

# Statistische Methoden der Datenaufbereitung und -auswertung

## Organisatorisches und Einführung

# Erwartungen (und Befürchtungen)

# Organisatorisches: Spezielle Arbeitsleistung

Bei umfangreichen Datensätzen

- ▶ Aufbereitung und Auswertung (eines Teils) des in der Gruppe gewählten Forschungsdatensatzes
- ▶ Präsentation der Ergebnisse aus der Gruppenarbeit am 2. Tag
- ▶ Hochladen **eines gemeinsamen** Jupyter Notebooks pro Gruppe mit der vollständigen Datenaufbereitung und -auswertung der Gruppe.

# Organisatorisches: Erreichbarkeit und Konsultationstermin

- ▶ Sprechstunde: nach Vereinbarung
- ▶ Erreichbarkeit:
  - ▷ Virtuell über [Zoom](#)
  - ▷ Telefonisch: 0331 580 1538
  - ▷ E-Mail: [julia.struss@fh-potsdam.de](mailto:julia.struss@fh-potsdam.de)
- ▶ Konsultationstermin nach der Präsenz:  
16.05.2023, 20:00 Uhr

# Lernziele

- ▶ Kennen quantitativer Erhebungsmethoden und verstehen der Einsatzbereiche und Grenzen, beurteilen von Studien und der dort eingesetzten Methoden
- ▶ Daten empirischer Untersuchungen für anschließende Auswertung aufbereiten können
- ▶ Auswerten empirischer Erhebungen mit statistischen Methoden (Fokus: deskriptiv)
- ▶ Probleme und Herausforderung bei der Arbeit mit Forschungsdatensätzen erkennen und für die spätere berufliche Praxis reflektieren können

# Themenübersicht

- ▶ statistische Grundlagen
- ▶ Erhebungsmethoden mit Fokus auf quantitative Methoden
- ▶ Stichprobenziehung / Repräsentativität
- ▶ Fragebogengestaltung
- ▶ Auswertung mit R: deskriptive Statistik
- ▶ Grafische Darstellung von Ergebnissen
- ▶ Ausblick auf: Testverfahren, schließende Statistik, multivariate Verfahren

# Was ist Statistik?

„Statistik ist die Kunst, Daten zu gewinnen, darzustellen, zu analysieren und zu interpretieren, um zu neuem Wissen zu gelangen“ (Sachs/Hedderich 2009,1).

# Warum brauchen wir Statistik?

„Statistik ist ein wichtiger Bestandteil empirisch-wissenschaftlichen Arbeitens. Statistik beschränkt sich nicht nur auf Zusammenfassung und Darstellung von Daten (dies ist Aufgabe der deskriptiven Statistik [...]), sondern sie ermöglicht empirischen Wissenschaften objektive Entscheidungen über die Brauchbarkeit der überprüften Hypothesen. Dieser Teilaspekt der Statistik, der sich mit der Überprüfung von Hypothesen befasst, wird häufig als analytische Statistik oder Inferenz- (schließende) Statistik bezeichnet.“ (Bortz 2005, 1)

# Deskriptive Statistik (Beschreibende Statistik)

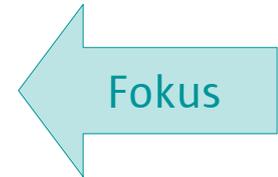
- ▶ Beschreibung von Zuständen und Vorgängen
- ▶ mit Hilfe von
  - ▷ Tabellen,
  - ▷ graphischen Darstellungen,
  - ▷ Verhältniszahlen (z.B. relative Häufigkeit)
  - ▷ typischen Kenngrößen wie Lagemaße (z.B. arithmetischer Mittelwert)
  - ▷ Streuungsmaßen (z.B. Varianz und Standardabweichung)

# Induktive Statistik (Schließende Statistik)

- ▶ auf Grundlage geeigneter Daten auf allgemeine Gesetzmäßigkeiten schließen
- ▶ basiert auf Wahrscheinlichkeitsrechnung
- ▶ (zufallsbedingte Experimente oder stochastische Experimente (z.B. Werfen eines Würfels))
- ▶ untersucht werden Stichproben
- ▶ geprüft werden Hypothesen

# Pardigmen

- ▶ Quantitatives Paradigma (z.B. Kritischer Rationalismus)
- ▶ Qualitatives Paradigma (z.B. Sozialkonstruktivismus)
- ▶ Mixed-Methods-Paradigma (z.B. Pragmatismus)



vgl. Döring/Bortz 2016, S. 32

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ Auswertungsphase
- ▶ Entscheidungsphase

