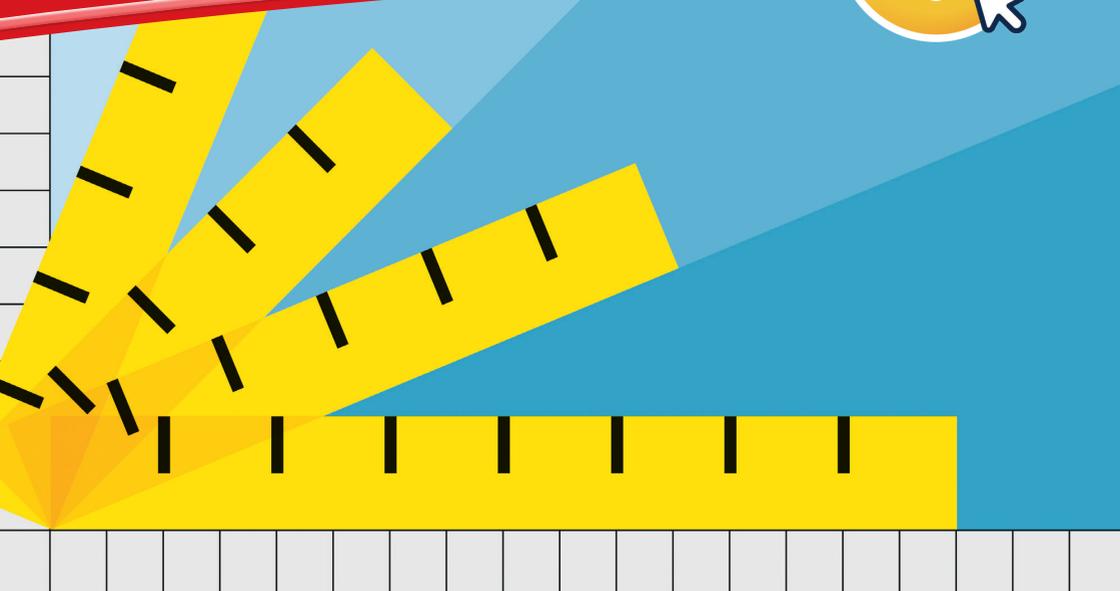


Claus Braunecker

# How to do Statistik und SPSS

Eine Gebrauchsanleitung

2. Auflage



## 5 | Analyse quantitativer Daten: Statistische Grundlagen

### ▼ Abstract (in diesem Kapitel geht's um ...) ▼

#### • Deskriptive Statistik:

beschreibt vorhandene Daten, geht aber nicht generalisierend darüber hinaus

**Analyse einzelner nominaler, ordinaler oder (quasi-)metrischer Variablen:** erläutert werden absolute Häufigkeiten | Arten von Prozentwerten (kumulierte, gültige) | Zentralmaße (Mittelwert, Median, Modus) | Lagemaße (Quantile – Terzile, Quartile, Quintile, Dezile usw.) | Minimum, Maximum | Modus, Spannweite, Quartilsabstand, Dezilabstand | Ausreißer, Extremwerte | Streuungsmaße (Varianz, Standardabweichung) | Normalverteilung (symmetrische Verteilung, Schiefe, Wölbung)

**Kreuztabelle zur Zusammenhangsdarstellung nominaler oder ordinaler Variablen:** erläutert werden absolute Häufigkeiten | Zeilen- und Spaltenprozent | abhängige und unabhängige Variablen | beobachtete und erwartete Werte | quadratische Kontingenz  $\chi^2$  | Zusammenhangsmaß Cramérs  $V$

**Mittelwertvergleich zur Zusammenhangs- bzw. Vergleichsdarstellung (quasi-)metrischer Variablen nach nominalen (ordinalen) Untergruppen ODER (quasi-)metrischer Variablen im Vergleich zueinander:** erläutert werden tabellarische und grafische Darstellung (Boxplot) | abhängige und unabhängige Vergleiche | abhängige und unabhängige Variablen

**Korrelation zur Zusammenhangs- bzw. Vergleichsdarstellung ordinaler oder (quasi-)metrischer Variablen:** erläutert werden Kovarianz | Korrelationskoeffizient nach Pearson | Rangkorrelation nach Spearman mit und ohne Bindungen

#### • Schließende Statistik:

schließt von Stichproben generalisierend auf dahinterstehende Grundgesamtheiten

- erläutert werden der Begriff „Signifikanz“ | inhaltliche Hypothese, statistische Null- und Alternativhypothese | Signifikanzniveau und Irrtumswahrscheinlichkeit | p-Wert | Signifikanz, Stichprobengröße und Unterschiedsstärke | ein- und zweiseitige Signifikanz | Freiheitsgrade

bei **Einzelvariablen:** statistische Schwankungsbreiten | Mittelwert-Konfidenzintervall • bei **Kreuztabellen:**  $\chi^2$ -Test • bei **Mittelwertvergleichen:** Fehlerbalkendiagramme | Mittelwertvergleichstest (Parameterverfahren und parameterfreie Verfahren)

