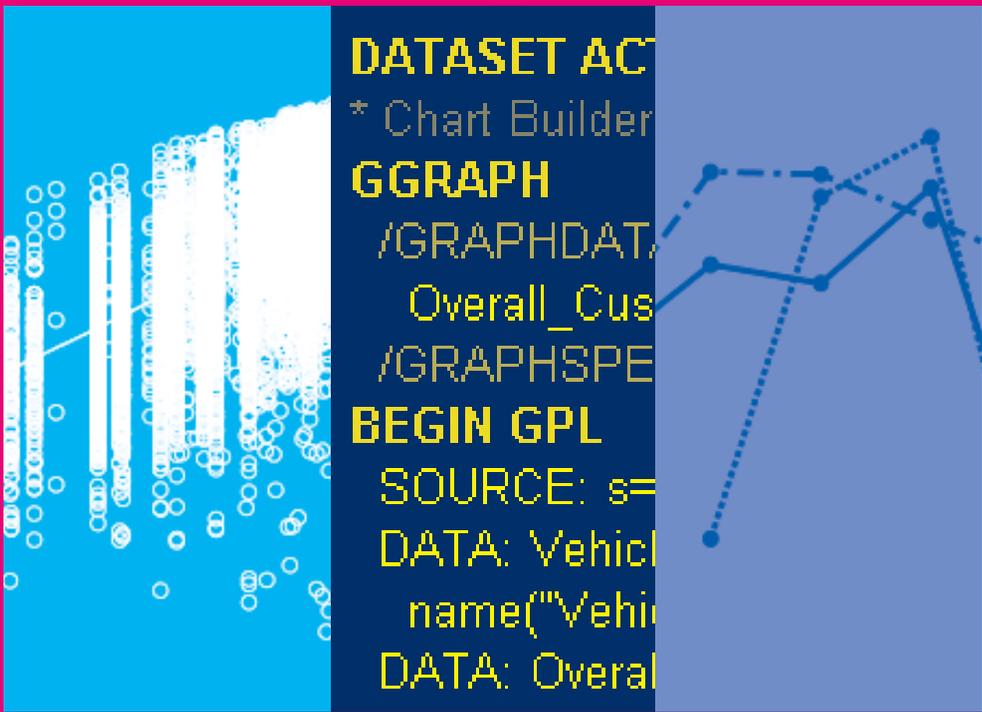


Marko Sarstedt · Tobias Schütz · Sascha Raithel



# IBM SPSS SYNTAX

Eine anwendungsorientierte Einführung

3. Auflage

Vahlen

## **Zum Inhalt:**

IBM SPSS Statistics gehört zu den populärsten Statistikprogrammen im Studium, in der Forschung und in der Praxis. Leider führen viele Anwender ihre Analysen ausschließlich mit Hilfe der grafischen Benutzeroberfläche von SPSS durch. Dabei können über die Steuersprache SPSS Syntax viele Prozeduren schneller und eleganter realisiert werden. Der souveräne Umgang mit der SPSS Syntax bietet einen unschätzbaren Vorteil für die tägliche Arbeit von Anwendern, die mit der Analyse von Daten zu tun haben.

Das Buch ist eine integrierte Einführung in die Steuersprache von IBM SPSS Statistics. Neben den notwendigen Syntax-Grundlagen behandelt es die Themengebiete Datenaufbereitung, Datentransformation und -modifikation sowie die Makro- und Matrixsprache, die in der 3. Auflage grundlegend überarbeitet wurden. Die Neuauflage wurde den Entwicklungen von SPSS angepasst, sprachlich verbessert und um weitere Anwendungsbeispiele ergänzt, die anhand realer Daten u. a. des J. D. Power and Associates Customer Satisfaction Index veranschaulicht werden. Das Buch legt besonderen Wert auf die gute Nachvollziehbarkeit der Beispiele durch begleitende Übungen. Die verwendeten Datensätze sind als kostenfreies Zusatzmaterial erhältlich. Das Buch bietet eine prägnante und umfassende Anleitung zur effizienteren Arbeit mit IBM SPSS Statistics und ist sowohl als Einstiegsliteratur für Programmieranfänger, als auch als Nachschlagewerk für fortgeschrittene Anwender geeignet.

Das Werk wurde auf Grundlage der Version 25.0 von IBM SPSS Statistics erstellt, kann aber auch für andere Versionen verwendet werden.

Es richtet sich an Studierende, Forscher (Wirtschaftswissenschaften, Wirtschaftsingenieurwesen, Soziologie, Psychologie und Statistik) sowie an Praktiker (Strategie, Consulting, Business Analytics, Data Science, etc.).

## **Zu den Autoren:**

**Prof. Dr. Marko Sarstedt** ist Professor für Marketing an der Otto-von-Guericke-Universität Magdeburg.

**Prof. Dr. Tobias Schütz** ist Professor für Marketing und Customer Science an der ESB Business School Reutlingen.

**Prof. Dr. Sascha Raithel** ist Professor für Marketing an der Freien Universität Berlin.

# IBM SPSS Syntax

Eine anwendungsorientierte Einführung

von

Prof. Dr. Marko Sarstedt, MBR

Prof. Dr. Tobias Schütz, MBR

und

Prof. Dr. Sascha Raithel, MBR

3., vollständig überarbeitete Auflage

Verlag Franz Vahlen München

*Unseren Großeltern*

## Vorwort zur dritten Auflage

“Intelligence is the ability to avoid doing work, yet getting the work done.”  
(Linus Torvalds)

Wofür ein Buch über SPSS Syntax?

Ganz einfach, weil die Verwendung der Syntax unglaublich praktisch ist! Auch wenn über die Menüsteuerung ein Großteil der Funktionalitäten von SPSS angesprochen werden kann, eröffnet die Syntaxsprache den Zugang zu neuen Verfahrensoptionen und ermöglicht zudem eine effizientere Datenanalyse. Zur praktischen Umsetzung umfangreicher Analysearbeiten ist die Beherrschung der Syntax unerlässlich. So sind Syntaxkenntnisse auch, aber nicht nur im Kreis routinierter Anwender nach wie vor weit verbreitet, zumal die Kenntnis der Steuersprache das Verständnis anderer syntax-getriebener Statistikpakete wie zum Beispiel *R*, *Matlab* oder *Stata* erheblich erleichtert.

Leider erweist sich die Erschließung dieser Materie alleine auf Basis der von IBM SPSS bereit gestellten Syntaxreferenz als hinreichend freudlos. Zwar ermöglicht die Referenz dem erfahrenen Nutzer die zielgerichtete Erweiterung seiner Syntaxkenntnisse – sie ist aber kaum für den Einstieg geeignet. Den interessierten Einsteiger zu einem erfahrenen Nutzer zu machen ist genau das Ziel dieser anwendungsorientierten Einführung. Sie setzt daher auch keinerlei Vorkenntnisse von Programmiersprachen voraus.

