Aufgabe 8:

a categorical
b quantitative
c categorical
d categorical
e quantitative
f quantitative

Aufgabe 9:

n_b n_bauern = 10 e_b = einkommen_bauern = 10.000
n_m n_millionaere = 1 e_m = einkommen_millionaere = 100.000.000

arithm. Mittel: $\left(\sum_{i=1}^{(n_b+n_m)} e_b + \sum_{i=1}^{n_m} e_m \right) \cdot \frac{1}{n_b + n_e}$

$= 9.100.000$

Median: Weil $n_{ges} = n_b + n_m = 11$ (und somit ungerade ist), gilt für die sortierte Liste $(e_b, e_b, \dots e_b, e_m)$ dass der Median an position $\left(\frac{n+1}{2}\right)$ steht und somit $e_b = 10.000$ entspricht.

Der Median ist das geeignete Maß für diesen Fall. Der Mittelwert eignet sich für symmetrische (z.B. normalverteilte) Daten. Beim Mittelwert fallen Ausreißer stark ins Gewicht (hier: der Millionär). Die Beschreibung über den Median ist bei schiefen Daten sinnvoller.

Aufgabe 10

♀

Stem	Leaves
0	
1	
2	9
3	1 8
4	6
5	
6	0 3 3
7	0 1 1 8
8	2 5 6
9	9 9 9
10	

♂

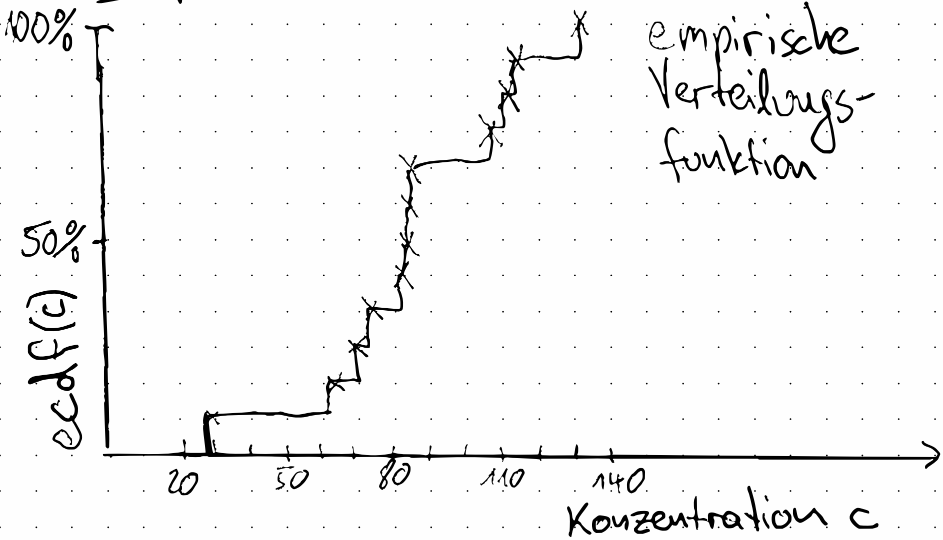
Stem	Leaves
0	
1	
2	
3	
4	
5	0
6	8 8
7	0 8
8	3 4 5 9
9	2 2 4 5 6
10	0 0 0

Deutlich zu sehen ist, dass die Verteilung der Alphabetisierungsgrade für Männer in höhere Prozentbereiche gelangt. Für Frauen gibt es einige Extremwerte im unteren Spektrum.

Leider <u>nicht</u> zu sehen ist die Relation zwischen den Datenpunkten über die beiden Stem-and-Leaf Plots. Z.B. dass die Werte bei den Frauen (99, 99 und 99 %) zu den Ländern gehören, in welchen die Männer 100, 100 und 100 % Alphabetisierung erreichen.

Aufgabe 11

$\boxed{a}$

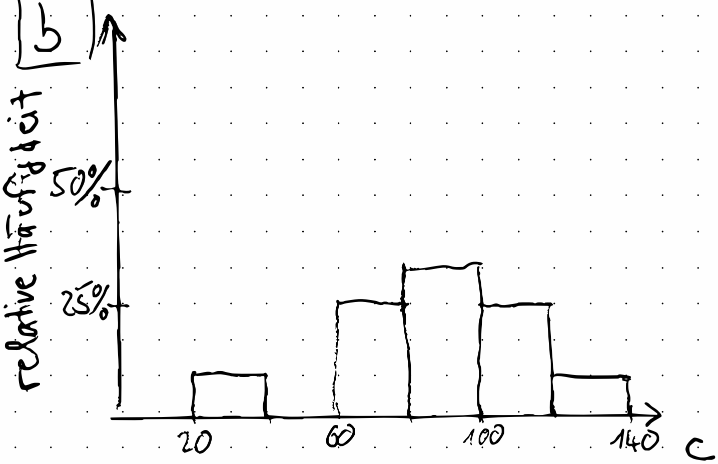


empirische Verteilungs-funktion

sortiert

29,63 x_1
63,79
70,05 x_3
73,30
82,76 x_5
83,07
83,60 x_7
85,68
108,32 x_9
112,97
115,18 x_{11}
132,53

