

2-1

STATISTIK FÜR ERZIEHUNGSWISSENSCHAFTLER:INNEN

Vorlesung 2

Dr. Katja Serova



2-2 Rückblick

Einstieg

- Was sind statistischen Methoden und wozu werden sie gebraucht?
- Was wird in dieser Veranstaltung betrachtet und wozu?

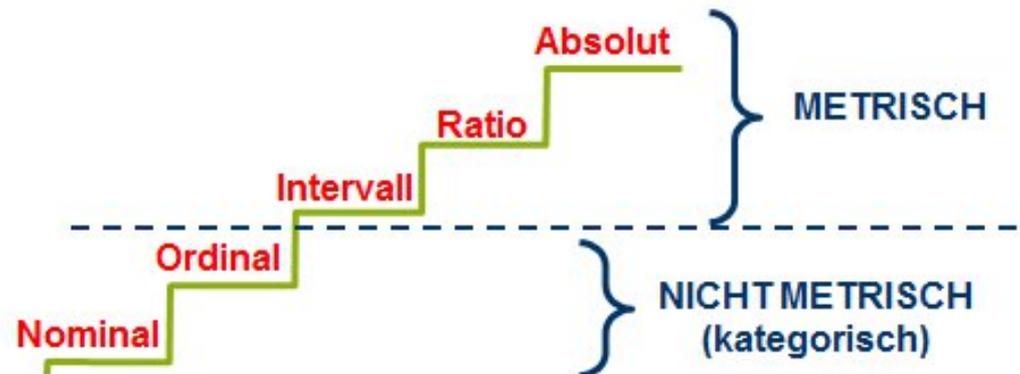


Kapitel: Statistische Variablen und Skalenniveau

- Statistische Variable
- Der Begriff des Messens
- Skalenniveaus
- Typ der statistischen Variablen



$$x_1 = 120$$



2-3

Bereich: Deskriptive Statistik

Kapitel: Häufigkeitsverteilungen

- Was ist eine Häufigkeitsverteilung?
- Tabellarische und graphische Darstellungen
 - Univariate Häufigkeitsverteilungen
 - Bivariate Häufigkeitsverteilungen
- Qualität der Datenpräsentation



2-4 Diskrete Häufigkeitsverteilung (1)



Daten-
erhebung

Urliste

- **Variable:** Zweites Studienfach der EW-Studierenden (nominalskaliert, diskret)
- **Merkmalsraum:** Germanistik (1), Anglistik (2), GeWi (3), MINT (4), Sonstige (5)
- **Gesamtanzahl** der Merkmalsträger (hier: Befragten): $n = 131$

Häufigkeitsverteilung

= Zuordnung der Häufigkeiten zu den Merkmalsausprägungen

Absolute Häufigkeiten

f_i ,ef i'

$$f_1 = 53$$

53 Befragten
studieren EW +
Germanistik

Relative Häufigkeiten

f'_i

$$f'_1 = \frac{53}{131} \cdot 100 \% = 40,46 \%$$

$$f'_i = \frac{f_i}{n} \cdot 100 \% = \frac{f_i}{\sum_{i=1}^k f_i} \cdot 100 \%$$

$k =$ Anzahl der
Merkmalswerte im
Merkmalsraum/
Gruppen / Klassen

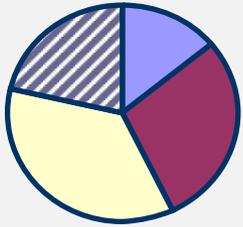
Summe aller
,ef i' Werte
von i gleich 1
bis i gleich k

Nummer	Zweites _Fach
35	1
36	2
37	1
38	4
39	3
40	1
41	5
42	2
43	1

Diese Folie (5) steht Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle wird in der Vorlesung die diskrete Häufigkeitsverteilung mit dem erhobenen Datensatz „Zweites Fach“ vorgestellt. Die Studierenden erhalten die Aufgabe die relativen Häufigkeiten und die prozentualen Anteile zu berechnen. In der Voll-Version sind alle Ergebnisse enthalten.

