

# Wirtschaftsstatistik

## Übungsblatt mit Lösungen Modul 4 & 5

### Lageparameter & Streuungsparameter

#### Aufgabe 1

- a) Worüber informiert ...
- a. ... die Standardabweichung?
  - b. ... das Quartil Q3?
  - c. ... der Variationskoeffizient?
  - d. ... der Interquartilsabstand?
  - e. ... das 5%-Quantil?
  - f. ... der Median?
  - g. ... der Modus?
  - h. ... die Spannweite?
  - i. ... das Quartil Q1?
- b) Was haben die statistischen Parameter Varianz, Standardabweichung und Variationskoeffizient gemeinsam? Wodurch unterscheiden sich die drei Parameter?

#### Lösung der Aufgabe 1

- a)
- a. Die Standardabweichung  $s$  informiert über die (absolute) Streuung einer Verteilung. Die Standardabweichung hat die gleiche Maßeinheit wie die Beobachtungswerte und die Mittelwerte.
  - b. Das Quartil Q3 informiert darüber, welcher Wert von 75% der Beobachtungswerte einer Verteilung unterschritten wird und von 25% der Beobachtungswerte überschritten wird.
  - c. Der Variationskoeffizient  $v$  informiert über die relative Streuung einer Verteilung.
  - d. Der Interquartilsabstand informiert über die Spannweite der mittleren 50% der Beobachtungswerte einer Verteilung.
  - e. Das 5%-Quantil informiert über den Wert, der von 5% der Beobachtungswerte einer Verteilung unterschritten wird und von 95% der Beobachtungswerte überschritten wird.
  - f. Der Median informiert über den Wert, der von der Hälfte der Beobachtungswerte überschritten und von der anderen Hälfte unterschritten wird.

- g. Der Modus informiert darüber, welche Merkmalsausprägung(en) in der Häufigkeitsverteilung am häufigsten vorkommen.
  - h. Die Spannweite informiert über den maximalen Abstand der Merkmalsausprägungen, die in der Häufigkeitsverteilung vorkommen. Sie ergibt sich aus der Differenz zwischen dem größten und kleinsten Merkmalswert der Verteilung.
  - i. Das Quartil  $Q_1$  informiert darüber, welcher Wert von 25% der Beobachtungswerte einer Verteilung unterschritten wird und von 75% der Beobachtungswerte überschritten wird.
- b) Varianz, Standardabweichung und Variationskoeffizient sind Streuungsparameter. Mit ihnen lässt sich die Streuung einer Häufigkeitsverteilung charakterisieren. Varianz und Standardabweichung informieren über die **absolute** Streuung, der Variationskoeffizient über die **relative** Streuung.

### Aufgabe 2

In einer Statistik über die Einkommen (Jahreseinkommen) von leitenden Angestellten im Rechnungswesen lesen Sie:

1. Quartil  $Q_1$ : 75.000 €, 3. Quartil  $Q_3$ : 150.000 €, Median  $\bar{x}_Z$ : 100.000 €, Mittelwert  $\bar{x}$ : 140.000 €.

Welche der folgenden Aussagen über die Einkommensverteilung sind richtig? (Zutreffendes ankreuzen!)

### Lösung der Aufgabe 2

	RICHTIG	FALSCH
(1) Die Einkommensverteilung ist linksschief.		X
(2) 25% der Befragten verdienen weniger als 75.000 €.	X	
(3) 50% der Befragten verdienen zwischen 75.000 und 150.000 €.	X	
(4) 75% der Befragten verdienen mehr als 150.000 €.		X
(5) Die Einkommensverteilung ist symmetrisch.		X
(6) 50% der Befragten verdienen weniger als 140.000 €.		X

### Aufgabe 3

Erstellen Sie auf der Basis der folgenden Angaben eine eindimensionale, unklassierte Häufigkeitsverteilung:

In einer Stadt haben vier Taxi-Unternehmen jeweils drei Wagen, ein Taxi-Unternehmen hat 26 Wagen. Die übrigen Taxi-Unternehmen in der Stadt sind kleiner und haben weniger

Wagen: acht Taxi-Unternehmen haben jeweils nur einen Wagen, sieben haben jeweils zwei Wagen.

