

**STATISTIK I – Lösung 07****Box-Plots und Stem-and-Leaf-Diagramme**

Eine Befragung von 20 willkürlich ausgewählten Professorinnen und Professoren der Hochschule Harz nach ihrem Alter ergab folgendes Bild:

Prof.	Alter	Prof.	Alter
1	44	11	48
2	61	12	56
3	62	13	66
4	54	14	53
5	55	15	39
6	50	16	42
7	51	17	46
8	44	18	45
9	40	19	60
10	33	20	52

1) Konstruieren Sie einen erweiterten Box-Plot.

Zur Berechnung der Quartile ordnen wir im ersten Schritt die Werte der Verteilung:

33; 39; 40; 42; 44; 44; 45; 46; 48; 50; 51; 52; 53; 54; 55; 56; 60; 61; 62; 66

Ergibt  $(n \cdot p)$  einen ganzzahligen Wert  $(k)$ , berechnet sich das Perzentil wie folgt:

$$x_p = \frac{1}{2}(x_{(k)} + x_{(k+1)})$$

$$(n \cdot p) = (20 \cdot 0,25) = 5 \rightarrow k = 5; k+1 = 6 \rightarrow x_p = (44+44)/2=44$$

$$(n \cdot p) = (20 \cdot 0,50) = 10 \rightarrow k = 10; k+1 = 11 \rightarrow x_p = (50+51)/2=50,5$$

$$(n \cdot p) = (20 \cdot 0,75) = 15 \rightarrow k = 15; k+1 = 16 \rightarrow x_p = (55+56)/2=55,5$$

$$\text{IQR} = 55,5 - 44 = 11,5$$

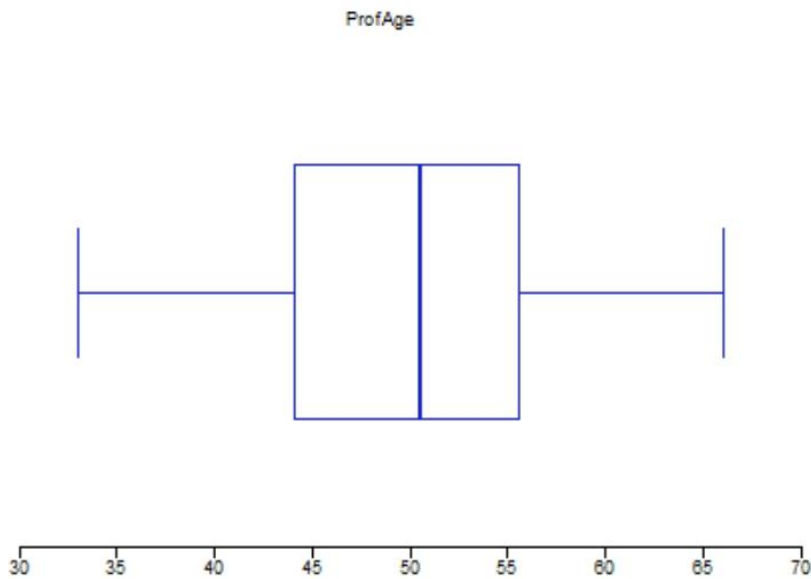
$$1,5 \text{ IQR} = 17,25$$

$$3,0 \text{ IQR} = 34,50$$

Auf Basis dieser Ergebnisse lässt sich der Box-Plot nun wie gewohnt konstruieren:

- Die Box verläuft von 44 (unteres Quartil) bis zu 55,5 (oberes Quartil).
- Der Median (mittleres Quartil) wird in diese Box beim Wert 50,5 eingezeichnet.
- Der obere Zaun könnte maximal bis zum Wert  $55,5 + 17,25 = 72,75$  reichen. Der größte in dem Bereich liegende Wert (und damit die Grenze des oberen Zauns) ist die 66. Da sich im Datensatz keine größeren Werte befinden, gibt es nach oben keine Ausreißer.
- Der untere Zaun könnte minimal bis zum Wert  $44 - 17,25 = 26,75$  reichen. Der kleinste in diesem Bereich liegende Wert (und damit die Grenze des unteren Zauns) ist die 33. Da sich im Datensatz keine kleineren Werte befinden, gibt es nach unten keine Ausreißer.
- Zu Übungszwecken sei an dieser Stelle noch aufgeführt, dass Ausreißer nach oben zwischen 72,75 und (einschließlich) 90, Ausreißer nach unten zwischen 26,75 und (einschließlich) 9,5 hätten liegen müssen. Werte oberhalb von 90 bzw. unterhalb von 9,5 hätten dagegen als Extremwerte gekennzeichnet werden müssen.

Die Kontrollgrafik wurde in diesem Fall nicht mit PAST, sondern mit SSP (seitlich liegend) erstellt.



2) Konstruieren Sie ein Stem-and-Leaf-Diagramm.

Für das Stem-and-Leaf-Diagramm können wir auf die geordnete Verteilung zurückgreifen:

33; 39; 40; 42; 44; 44; 45; 46; 48; 50; 51; 52; 53; 54; 55; 56; 60; 61; 62; 66

Da alle Zahlen zwischen 30 und 70 liegen (keine Ausreißer) und sich zudem kaum Werte wiederholen, bietet sich eine Stammbreite von 10 sowie die Eintragung eines Blatts je Wert der Verteilung an. Das sich ergebende Stem-and-Leaf-Diagramm sieht dann wie folgt aus:

```

3 | 3 9
4 | 0 2 4 4 5 6 8
5 | 0 1 2 3 4 5 6
6 | 0 1 2 6

```

Stammbreite: 10  
 Ein Wert pro Blatt  
 Keine Ausreißer

Zu Übungszwecken: So sieht das Stem-and-Leaf-Diagramm bei einer Stammbreite von 5 aus:

```

3 | 3
3 | 9
4 | 0 2 4 4
4 | 5 6 8
5 | 0 1 2 3 4
5 | 5 6
6 | 0 1 2
6 | 6

```

Stammbreite: 5  
 Ein Wert pro Blatt  
 Keine Ausreißer