- bei **Korrelationen:** t-Statistik

Eine quantitative Datenanalyse (vgl. Kapitel 6 ab Seite 105) ist nur mit Kenntnis statistischer Grundlagen sinnvoll durchführ- und interpretierbar. Das folgende Kapitel fasst deshalb die notwendigen statistischen Begriffe überblicksmäßig zusammen. Alle Analyseschritte referenzieren dabei auf die Messniveaus in Kapitel 1.2 ab Seite 20.

Die Fallzahlen sind in den meisten der folgenden Berechnungsbeispiele zur besseren Übersicht klein gehalten. In einem Realszenario wären sie deshalb für Ergebnisgeneralisierungen, die hier wiederholt zur Demonstration durchgeführt werden, NICHT geeignet!

Auch bei den geprüften „Zusammenhängen“ steht „leicht fassbar“ über wissenschaftlich!

Am Ende dieses Kapitels findet sich eine zusammenfassende Übersicht, bei welchem Messniveau und welcher Merkmalskombination welche statistischen Maßzahlen und Analyseverfahren die richtigen sind (vgl. Kapitel 5.3 ab Seite 102).

## 5.1 | Deskriptive Statistik

Datenanalysen beginnen meist rein deskriptiv. Die **deskriptive Statistik** beschreibt und ordnet Merkmalsverteilungen – sie stellt sie dar. Das dient rein zur Informationsverdichtung der Einzeldatensätze. Oft deckt deskriptives Auswerten auch Zusammenhänge zwischen einzelnen Merkmalen auf.

Alle Analysen beziehen sich auf die vorliegenden Datensätze: Sie beschreiben die vorhandene Datenlage und „gehen nicht darüber hinaus“.

Es erfolgt noch keine Projektion auf eine durch die Daten eventuell repräsentierte (vgl. Braunecker, 2023, S. 61–68) **Grundgesamtheit** (vgl. Braunecker, 2023, S. 45–47) – **Ergebnisdaten** entstammen ja oft einer Stichprobe (vgl. Braunecker, 2023, S. 68–76): Das Schließen von Stichproben auf dahinterstehende Grundgesamtheiten ist Aufgabe der **schließenden Statistik** (vgl. Kapitel 5.2 ab Seite 81).

### 5.1.1 | Häufigkeiten (absolut und prozentuell)

**Absolutzahlen** und **Prozentwerte** sind die einfachsten und in der Praxis beliebtesten Maßzahlen zur Beschreibung kategorialer (nominaler bzw. ordinaler) Daten. Häufigkeitszählungen ermitteln, wie oft jede Merkmalsausprägung in den Daten vorkommt.

Häufigkeitszählungen sollten zu Beginn jeder Auswertung bei allen Merkmalen erfolgen – noch vor jeder weiteren Datenanalyse. Dabei sind alle zulässigen und unzulässigen Codierungen (vgl. Kapitel 4.4 ab Seite 47) sowie die Missing Values (vgl. Kapitel 6.3.5 ab Seite 130) rasch überblickbar.

Zu unterscheiden ist zunächst zwischen absoluten und relativen Häufigkeiten. Statistisch formuliert: Eine absolute Häufigkeit gibt an, wie oft eine Ausprägung  $i$  des Merkmals  $X$  in einer Grundgesamtheit mit dem Umfang  $N$  oder einer Stichprobe mit dem Umfang  $n$  auftritt (= „Absolutzahlen“). Die relative Häufigkeit kennzeichnet, wie oft eine Ausprägung  $i$  des Merkmals  $X$  im Verhältnis zu einer Grundgesamtheit mit dem Umfang  $N$  oder im Verhältnis zu einer Stichprobe mit dem Umfang  $n$  auftritt (= „Prozentwerte“).

Dazu kommen – in SPSS immer standardmäßig ausgewiesen – gültige und kumulierte Prozentwerte. Abbildung 19 auf Seite 54 veranschaulicht und erläutert die einzelnen Arten von Häufigkeiten bzw. Prozentwerten. Die Tabelle zeigt eine in fünf Klassen (= Gruppen) zusammengefasste (vgl. Kapitel 6.4.1 ab Seite 133) Häufigkeitsauszählung der Frage 16 der BUCHdaten (vgl. Frage 16 des Fragebogens auf Seite 18).

Absolute **Häufigkeiten** zählen, wie oft jede Merkmalsausprägung in den Daten vorkommt (❶). „Normale“ **Prozentwerte** (❷) relativieren die Häufigkeiten an ALLEN Datensätze (❶). **Gültige Prozentwerte** (❸) schließen fehlende Werte aus (vgl. Kapitel 6.3.5 ab Seite 130).

