

Lösungsskizze Übung 1

1. Alter der Frauen sortiert:

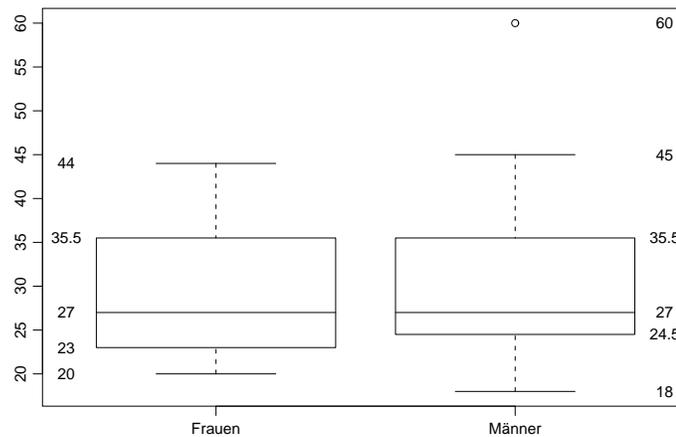
20	21	22	23	23	24	25	27	27	27	34	37	38	42	44
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Alter der Männer sortiert:

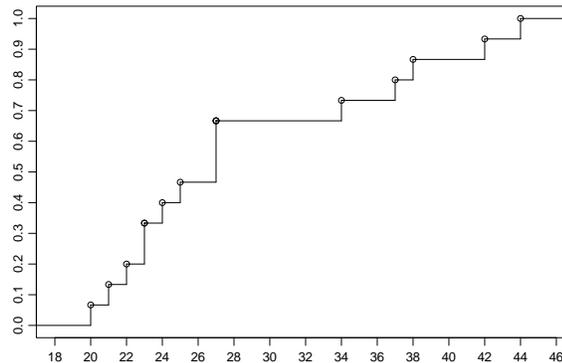
18	22	23	23	26	26	26	27	31	32	35	36	36	45	60
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

a) Für die Frauen haben wir: Median: 27, (interpoliertes) 1.Quartil: 23, (interpoliertes) 3.Quartil: 35.5. Für die Männer: Median: 27, (interpoliertes) 1.Quartil: 24.5, (interpoliertes) 3.Quartil: 35.5.

b) Die beiden Boxplots sehen folgendermassen aus:



c) Und die kumulative Verteilungsfunktion für die Daten der Frauen:



d) Das arithmetische Mittel ist bei den Männern tatsächlich höher (31.1), als bei den Frauen (28.9). Der Wert bei den Männern wird aber sehr stark durch das Alter des Mannes in Paar 15 beeinflusst. Betrachtet man aber den Median als Lagemass, stellt man fest, dass der bei beiden Gruppen gleich ist. Zudem, würde man das 15te Paar weglassen, würden beide Boxplots etwas gleich aussehen. Die Schlussfolgerung der Zeitung stützt sich also scheinbar allein auf das arithmetische Mittel, ohne die Daten genauer anzusehen.

e) $P[X < 25] = 0.4$

$$P[30 < X < 35] = P[X < 35] - P[X < 30] = 0.73 - 0.66 = 0.07.$$

2. a) **A2:** Sehr kleine Streuung, Median ein bisschen gegen die tiefen Werte verschoben.
B5: Lage gegen kleinere Werte hin verschoben, extreme Werte gegen oben, positive Schiefe.
C1: Grösste Streuung, Median gegen kleine Werte hin verschoben, keine extremen Werte.
D4: Lage gegen grössere Werte hin verschoben, extreme Werte gegen unten, negative Schiefe.
E3: Fast symmetrisch, grössere Streuung als A2, aber weniger extreme Werte als B5 und D4.
- b) Die fünf Stichproben lassen sich wie folgt zuordnen:
E4: 1.Quartil ist grösser als 6, alle andern sind kleiner.
C1 / D3: **C** und **D** haben beide einen Median wenig über 5. Die Stichprobe **1** hat aber eine viel kleinere Varianz als Stichprobe **3**.
A2 / B5: **A** und **B** haben fast gleichen Median und fast gleiches 3.Quartil, das 1.Quartil liegt aber bei **B** höher und **B** hat zudem extremere Werte gegen oben als **A**. Daher ist der Mittelwert bei **B** grösser.
Die gesuchten Kenngrössen kann man durch Vergleichen mit den gegebenen und der Achse schätzen:

	C	A	D	E	B
Stichprobe	1	2	3	4	5
Mittelwert	5.21	4.52	5.14	7.02	5.07
Varianz	0.84	2.18	2.91	0.76	2.03
1.Quartil	4.34	3.54	3.61	6.69	4.01
Median	5.16	4.65	5.13	7.14	4.74
3.Quartil	6.13	5.59	6.37	7.69	5.38

3. a) Es wurde jeweils bei den Männern und Frauen dieselbe Skala gewählt, damit man besser vergleichen kann.

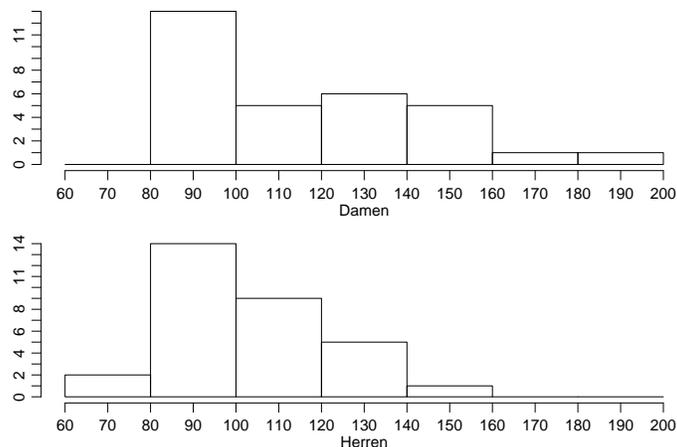


Abbildung 1: Histogramme der SiegerInnenzeiten in Minuten

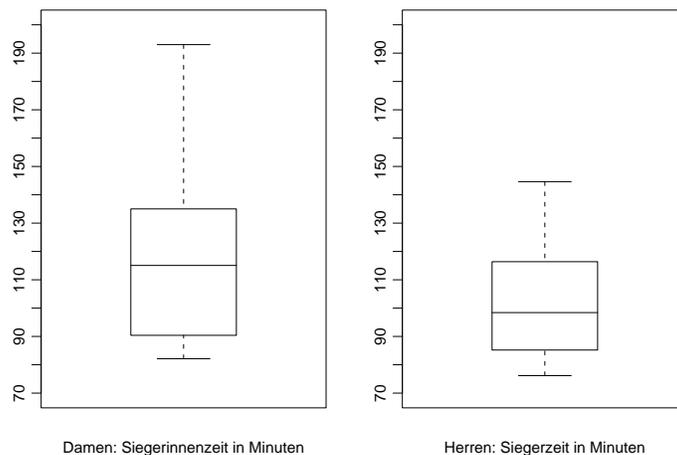
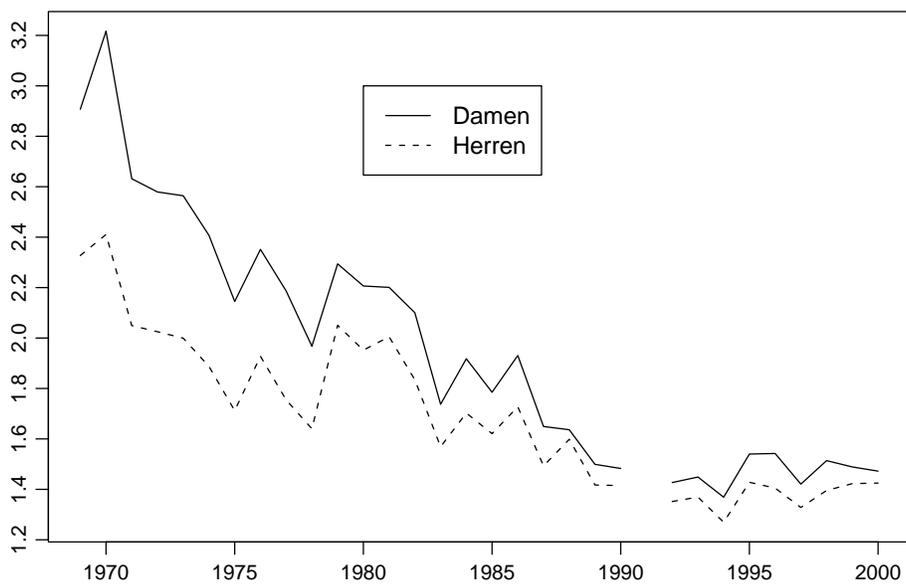


