

Onlinekurs Statistik – mit Lösung

Theorie: <http://juttgut.at/kurs/statistik1.htm>

VIDEOS:

- Skalenniveaus: <https://www.youtube.com/watch?v=TV4tTtW4UBU>
- Klasseneinteilung + Histogramm: <https://www.youtube.com/watch?v=uJ2DISBckps>
- Stängel-Blatt-Diagramm lesen: <https://www.youtube.com/watch?v=-JdpneFELeK>
- Modus, Median, Mittelwert: https://www.youtube.com/watch?v=7Z_nLMn9z0E
- Boxplot: <https://www.youtube.com/watch?v=ttz7Oo-7NhA>
- Standardabweichung: <https://www.youtube.com/watch?v=MJA0-MCAVtk>
- TI-30: <https://www.youtube.com/watch?v=SwhJlayiAoY>

STATISTIK 1: Erklärung vom Kursleiter (Nominalskala, Ordinalskala):

VIDEO: <https://www.youtube.com/watch?v=PLF-csIlogA>

Ausgehend von 18 zufälligen erzeugten Zahlen zwischen 20 und 70, die z.B. das Alter von KursteilnehmerInnen sein können (erzeugt mit dem Taschenrechner-Befehl RAND, der sich bei der PRB-Taste fast ganz rechts befindet) soll zuerst eine Stängel-Blatt-Diagramm erzeugt werden

- 18 Alters-Zahlen (mit $RAND*50+20$ erzeugt):
35,39,58,45,27, 63,48,60,38,34, 35,35,44,57,64, 41,56,57
- **Stängel-Blatt-Diagramm** (aus den Zahlen der Reihe nach erstellt)

Zehner	Einer
2	7
3	5,9,8,4,5, 5
4	5,8,4,1
5	8,7,6,7
6	3,0,4

- Daraus kann man die **sortierte Liste** machen:
27,34,35, 35,35,38, 39,41,44, **nix** 45,48,56, 57,57,58, 60,63,64
- **Nominalskala und Modus**

Von dieser Liste kann man den Modus (=häufigster Wert) recht gut ablesen: **MODUS = 35**

- **Ordinalskala: Min-Quartile-Max**

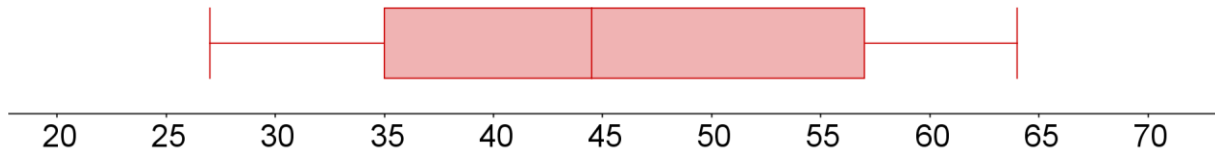
Außerdem sieht man am linken Ende der Liste das MINIMUM (=27) und am rechten Ende das MAXIMUM (=64), die Differenz der beiden Werte nennt man

SPANNWEITE = MAX – MIN = 64–27 = **37**

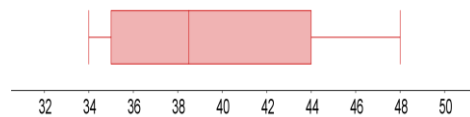
Weiters ist der Wert, der in der Mitte steht der MEDIAN = ZENTRALWERT = 2.Quartil. Da hier aber in der Mitte **nix** steht, nimmt man als MEDIAN den Mittelwert der Nachbarn = $(44+45):2 = 44,5$. Zum Schluss findet man noch in der Mitte der linken Liste das 1.QUARTIL (=35) und in der Mitte der rechten Liste das 3.QUARTIL (=57)

- **Boxplot**

Daraus lässt sich ein sogenanntes **BOXPLOT** erstellen: Zuerst muss eine Skala von MIN bis MAX gezeichnet werden und darüber die 5 Werte (MIN, 1.Quartil, MEDIAN, 3.Quartil, MAX) mit senkrechten Strichen eingezeichnet werden und zum Schluss waagrecht mit Strichen zu 2 Boxen und zwei Henkeln verbunden werden:

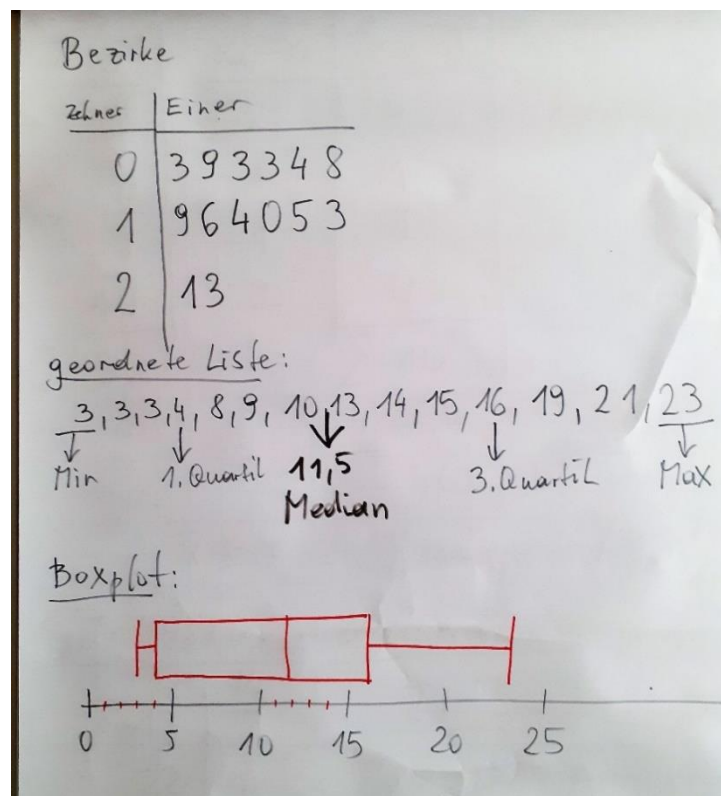


- **2.Boxplot zum Vergleich**



In diesem 2.Boxplot ist die Spannweite viel kleiner als im 1.Boxplot, das heißt, dass die Altersdaten sich enger um den Median (bei 38) gruppieren.

Übung: Erstellen Sie ein **Stängel-Blatt-Diagramm**, eine **sortierte Liste** und ein **Boxplot** mit den Bezirksangaben von 14 TeilnehmerInnen:
21,03,19,16,14, 09,03,23,10,15, 13,03,04,08



Oder: Beispiel 3 und Übung 3 von https://mathe-mit-manfred.at/math/TH_UE_Statistik_Skalen_Gurtner_SS04.pdf

STATISTIK 2: Erklärung vom Kursleiter (metrische Skala erlaubt Zwischenwerte):

<https://www.youtube.com/watch?v=LA1Tmi5Ut8Q>

- Gegeben sind Gästezahlen an 30 Tagen in einem Restaurant:
20,20,30,19,20, 28,15,36,28,25, 37,28,30,21,18,
16,26,33,25,26, 20,35,19,23,38, 23,37,22,36,22

Daraus wird eine **Klasseneinteilung** gemacht, um zu 5 Gruppen zu kommen – zur besseren Übersichtlichkeit – mit den Gruppen **15–19, 20–24, 25–29, 30–34, 35–39**

- Absolute Häufigkeit + relative Häufigkeit
- Mittelwert mit Tabelle

Danach wird aus der **Strichliste** die **absolute Häufigkeit**, nach Division durch die Summe der Daten (=30) die **relative Häufigkeit** (auch in %).

Zum Schluss wird mit Hilfe der **Klassenmitten** der Personenzahlen und den relativen Häufigkeiten das **Produkt** gebildet und summiert. Die Summe ist das der **Mittelwert** der Personenzahlen (kann auch mit den Produkten mit den absoluten Zahlen, deren Summe und zum Schluss Division durch 30 erreicht werden)

Klasse	Strichliste	Abs. Häuf.	Rel. Häufigkeit	Klassenmitte	Produkt
15–19	IIII	5	5/30 = 16,7%	17	2,83
20–24	IIII IIII	9	9/30 = 30%	22	6,6
25–29	IIII II	7	7/30 = 23,3%	27	6,3
30–34	III	3	3/30 = 10%	32	3,2
35–39	IIII I	6	6/30 = 20%	37	7,4
SUMMEN		30	1 = 100%		26,33

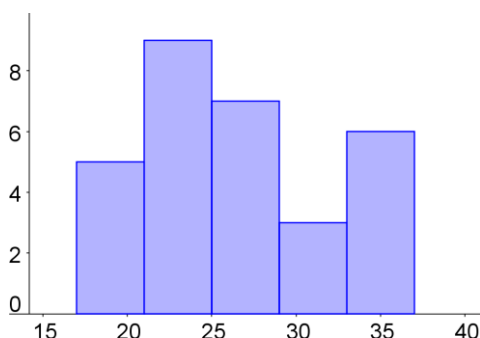
Der letzte Wert ist der **Mittelwert** \bar{x} ($=\mu$) = **26,33** der Anzahl der Personen, die im Gasthaus waren.

- Mittelwert der Urliste

Der Mittelwert der Urliste der 30 Zahlen ist $\frac{20+20+30+19+\dots+22+36+22}{30} = \mathbf{25,86}$. Das ist abweichend, da wir bei der Klasseneinteilung manche Zahlen größer oder kleiner gemacht haben und durch die Klassenmitten ersetzt haben.