## Arten von Häufigkeiten und Prozentwerten

- ① **Häufigkeit** = Anzahl der Fälle je Merkmalsausprägung  
▶ absolute Häufigkeiten
- ② **Prozent** = Anzahl der Fälle je Ausprägung, relativiert zu ALLEN Fällen  
▶ relative Häufigkeiten
- ③ **Gültige Prozente** = Anzahl der Fälle je Ausprägung, relativiert an nur jenen Fällen, die eine Merkmalsausprägung haben (also ohne „Fehlend“)
- ④ **Kumulierte Prozente** = Prozentwerte, summiert in steigender Reihenfolge der Merkmalsausprägungen, beginnend beim kleinsten Wert  
▶  $29,2 + 20,7 = 49,9$ ;  $49,9 + 15,0 = 64,9$  usw.

Wie viele Bücher lesen Sie im Schnitt pro Jahr? (in Klassen)

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1 Buch	138	27,3	29,2	29,2
	2 Bücher	98	19,4	20,7	49,9
	3 bis 5 Bücher	71	14,1	15,0	64,9
	6 bis 10 Bücher	63	12,5	13,3	78,2
	mehr als 10 Bücher	103	20,4	21,8	100,0
Gesamt	473	93,7	100,0		
Fehlend	System	32	6,3		
Gesamt		505	100,0		

Abbildung 19: Arten von Häufigkeiten und Prozentwerten

Die **kumulierten Prozentwerte** (④) zählen zum (gültigen) Prozentsatz der ersten Kategorie (⑤) den (gültigen) Prozentsatz der zweiten (⑥) dazu. Zu dieser Summe wird die dritte Kategorie (⑦) addiert. Zur neuen Summe kommt die vierte Kategorie (⑧) usw. ⑨ in Abbildung 19 kann z.B. wie folgt interpretiert werden:

*Rund die Hälfte der Befragten, die eine Antwort gegeben haben (49,9%), lesen bis zu zwei Bücher pro Jahr. Die Gegenmenge – die zweite Hälfte mit 51,1% – liest mehr als zwei Bücher.*

Für die in Abbildung 19 dargestellte Auswertung wurde  $f_{16}$  „klassiert“: Besitzt ein – metrisches (vgl. Braunecker, 2023, S. 106) – Merkmal für eine Häufigkeitszählung zu viele Ausprägungen, ist es möglich, die Ausprägungen zu Klassen zusammenzufassen. Dabei ist jedoch Vorsicht geboten: Je breiter die einzelnen Klassen definiert werden, desto weniger von ihnen sind letztendlich vorhanden – der Informationsgehalt sinkt.

*Jemand hätte  $f_{16}$  nicht wie im Beispiel zu den fünf Klassen – „1 Buch“ | „2 Bücher“ | „3 bis 5 Bücher“ | „6 bis 10 Bücher“ | „mehr als 10 Bücher“ – zusammenfassen können, sondern z.B. nur zu zwei: „1 bis 5 Bücher“ und „6 Bücher und mehr“: Dann ließe sich aus den Ergebnissen nur mehr herauslesen, dass knapp zwei Drittel (64,9 %) bis zu fünf Bücher lesen, der Rest mehr.*

Umgekehrt gilt aber auch: Je schmaler Klassen sind, desto mehr gibt es von ihnen – wieder sinkt der Informationsgehalt.

*Eine andere Person hätte aus  $f_{16}$  auch zehn oder mehr Klassen bilden können: „1 Buch“ | „2 Bücher“ | „3 oder 4 Bücher“ | „5 oder 6 Bücher“ | „7 oder 8 Bücher“ | „9 oder 10*

„Bücher“ | „11 oder 12 Bücher“ | „13 oder 14 Bücher“ | „15 oder 16 Bücher“ usw. Das stellt die gesamte Klassenbildung in Frage.

Die Wahl einer akzeptablen Klassenbreite wesentlich erleichtern können Quantile (vgl. Kapitel 5.1.2.5 ab Seite 59).

Eine sehr hohe Bedeutung kommt der **Basis von Prozentwerten** zu: Sie zu verschweigen oder gar nicht zu beachten, kann rasch zu großen Fehlinterpretationen führen.

Oft fehlt bei der Angabe von Prozentwerten die Information darüber, worauf sich die Relativierung bezieht. Dann bleibt die Menge unklar, die hinter dem Prozentwert steht.