Abbildung 2: Boxplots der SiegerInnenzeiten in Minuten

- b) Die besten Zeiten bei den Herren (unter 80 Min.) sind besser als die besten Zeiten der Damen (über 80 Min.) und die langsamste Siegerzeit ist bei den Frauen (ungefähr 190 Min.) schlechter als bei den Männern (ungefähr 140 Min.). Dass Frauen etwas langsamer sind als Männer ist allerdings keine besonders grosse Überraschung. Erstaunlicher ist, dass auch die Schwankungen bei den Frauen grösser sind. Kann sein, dass dies daran liegt, dass die Frauen früher noch wenig trainiert haben und viel schlechter waren als heute, aber das kann man aus diesen Graphiken nicht sehen.
- c) Aus dem obenstehenden Plots ist nicht ersichtlich, wie sich die Zeiten im Verlauf der Jahre verändert haben. Dies ist jedoch eine der interessantesten Fragestellungen und sollte unbedingt mituntersucht werden.

- d) Die Siegeszeiten wurden im Laufe der Jahre tendenziell immer besser. Wie in a) schon vermutet, gleichen sich die Zeiten der Damen und Herren immer mehr an. Die Verbesserungen sind besonders in den Jahren 1969-1990 ersichtlich. 1991 konnte der Engadiner wegen eines Wärmeeinbruchs nicht durchgeführt werden.



4. Daten der Grösse nach geordnet:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
547	710	721	740	850	895	945	961	980	1018	1075	1090	1150	1358	1480

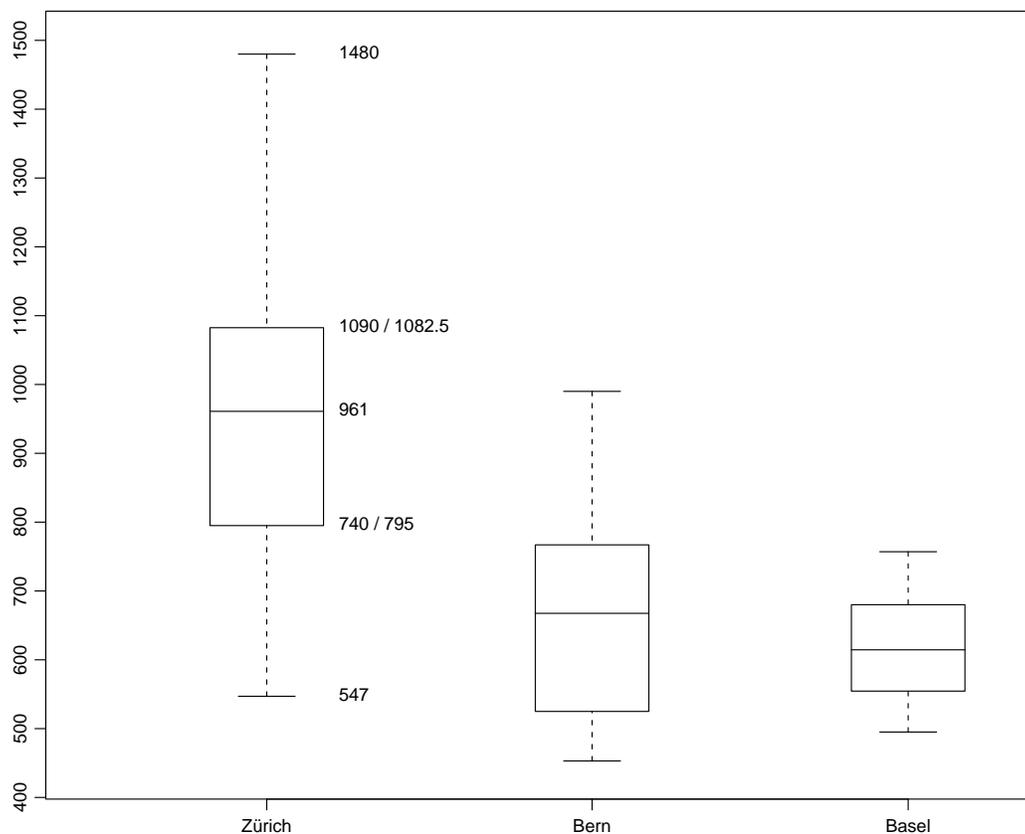
a) Median: 961.

