Prof. Dr. Rainer Schwabe
Dipl.-Math. Marius Schmidt
Otto-von-Guericke-Universität Magdeburg
Fakultät für Mathematik

5. Februar 2014

Prüfungsklausur - Explorative Datenanalyse

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Σ

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:

- Nationalität
- Prozentualer Gewinn eines Unternehmens
- Qualität eines Produkts (-1 =„gering“, 0 =„mittel“, 1 =„hoch“)
- Semesterzahl

A ☐ Qualität eines Produkts ist nominal und prozentualer Gewinn ist verhältnisskaliert.

B ☐ Prozentualer Gewinn ist metrisch und Nationalität ist ordinal.

C ☐ Nationalität und Semesterzahl sind quantitativ.

D ☐ Nationalität ist nominal und Qualität eines Produkts ist ordinal.

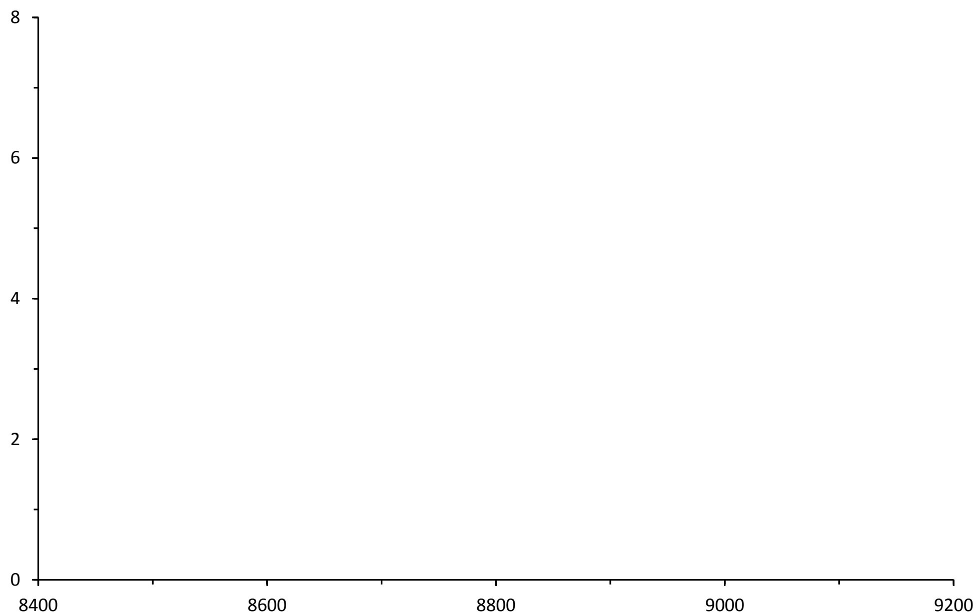
E ☐ Qualität eines Produkts und Semesterzahl sind metrisch diskret.

Aufgabe 2.

In der folgenden Tabelle ist der DAX-Kurs zu Handelsschluss für die Werktage vom 04.10.2013 bis 31.10.2013 angegeben.

8 623,0	8 591,6	8 555,9	8 516,7	8 685,8
8 724,8	8 723,8	8 804,4	8 846,0	8 812,0
8 865,1	8 867,2	8 947,5	8 919,9	8 980,6
8 985,7	8 978,7	9 022,0	9 010,3	9 033,9

Zeichnen Sie das zugehörige Histogramm mit Klassenbreite 100 Punkte, wobei die unterste Klasse bei 8500 Punkten beginnt.

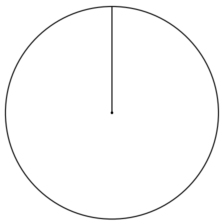


Aufgabe 3.

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	10
B	30
C	15
D	5

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

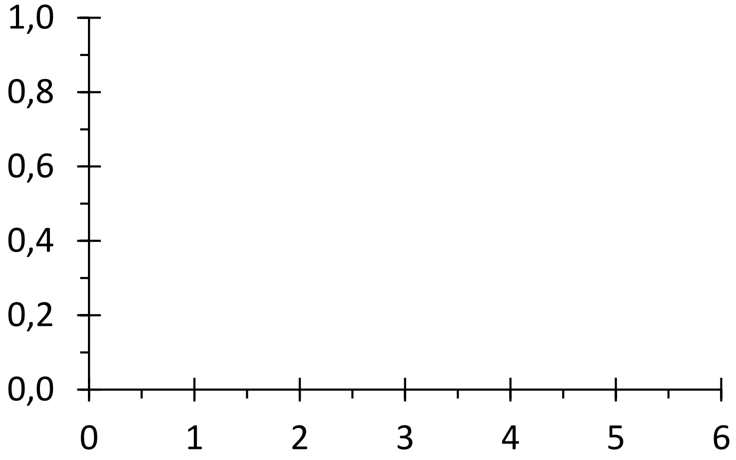


Aufgabe 4.