## Ziel

- Einordnen der eigenen Fragestellung in theoretischen Rahmen
- Untersuchungsstatus festlegen (Hypothesen prüfen/erkunden)

vgl. Bortz 2005, S. 2

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ Auswertungsphase
- ▶ Entscheidungsphase

Ziel:  
Prüfen der untersuchenden  
Theorie/Hypothese auf  
Präzision, Informationsgehalt,  
logische Konsistenz und  
Vereinbarkeit, empirische  
Überprüfbarkeit

vgl. Bortz 2005, S. 2

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ Auswertungsphase
- ▶ Entscheidungsphase

Ziel:  
Aufbau und Ablauf der  
Untersuchung festlegen  
(Variablen auswählen und  
klassifizieren, Labor- vs.  
Felduntersuchung,..., Planung  
der statistischen Auswertung)

vgl. Bortz 2005, S. 2

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ Auswertungsphase
- ▶ Entscheidungsphase

Ziel:  
Erheben der benötigten Daten  
für die Überprüfung der Theorie

vgl. Bortz 2005, S. 2

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ **Auswertungsphase**
- ▶ Entscheidungsphase

Ziel:  
Anwenden statistischer  
Methoden zur Untersuchung der  
aufgestellten Theorie.

vgl. Bortz 2005, S. 2

# Phasen der empirischen Forschung

- ▶ Erkundungsphase
- ▶ Theoretische Phase
- ▶ Planungsphase
- ▶ Untersuchungsphase
- ▶ Auswertungsphase
- ▶ Entscheidungsphase

Ergebnisse der Auswertung in Bezug auf die Theorie interpretieren.

vgl. Bortz 2005, S. 2

# Grundbegriffe und ihre Bedeutungen: Methoden

- ▶ Systeme von Handlungsanweisungen und Regeln für den Erkenntnisgewinn/das Sammeln von Informationen
- ▶ Grundmethoden der Datenerhebung in der Sozialforschung: Befragungen, Beobachtungen, Inhaltsanalysen
- ▶ Qualitative Methoden: Ermitteln und Analysieren individuellen Handelns
- ▶ Quantitative Methoden: Suche nach verallgemeinerbaren Aussagen

(vgl. Häder 2015, S. 13)

## Art der Studie

### Empirische Studie

Untersuchung inhaltlicher Forschungsfragen

#### Originalstudie

Eigenes Untersuchungsdesign

#### Replikationsstudie

Replikation des Untersuchungsdesigns einer Originalstudie

### Methodenstudie

Untersuchung methodischer Fragestellungen

### Theoretische Studie

Untersuchung des bisherigen Forschungsstands

#### Forschungsüberblicksstudie

(z.B. Systematic Review)

#### Metaanalyse

Statistische Auswertung von Studienergebnissen zu einem Forschungsthema

## Datenerhebung

### Befragung

**Schriftliche vs. mündliche Befragung**  
(z.B. Fragebogen, Interview)

#### Individuen vs. Gruppen

(z.B. Online-Befragung, Delphi-Befragung)

#### Grad der Strukturierung

(z.B. Vollstrukturierte schriftliche Befragung, halbstrukturierte mündliche Befragung (bspw. Leitfadeninterview))

### Beobachtung

**Teilnehmende vs. nicht-teilnehmende Beobachtung, indirekte Beobachtung**  
(z.B. Verhaltensspuren auf Webseiten: Logfile-Analysen)

#### Einsatz technischer Messinstrumente

(z.B. Eye-Tracking-Geräte)

#### Beobachtungsort

(z.B. Beobachtungen im Usability Labor, Feldbeobachtung)

### Dokumentenanalyse

**Grad der Strukturierung des Vorgehens**

#### Materialarten

(z.B. textuelle Materialien, audiovisuelle Materialien)

## Datenmaterial

### Qualitative Daten

#### Befragung

(z.B. Freitextantworten aus offenen Fragen im Fragebogen, Mitschriften oder Transkripte von Interviews)

#### Beobachtung

(z.B. Beobachtungsprotokolle, Audio- und Videoaufzeichnungen)

#### Dokumentenanalyse

(z.B. Blogbeiträge, Zeitungsartikel, Kundenrezensionen wissenschaftliche Artikel)

### Quantitative Daten

#### Befragung

(z.B. Angaben geschlossener Fragebogenitems)

#### Beobachtung

(z.B. Ausgefüllter, vollstandardisierter Beobachtungsbogen)

