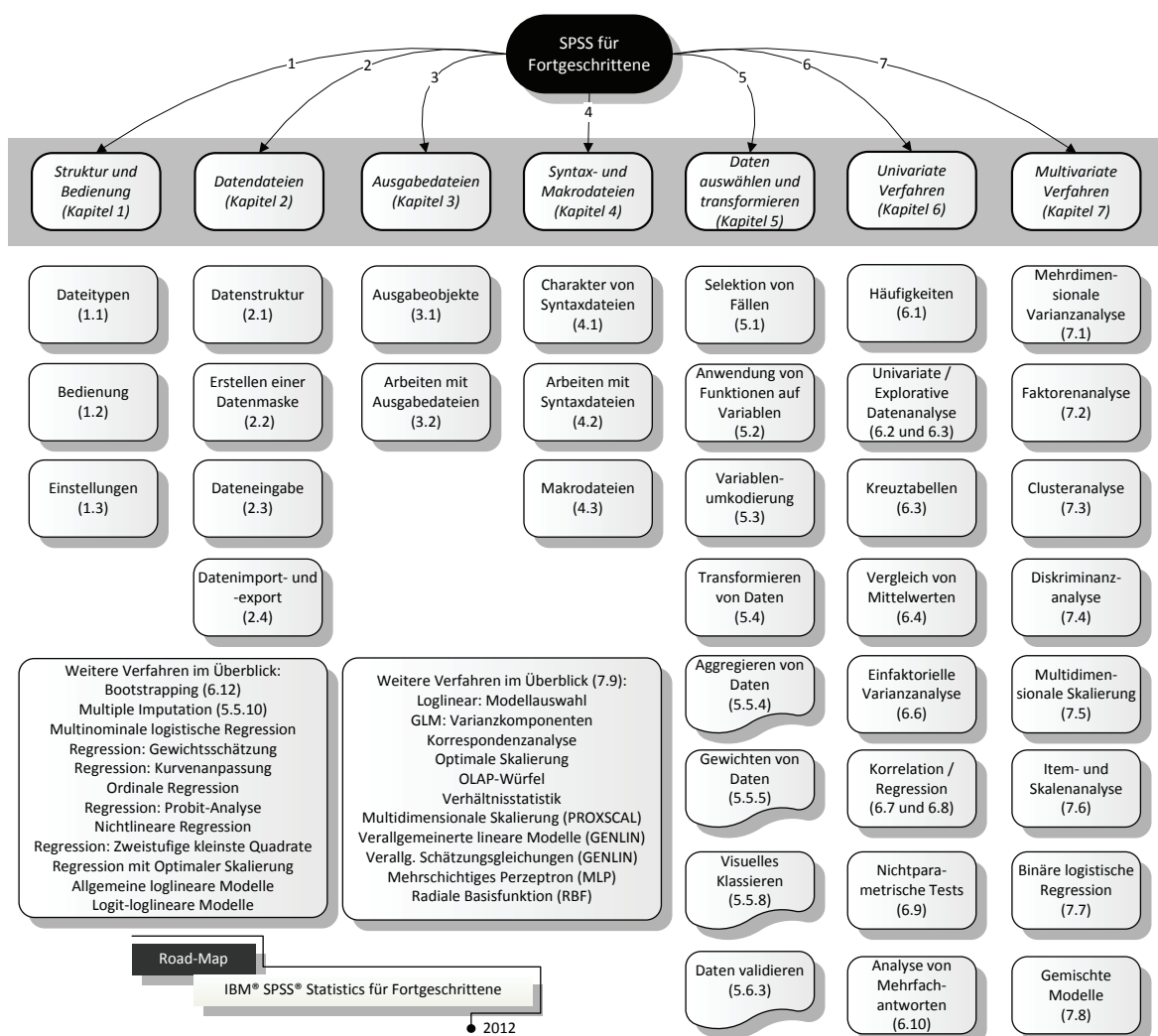


Prolog

IBM® SPSS® STATISTICS (im Folgenden kurz SPSS genannt) ist ein System zur statistischen Analyse und grafischen Darstellung empirisch quantitativer Daten. Dieses Handbuch richtet sich an den fortgeschrittenen Anwender von SPSS (Student/in, Doktorand/in und Dozent/in), der eine kompakte und übersichtliche Einführung in die wichtigsten statistischen Prozeduren im Bereich *deskriptive, univariate* und als besonderer Schwerpunkt *multivariate* Statistik sucht. Dabei wird eine gründliche Auseinandersetzung mit den Grundlagen und Anwendungen der Statistik vorausgesetzt.



