

Stefanie van Ophuysen
Bernd Fischer | Lars Behrmann
Bea Bloh

Statistik verstehen

Band 1: Deskriptive Statistik
für die Bildungswissenschaften

Prof. Dr. Stefanie van Ophuysen leitet die Arbeitsgruppe Forschungsmethoden/Empirische Bildungsforschung am Institut für Erziehungswissenschaft der Universität Münster. Dr. Bernd Fischer ist Lehrbeauftragter am Institut für Erziehungswissenschaft der Universität Münster.

Dr. Lars Behrmann ist Studienrat im Hochschuldienst in der AG Forschungsmethoden/Empirische Bildungsforschung am Institut für Erziehungswissenschaft der Universität Münster.

Prof. Dr. Bea Bloh lehrt Schulpädagogik der Sekundarstufe I am Institut für Erziehungswissenschaft an der Universität Paderborn.

Stefanie van Ophuysen, Bernd Fischer,
Lars Behrmann, Bea Bloh

Statistik verstehen

Band 1: Deskriptive Statistik
für die Bildungswissenschaften

Waxmann
Münster · New York

Online-Angebote oder elektronische Ausgaben sind erhältlich
unter www.utb-shop.de

Bibliografische Information der Deutschen Nationalbibliothek

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

utb 5585
ISBN 978-3-8252-5585-5

© Waxmann Verlag GmbH, 2021

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Atelier Reichert, Stuttgart
Satz: satz&sonders GmbH, Dülmen
Druck: CPI books
Gedruckt auf alterungsbeständigem Papier gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages
in irgendeiner Form reproduziert oder unter Verwendung elektronischer
Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

1	Einleitung	9
2	Daten und Messen	11
2.1	Objektbereich	12
2.2	Vom Objektbereich zum Datenbereich	16
2.3	Messniveaus	19
2.3.1	Nominalskala	19
2.3.2	Ordinalskala	20
2.3.3	Intervallskala	21
2.3.4	Verhältnisskala	23
2.3.5	Weitere Hinweise	23
2.4	Die Datenmatrix	26
3	Daten zusammenfassend darstellen: Häufigkeiten	31
3.1	Univariate Häufigkeitsverteilungen	31
3.1.1	Absolute und relative Häufigkeiten	31
3.1.2	Kumulierte Häufigkeiten	37
3.1.3	Häufigkeitsverteilungen bei metrischen Variablen	40
3.1.4	Zusammengefasste (klassierte) Daten	44
3.2	Bivariate Häufigkeitsverteilungen	49
4	Daten zusammenfassend darstellen: Verteilungskennwerte	57
4.1	Lagemaße	58
4.1.1	Nominalskalenniveau: der Modalwert	58
4.1.2	Ordinalskalenniveau: der Median	59
4.1.3	Metrisches Skalenniveau: das arithmetische Mittel	63
4.1.4	Vergleich der Maße der zentralen Tendenz	69
4.2	Weitere Lagemaße: die Perzentile	71
4.2.1	Bestimmen von \tilde{x}_p , wenn $n \cdot p$ ganzzahlig	71
4.2.2	Bestimmen von \tilde{x}_p , wenn $n \cdot p$ nicht ganzzahlig	72
4.2.3	Weitere Hinweise zu Perzentilen	74
4.3	Streuungsmaße	75
4.3.1	Die Spannweite	76
4.3.2	Der Quartilsabstand	76
4.3.3	Varianz und Standardabweichung	76
4.4	Nutzung von Verteilungskennwerten	81
4.5	Grafische Darstellung der Verteilung auf Basis statistischer Kennwerte	83

6 Inhalt

4.5.1	Das Perzentilband	83
4.5.2	Der Box-Plot	88
5	Vergleichende Analysen auf Basis von Häufigkeiten	91
5.1	Datenstrukturen	91
5.2	Vergleich der Häufigkeitsverteilungen bei verbundenen Messreihen .	93
5.3	Vergleich der Häufigkeitsverteilungen bei unverbundenen Messreihen	96
6	Vergleichende Auswertungsstrategien auf Basis statistischer Kennwerte . .	105
6.1	Mittelwertvergleiche bei unverbundenen Messreihen	106
6.2	Mittelwertvergleiche bei verbundenen Messreihen	112
6.3	Inner-Subjekt-Faktoren und Zwischen-Subjekt-Faktoren	116
6.4	Kombination von zwei Zwischen-Subjekt-Faktoren	119
6.5	Kombination von einem Zwischen-Subjekt-Faktor und einem Inner-Subjekt-Faktor	126
6.5.1	Mittelwertvergleiche im gemischten zweifaktoriellen Design	127
6.5.2	Mittelwertvergleich im gemischten Design mit Messwiederholung	130
6.6	Kombination von zwei Inner-Subjekt-Faktoren	133
7	Zusammenhangsmaße für kategoriale Variablen	139
7.1	(Bei Unabhängigkeit) Erwartete Häufigkeiten	140
7.2	Der χ^2 -Koeffizient	147
7.3	Weitere Koeffizienten auf Basis von χ^2	151
7.3.1	Der Phi-Koeffizient	152
7.3.2	Cramér's V	152
7.3.3	Der Kontingenzkoeffizient K	153
7.3.4	Der korrigierte Kontingenzkoeffizient K^*	154
8	Zusammenhangsmaße für metrische Variablen	155
8.1	Grafische Darstellung: Punktewolken im Streudiagramm	156
8.2	Der Korrelationskoeffizient	163
8.2.1	Der Korrelationskoeffizient nach Fechner	164
8.2.2	Der Korrelationskoeffizient nach Pearson	167
8.3	Zur Interpretation von Korrelationen	171
9	Einfache lineare Regression	181
9.1	Bestimmung der Regressionsgleichung	181
9.1.1	Lineare Funktionen	182
9.1.2	Residuen	183
9.1.3	Die Idee der Kleinst-Quadrat-Schätzung	184