*Ein Unternehmen verkauft zwei Arten von Fahrrädern – solche um 10.000 € und solche um 100 €. Da sich der Verkauf der billigen Fahrräder im vergangenen Jahr rückläufig entwickelt hat (minus 20%), der Verkauf der teuren Modelle hingegen sprunghaft angestiegen ist (plus 200%), stellt das Unternehmen die Produktion der billigen Räder ein und setzt nur noch auf Luxusmodelle. Ein Jahr später ist das Unternehmen nicht mehr am Markt.*

*Warum wohl? Im Vorjahr wurden 5.000 billige Räder (um 100 €) verkauft, heuer 4.000 (= minus 20%). Im Vorjahr wurde 1 Luxus-Rad (um 10.000 €) verkauft, heuer 3 (= plus 200%).*

50% aus 4 Fällen sind statistisch anders zu beurteilen als 50% aus 400!

An dieser Stelle soll noch einmal auf die Gefährlichkeit hingewiesen werden, bei kleiner Datenbasis Häufigkeiten oder Prozentwerte zu interpretieren: Bei weniger als 50 Datensätzen sollten Häufigkeitsauszählungen unterbleiben! Ebenso gefährlich sind Prozentwert-Interpretationen bei weniger als 70 Fällen oder nicht repräsentativen (vgl. Braunecker, 2023, S. 61–68) Daten.

Zu sehr wird damit eine hier NICHT mögliche Generalisierbarkeit der Ergebnisse suggeriert (vgl. die Ausführungen zu Fallzahlen von Stichproben in Kapitel 3.3 ab Seite 36). Die Darstellung von (qualitativen) Ergebnissen auf Basis kleiner Fallzahlen erfolgt deshalb besser verbalisierend (vgl. Braunecker, 2023, S. 155, linke Seite der Abbildung). Dieses Thema hängt auch stark mit der „Ergebnisgüte“ zusammen – Stichwort „statistische Schwankungsbreiten“ (vgl. Braunecker, 2023, S. 82–93). DAS ist aber ein zentrales Thema der schließenden Statistik, die von wenigen Datensätzen auf durch sie repräsentierte viele schließen möchte (vgl. Kapitel 5.2 ab Seite 81).

### 5.1.2 | Zentral- und Lagemaße

Zur Darstellung (quasi-)metrischer Daten dienen Zentral- und Lagemaße sowie Streuungsmaße (vgl. Kapitel 5.1.3 ab Seite 64). **Zentralmaße** sind Mittelwert, Median und Modus. Als **Lagemaße** bezeichnet werden Minimum, Maximum und Quantile – z.B. Terzile, Quartile, Quintile, Dezile usw.

#### 5.1.2.1 | Mittelwert (arithmetisches Mittel)

Mittelwert und Median (vgl. Kapitel 5.1.2.2 auf Seite 58) sind die in der Praxis beliebtesten Werte zur Beschreibung (quasi-)metrischer Verteilungen. Zur Berechnung eines **Mittelwerts (= arithmetisches Mittel)** werden alle Einzelwerte addiert und durch die Zahl der (gültigen) Messungen dividiert.



## 6 | Analyse quantitativer Daten mit SPSS (PSPP)

Der zweite Abschnitt dieses Buchs beschäftigt sich mit der softwaregestützten Auswertung empirischer Erhebungen.

Basis dazu bildet der **Fragebogen dieses Buchs** (vgl. Seite 17 und Seite 18) mit korrespondierendem **Datenfile** (BUCHdaten, beides frei zugänglich unter [howtodo.at](http://howtodo.at) bzw. auf [utb.de](http://utb.de)).



[www.utb.de/do/10.36198/9783838561615-m01](http://www.utb.de/do/10.36198/9783838561615-m01)

**Zusatzmaterial**

[howtodo.at/downloads/Fragebogen\\_BUCHdaten.pdf](http://howtodo.at/downloads/Fragebogen_BUCHdaten.pdf)





[www.utb.de/do/10.36198/9783838561615-m02](http://www.utb.de/do/10.36198/9783838561615-m02)

**Zusatzmaterial**

[howtodo.at/downloads/BUCHdaten.sav](http://howtodo.at/downloads/BUCHdaten.sav)



Die kapitelmäßige Gliederung folgt den Schritten üblicher Datenanalysen. Viele Querverweise verbinden zu den in vorderen Buchteilen skizzierten thematischen Grundlagen.

Der Beginn widmet sich der technischen Programmanwendung (Erläuterung der Fenster, vgl. Kapitel 6.2 ab Seite 106). In den nächsten Schritten werden verschiedene Möglichkeiten angesprochen, Datensätze zu handeln (vgl. Kapitel 6.3 ab Seite 120) und für die spätere Analyse eventuell noch besser zu strukturieren (vgl. Kapitel 6.4 ab Seite 133). Nach ausführlichen Darstellungen deskriptiver und schließender Analyseroutinen (vgl. Kapitel 6.5 ab Seite 143) gibt ein Entscheidungsbaum Überblick über alle Verfahren (vgl. Kapitel 6.6 ab Seite 193). Den Abschluss bilden beispielhafte Ausführungen zu Analyseplanung und Bericht (vgl. Kapitel 6.7 ab Seite 197).

