



i Hinweis: KI-generierte Musterlösung - bei Unklarheiten gerne nochmal im Buch nachschlagen.

AUFGABE 1a (8 Punkte)

Erwartung: (1) Fachsemesteranzahl: ratio-/verhältnisskaliert (natürlicher Nullpunkt, Verhältnisse sinnvoll interpretierbar). (2) Bevorzugte Sitzplatzreihe: ordinalskaliert, da eine natürliche Rangfolge vorne-Mitte-hinten besteht, die Abstände zwischen den Kategorien aber nicht quantifizierbar/nicht gleich groß sind. (3) Matrikelnummer: nominalskaliert, da die Zahl nur der Identifikation/Unterscheidung dient und weder Rangfolge noch Abstand oder Verhältnis eine inhaltliche Bedeutung haben. (4) Körpergröße: ratio-/verhältnisskaliert (natürlicher Nullpunkt bei 0 cm, Verhältnisse wie „doppelt so groß“ sinnvoll). (5) Note im Statistik-Vortest: ordinalskaliert, da Schulnoten eine Rangfolge der Leistung abbilden, die Abstände zwischen den Notenstufen aber nicht als inhaltlich gleich groß angenommen werden können (der Leistungsunterschied zwischen Note 1 und 2 muss nicht dem zwischen 4 und 5 entsprechen). (6) Religionszugehörigkeit: nominalskaliert (Kategorien ohne natürliche Rangfolge). (7) Reaktionszeit in Millisekunden: ratio-/verhältnisskaliert (natürlicher Nullpunkt bei 0 ms, Verhältnisse sinnvoll). (8) Raumtemperatur in Grad Celsius: intervallskaliert, da gleiche Abstände zwischen den Werten inhaltlich interpretierbar sind, der Nullpunkt (0 Grad Celsius) aber willkürlich gewählt ist und keine absolute Abwesenheit von Temperatur bedeutet, weshalb Verhältnisaussagen („doppelt so warm“) nicht sinnvoll sind.

Bewertung: 0,5 P. je korrektes Skalenniveau (8 x 0,5 P. = 4 P.); je 1 P. für die vier geforderten Begründungen bei (2), (3), (5) und (8) (4 x 1 P. = 4 P.)

AUFGABE 1b (4 Punkte)

Erwartung: Sowohl Intervall- als auch Verhältnisskala besitzen gleich große, interpretierbare Abstände zwischen den Werten. Der Unterschied liegt im Nullpunkt: Bei der Verhältnisskala ist der Nullpunkt ein absoluter, natürlicher Nullpunkt, der die vollständige Abwesenheit der Eigenschaft bedeutet (z. B. 0 cm = keine Länge). Dadurch sind Verhältnisaussagen („doppelt so viel“) zulässig. Bei der Intervallskala ist der Nullpunkt dagegen willkürlich festgelegt und bedeutet nicht die Abwesenheit der Eigenschaft. Die Raumtemperatur in Grad Celsius ist deshalb nicht verhältnisskaliert, weil 0 Grad Celsius lediglich der Gefrierpunkt von Wasser ist und keine „Abwesenheit von Temperatur“ darstellt (anders als 0 Kelvin); Aussagen wie „20 Grad Celsius ist doppelt so warm wie 10 Grad Celsius“ sind daher nicht sinnvoll interpretierbar.

Bewertung: 1 P. gleich große Abstände bei beiden Skalentypen; 1 P. Verhältnisskala besitzt absoluten/natürlichen Nullpunkt; 1 P. Intervallskala besitzt willkürlichen Nullpunkt, dadurch keine Verhältnisaussagen zulässig; 1 P. korrekte Erklärung am Beispiel Grad Celsius





AUFGABE 1c (3 Punkte)

Erwartung: Aussage 1 („Der Abstand zwischen Rangplatz 1 und 2 ist genauso groß wie zwischen Rangplatz 2 und 3.“): Falsch - bei ordinalskalierten Daten ist nur die Rangfolge, nicht aber die Gleichheit der Abstände zwischen den Rängen bekannt oder garantiert. Aussage 2 („Person A sitzt lieber weiter vorne als Person B.“): Richtig - dies ist lediglich eine Rangordnungsaussage (Vergleich „mehr/weniger“), die bei ordinalskalierten Daten zulässig und sinnvoll interpretierbar ist.

