

AXEL BAUMANN

Einstieg in die Daten- verarbeitung mit SPSS 23

Kurzanleitung für die Analyse
von Marketingforschungsdaten
mit SPSS 23



AVMpress

Einstieg in die Datenverarbeitung mit SPSS 23

AXEL BAUMANN

Einstieg in die Daten- verarbeitung mit SPSS 23

Kurzanleitung für die Analyse von
Marketingforschungsdaten mit SPSS 23



AVMpress

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

AVM - Akademische Verlagsgemeinschaft München 2016
© Thomas Martin Verlagsgesellschaft, München

Umschlagabbildung: © Robert Kneschke - Fotolia.com

Alle Rechte vorbehalten. Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urhebergesetzes ohne schriftliche Zustimmung des Verlages ist unzulässig und strafbar. Das gilt insbesondere für Nachdruck, auch auszugsweise, Reproduktion, Vervielfältigung, Übersetzung, Mikroverfilmung sowie Digitalisierung oder Einspeicherung und Verarbeitung auf Tonträgern und in elektronischen Systemen aller Art.

Alle Informationen in diesem Buch wurden mit größter Sorgfalt erarbeitet und geprüft. Weder Autoren noch Verlag können jedoch für Schäden haftbar gemacht werden, die in Zusammenhang mit der Verwendung dieses Buches stehen.

e-ISBN (ePDF) 978-3-96091-292-7
ISBN (Print) 978-3-86924-978-0

Verlagsverzeichnis schickt gern:
AVM - Akademische Verlagsgemeinschaft München
Schwanthalerstr. 81
D-80336 München

www.avm-verlag.de

Inhaltsverzeichnis

Vorworte	7
Vorbemerkung und Danksagung	11
1. Vorhaben und Umsetzung	13
2. Messung und Typen von Daten	17
2.1 Messung von Daten	17
2.2 Typen von Daten	21
3. Erfassung und Analyse von Daten	29
3.1 Erfassung von Daten	29
3.2 Anmerkung zur Datenanalyse	32
4. Variablenansicht und Datenansicht in SPSS	35
4.1 Datenverarbeitung starten	35
4.2 Variablen definieren	39
4.3 Daten eingeben	57
5. Variablenansicht und Datenansicht in SPSS erweitern	63
5.1 Daten einlesen	63
5.2 Variablen einlesen	72
6. Datensatz in SPSS Statistics bearbeiten	77
6.1 Daten klassieren	77
6.2 Fälle auswählen	81
7. Variablenset bearbeiten und Eingaben prüfen	85
7.1 Variablen transformieren	85
7.2 Variablen umcodieren	93
7.3 Eingaben prüfen	96
8. Viewer und Syntaxeditor in SPSS	99
8.1 Umgang mit der SPSS-Ausgabe	99

8.2 Umgang mit der SPSS-Syntax	103
9. Datenbasis beschreiben mit SPSS Statistics	107
9.1 Häufigkeiten mit SPSS Statistics ermitteln	107
9.2 Diagramme mit SPSS Statistics erstellen	121
10. Kreuztabellen und Mehrfachantworten	135
10.1 Kreuztabellen mit SPSS Statistics erstellen	135
10.2 Mehrfachantworten in SPSS Statistics verarbeiten	139
11. Einblick in ausgewählte Analysekonzepte in SPSS	147
11.1 Statistische Prüfverfahren	147
11.2 Korrelation und einfache lineare Regression	162
12. Abschlussbemerkungen und Quick References	171
12.1 Hinweis auf multivariate Datenanalyse mit SPSS	171
12.2 Resümee und Schlussbetrachtung	174
12.3 Quick Reference IBM SPSS Statistics Prozeduren	178
12.4 Checkliste Datenbearbeitung	182
Anhang 1: Auszug der Variablenansicht zum Nachstellen	183
Anhang 2: Auszug der Datenansicht zum Nachstellen	184
Anhang 3: Fragebogenbeispiel als Einarbeitungsübung	185
Anhang 4: Arbeiten mit SPSS Statistics Übungsvorschlag	187
Anhang 5: Beispiele statistischer Prüfverfahren mit SPSS	189
Abbildungen	193
Literatur	197
Internetquellen	209
Index	211
Hinweis	217

