

WIRTSCHAFTSSTATISTIK

MODUL 2: SKALEN UND KLASSIERUNG

WS 2020/21

DR. E. MERINS

MESSBARKEIT

- **Informationsbedarf → empirische (statistische) Untersuchung**

Bei einer empirischen Untersuchung messen wir Merkmale bei ausgewählten Untersuchungseinheiten mit einem Messinstrument auf einer Skala.

Ergebnis: Messwerte = Merkmalswerte = Beobachtungswerte

Wir messen bei Kind und seiner Mutter das Merkmal Körpergröße mit einem cm-Maß auf einer cm-Skala.

Messergebnisse:

Kind: 121 cm, Mutter: 168 cm.



GRUNDBEGRIFFE DER STATISTIK

Grundbegriffe der Statistik

Merkmalsträger	Einzelnes Objekt einer statistischen Untersuchung, Träger der Informationen, für die man sich interessiert. → Untersuchungseinheit → Erhebungseinheit → Unit
Statistische Masse	Menge aller Merkmalsträger, die <ul style="list-style-type: none">• mit dem Untersuchungsziel in Verbindung stehen,• unter sich mindestens eine übereinstimmende Eigenschaft haben,• sich exakt abgrenzen lassen, und zwar<ul style="list-style-type: none">- sachlich- räumlich- zeitlich → Kollektiv, Grundgesamtheit, Population Beispiele: Bevölkerung des Landes, Automobilproduktion
Merkmal	Im Rahmen der statistischen Erhebung relevante Eigenschaften der Merkmalsträger → Statistische Variable
Merkmalsausprägung	Grundsätzlich mögliche Ausformungen eines Merkmals → Wert der Variable, Beobachtungswert

MERKMALE

Merkmalstypen

Qualitative Merkmale

Merkmale lassen sich nicht mit Zahlen messen (Codierung und Rangordnung möglich)
z.B. Geschlecht, Güteklasse



diskrete Merkmale

Qualitative Merkmale sind immer diskret, da sie von Natur aus nur eine abzählbare Menge möglicher Merkmalswerte haben

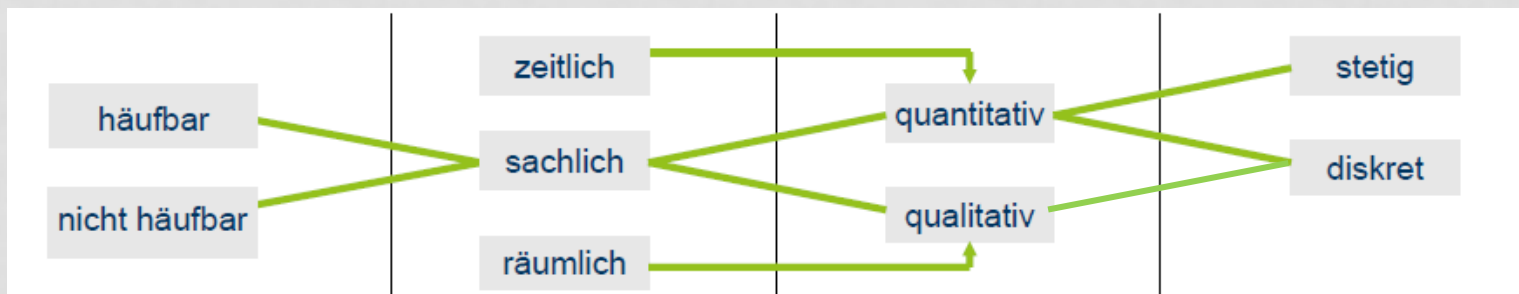
Quantitative = metrische Merkmale

stetige Merkmale

Menge der Merkmalsausprägungen überabzählbar, Intervall der reellen Zahlen (es gibt zwischen zwei Ausprägungen immer noch weitere Zwischenwerte)
z.B. Gewicht, Alter, Fahrzeit

diskrete Merkmale

Menge der Merkmalsausprägungen endlich bzw. abzählbar (i.d.R. ganze Zahlen)
z.B. Kinderzahl, Sitzplätze, das monatliche Gehalt



SKALENNIVEAU

Nach der Art des Merkmals richtet sich, auf welche Weise die Beobachtungswerte bei der statistischen Untersuchung gemessen werden können (Messung = Eindeutige Zuordnung einer Beobachtung zu einem Punkt auf einer Messskala)

Vom **Skalenniveau** hängt auch ab, welche Rechenoperationen mit den Beobachtungswerten und welche statistischen Auswertungsmethoden zulässig sind.