Bewertung: 1 P. richtiges Kreuz bei Aussage 1 (Falsch) mit Begründung; 1 P. richtiges Kreuz bei Aussage 2 (Richtig) mit Begründung; 1 P. für die inhaltlich korrekte Unterscheidung (Rangordnung erlaubt, Abstandsaussage nicht erlaubt)

AUFGABE 1d (3 Punkte)

Erwartung: Die Berechnung eines arithmetischen Mittelwerts setzt voraus, dass die Abstände zwischen den Werten gleich groß und inhaltlich interpretierbar sind (mindestens Intervallskalenniveau). Schulnoten sind jedoch nur ordinalskaliert: Es ist nicht gesichert, dass der Leistungsunterschied zwischen Note 1 und Note 2 genauso groß ist wie zwischen Note 4 und Note 5. Die Berechnung eines Mittelwerts unterstellt daher eine Gleichabständigkeit, die bei Ordinaldaten streng genommen nicht gegeben ist, weshalb das Verfahren methodisch umstritten ist, auch wenn es in der Praxis aus pragmatischen Gründen (z. B. zur Notendurchschnittsbildung) dennoch häufig verwendet wird. Als unstrittig zulässiges, dem Skalenniveau angemessenes Lagemaß eignet sich stattdessen der Median (alternativ auch der Modus).

Bewertung: 1 P. Voraussetzung des arithmetischen Mittels (Gleichabständigkeit/mind. Intervallskala) genannt; 1 P. Begründung, dass diese bei Ordinaldaten/Schulnoten nicht gesichert ist; 1 P. Nennung eines zulässigen Alternativmaßes (Median oder Modus)





AUFGABE 2a (5 Punkte)

Erwartung: Sortierte Daten: 3, 3, 3, 4, 4, 5, 5, 6, 7, 9, 20. Modus: Der Wert 3 kommt mit dreimaligem Auftreten am häufigsten vor, Modus = 3. Median: Bei $n = 11$ (ungerade) ist der Median der 6. Wert der sortierten Reihe: Median = 5. Arithmetisches Mittel: Summe = $3+3+3+4+4+5+5+6+7+9+20 = 69$; Mittelwert = $69 : 11 = 6,27$ (gerundet auf zwei Nachkommastellen).

Bewertung: 1 P. korrektes Sortieren der Werte; 1 P. Modus = 3 korrekt bestimmt; 1 P. Median = 5 korrekt bestimmt (mit Nennung der mittleren Position); 1 P. Summe der Werte korrekt berechnet (69); 1 P. Mittelwert korrekt berechnet und gerundet (6,27)

AUFGABE 2b (4 Punkte)

Erwartung: Das arithmetische Mittel wird am stärksten durch den Ausreißer beeinflusst, da in seine Berechnung jeder einzelne Wert (einschließlich des extremen Werts 20) direkt mit seiner vollen Größe eingeht; ohne den Ausreißer wäre der Mittelwert deutlich niedriger. Der Modus wird am wenigsten beeinflusst, da er lediglich von der Häufigkeit des Auftretens eines Werts abhängt und ein einzelner extremer Wert (20 kommt nur einmal vor) daran nichts ändert. Der Median liegt in seiner Robustheit dazwischen, da er nur von der Position in der sortierten Reihe abhängt, nicht vom konkreten Zahlenwert des Ausreißers.

