

Datenanalyse 2017

- Prüfungsdauer ■ 60 Minuten
- Hilfsmittel ■ Taschenrechner erlaubt, CAS-Rechner im Prüfungsmodus!
 ■ Formelsammlung
- Bedingungen ■ Dokumentieren Sie den Lösungsweg sauber.
 ■ Der Lösungsweg muss klar ersichtlich sein.
 ■ Das Resultat ist so weit wie möglich zu vereinfachen.
 ■ **Kontrollieren Sie Ihre Resultate!**
 ■ Falls der freie Platz bei den Aufgaben nicht ausreicht, be-
 nutzen Sie bitte eigene Zusatzblätter.
 Versehen Sie die Aufgabenseite mit einem Hinweis wie
 «Fortsetzung auf Zusatzblatt».

Ich wünsche Ihnen viel Erfolg!

Name und Vorname

Bewertungsübersicht

Aufgabe	1	2	3	4	5	6	7	Gesamtpunkte
Punkte	2	2	1	1	3	2	3	14

Note

Aufgabe 1

2 Punkte

Kreuzen Sie die richtigen Zuordnungen von Merkmalen an.

	nominal	ordinal	stetig	diskret
a. Buchbewertung (*, **, ***, ****)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Hautfarbe	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Verkaufspreise	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
d. Wohnort	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e. Lautstärke	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
f. Anzahl Schulbücher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
g. Vorkommen (oft, selten, nie)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h. Nettoeinkommen	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

a.	0.25
b.	0.25
c.	0.25
d.	0.25
e.	0.25
f.	0.25
g.	0.25
h.	0.25
Total 2	

Aufgabe 2

2 Punkte

Ein Unternehmen besitzt aktuell 40 Dienstwagen. Durch den Kauf von 10 neuen Wagen mit einem durchschnittlichen Benzinverbrauch von 3.5 Liter pro 100 km, soll der Durchschnittsverbrauch aller Dienstwagen um 1.1 Liter pro 100 km gesenkt werden.

Berechnen Sie den Durchschnittsverbrauch, den die bestehenden 40 Wagen haben!

Lösung:

$x = ?$ (Durchschnittsverbrauch in Liter pro 100 km der bestehenden 40 Wagen)

$$\begin{array}{l}
 \text{Verbrauch vor Kauf} + \text{Verbrauch der Neuwagen} = \text{Verbrauch nach Kauf} \\
 40 \cdot x + 10 \cdot 3.5 = 50 \cdot (x - 1.1) \qquad \text{je (0.5)} \\
 40x + 35 = 50x - 55 \qquad \qquad \qquad \text{(0.5)} \\
 x = \frac{90}{10} = 9 \qquad \qquad \qquad \text{(0.5)}
 \end{array}$$

Der Durchschnittsverbrauch der bestehenden 40 Wagen beträgt 9 Liter pro 100 km.

0.5	
0.5	
0.5	
0.5	
Total 2	

Aufgabe 3

1 Punkt

Die Daten in einer Stichprobe wurden in der untenstehenden Tabelle neu organisiert. Berechnen Sie die korrekten Einträge A, B, C und D in den grau schraffierten Feldern.

Klasse	absolute Häufigkeit	Klassenmitte	relative Häufigkeit
[15;20[9	17.5	0.045
A	57	22.5	
[25;30[B	27.5	0.335
[30;35[C		
[35;40[22	37.5	D

A	0.25
B	0.25
C	0.25
D	0.25
Total 1	

Lösung:

A: [20;25[ohne Berechnung

$$B: \frac{9}{0.045} = \frac{B}{0.335} \rightarrow B = \frac{9 \cdot 0.335}{0.045} = \underline{\underline{67}}$$

C: Stichprobenumfang $n = \frac{9 \cdot 1}{0.045} = \underline{\underline{200}}$
 somit: $C = 200 - 9 - 57 - 67 - 22 = \underline{\underline{45}}$

oder (falls D zuerst berechnet wurde)

relative Häufigkeit bei A: $h_A = \frac{0.045 \cdot 57}{9} = \underline{\underline{0.285}}$
 relative Häufigkeit bei C: $1 - 0.045 - 0.285 - 0.335 - 0.110 = \underline{\underline{0.225}}$
 somit: $C = \frac{9 \cdot 0.225}{0.045} = \underline{\underline{45}}$