2-6 Kreisdiagramm (1)



Kreisdiagramm

Graphische Darstellung einer univariaten, diskreten Häufigkeitsverteilung einer **nominalskalierten Variablen**

- Fläche der Sektoren symbolisiert den Anteil
- Gesamtfläche $360^\circ \approx n \approx 100\%$

Berechnung der **Sektorengröße** ...

... aus der absoluten Häufigkeit:

$$\text{Sektor}_1 = \frac{53}{131} \cdot 360^\circ \approx 146^\circ$$

... aus dem Anteil:

$$f'_1 = 0,4046 \approx 0,40$$

$$\text{Sektor}_1 = 0,40 \cdot 360^\circ = 144^\circ$$

Zweites Studienfach x_i	Absolute Häufigkeit f_i	Sektorengröße
Germanistik	53	146°
Anglistik	30	82°
GeWi	27	74°
MINT-Fächer	7	19°
Sonstige	14	38°
	131	360°

Diese Folien (7-10) stehen Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle werden in der Vorlesung weitere Beispiele für Kreisdiagramme und häufige Darstellungsfehler vorgestellt. Auf Folie 10 wird in der Vorlesung die diskrete Häufigkeitsverteilung von ordinalskalierten und metrischen Variablen behandelt.

2-11 Stabdiagramm (1)

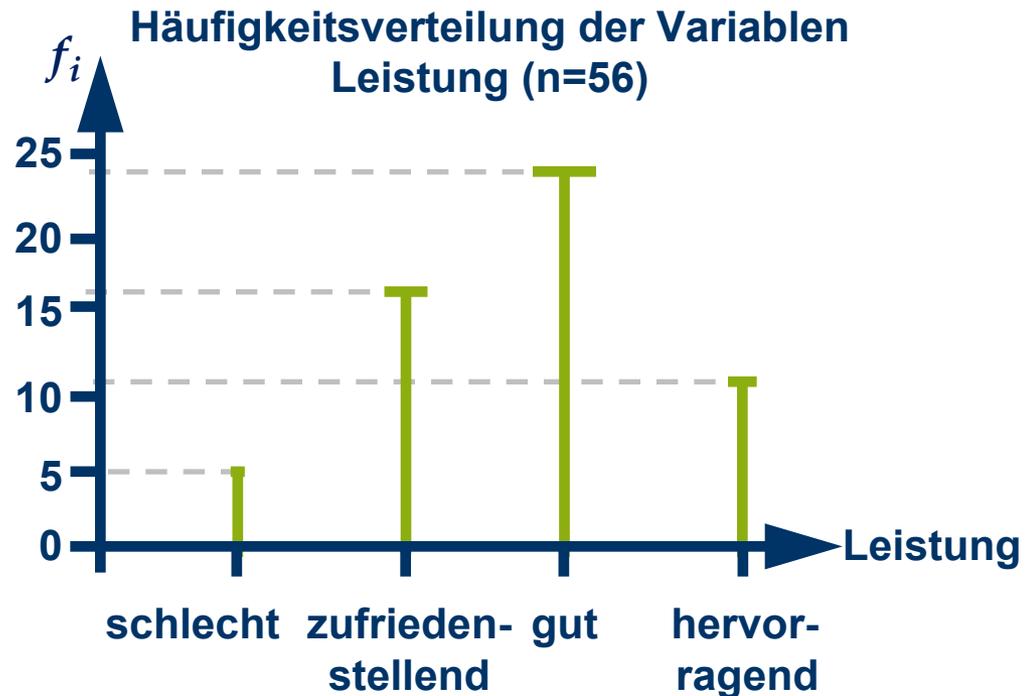


Graphische Darstellung einer Häufigkeitsverteilung

- einer **ordinalskalierten** Variablen
- einer **diskreten metrischen** Variablen

Leistung x_i	f_i	$f'_i(\%)$
schlecht	5	8,93 %
zufriedenstellend	16	28,57 %
gut	24	42,86 %
hervorragend	11	19,64 %
Summe	56	100 %

- Darstellung in einem Achsenkreuz
- **Höhe** der Stäbe symbolisiert absolute oder relative **Häufigkeit**



Diese Folien (12-14) stehen Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle werden in der Vorlesung die Varianten von Stabdiagrammen anhand von Beispielen vorgestellt.