$\boxed{b}$



$\boxed{C}$ $\min(c) = 29{,}63$

$\max(c) = 132{,}53$

$\text{median}(c) = 83{,}335$ (Da Anzahl der Elemente gerade, Mittelwert aus den beiden mittleren Werten der sortierten Liste)

$$\text{mean}(c) = \frac{1}{n} \sum_{i=1}^{n} c_i = 86{,}73167$$

$\tilde{x}_{0,25} \Rightarrow$ da $n \cdot p \in \mathbb{N}$ für $n = 12$ und $p = 0{,}25$

gilt $\tilde{x}_{0,25} = \frac{1}{2}(x_3 + x_4)$

$= \frac{1}{2}(70{,}05 + 73{,}30) = \frac{1}{2}(143{,}35)$

$= 71{,}675$ R: 72,4875

$\tilde{x}_{0,75} \Rightarrow$ gleiches gilt wie oben:

$= \frac{1}{2}(x_9 + x_{10})$

$= \frac{1}{2}(108{,}32 + 112{,}87)$

$= 110{,}595$ R: 109,4575

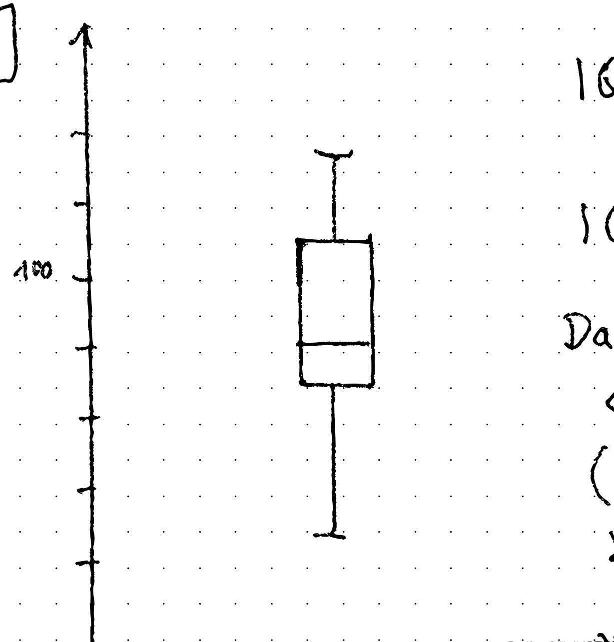
d

$$s^2 = s_n^2 = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \bar{x})^2$$