Bewertung: 1 P. arithmetisches Mittel als am stärksten beeinflusstes Maß benannt; 1 P. inhaltliche Begründung dafür (jeder Wert geht mit seiner Größe ein); 1 P. Modus als am wenigsten beeinflusstes Maß benannt; 1 P. inhaltliche Begründung dafür (nur Häufigkeit zählt)





AUFGABE 2c (6 Punkte)

Erwartung: Sortierte Daten: 3, 3, 3, 4, 4, 5, 5, 6, 7, 9, 20. Spannweite: Range = Maximum - Minimum = $20 - 3 = 17$. Quartile ($n = 11$): Median (Q2) ist der 6. Wert = 5. Unterer Teildatensatz (ohne Median, erste 5 Werte): 3, 3, 3, 4, 4 - Q1 ist der mittlere Wert davon = 3. Oberer Teildatensatz (letzte 5 Werte): 5, 6, 7, 9, 20 - Q3 ist der mittlere Wert davon = 7. Interquartilsabstand: $IQR = Q3 - Q1 = 7 - 3 = 4$.

Bewertung: 1 P. korrekte Spannweite (17); 1 P. korrekte Bestimmung von Q1 (3) mit nachvollziehbarem Vorgehen; 1 P. korrekte Bestimmung von Q3 (7) mit nachvollziehbarem Vorgehen; 1 P. korrekter IQR (4); 2 P. vollständiger, nachvollziehbarer Rechenweg/Sortiervorgang dokumentiert

AUFGABE 2d (4 Punkte)

Erwartung: Die Box reicht vom ersten Quartil $Q1 = 3$ bis zum dritten Quartil $Q3 = 7$ und umfasst damit die mittleren 50 Prozent der Daten. Innerhalb der Box wird der Median bei 5 als Strich eingezeichnet. Die Whisker (Antennen) würden sich regulär bis zum größten bzw. kleinsten Wert innerhalb des 1,5-fachen IQR oberhalb $Q3$ bzw. unterhalb $Q1$ erstrecken (hier: obere Grenze = $7 + 1,5 \cdot 4 = 13$, untere Grenze = $3 - 1,5 \cdot 4 = -3$). Der Wert 20 liegt oberhalb der oberen Grenze von 13 und würde daher in einem klassischen Boxplot nicht durch den Whisker erfasst, sondern separat als einzelner Punkt (Ausreißer) eingezeichnet.

Bewertung: 1 P. Box von $Q1$ bis $Q3$ korrekt beschrieben; 1 P. Median korrekt als Strich in der Box bei 5 verortet; 1 P. korrekte Erläuterung der Whisker-Grenze (1,5-facher IQR) bzw. dass der Whisker nur bis zum letzten Wert innerhalb dieser Grenze reicht; 1 P. Ausreißer 20 korrekt als separater Punkt außerhalb des Whiskers beschrieben

AUFGABE 2e (3 Punkte)

Erwartung: Die Spannweite wird allein aus Minimum und Maximum berechnet und reagiert daher extrem empfindlich auf einzelne Extremwerte/Ausreißer - ein einziger ungewöhnlich hoher oder niedriger Wert verändert die Spannweite vollständig. Der Interquartilsabstand bezieht sich dagegen nur auf den Bereich zwischen dem ersten und dritten Quartil, also auf die mittleren 50 Prozent der Daten, wodurch extreme Einzelwerte an den Rändern der Verteilung keinen Einfluss auf seinen Wert haben. Dadurch gilt der IQR als robusteres Streuungsmaß.

Bewertung: 1 P. Spannweite basiert nur auf Minimum/Maximum, daher ausreißeranfällig; 1 P. IQR bezieht sich nur auf mittlere 50 Prozent der Daten; 1 P. Schlussfolgerung, dass Extremwerte den IQR nicht beeinflussen, die Spannweite aber schon





AUFGABE 3a (6 Punkte)

Erwartung: Summe = $12+15+11+14+13+16+12+15 = 108$. Anzahl $n = 8$.

Mittelwert = $108 : 8 = 13,5$ Minuten.