Vorworte

Weder die Privatwirtschaft noch die öffentlichen Verwaltungen kommen heute bei zukunftsorientierten Entscheidungen ohne relevantes und gegenwartsnahes Wissen aus. Die Datenfülle, die den Entscheidungen zugrunde liegt, steigt stetig an und wird zudem immer komplexer. Angesichts dieser Herausforderungen wird ein leistungsfähiges und am Bedarf der Entscheidungsträger ausgerichtetes Berichtswesen zur Grundlage einer effektiven und effizienten Maßnahmensteuerung. „Mit der Hand am Arm“ wird es immer seltener möglich sein, fundierte Entscheidungen zu treffen. Wegen der ansteigenden Datenflut benötigt man deshalb eine leistungsfähige Software, da strukturierte Entscheidungen sonst kaum mehr möglich sind. SPSS ist in diesem Kontext die führende Datenanalysesoftware und bietet dafür in vielen Bereichen praxiserprobte Paketlösungen. Das Komprimieren der Datenpakete mit SPSS führt zum Beispiel zu entscheidungsrelevanten Kennzahlen, Tabellen und Grafiken sowie statistischen Auswertungen. Dankbar dürfen wir sein, dass es Menschen wie Axel Baumann gibt, die sich u. a. auf die Nutzung von SPSS im Rahmen von Analyse- und Forschungsarbeiten spezialisiert haben und ihre Erfahrungen gerne - wie auch in diesem Buch - weitergeben. Verständlich und anschaulich beschreibt Axel Baumann die praktische Anwendung von SPSS anhand eines Studienbeispiels. Der Leser kann mit Hilfe dieses Buches seine konkreten Datenanalysearbeiten mittels SPSS lösen. Mein Amtssitz ist in Bondorf - eine Gemeinde mit rund 5.900 Einwohnern zwischen Herrenberg, Rottenburg und Nagold im Landkreis Böblingen gelegen. Diese Gemeinde ist von Handwerk, Mittelstand, Dienstleistern und Landwirtschaft geprägt und zeichnet sich besonders durch eine vielfältige und aktive Bürgerschaft aus. Es freut mich, dass Axel Baumann in seinem Anschauungsbeispiel Bezug zur Landwirtschaft genommen hat, denn diese genießt in unserer Gemeinde noch immer einen hohen Stellenwert. Außerdem ist Axel Baumann selbst einer der aktiven Bürger, die sich im Gemeinwesen engagieren. In unserem Computertreff bringt er sich ehrenamtlich ein und unterstützt dabei vor allem ältere Menschen bei der Nutzung und praktischen Anwendung von Computern. Mit seinem Buchprojekt und den nützlichen Tipps liefert er einen weiteren Beleg für das Bürgerengagement in der Gemeinde Bondorf.

Bernd Dürr, Bürgermeister

Für unsere Studierenden ist der Umgang mit SPSS ein wichtiger Bestandteil der Ausbildung, denn es ist auch in der Praxis weit verbreitet. Die Studierenden setzen SPSS intensiv in empirischen Projekt-, Seminar- und Ab-

schlussarbeiten für die statistische Datenanalyse ein. So können die Studierenden auch im Unternehmen die unterschiedlichsten Geschäfts- und Forschungsprobleme mit SPSS lösen. Wir sind daher sehr froh, dass Axel Baumann die SPSS-Schulung an der Hochschule übernehmen konnte und zusätzlich dieses SPSS-Lehrbuch erstellt hat. In seinem Buch erklärt Herr Baumann anschaulich und verständlich die Nutzung des Programms anhand eines Studienbeispiels und bettet seine Ausführungen in einen forscherschen und statistischen Kontext ein. Die beschriebenen SPSS-Anwendungen können vom Leser mit dem Buch nachvollzogen und praktisch geübt werden - so dass die Bearbeitung empirischer Daten mit SPSS rasch gelingt. Herr Baumann beschreibt alle nötigen Arbeitsschritte in SPSS: die Erstellung von Datenmasken und Eingabe von Daten, die Datenauswertung und Datenanalyse sowie die Erstellung von Graphiken und Berichten. Die Beschreibungen werden mit zahlreichen Grafiken illustriert. Leser, die zum ersten Mal Marktforschungsdaten mit SPSS verarbeiten möchten, erhalten praktische Hilfen von der Vorbereitung über die Umsetzung der Datenerfassung bis hin zur Analyse und Darstellung der Ergebnisse eines Datenverarbeitungsprojekts. SPSS-Einsteiger nützt es, wenn ihnen der Stoff in einer überschaubaren und leicht erfassbaren Form dargeboten wird und ihnen zentrale Elemente des Fachgebietes mit Praxis- und Forschungsbezug vorgestellt werden. Darauf zielt dieses Buch ab. Dem Autor sei Dank, dass er mit diesem Buch den Einstieg in dieses komplexe Fachgebiet ermöglicht. **Professor Dr. Thomas Berger, Duale Hochschule Baden Württemberg, Stuttgart**

Die Geschichte der maschinell-gestützten Datenanalyse ist untrennbar mit dem Computerprogramm SPSS verbunden. 1968 an der Universität Stanford als Großrechneranwendung zur Unterstützung von Datenauswertungen in den Sozialwissenschaften (Statistical Package for the Social Sciences) entstanden, ist das Programm heute unverzichtbar für die Datenanalyse in den Sozial- und Humanwissenschaften und der betrieblichen Praxis in Forschung, Controlling, Marketing etc.. In der vorliegenden Einführung vermittelt Axel Baumann interessierten Lernenden und Praktikern die Grundlagen der Datenverarbeitung mit SPSS. Abgedeckt werden neben den Basics der statistischen Datenanalyse vor allem auch die oft unterschätzten und vernachlässigten Gebiete des Datenmanagements und der notwendigen Datenaufbereitung. Der Leser gewinnt einen fundierten Einblick in den schrittweisen Prozess der Transformation von Daten in Wissen und wird in die Lage versetzt, durch den Einsatz von SPSS eigenständig Aufgaben der Datenanalyse in Lehre, Forschung und betrieblichem Alltag durchzuführen.

