

Datenerhebung, Datenaufbereitung und Darstellung.

In der beschreibenden Statistik werden Daten erhoben, aufbereitet und analysiert.

Beispiel einer Datenerhebung mit Begriffserklärungen (Vokabel)

Was ist an der Behauptung dran „Unsere Jugend wird immer dicker und träger“?
Um eine Behauptung dieser Art zu überprüfen kann man nicht alle Jugendlichen befragen, sondern nur einen Teil der Gesamtheit.

Stichprobe: Wird der Teil einer Gesamtheit befragt, dann spricht man bei der Datenerhebung von einer Stichprobe.

Urliste: Das Ergebnis der Stichprobe wird in einer Urliste festgehalten.

Beispiel einer Urliste:

Schüler Nr.	Geschlecht		Körpergröße in cm	Gewicht in kg	Raucher		Sportart
	w	m			ja	nein	
01		x	160	52		x	Handball
02	x		172	67	x		Fußball
03	x		180	60	x		Golf
...
xx		x	170	65		x	Judo

(Eine entsprechende Urliste ist in der jeweiligen Klasse anonym zu erstellen)

Rohdaten: Sind alle in der Urliste enthaltenen Daten.

Erhebungsumfang: Ist die Anzahl der untersuchten Objekte (hier Schüler)
Werden z.B. 27 Schüler befragt, so sagt man,
“Die Anzahl der Merkmalsträger ($n = 27$) bildet den Erhebungsumfang“.

Merkmale: Sind die Eigenschaften der Objekte.
(z.B. Geschlecht, Körpergröße, Gewicht, Raucher, Sportart, ...)

Merkmalsausprägung x_i : Ein Merkmal kann in verschiedenen Ausprägungen vorkommen. (z.B. Geschlecht m oder w)

Beispiele für Merkmale und deren Ausprägungen:

Merkmal	Merkmalsausprägung x_i
Geschlecht	männlich, weiblich
Körpergröße	160 cm, 182 cm, 154 cm, 163 cm, ...
Körpergewicht	52 kg, 81 kg, 71 kg,
Raucher	ja, nein
Sportart	Judo, Tischtennis, Fußball, ...

Bemerkung: Durch eine Erhebung soll festgestellt werden, wie die verschiedenen Ausprägungen eines Merkmals in einer Stichprobe verteilt sind. Vor der Erhebung sind einige Fragen zu klären, z.B.
Wonach soll gefragt werden?
Welche Antworten sind möglich?
Welchen Umfang soll die Stichprobe haben?

Aufbereiten der Daten und ihre Darstellung.

Ermittlung der Häufigkeit bestimmter Merkmalsausprägungen

In einer **Strichliste** werden Merkmalsträger mit gleichen Merkmalsausprägungen für ein bestimmtes Merkmal zusammengefasst.

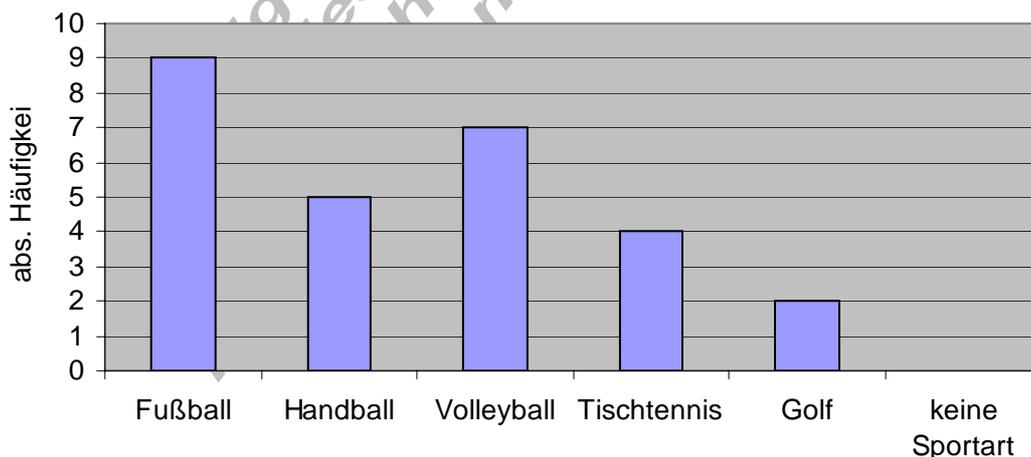
Bearbeitung des Merkmals *Sportart* (eventuell in Klassen zusammenfassen)

Merkmalsausprägung		absolute Häufigkeit
Fußball	/// IIII	9
Handball	///	5
Volleyball	/// II	7
Tischtennis	IIII	4
Golf	II	2
keine Sportart		0

Der Merkmalsausprägung z.B. Fußball wird die absolute Häufigkeit 9 zugeordnet.