9.1.4	Bestimmung der Regressionsparameter	188
9.1.5	Interpretation der Regressionsparameter	190
9.1.6	Zum Zusammenhang von Korrelation und Steigungsparameter	191
9.2	Wie gut repräsentiert die Regressionsgerade die empirischen Daten? .	192
9.2.1	Minimum und Maximum für SS_R	193
9.2.2	Das Bestimmtheitsmaß R^2	194
9.2.3	Interpretation des Bestimmtheitsmaßes R^2	195
9.3	Lassen sich auch x -Werte über eine Regression vorhersagen?	197
10	Feierabend!	201
11	Literatur	202
12	Abbildungsverzeichnis	204
13	Tabellenverzeichnis	207
14	Aufgabenverzeichnis	209
15	Index	212

1 Einleitung

Statistik verstehen! Dabei möchten wir Sie mit diesem Buch unterstützen.

Statistik ist aus keiner (bildungs-)wissenschaftlichen Disziplin mehr wegzudenken. Egal, ob Sie als Wissenschaftler*in in Forschungsprojekten tätig sind, als Student*in im Kontext des Forschenden Lernens an Studien- oder Lehrforschungsprojekten beteiligt sind, als Lehrkraft datengestützte Schul- oder Unterrichtsentwicklung betreiben oder in irgendeinem anderen sozialwissenschaftlichen Berufsfeld die Wirksamkeit von Programmen evaluieren: Die Statistik liefert uns ein umfangreiches Repertoire an Methoden, mit denen wir unterschiedliche, in Daten verpackte Informationen zusammenfassen, veranschaulichen und analysieren können. Statistik hilft uns so, Erkenntnisse über den uns jeweils interessierenden (Forschungs-)Gegenstand zu generieren.

Wir möchten Ihnen zeigen, dass Statistik kein Hexenwerk ist. Mit recht einfacher Mathematik, einer guten Prise Logik und gesundem Menschenverstand lassen sich die Grundideen der deskriptiven (= beschreibenden) Statistik gut durchschauen. Mit unserem Buch möchten wir Sie mit zentralen Grundbegriffen vertraut machen und Ihnen Verfahren vorstellen, mit denen Sie typische bildungswissenschaftlich relevante Fragestellungen beantworten können.

Kleine Vorwarnung

Statistik basiert auf mathematischen Grundlagen und die Sprache der Mathematik erlaubt, Sachverhalte sehr knapp und gleichzeitig präzise darzustellen. Diese Möglichkeit nutzen wir und somit ist dieses Buch keines, in dem Statistik „ohne Formeln“ erklärt wird. Aber wir erläutern die Formeln sehr genau und begründen, warum Kennwerte auf die jeweilige Art gebildet und warum Verfahren auf die jeweilige Art durchgeführt werden.

Um neue Sachverhalte nachhaltig zu lernen, reicht das einfache Lesen eines Textes in der Regel nicht aus. Lernen ist ein aktiver Prozess, der die zielgerichtete Auseinandersetzung mit den Inhalten erfordert. In der Psychologie werden unterschiedliche kognitive Strategien unterschieden, mit denen der Lernprozess erfolgreich gestaltet werden kann. Einfache Wiederholungsstrategien (z. B. Texte abschreiben, Definitionen auswendig lernen, Routinen wiederholt anwenden) dienen der Festigung von Wissen im Langzeitgedächtnis. Elaborationsstrategien (z. B. Sachverhalte in eigenen Worten formulieren, neues Wissen auf ein bekanntes Beispiel transferieren, Wissensbestandteile verknüpfen) unterstützen die Vernetzung mit bestehendem Vorwissen. Organisationsstrategien (z. B. Zusammenfassungen schreiben, Mindmap erstellen) helfen, dem neuen Wissen Struktur zu geben und dieses in ein kohärentes Ganzes zu überführen. Wenn Sie mit diesem Buch nachhaltig lernen wollen, sollten Sie alle drei Typen von Strategien nutzen.

Vorschläge zum Gebrauch dieses Buches

Dieses Buch ist ein Arbeitsbuch und es ist so konzipiert, dass Sie sich die Inhalte mit seiner Hilfe eigenständig aneignen können.

 Wir unterstützen Sie dabei, indem wir Ihnen im Text zahlreiche Übungsaufgaben zur Verfügung stellen. Dabei gibt es Aufgaben, die der Wiederholung und Festigung von Wissen und Routinen dienen. Vielleicht möchten Sie – wie in der Schule – ein Heft führen, in dem Sie die Aufgaben bearbeiten. Teilweise können die Aufgaben auch direkt im Buch gelöst werden.

 Hervorheben möchten wir aber ganz besonders die sogenannten Denk-mit!-Aufgaben, die das aktive Nach- bzw. Weiterdenken anregen und damit im Sinne der Elaboration das neue Wissen mit bestehendem verknüpfen. Bitte beachten Sie: Lösungen und Erläuterungen zu diesen Denk-mit!-Aufgaben folgen unmittelbar dahinter im Text. Für einen bestmöglichen Lernerfolg sollten Sie also nicht direkt weiterlesen, sondern sich erst einmal ein wenig Zeit nehmen, um eigene Überlegungen anzustellen.

Versuchen Sie den Überblick über den Stoff insgesamt und die Verbindungen untereinander zu wahren, indem Sie Zusammenfassungen erstellen oder Begriffe in Mindmaps visualisieren. Im Rahmen der Statistik-Vorlesungen, auf deren Basis dieses Buch entstanden ist, dürfen die Studierenden einen einseitig beschriebenen DIN-A4-Zettel als „Spickzettel“ mit in die Klausur nehmen. Vielleicht ist die schrittweise Entwicklung eines solchen Zetts auch für Sie eine Möglichkeit, Ihr Wissen zu sortieren.

