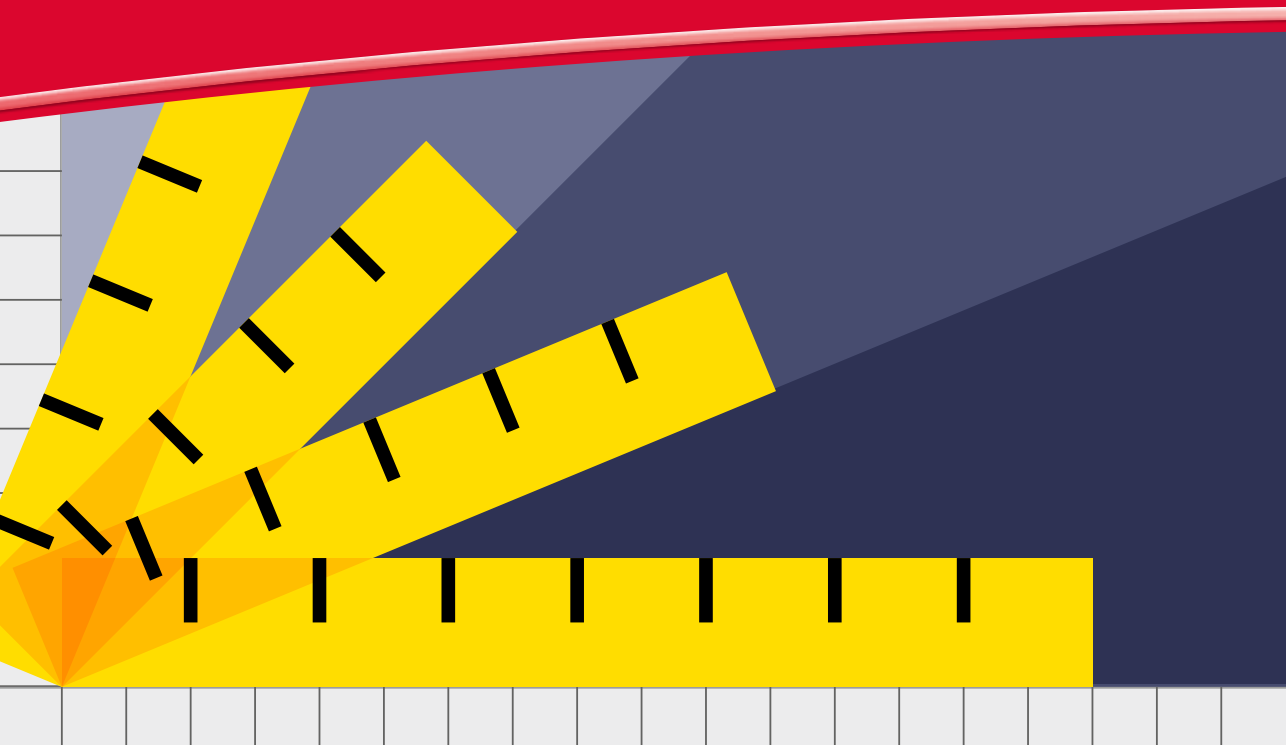


utb.

Claus Braunecker

# How to do Empirie, how to do SPSS

Eine Gebrauchsanleitung



## B 4 | Analyse quantitativer Daten (mit SPSS bzw. PSPP)

... in diesem Kapitel geht's um:

- |  |
|--|
| <ul style="list-style-type: none"> <li>• Im Mittelpunkt dieses Kapitels stehen die <b>softwaregestützte Datenanalyse</b> und das <b>Verständnis allgemeiner statistischer Zusammenhänge</b></li> </ul> |
| <ul style="list-style-type: none"> <li>• Basis ist ein <b>eigener Fragebogen mit korrespondierendem Datenfile</b> • downloadbar unter <a href="http://howtodo.at">howtodo.at</a></li> </ul>            |
| <ul style="list-style-type: none"> <li>• Abbildungen und Erläuterungen gelten für SPSS • sind fast immer auch auf PSPP anwendbar</li> </ul>  |

### B 4.1 | Quantitative Analysesoftware

Qualitative Datenanalysen sind (mit Einschränkungen) oft auch ohne Software möglich. Da **quantitative Auswertungen** und das Darstellen von Häufigkeiten, Prozentzahlen, Mittelwerten und anderen Maßzahlen hingegen meist auf großen Datenmengen basieren, kommt man ohne Datenanalysesoftware nicht weit.