Grafische Darstellung durch ein **Säulendiagramm**.

Säulendiagramm



Berechnung des Body - Mass - Index (BMI):
$$\text{BMI} = \frac{\text{Gewicht in kg}}{(\text{Körpergröße in m})^2}$$

Wertung:

I) $0 < x < 20$ Untergewicht	II) $20 \leq x \leq 25$ normal
III) $25 < x \leq 30$ Übergewicht	IV) $30 < x \leq 35$ moderate Fettsucht
V) $35 < x$ morbide Fettsucht	

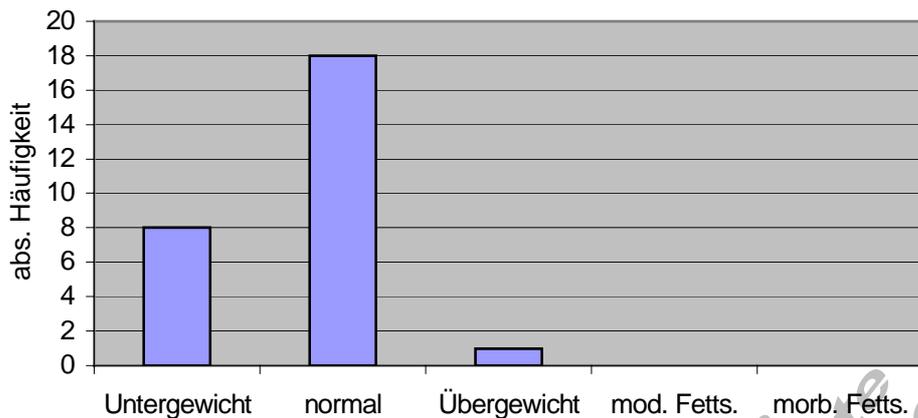
Auszug aus der Urliste mit Berechnung des BMI

Schüler Nr.	Geschlecht		Körpergröße in cm	Gewicht in kg	BMI	Wertung
	m	w				
01		x	160	52	20,3	normal
02	x		172	67	22,6	normal
03	x		180	60	18,5	Untergewicht
04		x	167	55	19,7	Untergewicht
05		x	178	63	19,9	Untergewicht
06		x	175	63	20,6	normal
07	x		183	70	20,9	normal
08	x		188	78	22,1	normal
09	x		181	84	25,6	Übergewicht
10	x		183	68	20,3	normal
11		x	162	63	24,0	normal
12		x	171	57	19,5	Untergewicht
13	x		177	67	21,4	normal
14		x	165	58	21,3	normal
15	x		174	70	23,1	normal
16	x		179	73	22,8	normal
17		x	175	55	18,0	Untergewicht
18	x		183	72	21,5	normal
19		x	163	51	19,2	Untergewicht
20		x	163	60	22,6	normal
21		x	165	64	23,5	normal
22		x	171	51	17,4	Untergewicht
23		x	175	54	17,6	Untergewicht
24	x		176	68	22,0	normal
25	x		184	75	22,2	normal
26	x		185	76	22,2	normal
27		x	169	59	20,7	normal

BMI – Auswertung:

Wertung	Untergewicht	normal	Übergewicht	mod. Fetts.	morb. Fetts.
abs. Häufigkeit	8	18	1	0	0

Body-Mass-Index (BMI)

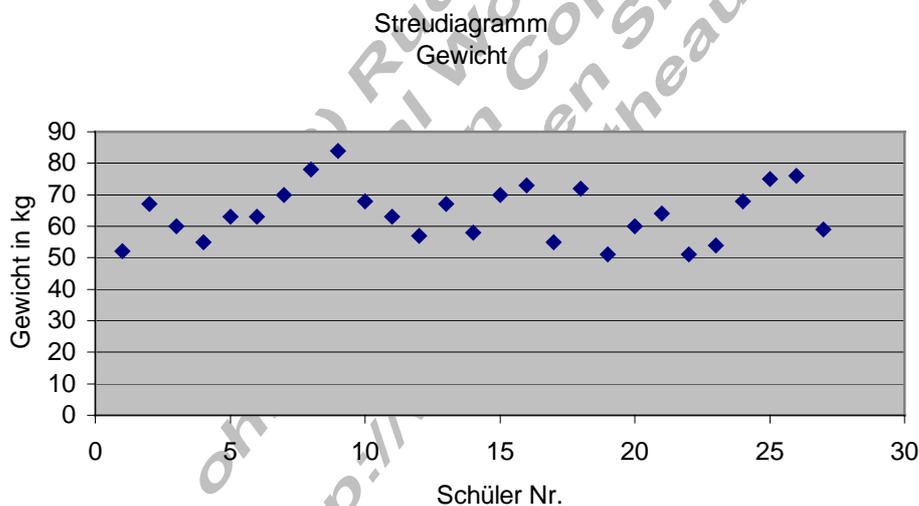


Bearbeitung des Merkmals *Gewicht*.