### 6.1 | Datenanalysesoftware – SPSS und PSPP

Alle folgenden Erläuterungen basieren auf der seit Ende 2022 aktuellen **SPSS-Version 29**. Die meisten Ausführungen sind auch auf die **Freeware PSPP** anwendbar.

PSPP ist eine SPSS „nachgebaute“ Software. Details zu beiden Produkten finden sich gleich zu Beginn dieses Buchs im Kapitel „(Quantitative) Datenanalysesoftware“ ab Seite 15.



[ibm.com/de-de/products/spss-statistics](http://ibm.com/de-de/products/spss-statistics)



[gnu.org/software/pspp/](http://gnu.org/software/pspp/)

## 6.2 | SPSS anwenden

### ▼ Abstract (in diesem Kapitel geht's um ...) ▼

- **SPSS** ist stark **fensterorientiert** • der **Dateneditor** zeigt die Daten (Datenansicht) und Variableneinstellungen (Variablenansicht) • das **Ausgabefenster (Viewer)** liefert Ergebnisse, Programmmeldungen und Warnhinweise • der optionale **Syntaxeditor** enthält den speicherbaren und wiederverwendbaren Befehltext, den die Menü-Klicks hervorrufen
- Die **Statusleiste** am unteren Fensterrand gibt 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 erfolgen am unteren Rand
- Alle Programmeinstellungen sind über **Bearbeiten** ➤ **Optionen...** steuerbar
- Das **Hilfe-System** ist vielseitig einsetzbar • **Tabellenvorlagen** erlauben die Definition eigener Layouts • eine **Menü-Übersicht** am Kapitelende kategorisiert die gebräuchlichsten Befehle

SPSS öffnet sich mit einem Fenster (① in Abbildung 40), das die zuletzt geöffneten Datenfiles anzeigt (①) oder neue Daten anlegen lässt (②). Dieses Start-Dialogfeld lässt sich deaktivieren (③).

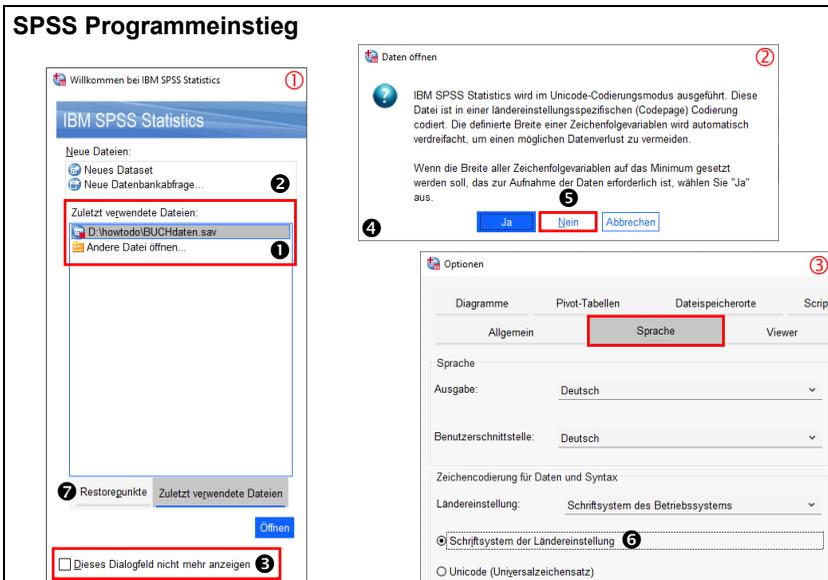


Abbildung 40: Programmeinstieg in SPSS

SPSS wird von seiner Grundeinstellung her im internationalen Zeichensatz (**Unicode-Modus**) ausgeführt (②). Bei Öffnen jeder länderspezifischen Datendatei (z.B. deutsch, mit Umlauten) wird eine Warnmeldung ausgegeben (④ in Abbildung 40).

 Um an keiner Stelle Probleme mit Umlauten auftreten zu lassen, empfiehlt es sich, hier auf **Nein** zu klicken (⑤).

Soll diese Meldung überhaupt vermieden werden, kann SPSS generell auf das „Schriftsystem der Ländereinstellung“ eingestellt werden (③ in Abbildung 40).

 Das ist unter **Bearbeiten** → **Optionen**<sup>61</sup> im Karteireiter **Sprache** möglich (⑥).

**Achtung:** Diese Einstellung ist nur zugänglich, wenn gerade kein Datenfile geöffnet ist! Sind Daten geöffnet, müssen diese – und damit SPSS – zuerst geschlossen und in weiterer Folge ohne Daten noch einmal geöffnet werden.