D: $D = \frac{0.045 \cdot 22}{9} = \underline{\underline{0.110}}$

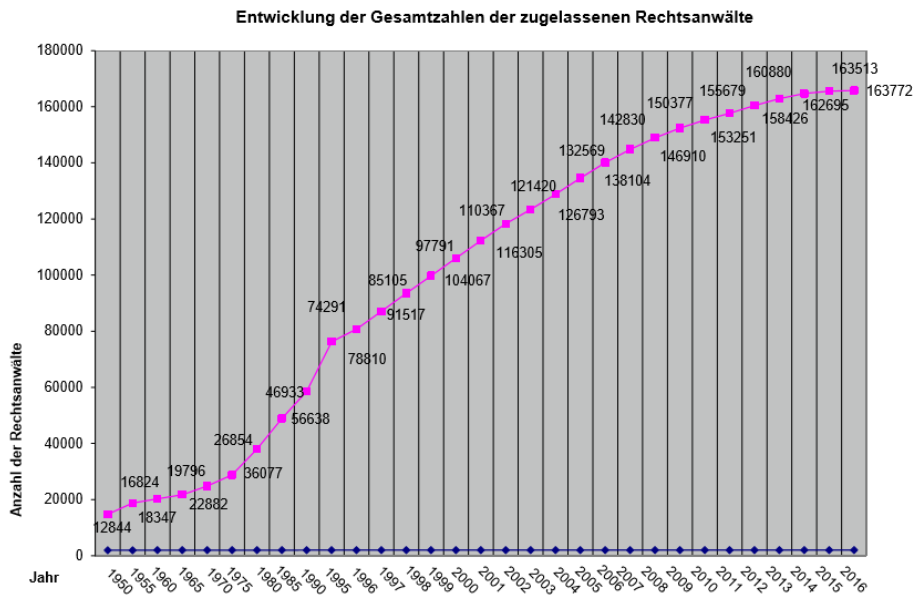
Aufgabe 4

1 Punkt

In diversen Medien finden sich immer wieder Diagramme und Graphiken, die fehlerhaft sind oder manipulieren sollen.

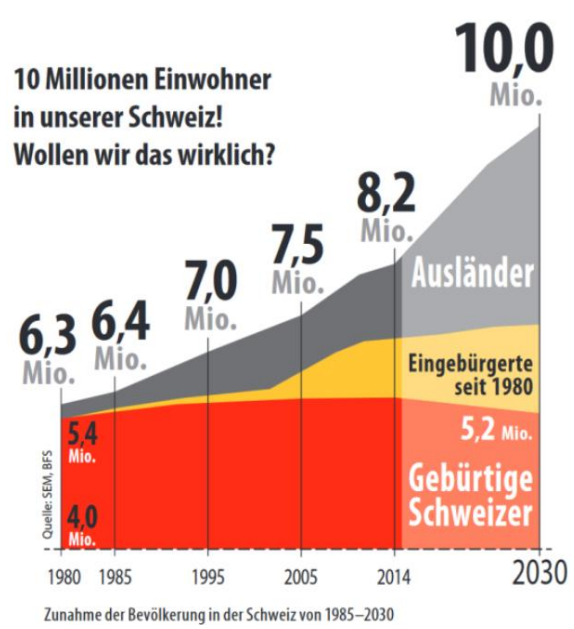
Geben Sie je ein Argument an, weshalb folgende Graphiken zu kritisieren sind:

a.	0,5
b.	0,5
Total 1	



Quelle: Bundesrechtsanwaltskammer Deutschland

a. Die Einteilung der x-Achse (Massstab) ist nicht konstant (5er- und 1er- Intervalle)!



Quelle: SVP Flyer

b. Die y-Achse beginnt erst bei 4 Millionen. Unterer Teil der Grafik wurde abgeschnitten.

Aufgabe 5

3 Punkte

Der Haemocult-Test spürt Blut im Stuhl auf, ein Hinweis auf Darmkrebs. Der Test spricht in etwa 50 Prozent der Fälle an (Sensitivität) und liefert bei 3 Prozent falsch positive Ergebnisse (Spezifität). 0.3 Prozent der Bevölkerung erkranken überhaupt an Darmkrebs. Insgesamt werden 100'000 Personen untersucht. Lösen Sie die Aufgabe mit Hilfe einer Vierfeldertafel.

- a. Wie viel Prozent der Personen, von denen man weiss, dass sie ein positives Testergebnis haben sind nicht krank?
- b. Wie viel Prozent der Personen, von denen man weiss, dass sie nicht krank sind, hatten ein negatives Testergebnis?