Ganz besonders möchten wir Sie ermuntern, dieses Buch zur Arbeit in einer Lerngruppe zu nutzen. Vergleichen Sie Ihre Lösungen für die Übungsaufgaben und vor allem: Lösen Sie gemeinsam die Denk-mit!-Aufgaben. Sprechen Sie über die Inhalte, tauschen Sie Ihre Ideen aus. Die eigenen Ideen zu verbalisieren, ist eine weitere Möglichkeit, das Wissen nachhaltig zu verankern. Daneben macht das Arbeiten im Team in der Regel mehr Spaß und mit Spaß lernt es sich bekanntlich besser.

Wir können und wollen Ihnen also nicht versprechen, dass Sie sich die Inhalte mit Hilfe dieses Buches ganz mühelos und in kürzester Zeit aneignen können. Lernen bedarf immer auch der kognitiven Anstrengung. Aber die hier gewählte, sehr kleinschrittige Herangehensweise mit ausführlichen Erläuterungen liefert hoffentlich auch Ihnen einen Weg über den (Statistik-)Berg, der nicht zu steil und nicht zu steinig ist. Selbst wenn Sie nicht zu den Personen gehören, denen Mathematik zufliegt, sind wir optimistisch, dass Sie nach der Erarbeitung dieses Buches keinen Grund (mehr) haben, Statistik zu fürchten, sondern dass Sie Statistik als ein nützliches Werkzeug für Studium und Forschung schätzen werden. Vielleicht macht Ihnen Statistik danach ja sogar Spaß – so wie uns!

Wir wünschen Ihnen eine erkenntnisreiche Arbeit mit diesem Buch!

2 Daten und Messen

Wir verstehen unter Statistik diejenigen mathematisch fundierten Verfahren, die der quantitativen, zusammenfassenden Darstellung von Daten und deren Analyse zur Beantwortung von Untersuchungsfragen dienen. Bevor wir uns in den folgenden Kapiteln dieses Buches der statistischen Beschreibung und Auswertung von Daten zuwenden können, sind einige grundsätzliche Überlegungen und Erläuterungen zum Thema „Daten“ erforderlich. Besondere Beachtung widmen wir dabei dem Zusammenhang zwischen den untersuchten Merkmalen in der „Realität“ und den Daten, die diese Merkmale in der statistischen Analyse repräsentieren.

Die Basis empirischer wissenschaftlicher Studien sind Informationen über ausgewählte Merkmale bei bestimmten Merkmalsträgern: zum Beispiel die Schulfreude bei Kindern in der Grundschule, das Unterrichtsklima in Schulklassen oder die Belastung pädagogischer Fachkräfte in Kitas. Diese Informationen werden mit Hilfe wissenschaftlicher Erhebungsmethoden wie beispielsweise Beobachtung oder Befragung¹ erfasst und stehen der statistischen Analyse als Daten, meist in Form von Zahlen und seltener in Form anderer Symbole wie Buchstaben,² zur Verfügung. Abstrakt gesprochen wird der in einer Untersuchung interessierende Realitätsausschnitt, der Objektbereich, in den Datenbereich abgebildet. Die Daten stellen als Ergebnis des Abbildungsprozesses ein reduziertes, auf die interessierenden Aspekte fokussiertes Abbild der Realität dar. Um sinnvolle Aussagen über die Realität zu ermöglichen, muss das Datenmaterial den interessierenden Realitätsausschnitt in seinen wichtigen Eigenschaften möglichst exakt widerspiegeln.

Um den Zusammenhang von Objekt- und Datenbereich zu verdeutlichen, betrachten wir zunächst den Objektbereich und seine für die Statistik bedeutsame Struktur. Danach betrachten wir, wie Informationen aus der Realität in Daten überführt werden können, und schließlich beleuchten wir die resultierenden Konsequenzen für das statistische Vorgehen.

1 Siehe zu Forschungsmethoden allgemein z. B. Diekmann (2010), Döring & Bortz (2016), Pfeiffer & Pittman (2018), Schnell, Hill & Esser (2018), zu Befragung z. B. Porst (2011) und zu Beobachtung van Ophuysen, Bloh & Gehrau (2017).

2 Daten in Form von Texten oder Bildern sind für die statistische Analyse nur schwer zugänglich und können durch sogenannte qualitative Auswertungsstrategien analysiert werden; zum Beispiel mit Hilfe der qualitativen Inhaltsanalyse, die auch computergestützt durchgeführt werden kann (z. B. mit MAXQDA, vgl. Kuckartz, 2018).

2.1 Objektbereich

Der Objektbereich stellt denjenigen Ausschnitt der Welt dar, über den Aussagen getroffen werden sollen. Im bildungswissenschaftlichen Kontext ist dies häufig der schulische Bereich, in dem Informationen über einzelne Akteur*innen, über Gruppen oder ganze Organisationen erfasst werden. Im Objektbereich können wir Merkmalsträger, Merkmale und Merkmalsausprägungen unterscheiden.

Merkmalsträger (oder Untersuchungsobjekte) sind die Einheiten, die untersucht werden. In bildungswissenschaftlichen Studien handelt es sich häufig um Einzelpersonen (Schüler*innen, Lehrkräfte, Studierende, Dozent*innen, Erzieher*innen in Kindergärten, Väter, Mütter etc.), manchmal sind aber auch Kollektive von Personen wie Schulklassen, Kollegien oder Familien Gegenstand der Untersuchung. Neben Personen können auch Produkte materieller oder immaterieller Art (z. B. Schultaschen, Lehrpläne) Objekte sein. Sie interessieren nicht in ihrer Gänze, sondern mit Blick auf spezifische, für die Untersuchungsfrage relevante Merkmale, deren „Träger“ sie sind.