SPSS arbeitet auch mit Wiederherstellungspunkten (**Restorepunkte**, ⑦): Diese speichern Daten, wenn Programminstanzen unerwartet enden (ähnlich der aus Officeprogrammen bekannten automatischen Dateiwiederherstellung).

Daneben ist es möglich, Wiederherstellungspunkte als Momentaufnahme einer Programminstanz (mit allen begleitenden Dateien, vgl. die Kapitel 6.2.1 ab Seite 107, 6.2.2 ab Seite 110, und 6.2.3 ab Seite 112) direkt zu speichern bzw. zu öffnen.

 Das ist unter **Datei** → **Restorepunkt speichern...** bzw. **Datei** → **Restorepunkt öffnen...** möglich.

SPSS ist **stark „fensterorientiert“** – die Anwendung wird durch drei im Folgenden beschriebene Fenster strukturiert: die Daten (Dateneditor, in zwei Ansichten), das Ergebnis- und Meldungsfenster (Ausgabefenster, Viewer) sowie die Möglichkeit, mit Befehlssprache zu arbeiten (Syntaxeditor).

### 6.2.1 | Dateneditor

SPSS-Daten besitzen die Erweiterung **\*.sav**. Sie lassen sich bei Programmstart über das Start-Dialogfeld öffnen (① in Abbildung 40) oder neu anlegen (②).

 Darüber hinaus öffnen sich Datenfiles auch durch Doppelklick oder aus SPSS heraus über den Menübefehl **Datei** → **Öffnen** → **Daten...**

 Einen neuen Datenfile erzeugt **Datei** → **Neu** → **Daten** (vgl. Kapitel 6.3.1 ab Seite 120).

Der **Dateneditor** ist das erste „bleibende“ Fenster nach dem Einstiegsfenster beim Programmstart. Er besitzt eine Datenansicht (vgl. Abbildung 41 auf Seite 108) und eine Variablenansicht (vgl. Abbildung 42 auf Seite 108).

<sup>61</sup> Weitere Programmeinstellungsmöglichkeiten finden sich ab Seite 114 im Kapitel 6.2.4.

### SPSS Dateneditor: Datenansicht

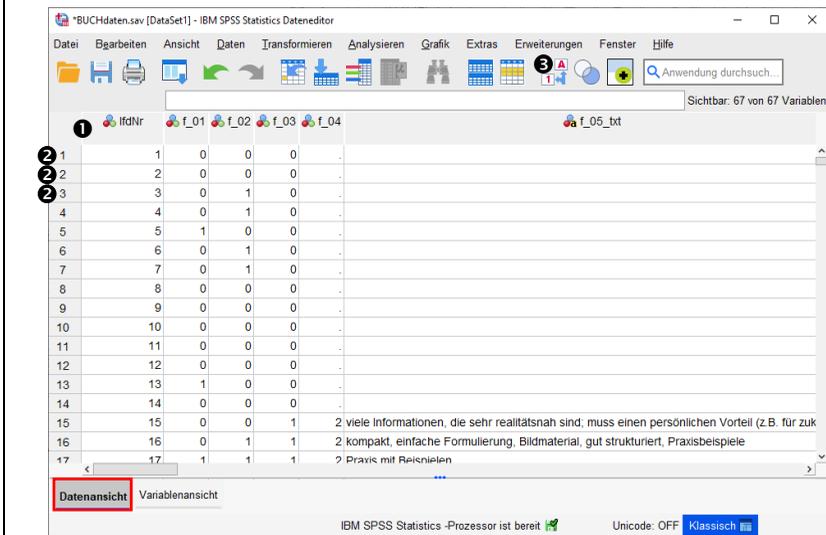


Abbildung 41: SPSS Dateneditor (Datenansicht)

### SPSS Dateneditor: Variablenansicht

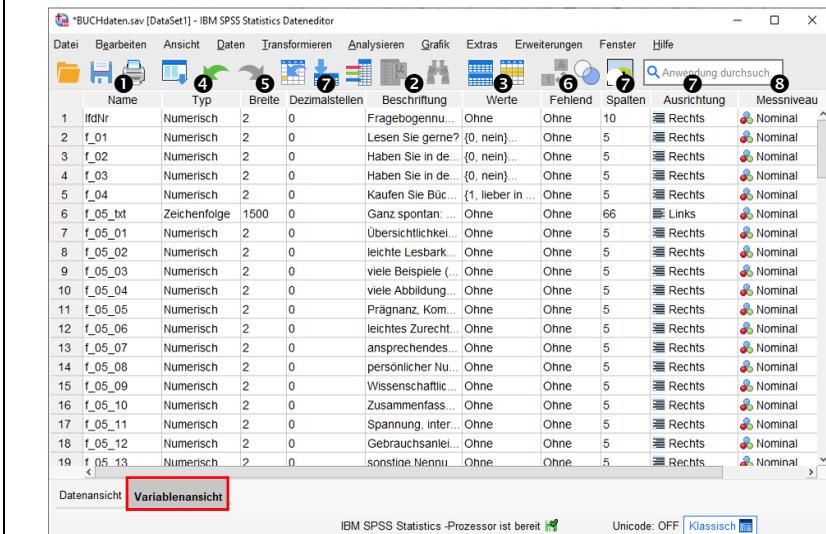


Abbildung 42: SPSS Dateneditor (Variablenansicht)

Die **Datenansicht** zeigt den Inhalt der Datendatei an – die Daten: 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/alfanumerisch. Jede Zeile darunter enthält jeweils einen Datensatz (❷).