Um den Einstieg zu erleichtern, werden in den Kapiteln 1 bis 4 die Grundlagen im Umgang mit Daten-, Ausgabe-, Syntax- und Makrodateien zusammenfassend dargestellt.

Kapitel 5 beschäftigt sich mit den Möglichkeiten der Veränderung und Selektion in bestehenden Datensätzen und zeigt verschiedene Varianten der Umstrukturierung, Umkodierung, Sortierung, Verrechnung, Aggregation von Variablen auf. Fortgeschrittene Arbeitstechniken, wie z.B. der Import und Export von Daten- oder Ausgabedateien, die Aufbereitung und Validierung von Daten sowie die multiple Imputation fehlender Werte werden hierbei vertieft.

Im 6. Kapitel folgt eine kompakte Übersicht der univariaten statistischen Verfahren unter besonderer Berücksichtigung ihrer Funktion im Kontext multivariater Statistik; hierzu werden im Folgenden diskutiert:

1. Häufigkeiten und Histogramme
 - absolut, prozentual, kumuliert.
2. Deskriptive Lage- und Dispersionsmaße
 - Lagemaße: Mittelwert, Median, Modalwert, Maximum, Minimum, Perzentile, Quartile, Spannweite;
 - Streuungsmaße: Varianz, Standardabweichung, Standardfehler;
 - Verteilungsmaße: Schiefe, Exzess;
 - Erzeugung standardisierter Werte (z-Werte).
3. Verfahren der explorativen Datenanalyse
 - Grafische Darstellungsmöglichkeiten: Histogramm, Stem-and-Leaf-Plot, Box-and-Whiskers-Plot;
 - Spezielle Lagemaße: 5% getrimmtes Mittel, M -Schätzer;
 - Prüfung auf Normalverteilung: Normalverteilungsdiagramm, Normal-Q-Q-Plot, Trendbereinigter Q-Q-Plot, Kolmogorov-Smirnov-Test, Lilliefors-Test, Shapiro-Wilks-Test.
4. Kreuztabellen und deren statistische Prüfung
 - Häufigkeitsverteilungen in Kreuztabellen und Chi-Quadrat-Statistik;
 - Zusammenhangsmaße auf Nominalskalenniveau: Kontingenzkoeffizient C , Phi-Koeffizient, Cramers V , Lambda, Goodman & Kruskal-Tau;
 - Zusammenhangsmaße auf Ordinalskalenniveau: Kendalls tau- b und $-c$, Gamma, Somers- D ;
 - Zusammenhangsmaße auf Intervallskalenniveau: Pearsonscher Korrelationskoeffizient, Eta, Kappa-Koeffizient, Relatives Risiko.
5. Univariate inferenzstatistische Verfahren
 - Vergleich von Mittelwerten: t -Test bei einer Stichprobe, t -/Welch-Test bei zwei abhängigen und unabhängigen Stichproben (einschließlich Levene-Test auf Varianzhomogenität), einfaktorielle Varianzanalyse, Post-Hoc-Tests, Kontrastverfahren;
 - Korrelationen: Bivariate Korrelation nach Pearson, Kendall und Spearman, Partielle Korrelation;
 - Lineare Regression: Multipler Korrelationskoeffizient R , Bestimmtheitsmaß R -Quadrat, Korrigiertes R -Quadrat, Standardfehler des Schätzers, Parameter der Regressionsgleichung und deren Signifikanzprüfung, Kollinearitätsdiagnose über den *Variance Inflation Factor* und Konditionsindex, Methoden der Prädiktorberücksichtigung;
 - Vergleich von Rangreihen: Chi-Quadrat-Test, Binomialtest, Sequenzanalyse, Ein-Stichproben-Kolmogorov-Smirnov-Test, Mann-Whitney- U -Test, Wilcoxon-Test, Rand-Homogenitätstest, Kolmogorov-Smirnov- Z -Test, Wald-Wolfowitz-Sequenzen-Test, Moses-Test, Kruskal-Wallis-Test, Jonckheere-Terpstra-Test, Vorzeichen-Test, Friedman-Test, Kendalls W , Cochrans Q .
6. Analyse von Mehrfachantworten
7. Analyse von fehlenden Werten
8. Varianzanalytische Verfahren
 - Mehrfaktorielle univariate Varianzanalyse einschließlich der Berechnung von Zwischensubjekteffekten, Post-Hoc-Tests, Kontrasten und der Analyse geschätzter Mittel von Randverteilungen;
 - Multivariate Varianzanalyse nach dem Allgemeinen Linearen Modell (*GLM*) einschließlich der Berechnung von multivariaten Tests, Zwischensubjekteffekten und der Analyse geschätzter Randmittel;
 - Varianzanalyse mit Messwiederholungen nach dem Allgemeinen Linearen Modell (*GLM*) einschließlich der Berechnung von Innersubjekteffekten, multivariaten Tests, der Sphärizität, von Tests der Innersubjekteffekte, Innersubjektkontraste, Zwischensubjekteffekte und der Analyse geschätzter Randmittel.

