

### Prüfungsklausur - Explorative Datenanalyse

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\Sigma$

Name: \_\_\_\_\_, Vorname: \_\_\_\_\_

Matr.-Nr. \_\_\_\_\_ Studiengang: \_\_\_\_\_

**Hinweise:** (Bitte lesen Sie diese Hinweise genau durch.)

- Bitte tragen Sie als erstes Ihre persönliche Daten auf diesem Blatt ein.
- Zugelassene Hilfsmittel: Ein beidseitig mit der Hand beschriebenes DIN-A4-Blatt (Markierungen erlaubt), ein (vom WiWi-Prüfungsamt erlaubter) Taschenrechner und ein Geodreieck (oder Lineal und Winkelmesser).
- Die Klausur besteht aus fünfzehn (15) Aufgaben.
- Bei den Multiple-Choice-Aufgaben ist jeweils **genau eine** Antwort richtig.
- Für eine korrekte Antwort erhalten Sie einen Punkt. Für eine falsche Antwort oder eine nicht beantwortete Frage erhalten Sie weder einen Punkt noch wird Ihnen etwas abgezogen.
- Bitte markieren Sie bei jeder Frage die von Ihnen ausgewählte Antwort klar erkennbar. Mehrfache Antworten bei einer Frage werden als falsch beantwortet bewertet. Kennzeichnen Sie daher Korrekturen deutlich.
- **Beschriften Sie die Grafiken sinnvoll.**
- Die Klausur ist bestanden, wenn mindestens **acht** (8) Punkte erzielt werden.

## Aufgabe 1.

Um welchen Typ von Daten handelt es sich bei den folgenden Merkmalen:

- Kaffeepreis in €/kg
- Berufsgruppenzugehörigkeit
- Anzahl der Kinder in einer Familie
- Kundenzufriedenheit („zufrieden“, „teils/teils“, „unzufrieden“)

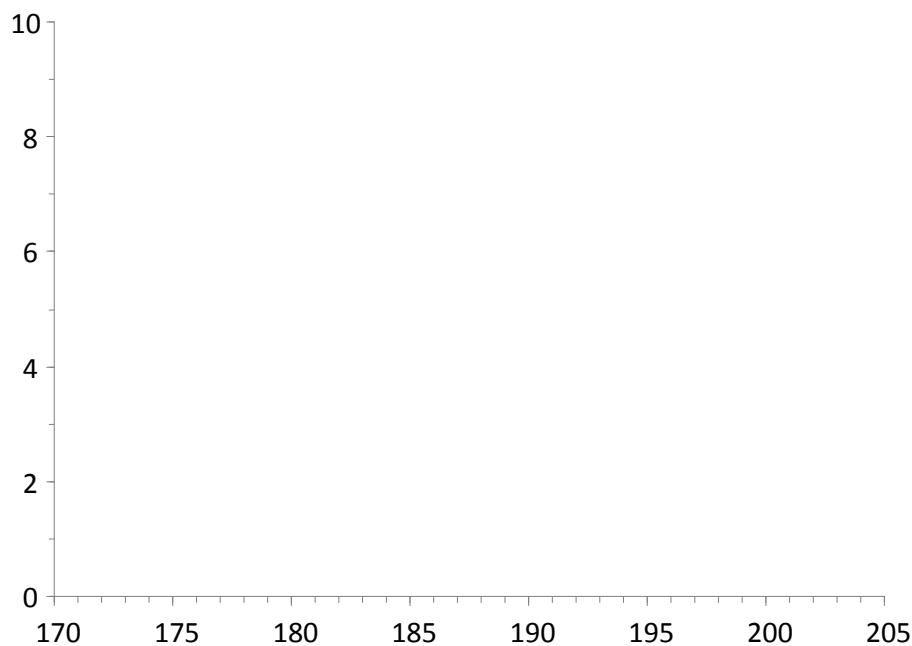
- A  Anzahl der Kinder ist quantitativ und Kaffeepreis ist ordinal  
B  Anzahl der Kinder und Kundenzufriedenheit sind metrisch diskret  
C  Kaffeepreis und Berufsgruppenzugehörigkeit sind qualitativ  
D  Berufsgruppenzugehörigkeit und Anzahl der Kinder sind intervallskaliert  
E  Kundenzufriedenheit ist ordinal und Berufsgruppenzugehörigkeit ist nominal
- 

## Aufgabe 2.

In der folgenden Tabelle sind die Weizenpreise in €/Tonne für die Werkstage im November 2011 angegeben.