In den Datenzellen werden wahlweise die Codierungen oder deren Beschriftungen (vgl. Kapitel 6.3.4 ab Seite 127) angezeigt. Das Umschalten zwischen Codierungs- und Beschriftungsanzeige erfolgt mit Ⓜ in Abbildung 41 auf Seite 108. In alle Zellen (ab der ersten Zeile ❷) können Daten eingegeben werden.

Die Position einer Variablen lässt sich verändern:



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

In dieser Ansicht ist der direkte Zugriff auf die Daten möglich: ändern, löschen, ausschneiden, kopieren, einfügen usw. (Numerisch codierte) leere Zellen sind **fehlende Werte** („Missings“, sie enthalten einen Punkt „.“ – vgl. Kapitel 6.3.5 ab Seite 130).

Die **Variablenansicht** des Dateneditors (vgl. Abbildung 42 auf Seite 108) zeigt und ermöglicht wichtige Einstellungen für alle Variablen des Datenfiles.

Die **Variablennamen** (❶) sind meist knappe, kurze 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.

*So ist es z.B. besser, eine Variable **Alter** oder **f\_01** zu nennen anstelle **Variable322**.*

Strukturierte Formulierungen der Variablennamen helfen dabei, die Übersicht zu behalten: ❶ – die Benennung der Variablen im Beispiel erfolgte nach den Nummern der Fragen im Fragebogen (vgl. Seite 17 und Seite 18) zu den BUCHdaten: **f\_01** bis **f\_17**.

Variablennamen 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 ...): Sie sollten bzw. dürfen nicht in Variablennamen vorkommen.

Kurze Variablennamen erleichtern zwar die Auswertung, der Ergebnisoutput wird dadurch aber schwer lesbar. Deshalb werden die **Variablennamen** in der Regel mit **Beschriftungen** (❷) versehen: Damit enthält der Output nicht länger unverständliche Kurz-Bezeichnungen. Und nicht nur die Namen der Variablen, auch ihre **Ausprägungen** sind beschriftbar (❸). Dann zeigt der Ergebnisoutput auch an, was sich hinter den Codierungen verbirgt.

*Die Beschriftung der Variable **f\_01** der BUCHdaten könnte „Lesen Sie gerne?“ lauten. Code 0 würde mit „nein“, 1 mit „ja“ bezeichnet werden.*

Weitere Details zu Variablen- und Wertebezeichnungen finden sich in Kapitel 6.3.4 ab Seite 127.

Als **Variablentyp** (❹) lässt sich eine Vielzahl von Formaten einstellen. Es ist empfehlenswert, sich auf die Typen **Numerisch** und **Zeichenfolge** zu beschränken. Andere Formate können die spätere Auswertung verkomplizieren.

**Achtung:** Enthalten als Zeichenfolge definierte Variablen (auch) Zahlen, werden auch diese von SPSS als Text interpretiert: Mit Textvariablen können aber keine Berechnungen durchgeführt werden (z.B. Mittelwerte). Wird eine Zeichenvariable deshalb als numerisch umdefiniert, ist sie zwar berechenbar, es verschwinden jedoch alle Texteinträge und nur reine Zahleneinträge bleiben bestehen!

Die **Breite** der Spalten (⑤) hat weniger Relevanz, sollte aber nicht verändert werden. Eine – v.a. bei Textvariablen – zu gering bemessene Zahl zerstört Daten: Alle längeren Einträge werden „abgeschnitten“ und unbrauchbar.

Die Spalte **Fehlend** (⑥) lässt die Eingabe von Werten zu, die bei der Auswertung wie nicht vorhandene Werte („fehlende Werte“) behandelt und nicht in die Ergebnisermittlung einbezogen werden sollen (vgl. Kapitel 6.3.5 ab Seite 130).

Angaben zu Dezimalstellen, Spaltenbreite und Ausrichtung (⑦) sind eher vernachlässigbar. Keine besondere Relevanz besitzt an dieser Stelle auch die Einstellung des Messniveaus (⑧): 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 Variablennamen** verändern.

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

VOR der ersten Datenanalyse müssen die Daten unbedingt **bereinigt** (= plausibilisiert) vorliegen (vgl. dazu Kapitel 4.4 ab Seite 47).