Neben der Vermittlung der notwendigen Kenntnisse in den Bereichen des Datenmanagements und der Datenanalyse mit SPSS ist das Buch dazu geeignet, den Leser in konkreten Projekten als Handbuch zu begleiten. Axel Baumann ist ein Mann der Theorie und Praxis und kann auf eine langjährige Tätigkeit in Forschung, betrieblicher Praxis und Lehre zurückschauen. Sein Hauptinteresse liegt auf den Gebieten des Marketings und der Marketingforschung und der Schulung von Softwarelösungen aus dem Bereich der Business Intelligence. **Klaus-Michael Böhme, Sozialwissenschaftler**

Bereits im Rahmen meines Studiums hatte ich durch die Nutzung für wissenschaftliche Erhebungen umfangreiche Berührungspunkte mit der Software IBM SPSS Statistics. In diesem Zuge durfte ich Axel Baumann als kompetenten und zuverlässigen Ansprechpartner in allen Fragen zu IBM SPSS Statistics kennenlernen. Als Dozent der Dualen Hochschule Baden-Württemberg Stuttgart unterrichtet er die Anwendung von IBM SPSS Statistics unter anderem in seinen Vorlesungsmodulen zur Marketingforschung und Datenanalyse. Die Software bietet insbesondere Studierenden die Möglichkeit ihre wissenschaftlichen Arbeiten mit individuell erstellten Auswertungen anzureichern und Ergebnisse grafisch und anschaulich darzustellen. Anhand des Studienbeispiels eines Forschers leitet Axel Baumann in seinem Buch anschaulich durch die einzelnen Themengebiete „Messung und Erfassung von Daten“, „Bearbeitung von Datensätzen“ und „Datenbeschreibung und -analyse“. Dieses Buch stellt damit sämtliche Anwendungsbereiche der Datenverarbeitung mit IBM SPSS Statistics dar und ermöglicht sowohl Neuanfängern und insbesondere Studierenden als auch Praktikern die Software sukzessive mit all ihren Möglichkeiten kennenzulernen und Spezialfälle zu bearbeiten. Dabei verknüpft Axel Baumann praxisorientierte Sachverhalte mit statistischem und wissenschaftlichem Hintergrundwissen und bietet allen Nutzern und Interessierten eine große Unterstützung in der Datenverarbeitung und -aufbereitung mit IBM SPSS Statistics. **Verena Groß, Bachelor of Arts**

Vorbemerkung und Danksagung

Die computergestützte Verarbeitung und Analyse von Daten zum Beispiel mittels der Software IBM SPSS Statistics werden **in Kursen der Markt- und Marketingforschung** für Fortgeschrittene¹ vermehrt zur Pflicht. Die grafische Benutzeroberfläche des Computerprogrammes IBM SPSS Statistics unterstützt die **intuitive Umsetzung** komplexer **statistischer Analyseverfahren**. Exzellente Kenntnisse über die Voraussetzungen für die einzelnen statistischen Verfahren sind nicht zwingend erforderlich. Eine intensive statistische Grundausbildung, die das Verständnis für den Einsatz statistischer Methoden schärft, wird jedoch empfohlen.² Zu leicht kann es zum Einsatz nicht adäquater Analyseverfahren und zu verfehlten Analyseergebnissen kommen.

SPSS bedeutet Superior Performing Software System (früher Statistical Package for the Social Sciences).³ SPSS ist eine Marke der International Business Machines Corporation (IBM) und wurde im Jahr 2009 von IBM übernommen.⁴ Das Programm wird in dieser Publikation aus Vereinfachungsgründen häufig nur SPSS oder SPSS Statistics genannt. Die Version 18 wurde mit der Bezeichnung PASW (Predictive Analysis Software) veröffentlicht.⁵ Im Jahr 2015 wurde die Version IBM SPSS Statistics 23 auf den Markt gebracht.

Die vorliegende Ausarbeitung beinhaltet Ausführungen über ausgewählte, relevante statistische Grundlagen und Darstellungen von SPSS Statistics 23 Programmkomponenten aus Anwendersicht. Der Ablauf der behandelten Themen entspricht dem Arbeitsprozess der Datenverarbeitung in der Praxis. Die Publikation ist als **Einstiegswerk** gedacht und kein SPSS-Handbuch mit allen Programmfeatures. Grundlagen und Eckpfeiler der Datenanalyse mit IBM SPSS Statistics werden den Lesern zugänglich, anwendungsorientiert und bewusst fokussiert vermittelt. Der Transfer in die Praxis soll dadurch unterstützt werden. Der Autor erhebt jedoch keinen Anspruch auf Vollständigkeit aller Methoden, wissenschaftlichen Ausführungen und Möglichkeiten der Datenanalyse mit IBM SPSS Statistics. Dieser

¹ Vgl. hierzu zum Beispiel in den Studienplänen der Dualen Hochschule Baden-Württemberg, Fakultät Wirtschaft.

² Vgl. Baumann, A. (2015), S. 13 ff..

³ Vgl. Hüttner, M., u. a. (2002), S. 215.

⁴ © Copyright IBM Corp. 2013. Alle Rechte vorbehalten.

⁵ Mehr zur Geschichte von SPSS und der Begrifflichkeit SPSS vgl. zum Beispiel unter Martens, J. (2003), S. 1 ff..

Aspekt ist im Bedarfsfall durch vertiefende Literatur zu ergänzen.⁶ **Zielgruppen** dieser Publikation sind **Auszubildende, Studierende, Forschende und Interessierte**, die sich mit der Datenanalyse mit IBM SPSS Statistics zum ersten Mal beschäftigen und hierfür eine Anleitung sowie ein knappes Nachschlagewerk benötigen. Diese Veröffentlichung ist ein Buch für SPSS-Anwender/innen und wendet sich nicht an Mathematiker/innen. Auf ausgewählte theoretische Grundlagen wird in reduzierter Form und so weit als unbedingt nötig eingegangen. Die Inhalte dieser Veröffentlichung stammen aus den Workshops des Autors und sind praxiserprobt. Diese Publikation ist konzipiert als Textbuch zur eigenständigen Erschließung der Arbeiten mit dem System SPSS sowie als Textbuch für entsprechende anwendungsbezogene Lehrveranstaltungen. Die Gliederung dieser Publikation orientiert sich daher an Hauptaufgaben der Praxis. Durchaus von Vorteil für Leserinnen und Leser ist es zwar, wenn bereits erste Kompetenzen in Markt- und Marketingforschung sowie Statistik vorhanden sind, umfassendes mathematisches Vorwissen ist jedoch nicht erforderlich.