177,50	177,50	179,25	180,00	180,00
180,25	180,75	183,00	183,25	183,75
184,50	185,00	185,50	186,50	187,50
188,00	190,00	193,00	193,25	195,25

Zeichnen Sie das zugehörige Histogramm mit Klassenbreiten 5 €/Tonne, wobei die unterste Klasse bei 175 €/Tonne beginnt.

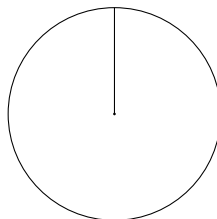


### Aufgabe 3.

Bei einem Marketing-Experiment wurden Verkaufsmengen von Produkten A, B, C und D während eines festgesetzten Zeitraums beobachtet. Folgende Tabelle enthält die Verkaufsmengen der einzelnen Produkttypen.

Produkttyp	Verkaufsmenge
A	20
B	10
C	60
D	30

Vervollständigen Sie das nachfolgende Tortendiagramm unter Berücksichtigung des Pareto-Prinzips.



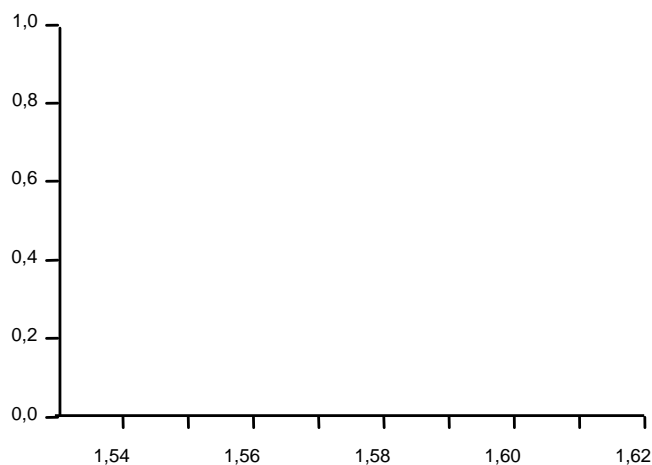
---

### Aufgabe 4.

In der folgenden Tabelle sind Kraftstoffpreise in € pro Liter für den Kraftstoff Super E10 an Tankstellen in Magdeburg im Zeitraum vom 26.01. bis 30.01.2012 aufgeführt.

Kraftstoffpreis	absolute Häufigkeit
1,539	4
1,559	5
1,569	6
1,579	3
1,609	2

Zeichnen Sie die zugehörige empirische Verteilungsfunktion.



### Aufgabe 5.

An aufeinander folgenden Werktagen wurden folgende Kaffeepreise in US-\$ pro Pfund (lb.) notiert:

3. Jan.	4. Jan.	5. Jan.	6. Jan.	9. Jan.	10. Jan.	11. Jan.	12. Jan.
2,27	2,27	2,20	2,22	2,22	2,26	2,35	2,34

Der mittlere Preis für diese Periode beträgt 2,27 US-\$ pro Pfund. Berechnen Sie den mittleren Preis in Euro pro kg Kaffee. Verwenden Sie den Umrechnungskurs 0,759 € für 1 US-\$ und 1 Pfund (lb.) entspricht 0,453 kg.

- A  0,78
  - B  1,35
  - C  1,72
  - D  3,80
  - E  6,60
- 

### Aufgabe 6.

Folgende Tabelle enthält die Preise von Daimler-Aktien an zufällig ausgewählten Tagen des Zeitraums vom 1.7. bis 31.12.2011.