Bewertung: 2 P. korrekte Summenbildung (108); 2 P. korrekte Division durch $n = 8$; 2 P. korrektes Endergebnis mit Einheit (13,5 Minuten)

AUFGABE 3b (8 Punkte)

Erwartung: Tabelle mit Mittelwert 13,5: Wert 12 - Abweichung -1,5 - Quadrat 2,25. Wert 15 - Abweichung +1,5 - Quadrat 2,25. Wert 11 - Abweichung -2,5 - Quadrat 6,25. Wert 14 - Abweichung +0,5 - Quadrat 0,25. Wert 13 - Abweichung -0,5 - Quadrat 0,25. Wert 16 - Abweichung +2,5 - Quadrat 6,25. Wert 12 - Abweichung -1,5 - Quadrat 2,25. Wert 15 - Abweichung +1,5 - Quadrat 2,25. Summe der quadrierten Abweichungen = $2,25+2,25+6,25+0,25+0,25+6,25+2,25+2,25 = 22$. Stichprobenvarianz $s^2 = 22 : (8-1) = 22 : 7 = 3,14$. Stichprobenstandardabweichung $s = \sqrt{3,14} = 1,77$.

Bewertung: 2 P. korrekte Abweichungen in der Tabelle; 2 P. korrekte quadrierte Abweichungen in der Tabelle; 1 P. korrekte Summe der quadrierten Abweichungen (22); 1 P. korrekte Division durch $n-1 = 7$ als Ansatz; 1 P. korrekte Varianz (3,14); 1 P. korrekte Standardabweichung (1,77)

AUFGABE 3c (3 Punkte)

Erwartung: Da der Mittelwert selbst aus den Stichprobendaten geschätzt wurde, ist nur $n-1$ der n Abweichungen frei wählbar - die letzte Abweichung ergibt sich zwangsläufig aus den übrigen, da die Summe aller Abweichungen vom Mittelwert stets 0 ergeben muss. Man sagt daher, es stehen nur $n-1$ Freiheitsgrade zur Verfügung. Die Division durch $n-1$ statt durch n korrigiert zudem die sonst systematische Unterschätzung (Verzerrung) der wahren Populationsvarianz, wenn man von einer Stichprobe auf die Grundgesamtheit schließen will, und liefert dadurch einen erwartungstreuen (unverzerrten) Schätzer.

Bewertung: 1 P. Begriff der Freiheitsgrade korrekt erläutert ($n-1$ frei wählbare Werte, da Summe der Abweichungen 0 ergeben muss); 1 P. Bezug zur Schätzung des Mittelwerts aus der Stichprobe; 1 P. Hinweis auf Korrektur der Verzerrung/Erwartungstreue der Varianzschätzung





AUFGABE 3d (3 Punkte)

Erwartung: Ein direkter Vergleich der Standardabweichungen ist irreführend, da die Bearbeitungszeiten in Minuten und die Reaktionszeiten in Millisekunden gemessen werden - beide Variablen haben unterschiedliche Maßeinheiten und unterschiedlich große Skalenwerte (absolute Streuungsmaße hängen von der Einheit und dem allgemeinen Größenniveau der Werte ab). Eine Standardabweichung von z. B. 2 Minuten kann relativ zum jeweiligen Mittelwert eine ganz andere Bedeutung haben als eine Standardabweichung von 2 Millisekunden. Als geeignetes relatives (einheitenfreies) Streuungsmaß für einen fairen Vergleich eignet sich der Variationskoeffizient (Standardabweichung geteilt durch den Mittelwert, meist in Prozent angegeben).

Bewertung: 1 P. Hinweis auf unterschiedliche Maßeinheiten/Größenordnungen als Problem; 1 P. Begründung, dass absolute Streuungsmaße dadurch nicht direkt vergleichbar sind; 1 P. korrekte Nennung des Variationskoeffizienten als relatives Streuungsmaß

AUFGABE 4a (4 Punkte)

Erwartung: Relative Häufigkeiten (bezogen auf $n = 50$): Wert 1: $3 : 50 = 0,06$. Wert 2: $7 : 50 = 0,14$. Wert 3: $10 : 50 = 0,20$. Wert 4: $20 : 50 = 0,40$. Wert 5: $10 : 50 = 0,20$. Kontrolle: Summe = $0,06 + 0,14 + 0,20 + 0,40 + 0,20 = 1,00$.