Man unterscheidet folgende **Skalenniveaus**:

I. Nicht metrische Skalen → Anwendung bei qualitativen Merkmalen. Keine Rechenoperationen mit den Merkmalsausprägungen zulässig:

- **Nominalskala**
- **Ordinalskala**

II. Metrische Skalen (Kardinalskalen) → Anwendung bei quantitativen Merkmalen.

Skala hat Nullpunkt und Maßeinheit. Rechenoperationen sind zulässig:

- **Intervallskala**
- **Verhältnisskala** (Ratioskala)
- **Absolutskala**

SKALIERUNG

Skalenart	Besonderheiten	zulässige Operationen	Beispiel für Merkmale	Beispiel für Operationen
Nominalskala	Daten haben nur eine endliche Menge von Ausprägungen, unterliegen keiner Rangfolge und sind nicht vergleichbar. Zuordnung von Zahlen ist lediglich eine Kodierung der Merkmalsausprägungen	$=, \neq$	Geschlecht, Familienstand, Steuerklasse, PLZ	Geschlecht von Claudia \neq Geschlecht von Peter
Ordinalskala = Rangskala	Daten haben nur eine endliche Menge von Ausprägungen, können in eine natürliche Rangfolge gebracht werden. Ordnungsprinzip ist die Stärke bzw. der Grad der Intensität, man kann hier allerdings keine Abstände zwischen den einzelnen Ausprägungen interpretieren	$=, \neq, <, >$	Konfektionsgröße, Schulnoten, Windstärke	XXL > XL > L > M > S > XS
Intervallskala	Besitzt <u>keinen</u> natürlichen Nullpunkt, keine Verhältnisse können gebildet werden. Daten können alle (<i>unendlich viele</i>) Ausprägungen innerhalb eines Intervalls annehmen.	$=, \neq, <, >, +, -$	Längendifferenzen, IQ, Temperatur in Celsius	morgen wird es 10 Grad kälter als heute
Verhältnisskala = Ratioskala	Besitzt natürlichen Nullpunkt Quotienten (das Verhältnis) gemessener Werte werden verglichen	$=, \neq, <, >, +, -, \times, /$	Umsatz, Körpergröße, Einkommen, Temperatur in Kelvin	Der Umsatz ist um 7% gegenüber dem Vorjahr gestiegen oder doppelt so hoch wie...
Absolutskala	Ausprägungen absolut skaliert Merkmale sind Anzahlen und Stückzahlen. Allgemein: Häufigkeiten oder alles, was man zählen kann	$=, \neq, <, >, +, -, \times, /$	Zahl der Beschäftigten	150 Beschäftigte sind 3 mal so viel wie 50 Beschäftigte

SKALIERUNG, BEISPIELE

Merkmal	Menge der Merkmalsausprägungen	Messinstrument	Skala	Merkmalstyp
Familienstand	{ledig, verheiratet, verwitwet, geschieden}	Frage	Nominalskala	qualitatives Merkmal
Hotelgütekategorie	{*****, ****, ***, **, *, }	Fragebogen	Rangskala = Ordinalskala	qualitative Merkmale = Rangmerkmale
Klausurnote	{1,0 1,3 1,7 2,0 2,3 2,7 3,0 3,3 3,7 4,0 5,0}	Klausur		
Temperatur (°C)	\mathbb{R}	Thermometer	Metrische Skala = Intervallskala	Quantitative Merkmale = metrische Merkmale
Körpergröße	$\{x \mid x \in \mathbb{R} \text{ und } x > 0\}$	cm-Maß	Metrische Skala = Verhältnisskala	
Kinderzahl	$\mathbb{N} \cup \{0\}$	Frage	Metrische Skala = Absolutskala	

KLASSIERUNG BEI QUALITATIVEN MERKMALEN

Beispiel:

Frage: Beruf?