Tag	1	2	3	4	5	6
Aktienpreis	32,20	34,80	36,58	33,25	47,10	50,77

Das arithmetische Mittel beträgt:

- A  30,66
  - B  234,70
  - C  33,32
  - D  39,12
  - E  46,94
- 

### Aufgabe 7.

Bestimmen Sie die Stichprobenstandardabweichung zu den Daten aus Aufgabe 6.

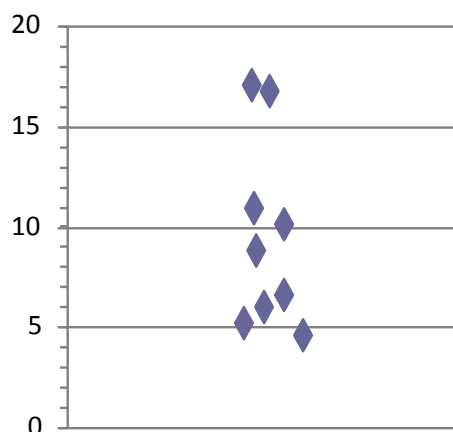
- A  61,37
  - B  7,83
  - C  7,15
  - D  125,28
  - E  51,14
-

### Aufgabe 8.

In Magdeburg wurden im Januar 2012 folgende Windgeschwindigkeiten in m/sec gemessen:

Tag	1	2	3	4	5	6	7	8	9
Maximale Windgeschwindigkeit	6,6	5,2	10,2	8,8	4,6	6,0	11,0	16,8	17,1

Zeichnen Sie einen Boxplot für diese Daten.



---

### Aufgabe 9.

Für Lage- und Steuungsmaßzahlen gilt:

- A  Das arithmetische Mittel  $\bar{x}$  und der Median  $x_{\text{med}}$  sind immer gleich.
- B  Die Stichprobenstandardabweichung  $s_x$  ist immer kleiner als der Interquartilsabstand.
- C  Median und Interquartilsabstand sind robust gegen Ausreißer.
- D  Das geometrische Mittel kann auch negativ sein.
- E  Das arithmetische Mittel kann aus der empirischen Verteilungsfunktion abgelesen werden.

---

### Aufgabe 10.

Für die Maßzahl  $\chi^2$  und den Kontingenzkoeffizienten  $K$  gilt:

- A   $K > 0$  bedeutet, dass ein positiver linearer Zusammenhang besteht.
- B   $K < 0$  bedeutet, dass ein negativer monotoner Zusammenhang besteht.
- C   $K$  liegt immer zwischen Null und Eins ( $0 \leq K \leq 1$ ).
- D   $K$  ist immer größer als  $\chi^2$ .
- E   $\chi^2$  kann auch negative Werte annehmen.

---

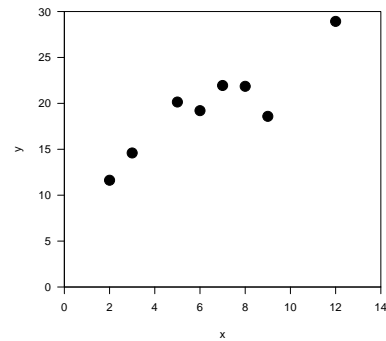
### Aufgabe 11.

Gegeben sind folgende Daten:

$i$	$x_i$	$y_i$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	3	14,60	12,25	25,15	17,55
2	6	19,21	0,25	0,16	0,20
3	2	11,63	20,25	63,76	35,93
4	8	21,86	2,25	5,04	3,37
5	12	28,93	30,25	86,77	51,23
6	5	20,15	2,25	0,29	-0,80
7	7	21,95	0,25	5,45	1,17
8	9	18,59	6,25	1,05	-2,56
Summe	52	156,92	74,00	187,67	106,09

Der Pearsonsche Korrelationskoeffizient ist:

- A  0,90
- B  1,03
- C  0,01
- D  -0,81
- E  0,13



---

### Aufgabe 12.

Bestimmen Sie den Achsenabschnitt  $a$  und die Steigung  $b$  der Regressionsgeraden zu den Daten aus Aufgabe 11.