Wir freuen uns sehr, dass auch die zweite Auflage eine sehr gute Aufnahme erfahren hat, und wir möchten uns bei allen Lesern für ihr Vertrauen und das positive, aber auch kritische Feedback bedanken. Im Rahmen der Neuauflage haben wir eine Reihe von Anpassungen vorgenommen:

- Zunächst haben wir die Texte der Kapitel sprachlich überarbeitet und diese einfacher und verständlicher gestaltet. Dies war eine häufige und wichtige Anregung aus den Feedbacks, die wir sehr gerne aufgenommen haben.
- Alle Beispiele wurden aktualisiert und mit Blick auf deren praktische Relevanz überarbeitet. So haben wir beispielsweise die Beschreibungen der Anweisungen zur Datentransformation und (bedingten) Datenmodifikation in den Kapiteln 3 und 4 erweitert und gleichzeitig Beispiele mit eingeschränkter praktischer Relevanz gestrichen.
- Die Beschreibungen der teilweise sehr umfangreichen Beispiele in Kapitel 7 und vor allem Kapitel 8 wurden deutlich erweitert. Die einzelnen Schritte der Beispiele wurden neben einer viel detaillierteren Beschreibung im Text um zahlreiche Abbildungen und Tabellen erweitert. Dadurch kann der Leser diese Beispiele deutlich besser nachvollziehen. Gleichzeitig wurden kleine Fehler der Syntax-Programme beseitigt und diese um einige hilfreiche Elemente wie das automatische Löschen von fehlenden Werten erweitert.
- Die Neuauflage beinhaltet mehr Boxen, um wichtige Anweisungen und deren Funktionsweise hervorzuheben.

- Alle Abbildungen basieren nun auf Version 25 von IBM® SPSS® Statistics für Mac. Gleichzeitig können aber auch Anwender älterer Versionen sowie der PC-Version alle Beispiele (bis auf wenige Ausnahmen) mit ihrer Software laufen lassen.

Auch die Beispiele in der dritten Auflage basieren auf einem von J.D. Power & Associates zur Verfügung gestellten Datensatz. Weitere Informationen zum Datensatz und J.D. Power & Associates finden sich im Abschnitt „Hinweise zum verwendeten Datensatz“. Der Datensatz kann ebenso wie weitere Syntaxdateien und Ergänzungen auf der Webseite zum Buch heruntergeladen werden:

**<http://www.das-syntax-buch.de>**

Ohne die Hilfe zahlreicher Personen wäre die umfassende Überarbeitung nicht möglich gewesen. Neben Hermann Schenk vom Verlag Vahlen, der uns auch bei der dritten Auflage tatkräftig zur Seite stand, möchten wir uns bei Mitarbeitern und Kollegen bedanken, die uns an vielen Stellen unterstützt haben. Hierzu gehören insbesondere Kati Barth, Frauke Kühn, Doreen Neubert, Mandy Pick, Victor Schliwa (alle Otto-von-Guericke-Universität Magdeburg), Barbara Wagner-Horch (ESB Business School Reutlingen), Erik A. Mooi (University of Melbourne), Christian M. Ringle (Technische Universität Hamburg Harburg) und Manfred Schwaiger (Ludwig-Maximilians-Universität München).

Um auch zukünftig von den Ideen der Leser zu profitieren, sind wir für Vorschläge, wie die Inhalte ergänzt und verbessert werden können, immer dankbar!

Magdeburg, Reutlingen und Berlin im Juli 2018

*Marko Sarstedt  
Tobias Schütz  
Sascha Raithel*

# Hinweis zum verwendeten Datensatz

## **J.D. Power & Associates**

J.D. Power & Associates ist ein Marktforschungsunternehmen mit Hauptsitz in Costa Mesa, Kalifornien, USA und zählt zu den führenden Anbietern von Marketinginformationen. Die von J.D. Power zur Verfügung gestellten Daten ermöglichen es Unternehmen, ihre Produkte und Dienstleistungen exakter an die Wünsche und Bedürfnisse der Kunden anzupassen. Dadurch lassen sich Kundenbindung, Kundenloyalität und in der Folge auch Profitabilität nachhaltig steigern.

Befragen, Beobachten, Zuhören – J.D. Power deckt das gesamte Spektrum der Informationserhebung ab. So befragt J.D. Power Kunden und erstellt auf Basis der Ergebnisse eigene Benchmarkingstudien, wie auch Trackingstudien im Kundenauftrag. Außerdem beobachtet J.D. Power Veränderungen im Kundenverhalten, um neue Entwicklungen und Trends von Anfang an zu verstehen und zu nutzen. Speziell entwickelte Technologie ermöglicht die Auswertung benutzergenerierter Inhalte in sozialen Medien und Netzwerken.

J.D. Power bietet darüber hinaus Beratungsdienstleistungen an, die es Unternehmen ermöglichen, die vielen zur Verfügung stehenden Informationen nicht nur zu verstehen, sondern auch optimal gewinnbringend zu nutzen.

## **Germany Customer Satisfaction Index Study<sup>SM</sup>**

### **Ziel:**

Die J.D. Power & Associates Germany Customer Satisfaction Index Study bietet sowohl Automobilherstellern, als auch Kaufinteressenten unabhängige Informationen darüber, wie zufrieden Autobesitzer in Deutschland mit Ihren Fahrzeugen insgesamt sind. Die Studie bietet Informationen zu verschiedenen Aspekten der Kundenzufriedenheit:

- Ranglisten, welche Marken die insgesamt zufriedensten Käufer aufweisen
- Die relativen Stärken und Schwächen jedes Herstellers hinsichtlich Fahrzeugqualität, Fahrzeugattraktivität, Händlerservice und Unterhaltskosten des Fahrzeugs
- Unterschiede in der Zufriedenheit der Kunden desselben Herstellers mit dessen unterschiedlichen Modellen
- Verbesserungspotentiale der jeweiligen Hersteller in allen genannten Bereichen der Kundenzufriedenheit

### **Datenerhebung und Methodik:**

Für die Customer Satisfaction Index Study wurden Neuwagenkäufer befragt, die Ihre Fahrzeuge 2 bis 3 Jahre vor Studiendurchführung zugelassen haben. Für den vorliegenden Datensatz wurden 73.069 Fragebögen verschickt. Die Studienteilnahme wurde nicht vergütet. Insgesamt wurden 22.265 auswertbare Fragebögen erfasst, was einer Antwortquote von 30% entspricht. Auf diese Weise wurde für 28 Marken die nötige Mindestzahl von 100 Rücksendungen erreicht. Diese 28 Marken machen in Summe 95% der Neuwagenverkäufe auf dem deutschen Markt im untersuchten Zeitraum aus. Die

Richtigkeit der Angaben bezüglich Marke und Modell wurde bei allen Rücksendungen mit den vorliegenden Informationen abgeglichen. Zugleich wurde sichergestellt, dass die Fahrzeuge auch tatsächlich im relevanten Zeitraum erstzugelassen worden waren.