Die Darstellungen der IBM SPSS Statistics Programmfunktionalitäten beziehen sich auf die Software-Version von IBM SPSS Statistics 23, die von den Entwicklern im Jahre 2015 auf den Markt gebracht wurde. Es geht allerdings in der vorliegenden Publikation um grundsätzliche Funktionsweisen, Nutzungskonzepte und Prinzipien von IBM SPSS und nicht um sämtliche Programmfeatures der Programmversionen. Der Großteil der Anleitungen zur Nutzung der neueren Version von IBM SPSS lässt sich auch auf frühere SPSS Softwareversionen übertragen. Das Betriebssystem des Computers mit der installierten SPSS Software - hier im Sprachmodus Deutsch - ist in dieser Publikation eine neuere Version von Microsoft Windows.

Einen besonderen Dank muss Frau Gabriele Baumann ausgesprochen werden, die mit viel Geduld und Ausdauer die Texte dieser Publikation gelesen hat. Ihre kritische Lektoren-Arbeit hat zu wertvollen Hinweisen und Vorschlägen geführt. Den Lesern danke ich für das Interesse an diesem Buch.

⁶ Vgl. zum Beispiel Backhaus, K., u. a. (2011); Bühl, A. (2012) und vgl. die Software-Handbücher von IBM SPSS, die bei Lieferung der Software von IBM mitgeliefert werden oder auf der Homepage von IBM SPSS heruntergeladen werden können.

1. Vorhaben und Umsetzung

Die vorliegende **Publikation** beinhaltet **ausgewählte Grundlagen** der Datenerhebung, Datenerfassung und -auswertung sowie **ausgewählte Programmfunktionen** des Programms **IBM SPSS Statistics**. Mit dieser Statistiksoftware wird die computergestützte Verarbeitung und Analyse von Daten in dieser Veröffentlichung durchgeführt. Die Grundlagen schaffen die Basis für das nötige Verständnis für die jeweiligen statistischen und computergestützten Arbeiten. Die Datenverarbeitung hat die Aufgabe die Informationen, die in den Daten stecken, zu erschließen. Mit Hilfe von IBM SPSS Statistics können Daten so komprimiert werden, dass sich relevante Einsichten und Erkenntnisse herauskristallisieren. Statistischen Methoden werden genutzt, um Daten in Informationen umzuformen und die Interpretationsbasis und Entscheidungsgrundlage zu fundieren.

Mathematische Formeln und formale Prozeduren stehen nicht im Vordergrund dieser Veröffentlichung sondern die Beschreibung von Datenverarbeitungs- und Datenanalysekonzepten in Kombination mit IBM SPSS Statistics. Statt der formalen Prozeduren sollen Verständnis und Gespür für Konzepte der Datenverarbeitung und Datenanalyse im Sinne eines Workshops erarbeitet werden. Anhand von praxisorientierten Aufgaben wird beschrieben, wie die Datenbearbeitung in der Praxis aussehen kann.

Zum Beispiel wird dargestellt:

- wie Daten computergerecht erfasst und für die Weiterverarbeitung aufbereitet werden,
- welche statistischen Methoden für welche Fragestellung eingesetzt werden,
- wie mit IBM SPSS Statistics die Datenverarbeitung und -analyse realisiert wird.

Die Leserinnen und Leser sollen nach der Lektüre dieser Veröffentlichung ausgewählte statistische Methoden anwenden, mit Hilfe von IBM SPSS Statistics umsetzen und die Inhalte der Veröffentlichung auf ihre eigene Praxis übertragen können.

Mit Hilfe der graphischen Benutzeroberfläche von IBM SPSS lassen sich Tabellen, Diagramme und eine Vielzahl von statistischen Auswertungen mit wenigen Klicks realisieren. Das Benutzerinterface ist für Statistik-Einsteiger einfach und intuitiv zu bedienen und erleichtert den Zugang zur

statistischen Datenanalyse. Ein Umstand, der in unserer modernen Informationsgesellschaft und dem Drang nach datengestützter Entscheidungsfindung von großer Bedeutung ist. Die Arbeitsergebnisse der Analysen werden in SPSS übersichtlich in präsentationsfähigen Tabellen und Grafiken dargestellt. Die SPSS-Ergebnisausgaben können anschließend weiter bearbeitet und in andere PC-Programme exportiert werden. Die Vielzahl der Programm-Module von IBM SPSS⁷ bieten ein beeindruckendes Repertoire statistischer Funktionen und unterstützen dadurch die Primär- und Sekundärforschungen technisch und methodisch umfangreich. Das SPSS-Basismodul ermöglicht das grundlegende Management der Daten sowie die statistische und grafische Analyse der Daten und kann um speziellere statistische Verfahren ergänzt werden. Der ausgezeichnete Bedienkomfort mit einer eingängigen charakteristischen Systematik und einer eigenen Programm-Arbeitslogik sowie die umfassenden Funktionen machen SPSS zu einem sehr beliebten und etablierten Statistikprogramm von Forschung, Wirtschaftspraxis und Lehre.⁸