- A   $a = 0,00, \quad b = 1,37$
  - B   $a = 28,92, \quad b = -1,43$
  - C   $a = 2,33, \quad b = 0,59$
  - D   $a = -60,53, \quad b = 12,33$
  - E   $a = 10,30, \quad b = 1,43$
-

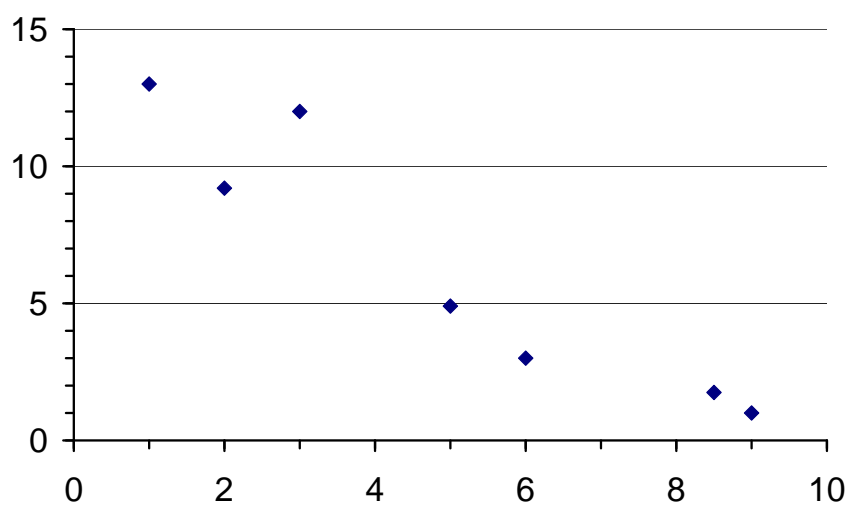
### Aufgabe 13.

Für die Regressionsgerade  $y = a + bx$  gilt:

- A  Die Achsenabschnitt  $a$  und der Pearsonsche Korrelationskoeffizient  $r_{xy}$  haben immer dasselbe Vorzeichen.
  - B  Die Steigung  $b$  ist immer gleich dem Pearsonschen Korrelationskoeffizienten  $r_{xy}$ .
  - C  Aus  $b < 0$  folgt  $r_{xy} > 0$ .
  - D  Der Schwerpunkt  $(\bar{x}, \bar{y})$  liegt immer auf der Regressionsgerade.
  - E  Ein positiver Achsenabschnitt  $a$  bedeutet, dass mit größeren  $x$ -Werten auch größere  $y$ -Werte auftreten.
- 

### Aufgabe 14.

Tragen Sie in das nachfolgende Streudiagramm die entsprechende Regressionsgerade ein. Die Regressionsparameter  $a$  und  $b$  betragen hier 13,8 und  $-1,5$ .



### Aufgabe 15.

Für den Spearmanschen Rangkorrelationskoeffizienten  $r_{S,xy}$  gilt:

- A   $r_{S,xy} = -1$  bedeutet, dass alle Daten auf einer Gerade liegen.
  - B  Positive Rangkorrelation ( $r_{S,xy} > 0$ ) bedeutet, dass ein positiver linearer Zusammenhang vorhanden ist.
  - C   $r_{S,xy} < 0$  bedeutet, dass ein negativer monotoner Zusammenhang vorhanden ist.
  - D  Wenn ein Zusammenhang zwischen den zwei Merkmalen besteht, dann ist die Rangkorrelation von Null verschieden ( $r_{S,xy} \neq 0$ ).
  - E   $r_{S,xy}$  ist immer größer als der Pearsonsche Korrelationskoeffizient  $r_{xy}$ .
-