Beantworten Sie auf der Basis der erstellten Tabelle die folgenden Fragen (mit 1 Nachkommastelle):

- a) Wie viele Taxi-Unternehmen gibt es in dieser Stadt?
- b) Wie viele Wagen bieten in dieser Stadt ihre Leistungen an?
- c) Bestimmen / berechnen Sie für die obige Verteilung
  - a. den Modus
  - b. den Median
  - c. das arithmetische Mittel
- d) Machen Sie Aussagen über die Schiefe der Verteilung
- e) Berechnen Sie die Standardabweichung
- f) Wie viel % der Taxi-Unternehmen haben mehr als einen Wagen
- g) Wie viel % der Taxi-Unternehmen haben weniger als 3 Wagen

Tipp: Bitte beachten Sie, dass in dieser Aufgabe Unternehmen nach der Anzahl ihrer Wagen

(=Merkmalsausprägung) charakterisiert werden und somit werden die Unternehmen gezählt, nicht die Wagen.

### Lösung der Aufgabe 3

i	Anzahl Wagen $x_i$	Anzahl Taxi- Unternehmen $h(x_i)$	$f(x_i)$ (%)	$H(x_i)$	$F(x_i)$ (%)
1	1	8	40,0%	8	40,0%
2	2	7	35,0%	15	75,0%
3	3	4	20,0%	19	95,0%
4	26	1	5,0%	20	100,0%
$\Sigma$		20	100,0%	-	-

- a) Es gibt 20 Taxi-Unternehmen in dieser Stadt.
- b) 60 Wagen stehen in der Stadt zur Verfügung ( $1 \cdot 8 + 2 \cdot 7 + 3 \cdot 4 + 26 \cdot 1 = 60$ ).
- c)

a.  $\bar{x}_D = 1$  (da Anz. Unternehmen mit 1 Wagen am höchsten ist)

b.  $\bar{x}_Z = 2$  (s. Folie 11 im Modul 4)

Reihe: 1 1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 26

c.

$$\bar{x} = \frac{1 \cdot 8 + 2 \cdot 7 + 3 \cdot 4 + 26 \cdot 1}{20} = \frac{60}{20} = 3$$

d) rechtsschiefe Verteilung

e)

$$\begin{aligned}
 s^2 &= \frac{1}{20} * ((1 - 3)^2 * 8 + (2 - 3)^2 * 7 + (3 - 3)^2 * 4 + (26 - 3)^2 * 1) \\
 &= \frac{1}{20} * (32 + 7 + 0 + 529) = \frac{568}{20} = 28,4 \\
 s &= \sqrt{s^2} = \sqrt{28,4} = 5,33
 \end{aligned}$$

f)  $100 - 40 = 60\%$

g)  $40 + 35 = 75\%$

#### Aufgabe 4

- a) Sie lesen in einer Studie über die Altersverteilung in einer Gruppe, dass  $\bar{x}_Z = 32$  Jahre und  $\bar{x} = 40$  Jahre ist. Welche Schlüsse können Sie daraus über die Altersverteilung ziehen?
- b) Sie lesen in einer Studie über die Einkommensverteilung einer Berufsgruppe:  $\bar{x}_Z = 30.000$  €,  $\bar{x} = 40.000$  €,  $Q_1 = 25.000$  €,  $Q_3 = 45.000$  €. Welche Informationen erhalten Sie aus diesen 4 statistischen Kennzahlen über die Einkommensverteilung? Erhalten Sie auch Informationen über die Streuung der Verteilung?