2-15 Kumulierte Häufigkeiten

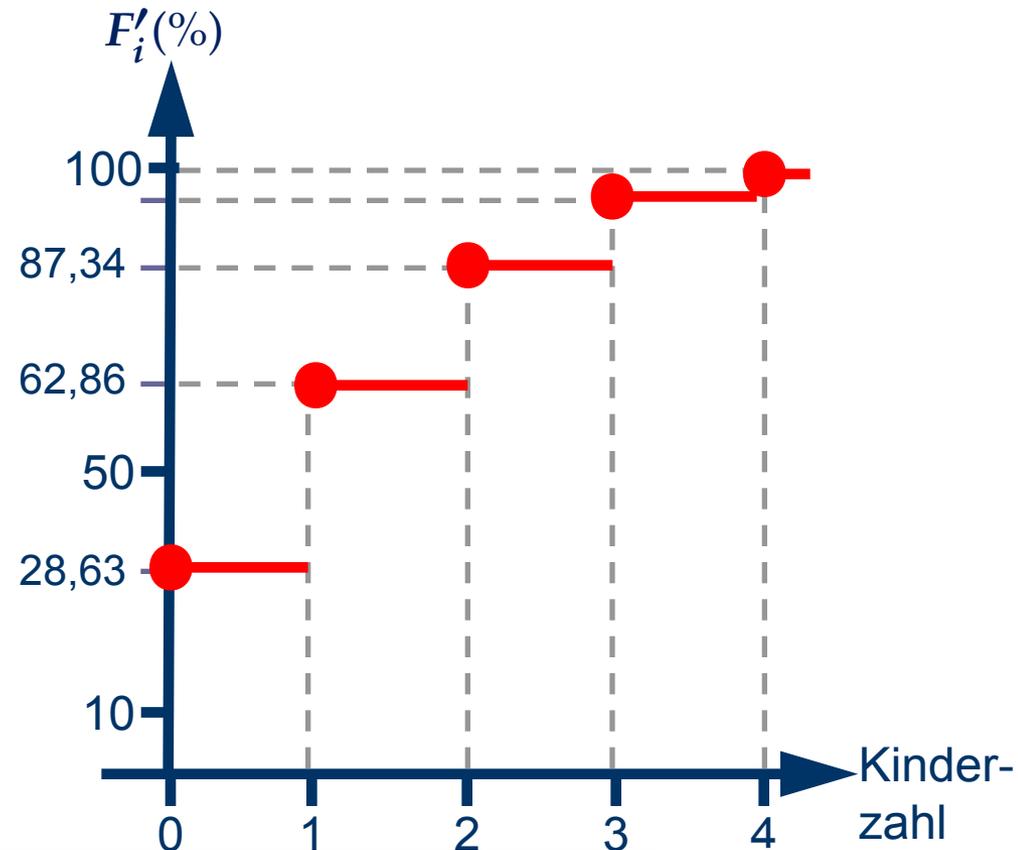
Kinderzahl x_i	Absolute Häufigkeit f_i	Relative Häufigkeit (%) f'_i	Kumulierte Häufigkeit (%) F'_i
0	138	28,63 %	28,63 %
1	165	34,23 %	62,86 %
2	118	24,48 %	87,34 %
3	55	11,41 %	98,75 %
4	6	1,24 %	99,99 %
Summe	482	99,99 %	

