

LÖSUNG 2D

- Bei der Berechnung des Einkommens je Beschäftigungsstunde ist zu berücksichtigen, dass das Einkommen je Monat und die Arbeitsstunden je Woche ausgewiesen werden. Daher ergibt sich die Gleichung ($EINKJESTD = EINK/ARBSTD/4$). Die Berechnung erfolgt im Menü "Transformieren", "Variable berechnen" durch Eingabe des Variablennamens im Feld „Zielvariable:“ und der Gleichung im Feld „Numerischer Ausdruck:“.
- Eine getrennte Berechnung des durchschnittlichen Einkommens je Arbeitsstunde für Befragte aus Gemeinden unter 50 Tsd. Einwohner (und später ab 50 Tsd. Einwohner) erreicht man im Menü "Daten", "Fälle auswählen" und Wahl von "Falls Bedingung zutrifft", in dem man nach Klicken der Schaltfläche "Falls" die Bedingung $GEM \leq 4$ (Gemeinden unter 50 Tsd. Einwohner) formuliert und aktiviert. Nachdem man diese erste Bedingung formuliert hat, kann man im Menü "Deskriptive Statistik" mit Auswahl der Option "Mittelwert" und Auswahl der Variablen EINKJESTD das arithmetische Mittel berechnen. Danach werden die Gemeinden ab 50 Tsd. Einwohner analog ausgewählt $GEM \geq 5$ (Gemeinden ab 50 Tsd. Einwohner) und anschließend ebenfalls mit im Menü "Deskriptive Statistik" mit Auswahl der Option "Mittelwert" das arithmetische Mittel berechnet. (Alternativ: man kodiert GEM in eine neue Variable um (z.B. GEM1) mit den Werten 1 (alte Werte: 1 bis 4) und 2 (alte Werte: 5 bis größter). Mit "Daten", "Aufgeteilte Datei", "Gruppen vergleichen" (oder "Ausgabe nach Gruppen aufteilen") und Eingabe der neuen Variable GEM1 in das Eingabefeld legt man eine getrennt Auswertung für die Gruppen fest. Danach kann man wieder im Menü "Deskriptive Statistik" und Auswahl der Option "Mittelwert" das arithmetische Mittel berechnen).

Durchschnittliches Einkommen je Arbeitsstunde:

- Für Gemeinde mit einer Größe kleiner 50 Tsd. Einwohner:

	N	Mittelwert
EinkjeStd = eink/arbstd/4	44	15,3525
Gültige Werte (Listenweise)	44	

- Für Gemeinde mit einer Größe gleich bzw. größer 50 Tsd. Einwohner:

	N	Mittelwert
EinkjeStd = eink/arbstd/4	32	16,2732
Gültige Werte (Listenweise)	32	

Boxplotdiagramm:

- Erzeugen: „Grafik“, „Diagrammerstellung...“, „Boxplot“ wählen. Durch Doppelklicken auf das Symbol „Gruppiertes Boxplot“ dieses in die Diagrammvorschau übertragen. SCHUL2 auf „X-Achse?“, EINKJESTD auf „Y-Achse?“ und GESCHL auf „Clustervariable auf X: Farbe festlegen“ ziehen.
- Die Boxen des Diagramms werden durch das erste (Q_1) und dritte Quartil (Q_3) begrenzt. Folglich liegen im Boxenbereich 50 % der Fälle. Der schwarze Balken in der Box ist der Median (Zentralwert, Q_2). Die Lage des Zentralwerts in der Box gibt Auskunft über die Schiefe der Ver-

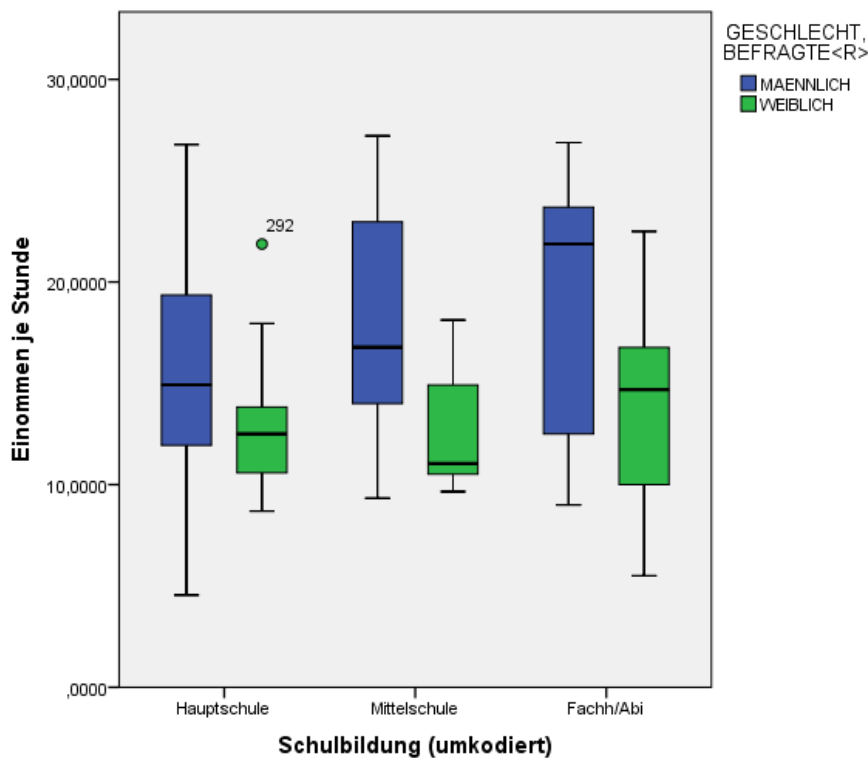
teilung. Bei z. B. weiblichen Erwerbstätigen mit mittlerer Reife ist die Verteilung von EINKJESTD ausgeprägt linkssteil.