**Indexberechnung:**

Um einen einfachen Vergleich der Gesamtheit der zahlreichen erhobenen Aspekte der Kundenzufriedenheit zu ermöglichen, wird ein Gesamtzufriedenheitsindex gebildet. Dieser „Customer Satisfaction Index“ (CSI) ermöglicht Automobilherstellern einen verlässlichen Vergleich der Zufriedenheit Ihrer Kunden relativ zu den Kunden anderer Hersteller. Der Index besteht dabei aus den folgenden vier, ihrem Beitrag zur Gesamtzufriedenheit nach gewichteten Kategorien:

- Fahrzeugqualität/-zuverlässigkeit 30%
- Fahrzeugattraktivität 25%
- Erfahrung beim Servicebesuch 23%
- Unterhaltskosten 22%

Die aufgeführten vier Kategorien setzen sich aus insgesamt 77, als Attribute bezeichneten, Einzelbewertungen zusammen. Die Kategorien stellen den wiederum gewichteten Durchschnitt der darin jeweils enthaltenen Attribute dar. Die Attribute werden von den befragten Personen auf einer Skala von 1-10 Punkten bewertet. Zur Indexberechnung wird demnach je Fahrzeugmodell und Attribut zunächst der Durchschnitt aller Antworten berechnet. Im Anschluss wird dieser mit dem Gewicht multipliziert, das das jeweilige Attribut innerhalb seiner Kategorie hat. Das Ergebnis wird in der Folge mit dem Gewicht der jeweiligen Kategorie innerhalb des Index multipliziert. Durch Multiplikation mit 100 wird die 1-10 Punkte Skala der Attribute auf die 1.000 Punkte Skala des Index transformiert. Der Index errechnet sich durch Summierung der dadurch gewonnenen 77 Zwischenergebnisse.

# Inhaltsverzeichnis

Vorwort zur dritten Auflage .....	VII
Hinweis zum verwendeten Datensatz .....	IX
<b>1 Grundlagen der SPSS-Befehlssyntax .....</b>	<b>1</b>
1.1 Vorteile der Syntax .....	1
1.2 Erzeugen und Ausführen der Syntax .....	1
1.3 Hilfeoption und Syntax-Referenz .....	7
1.4 Aufbau und prinzipielle Regeln der Befehlssyntax .....	8
1.4.1 Unterbefehle .....	10
1.4.2 Schlüsselwörter .....	10
1.4.3 Variablennamen .....	10
1.4.4 Zahlen und Zeichenfolgen .....	11
1.4.5 Arithmetische Operatoren und spezielle Begrenzungszeichen .....	12
1.5 Konventionen zur Darstellung von Syntax-Befehlen .....	13
1.6 Grundlegende SPSS-Befehle .....	13
1.6.1 Ausführen von Anweisungen: EXECUTE .....	14
1.6.2 Einfügen von Kommentaren: COMMENT .....	14
1.7 Fehler im Syntax-Programm .....	15
1.7.1 Einige typische Fehler beim Programmieren der Syntax .....	16
1.7.2 Fehlerbehandlung .....	17
<b>2 Dateneingabe und Datenaufbereitung .....</b>	<b>21</b>
2.1 Dateneingabe im Syntax-Editor .....	21
2.2 Variablendefinition mit STRING und NUMERIC .....	24
2.3 Einlesen von externen Datendateien mit GET DATA .....	25
2.4 Datendateien .....	29
2.4.1 Erstellen einer SPSS-Datendatei: SAVE .....	29
2.4.2 Einlesen einer SPSS-Datendatei: GET .....	31
2.4.3 Simultanes Arbeiten mit Datensätzen: DATASET .....	31
2.5 Datenaufbereitung – Erweiterte Variablendefinition .....	35
2.5.1 Variablen umbenennen: RENAME VARIABLES .....	37
2.5.2 Variablenetiketten: VARIABLE LABELS .....	37
2.5.3 Wertetiketten: VALUE LABELS/ADD VALUE LABELS .....	38
2.5.4 Skalenniveau: VARIABLE LEVEL .....	40
2.5.5 Definieren von Ausgabeformaten für Werte: FORMATS .....	40
2.5.6 Fehlende Werte .....	42
2.5.6.1 Benutzerdefinierte fehlende Werte: MISSING VALUES .....	43
2.5.6.2 Systemdefinierte fehlende Werte .....	45

<b>3</b>	<b>Datentransformation und -modifikation</b>	47
3.1	Umcodieren von Variablen: RECODE-Anweisung	47
3.2	Variablen berechnen: COMPUTE-Anweisung	52
3.2.1	Arithmetische & statistische Funktionen	54
3.2.2	Logische Funktionen	57
3.2.3	Funktionen zur Erzeugung von Zufallszahlen	58
3.2.4	Funktionen für fehlende Werte	58
3.2.5	Funktionen für Datums- und Zeitvariablen	59
3.2.6	Funktionen für Zeichenfolgevariablen	64
3.2.7	Weitere Funktionen	67
3.3	Ermitteln der Häufigkeit spezifischer Werte: COUNT	68
3.4	Permanente und temporäre Datenmodifikation: TEMPORARY	69
3.5	Gewichtung der Beobachtungen: WEIGHT	71
<b>4</b>	<b>Bedingte Datenmodifikationen</b>	73
4.1	Einfache bedingte Datenmodifikationen: IF	73
4.1.1	Relationale Operatoren	75
4.1.2	Logische Operatoren	76
4.2	Komplexe bedingte Datenmodifikation: DO IF	79
4.3	Wiederholung: DO REPEAT	87
4.4	Iterative Anweisungen: LOOP	90
4.4.1	Schleifensteuerung: Anzahl der Schleifendurchläufe	91
4.4.2	Schleifensteuerung: Einhalten einer Bedingung	94
4.4.3	Schleifensteuerung: Erfüllen eines Abbruchkriteriums	97
4.5	Exkurs: Definieren von Vektoren mit VECTOR	99
4.6	Zusammenführendes Beispiel	100
<b>5</b>	<b>Datenselektion</b>	103
5.1	Datenselektion in Abhängigkeit einer Variablen: FILTER	103
5.2	Datenselektion in Abhängigkeit einer Bedingung: SELECT IF	106
5.3	Exkurs: Systemvariablen	108
5.4	Ziehung einer Zufallsstichprobe: SAMPLE	109
<b>6</b>	<b>Datendateien</b>	111
6.1	Datendatei sortieren: SORT CASES	111
6.2	Datendatei aufteilen: SPLIT FILE	112
6.3	Aggregieren von Daten: AGGREGATE	115
6.4	Zusammenführendes Beispiel	119
6.5	Zusammenfügen von Datendateien	121
6.5.1	Verschiedene Beobachtungen mit identischen Variablen: ADD FILES	122
6.5.2	Verschiedene Variablen bei gleichen Merkmalsträgern: MATCH FILES	129
6.6	Zusammenführendes Beispiel	134