Merkmale sind charakterisierende Eigenschaften von Merkmalsträgern, die einen Vergleich mit anderen Merkmalsträgern ermöglichen. In der Regel handelt es sich um Merkmale, die einem Individuum (z. B. Schüler*innen: Geschlecht, Deutschnote, Alter) oder einem Produkt (z. B. Schultaschen: Material, Gewicht) zugeordnet sind. Aber auch die Gruppenmerkmale eines Kollektivs können untersucht werden; sie können anhand von Individualmerkmalen aggregiert (z. B. Frauenanteil, Durchschnittsalter eines Kollegiums) oder global zugewiesen werden (z. B. Einschätzung des Arbeitsklimas in einem Kollegium durch externe Expert*innen). Ein wichtiges Kennzeichen von Merkmalen ist ihre Variabilität: Damit Merkmale zum Vergleich von Merkmalsträgern geeignet sind, müssen sie verschiedene Ausprägungen besitzen.

Ausprägungen sind die unterschiedlichen Zustände, die ein Merkmal annehmen kann. Während manche Merkmale nur wenige Ausprägungen haben (z. B. Fächerbereich der Leistungskurse in der gymnasialen Oberstufe: sprachlich-künstlerisch, gesellschaftswissenschaftlich, mathematisch-naturwissenschaftlich), sind bei anderen Merkmalen viele Ausprägungen möglich (z. B. Gewicht eines Neugeborenen gemessen in Gramm). Merkmale unterscheiden sich aber nicht nur in der Anzahl der Ausprägungen, sondern auch im Informationsgehalt, der in ihren Ausprägungen steckt: Für das Merkmal „Fächerbereich“ kann nur überprüft werden, ob zwei Schüler*innen ihren Leistungskurs im gleichen oder in unterschiedlichen Bereichen gewählt haben. Bei „Gewicht“ kann zusätzlich verglichen werden, welches Kind mehr wiegt als das andere.

Die Vergleichsmöglichkeiten zwischen Merkmalsträgern werden mit dem folgenden Beispiel verdeutlicht. In einer Studie zur Qualität von Kindertagesstätten werden vier Merkmale betrachtet: Träger, Gütesiegel, Gründungsjahr und Anzahl betreuter Kinder.

Tabelle 2-1: Merkmale und ihre (realisierten) Ausprägungen bei verschiedenen Kitas

Kita	Träger	Gütesiegel	Gründungsjahr	Anzahl Kinder
Brommberg	Stadt	A-	1972	50
Rappelkiste	AWO	B	1972	100
Hasenheide	AWO	C+	2001	50
St. Josef	Kath. Kirche	B	1958	75

Aufgabe 2-1: Formulieren vergleichender Aussagen

In Tabelle 2-1 finden Sie für jedes Merkmal bereits eine Reihe von vergleichenden Aussagen. Notieren Sie für jedes Merkmal je eine oder zwei weitere vergleichende Aussagen über die vier Einrichtungen.

Merkmalsausprägung	Aussage
Träger	Brommberg und St. Josef haben unterschiedliche Träger.
Gütesiegel	Rappelkiste und St. Josef haben das gleiche Gütesiegel. Rappelkiste hat ein besseres Gütesiegel als Hasenheide.
Gründungsjahr	Brommberg und Rappelkiste wurden im gleichen Jahr gegründet. Rappelkiste wurde früher gegründet als Hasenheide. Brommberg und Rappelkiste wurden 14 Jahre später als St. Josef gegründet.
Anzahl Kinder	Brommberg und St. Josef betreuen unterschiedlich viele Kinder. Rappelkiste betreut mehr Kinder als St. Josef. St. Josef betreut 25 Kinder mehr als Hasenheide. Rappelkiste betreut doppelt so viele Kinder wie Brommberg.

Die genauere Betrachtung der Vergleichsaussagen zeigt, dass vier Arten von Vergleichen möglich sind. So gibt es zunächst Aussagen über Gleichheit oder Ungleichheit der Ausprägungen in verschiedenen Kitas (z. B. Hasenheide und Rappelkiste haben denselben Träger, unterschiedliche Gütesiegel und unterschiedlich viele betreute Kinder). Weiterhin gibt es Aussagen über die Anordnung oder Reihenfolge der Merkmalsausprägungen (z. B. Brommberg hat das beste Gütesiegel, gefolgt von Rappelkiste und St. Josef; Hasenheide hat das schlechteste Gütesiegel; Hasenheide wurde später



gegründet als Brommberg; St. Josef ist größer als Hasenheide und Brommberg, aber kleiner als Rappelkiste). Schließlich können „mathematische“, zahlenbasierte Aussagen formuliert werden. Diese numerischen Aussagen können noch einmal untergliedert werden. Einerseits gibt es Aussagen, die sich auf den Abstand zwischen zwei Ausprägungen beziehen (z.B. Hasenheide wurde 29 Jahre später als Rappelkiste gegründet). Andererseits gibt es Aussagen, die sich auf das numerische Verhältnis zwischen zwei Ausprägungen beziehen (z.B. St. Josef betreut 1,5-mal so viele Kinder wie Hasenheide).



Aufgabe 2-2: Ordnen von Vergleichstypen

Denk mit! Ordnen Sie die Aussagen aus Aufgabe 2-1 den verschiedenen Vergleichstypen zu. Welche Aussage bezieht sich auf Gleichheit, Anordnung, numerischen Abstand oder numerisches Verhältnis der Merkmalsausprägungen? Was fällt Ihnen auf?