- Skalenqualität: **ordinal** oder **metrisch**
- Kumuliert werden **relative** und **absolute** Häufigkeiten
- Kumuliert wird **aufwärts** (von klein nach groß, Intensität steigt) und **abwärts** (von groß nach klein, Intensität sinkt)

2-16 Treppenfunktion

- Graphische Darstellung einer **kumulierten** Häufigkeitsverteilung wird als **Verteilungsfunktion** bezeichnet
- Die Verteilungsfunktion einer **kumulierten diskreten** Häufigkeitsverteilung ist eine **Treppenfunktion**

Kinderzahl X_i	Relative Häufigkeit (%) f'_i	Kumulierte Häufigkeit (%) F'_i
0	28,63 %	28,63 %
1	34,23 %	62,86 %
2	24,48 %	87,34 %
3	11,41 %	98,75 %
4	1,24 %	99,99 %
Summe	99,99 % (*)	



Diese Folien (17-18) stehen Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle wird in der Vorlesung die stetige Häufigkeitsverteilung vorgestellt.

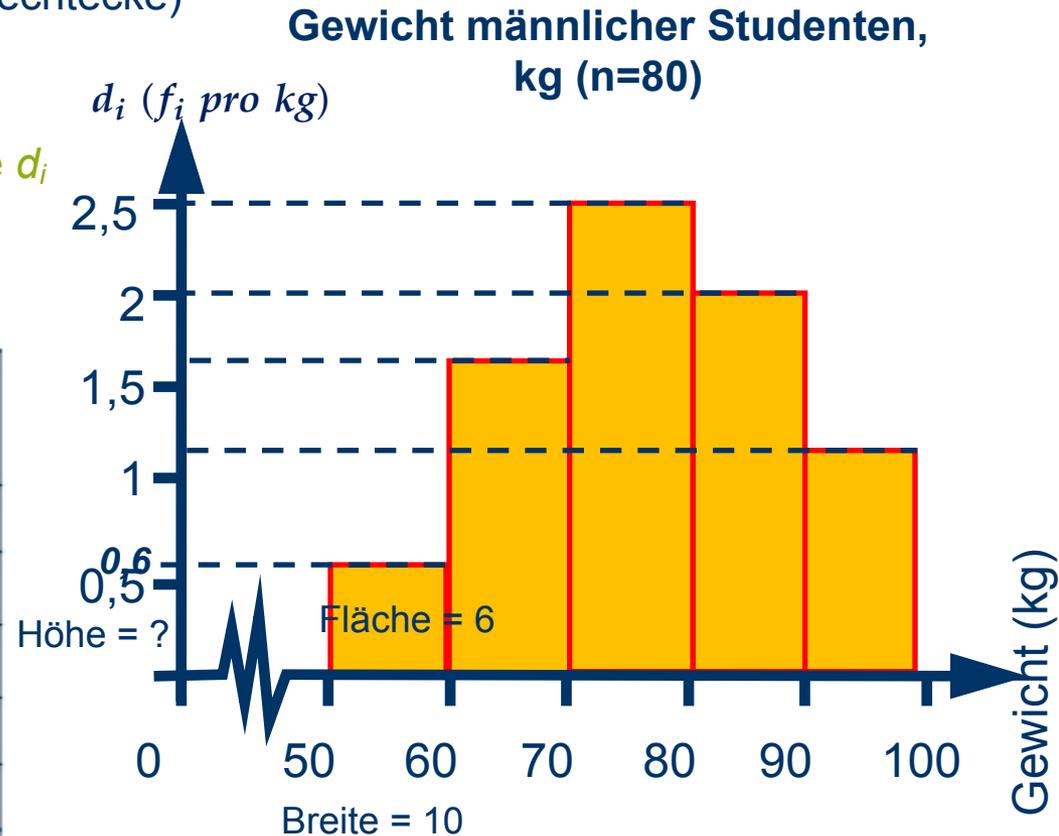
2-19 Histogramm (1)