Das gesamte **KAPITEL B 4** widmet sich deshalb der **softwaregestützten Datenanalyse**. Basis dazu bildet ein **eigener Fragebogen** mit korrespondierendem **Datenfile** (**BUCHDATEN**, downloadbar unter [howtodo.at](http://howtodo.at)). Um effizient arbeiten zu können, wird auch dem Verständnis allgemeiner statistischer Zusammenhänge die notwendige Aufmerksamkeit gewidmet.

Auswertungssoftware gibt es in beträchtlicher Zahl. Die Bandbreiten reichen von sehr einfacher bis hochkomplexer Bedienung, von Freeware bis zu Lizenzkosten von mehreren 1.000 Euro, von weiter Verbreitung bis zum Einsatz für (sehr) spezielle Nutzergruppen.

Dieses Buch demonstriert quantitatives Auswerten am Statistik-Paket **SPSS** (**S**tatistical **P**ackage for the **S**ocial **S**ciences). Die Software startete in den 1960er-Jahren ihre weltweite Verbreitung in San Francisco und wurde 2009 von IBM übernommen. Für dieses Buch wurde die seit März 2016 aktuelle **Version 24** herangezogen. SPSS wird sowohl in der Wirtschaft als auch in vielen (universitären) Forschungsbereichen eingesetzt. Lizenzen sind direkt bei IBM oder über Vertriebspartner erhältlich. Für Studierende gibt es spezielle Angebote einzelner Bildungseinrichtungen. Auch im Netz sind Studierendenlizenzen ([studyhouse.de](http://studyhouse.de)), auf der SPSS-Homepage Testversionen ([spss.com](http://spss.com)) zum Download verfügbar.

An dieser Stelle soll die **Freeware PSPP** nicht unerwähnt bleiben. Das Spiel mit den Buchstaben kommt nicht von ungefähr: Bei dem Programm handelt es sich um eine SPSS „nachgebaute“ Software, die laufend ergänzt und verfeinert wird. Natürlich ist das Original-Produkt SPSS deutlich ausgereifter, es bietet eine attraktivere Benutzeroberfläche und weit größere Programmviefalt. PSPP kann da und dort ein wenig „buggy“ sein, der Ergebnisoutput ist sehr einfach gehalten. Wenn aber jemand erste Auswertungsschritte unternehmen möchte, stellt das Programm zwar einen reduzierten, jedoch kompatiblen Ersatz dar: Es verwendet das SPSS-Datenformat und vergleichbare Menübefehle.

**Alle Ausführungen und Screenshots der folgenden Kapitel gelten für SPSS. Wegen seiner Nähe zu SPSS sind fast alle Erläuterungen auch auf PSPP anwendbar.**

## B 4.2 | SPSS Programm-Basics

... in diesem Kapitel geht's um:

- **SPSS** ist stark **fensterorientiert**: Der **Dateneditor** zeigt die Daten (Datenansicht) und Variableneinstellungen (Variablenansicht) • das **Ausgabefenster (Viewer)** liefert Ergebnisse, Programm Meldungen und Warnhinweise • der optionale **Syntaxeditor** enthält den abspeicherbaren und wiederverwendbaren Befehlstext, den die Menü-Klicks hervorrufen
- Die **Statusleiste** am unteren Fensterrand zeigt Infos zum Programmablauf sowie ob die Daten gefiltert oder gewichtet vorliegen
- **Dialogfelder** listen links die vorhandenen, rechts die für die jeweilige Prozedur ausgewählten Variablen • Verzweigungen zu Untermenüs finden sich am rechten Rand • Einfügen der Syntax und Zurücksetzen des Dialogfelds erfolgt am unteren Rand
- Alle Programmeinstellungen sind über **Bearbeiten** → **Optionen...** steuerbar
- Das **Hilfe-System** ist vielseitig einsetzbar • **Tabellenvorlagen** erlauben die Definition eigener Auswertungslayouts • die **Menü-Übersicht** am Kapitelende listet die gebräuchlichsten Befehle

Öffnet man SPSS, zeigt sich zunächst ein **Start-Dialogfeld** (VGL. ABBILDUNG 65 AUF SEITE 204): Dort lassen sich neue Dateien anlegen (❶), zuletzt verwendete öffnen (❷) und Neuerungen zur aktuellen Programmversion, zur SPSS-Community bzw. Lernprogramme anzeigen (❸).

Jene, die nicht bei jedem Programmstart das Start-Dialogfeld sehen möchten, können es mit ❹ wegschalten. Andere, die gerne mit einem Klick auf zuletzt geöffnete Dateien zugreifen wollen, arbeiten bei jedem Programmstart mit ❷.