#### Lösung der Aufgabe 4

- a) Wegen  $\bar{x}_Z = 32 < 40 = \bar{x}$ , liegt hier eine rechtsschiefe Altersverteilung vor. Es gibt Ausreißer im oberen Altersbereich, die das arithmetische Mittel nach oben ziehen, den Median aber nicht beeinflussen.
- b) Die Einkommensverteilung ist rechtsschief, da  $\bar{x}_Z < \bar{x}$ . Es gibt Ausreißer im oberen Einkommensbereich.  
 25% der Personen verdienen weniger und 75% mehr als 25.000 €. 75% verdienen weniger und 25% mehr als 45.000 €. Ein Einkommen von 30.000 € wird von 50% über- und von den anderen 50% unterschritten.  
 Der Interquartilsabstand  $IQR = Q_3 - Q_1 = 20.000$  € gibt die Spannweite bei den mittleren 50% an und informiert somit über die Streuung des "mittleren" Teils der Einkommensverteilung.

#### Aufgabe 5

Auf die Frage „Wie viel Stück des Produktes ABC haben Sie im letzten Monat gekauft?“ gab es bei der Hauptuntersuchung unterschiedliche Antworten zur Zahl der gekauften Stücke von ABC. Die statistische Reihe wurde zusammengefasst in der folgenden

Häufigkeitsverteilung:

- a) Bestimmen Sie den Modus, den Median und das arithmetische Mittel.

- b) Bestimmen Sie die Streuungsparameter Spannweite  $w$ , Varianz  $s^2$ , Standardabweichung  $s$ , den Variationskoeffizient  $v$
- c) Welche Aussagen können Sie auf der Basis der Werte der Lageparameter über die Form der Verteilung machen?

Lösung der Aufgabe 5

i	gekaufte Stückzahl von ABC $x_i$	Anzahl der Nennungen $h(x_i)$	Anteil $f(x_i)$ (%)	H( $x_i$ )	F( $x_i$ ) (%)	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 h(x_i)$
1	0	100	20%	100	20%	1,44	144
2	1	300	60%	400	80%	0,04	12
3	2	50	10%	450	90%	0,64	32
4	3	20	4%	470	94%	3,24	64,8
5	4	20	4%	490	98%	7,84	156,8
6	5	5	1%	495	99%	14,44	72,2
7	7	5	1%	500	100%	33,64	168,2
$\Sigma$		500	100%				650

a)  $\bar{x}_D = 1 \quad \bar{x}_Z = 1$

$$\bar{x} = \frac{0 \cdot 100 + 1 \cdot 300 + 2 \cdot 50 + 3 \cdot 20 + 4 \cdot 20 + 5 \cdot 5 + 7 \cdot 5}{500} = \frac{600}{500} = 1,2$$

b)  $w = x_{\max} - x_{\min} = 7 - 0 = 7$

1. Variante (Formel (2), s.Folie 20 Modul 5)

$$\begin{aligned}
 s^2 &= \left( \sum_{i=1}^j x_i^2 * f(x_i) \right) - \bar{x}^2 \\
 &= (0 + 1 \cdot 0,6 + 4 \cdot 0,1 + 9 \cdot 0,04 + 16 \cdot 0,04 + 25 \cdot 0,01 + 49 \cdot 0,01) - 1,2^2 \\
 &= 0,6 + 0,4 + 0,36 + 0,64 + 0,25 + 0,49 - 1,44 = 1,3
 \end{aligned}$$

2. Variante (Formel (1), s. Folie 20 Modul 5)

$$s^2 = \frac{1}{n} \sum_{i=1}^j (x_i - \bar{x})^2 * h(x_i) = \frac{650}{500} = 1,3$$

$$s = \sqrt{s^2} = \sqrt{1,3} = 1,14$$

$$v = \frac{s}{\bar{x}} = 1,14/1,2 = 0,95$$

c) Die Verteilung ist rechtsschief wegen  $\bar{x}_Z = 1 < 1,2 = \bar{x}$

Aufgabe 6

Für 200 Unternehmen liegt für das Jahr 2006 die folgende Umsatzverteilung vor.

Vervollständigen Sie die klassierte Häufigkeitsverteilung.

- Bestimmen Sie den Modus, den Median und das arithmetische Mittel
- Bestimmen Sie die Streuungsparameter Spannweite  $w$ , Varianz  $s^2$ , Standardabweichung  $s$ , Variationskoeffizient  $v$ .