... graphische Darstellung **stetiger** Häufigkeitsverteilungen

- Häufigkeiten werden durch **Flächen** (Rechtecke) dargestellt
- Darstellung im **Achsenkreuz**
- Häufigkeit = Klassenbreite b_i * **Dichte** d_i

$$d_i = \frac{\text{Häufigkeit}}{\text{Klassenbreite}} = \frac{f_i}{b_i}$$

Gewicht männlicher Studenten (kg)	Absolute Häufigkeit f_i	Klassenbreite b_i	Dichte d_i
50 bis unter 60	6	10	0,6
60 bis unter 70	17	10	1,7
70 bis unter 80	25	10	2,5
80 bis unter 90	20	10	2,0
90 bis unter 100	12	10	1,2
Summe	80		



Diese Folien (20-22) stehen Ihnen in dieser Version nicht zur Verfügung.

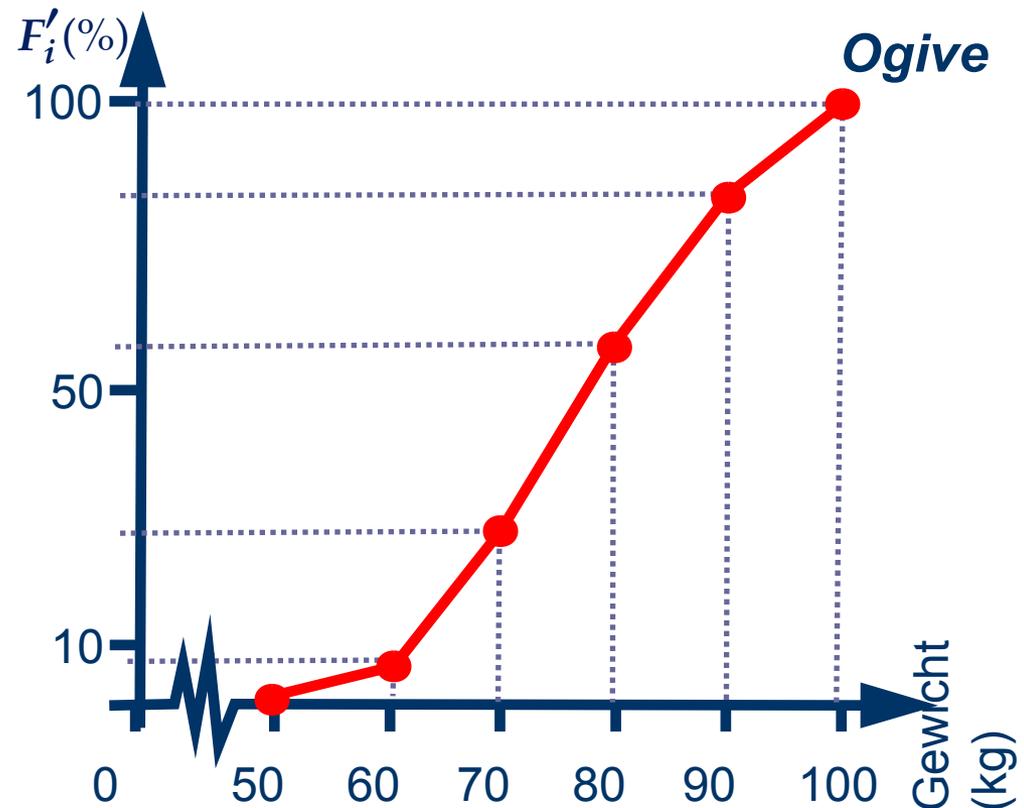
An dieser Stelle wird in der Vorlesung das Histogramm mit Beispielen weiter vertieft.