Antworten:

- Maurer
- Dachdecker
- Schreiner
- Fliesenleger

→ Berufsgruppe: Handwerker

Zielkonflikt: Übersichtlichkeit versus Informationsverlust

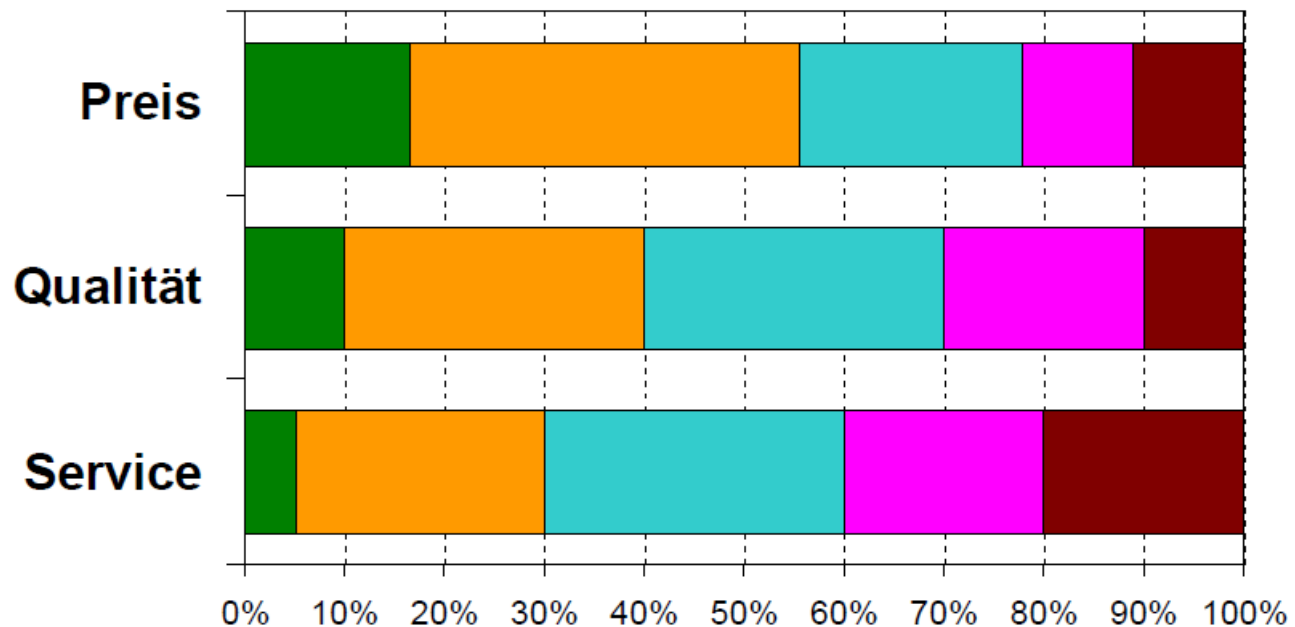
KLASSIERUNG BEI RANGMERKMALEN

Beispiel:

Frage: wie zufrieden sind Sie mit einem Produkt bzgl. ...?

Antworten:

- | | |
|---------------------------------|-----------------|
| ■ 1 sehr zufrieden | ■ 2 zufrieden |
| ■ 3 mehr oder weniger zufrieden | ■ 4 unzufrieden |
| ■ 5 sehr unzufrieden | |