<b>7 Matrix-Programme</b> .....	137
7.1 Matrizeneingabe .....	138
7.1.1 Explizite Matrizeneingabe: COMPUTE .....	138
7.1.2 Einlesen einer von SPSS erzeugten Matrix: MGET .....	139
7.1.3 Definition von Matrizen: GET .....	141
7.2 Matrizenausgabe: PRINT .....	143
7.3 Arithmetische Operatoren .....	144
7.4 Matrix-Funktionen .....	147
7.5 Zusammenführende Beispiele .....	149
<b>8 Makros</b> .....	157
8.1 Der Aufbau eines Makros .....	158
8.2 Flexibilität durch Argumente .....	160
8.2.1 Schlüsselwortargumente .....	161
8.2.2 Positionale Argumente .....	163
8.3 Die Spezifikation von Argumenten .....	164
8.3.1 !TOKENS .....	165
8.3.2 !CHAREND .....	166
8.3.3 !CMDEND .....	167
8.4 Hilfreiche Befehle für die Verwendung in Makros .....	168
8.4.1 Befehle für die Manipulation von Textvariablen und Text in Makro-Programmen .....	168
8.4.2 Befehl für die bedingte Datenmodifikation: !IF .....	169
8.4.3 Befehle für wiederholte Anweisungen: !DO .....	171
8.5 Einfache Makroideen .....	173
8.5.1 Bündeln von Variablen .....	173
8.5.2 Makros kombinieren .....	174
8.5.3 Variablenbündel im Makro-Aufruf .....	175
8.5.4 Dateipfade als Makro .....	175
8.6 Komplexe Makro-Programme .....	176
8.6.1 Regression mit einem rollierenden Zeitfenster .....	176
8.6.2 Identifikation von Heteroskedastizität: Der Breusch-Pagan-Test .....	185
8.6.3 Lineare Regressionen mit heteroskedastizitätskonsistenten Standardfehlern .....	197
8.6.4 Bootstrapping mit dem Ausgabeverwaltungssystem OMS .....	208
<b>Literaturverzeichnis</b> .....	221
<b>Stichwortverzeichnis</b> .....	225

# 1 Grundlagen der SPSS-Befehlssyntax

## 1.1 Vorteile der Syntax

Im Vergleich zur grafischen SPSS-Benutzeroberfläche wirkt die Eingabe von Befehlen nach syntaktischen Regeln ziemlich umständlich – zugegeben. Gibt es die praktischen Icons denn nicht gerade damit wir keine Steuersprache erlernen müssen? Erneut: Ja, das ist die Idee! Einerseits...andererseits hat die Syntax in vielen Bereichen aber auch große Vorteile. Zum Beispiel können in der Syntax Funktionalitäten genutzt werden, die über die Menüführung einfach nicht verfügbar sind. Anders als Dialogfeldeinstellungen können Syntaxbefehle zudem dauerhaft gespeichert werden. Dadurch entsteht eine lückenlose Dokumentation des eigenen Vorgehens, was insbesondere bei Datentransformationen wichtig ist. Außerdem können die definierten Auswertungsschritte problemlos zu einem späteren Zeitpunkt erneut ausgeführt werden. Das erspart Arbeit, da Modifikationen einzelner Einstellungen schnell und einfach möglich sind. Die Formatierung der Syntax-Befehle im *.txt*- oder *.docx*-Format, und die Möglichkeit, Kommentare einzufügen (COMMENT-Anweisung, vgl. Kapitel 1.6.2) erlaubt den bequemen Austausch über vorgenommene Auswertungsschritte mit Projektpartner und Kollegen – direkt in SPSS und völlig ohne zusätzlichen Mailverkehr. Durch das Zusammenstellen mehrerer Befehle zu einem zusammenhängenden Programm lassen sich auch komplexe Aufgaben automatisieren.

Weitere Vorteile sind:

- Übersichtliche und schnellere Auswertung von Datensätzen mit besonders vielen Variablen.
- Steigerung des Bewusstseins für die geplanten Auswertungen, da die einzelnen Schritte dezidiert formuliert werden müssen.
- Interoperabilität, d.h. Syntax-Befehle können unabhängig von Betriebssystemen genutzt werden.

Es lohnt sich auf lange Sicht also durchaus die SPSS-Syntax zu erlernen. Doch auch wer nicht gleich ganz tief einsteigen möchte, kann viele Vorteile der Syntax nutzen: Eine Syntax kann nämlich auch einfach aus dem Ausgabefenster einer „erklickten“ Analyse kopiert und in ein Syntax-Fenster eingefügt oder (wer es noch einfacher möchte) durch sogenannte Point-and-click Operationen erzeugt werden.

## 1.2 Erzeugen und Ausführen der Syntax

Mit dem **Syntax-Editor** werden einzelne SPSS-Befehle oder vollständige SPSS-Programme im **Syntax-Fenster** erstellt und ausgeführt. SPSS-Befehle können hierbei entweder manuell oder mit Hilfe von Point-and-click-Operationen in den Syntax-Editor eingegeben werden, der beliebig viele Syntax-Befehle verarbeiten kann.

Das Öffnen eines Syntax-Eingabefensters erfolgt über die Menüfolge

► *Datei* ► *Neu* ► *Syntax*

**Point-and-click-Operationen** sind menügesteuerte Auswertungen, die in den Syntax-Editor eingesetzt werden können. Dabei ist es nicht einmal notwendig, vorher ein neues Syntax-Fenster zu öffnen. SPSS erledigt dies von selbst. Wir müssen nach einer definierten Analyse lediglich statt dem gewohnten *OK* das Steuerfeld *Einfügen* klicken. Die so erzeugte Syntax kann nun im Editor gespeichert oder modifiziert werden. Abbildung 1.1 zeigt beispielhaft das Menüfenster einer Korrelationsanalyse und die entsprechende Syntax im Syntax-Fenster.