Bewertung: 0,5 P. je korrekt berechnete relative Häufigkeit ($5 \times 0,5 \text{ P.} = 2,5 \text{ P.}$); 1,5 P. für korrekte Rechenwege/Darstellung und Kontrolle, dass die Summe 1 ergibt

AUFGABE 4b (5 Punkte)

Erwartung: Modus: Der Wert 4 hat mit 20 Nennungen die größte absolute Häufigkeit, Modus = 4. Median: Bei $n = 50$ wird zunächst die kumulierte Häufigkeit gebildet: bis Wert 1: 3, bis Wert 2: 10, bis Wert 3: 20, bis Wert 4: 40, bis Wert 5: 50. Der Median liegt zwischen dem 25. und 26. Wert der sortierten Reihe. Da die kumulierte Häufigkeit bis einschließlich Wert 3 erst 20 beträgt und bis einschließlich Wert 4 bereits 40, liegen sowohl der 25. als auch der 26. Wert in der Kategorie 4. Median = 4. Vorgehen allgemein: Man bildet die kumulierten (absoluten oder relativen) Häufigkeiten und bestimmt, in welcher Kategorie die 50-Prozent-Grenze (bzw. bei geradem n die mittleren beiden Positionen) erstmals erreicht bzw. überschritten wird.

Bewertung: 1 P. Modus korrekt bestimmt (4); 2 P. Median korrekt bestimmt (4) mit kumulierter Häufigkeit als Nachweis; 2 P. allgemeine Erläuterung des Vorgehens bei der Medianbestimmung aus einer Häufigkeitstabelle





AUFGABE 4c (5 Punkte)

Erwartung: Gewichtetes Mittel = $(1 \cdot 3 + 2 \cdot 7 + 3 \cdot 10 + 4 \cdot 20 + 5 \cdot 10) : 50 = (3 + 14 + 30 + 80 + 50) : 50 = 177 : 50 = 3,54$.

Bewertung: 1 P. korrekter Ansatz mit Gewichtung x_i mal f_i ; 2 P. korrekte Einzelprodukte (3, 14, 30, 80, 50); 1 P. korrekte Summe (177); 1 P. korrektes Endergebnis (3,54)

AUFGABE 4d (6 Punkte)

Erwartung: Streng genommen handelt es sich bei der Zufriedenheitsskala um eine ordinalskalierte Variable (Rangfolge von „sehr unzufrieden“ bis „sehr zufrieden“), bei der nicht gesichert ist, dass die Abstände zwischen den Antwortstufen gleich groß sind - die Berechnung eines arithmetischen Mittels setzt aber genau diese Gleichabständigkeit voraus und ist damit methodisch nicht streng zulässig. In der psychologischen Forschungspraxis werden solche mehrstufigen Rating-/Likert-Skalen dennoch sehr häufig als quasi-intervallskaliert behandelt und Mittelwerte berechnet, da dies etablierte Analyseverfahren (z. B. Mittelwertvergleiche, Varianzanalysen) ermöglicht und sich in der Praxis oft als robust gegenüber dieser Verletzung erweist. Eine methodisch strengere Alternative wäre die Berichterstattung von Median und Modus zusätzlich oder anstelle des Mittelwerts.