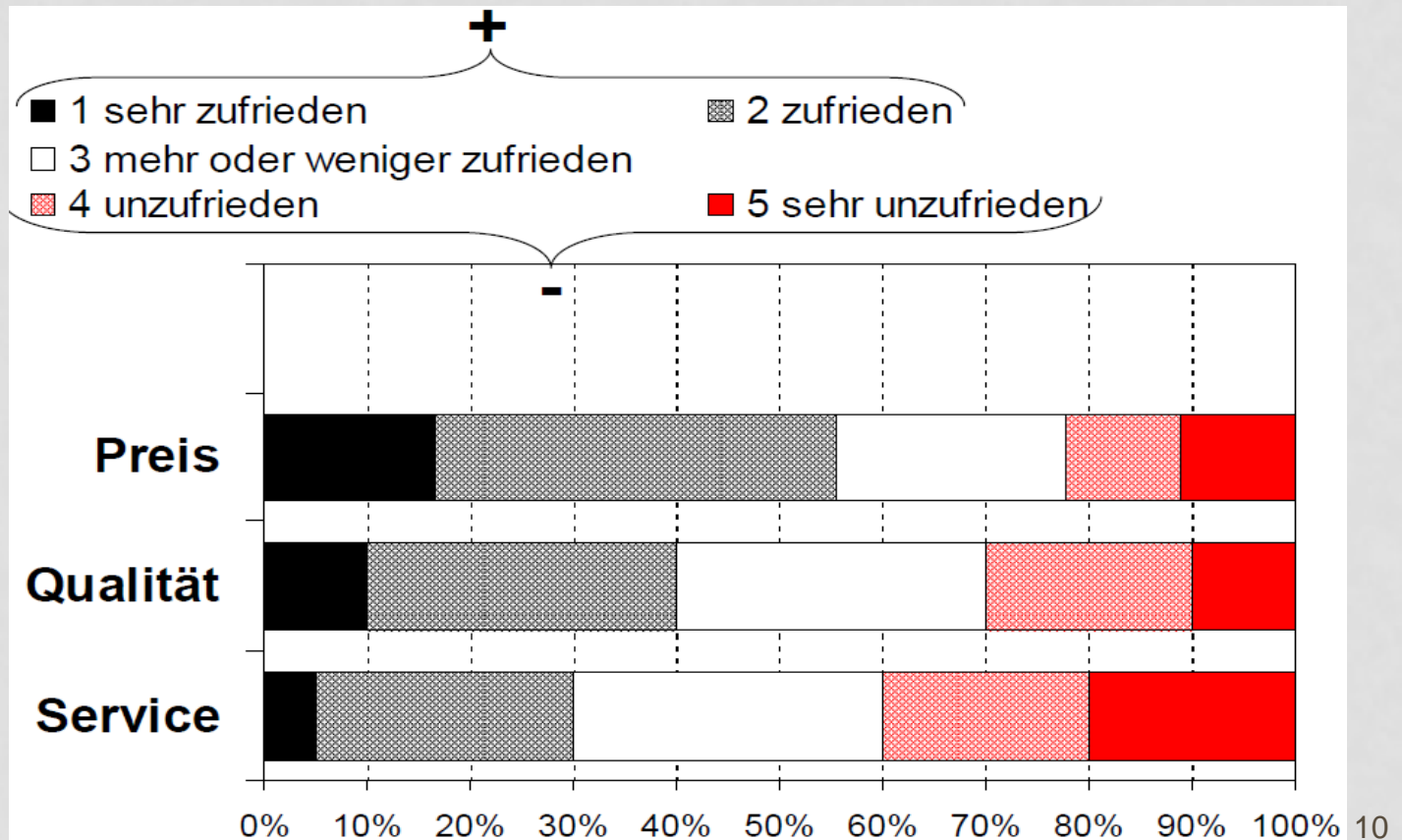


KLASSIERUNG BEI RANGMERKMALEN

Beispiel:

Frage: wie zufrieden sind Sie mit einem Produkt bzgl. ...?