Lösung der Aufgabe 6

Klasse Nr. $i$	Umsatzklasse (Mio €)	Anzahl Unternehmen $h_i$	Anteil $f_i$ (%)	$H_i$	$F_i$ (%)	Klassenmitte $m_i$
1	0 b.u. 1	60	30%	60	30%	0,5
2	1 b.u.2	80	40%	140	70%	1,5
3	2 b.u.5	40	20%	180	90%	3,5
4	5 b.u.10	10	5%	190	95%	7,5
5	10 b.u.20	10	5%	200	100%	15
$\Sigma$		200	100%			

- a)  $\bar{x}_D = 1,5$  Mio € (Mitte der Klasse 2 mit größtem  $h_i$ )

Einfallsklasse  $k = 2$

(s. Bsp. 2 Folie 15 Modul 4)

$$\bar{x}_Z = 1 + (2 - 1) \frac{0,5 - 0,3}{0,4} = 1 + \frac{0,2}{0,4} = 1,5 \text{ (Mio €)}$$

$$\begin{aligned} \bar{x} &= \sum_{i=1}^k m_i * f_i = 0,5 \cdot 0,3 + 1,5 \cdot 0,4 + 3,5 \cdot 0,2 + 7,5 \cdot 0,05 + 15 \cdot 0,05 \\ &= 0,15 + 0,6 + 0,7 + 0,375 + 0,75 + 2,575 \text{ (Mio €)} \end{aligned}$$

- b)  $w = 20 - 0 = 20$

$$\begin{aligned} s^2 &= \sum_{i=1}^k (m_i - \bar{x})^2 * f_i \\ &= (0,5 - 2,575)^2 \cdot 0,3 + (1,5 - 2,575)^2 \cdot 0,4 \\ &\quad + (3,5 - 2,575)^2 \cdot 0,2 + (7,5 - 2,575)^2 \cdot 0,05 \\ &\quad + (15 - 2,575)^2 \cdot 0,05 = 10,86 \end{aligned}$$

$$s = \sqrt{s^2} = \sqrt{10,86} = 3,29$$

$$v = \frac{s}{\bar{x}} = 3,29 / 2,575 = 1,28$$

Aufgabe 7 (Klausuraufgabe WS17/18 mit 18 Punkten):

In der folgenden Tabelle ist die Verteilung der männlichen Teilnehmer bei einer Umfrage auf Altersklassen dargestellt. Dabei wurde zwischen Personen mit und ohne Migrationshintergrund unterschieden.

- Berechnen Sie approximativ die Alters**quartile** ( $Q_1$ ,  $Q_2$ ,  $Q_3$ ) für die beiden Gruppen. Alle Quartile auf ganze Zahlen runden.  
TIPP: vervollständigen Sie zuerst die Häufigkeitstabelle.
- Zeichnen Sie die Boxplots für beide Gruppen in einem Diagramm.  
Bitte denken Sie an die „Lesbarkeit“ des Boxplots (Achsenbeschriftung und Legende nicht vergessen!)
- Wie groß ist der Anteil der Personen im Alter zwischen 15 und 75 Jahre für jede Gruppe? (gemeint ist das Intervall [15;75))

Lösung der Aufgabe 7

Klassen-Nr.	Alter (von – b.u.) in Jahren	mit Migrations- hintergrund $f(x)$ in %	ohne Migrations- hintergrund $f(x)$ in %	mit Migrations- hintergrund kumm. $F(x)$ in %	ohne Migrations- hintergrund kumm. $F(x)$ in %
1	b.u. 15	22	12	22	12
2	15 – 35	31	23	53	35
3	35 – 55	29	33	82	68
4	55 – 75	15	25	97	93
5	75 und älter	3	7	100	100

- a) Median =  $Q_2 \rightarrow$  Median ist auch Quartil (s. Folie 38 Modul 4).