Die vorliegende Publikation erläutert im Kapitel 2 ausgewählte zentrale Grundlagen der Statistik, die für das Verständnis der Arbeit mit IBM SPSS und der statistischen Analyse nötig sind. In diesem Kapitel werden grundlegende Terminologien, Herausforderungen bei Messungen und der Verarbeitung von Daten erläutert. Das Kapitel 3 und 4 gibt einen Überblick über wesentliche Arbeitsschritte in IBM SPSS Statistics und handelt davon, wie Daten und Variablen in SPSS eingearbeitet werden. Danach werden in den anschließenden Kapiteln dieses Einstiegswerks charakteristische Problemstellungen der Praxis behandelt, welche sich aufgrund der Datensätze mit Daten unterschiedlichen Datentyps ergeben. Die Kapitel 5, 6 und 7 enthalten Erläuterungen der zentralen, immer wiederkehrenden Fertigkeiten der Datenbearbeitung mit IBM SPSS Statistics. Es geht im Kapitel 6 darum, wie Daten in Klassen eingeteilt und spezielle Fälle des Datensatzes ausgewählt werden. Beschrieben wird in Kapitel 7, wie aus vorhandenen Variablen bestimmte andersartige Variablen mit dem System erzeugt werden können und wie die Richtigkeit der gemachten Dateneingaben geprüft werden kann. In Kapitel 8 wird der Umgang mit verschiedenen Editoren von SPSS und in Kapitel 9 das Erstellen von Datenübersichten sowie die grafische Aufbereitung vorhandener Daten besprochen. Die Darstellung der erweiter-

⁷ Vgl. IBM (o. J.), SPSS Statistics-Produkte, www-01.ibm.com/software/de/analytics/spss/products/statistics/products.html (Stand: 22.11.2015).

⁸ Vgl. Wilke, H. (1986), S. 3. Möchte man SPSS nutzen, müssen Lizenzgebühren, gewisse Hardwareanforderungen und eine umfangreiche Programmaktivierung akzeptiert werden. Lizenzkostenfreie Alternativen zu IBM SPSS sind zum Beispiel: Gretl, Instat+, Gnu_R, Statistiklabor, OpenStat.

ten ausgewählten Datenanalyse mit IBM SPSS Statistics findet ab Kapitel 10 statt. In diesen Kapiteln geht es um die simultane Betrachtung und Analyse mehrerer Variablen einer Datenerhebung. Es folgen Anleitungen für Hypothesen- und Struktur-Überprüfungen in Kombination mit einem etwas umfangreicheren Theorieanteil und anschließend gerafft Korrelations- und Regressionsanalysen in Kapitel 11 sowie Hinweise für multivariate Datenanalysen in Kapitel 12. Das Kapitel 12 bildet mit einem Resümee und Quick-References den Abschluss.

Die Problemstellungen werden mit ausgewählten statistischen Methoden sowie mit IBM SPSS Statistics 23 (vgl. Gliederungspunkt „Vorbemerkung und Danksagung“) bearbeitet. Die ausgewählten Funktionalitäten und Menüpunkte der Statistiksoftware IBM SPSS Statistics werden zusammen mit den Nutzungsanleitungen der statistischen Analyseverfahren erläutert. Mathematische Formeln und formale Prozeduren werden nur dann aufgeführt, wenn sie zwingend für das Verständnis der Erläuterungen erforderlich sind. Die Darstellungen der **Arbeitskonzepte der Praxis** stehen im **Vordergrund** der Ausführungen. Hierfür kommt ein Fragebogen zum Einsatz, der so ausgelegt ist, dass die dargestellten Arbeitskonzepte trainiert werden können. Die erläuterten Erhebungen sollen die typischen Arbeitskonzepte unterstützen und Übungszwecken dienen. Objektivität, Reliabilität und Validität einer Erhebung müssen im Rahmen dieser Einführungsschrift vernachlässigt bleiben. Die dargestellten Anschauungsbeispiele und Datengrundlagen sind fiktiv gewählt. Anhand der verwendeten Interviewfragen (vgl. zum Beispiel Anlage) und einer selbstgewählten Datenbasis sollte der Leser die dargestellten SPSS-Prozeduren nachbilden können. Zu Übungszwecken lassen sich Dateneingaben technisch vervielfachen und damit die Analyseübungsbasis erweitern. Empfehlenswert ist, unvertraute SPSS-Prozeduren anhand von kleinen, sehr übersichtlichen Datensätzen - die sich aufgrund ihrer Größe auch zur Sichtkontrolle eignen - auszuprobieren.

Für bestimmte Fachbegriffe, Bezeichnungen, Beispiele, Steuerelemente, technische Komponenten und Anweisungen werden im Text bestimmte Kennzeichnungen und Schriften zur besseren Orientierung eingesetzt. Ein Variablenname erhält die Schrift *variablenname*, Menüpunkt, Fenstertitel und Schaltfläche der Software erhalten die Schrift **MENÜPUNKT**. Die Darstellung der Bedienung der grafischen Benutzeroberfläche von IBM SPSS folgt in der Regel einem einheitlichen Prinzip. Zum Beispiel „**Daten einer Datei öffnen**“ wird im Text folgendermaßen dargestellt.

1. **DATEI** anklicken,
2. **ÖFFNEN** anwählen,
3. **DATEN...** anklicken,
4. im Fenster gewünschte Datei mit den Daten anklicken,
5. **ÖFFNEN** anklicken.