Auszug aus der Urliste: (Wertetabelle)

Nr.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
kg	52	67	60	55	63	63	70	78	84	68	63	57	67	58	70	73	55	72	51	60	64	51	54	68	75	76	59

Darstellung im *Punktdiagramm* (Streudiagramm).



Durch die Zuordnung Schüler – Gewicht wird die Übersichtlichkeit nicht wesentlich erhöht. Es kann jedoch unmittelbar abgelesen werden, dass das Gewicht aller Schüler zwischen 40 kg und 100 kg liegt.

Eine Strichliste zur Bestimmung der Häufigkeit einer bestimmten Merkmalsausprägung macht hier so keinen Sinn.

Eine Einteilung der Merkmalsausprägungen in **Klassen** wird nun durchgeführt. Es werden 4 Klassen mit einer Klassenbreite von 10 kg gewählt, in welche die Körpergewichte eingeordnet werden.

Für das Merkmal Gewicht gibt es als Ausprägungen nun die Klassen I bis IV.

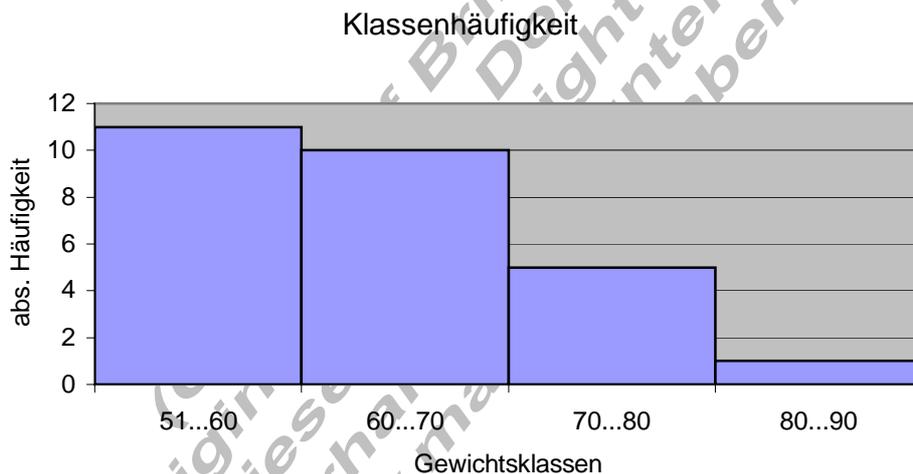
Klasse (in kg)		absolute Klassenhäufigkeit
I) $51 \leq x < 60$	III III I	11
II) $60 \leq x < 70$	III III	10
III) $70 \leq x < 80$	III	5
IV) $80 \leq x < 90$	I	1

Klasseneinteilung Werden verschiedene Merkmalsausprägungen zu einer neuen Ausprägung zusammengefasst, so spricht man von einer Klasseneinteilung der Stichprobenwerte. Die Darstellung erfolgt in einem Säulendiagramm ohne Lücken.

Häufigkeitstabelle

Klasse in kg x_i	$51 \leq x < 60$	$60 \leq x < 70$	$70 \leq x < 80$	$80 \leq x < 90$
absolute Häufigkeit n_i	11	10	5	1

Grafische Darstellung der Häufigkeitsverteilung der Klassen.