Berechnung von  $Q_1$ ,  $Q_2$ ,  $Q_3$  wie Median für klassierte Daten! (s. Folie 13, 15, 16 Modul 4)

$$\bar{x}_Z := x_{k-1} + (x_k - x_{k-1}) * \frac{0,5 - F_{k-1}}{f_k}$$

mit Migrationshintergrund:

für  $Q_1$  (25%)  $k=2$ ,  $Q1mM=15+(35-15)*(0,25-0,22)/0,31=17$

für  $Q_2$  (50%)  $k=2$ ,  $Q2mM=15+(35-15)*(0,5-0,22)/0,31=33$

für  $Q_3$  (75%)  $k=3$ ,  $Q3mM=35+(55-35)*(0,75-0,53)/0,29=50$

ohne Migrationshintergrund:

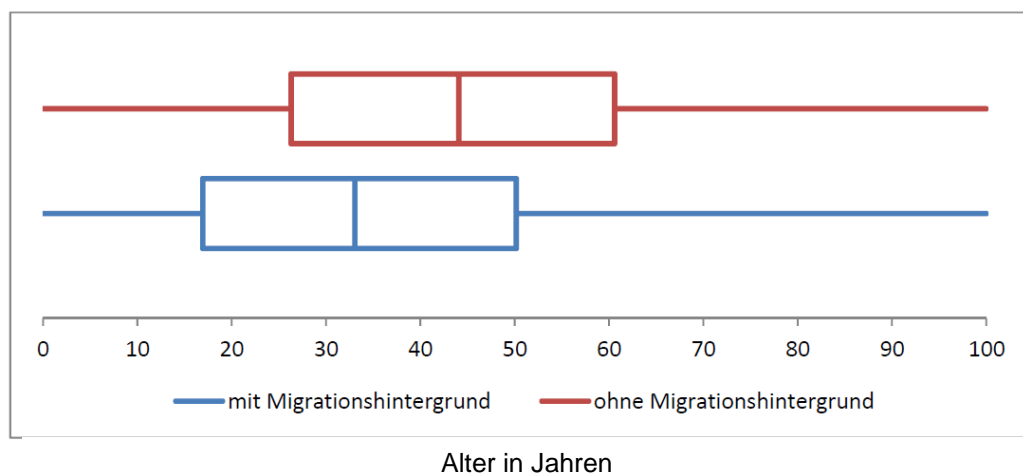
für  $Q_1$  (25%)  $k=2$ ,  $Q1oM=15+(35-15)*(0,25-0,12)/0,23=26$

für  $Q_2$  (50%)  $k=3$ ,  $Q2oM=35+(55-35)*(0,5-0,35)/0,33=44$

für  $Q_3$  (75%)  $k=4$ ,  $Q3oM=55+(75-55)*(0,75-0,68)/0,25=61$

b)

Boxplot:



- c) Mit  $M = 31 + 29 + 15 = 75\%$  (ODER:  $97 - 22 = 75\%$ )

$\rightarrow$ Anteile Personen aus den Klassen 2, 3, 4 addieren ODER subtrahieren Anteil Personen der Klasse 1 von kumuliertem Anteil Personen der Klasse 4.

Ohne  $M = 23 + 33 + 25 = 81\%$  (ODER:  $93 - 12 = 81\%$ )



Aufgabe 8

Die Tabelle zeigt die Zahl der Eheschließungen bzw. die Zahl der Ehescheidungen je 10.000 Ehen in Deutschland auf (Quelle: Statistisches Bundesamt):

Jahr	Eheschließungen	Ehescheidungen je 10.000 Ehen
2001	389.000	198,2
2000	418.550	194
1999	430.674	187,7
1998	417.420	191,4
1997	422.776	181,2
1996	427.297	161
1995	430.534	153,8
1994	440.244	150
1993	442.605	135,6
1992	452.428	104,8
1991	454.291	104

- Bestimme den Modalwert, den Zentralwert und das arithmetische Mittel sowohl von den Eheschließungen als auch von den Ehescheidungen. Was fällt dir an den Ergebnissen auf?
- Bestimme die Varianz und die Standardabweichung sowohl von den Eheschließungen als auch von den Ehescheidungen.
- Welche Veranschaulichungsmöglichkeiten für solch einen tabellarischen Zusammenhang hast du bereits kennen gelernt? Wähle zwei davon aus und realisiere sie! Welche Visualisierungsform ist in diesem Fall besonders geeignet bzw. ungeeignet und warum?