2-23 Ogive

- Graphische Darstellung einer **kumulierten** Häufigkeitsverteilung wird als **Verteilungsfunktion** bezeichnet
- Die Verteilungsfunktion einer **kumulierten stetigen** Häufigkeitsverteilung ist eine **Ogive (Summenkurve)**

Gewicht männlicher Studenten,
kg (n=80)

Gewicht (kg)	Kumulierte Häufigkeit (%) F'_i
50 bis unter 60	7,50 %
60 bis unter 70	28,75 %
70 bis unter 80	60,00 %
80 bis unter 90	85,00 %
90 bis unter 100	100,00 %



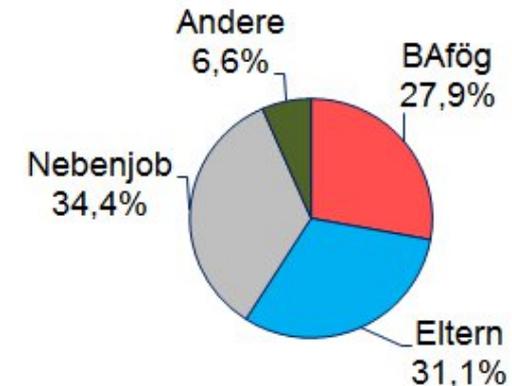
Diese Folie (24) steht Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle erfolgen in der Vorlesung Beispiele zu Histogrammen und Ogiven.

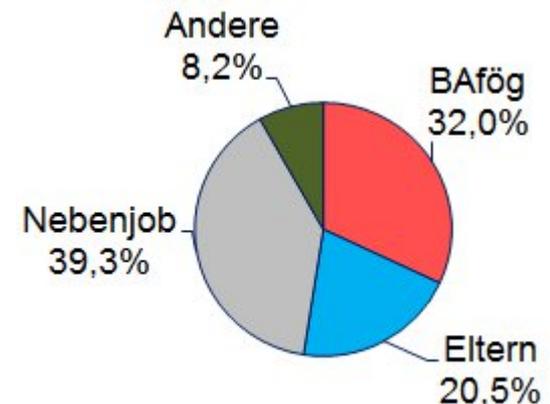
2-25 Bivariate Häufigkeitsverteilung (1)

Hauptfinanzierungsquelle	Studienfach		Summe
	WiWi und Jura	SoWi	
BAfög	17	39	56
Eltern	19	25	44
Nebenjob	21	48	69
Andere	4	10	14
Summe	61	122	183

Hauptfinanzierungsquelle



WiWi und Jura (n = 61)



SoWi (n = 122)

Relative Häufigkeiten (%)

- **Gesamt:** **26,23 %** der Befragten haben einen Nebenjob und studieren SoWi
- **Spaltenprozente:** **39,34 %** der SoWi-Studierenden haben einen Nebenjob
- **Zeilenprozente:** **69,56%** der Studierenden mit Nebenjob studieren SoWi