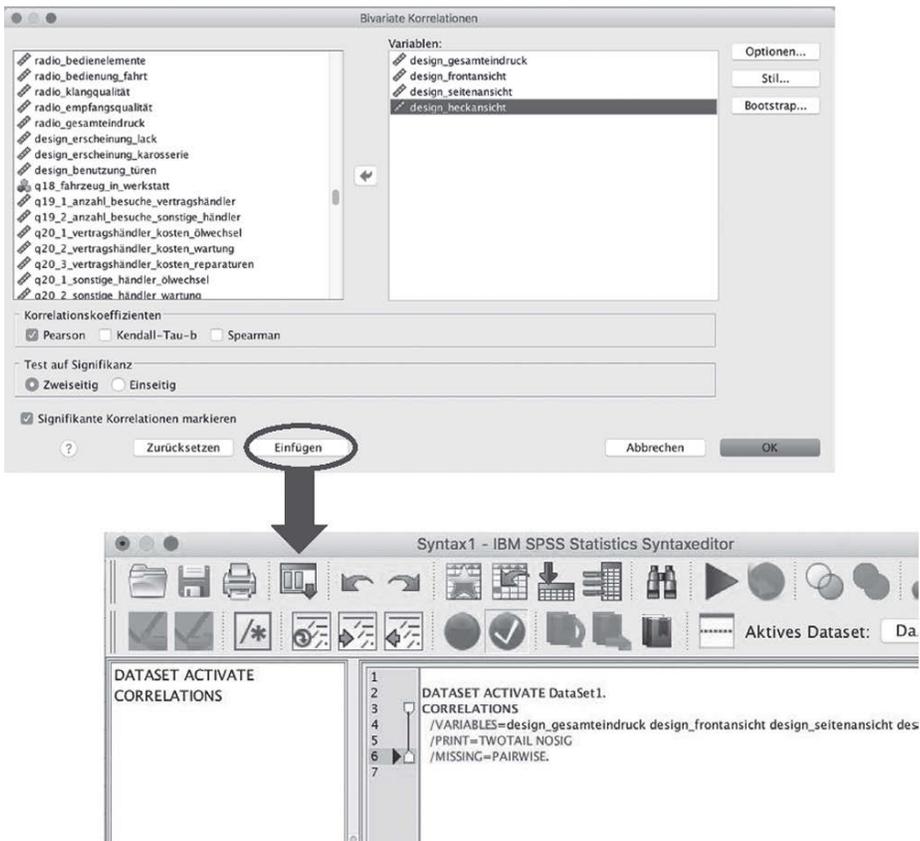


Abbildung 1.1: Point-and-Click Operation

Damit die im Syntax-Fenster formulierten Befehle dann ausgeführt werden, müssen sie zur Verarbeitung an SPSS übergeben werden. Hierfür gibt es diverse Optionen (Abbildung 1.2).

Zum **Ausführen eines einzelnen Befehls** positionieren wir den Cursor an einer beliebigen Stelle innerhalb des Befehls und wählen die Menüfolge

► *Ausführen* ► *Auswahl*

beziehungsweise betätigen die Schaltfläche mit dem grünen Pfeil .

Zum **Ausführen mehrerer Befehle**, die innerhalb der Syntax-Datei unmittelbar hintereinanderstehen, markieren wir den auszuführenden Bereich und wählen den bereits bekannten Menübefehl

► *Ausführen* ► *Auswahl*

beziehungsweise betätigen die Schaltfläche .

Sollen **alle Befehle bis zum Ende der Syntax-Datei** ausgeführt werden, so positionieren wir den Cursor an einer beliebigen Stelle innerhalb der Datei. Der Menübefehl

► *Ausführen* ► *Bis Ende*

führt den markierten sowie alle folgenden Befehle der Syntax-Datei der Reihe nach aus.

Um **alle Befehle** einer Syntax-Datei auszuführen, wählen wir den Menübefehl

► *Ausführen* ► *Alle*



**Abbildung 1.2:** Ausführen der Syntax

Auf Grundlage des Beispieldatensatzes *JDPA Germany CSI.sav* wollen wir untersuchen, ob ein Zusammenhang zwischen den Zufriedenheiten mit den einzelnen Aspekten des Fahrzeugdesigns besteht. Die Variablen *design\_frontansicht*, *design\_seitenansicht* und *design\_heckansicht* geben die Zufriedenheit des Probanden mit dem jeweiligen Aspekt des Außendesigns seines Fahrzeugs wider. Die Variable *design\_gesamteindruck* gibt Aufschluss über die Gesamtzufriedenheit mit dem Fahrzeugdesign.



Eine Trainingsversion des Datensatzes ist auf der Webseite zum Buch (<http://www.das-syntax-buch.de>) als Download verfügbar.

In der J.D. Power & Associates Germany Customer Satisfaction Index (CSI) Study<sup>SM</sup> werden Zufriedenheiten stets auf einer Skala von 1 (unakzeptabel) bis 10 (herausragend)

abgefragt. Unter der Annahme äquidistanter Abstände kann die Skala als quasi-metrisch interpretiert werden und wir können den linearen Zusammenhang mit dem Pearson'schen Korrelationskoeffizienten messen (siehe z.B. Sarstedt und Mooi 2019). Dieser wird in der Syntax durch die Prozedur **CORRELATIONS** (alternativ **PEARSON CORR**) angesteuert. Die folgende Menüsequenz führt zum gewünschten Test:

- ▶ *Analysieren* ▶ *Korrelation* ▶ *Bivariat...*
- ▶ *design\_gesamteindruck...* *design\_heckansicht* ▶ *Variablen*
- ▶ *Korrelationskoeffizienten* ▶  *Pearson*
- ▶ *Einfügen*

Durch die Betätigung des Steuerfeldes *Einfügen* wird folgende Programm-Syntax im (bereits geöffneten oder sich sodann öffnenden) Syntax-Fenster angezeigt:

```
CORRELATIONS
  /VARIABLES = design_gesamteindruck design_frontansicht
              design_seitenansicht design_heckansicht
  /PRINT = TWOTAIL NOSIG
  /MISSING = PAIRWISE.
```

Der resultierende Syntax-Befehl (auch Anweisung genannt; beide Bezeichnungen werden im Rahmen dieses Buches synonym verwendet) besteht aus dem Befehlsnamen (hier: **CORRELATIONS**), verschiedenen Unterbefehlen (hier: **VARIABLES**, **PRINT**, **MISSING**) sowie deren Spezifikation in Form von Variablenamen (hier: z.B. *design\_gesamteindruck*) und Schlüsselwörtern (hier: **TWOTAIL**, **NOSIG**, **PAIRWISE**). Ausführliche Erklärungen zum Aufbau der Syntax und Syntax-Regeln folgen in Kapitel 1.4.

Zum Ausführen des Syntax-Befehls wählen wir innerhalb des Syntax-Fensters die Menüsequenz

- ▶ *Ausführen* ▶ *Alles*.

Durch diese Syntax werden die bivariaten Korrelationskoeffizienten nach Pearson berechnet und in einer symmetrischen Korrelationsmatrix ausgegeben (Tabelle 1.1).

Korrelationen					
		design_ gesamteindruck	design_ frontansicht	design_ seitenansicht	design_ heckansicht
design_gesamteindruck	Korrelation nach Pearson	1	.796**	.821**	.778**
	Signifikanz (2-seitig)		.000	.000	.000
	N	17486	17387	17365	17367
design_frontansicht	Korrelation nach Pearson	.796**	1	.815**	.696**
	Signifikanz (2-seitig)	.000		.000	.000
	N	17387	17443	17408	17406
design_seitenansicht	Korrelation nach Pearson	.821**	.815**	1	.763**
	Signifikanz (2-seitig)	.000	.000		.000
	N	17365	17408	17421	17396
design_heckansicht	Korrelation nach Pearson	.778**	.696**	.763**	1
	Signifikanz (2-seitig)	.000	.000	.000	
	N	17367	17406	17396	17422
**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.					