$$= \frac{1}{11} \sum_{k=1}^{12} (x_k - 86{,}73167)^2$$

$$= 758{,}3425$$

e



$IQR = 110{,}6 - 71{,}7$
$= 38{,}9$

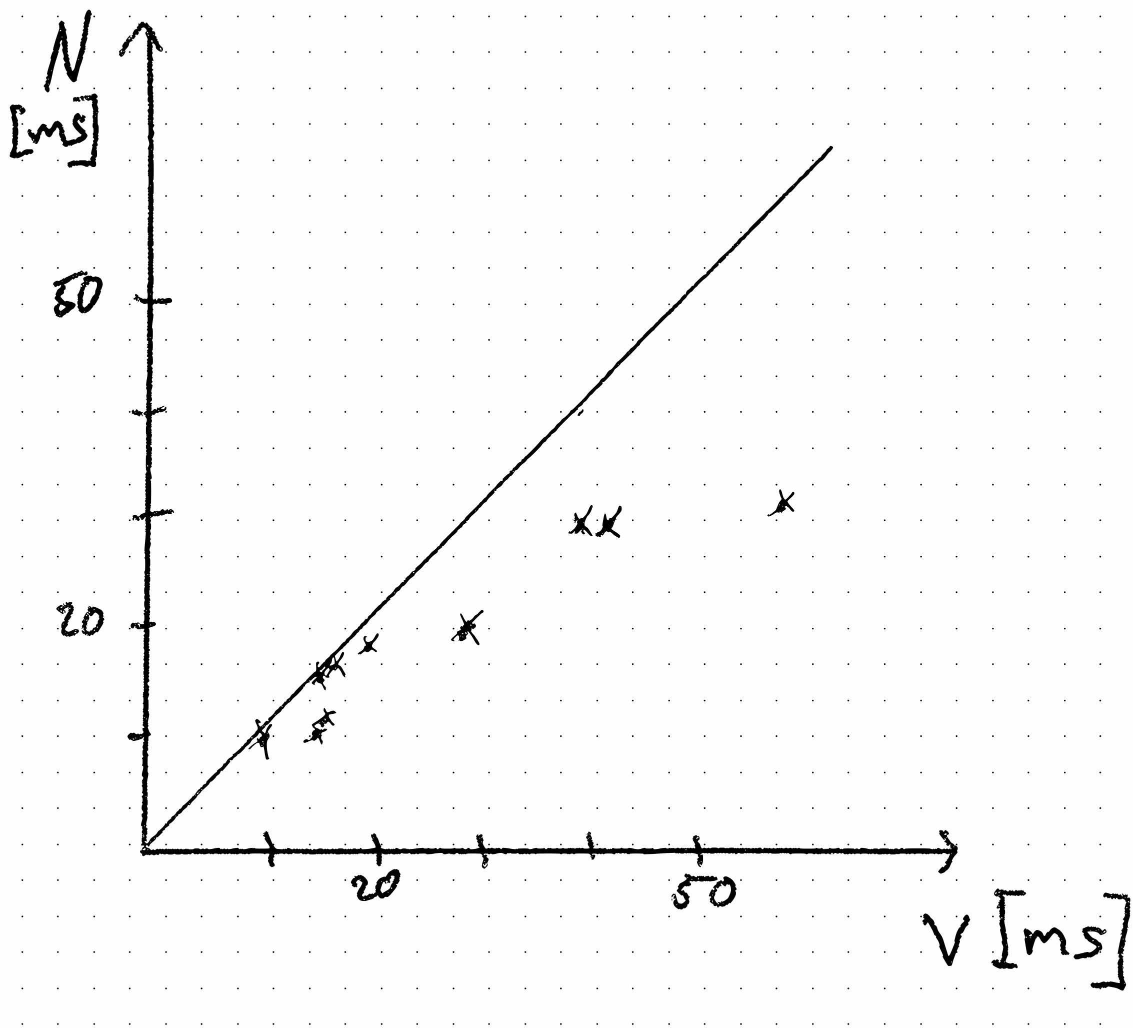
$IQR \cdot 1{,}5 = 58{,}35$

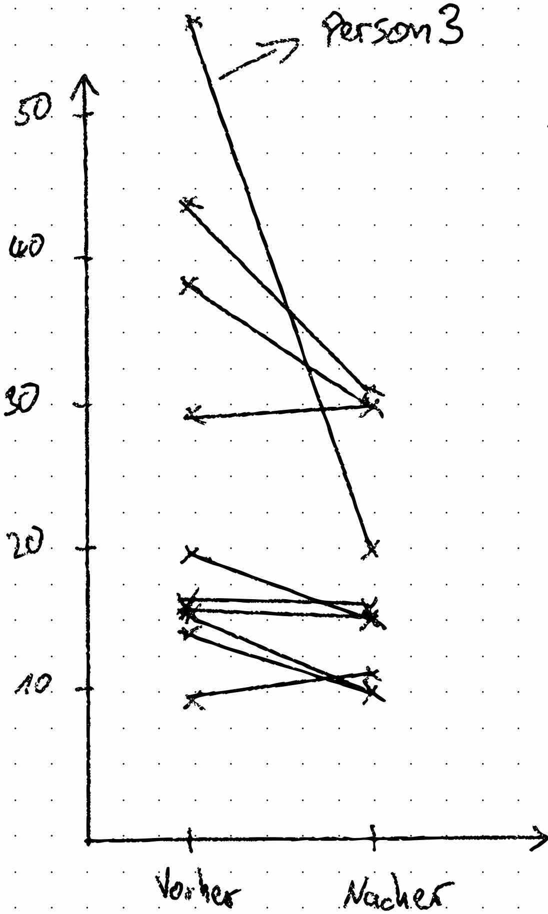
Da $(\max(x) - IQR \cdot 1{,}5) < \tilde{x}_{0{,}75}$ und $(\min(x) + IQR \cdot 1{,}5) > \tilde{x}_{0{,}25}$ sind, werden keine Ausreiser eingezeichnet.

Aufgabe 12

$V = \{9, 14, 15, 16, 17, 19, 29, 38, 43, 57\}$

$N = \{10, 10, 11, 15, 16, 17, 20, 30, 30, 32\}$





Trägt man die Probanden paarweise auf, sieht man die Tendenz zur Verkleinerung der Reaktionszeit nach Training. Jedoch sieht man auch einen großen Ausreiser, welcher große Einflüsse auf statistische Maße nehmen kann.