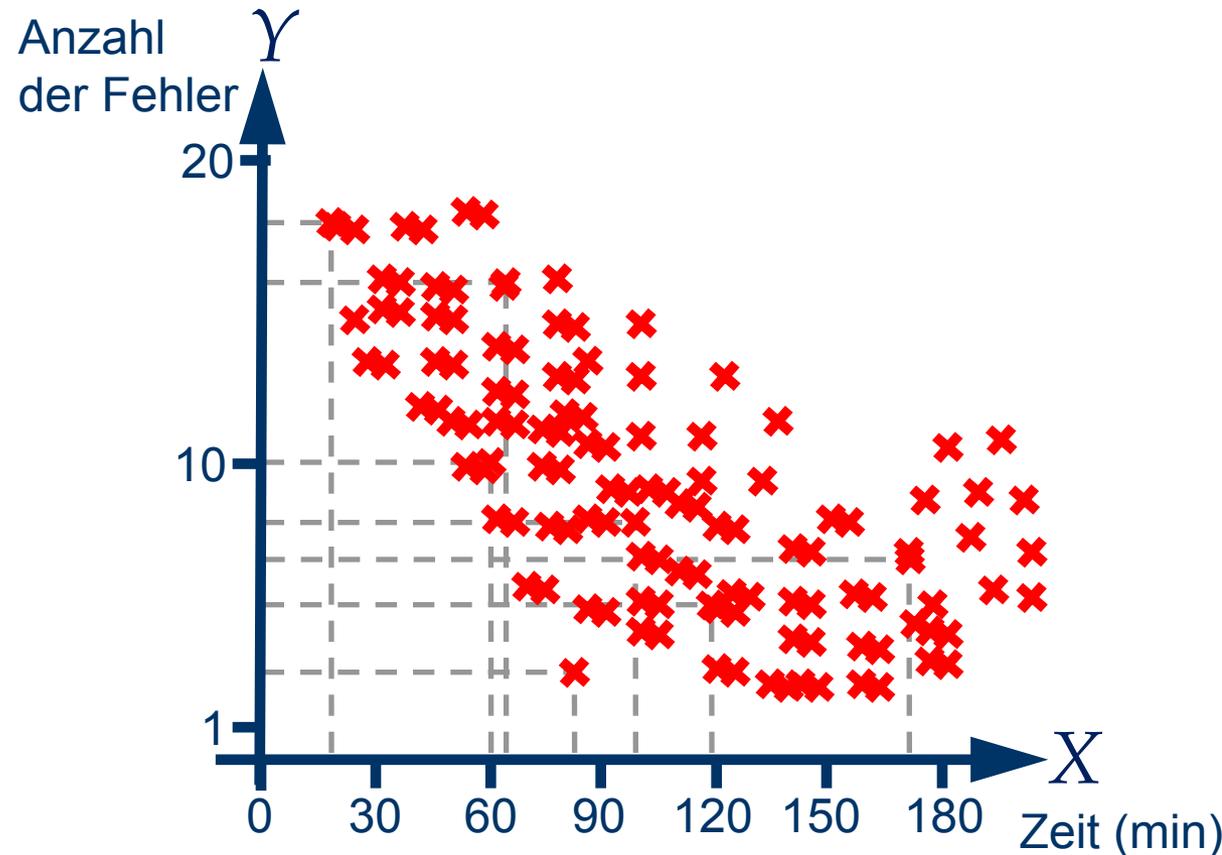
Diese Folie (26) steht Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle wird die bivariate Häufigkeitsverteilung in verschiedenen Säulendiagrammen dargestellt.

2-27 Streudiagramm

Graphische Darstellung von **zwei metrischen** Variablen (Punktewolke)

Vorbereitungs- zeit (Min) x_i	Anzahl der Fehler y_i
60	10
120	5
20	18
65	16
100	8
170	7
85	3



Diese Folie (28) steht Ihnen in dieser Version nicht zur Verfügung.

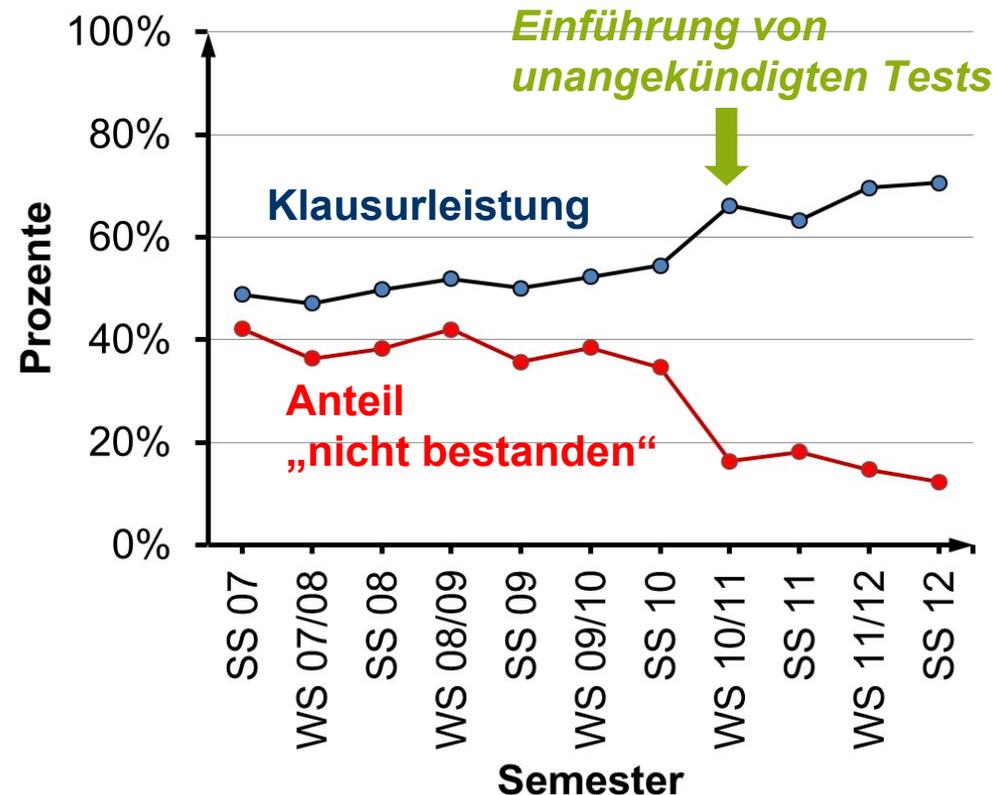
An dieser Stelle wird in der Vorlesung das Streudiagramm anhand eines Beispiels näher betrachtet.