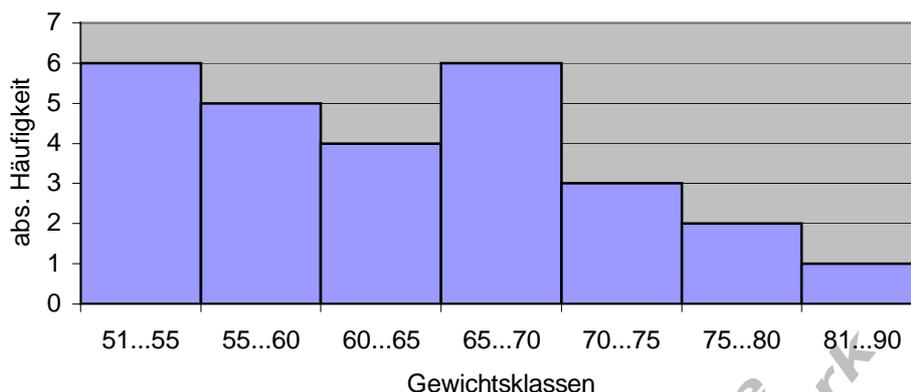
Die Klassenbreite (10 kg) wurde willkürlich festgelegt. Nun soll die Klassenbreite halbiert werden. Damit vergrößert sich die Anzahl der Klassen.

Häufigkeitstabelle

Klasse in kg x_i	51... 55	55... 60	60... 65	65... 70	70... 75	75... 80	80... 85
Häufigkeit n_i	6	5	4	6	3	2	1

Bemerkung: 51 ... 55 bedeutet: $51 \leq x < 55$ (halboffenes Intervall)

Klassenhäufigkeit



Grafiken im Vergleich:

Je geringer die Anzahl der Klassen, desto besser ist die Übersichtlichkeit. Der Informationsgehalt wird aber geringer (Hier kann manipuliert werden). Es gibt keine allgemeingültige Festlegung der Klassenbreite, deshalb muss eine sinnvolle Klassenbreite für die Häufigkeitstabelle gewählt werden. Üblich ist ein Wert zwischen 5 und 10.

Weitere Darstellungsarten im Säulendiagramm

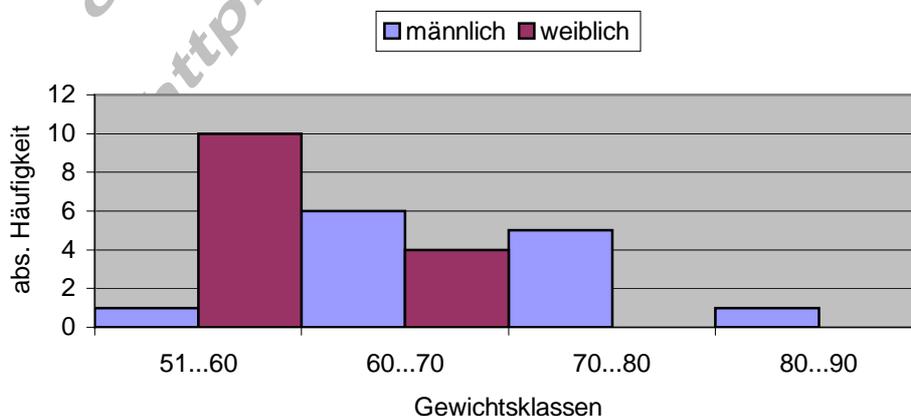
Um die Unterschiede der Körpergewichte von Schülern und Schülerinnen aufzuzeigen, wird die Häufigkeitstabelle mit der Klassenbreite 10 kg nach männlich und weiblich aufgelistet.

Häufigkeitstabelle

Klasse in kg x_i	51 ... 60	60 ... 70	70 ... 80	80 ... 90
Häufigkeit (männlich) n_1	1	6	5	1
Häufigkeit (weiblich) n_2	10	4	0	0

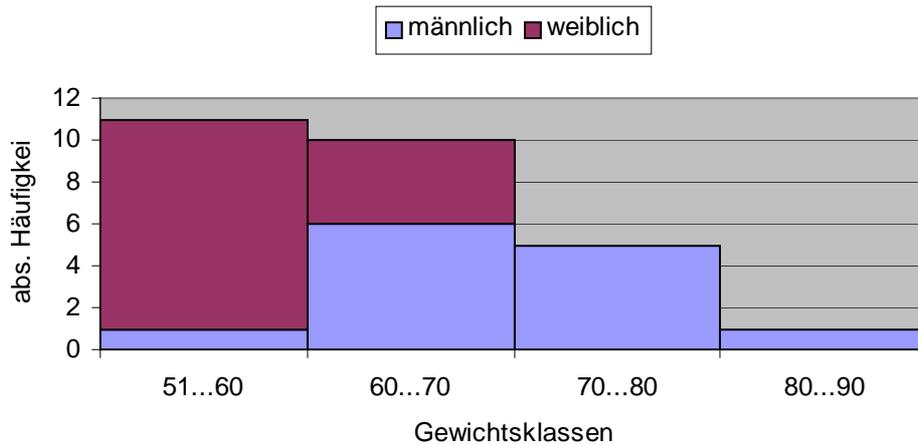
Säulendiagramm für zwei Verteilungen

Häufigkeitsdiagramm männlich/weiblich



Doppelsäulendiagramm

Häufigkeitsverteilung männlich/weiblich



Die Körpergewichte der Schülerinnen verteilen sich nur auf die ersten beiden Klassen, während 6 von 13 Schülern ein Gewicht von mehr als 70 kg haben. Das sind $(100:13) \cdot 6 \approx 46\%$.