Es wird deutlich, dass bei allen Merkmalen ein Vergleich hinsichtlich Gleichheit oder Ungleichheit der Ausprägungen möglich ist. Für das Merkmal Träger kommen allerdings keine anderen Vergleichsformen in Betracht. Sinnvolle Aussagen über die Anordnung der Merkmalsausprägungen können für das Gütesiegel, das Gründungsjahr und die Anzahl betreuter Kinder gemacht werden, aber numerische Aussagen über Abstände zwischen den Merkmalsausprägungen sind nur für das Gründungsjahr und die Anzahl betreuter Kinder möglich. Schließlich lassen sich nur die Ausprägungen des Merkmals Anzahl Kinder in ihrem numerischen Verhältnis zueinander interpretieren.

Wir können erkennen, dass die Vergleichsarten hierarchisch sortiert sind: Wenn das numerische Verhältnis interpretiert werden kann, sind auch Aussagen über den Abstand möglich. Aussagen über den Abstand implizieren auch Aussagen über die Anordnung und diese wiederum Aussagen über Gleichheit. Wir sehen, dass die Merkmalsausprägungen sich in ihrem Informationsgehalt unterscheiden, weil sie unterschiedliche Aussagen über Beziehungen zwischen den Ausprägungen erlauben. Diese Beziehungen zwischen den Merkmalsausprägungen werden als Relationen bezeichnet.

- (1) Die **Äquivalenz- oder Gleichheitsrelation** macht Aussagen über die Gleichheit oder Ungleichheit der Ausprägungen (z.B. Geschlecht mit den Ausprägungen männlich, weiblich, divers; Familiensprache mit den Ausprägungen nur Deutsch, Deutsch und eine andere Sprache, kein Deutsch).
- (2) Die **Ordinal- oder Ordnungsrelation** sagt über die Äquivalenzrelation hinaus etwas über die Rangfolge der Merkmalsausprägungen (z.B. Schulleistung in Deutsch mit den Ausprägungen sehr gut bis ungenügend; Energieeffizienzklassen mit den Ausprägungen A+++ bis G). Beachten Sie: Die Abstände zwischen den Ausprägungen können nicht verglichen werden. Beispielsweise implizieren benachbarte

Noten nicht unbedingt auch identische Leistungsunterschiede; der Leistungsunterschied zwischen einer sehr guten (1) und einer guten Note (2) stimmt nicht unbedingt mit dem Unterschied zwischen gut (2) und befriedigend (3) überein.

- (3) Die **Distanz- oder Abstandsrelation** präzisiert die Information der Ordinalrelation durch die Angabe interpretierbarer Abstände zwischen den Ausprägungen – dafür ist die Angabe einer Einheit notwendig (z. B. Intelligenz in IQ-Punkten; Temperatur in Grad Celsius). In einem Raum mit 20 °C Raumtemperatur ist es um 4 Grad kälter als in einem Raum mit 24 °C.
- (4) Die **Proportional- oder Verhältnisrelation** setzt wie die Distanzrelation voraus, dass die Merkmalsausprägungen mit Hilfe einer Maßeinheit beschrieben werden. Damit das Verhältnis zweier Ausprägungen sinnvoll interpretierbar ist, ist ein absoluter Nullpunkt erforderlich. Nur dann sind Aussagen möglich wie z. B. beim Gewicht: 60 kg ist 1,5-mal so schwer wie 40 kg.

Die erwähnte hierarchische Ordnung der vier Relationen gibt Tabelle 2-2 wieder. An der Spitze der Hierarchie steht die Proportionalrelation. Sie schließt die Distanzrelation ein, die ihrerseits auch die Ordinalrelation und diese wiederum die Nominalrelation umfasst. Je mehr Relationen für die Ausprägungen eines Merkmals gelten, desto größer ist deren Informationsgehalt.

Tabelle 2-2: Relationen im Objektbereich

Relation (Kürzel)	eingeschlossene Relationen	Informationsgehalt
Verhältnisrelation (V)	(A) (O) (G)	hoch
Abstandsrelation (A)	(O) (G)	
Ordnungsrelation (O)	(G)	
Gleichheitsrelation (G)	keine	niedrig

Bei der Entscheidung, welche Relationen ein Merkmal erfüllt, ist die Beantwortung der folgenden Fragen hilfreich:

- Können die Ausprägungen eindeutig voneinander unterschieden werden?
(wenn ja, gilt die Gleichheitsrelation)
- Können alle Ausprägungen eindeutig in einer Rangfolge angeordnet werden?
(wenn ja, gilt die Ordnungsrelation)
- Können die Abstände zwischen den Ausprägungen in einer Einheit angegeben werden? (wenn ja, gilt die Abstandsrelation)
- Gibt es einen natürlichen Nullpunkt? (wenn ja, gilt die Verhältnisrelation)



Aufgabe 2-3: Identifizieren von Relationen im Objektbereich

In einer bundesweiten Untersuchung zum Sport- und Gesundheitsverhalten von Jugendlichen werden u. a. die folgenden Merkmale erhoben. Welche Relationen gelten jeweils?

Merkmal	Ausprägungen	Geltende Relationen?
Sozialstatus	hoch – mittel – niedrig	
Rauchen	ja – nein	
Zigarettenkonsum	Anzahl Zigaretten pro Tag	
Sportverein	Mitglied – kein Mitglied	
Dauer der Vereinsmitgliedschaft	Anzahl Monate	
Selbsteinschätzung Fitness	sehr fit – eher fit – weniger fit – nicht fit	
Wohnregion	städtisch – ländlich	
Sport in der Woche	Anzahl Stunden	