- Histogramm (oder Staffeldbild)

Nun kann man mit den Klassen (auf der x-Achse) und den absoluten Häufigkeiten (auf der y-Achse) ein Histogramm (auch Staffeldbild genannt) zeichnen:



- **Standardabweichung** <https://www.youtube.com/watch?v=T6uTJvYa9O4>
Um nun eine mittlere Abweichung der Daten vom Mittelwert zu bekommen = **Standardabweichung** s ($= \sigma$) zeige ich es an Hand der 4 Daten 6,7,9,10

	Daten x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	
	6	$6 - 8 = -2$	4	
	7	$7 - 8 = -1$	1	
	9	$9 - 8 = 1$	1	
	10	$10 - 8 = 2$	4	
SUMME	32	0 (fad)	10	
Mittelwert	$32/4 = 8$		$10/4 = 2,5$	= VARIANZ
Wurzel			$\sqrt{2,5} = 1,6$	= Standardabweichung

Zuerst wird die Summe der Daten ($=32$) berechnet, dann der Mittelwert ($= 32/4 = 8$) in der ersten Spalte.

In der 2.Spalte wird die Differenz zwischen Daten und Mittelwert berechnet, das ergibt die Abweichung vom Mittelwert. Die kann man nicht sinnvoll summieren, da dann 0 herauskommt (fad).

Aber man kann diese Differenzen quadrieren (dritte Spalte), dann summieren, durch 4 dividieren und zum Schluss noch die Wurzel ziehen, dann hat man die Standardabweichung ($= 1,6$)

2.Verfahren

Es gibt ein 2.Verfahren, das zum gleichen Resultat kommt: Hier werden die Daten zuerst quadriert, summiert, durch Anzahl dividiert, **MINUS Quadrat des Mittelwerts** und die Wurzel daraus gezogen

	Daten x_i	x^2	
	6	36	
	7	49	
	9	81	
	10	100	
SUMME	32	266	
Mittelwert (der Quadrate)	$32/4 = 8$	$266/4 = 66,5$	
Minus Quadrat des Mittelwertes		$66,5 - 8^2 = 2,5$	= VARIANZ
Wurzel		$\sqrt{2,5} = 1,6$	= Standardabweichung

MERKSPRUCH: *Mittelwert der Quadratsumme minus Quadrat des Mittelwerts* → VARIANZ

- Eingabe im TI-30

Für die Daten aus dem 2. Beispiel zeige ich die Eingabe im Taschenrechner TI-30, die praktischer ist für die Berechnung der Standardabweichung:

Klassenmitten x_i	Abs. Häuf. H_i	Quadrate x_i^2	Produkt $x_i^2 * H_i$
17	5	289	1445
22	9	484	4356
27	7	729	5103
32	3	1024	3072
37	6	1369	8214
SUMMEN	30		22190
Geteilt durch 30			739,66
Minus Quadrat des Mittelwerts			$739,66 - 26,33^2 = 46,39$
Wurzel = Standardabweichung			$\sqrt{46,39} = 6,8$

Start der Berechnung: **2nd DATA** 1-Var =

DATA: X1 = 17 → dann Taste **Scrollen nach unten**

FRQ = 5 → dann Taste **Scrollen nach unten**

X2 = 22 → dann Taste **Scrollen nach unten**

FRQ = 9 → dann Taste **Scrollen nach unten**

.....
Nach der Eingabe aller Daten kommt **STATVAR:** $n=30$, $\bar{x} = 26,33$, $\sigma_x = 6,79$

Nicht vergessen : 2nd STATVAR = Exit STAT drücken, damit man wieder scrollen kann !

Übungen:

Übung 6 von https://mathe-mit-manfred.at/math/TH_UE_Statistik_Skalen_Gurtner_SS04.pdf

Übung 13 und 15 von <http://juttgut.at/kurs/statistik.pdf>

=====
=

STATISTIK 3: Lineare Regression:

Theorie: <http://juttgut.at/kurs/statistik2.htm>

VIDEOS

- vom Kursleiter: https://www.youtube.com/watch?v=oU_b4j7Fau8
- [Korrelationskoeffizient-Bedeutung](#)

Erklärung vom Kursleiter: TI-30 -Eingabe

Start der Berechnung: **2nd DATA** 2-Var =

DATA: X1 = ... → dann Taste **Scrollen nach unten**

Y1 = ... → dann Taste **Scrollen nach unten**

X2 = ... → dann Taste **Scrollen nach unten**

Y2 = 9 → dann Taste **Scrollen nach unten**

.....
Nach der Eingabe aller Daten kommt **STATVAR:** ganz rechts a=..., b=..., r=...

Daraus entsteht die LINEARE FUNKTION $y = a \cdot x + b$ mit Güte r^2

Nicht vergessen : 2nd STATVAR = Exit STAT drücken, damit man wieder scrollen kann !

Übungen: 2, 3 ev. 7 von

https://mathe-mit-manfred.at/math/TH_UE_Regression_Gurtner_SS09.pdf

HAUSÜBUNG:

[Statistik:](#) 1,2, 4, 8, 16

[Regression:](#) 1,9