Die feine Linie mit dem Querbalken ("whisker") dient dazu, die größten und kleinsten Werte (mit Ausnahme von Ausreißern und Extremwerten) anzuzeigen.

Ausreißer (Fälle mit Werten, die 1,5 bis 3 Boxenlängen vom unteren bzw. oberen Boxenrand entfernt sind) und Extremwerte (Fälle mit Werten, die mehr als 3 Boxenlängen vom unteren bzw. oberen Boxenrand entfernt sind) werden mit der Fallnummer oder mit dem Wert der Fallbeschriftungsvariable angezeigt. Fall 292 zeigt für weibliche Befragte mit Hauptschulabschluss einen Ausreißer.

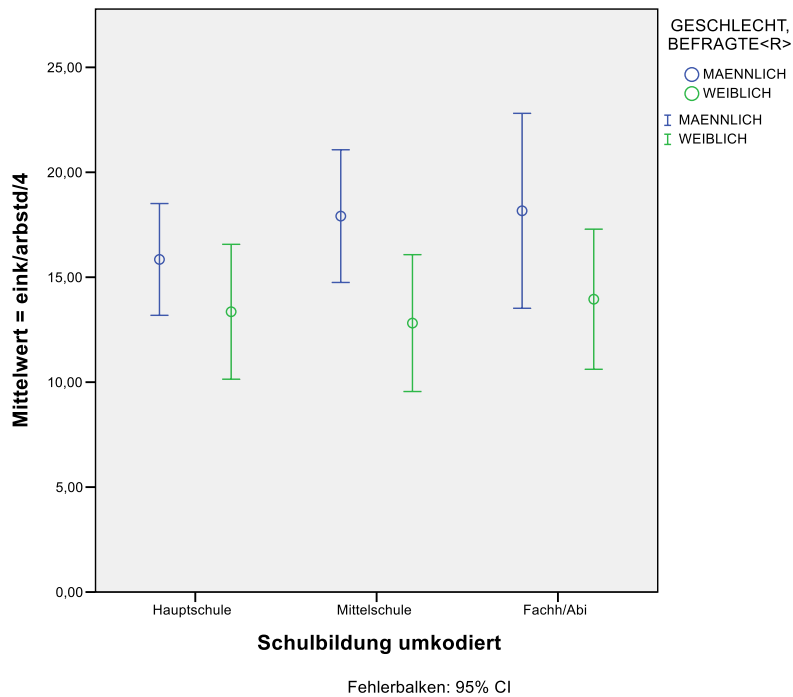
Im Boxplot werden die mittleren Werte sowie die Streuungen von EINKSTD für die einzelnen Gruppen grafisch veranschaulicht und damit vergleichend sichtbar gemacht. Man sieht z.B., dass der Median von EINKJESTD für Männer mit höherem Schulabschluss steigt und die Streuung von EINKJESTD bei Männern größer ist als bei Frauen.

Ein Boxplotdiagramm eignet sich auch gut zum Auffinden von Extremwerten und Ausreißern (s. Lösung zu Aufgabe 2c).



Fehlerbalkendiagramm:

- Erzeugen des Diagramms: „Grafik“, „Diagrammerstellung...“. „Balken“ wählen, durch Doppelklicken auf das Symbol für „Gruppierte Fehlerbalken“ dieses in die Diagrammvorschau übertragen. SCHUL2 auf „X-Achse?“, EINKJESTD auf „Y-Achse?“ und GESCHL auf „Clustervariable auf X: Farbe festlegen“ ziehen.



- Für die unterschiedlichen Schulabschlüsse, untergliedert nach dem Geschlecht, werden das arithmetische Mittel von EINKJESTD (angezeigt durch einen Kreis auf dem Fehlerbalken) sowie ein oberer und unterer Grenzwert als Balken dargestellt. Die Grenzwerte sind 95 %-Konfidenzintervalle für den Mittelwert $[\bar{x} \pm t_{\alpha/2, fg=n-1} * s_x / \sqrt{n}]$, entspricht der Standardeinstellung (angezeigt durch „Error Bars: 95 % CI, CI = confidence interval“). Es können andere Grenzen gewählt werden, s. auch Lösung zu Aufgabe 2f und Kap. 29.3].