Antworten:

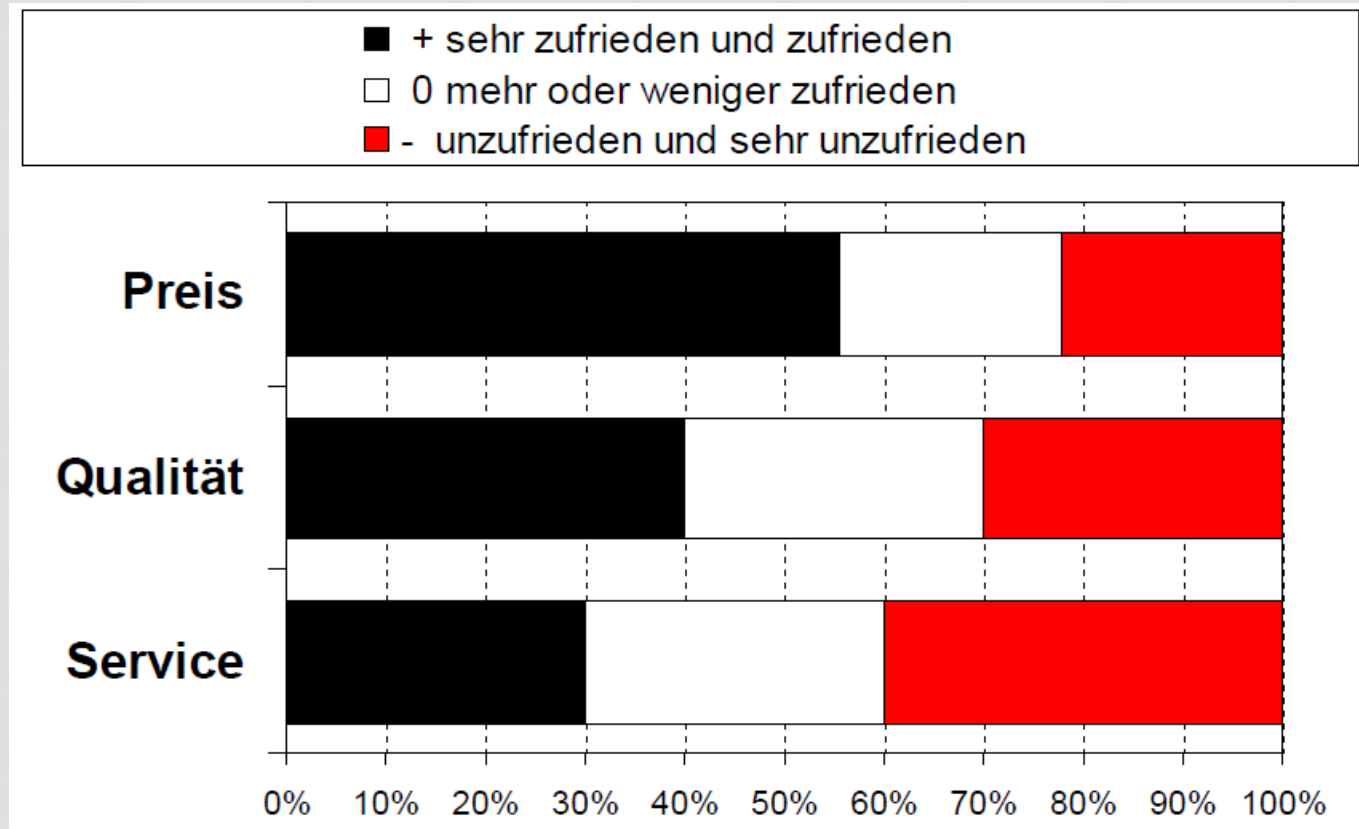


KLASSIERUNG BEI RANGMERKMALEN

Beispiel:

Frage: wie zufrieden sind Sie mit einem Produkt bzgl. ...?

Antworten:



KLASSIERUNG BEI METRISCHEN MERKMALEN

Metrische Merkmale (vgl. Folie 4)

diskret

z.B. Einwohnerzahl

Klassierung

1. 0 – 19.999
 2. 20.000 – 49.999
 3. 50.000 – 99.999
 4. 100.000 – 249.999
- usw.

stetig

z.B. Körpergröße

Klassierung

1. 0 – 99 cm
 2. 100 – 139 cm
 3. 140 – 159 cm
 4. 160 – 169 cm
- usw.