Bewertung: 1 P. korrekte Einordnung als ordinalskalierte Variable; 1 P. Erläuterung, dass Mittelwertbildung Gleichabständigkeit voraussetzt, die hier nicht gesichert ist; 2 P. Hinweis auf übliche Praxis, Rating-/Likert-Skalen als quasi-intervallskaliert zu behandeln und dennoch Mittelwerte zu berechnen; 2 P. kritische Einordnung/Alternative (Median/Modus als methodisch strengere Option) und abgewogenes Fazit





AUFGABE 5a (6 Punkte)

Erwartung: Formel: $z = (x - \text{Mittelwert}) : \text{Standardabweichung}$. Klausur A: $z_A = (72 - 60) : 8 = 12 : 8 = 1,5$. Klausur B: $z_B = (68 - 55) : 6 = 13 : 6 = 2,17$.

Bewertung: 1 P. korrekte Formel angegeben; 2 P. korrekte Berechnung $z_A = 1,5$; 2 P. korrekte Berechnung $z_B = 2,17$; 1 P. saubere Darstellung des Rechenwegs für beide Klausuren

AUFGABE 5b (5 Punkte)

Erwartung: Der Student hat sich in Klausur B relativ besser platziert, da sein z-Wert dort mit 2,17 höher ist als der z-Wert von 1,5 in Klausur A. Das bedeutet, dass seine Leistung in Klausur B um 2,17 Standardabweichungen über dem Durchschnitt aller Teilnehmenden lag, während sie in Klausur A nur 1,5 Standardabweichungen über dem Durchschnitt lag - obwohl die absolute Rohpunktzahl in Klausur A mit 72 Punkten höher war als in Klausur B mit 68 Punkten.

Bewertung: 1 P. richtige Schlussfolgerung (Klausur B relativ besser); 2 P. Begründung anhand des Vergleichs der z-Werte (2,17 gegenüber 1,5); 2 P. Hinweis, dass dies trotz niedrigerer absoluter Rohpunktzahl in B gilt und der Vergleich sich auf die Lage relativ zur jeweiligen Verteilung bezieht

AUFGABE 5c (4 Punkte)

Erwartung: Die z-Transformation dient dazu, Werte aus unterschiedlichen Verteilungen (mit unterschiedlichem Mittelwert und unterschiedlicher Streuung) vergleichbar zu machen, indem sie angibt, wie viele Standardabweichungen ein Wert über oder unter dem jeweiligen Mittelwert liegt. Ein direkter Vergleich der Rohpunktzahlen 72 und 68 wäre irreführend, da die beiden Klausuren unterschiedliche Mittelwerte und unterschiedliche Standardabweichungen aufweisen, also unterschiedlich schwer bzw. unterschiedlich stark streuend waren - eine gleiche Punktzahl kann in einer leichteren Klausur mit hohem Mittelwert eine schlechtere relative Leistung bedeuten als in einer schwereren Klausur mit niedrigerem Mittelwert.

Bewertung: 2 P. Zweck der z-Transformation korrekt erläutert (Vergleichbarkeit von Werten aus unterschiedlichen Verteilungen herstellen); 2 P. Begründung, warum der direkte Rohwertvergleich wegen unterschiedlicher Mittelwerte/Streuungen der beiden Klausuren irreführend ist





AUFGABE 5d (5 Punkte)

Erwartung: Eine z-standardisierte Variable hat per Definition stets einen Mittelwert von 0 und eine Standardabweichung (bzw. Varianz) von 1. Dass der Mittelwert 0 ist, folgt daraus, dass im Zähler der z-Formel von jedem Wert der Mittelwert subtrahiert wird, sodass die mittlere Abweichung über alle Werte hinweg im Durchschnitt 0 ergibt. Dass die Standardabweichung 1 ist, folgt daraus, dass die Abweichungen zusätzlich durch die Standardabweichung der Ausgangsverteilung geteilt werden, wodurch die neue Streuung auf genau eine Einheit skaliert wird.

Bewertung: 1 P. Eigenschaft Mittelwert = 0 genannt; 1 P. Eigenschaft Standardabweichung = 1 genannt; 1,5 P. Begründung für Mittelwert = 0 aus der Formel; 1,5 P. Begründung für Standardabweichung = 1 aus der Formel