1.Quartil: 740 oder 795 (je nach Berechnungsmethode)

3.Quartil: 1090 oder 1082.5 (je nach Berechnungsmethode)

Aber: Beide Quartile sollten mit der gleichen Methode berechnet werden, d.h. entweder hat man die beide ersten Werte oder die beiden zweiten, aber nicht gemischt.

b)



c) Die wichtigsten Erkenntnisse sind:

- Die billigste Wohnung befindet sich in Bern, die teuerste in Zürich.
- Der Median der Preise in Zürich liegt fast so hoch, wie die teuerste Wohnung in Bern, d.h. also fast 50% der Wohnungen in Zürich sind teurer als die Teuerste in Bern.
- Die Streuung der Preise ist in Zürich am Grössten, die billigste Wohnung liegt etwa auf dem Niveau des 1.Quartils von Bern, die Teuerste weit, weit oberhalb der Teursten von Bern und Basel.

- Basel hat eine sehr kleine Streuung der Preise, alle Preise liegen im Bereich zwischen dem 1. und 3. Quartil der Preise von Bern.
- d)
- Der Boxplot von Basel ist fast symmetrisch (obere “Antenne” ist ein klein bisschen länger). Also muss das arith. Mittel ungefähr dem Median sein: ca. 620.
 - Auch bei Bern ist das arithm. Mittel ungefähr dem Median: Die Box oberhalb des Median ist zwar kleiner als die unterhalb, dafür ist der obere Stab länger als der untere, so dass sich die Werte wieder ungefähr ausmitteln. Daher schätzen wir mal: 670.
 - Als Anhaltspunkte kennen wir die Standardabweichungen von Bern und Basel. Im Boxplot selber, ist die Streuung durch die Distanz der Enden der “Antennen” (und allfälliger extremer Punkte) zum Median visualisiert. Gehen wir einmal ganz banal vor: Die Distanz vom Median zu den Enden der Stäbe von Zürich ist gerade ungefähr die Distanz vom Median zu den Enden der Stäbe von Bern plus die Distanz vom Median zu den Enden der “Antennen” von Basel, also $79 + 163 = 242$.

Unten folgt die Tabelle mit den wahren Werten. Mit unseren Schätzungen liegen wir also ziemlich genau! In Klammern steht noch der Bereich der “vernünftigen” Schätzungen:

Stadt	Median	arithm. Mittel		Standardabweichung
Zürich	a)	968		246.83 [200 - 300]
Bern	667.5	674.08	[620 - 720]	163.60
Basel	614.5	621.25	[600 - 650]	78.88