2-29 Liniendiagramm

- **Liniendiagramm:** graphische Darstellung von metrischen Variablen im **Zeitverlauf**

Semester	Klausurleistung im Prozent	Anteil "nicht bestanden"
SS 07	48,8%	42,1%
WS 07/08	47,1%	36,4%
SS 08	49,7%	38,3%
WS 08/09	51,9%	42,0%
SS 09	50,1%	35,7%
WS 09/10	52,3%	38,5%
SS 10	54,5%	34,6%
WS 10/11	66,1%	16,3%
SS 11	63,3%	18,1%
WS 11/12	69,6%	14,7%
SS 12	70,5%	12,3%

Klausurleistung und Anteil der nicht bestandenen Klausuren im Prozent vom SS 2007 bis SS 2012



Diese Folien (30-31) stehen Ihnen in dieser Version nicht zur Verfügung.

An dieser Stelle werden in der Vorlesung weitere Beispiele von Liniendiagrammen vorgestellt.

2-32 Qualität der Datenpräsentation

Eigenschaften **guter** tabellarischer und graphischer Darstellungen:

- informativ
- korrekt
- klar
- übersichtlich

Die **Art** der graphischen
Darstellung



Skalenniveau und **Typ** (diskret oder stetig)
der Variablen

2-33

Nächste Übung: **Aufgabenblatt 02**

Nächste Vorlesung

Bereich: Deskriptive Statistik

Kapitel: Maße der zentralen Tendenz



Bis zum nächsten Mal
oder bis Freitag!

2-34 Texte zu diesem Kapitel



- **Eindimensionale Häufigkeitsverteilung & Graphische Darstellung von Daten** - aus Bourier, Günther (2008): Beschreibende Statistik. Praxisorientierte Einführung. Mit Aufgaben und Lösungen. 7., überarbeitete Auflage. Wiesbaden: Gabler Verlag, S. 38-64 (Auszüge).
- **Ten Ways to a Great Figure** - aus Salkind, Neil J. (2000): Statistics for people who (think they) hate statistics. London: Sage publications, S. 54.

Zusätzliche Texte

- **Häufigkeitsverteilungen** - aus Oestreich, Markus; Romberg, Oliver (2010): Keine Panik vor Statistik. Erfolg und Spaß im Horrorfach nichttechnischer Studiengänge. 3., überarbeitete Auflage. Wiesbaden: Vieweg+Teubner Verlag, S.49-68.

Quellen für die Bilder in den Folien:

- www.freepik.com (designed by  freepik)
- www.icon-icons.com
- Spiegel und ADAC