Weiteres Beispiel zur Klasseneinteilung

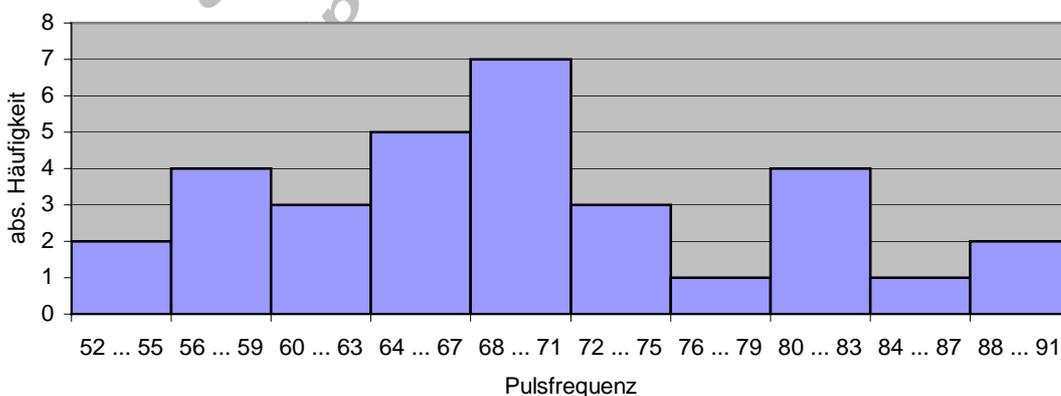
In einer Klasse mit 32 Schülern misst jeder Schüler seine Pulsfrequenz. Die Erhebungsdaten werden in eine Urliste eingetragen. Zur Bestimmung der Klassenbreite und damit der Anzahl der Klassen ist es sinnvoll, in der Urliste den kleinsten und den größten Wert zu markieren.

64	65	70	80	88	58	60	68	63	64	57	77	74	73	62	52
72	84	63	90	68	59	58	71	80	82	81	69	53	65	69	71

Klassenbreite 4 ergibt 10 Klassen.

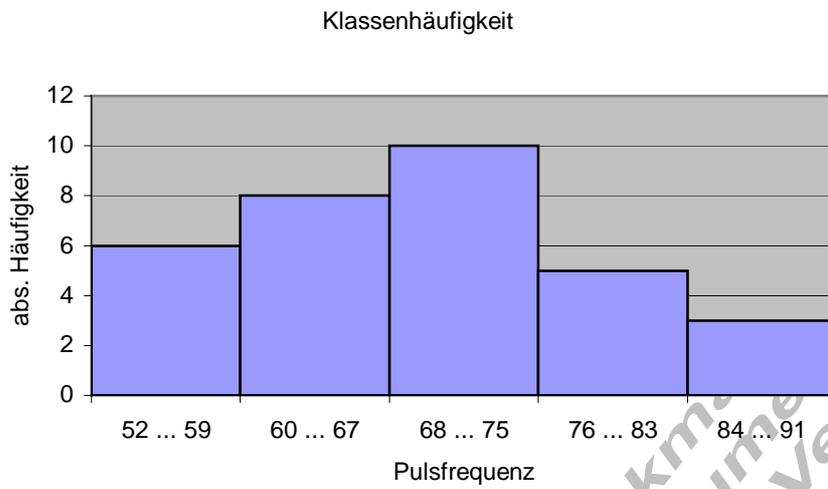
Klasse x_i	52 ... 55	56 ... 59	60 ... 63	64 ... 67	68 ... 71	72 ... 75	76 ... 79	80 ... 83	84 ... 87	88 ... 91
Häufigkeit n_i	2	4	3	5	7	3	1	4	1	2

Klassenhäufigkeit der Pulsfrequenz



Klassenbreite 8 ergibt 5 Klassen.

Klasse x_i	52 ... 59	60 ... 67	68 ... 75	76 ... 83	84 ... 91
Häufigkeit n_i	6	8	10	5	3



(C) Rudolf Brinkmann
Original Word-Dokumente
ohne diesen Copyright-Vermerk
<http://www.matheaufgaben-du.de>