen.
 - Quellvariablenliste enthält alle Variablen.
 - Wähle Variable(n) aus. Bei mehreren: gleicher Typ.

Prozedur

- Wählen: Transformieren → Variable berechnen
- Dialogbox: Variable berechnen
- Klassische Anwendung: Summenskala berechnen
- Wichtig: Vorher fehlende Werte bereinigen!
- Die Dialogbox gibt uns vielfältige Möglichkeiten, die verschiedensten Operatoren zu nutzen