In der folgenden Tabelle sind die Anzahlen von Autos pro Haushalt und deren absolute Häufigkeiten angegeben.

Anzahl der Autos	absolute Häufigkeit
0	2
1	8
2	6
3	3
5	1

Zeichnen Sie die zugehörige empirische Verteilungsfunktion.



Aufgabe 5.

An aufeinander folgenden Werktagen wurden folgende Silberpreise in US-$ pro Feinunze notiert:

6. Jan.	7. Jan.	8. Jan.	9. Jan.	10. Jan.	13. Jan.	14. Jan.	15. Jan.
20,25	19,63	19,64	19,58	20,15	20,40	20,25	20,18

Der mittlere Preis für diese Periode beträgt 20,01 US-$ pro Feinunze. Berechnen Sie den mittleren Preis in Euro pro g Silber. Verwenden Sie den Umrechnungskurs 1,362 US-$ ist 1 Euro und 1 Feinunze entspricht 31,103 g.

A ☐ 0,47
B ☐ 0,88
C ☐ 14,69
D ☐ 456,95
E ☐ 847,67

Aufgabe 6.

Folgende Tabelle enthält die Preise in Euro von BMW-Aktien an zufällig ausgewählten Tagen des Zeitraums vom 1.12. bis 31.12.2013.

Tag	1	2	3	4	5	6
Aktienpreis	82,00	81,68	80,96	79,89	84,09	85,06

Das arithmetische Mittel beträgt:

A ☐ 81,84
B ☐ 89,74
C ☐ 82,28
D ☐ 493,68
E ☐ 80,96

Aufgabe 7.

Bestimmen Sie die Stichprobenstandardabweichung zu den Daten aus Aufgabe 6.

A ☐ 4,35
B ☐ 201,54
C ☐ 3,78
D ☐ 1,77
E ☐ 1,94

Aufgabe 8.

In Magdeburg wurden für elf aufeinanderfolgende Tage im Januar 2014 folgende Tageshöchsttemperaturen bestimmt:

Tag	1	2	3	4	5	6	7	8	9	10	11
Tageshöchsttemperatur	7,9	11,6	11,0	12,7	7,6	8,0	5,7	5,3	4,7	4,2	5,8

Zeichnen Sie für diese Daten einen Boxplot in folgendes Diagramm:

Aufgabe 9.

Für Lage- und Steuungsmaßzahlen sowie Quantile gilt:

A ☐ Für positive Daten gilt immer $x_{\text{geom}} \leq x_{\text{harm}} \leq \bar{x}$.

B ☐ Das arithmetische Mittel $\bar{x}$ und der Median x_{med} sind immer gleich.

C ☐ Für Stichprobenvarianz s_x^2 und Populationsvarianz d_x^2 gilt immer $s_x^2 \geq d_x^2$.

D ☐ Quantile sind immer positiv.

E ☐ Für den Interquartilsabstand gilt immer $IQR = 2(x_{\text{med}} - x_{0,25})$.

Aufgabe 10.

Für die Maßzahl χ^2 und den Kontingenzkoeffizienten K gilt:

A ☐ $\chi^2 = 1$ bedeutet, dass es einen starken Zusammenhang zwischen zwei Merkmalen gibt.

B ☐ χ^2 kann bei qualitativen Merkmalen angewendet werden.

C ☐ K ist immer größer als χ^2.

D ☐ Bei $K = 0$ ist ein quadratischer Zusammenhang zwischen zwei Merkmalen noch möglich.

E ☐ K kann im Gegensatz zu χ^2 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	6	21,72	0,25	72,44	-4,26
2	8	6,04	2,25	51,39	-10,75
3	3	19,43	12,25	38,70	-21,77
4	13	2,86	42,25	107,10	-67,27
5	10	9,93	12,25	10,75	-11,48
6	1	18,15	30,25	24,42	-27,18
7	7	10,95	0,25	5,10	-1,13
8	4	16,59	6,25	11,43	-8,45
Summe	52	105,67	106,00	321,33	-152,29

Der Pearsonsche Korrelationskoeffizient ist:

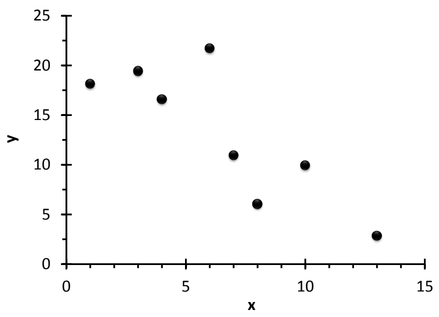
A ☐ 0,83

B ☐ -0,65

C ☐ -0,83

D ☐ 0,05

E ☐ 1,08



Aufgabe 12.