Lösung:

		Erkrankung		Summe
		Ja	Nein	
Test	positiv	150 (0.25)	2'991 (0.25)	3'141 (0.25)
	negativ	150 (0.25)	96'709 (0.25)	96'859 (0.25)
Summe		300 (0.25)	99'700 (0.25)	100'000

0.25
0.25
0.25
0.25
0.25
0.25
0.25
0.25
a. 0.5
b. 0.5
Total 3

a. $P(\text{nicht krank} | \text{positiver Test}) = \frac{2'991}{3'141} = 0.9522 \approx \underline{\underline{95.2\%}}$ (0.5)

b. $P(\text{negativer Test} | \text{nicht krank}) = \frac{96'709}{99'700} = 0.97 = \underline{\underline{97\%}}$ (0.5)
oder 100% - 3%

Aufgabe 6

2 Punkte

Die Befragung «Wie viele Geschwister haben Sie?» hat folgende Häufigkeitstabelle ergeben:

Merkmal: Anzahl Geschwister

i	x_i	n_i	h_i	F_i
1	0	3	0.1	0.1
2	1	12	0.4	0.5
3	2	7	0.2333...	0.7333...
4	3	5	0.1666...	0.9
5	4	2	0.0666...	0.9666...
6	5	1	0.0333...	1

a.	0.25
b.	0.25
c.	0.25
d.	0.25
e.	0.25
f.	0.25
g.	0.25
h.	0.25
Total 2	

Beantworten Sie aufgrund obiger Häufigkeitstabelle die folgenden Fragen:

- Wie viele Befragte haben mindestens 3 Geschwister?
- Wie viele Prozent der Befragten haben höchstens 2 Geschwister?
- Wie viele Befragte haben genau 3 Geschwister?
- Wie viele Befragte haben höchstens 2 Geschwister?
- Wie viele Befragte sind Einzelkinder?
- Wie viele verschiedene Anzahlen wurden erhoben?
- Wie viele Prozent der Befragten sind keine Einzelkinder?
- Wie viele Prozent der Befragten haben mindestens 2 Geschwister?

a.	8	$n_4 + n_5 + n_6$
b.	73.3 %	F_3
c.	5	n_4
d.	22	$n_1 + n_2 + n_3$
e.	3	n_1
f.	6	i_6
g.	$100 \% - 10 \% = 90 \%$	$F_6 - F_1$
h.	$100 \% - 50 \% = 50 \%$	$F_6 - F_2$

Aufgabe 7

3 Punkte

Ein Unternehmen möchte den Einfluss seiner Werbeausgaben auf den erzielten Umsatz quantifizieren. Hierfür werden die jährlichen Ausgaben für Werbung y (in 1'000 CHF) und die jährlich erzielten Umsätze x (in 100'000 CHF) über einen Zeitraum von acht Jahren erfasst. Die folgenden Werte wurden bereits berechnet:

$$\sum y = 110, \quad \sum y^2 = 1'700, \quad \sum x = 88, \quad \sum x^2 = 1'020, \quad \sum xy = 1'288$$

- a. Angenommen das Unternehmen plant jährlich 14'000 CHF für die Werbung auszugeben. Mit welchem Umsatz kann es rechnen?
- b. Berechnen und interpretieren Sie eine geeignete statistische Kennzahl zur Messung des linearen Zusammenhangs zwischen den jährlichen Werbeausgaben und dem Umsatz.

Lösung:

$$a. \quad m = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum x_i^2 - n \cdot \bar{x}^2} = \frac{\overbrace{1'288 - 8 \cdot \frac{88}{8} \cdot \frac{110}{8}}^{(0.25)}}{\underbrace{1'020 - 8 \cdot \left(\frac{88}{8}\right)^2}_{(0.25)}} = \frac{3}{2} \quad (0.25)$$

$$b = \bar{y} - m\bar{x} = \frac{110}{8} - \frac{3}{2} \cdot \frac{88}{8} = -\frac{11}{4} \quad (0.25)$$

$$\text{somit: } y = f(x) = \frac{3}{2} \cdot x - \frac{11}{4}$$

$$y = f(x) = 14 \rightarrow \frac{3}{2} \cdot x - \frac{11}{4} = 14 \rightarrow x = \frac{14 + \frac{11}{4}}{\frac{3}{2}} = \underline{\underline{11.1\bar{6}}} \quad (\text{in 100'000 CHF}) \quad (0.25)$$

Das Unternehmen kann mit einem Umsatz von 1'116'667 CHF rechnen! (0.5)

$$b. \quad r = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sqrt{(\sum x_i^2 - n \cdot \bar{x}^2) \cdot (\sum y_i^2 - n \cdot \bar{y}^2)}} = \frac{\overbrace{1'288 - 8 \cdot \frac{88}{8} \cdot \frac{110}{8}}^{(0.25)}}{\sqrt{\underbrace{\left[1'020 - 8 \cdot \left(\frac{88}{8}\right)^2\right]}_{(0.25)} \cdot \underbrace{\left[1'700 - 8 \cdot \left(\frac{110}{8}\right)^2\right]}_{(0.25)}}} = \underline{\underline{0.7899}} \quad (0.25)$$

Es liegt eine positive mittelstarke Korrelation vor. (0.25)

a.	0.25
	0.25
	0.25
	0.25
	0.25
	0.5
b.	0.25
	0.25
	0.25
	0.25
	0.25
Total 3	