2.2 Vom Objektbereich zum Datenbereich

Der Objektbereich (Merkmalsträger, Merkmale und ihre Ausprägungen) kann nicht unmittelbar statistisch analysiert werden: Die Statistik arbeitet nicht mit den realen Ausprägungen der Merkmale von Schüler*innen, Schultaschen oder Kollegien, sondern mit deren numerischen oder symbolischen Repräsentationen. Dazu müssen – wie schon erwähnt – die Informationen über die realen Merkmalsträger bzw. die interessierenden Merkmale und die jeweiligen Ausprägungen zunächst in Daten – in der Regel Zahlen – überführt werden. Man sagt: Es ist eine Abbildung des Objektbereiches in den Datenbereich erforderlich. Um die Unterscheidung zwischen Objekt- und Datenbereich auch sprachlich zu verdeutlichen, nutzen wir unterschiedliche Begrifflichkeiten.³

Tabelle 2-3: Unterscheidung von Begriffen im Objekt- und Datenbereich

Objektbereich	Datenbereich
Merkmalsträger/Objekt	Fall
Merkmal	Variable
Mögliche Merkmalsausprägung	Variablenwert
Realisierte Merkmalsausprägung	Messwert

³ Wenn im Weiteren Variablen bzw. deren Ausprägungen beschrieben werden, so verwenden wir zur Kenntlichmachung im Text KAPITÄLCHEN bzw. setzen diese *kursiv*.

Die Übertragung des Objektbereiches in den Datenbereich wird als Messprozess bezeichnet. Dabei meint messen die regelgeleitete Zuordnung von Zahlen (Kromrey, Roose & Strübing, 2016) zu Merkmalsausprägungen, wobei (und das ist ganz wichtig) die im Objektbereich vorhandenen Relationen zwischen den Ausprägungen in den korrespondierenden Messwerten des Datenbereichs erhalten bleiben. Dieses Verständnis von Messen ist deutlich breiter gefasst als unser alltägliches, naturwissenschaftlich geprägtes Konzept des Messens, bei dem eine Merkmalsausprägung in Einheiten (z. B. Meter, Kilogramm, Sekunde) quantifiziert wird.

Zentral für naturwissenschaftliche Messungen ist der Abgleich einer empirisch möglichen Merkmalsausprägung mit einer abstrakten, per Konvention festgelegten Einheit. Zur Erfassung von Länge wird beispielsweise die Einheit Meter genutzt. Diese Einheit wurde ursprünglich als der zehnmillionste Teil des Abstands vom Pol zum Äquator definiert. Zur Längenmessung wird eine solche Einheit mit ihren Unterteilungen (z. B. Zentimeter) auf ein Messinstrument (z. B. einen Zollstock) übertragen. Mit Hilfe dieses Instruments wird die Länge eines Objektes gemessen, indem der Nullpunkt des Zollstocks an einer Seite angelegt und die Objektlänge an der anderen Seite des Objekts an den Einheiten des Zollstocks abgelesen wird. Die physikalische Längenmessung stellt also fest, wie oft eine im Vorfeld festgelegte Einheit in der Länge eines Objekts enthalten ist. Das Messergebnis wird in Form einer Zahl festgehalten. So besagt die Aussage „Der Tisch hat eine Länge von 1,50 m“, dass sich die Einheit Meter genau 1,5mal in der Länge des Tisches wiederfindet.

Die Mehrheit der Messvorgänge in den Bildungs- und Sozialwissenschaften erfolgt hingegen ohne eine solche Einheit. Statt von Messen spricht man hier häufig auch von Quantifizieren. Die Messung des Merkmals **GESCHLECHT** bedeutet in diesem Sinne einfach nur, dass den verschiedenen interessierenden Ausprägungen, die durch eine wissenschaftliche Erhebungsmethode (z. B. durch den Einsatz eines Fragebogens) erfasst werden, Zahlen zugeordnet werden (z. B. *weiblich* = 1, *männlich* = 2, *divers* = 3). Damit die resultierenden Daten auswertbar sind, muss die gewählte Zuordnungsregel für den Messvorgang innerhalb einer Untersuchung in immer gleicher Weise genutzt werden. Diese „eindeutige Zuordnung“ gewährleistet, dass in dem resultierenden Datensatz eine bestimmte Zahl durchgängig für dieselbe Merkmalsausprägung steht (z. B. 1 steht im Datensatz immer für *weiblich* und repräsentiert nie die Ausprägung *männlich*).

Die sozialwissenschaftliche Methodenlehre beschäftigt sich ausführlich mit der Frage nach der Beschaffenheit von Erhebungsinstrumenten, da die interessierenden Variablen häufig keine direkt beobachtbaren, sondern latente Variablen (sogenannte Konstrukte wie etwa Schulfreude) sind. Die spannende Frage nach der Messung (Operationalisierung) von Konstrukten kann und soll hier nicht weiter behandelt werden. Stattdessen fokussieren wir uns auf die Forderung, dass beim Messvorgang die Relationen zwischen den Merkmalsausprägungen in den Messwerten erhalten bleiben sollen.

In Kapitel 2.1 haben Sie die vier zentralen Relationen kennengelernt, mit denen sich Merkmalsausprägungen vergleichen lassen. Der Erhalt der Gleichheitsrelation durch die Messung bedeutet nun: Wenn zwei Merkmalsträger identische Merk-

merkmalsausprägungen haben, sollen auch die zugeordneten Zahlen, also die Messwerte, übereinstimmen. Die Ordnungsrelation bleibt erhalten, wenn gilt: Je „höher“ eine Merkmalsausprägung ist, desto größer⁴ soll die zugewiesene Zahl sein. Der Erhalt der Abstandsrelation ist gewährleistet, wenn Folgendes gesichert ist: Unterscheiden sich je zwei Ausprägungen in gleicher Weise, so sollen sich auch die zugewiesenen Zahlen um eine identische Zahl an Einheiten unterscheiden. Und schließlich wird die Verhältnisrelation dann durch die Messung angemessen abgebildet, wenn der Quotient zweier Merkmalsausprägungen dem Quotienten der Messwerte entspricht. Ist beispielsweise Merkmalsträger A doppelt so groß wie Merkmalsträger B, so sollten sich ihre Messwerte auch um den Faktor 2 unterscheiden (Messwert von A = 2 mal Messwert von B).