Lösung der Aufgabe 8

- Der Modalwert der Eheschließungen in Deutschland und Ehescheidungen je Ehe existiert nicht, denn jeder Wert kommt nur ein einziges Mal vor. Kein Wert der Datenreihe ist ein Modalwert!

Zentralwert=Median → ist der 6. Element der sortierten Reihe

für Eheschließungen  $\bar{x}_Z = 430.534$

für Ehescheidungen  $\bar{x}_Z = 1.610.000$  oder 161 je 10.000 Ehen

$$\begin{aligned}\bar{x}_{Eheschließungen} &= \frac{389.000 + \dots + 454.291}{11} = \frac{4.725.819}{11} \\ &= 429.619,909 \\ &\approx 429.620\end{aligned}$$

$$\begin{aligned}\bar{x}_{Ehescheidungen} &= \frac{198,2 + \dots + 104}{11} = \frac{1.762}{11} = 160,1545455 \\ &\approx 160 \text{ je } 10.000 \text{ Ehen}\end{aligned}$$

Das arithmetische Mittel entspricht ungefähr dem Zentralwert der Datenmenge sowohl bei den Eheschließungen als auch bei den Ehescheidungen, d.h. dass die Entwicklung der Zahlen gleichmäßig ist und keine Ausreißer aufweist.

b) Berechnung der Varianz über die absolute Häufigkeit:

$$\begin{aligned}s_{Eheschließungen}^2 &= \frac{1}{11} \sum_{i=1}^{11} (x_i - \bar{x})^2 \\ &= \frac{1}{11} ((389.000 - 429.620)^2 + \dots \\ &\quad + (454.291 - 429.620)^2) = 307.808.635\end{aligned}$$

$$s_{Eheschließungen} = \sqrt{s^2} = \sqrt{307.808.635} = 17.545$$

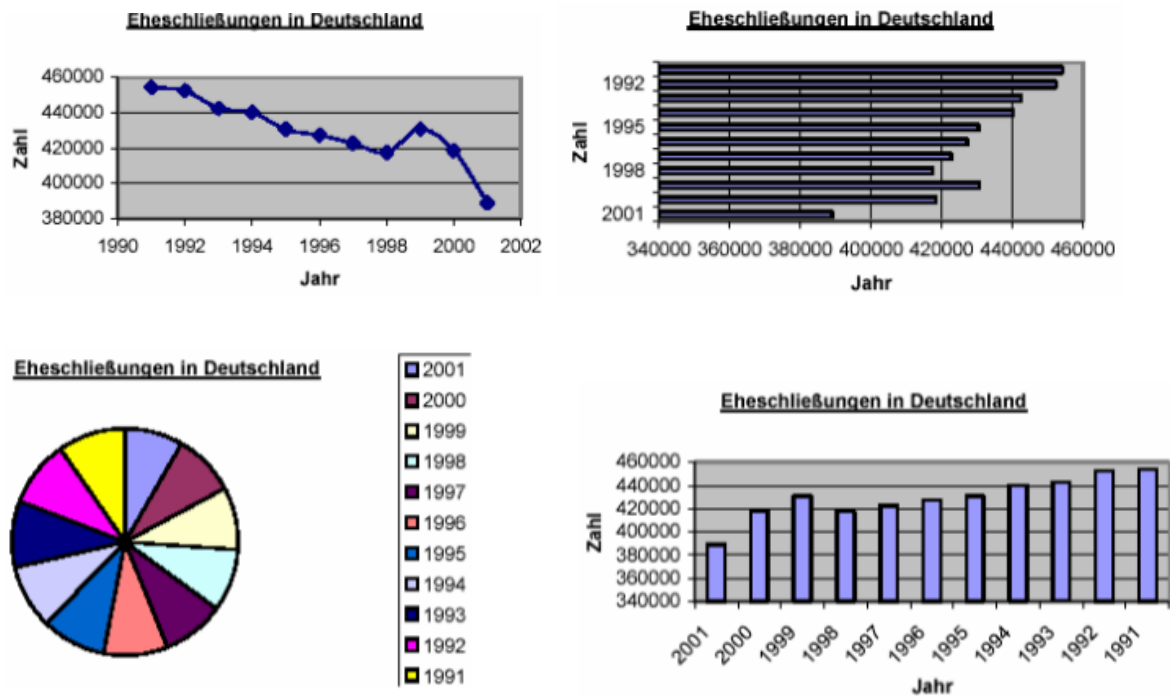
$$\begin{aligned}s_{Ehescheidungen}^2 &= \frac{1}{11} \sum_{i=1}^{11} (x_i - \bar{x})^2 \\ &= \frac{1}{11} ((198,2 - 160)^2 + \dots + (104 - 160)^2) \\ &= 1.066,89 \text{ je } 10.000 \text{ Ehen}\end{aligned}$$