#### Dokumentenanalyse

Nach Transformation mit quantitativer Inhaltsanalyse Messwerte formaler und inhaltlicher Merkmale

## Datenanalyse

### Qualitative Datenanalyse

#### Qualitative Inhaltsanalyse

Verbale Kodierung der Bedeutung von Dokumenten (induktiv oder deduktiv)

#### Kodierung gemäß Grounded Theory-Methodology

Kodiermethode des permanenten Vergleichs mit offenem, axialem und selektivem Kodieren

#### Diskursanalyse

(z.B. kritische Diskursanalyse)

### Quantitative Datenanalyse

#### Deskriptive statistische Analyse

(z.B. Mittelwerte, Varianzmaße, Korrelationen)

#### Inferenzstatistische Analyse

(z.B. t-Test, Varianzanalyse, Regression)

Basierend auf Ausführungen von Döring & Bortz 2016, Kap. 7 & 10, Idee in Anlehnung an Abb. 10.39, S. 535

# Grundbegriffe und ihre Bedeutungen: Merkmale und Merkmalsträger

- ▶ Merkmale:  
betrachtete Charakteristika der  
Untersuchungsobjekte (Merkmalsträger)
- ▶ Merkmalsausprägung:
  - ▷ Wert oder Eigenschaft, die ein Merkmal  
annehmen kann
  - ▷ Ausprägungen müssen überschneidungsfrei  
und erschöpfend sein

(vgl. Sachs/Hedderich 2009, S. 12; Häder 2015, S. 6)