Abbildung 1: Daten einer Datei öffnen⁹

Dies bedeutet, dass im aufgerufenen und aktiven System IBM SPSS Statistics in der Menüleiste **1.** der Menüpunkt „**DATEI**“ angeklickt, dann in der erscheinenden Dropdown-Liste **2.** der Menüpunkt „**ÖFFNEN**“ angewählt und anschließend in der nächsten Liste **3.** der Menüpunkt „**DATEN...**“ angeklickt wird, was daraufhin **4.** ein Dialogfenster öffnet, in dem die gewünschte Datei mit den gewünschten Daten ausgewählt, angeklickt und **5.** durch ein weiteres Klicken auf die Schaltfläche „**ÖFFNEN**“ aktiviert und geöffnet wird. Aus Komprimierungsgründen kommt es im Text der Publikation vor, dass die SPSS Prozeduren in Form einer chronologischen Sequenz der SPSS Kommandos ausgeführt werden (**DATEI**, **ÖFFNEN**, **DATEN...**, **ÖFFNEN**). Zusätzlich werden zur Veranschaulichung der Vorgehensweise in SPSS entsprechende Screenshots der Anwendung in den Text eingebaut.

Manche Problemstellungen lassen sich in SPSS auf unterschiedliche Weise lösen. Die dargestellten Lösungen mit SPSS sind daher als bewährte Vorschläge zu verstehen.

⁹ Eigene Darstellung.

2. Messung und Typen von Daten

2.1 Messung von Daten

Daten sind **zusammengefasste Zeichen**,¹⁰ die aufgrund von Übereinkünften Informationen - das heißt Angaben über Sachverhalte - darstellen.¹¹

Beispiel 1 - Daten:

→ 28 Jahre; 150 €; 180 cm; weiblich/männlich; viel/wenig/nichts etc..

Abbildung 2: Datenbeispiel¹²

Zur näheren Veranschaulichung der weiteren Erläuterungen stelle man sich zum Beispiel einen Bio-Landwirt vor, der in Deutschland eine noch wenig verbreitete Bio-Knolle (Topinambur) im Handel einführen möchte. Diese Knolle kann in der Küche als Alternative für die Kartoffel verwendet werden und bietet dem Landwirt zahlreiche Vorteile.¹³

Vor der Markteinführung dieser Knolle möchte der Landwirt eine **empirische Datenbasis** als **Entscheidungsgrundlage** für seine **Markteinführungsmaßnahmen** aufbauen.¹⁴ Der Landwirt hat verschiedene Vermutungen und Hypothesen, wie sich seine Zielgruppe wohl verhalten wird. Zum Beispiel vermutet der Landwirt, dass die Zahlungsbereitschaft bei Frauen signifikant ausgeprägter ist als jene bei Männern oder der Landwirt vermutet, dass das Interesse an der Bio-Knolle ausgesprochen hoch ist. Den Landwirt interessieren daher explizit verschiedene Sachverhalte, die er bei der Zielgruppe gezielt beobachten und erfassen möchte. Diese Sachverhalte sind zum Beispiel die Höhe der Zahlungsbereitschaft der potentiellen Ziel-

¹⁰ Zahl oder Symbol.

¹¹ Vgl. Sedlmeier, P., u. a. (2013), 56 ff.

¹² Eigene Darstellung.

¹³ Vgl. Baumann, A. (2011), Wer kennt Topinambur, www.online-artikel.de/article/wer-kennt-topinambur-87156-1 (Stand: 22.11.2015): Topinambur ähnelt im Aussehen einer Kartoffel und blüht im Herbst wie eine Sonnenblume. Topinambur enthält viele Mineralien sowie Vitamine, aber keine Stärke, sondern das Zucker-„Äquivalent“ Inulin, das beim Verzehr in einen für Diabetiker/Diabetikerinnen besonders verträglichen Fruchtzucker umgewandelt wird. Außerdem führt der Verzehr von Topinambur dazu, dass sich schnell ein Sättigungsgefühl einstellt (Appetitzügler). Die Knolle kann roh und gekocht in vielerlei Speisevariationen gegessen werden. Die Knolle ist frostunempfindlich und wird im Spätherbst, nachdem das Blattwerk und die gelben Blüten der Pflanze verwelkt sind, geerntet. Mit dem Austrieb der Knolle im Frühjahr endet die Erntezeit der Knolle. Die dünnschalige Knolle kann im Kühlschrank bis zu zwei Wochen gelagert werden.

¹⁴ Vor der Umsetzung von Erhebungen empfiehlt sich die Formulierung einer Projektskizze (vgl. Baumann, A., u. a. (2010), S. 25 ff..

gruppe in Euro, die Bewertung des Aussehens der neuartigen Knolle, das Interesse an der Bio-Knolle, die Bekanntheit der Knolle bei der potentiellen Zielgruppe. Mit Hilfe von Befragungen¹⁵ von 250 Zielpersonen werktags an ausgewählten Kontaktpunkten seines Verkaufsgebietes, plant der Landwirt mit Forschungskollegen/innen Anhaltspunkte und Größen der interessierenden Sachverhalte zu erheben, um seine Hypothesen zu überprüfen. Dafür entwickelt der Landwirt einen Fragebogen mit überwiegend geschlossenen Fragen, die er in den Befragungen einsetzen möchte (Fragen zu Zahlungsbereitschaft, Jahres-Nettoeinkommen, Geschlecht, Alter, Ausbildung, Einstellungen, Gewohnheiten, Interessen etc.). Die Antworten der Testpersonen werden auf den Fragebögen notiert und danach mit IBM SPSS Statistics erfasst.