Tabelle 1.1: Korrelationsmatrix

Nehmen wir an, dass wir aber gar nicht an allen Korrelationen interessiert sind und wir uns primär mit der Variable *design\_gesamteindruck* beschäftigen wollen. *design\_gesamteindruck* umfasst die Zufriedenheit des Probanden mit allen Aspekten des Fahrzeugdesigns und sollte somit mit den Teilaspekten des Außendesigns korrelieren. Wenn wir diese Vorahnung (um den Begriff *Hypothese* zu vermeiden) überprüfen wollen, könnten wir nun umständlich 3 Mal in Folge die Korrelation zwischen *design\_gesamteindruck* und jeder der anderen Variablen über die SPSS-Menüführung berechnen lassen. Eine leichte Modifikation der Syntax durch Einfügen des Schlüsselwortes **WITH** löst das Problem deutlich eleganter:

```
CORRELATIONS
  /VARIABLES = design_gesamteindruck WITH design_frontansicht
              design_seitenansicht design_heckansicht
  /PRINT = TWOTAIL NOSIG
  /MISSING = PAIRWISE.
```

Das Ausführen dieser Syntax liefert eine übersichtliche, reduzierte Korrelationsmatrix in Tabelle 1.2.

Korrelationen				
		design_ frontansicht	design_ seitenansicht	design_ heckansicht
design_ gesamteindruck	Korrelation nach Pearson	.796**	.821**	.778**
	Signifikanz (2-seitig)	.000	.000	.000
	N	17387	17365	17367
**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.				

**Tabelle 1.2:** Reduzierte Korrelationsmatrix

Die neu generierte Syntax kann einfach durch die Menüsequenz

► *Datei* ► *Speichern*

abgespeichert und zu einem späteren Zeitpunkt wieder geöffnet und verwendet werden. Die dafür notwendige Menüsequenz im Syntax-Fenster ist:

► *Datei* ► *Öffnen*

Anfänger werden ein Syntax-Programm in der Regel nicht von Grund auf in der Syntax-Sprache schreiben, sondern eher den oben beschriebenen Weg gehen und bestehende Syntax-Programme modifizieren. Neben der Verwendung des Steuerfeldes *Einfügen* können wir eine Syntax auch aus dem SPSS-Ausgabefenster per copy-paste in ein Syntax-Fenster übertragen. Die Syntax befindet sich in einem Textfeld, das einfach über einen Mausklick auf den Gliederungspunkt *Log* in der linken Navigationsleiste des Ausgabefensters angesteuert werden kann (Abbildung 1.3).

The screenshot shows the SPSS Output window with the following content:

```

DATASET ACTIVATE DataSet1.
CORRELATIONS
/VARIABLES=design_gesamteindruck WITH design_frontansicht design_seitenansicht
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

**Korrelationen**

		Design: Frontansicht	Design: Seitenansicht	Design: Heckansicht
Design: Gesamteindruck	Korrelation nach Pearson	,796**	,821**	,778**
	Signifikanz (2-seitig)	,000	,000	,000
	N	17387	17365	17367

\*\* . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 1.3: Syntax im Ausgabefenster

In SPSS ist die Ausgabe der Syntax im Outputfenster standardmäßig voreingestellt. Sollte die Syntax einmal nicht mit ausgegeben werden, so liegt dies daran, dass die Darstellung vorher durch den Syntax-Befehl **SET PRINTBACK** unterbunden wurde:

```
SET PRINTBACK OFF.
```

In diesem Fall können wir die Syntax-Ausgabe durch den folgenden Befehl einfach wieder einschalten.