Als problematisch erweist es sich, dass in der Statistik fast immer mit Zahlen gearbeitet wird. Selbst bei Merkmalen, deren Ausprägungen nur hinsichtlich gleich vs. ungleich unterschieden werden können, werden üblicherweise Zahlen als Messwerte zugewiesen. Unser „normales“ Verständnis von Zahlen sagt, dass man diese sowohl sinnvoll auf Gleichheit überprüfen, der Größe nach anordnen als auch ihre Abstände und Verhältnisse interpretieren kann. Bei sozialwissenschaftlichen Daten dürfen die Messwerte aber nicht in dieser Vollständigkeit als Zahl interpretiert werden. Wenn Messwerte beispielsweise die verschiedenen Ausprägungen des Merkmals „Hobby“ repräsentieren, dürfen sie nur hinsichtlich Gleichheit oder Ungleichheit interpretiert werden. Oder negativ formuliert: Da bei den Ausprägungen Fußball, Volleyball und Handball des Merkmals Hobby weder die Reihenfolge noch der Abstand als Relationen gegeben sind, kann man auch die Messwerte weder in eine Reihenfolge bringen noch mit ihnen „rechnen“.⁵

Um sozialwissenschaftliche Daten sinnvoll zu interpretieren, muss im Vorfeld geklärt werden, welche Relationen durch die Messung abgebildet werden. Es ist zu klären, welches Messniveau⁶ die Variablen haben:

- Nominalskala: Die Messwerte bilden nur die Gleichheitsrelation ab;
- Ordinalskala: Die Messwerte bilden die Gleichheitsrelation und die Ordnungsrelation ab;
- Intervallskala: Die Messwerte bilden zusätzlich zur Gleichheits- und Ordnungsrelation auch die Abstandsrelation ab;
- Verhältnisskala: Die Messwerte bilden alle vier möglichen Relationen ab.

⁴ Die Reihenfolge der Symbole kann auch umgepolt werden: Die deutschen Schulnoten drücken beispielsweise die höhere, also bessere Leistung in einer kleineren Ziffer aus.

⁵ Merkmalen, für die nur die Gleichheits- oder die Ordnungsrelation gelten, könnte man im Datenbereich nichtnumerische Symbole wie etwa Buchstaben zuordnen. Dann aber werden die Messwerte unhandlich, z. B. für die Auswertung mit Hilfe von Computerprogrammen.

⁶ Als Synonym für den Begriff Messniveau wird in der Literatur der Begriff Skalen- oder Datenniveau verwendet.

2.3 Messniveaus

Durch den Messvorgang werden die Relationen zwischen den Merkmalsausprägungen aus dem Objektbereich in den Datenbereich übertragen. Je nachdem, welche Relationen im Objektbereich gegeben sind und welche dieser Relationen die Messung abbildet, resultieren Variablen mit unterschiedlichen Messniveaus. Das Messniveau einer Variablen entscheidet darüber, welche Transformationen und welche statistischen Verfahren sinnvoll auf die Daten angewendet werden können. Dies wird für die vier genannten Messniveaus im Folgenden noch einmal ausführlich dargestellt.

2.3.1 Nominalskala

Die Ausprägungen einiger Merkmale im Objektbereich können nur in Hinblick auf ihre Gleichheit verglichen werden. Beispiele dafür sind Geschlecht, Lieblingssport, Nationalität oder auch Schulform. Diese Merkmale können nur als nominalskalierte Variablen abgebildet werden. Bei der Messung wird jeder Merkmalsausprägung ein Zahlenwert eindeutig zugeordnet, so dass anhand der Zahlen entschieden werden kann, ob die Merkmalsträger gleiche oder verschiedene Merkmalsausprägungen haben. Darauf hinaus sind die Messwerte inhaltlich nicht bedeutungstragend, sie können beliebig angeordnet werden.

Beispielsweise werden die Ziffern 1, 2 und 3 in der Praxis häufig als Messwerte für die Ausprägungen männlich, weiblich und divers des Merkmals Geschlecht gewählt. Statt dieser Zuordnung hätte man die Ziffern auch vertauschen oder die Werte -30, 22 und 2.250 wählen können. Solange eine Zuordnung eindeutig ist, bilden die Messwerte die Gleichheit der Merkmalsausprägungen strukturtreu ab und sind formal für die Messung des Geschlechts geeignet.

Falls erforderlich können die gewählten Zahlen also beliebig umgewandelt oder transformiert werden, solange die Transformation eindeutig ist und die Äquivalenzrelation erhalten bleibt. Wenn in einer Untersuchung für *männlich*, *weiblich* und *divers* die Zahlen 1, 2 und 3, in einer anderen Erhebung aber die Zahlen 0, 1 und 2 gewählt wurden, dann ist es problemlos möglich, die Daten zusammenzuführen. Dazu können beispielsweise die Messwerte aus der zweiten Untersuchung gemäß der Zuordnung der ersten Untersuchung umcodiert werden ($0 \Rightarrow 1$, $1 \Rightarrow 2$, $2 \Rightarrow 3$).

Andererseits beschränkt das Nominalskalenniveau die Analysemöglichkeiten. Die „Verrechnung“ von Messwerten einer nominalskalierten Variablen ergibt keinen Sinn, da das Ergebnis inhaltlich nicht interpretiert werden kann. So hat die Summe aus den Messwerten von 50 Jungen (Messwert 0) und 75 Mädchen (Messwert 1) und 3 Kindern mit diverser Geschlechtszugehörigkeit (Messwert 2) ebenso wenig eine inhaltliche Bedeutung wie deren Mittelwert. Die statistische Analyse eines nominalen Merkmals beschränkt sich auf das (Aus-)Zählen, wie häufig die verschiedenen Ausprägungen vorkommen.