SPSS ist ein **stark „fensterorientiertes“ Programm**. Man findet sich in der Bedienung nur dann gut zurecht, wenn man die Bedeutung von insgesamt drei Windows-Fenstern kennt, durch die das Programm strukturiert ist. Deshalb werden an dieser Stelle zunächst die SPSS-Programmfenster beschrieben. Dazu wird das Start-Dialogfeld aus ABBILDUNG 65 zunächst weggeklickt (❺).

Eine SPSS-Sitzung beginnt mit dem Öffnen oder Neuerstellen von Daten. SPSS-Daten besitzen die Erweiterung **\*.sav** und lassen sich durch Doppelklick auf den Datenfile oder den Menübefehl **Datei** → **Öffnen** → **Daten...** öffnen. Die Genese eines neuen Datenfiles wird in KAPITEL „B 4.3.1 | DATEN ÖFFNEN, IMPORTIEREN ODER NEU ERFASSEN“ AUF SEITE 216 F. behandelt.

### B 4.2.1 | SPSS Dateneditor

Der **Dateneditor** ist das erste Fenster bei Programmstart. Er besitzt eine Datenansicht (VGL. ABBILDUNG 66 AUF SEITE 204) und eine Variablenansicht (VGL. ABBILDUNG 67 AUF SEITE 205).

In der **Datenansicht** wird der Inhalt der Datendatei angezeigt. In der ersten Zeile befinden sich die Variablennamen (❶), das kleine Symbol vor jedem Variablennamen kennzeichnet dessen in der Variablenansicht einstellbares Messniveau – nominal, ordinal, metrisch oder Text/alphanumerisch. Die Zeilen darunter enthalten jeweils pro Zeile einen Datensatz (❷).