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

A ☐ $a = 16{,}29, \quad b = -0{,}47$

B ☐ $a = 22{,}55, \quad b = -1{,}44$

C ☐ $a = 10{,}13, \quad b = 0{,}47$

D ☐ $a = 18{,}57, \quad b = -0{,}83$

E ☐ $a = 17{,}73, \quad b = -0{,}70$

Aufgabe 13.

Für die Regressionsgerade $y = a + bx$ und den Pearsonschen Korrelationskoeffizienten r_{xy} gilt:

A ☐ Ist r_{xy} klein, so ist auch b klein.

B ☐ Der Pearsonsche Korrelationskoeffizient r_{xy} lässt sich mit Hilfe von a besonders leicht berechnen.

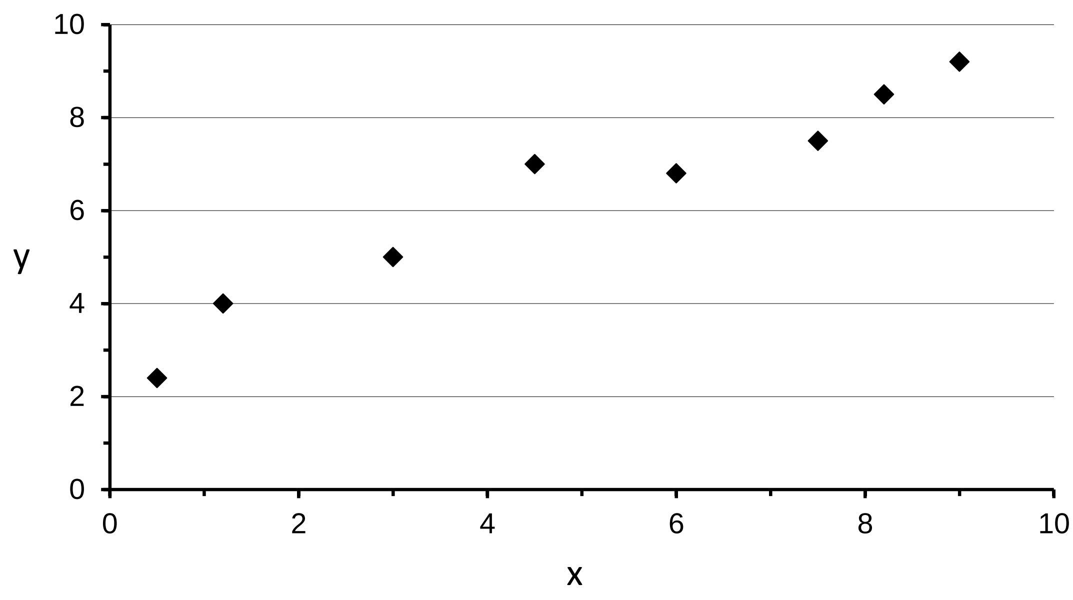
C ☐ Wenn alle Punkte (x_i, y_i) auf der Regressionsgeraden liegen, dann ist $r_{xy} = -1$.

D ☐ Wenn $|r_{xy}| \neq 1$ ist, dann ist es möglich, dass kein Punkt (x_i, y_i) auf der Regressionsgeraden liegt.

E ☐ Falls $r_{xy} \neq 0$ ist, liegt mindestens ein Punkt (x_i, y_i) auf der Regressionsgeraden.

Aufgabe 14.

Tragen Sie in das nachfolgende Streudiagramm die entsprechende Regressionsgerade $y(x) = a + bx$ ein. Die Regressionsparameter a und b betragen hier 2,81 und 0,70.



Aufgabe 15.

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

A ☐ Der Rangkorrelationskoeffizient $r_{S,xy}$ ist immer größer als der Pearsonsche Korrelationskoeffizient r_{xy}.

B ☐ Ist $r_{S,xy} = 1$, dann ist es möglich, dass alle Datenpunkte auf einer streng monoton wachsenden Geraden liegen.

C ☐ r_{xy} und $r_{S,xy}$ haben immer das gleiche Vorzeichen.

D ☐ Es gilt immer $0 \leq r_{S,xy} \leq 1$.

E ☐ Wenn ein Zusammenhang zwischen den zwei Merkmalen besteht, dann ist die Rangkorrelation von Null verschieden ($r_{S,xy} \neq 0$).