## 6.2.2 | Ausgabefenster

SPSS legt bei jeder Sitzung automatisch ein **Ausgabefenster (Viewer)** an. Dort werden die Ergebnisse aller Berechnungen, Programm Meldungen und Warnhinweise angezeigt.

Nach dem Ausführen jeder Auswertungsroutine können deren Ergebnisse (meist Tabellen oder Diagramme) im Ausgabefenster bearbeitet und als eigene Datei gespeichert werden. Auch ein Export in andere Dateiformate (Text, HTML, PDF, Word, Excel, Powerpoint) oder das direkte Übertragen von Elementen via **copy & paste** ist einfach umsetzbar.

In Abbildung 43 auf Seite 111 ist der Beginn der Ergebnisauflistung einer Häufigkeitszählung der gesamten BUCHdaten dargestellt.

 Eine Häufigkeitszählung des gesamten Datenfiles ist über den Menübefehl **Analisieren** → **Deskriptive Statistiken** → **Häufigkeiten...** möglich.

 Alle Variablen, die sich auf der linken Seite des Dialogfeldes befinden, werden (am schnellsten mit **strg & a**) markiert und per **drag & drop** nach rechts in den Auswahlbereich gezogen.

 Ein Klick auf **OK** startet die Auswertung.

## SPSS Ausgabefenster

Im Ausgabefenster werden die Ergebnisse aller Berechnungen angezeigt.

- linke Seite:**  
Übersicht, was alles ausgewertet wurde
- rechte Seite:**  
alle Ergebnisse im Detail

Ergebnisse ...

- ein-/ausblenden ③
- direkt anwählen, oder per drag & drop verschoben
- mit copy & paste in andere Programme übertragen

Die Ausgabe kann ...

- gespeichert (\*.SPV) oder
- exportiert werden (z.B. \*.PDF)

Gültig	0 nein	1 ja	Gesamt	Häufigkeit	Prozent
	383	122	505	75,8	24,2
			505	100,0	

Gültig	0 nein	1 ja	Gesamt	Häufigkeit	Prozent
	207	298	505	41,0	59,0
			505	100,0	

Gültig	0 nein	1 ja	Gesamt	Häufigkeit	Prozent
	372	133	505	73,7	26,3
			505	100,0	

Abbildung 43: SPSS Ausgabefenster

Die linke Seite des Ausgabefensters (① in Abbildung 43) enthält die sogenannte **Gliederung**: Sie zeigt an, was alles ausgewertet wurde. Auf der rechten (und breiteren) Seite der Ausgabe werden die **Ergebnisse** aufgelistet (②).<sup>62</sup> Diese sind per Mausklick direkt anwählbar.

Ein **Doppelklick** auf eine Ergebnisdarstellung oder einen Kommentar auf der rechten Seite des Ergebnisviewers öffnet dieses Element und ermöglicht seine Bearbeitung.

*So können z.B. aus Tabellen Diagramme erzeugt werden. Tabellen und Grafiken lassen sich verschönern, Texte einfügen, Seitenumbrüche und -ränder definieren usw.*

Alle Einzelelemente der Gliederung sind ein- und ausblendbar (**Doppelklick** auf die kleinen Buchsymbole ③) oder über das Menü **Ansicht** → **Ausblenden** (④ in Abbildung 43). Mit der **Entfernen**-Taste auf der Tastatur werden sie dauerhaft gelöscht.

SPSS erlaubt es, gleichzeitig mehrere Ausgabefenster geöffnet zu haben. Das kann aber zu Verwirrung führen („Wo ist das Ergebnis?“). Viewer-Dateien enden mit **\*.spv**.<sup>63</sup>

Nicht zuletzt beinhaltet das Ausgabefenster auch Fehlermeldungen und den Befehlstext (vgl. das folgende Kapitel 6.2.3) sämtlicher Ereignisse der aktuellen SPSS-Sitzung. Diese finden sich beim wiederkehrenden Eintrag „Log“ (⑤) in der Gliederung.

<sup>62</sup> Im Sinne einer besseren Darstellbarkeit und Übersichtlichkeit wurde das Erscheinungsbild der Ergebnistabellen in allen Abbildungen gegenüber den Standardeinstellungen adaptiert. Die Möglichkeiten, die SPSS dazu bietet, erläutert Kapitel 6.2.4 ab Seite 114 im Rahmen der Beschreibung der Anwendung von Tabellenvorlagen.

<sup>63</sup> Bis Version 15 besitzen SPSS-Outputs die Dateierweiterung **\*.spo**. Diese können mit aktuellen SPSS-Versionen nicht geöffnet werden. Zum Öffnen älterer Outputs ist ein eigenes, gratis online verfügbares Hilfsprogramm notwendig.