2.3.2 Ordinalskala

Bei Merkmalen, die neben der Gleichheitsrelation auch die Ordnungsrelation erfüllen, sind die Ausprägungen gemäß ihrer Größe, Qualität, Richtigkeit oder Ähnlichem angeordnet. Den Ausprägungen des Merkmals liegt eine Rangreihe im Sinne eines „mehr oder weniger“ zugrunde. Diese Beziehung wird durch das Messniveau einer Ordinalskala abgebildet. Ein klassisches Beispiel für eine Ordinalskala ist die Schulnotenskala. Einer „weit über den Anforderungen“ liegenden Leistung wird die Note 1 zugeordnet, einer besonders schwachen Leistung die Note 6. Auch das in den Sozialwissenschaften beliebte Instrument „Ratingsskala“ fragt die Einschätzung – beispielsweise die Zustimmung – zu Aussagen (Items) ab. So könnte die Zustimmung zu der Aussage „Ich gehe gerne zur Schule“ mit Hilfe einer vierstufigen Ratingsskala mit den Antwortalternativen *stimme gar nicht zu* (= 1), *stimme eher nicht zu* (= 2), *stimme eher zu* (= 3) und *stimme genau zu* (= 4) erfasst werden.

Die Werte auf dem Messniveau einer Ordinalskala können nach ihrer Größe sortiert werden. Bei einer Ratingsskala repräsentieren höhere Messwerte beispielsweise eine höhere Zustimmung zu der Aussage. Allerdings sind die zugeordneten Zahlen beliebig, solange die Zuordnung die gleiche Rangfolge ergibt. So könnte auch die Zuordnung *stimme gar nicht zu* = -10, *stimme eher nicht zu* = 0, *stimme eher zu* = 100,45 und *stimme genau zu* = 9999 gewählt werden. Auch diese Zuordnung gewährleistet, dass höhere Messwerte eine stärkere Zustimmung signalisieren. In der Praxis werden aufeinanderfolgende, ganze Zahlen als Messwerte verwendet. Aber das ist eine Frage der Konvention und Vereinfachung ohne inhaltliche Bedeutung.

Im Vergleich zu nominalen Variablen sind die Transformationsmöglichkeiten von ordinalen Variablen eingeschränkt, da neben der Äquivalenz- auch die Ordnungsrelation erhalten bleiben muss. Damit sind nur Transformationen erlaubt, bei denen die Reihenfolge der Ausprägungen erhalten bleibt. Ein praktisches Beispiel für Transformationen ordinaler Variablen ist die Umrechnung der in Deutschland verwendeten Notenskala in die Notensysteme anderer Länder (vgl. Tabelle 2-4). Die Rangfolge der Notenstufen bleibt gleich, auch wenn die Polung der Skala in Deutschland und Österreich anders ist als bei den niederländischen und bei den polnischen Notenskalen. Das norwegische Schulsystem verwendet keine Ziffernoten, sondern notiert die Leistungsstufen mit Buchstaben, die verdeutlichen, dass Schulnoten geordnete Kategorien sind, mit denen man eigentlich nicht rechnen kann.

Tabelle 2-4: Vergleich europäischer Schulnoten

Land	Notenstufe						
	aus-gezeichnet	sehr gut	gut	befriedi-gend	ausrei-chend	mangel-haft	ungenü-gend
Deutschland		1	2	3	4	5	6
Österreich		1	2	3	4	5	
Niederlande	10	9–8,5	8	7	6	5–4	3–1
Polen	6	5	4	3	2		1
Norwegen		A	B; C	D	E	F	

Für die statistische Auswertung ordinalskalierter Variablen ergeben sich mehr Möglichkeiten als für nominalskalierte Variablen. Die Messwerte dürfen immer noch nicht verrechnet werden, aber es sind Aussagen über die relative Stellung von Fällen mit einer bestimmten Ausprägung möglich. Haben in einer Stichprobe von 180 Schülerinnen und Schülern im Fach Deutsch zehn die Note *sehr gut* und zwanzig die Note *gut*, dann gehört eine Schülerin mit der Note *gut* zum leistungsstärksten Sechstel dieser Stichprobe.

Aufgabe 2-4: Verständnis der Begriffe „Merkmal“ und „Variable“

Was ist der Unterschied zwischen den Begriffen „Merkmal“ und „Variable“? Versuchen Sie zu erklären, warum bei nominalen Variablen die Angabe der relativen Position nicht möglich ist. Warum ist es mathematisch nicht zulässig, Schulnoten zu verrechnen (z. B. einen Durchschnittswert zu bilden)?



2.3.3 Intervallskala

Bei Merkmalen, für die neben Gleichheits- und Ordnungsrelation auch die Abstandsrelation erfüllt ist, können Unterschiede zwischen den Ausprägungen verglichen und numerisch interpretiert werden. Die Interpretation der Abstände wird durch die Messung auf dem Niveau einer Intervallskala möglich, die für die Zuweisung von Messwerten ein Messinstrument mit vorab definierten „Einheiten“ nutzt. Beachten Sie, dass die Maßeinheit nicht im Objektbereich gegeben sein muss. Sie ist eine Konvention, um die im Objektbereich vorhandene Distanzrelation adäquat in den Datenbereich zu übertragen. So sind physikalisch vorhandene Temperaturunterschiede zwar real existent, aber erst durch die Nutzung eines Thermometers werden sie in Grad Celsius (oder Fahrenheit) quantifiziert.

Um die Abstandsrelation strukturtreu abzubilden, muss Folgendes gewährleistet sein: Wenn je zwei Ausprägungen den gleichen Abstand voneinander haben, müssen