```
SET PRINTBACK ON.
```

### 1.3 Hilfeoption und Syntax-Referenz

Die Fülle an Syntax-Befehlen, die im Laufe der Programmversionen entwickelt wurden, ist selbst für routinierte Anwender kaum überschaubar. Eine offizielle Syntax-Referenz wurde daher direkt in SPSS integriert, so dass einzelne Befehle schnell und komfortabel nachgeschlagen werden können. Die Syntax-Referenz beinhaltet zudem eine umfassende Erläuterung aller Befehle. Allerdings ist die Syntax-Referenz eben ein Befehlskatalog und verfolgt kein didaktisches Konzept. Zum Erlernen der SPSS-Syntax ist sie darum wenig geeignet. Erfahrenen Anwendern bietet die Befehlsreferenz jedoch eine gute Unterstützung. Wir erreichen die **Befehlssyntaxreferenz** (*Command Syntax Reference*) über das Hilfe-Drop-Down Menü des Syntax-Fensters (Abbildung 1.4).

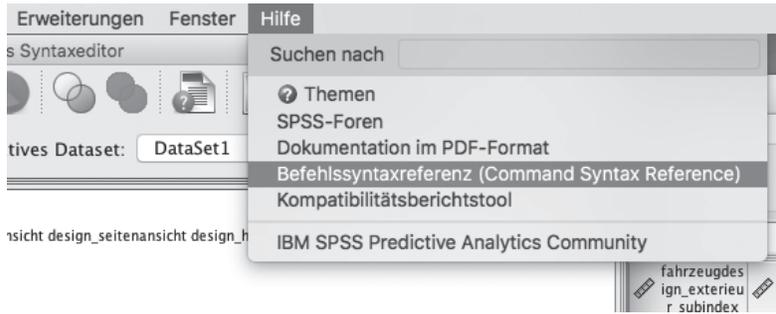


Abbildung 1.4: Befehlssyntaxreferenz

Möchten wir einen einzelnen Befehl einer bestehenden Syntax nachschlagen, so geht dies noch einfacher. Dazu muss innerhalb des Syntax-Fensters lediglich der Cursor vor den Befehl gestellt und das Icon *Hilfe zur Syntax* (stilisiertes Syntaxfenster mit einem Fragezeichen in der linken unteren Ecke) betätigt werden (Abbildung 1.5)



Abbildung 1.5: „Hilfe zur Syntax“ Icon

Es öffnet sich ein Hilfefenster, welches die Rohsyntax des aktuell gewählten Befehls sowie eine kurze Beispielsyntax abbildet.

## 1.4 Aufbau und prinzipielle Regeln der Befehlssyntax

Ein **SPSS-Befehl** beginnt immer mit einem **Befehlsnamen**, der aus einem oder mehreren Wörtern bestehen kann und die durchzuführende Aktion spezifiziert. Einem Befehlsnamen folgen in der Regel weitere **Spezifikationen**, die den zu Grunde liegenden Befehl konkretisieren. Spezifikationen können sich aus den folgenden Elementen zusammensetzen:

- Unterbefehle
- Schlüsselwörter
- Variablennamen
- Wertelabels
- Zahlen und Zeichenfolgen
- Arithmetische Operatoren und Begrenzungszeichen

Beim Schreiben der Syntax sind folgende **Syntax-Editierungsregeln** zu beachten:

- Jeder Befehl muss in einer neuen Zeile beginnen und mit einem Punkt abgeschlossen werden.
- Fortsetzungszeilen sollten aus Gründen der Übersichtlichkeit eingerückt werden.
- Ein in Apostrophe gesetzter Text (z.B. bei Variablenetiketten, vgl. Kapitel 2.5.2) muss sich in einer Zeile befinden.
- An jeder Stelle, an der ein einzelnes Leerzeichen erlaubt ist, können zusätzliche Leerzeichen eingefügt...
- ... oder die Anweisung auf eine neue Zeile umgebrochen werden ohne die Syntax inhaltlich zu verändern.

Der Syntax-Editor identifiziert und kategorisiert Spezifikationselemente (Befehlsnamen, Unterbefehle, Wertelabels, Kommentare und Schlüsselwörter) bereits während der Eingabe und weist den Kategorien Textfarben zu. Dies erfolgt in dem Moment, in dem ein Spezifikationselement als vollständig erkannt wird. Unterbefehle werden in der Syntax grün dargestellt, Schlüsselwörter erscheinen dunkelrot und Kommentare hellgrau. Befehlsnamen werden zunächst hellrot und fett gedruckt und wechseln die Farbe zu blau, sobald der den Befehl abschließende Punkt gesetzt wurde. Zahlen und Zeichenfolgen, sowie arithmetische Operatoren und Begrenzungszeichen bleiben schwarz. Das ist sehr praktisch, da diese Funktion die Orientierung in der Syntax extrem erleichtert. Außerdem werden wir im Prozess des Editierens einer Syntax vor Tipp- und Logikfehlern bewahrt. Nimmt ein Befehlsname, Unterbefehl oder Schlüsselwort nach der Eingabe des letzten Buchstabens nicht die entsprechende Farbe an, so liegt entweder ein Tippfehler vor, oder das Spezifikationselement steht logisch nicht im richtigen Kontext (z.B. Definition eines Unterbefehls ohne vorherigen Befehlsnamen).

Eine weitere Funktion des Editors macht uns das Bearbeiten einer Syntax sogar noch leichter: Wann immer wir mit dem Schreiben eines Spezifikationselements beginnen, schlägt SPSS in einem Menü mehrere Spezifikationsoptionen vor (Abbildung 1.6). Dabei steht bei Befehlsnamen eine alphabetisch geordnete Gesamtliste der Syntax-Befehle zur Verfügung. Bei Unterbefehlen oder Schlüsselwörtern gibt SPSS eine Kurzliste der im Kontext sinnvollen Elemente vor. Voraussetzung ist allerdings, dass der Anfangsbuchstabe mindestens eines geeigneten Unterbefehls oder Schlüsselwortes eingegeben wurde.

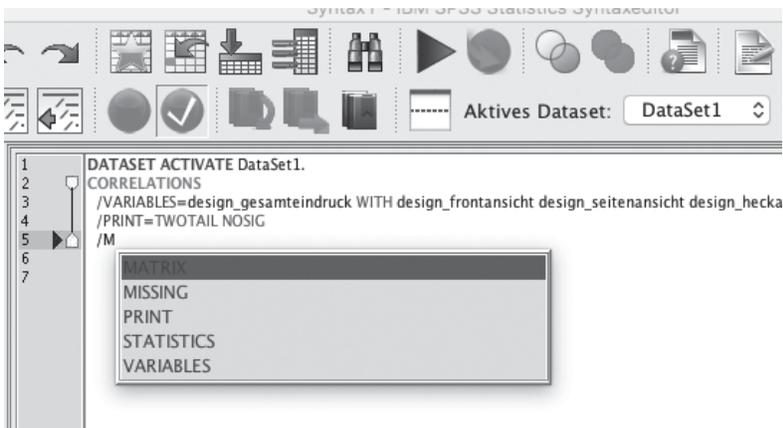


Abbildung 1.6: SPSS schlägt Spezifikationsoptionen vor

SPSS unterscheidet grundsätzlich nicht zwischen Groß- und Kleinschreibung, jedoch wird die Syntax aus optischen Gründen in Großbuchstaben angegeben. Die Namen von Variablen sind die Ausnahme. Diese werden stets in der Form dargestellt, wie sie im Datensatz erfasst sind. Im Folgenden wollen wir die verschiedenen Spezifikations-elemente genauer betrachten.

### 1.4.1 Unterbefehle

Für Syntax-Befehle existieren in der Regel ein oder mehrere Zusatzanweisungen, so genannte Unterbefehle, die weitere Spezifikationen enthalten können. Für Unterbefehle gilt:

- Unterbefehle beginnen mit einem fest vorgegebenen Befehlsnamen.
- Mehrere Unterbefehle innerhalb eines Befehls werden durch je einen Schrägstrich (/) voneinander getrennt.
- Unterbefehlen, die weitere Spezifikationen enthalten, folgt ein Gleichheitszeichen.

```
CORRELATIONS
  /VARIABLES = design_gesamteindruck WITH design_frontansicht
              design_seitenansicht design_heckansicht
  /PRINT = TWOTAIL NOSIG
  /MISSING = PAIRWISE.
```

Die Elemente VARIABLES, PRINT und MISSING sind Unterbefehle.

### 1.4.2 Schlüsselwörter

Schlüsselwörter spezifizieren Befehle oder Unterbefehle und werden für Funktionen, Operatoren und andere Angaben benötigt. Ihre Schreibweise ist fest vorgegeben. Die in der folgenden Syntax fett gedruckten Elemente WITH, TWOTAIL, NOSIG und PAIRWISE sind Schlüsselwörter.