Zur Erhebung von interessierenden Größen der Zahlungsbereitschaft entwickelt der Landwirt zum Beispiel folgende Fragen für seine Interviews:¹⁶

- Frage X: Wenn Sie die Knollen kaufen könnten (Interviewer/in: Knollen vorlegen), was wären Sie bereit für 1 kg dieser Knollen in € auszugeben? _____ €
- Frage Y: Wenn Sie die Knollen mit Kartoffeln vergleichen (Interviewer/in: Knollen und Kartoffeln vorlegen), wären Sie bereit mehr oder weniger Geld für die Knollen als für Kartoffeln auszugeben? Bitte treffen Sie eine Entscheidung:
___ = würde mehr ausgeben,
___ = würde weniger ausgeben.

Die Angaben der Interviewpartner/innen¹⁷ zur Zahlungsbereitschaft sind für den Bio-Landwirt die ersten Anhaltspunkte dafür, mit welchem Preis die Knolle bei Markteinführung im Handel wohl ausgezeichnet werden könnte. Sie liefern wichtige Informationen für zahlreiche betriebswirtschaftliche Überlegungen und Entscheidungen des Landwirts.

¹⁵ Beispielsweise als Face-to-Face-Interviews; vgl. Baumann, A., u. a. (2010), S. 66 ff.; andere Methoden sind denkbar, werden hier jedoch nicht behandelt.

¹⁶ Die Betrachtung der Güte der Fragen wird für dieses Anschauungsbeispiel nicht vorgenommen. Vgl. hierzu beispielsweise Baumann, A., u. a. (2010), S. 31 ff.; vgl. Schaich, E., u. a. (1984), S. 5 f.

¹⁷ Interviewpartner/innen sind gleichbedeutend mit Testpersonen, Probanden/innen, Fälle, Beobachtungseinheiten, Merkmalsträger, Untersuchungsobjekten.

Beispiel 2 - Zahlungsbereitschaft:

Frage X: Wenn Sie die Knollen kaufen könnten (Interviewer/in: Knollen vorlegen), was wären Sie bereit für 1 kg dieser Knollen in € auszugeben?
_____ €

→ Antwort 8 €;

→ Herstellungskosten für den Landwirt liegen bei 2 €;

→ $8 \text{ €} - 2 \text{ €} = 6 \text{ €}$ ergibt einen Überschuss von 6 €.

Abbildung 3: Entscheidungsgrundlage erarbeiten¹⁸

Dabei stellen die Interviewfragen die **Messinstrumente** dar, mit denen der Landwirt Daten über die entsprechende Zahlungsbereitschaft der potentiellen Kunden erfasst. Die **erfassten Daten** der Studie **ergeben** zusammengestellt den analysierbaren **Datensatz**. Dieser Datensatz ist die Informationsbasis und Entscheidungsgrundlage für Maßnahmen des Landwirts. Die **interessierenden Größen** der zu ergründenden Sachverhalte und Vorgänge werden als **Merkmale** oder **Variablen**¹⁹ bezeichnet und die **erfassten Zahlen** als **Ausprägungen** der Merkmale. Jede Frage in einem Interview zielt daher auf die Ergründung eines Merkmals oder einer Variablen und deren Ausprägungen ab (vgl. Beispiel oben). Die Eigenschaften der Untersuchungsobjekte, die im Zusammenhang mit der Problemstellung interessieren, sollen aufgedeckt werden. Zum Beispiel geht es in Frage X um Ausprägungen des Merkmals „Zahlungsbereitschaft“. **Zur Messung** einer **Merkmalsausprägung** wird dabei ein geeigneter **Maßstab** - eine **Skala** - entwickelt. Jeder Antwort auf eine Frage wird anhand der entwickelten Skala ein Zeichen²⁰ zugeordnet. Das Zeichen stellt das Äquivalent der jeweiligen Ausprägung des Merkmals dar. Eine Skala umfasst daher die Menge möglichst aller denkbaren Merkmalsausprägungen.

Beispiel 3 - Merkmal/Variable, Ausprägung, Maßstab/Skala:

→ „Zahlungsbereitschaft“ in Frage X = Merkmal/Variable;

→ „8 €“ (Antwort auf Frage X) = Ausprägung/Merkmalsausprägung;

→ „_____ €“ (in Frage X) = Maßstab/Skala.

Abbildung 4: Terminologie²¹

¹⁸ Eigene Darstellung.

¹⁹ Bei Variablen werden unabhängige und abhängige Variablen unterschieden. Ursache für eine abhängige Variable ist eine unabhängige Variable.

²⁰ Zahl oder Symbol; vgl. Sedlmeier, P., u. a. (2013), S. 56 ff.; vgl. Schaich, E., u. a. (1984), S. 4 f..

²¹ Eigene Darstellung.

Das heißt:²²

- **Messungen** sind planmäßige Beobachtungen interessierender Variablen,
- deren Resultate die strukturierte Zuordnung von Zeichen - Zahlen oder Symbolen - zu **beobachteten Ausprägungen** der Variablen sind,
- wobei die Relationen zwischen den Beobachtungen durch die Relationen zwischen den zugeordneten Zeichen repräsentiert werden (Homomorphismus).

Die Messung des Landwirts ist damit ein wesentliches Element seiner **Markt- und Marketingforschung** - also der **systematischen Suche, Sammlung, Aufbereitung und Interpretation** der **relevanten Informationen**, die sich **auf alle Probleme** des **Marktes** und des **Marketings** dieser neuartigen Knolle **beziehen**.²³ Anhand der empirischen Datenbasis versucht der Landwirt Gesetzmäßigkeiten zu entdecken, zu erläutern und zu prognostizieren und nutzt dafür die Methoden der Statistik.