9. Dimensionsreduzierende Verfahren

- Faktorenanalyse/Hauptkomponentenanalyse einschließlich der Berechnung von Kommunalitäten, Eigenwerten, Faktorladungsmatrizen, verschiedenen Methoden der Faktorextraktion und verschiedenen Formen der Faktorrotation;
- (Hierarchische, *k*-Means, Two-Step-) Clusteranalyse einschließlich der Darstellung der verschiedenen Distanz- und Ähnlichkeitsmaße für verschiedene Skalenniveaus, der verschiedenen Agglomerierungsverfahren und der Visualisierung in Form von Eiszapfendiagrammen und Dendrogrammen;
- (Einschließende und schrittweise) Diskriminanzanalyse einschließlich der Interpretation der fallweisen Einzelwahrscheinlichkeiten, der Eigenwerte, von Wilks-Lambda und den Funktionswerten der Diskriminanzfunktion(en) bzw. der kanonischen Diskriminanzfunktion(en) einschließlich ihrer grafischen Darstellung;
- Multidimensionale Skalierung für symmetrische und asymmetrische Datenmatrizen und Rohdaten einschließlich der zugrunde liegenden Daten- und Skalierungsmodelle und Konfigurations- und Streudiagramme;
- Item- und Skalenanalyse einschließlich Cronbachs Alpha, verschiedener Split-Half-Modelle und Paralleltest-Modelle.

10. Prädiktive Verfahren

- Binär logistische Regression einschließlich verschiedener Methoden der Berücksichtigung von Prädiktoren und deren grafische Veranschaulichung;
- Gemischte Modelle für Daten mit korrelierter und nicht konstanter Variabilität. Dabei werden nicht nur Mittelwerte, sondern auch Varianzen und Kovarianzen analysiert. Die Prozedur dient ferner der Modellanpassung im Kontext von Mehrebenenmodellen, hierarchisch linearen Modellen und Zufallskoeffizientenmodellen.

11. Weitere Verfahren im Überblick

Den Anhang bilden u.a. drei *Concept Maps*, die als Entscheidungs- und Strukturierungshilfe bei der Auswahl adäquater statistischer Verfahren herangezogen werden können. Abschließend folgt für den interessierten Leser die Möglichkeit, seinen Kenntnisstand anhand einiger Übungsaufgaben zu überprüfen. Den Abschluss liefert eine Literaturübersicht zur Vertiefung.

Die für die Berechnungen erforderlichen Datensätze (Beispiele und Übungen) können aus dem Internet unter der folgenden Adresse bezogen werden:

www.uni-saarland.de/fak5/psy/spssdat.zip

Die angegebene Datei *spssdat.zip* beinhaltet die verschiedenen Datensätze im ZIP-Format; vor der Verwendung ist die Datei zu dekomprimieren (z.B. mit dem Programm *winzip.exe*).

Ich danke auf diesem Wege besonders den Rezensenten – Monika Faßler (Universität Jena), Dr. Hans Grüner (Freie Universität Berlin), Dr. Manfred Heisterkamp (Universität Freiburg) und Dr. Ulrich Rohland (Universität Leipzig) – deren Korrekturen und Anmerkungen in die 4. und 7. Auflage eingeflossen sind. Ferner danke ich Herrn Andreas Zins (Student der Psychologie, Universität des Saarlandes) für seine Mithilfe bei der Aktualisierung der *Concept-Maps* in der 8. Auflage. Und schließlich gilt mein Dank den weiteren Rezensenten – Dr. Carina Ortseifen (Universität Heidelberg) und Dr. Steffen Krause (Universität der Bundeswehr, München) – die sich der 9. Auflage mit zahlreichen Kommentaren angenommen haben.

Schließlich gilt mein besonderer Dank Herrn Dipl.-Math. Wilhelm Noack vom Regionalen Rechenzentrum für Niedersachsen (RRZN) der Universität Hannover, für die anhaltend exzellente Betreuung in den letzten zehn Jahren, die mit dieser Auflage (ruhestandsbedingt) nun leider endet.

Saarbrücken, April 2012

PD Dr. Markus Pospeschill

*Email des Autors:
pospeschill@mx.uni-saarland.de*