$$s_{Ehescheidungen} = \sqrt{s^2} = \sqrt{1.066,89} = 32,66 \text{ je } 10.000 \text{ Ehen}$$

c) Darstellungsformen: Graph, Säulendiagramm, Stabdiagramm, Balkendiagramm, Kreisdiagramm.

Besonders geeignete Darstellungsformen: Graph, Balkendiagramm, Säulendiagramm und Stabdiagramm; ungeeignete Darstellungsform ist Kreisdiagramm. Begründung: die wichtigsten Informationen sollen der Graphik unmittelbar entnommen werden können. Dies ist im Fall beim Kreisdiagramm nicht möglich.

Beispiele für mögliche Darstellungsformen:



#### Aufgabe 9 (Klausuraufgabe WS18/19 mit 9 Punkten):

Um die Entwicklung der Telefonkosten der letzten 6 Monate des vergangenen Jahres zu analysieren, wird Claudia von ihrem Vater beauftragt, die mittleren Telefonkosten sowie deren Streuung zu berechnen. Die Telefonkosten (in €) sind in der folgenden Tabelle aufgeführt.

*Berechnung mit 2 Nachkommastellen!*

Monat	Juli	August	September	Oktober	November	Dezember
Kosten (€)	31,44	30,18	31,04	33,60	38,16	132,40

#### a) (5 Punkte)

Berechnen Sie das arithmetische Mittel sowie die Streuung der Telefonkosten.

#### b) (4 Punkte)

Claudia, die im Dezember häufig bei teuren Hotlines angerufen hat, ist entsetzt über den hohen Mittelwert und befürchtet Taschengeldentzug durch ihren Vater. Helfen Sie Claudia aus der Patsche, indem Sie ein alternatives Lagemaß, zu Claudias Gunsten, vorschlagen. Begründen Sie Ihren Vorschlag kurz (maximal drei Sätze) und berechnen Sie den Wert Ihres vorgeschlagenen Lagemaßes.

Lösung der Aufgabe 9

a) (5 Punkte)

$$\bar{x} = \frac{1}{6} * (31,44 + 30,18 + 31,04 + 33,6 + 38,16 + 132,4) = \frac{296,82}{6} \\ = 49,47$$

$$s^2 = \frac{1}{6} * (31,44^2 + 30,18^2 + 31,04^2 + 33,6^2 + 38,16^2 + 132,4^2) - 49,47^2 \\ = \frac{22977,69}{6} - 2447,28 = 1382,34 \\ s = \sqrt{s^2} = \sqrt{1382,34} = 37,18$$

b) (4 Punkte)

„Besserer“ Vorschlag: Median – weniger empfindlich gegenüber Ausreißern.  
n gerade, Reihe erst sortieren!

$$\bar{x}_Z = \frac{1}{2} (31,44 + 33,60) = 32,52$$