Statistiken sind **mittels Zahlen systematisch beschriebene und analysierte Massenerscheinungen**. Massenerscheinungen sind Ereignisse, die unter gleichen Bedingungen immer wieder beobachtet werden können. Statistiken sind notwendig, wenn **empirische Befunde zahlenmäßig charakterisiert werden sollen**.²⁴ Mit IBM SPSS Statistics werden die Daten übersichtlich in Tabellen und Grafiken überführt. Daraus wird zum Beispiel die Höhe der Zahlungsbereitschaft abgelesen und ermittelt, ob sich die Einführung der Knolle wirtschaftlich empfiehlt.

In der vorliegenden Publikation werden mit IBM SPSS Statistics ausgewählte statistische Methoden der **deskriptiven** (beschreibenden), der **explorativen** und der **induktiven** (inferentiellen, schließenden) Statistik umgesetzt.²⁵ Die deskriptive Statistik stellt die gesammelten Daten übersichtlich in Form von Tabellen und Grafiken dar. Die explorative Statistik untersucht Zusammenhänge in den Daten und die induktive (inferentielle) Statistik begründet mittels wahrscheinlichkeitstheoretischer Methoden (stochastische Modelle) Allgemeingültigkeiten. Die induktive Statistik

²² Vgl. Baumann, A., u. a. (2010), S. 26 f.; vgl. Berekoven, L., u. a. (2006), S. 60; vgl. Altobelli, C. F. (2007), S. 162 f.; vgl. Schaich, E., u. a. (1984), S. 3 f.; Sedlmeier, P., u. a. (2013), S. 56 ff..

²³ Vgl. Altobelli, C. F. (2007), S. 5; vgl. Böhrler (2004), S. 19; vgl. zum Forschungsprozess Baumann, A., u. a. (2010); vgl. Berekoven, L., u. a. (2006), S. 34 f.; vgl. Altobelli, C. F. (2007), S. 17 f..

²⁴ Vgl. Schaich, E., u. a. (1984), S. 1.

²⁵ Vgl. Bühner, M., u. a. (2009), S. 13 ff..

macht damit Aussagen über die **Grundgesamtheit**.²⁶ Die interessierende Grundgesamtheit des Bio-Landwirts ist hier zum Beispiel die Menge aller Frauen und Männer seines Verkaufsgebiets. Generell umfasst die Grundgesamtheit alle Untersuchungsobjekte, die sich durch (sachlich, zeitlich, räumlich) definierte, gemeinsame Merkmale zusammenfassen lassen.²⁷ Häufig ist es nun unmöglich oder ineffizient, alle Personen der Grundgesamtheit zu befragen (**Vollerhebung**). Aus diesem Grund wählt der Bio-Landwirt auch aus seiner Grundgesamtheit nach bestimmten Kriterien und Regeln 250 der insgesamt 2.500 Personen seines Verkaufsgebietes aus (**Stichprobe**).²⁸ Anhand der Befragung dieser 250 **Fälle (Stichprobenumfang)** werden dann die Ausprägungen der interessierenden Variablen beobachtet und Informationen über die Zahlungsbereitschaft der definierten Menge von Beobachtungseinheiten oder Merkmalsträger gewonnen (**Stichprobenerhebung**). Die Stichprobe als Teilmenge der Grundgesamtheit sollte daher ein möglichst getreues (**repräsentatives**) Abbild der Grundgesamtheit darstellen.²⁹ Die aus einer Stichprobe abgeleiteten Erkenntnisse besitzen Allgemeingültigkeit und sind repräsentativ,³⁰ wenn die Generierung der Stichprobeneinheiten aus der Grundgesamtheit mit Hilfe eines qualifizierten Zufallsmechanismus erfolgte (**Zufallsstichprobe**).³¹

2.2 Typen von Daten

Die Messung der Ausprägungen der Zahlungsbereitschaft erfolgt im Beispiel mit Hilfe der Frage X und der Frage Y.

Zur Erinnerung:

- Frage X: Wenn Sie die Knollen kaufen könnten (Interviewer/in: Knollen vorlegen), was wären Sie bereit für 1 kg dieser Knollen in € auszugeben? _____ €.
- Frage Y: Wenn Sie die Knollen mit Kartoffeln vergleichen (Interviewer/in: Knollen und Kartoffeln vorlegen), wären Sie bereit mehr oder weniger Geld für die Knollen als für Kartoffeln auszugeben?

²⁶ Vgl. Bühner, M., u. a. (2009), S. 101 ff., 139 ff.

²⁷ In der Praxis ist es häufig unmöglich die reale Grundgesamtheit vollständig festzulegen. Vgl. Hammann, P., u. a. (1994), S. 105 ff.; vgl. Decker, R. u. a. (2002), S. 183 f.

²⁸ Zu Stichprobenverfahren (zum Beispiel Zufallsstichprobe) vgl. Baumann, A., u. a. (2010), S. 49 ff.; Altobelli, C. F. (2007), S. 182 ff.; Gabler, S., u. a. (2002).

²⁹ Vgl. Baumann, A., u. a. (2010), S. 45 ff.; vgl. Altobelli, C. F. (2007), S. 182 ff.

³⁰ Wird auf die Grundgesamtheit übertragen.

³¹ Vgl. Altobelli, C. F. (2007), S. 182 ff.; vgl. Baumann, A., u. a. (o. J.), S. 49 ff.; vgl. Decker, R. u. a. (2002), S. 183 ff.; vgl. Schaich, E., u. a. (1984), S. 14 f.