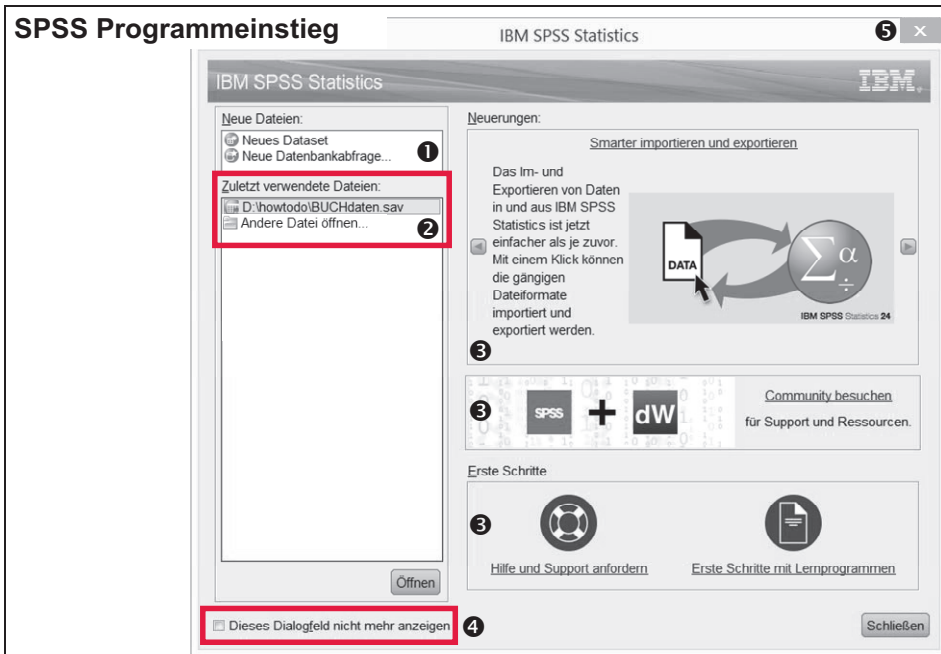


Abbildung 65: Programmeinstieg in SPSS

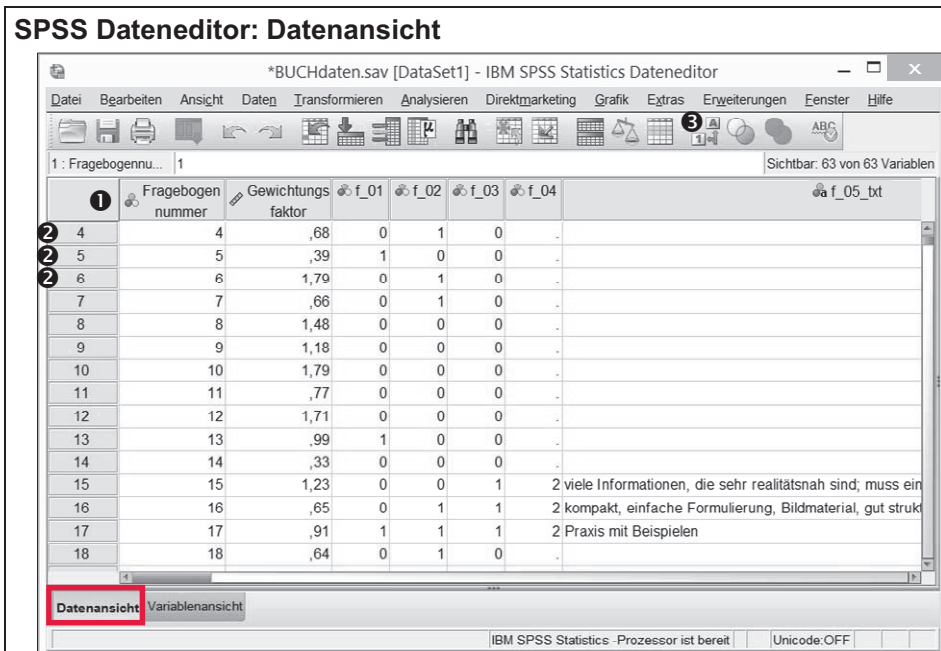


Abbildung 66: SPSS Dateneditor (Datenansicht)

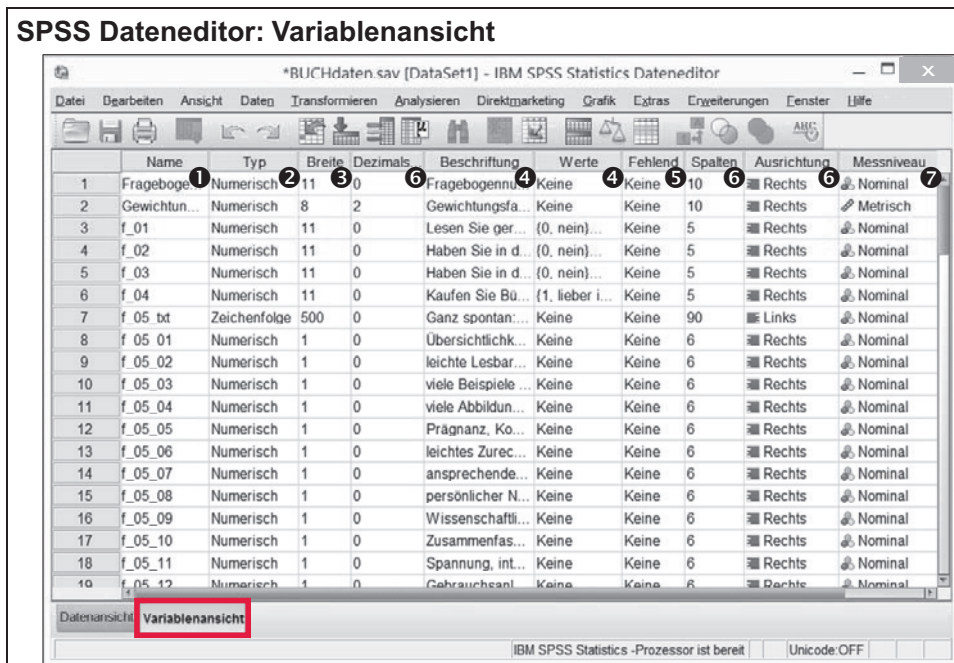



Abbildung 67: SPSS Dateneditor (Variablenansicht)

In jeder Datenzeile werden wahlweise die Codierungen oder deren Beschriftungen (VGL. KAPITEL „B 4.3.3 | VARIABLEN- UND WERTEBESCHRIFTUNGEN (LABELS)“ AB SEITE 224) angezeigt. Das Umschalten zwischen Codierungs- und Beschriftungsanzeige erfolgt mit ③ IN ABILDUNG 66 AUF SEITE 204. In alle Zellen können Daten eingegeben werden.

Die Position einer Variablen lässt sich verändern:

 Dazu wird die gesamte Spalte in der Spaltenüberschrift ① markiert, mit der linken Maustaste festgehalten und per drag & drop verschoben.

Daten kann man ändern, löschen, ausschneiden, kopieren, einfügen.

Leere (numerisch codierte) Zellen werden zu **fehlenden Werten** („Missings“, sie enthalten einen Punkt „.“; VGL. KAPITEL „B 4.3.4 | FEHLLENDE WERTE AB SEITE 226).

Die **Variablenansicht** des Dateneditors (VGL. ABILDUNG 67) zeigt und ermöglicht wichtige Einstellungen für alle Variablen des Datenfiles.