# Beispiel: Merkmale, -träger und -ausprägung

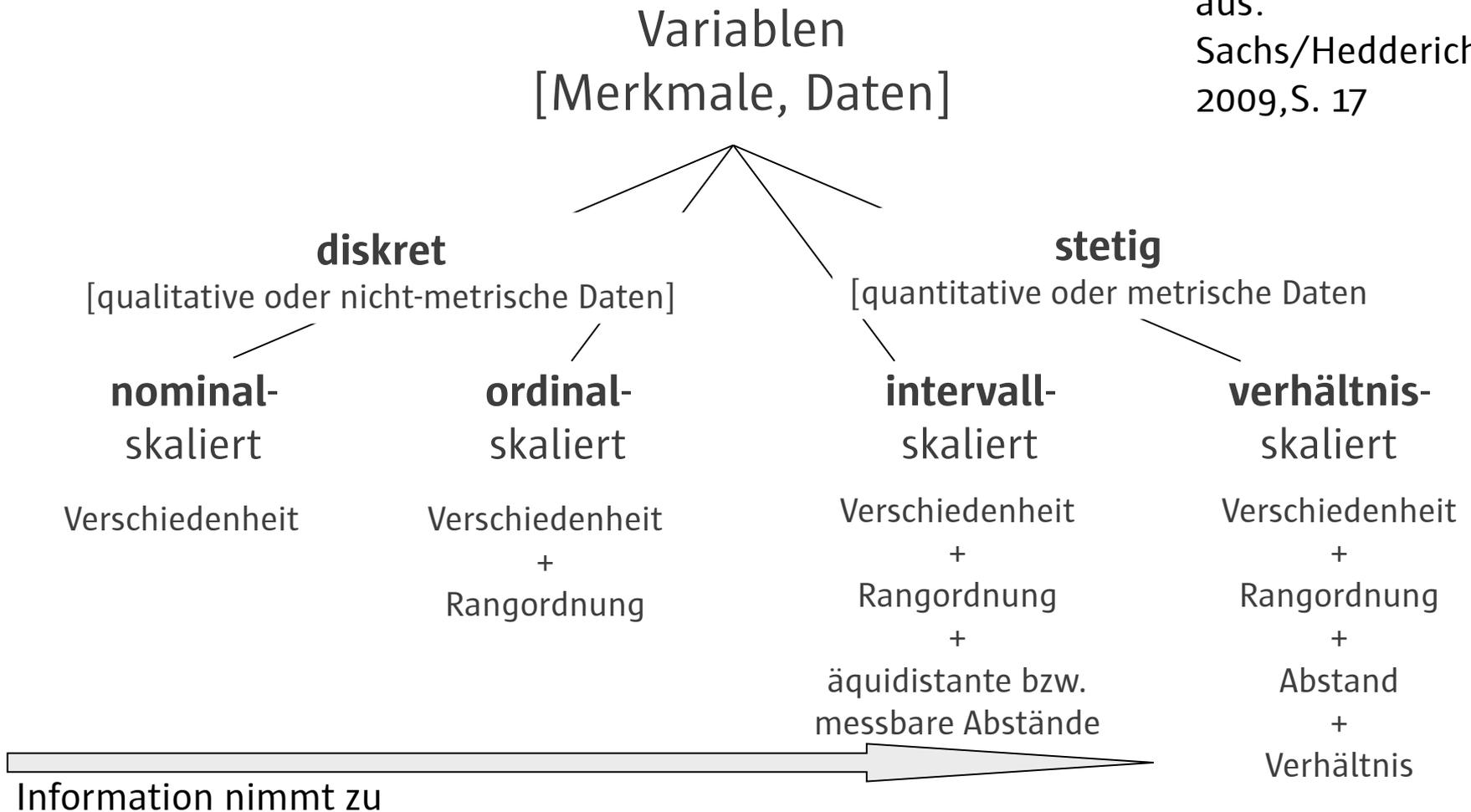
**Merkmal**  
Augenfarbe

bei **Merkmalsträger**  
Menschen, z.B.

**Merkmals-  
ausprägung**  
blau

# Skalenniveau von Variablen/Merkmalen

aus:  
Sachs/Hedderich  
2009, S. 17



# Übung

- ▶ Bestimmen Sie das Skalenniveau, mit dem die Variablen bei den einzelnen Items im Fragebogen Ihres ausgewählten Forschungsdatensatzes erhoben werden.
- ▶ Identifizieren Sie Variablen, bei denen die Erhebung auch mit einem höheren Skalenniveau möglich gewesen wäre.

# Grundbegriffe und ihre Bedeutungen: Gütekriterien

- ▶ Objektivität: Sind die Ergebnisse unabhängig von den durchführenden Personen?
- ▶ Reliabilität: Sind die Ergebnisse reproduzierbar?
- ▶ Validität: Misst das Testinstrument das, was es vorgibt zu messen?
  - ▷ Konstruktvalidität
  - ▷ Interne Validität
  - ▷ Externe Validität
  - ▷ Statistische Validität

# Grundbegriffe und ihre Bedeutungen: Stichproben

- ▶ Teilmenge der Grundgesamtheit/Population
- ▶ je besser die Stichprobe die Population repräsentiert (hinsichtlich der bedeutsamen Merkmale), desto präziser die darauf basierenden inferenzstatistischen Aussagen

(vgl. Bortz & Schuster 2010, S. 79ff.)

- ▶ Repräsentativität einer Stichprobe:  
*„[...] gibt an, wie gut bzw. unverzerrt die Merkmalszusammensetzung in der Stichprobe die Merkmalszusammensetzung in der Population widerspiegelt. Zu unterscheiden sind dabei merkmalspezifisch-repräsentative sowie global-repräsentative Stichproben.“*

Döring & Bortz 2016, S. 298

# Verschiedene Arten der Stichprobenziehung

- ▶ Probabilistische Stichproben
  - ▷ einfache Zufallsstichprobe
  - ▷ Klumpenstichprobe
  - ▷ Geschichtete Stichprobe
- ▶ nicht-probabilistische Stichproben
  - ▷ Gelegenheitsstichprobe/ Ad-hoc Stichprobe
  - ▷ Quotenstichprobe

Jedem Untersuchungsobjekt kann eine Wahrscheinlichkeit zugeordnet werden, mit der es Teil der Stichprobe wird.

Nicht jedes Untersuchungsobjekt hat eine Chance in die Stichprobe zu kommen.

# Übung

- ▶ Skizzieren Sie die Details der Stichprobenziehung, die bei der Erhebung zu dem von Ihnen gewählten Forschungsdatensatz zum Einsatz kamen.
- ▶ Welche Implikationen bringt die Methode der Stichprobenziehung für die spätere Auswertung bzw. den Umgang mit den Ergebnissen mit sich?

# Literatur

- ▶ Bortz, Jürgen (2005): Statistik für Human- und Sozialwissenschaftler. 6. vollst. überarb. u. akt. Aufl., Heidelberg: Springer.
- ▶ Döring, Nicola; Bortz, Jürgen (2016): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Berlin, Heidelberg: Springer Berlin Heidelberg.
- ▶ Field, Miles, Field (2012): Discovering Statistics Using R. London: Sage.
- ▶ Häder, Michael (2015): Empirische Sozialforschung. Wiesbaden: Springer Fachmedien Wiesbaden.
- ▶ Sachs, Lothar; Hedderich, Jürgen (2009): Angewandte Statistik. Methodensammlung mit R. 13. akt. u. erw. Aufl. Berlin [u.a.]: Springer.