wg. Lücken

Klassierung

1. 0 – 100 cm
 2. 100 – 140 cm
 3. 140 – 160 cm
 4. 160 – 170 cm
- usw.

wg. Überschneidungen

Klassierung

1. 0 bis unter 100 cm
 2. 100 b.u. 140 cm
 3. 140 b.u. 160 cm
 4. 160 b.u. 170 cm
- usw.

richtig !!!

$$100 \text{ b.u. } 140 \text{ cm} = \{x \mid x \in \mathbb{R}, 100 \leq x < 140 \text{ cm}\}$$

ENTSCHEIDUNGEN BEI KLASSIERUNG

- **Klassenbreite(n)**

→ alle gleich oder unterschiedlich

- **Anzahl der Klassen**

- **Klassengrenzen**

→ untere Klassengrenzen, obere Klassengrenzen

- **untere/obere offene Randklasse?**

→ „bis unter 50 kg“ bzw. „120 kg und schwerer“

„OPERATIONALISIERUNG“ EINES BEGRIFFS

"Operationalisierung" eines Begriffs ist die Angabe derjenigen Vorgehensweisen und Forschungs**operationen**, mit deren Hilfe zu entscheiden ist, ob und in welchem Ausmaß der mit dem Begriff bezeichnete Sachverhalt in der Realität vorliegt, was bedeutet, dass man beobachtbare Kriterien dafür anzugeben hat, wann ein Sachverhalt vorliegt bzw. je nach Skalenniveau auch in welcher Ausprägung er auftritt.

Etwas weniger abstrakt:

„Operationalisierung“ definiert, wie man den Begriff konkret misst. Die Operationalisierung ist besonders wichtig bei Begriffen ohne direkten empirischen Bezug (so genannte „latente Variable“), z.B. Kundenzufriedenheit, Teamfähigkeit, Intelligenz, Werbewirkung u.a.

Der Ausdruck Operationalisierung bezeichnet im weitesten Sinne die Entwicklung eines Forschungsdesigns für eine konkrete Fragestellung, während es im engeren Sinne um die Formulierung von Messvorschriften geht, d.h., um die Bestimmung von Indikatoren, mit deren Hilfe ein Konstrukt gemessen werden kann.

→ die Festlegung der Vorgehensweise (Operation) bei der Definition der Untersuchungsvariablen in einer Untersuchung.

Beispiel: Intelligenz kann operational durch die Anzahl der Lösungen von Intelligenzaufgaben in einem konkreten Intelligenztest definiert werden.

„OPERATIONALISIERUNG“ EINES BEGRIFFS

Beispiel: (Quelle der Geschichte: Becker, B.: Statistik, S. 40 und 79)

Ein Mitarbeiter des Statistischen Bundesamtes wird eines Tages von seiner kleinen Tochter gefragt, warum er so viel Geld verdient, wo er doch nur Zahlen addiert. Der Vater lächelt und sagt: „Mein Liebes, geh´ hinaus in den Garten und zähle die Bäume“. Nach einigen Minuten kommt die Tochter zurück, aufgelöst in Tränen: „Papi, ich kann die Bäume nicht zählen, denn da sind große Bäume, kleine Bäume, Büsche, einige haben Nadeln, andere Blätter, einige haben einen Stamm, andere haben 2 Stämme...“. Ihr Vater lächelt wieder und nimmt sie sanft auf seinen Arm: „Siehst du, das ist es, wofür ich bezahlt werde, ich zähle nicht einfach nur, ich mache Entscheidungen, **was ein Baum ist, wie man Bäume zählt** usw.

Wenn ein Statistiker mit der Arbeit beginnt, muss er geklärt haben, was er überhaupt quantifizieren will. Will er z.B. Bäume zählen, so muss er wissen, was mit einem Baum gemeint ist, wie man also von einem mehr oder weniger abstrakten Begriff zu einer „operationalisierbaren“ bzw. beobachtbaren Kategorie kommt. Es muss also gesagt werden, ob man z.B. nur Laubbäume meint oder auch Nadelbäume eingeschlossen sein sollen, ob man jeden Baum unabhängig von seiner Größe zählen soll usw..