Die **Variablenamen** ① sind meist knappe, eher kryptische Bezeichnungen, die sich an Fragennummern in Fragebögen oder der Kurzbeschreibung von Merkmalen orientieren. Sie sollten bewusst kurz und prägnant formuliert werden, da mit ihnen alle Auswertungen „gesteuert“ werden: Strukturierte Formulierungen der Variablenamen helfen dabei, die Übersicht zu behalten (VGL. ① – DIE BENENNUNG DER VARIABLEN NACH DEN NUMMERN DER FRAGEN BEI DEN BUCHDATEN: f\_01 BIS f\_17).

*So ist es z.B. besser, eine Variable ALTER oder f\_01 zu nennen anstatt VAR322.*

Variablenamen müssen eindeutig sein und mit einem Buchstaben beginnen, das letzte Zeichen darf kein Punkt sein. Sie dürfen keine Leer- und Sonderzeichen enthalten. Bestimmte Buchstabenkombinationen sind für SPSS-Programmbezeichnungen reserviert (z.B. all, ne, eq, to, with ...) und dürfen nicht in Variablenamen vorkommen.

Kurze Variablenamen erleichtern zwar die Auswertung, der Ergebnisoutput wird dadurch aber schwer lesbar. Deshalb müssen die Variablenamen mit **Beschriftungen** (④) versehen werden. Damit enthält der Output nicht länger unverständliche Kurz-Bezeichnungen.

Soll der Ergebnisoutput anstelle aussagenloser numerischer Codierungen auch anzeigen, was sich „hinter den Zahlen verbirgt“, können auch die **Werte** (④) beschriftet werden.

*Die Beschriftung der Variable f\_12 der BUCHDATEN könnte „Geschlecht der Befragten“ lauten. Code 1 würde mit „weiblich“, Code 2 mit „männlich“ bezeichnet werden.*

Weitere Details zu Variablen- und Wertebeschriftungen finden sich in KAPITEL „B 4.3.3 | VARIABLEN- UND WERTEBESCHRIFTUNGEN (LABELS)“ AB SEITE 224.

Als **Variablentyp** (②) lässt sich eine Vielzahl von Formaten einstellen. Empfehlenswert ist es jedoch, sich auf die Typen **Numerisch** und **Zeichenfolge** zu beschränken. Andere Formate können die spätere Auswertung deutlich verkomplizieren. Gibt man in Textvariablen Zahlen ein, werden diese als Text interpretiert: Mit diesen Variablen können dann keine numerischen Operationen wie z.B. Mittelwertberechnungen durchgeführt werden.

Die **Breite** der Spalten (③) hat weniger Relevanz und sollte nicht verändert werden. Gibt man hier – v.a. bei Textvariablen – eine zu kleine Zahl ein, werden die Einträge „abgeschnitten“ und damit unbrauchbar.

Die Spalte **Fehlend** (⑤) lässt die Eingabe von Werten zu, die bei der Auswertung wie nicht vorhandene Werte („fehlende Werte“) behandelt werden sollen (VGL. KAPITEL „B 4.3.4 | FEHLENDE WERTE“ AB SEITE 226).

Die Angaben zu Dezimalstellen, Spaltenbreite und Ausrichtung (⑥) sind eher vernachlässigbar. Keine besondere Relevanz besitzt auch die Einstellung des Messniveaus (⑦): Dieses wird vom Programm zumeist aufgrund der Datenlage richtig erkannt. Außerdem verlangen nur wenige Auswertungsdialoge nach einem „korrekten“ Skalenniveau. Ist in speziellen Dialogfeldern das Messniveau dann doch von Relevanz, lässt es sich an diesen Stellen bequem per „**rechter Maustaste über dem Variablenamen**“ verändern.

Im Menü **Ansicht** kann das Aussehen des Dateneditors (Schriftarten, Gitterlinien, Wertelabels) beeinflusst werden. Informationen zum aktuell geöffneten Datenfile (Variablen, deren Formate, Beschriftungen usw.) bekommt man über **Datei** → **Datendatei-Informationen anzeigen**, indem man dort **Arbeitsdatei** auswählt.

VOR der ersten Datenanalyse müssen die Daten unbedingt **bereinigt** (= plausibilisiert) vorliegen (VGL. DAZU KAPITEL „B 2.3 | DATEN PLAUSIBILISIEREN (SCREENEN)“ AB SEITE 190)!

## B 4.2.2 | SPSS Ausgabefenster

SPSS legt bei jeder Sitzung automatisch ein **Ausgabefenster (Viewer)** an. Die Ergebnisse