```
CORRELATIONS
  /VARIABLES = design_gesamteindruck WITH design_frontansicht
              design_seitenansicht design_heckansicht
  /PRINT = TWOTAIL NOSIG
  /MISSING = PAIRWISE.
```

### 1.4.3 Variablennamen

Variablennamen müssen in SPSS folgenden Regeln entsprechen:

- Variablennamen müssen mit einem Buchstaben (A bis Z oder Umlaute Ä, Ö, Ü) oder dem Sonderzeichen @ beginnen.
- Variablen, die mit dem Sonderzeichen # beginnen, bezeichnen **Hilfs- oder Arbeitsvariablen** (in SPSS als „scratch variables“ bezeichnet), die nur temporär verfügbar sind und nicht in die Datendatei übernommen werden (vgl. Kapitel 4.3).

- Variablen, die mit einem Dollarzeichen (\$) beginnen, kennzeichnen so genannte **Systemvariablen** (in SPSS als „system variables“ bezeichnet), die von SPSS automatisch gebildet werden und nicht durch den Benutzer verändert werden können (vgl. Kapitel 5.3).
- Die Sonderzeichen # und \$ können vom Nutzer nicht als erste Zeichen vergeben werden, wenn die Variable in der Datendatei gespeichert werden soll. Sie können aber trotzdem als erste Zeichen eines Variablennamens vorkommen.
- Folgezeichen können Buchstaben, Ziffern (0 bis 9), der Punkt (.), das Unterstreichungszeichen (\_) und die drei Sonderzeichen @, # und \$ sein.
- Variablennamen dürfen keine Leerzeichen enthalten und nicht mit einem Punkt oder Unterstreichungszeichen enden.
- Die maximale Länge eines Variablennamens beträgt 64 Zeichen.



Einige Schlüsselwörter dürfen nicht als Variablennamen verwendet werden. Hierzu gehören

ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO und WITH

Folgende Variablennamen wären somit gültig:

Alter	m	v.1	exp\$82	@fixed	c_block123
-------	---	-----	---------	--------	------------

Folgende Variablennamen wären ungültig:

Eq	geb-ort	einkommen/netto	berufliche ausbildung
----	---------	-----------------	-----------------------

#### 1.4.4 Zahlen und Zeichenfolgen

SPSS verarbeitet sowohl Ganzzahlen, als auch Dezimalzahlen, bei denen das in Kontinentaleuropa übliche Dezimalkomma durch einen Dezimalpunkt ersetzt wird. Ist kein Vorzeichen angegeben, so ist die Zahl positiv.

**Zeichenfolgen** (auch **Strings** genannt) sind Folgen von Buchstaben, Ziffern und Sonderzeichen, die von SPSS als Text interpretiert werden. Auch eine als Zeichenfolge definierte Abfolge von Zahlen wird somit als Text und nicht als Zahl eingestuft. Eine Zeichenfolge ist in Apostrophe ('Zeichenfolge') oder in Anführungszeichen ("Zeichenfolge") zu setzen. Sie darf maximal 32.767 Zeichen lang sein und kann in der Syntax ohne Umbruch durchgeschrieben werden.

### 1.4.5 Arithmetische Operatoren und spezielle Begrenzungszeichen

SPSS verwendet die in Tabelle 1.3 dokumentierten fünf **arithmetischen Operatoren**.

Operator	Bedeutung
**	Potenzierung
*	Multiplikation
+	Addition
-	Subtraktion
/	Division

**Tabelle 1.3:** *Arithmetische Operatoren*

Das Gleichheitszeichen dient im Zusammenhang mit arithmetischen Operationen der Wertzuweisung. Als spezielle **Begrenzungszeichen** verwendet SPSS die folgenden Zeichen:

- ( ) Mit runden Klammern werden Argumente von Funktionen, Wertebereiche in bestimmten Prozeduren und Schlüsselwörter eingeschlossen.
- ' und " Hochkommata und Anführungszeichen dienen der Abgrenzung von Zeichenfolgen.
- / Der Schrägstrich dient dazu, Unterbefehle oder mehrere Analyselisten in Prozeduren voneinander abzugrenzen.
- = Das Gleichheitszeichen dient dazu, Äquivalenz anzuzeigen oder trennt ein Schlüsselwort eines Unterbefehls von einem zugehörigen Wert.

Vor und nach Begrenzungszeichen sind keine Leerzeichen notwendig, es ist jedoch erlaubt, beliebig viele Leerzeichen einzufügen. Abbildung 1.7 zeigt abschließend den Aufbau eines einfachen Syntax-Befehls.

CORRELATIONS	→ Befehlsname
/VARIABLES=design_gesamteindruck WITH	→ Unterbefehl
design_frontansicht design_seitenansicht	→ Variablen
/PRINT=TWOTAIL SIG	→ Schlüsselwörter
/MISSING=PAIRWISE.	

**Abbildung 1.7:** *Aufbau eines Syntax-Befehls*

Dieses sehr kleine Beispiel lässt erahnen, dass die vollständige Dokumentation eines Syntax-Befehls in Anbetracht der Fülle von Analysemöglichkeiten schnell unübersichtlich werden kann. Im Laufe der Jahre hat sich daher eine Konvention zur Darstellung von Syntax-Befehlen herausgearbeitet, die wir im folgenden Abschnitt betrachten wollen.

## 1.5 Konventionen zur Darstellung von Syntax-Befehlen

Die folgende, detaillierte Darstellung von Syntax-Befehlen entspricht der im aktuellen Referenzwerk für SPSS-Syntax (Command Syntax Reference) verwendeten Illustration:

- Elemente in Großbuchstaben kennzeichnen Stichwörter, die von SPSS verwendet werden, um Befehle, Unterbefehle, Funktionen, Operatoren und andere Spezifikationen zu identifizieren.
- Elemente in Kleinbuchstaben kennzeichnen Variablen, die vom Anwender zu spezifizieren sind. In der unten angeführten Syntax zeigt beispielsweise der Terminus `varlist` an, dass der Anwender eine oder mehrere Variablen aufführen muss.
- Der Ausdruck `varname` steht für eine Variable, `varlist` für eine Variablenliste.
- Elemente in eckigen Klammern (`[]`) sind optional.
- Geschweifte Klammern (`{ }`) zeigen eine Wahlmöglichkeit zwischen mehreren Angaben an, die für jeden Unterbefehl untereinander aufgeführt werden.
- Fettgedruckte Elemente kennzeichnen Voreinstellungen, wobei SPSS zwischen zwei Arten von Voreinstellungen differenziert: Während eine mit zwei Sternen (`**`) gekennzeichnete Option bei Aufruf des entsprechenden (Ober-) Befehls in Kraft tritt, wird diese bei fehlenden Sternen erst mit Angabe des entsprechenden Unterbefehls aktiviert. Im nachfolgenden Beispiel werden durch den Aufruf des `CORRELATIONS`-Befehls die Unterbefehle `MISSING = PAIRWISE` und `PRINT = TWOTAIL` per Voreinstellung aktiviert.
- Notizen erscheinen unterhalb der eigentlichen Syntax-Angabe.

Abbildung 1.8 verdeutlicht beispielhaft die Konventionen zur Darstellung von Befehlen.

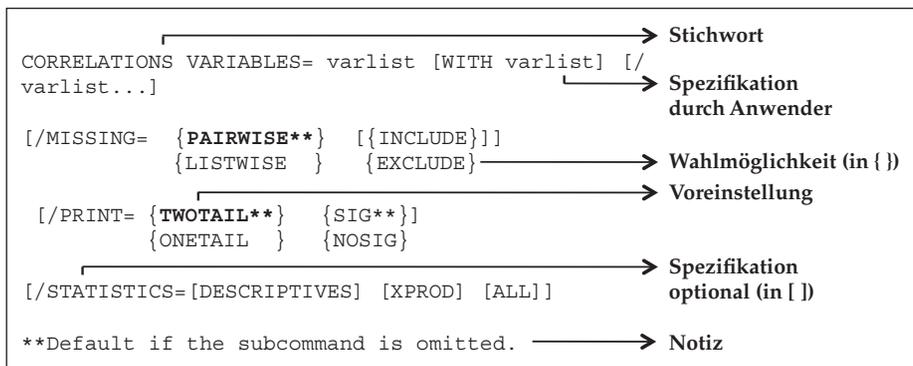


Abbildung 1.8: Konventionen zur Darstellung von Befehlen

## 1.6 Grundlegende SPSS-Befehle

Die Syntax-Befehle `EXECUTE` und `COMMENT` haben allgemeinen Charakter und werden daher einleitend behandelt.