

Statistik – Übungen WS 2026

Blatt 1: Beschreibende Statistik

Daten und Merkmale

1. Die folgende Tabelle zeigt die Ergebnisse einer Befragung von 15 Mitarbeiterinnen und Mitarbeitern eines großen Universitätsinstituts mit insgesamt 132 Beschäftigten. Folgende Merkmale wurden erhoben: **Alter** (in vollendeten Lebensjahren), **Tätigkeitsbereich** (Forschung, Lehre, Administration), **Postleitzahl** (Wohnort), **Körpergröße** (in cm), **Familienstand** (0: ledig, 1: verheiratet, 2: geschieden), **Sport** (wie oft betreiben Sie Sport?) (0: nie, 1: selten, 2: mäßig, 3: häufig).

ID	Alter	Tätigkeit	Postleitzahl	Körpergröße	Familienstand	Sport
1	42	Forschung	8302	178	1	2
2	30	Forschung	1190	185	0	0
3	39	Lehre	8075	179	1	1
4	63	Lehre	8041	175	0	3
5	45	Forschung	8700	176	1	1
6	31	Forschung	8044	161	0	3
7	25	Forschung	8010	173	0	2
8	45	Lehre	8020	174	1	1
9	28	Forschung	8045	171	0	0
10	43	Administration	8010	167	2	3
11	32	Administration	8112	184	1	1
12	36	Lehre	8020	172	0	3
13	59	Forschung	8043	168	2	3
14	45	Administration	8020	181	1	1
15	36	Lehre	8042	172	0	3

- a) Erklären Sie anhand der vorliegenden Datentabelle die Begriffe
- Grundgesamtheit
 - Stichprobe
 - Merkmalsträger (Objekt)
 - Merkmal (Variable)
 - Merkmalsausprägung
 - Realisation
- b) Geben Sie bei jedem Merkmal das dazugehörige Skalenniveau an und bestimmen Sie, ob das Merkmal stetig oder diskret ist.

2. **P 1** Geben Sie bei jedem Merkmal ein Beispiel für eine konkrete Ausprägung und das dazugehörige Skalenniveau an und bestimmen Sie, ob das Merkmal stetig oder diskret ist:

Merkmal	Ausprägung	Skalenniveau	diskret/stetig
Tägliche Anzahl der Besucher einer Website			
Ladezeit einer Website			
Bewertung einer App im App-Store (1 bis 5 Sterne)			
Kundensegment einer Bank			
Umsatzrendite eines Unternehmens in % im vergangenen Geschäftsjahr			
Mitgliedschaftstyp in einem Fitnessstudio (Basic, Plus, Premium)			
Position eines Mitarbeiters in der internen Rangliste nach Verkaufszahlen			
Land, in dem der Firmensitz eines Unternehmens liegt			

Verteilung eines Merkmals, Graphische Darstellungen

3. In einer Umfrage wurden 60 Studierende der TU-Graz nach der Anzahl Ihrer Prüfungsantritte in Mathematik I gefragt. Die Antworten finden Sie in nachfolgender Tabelle zusammengefasst:

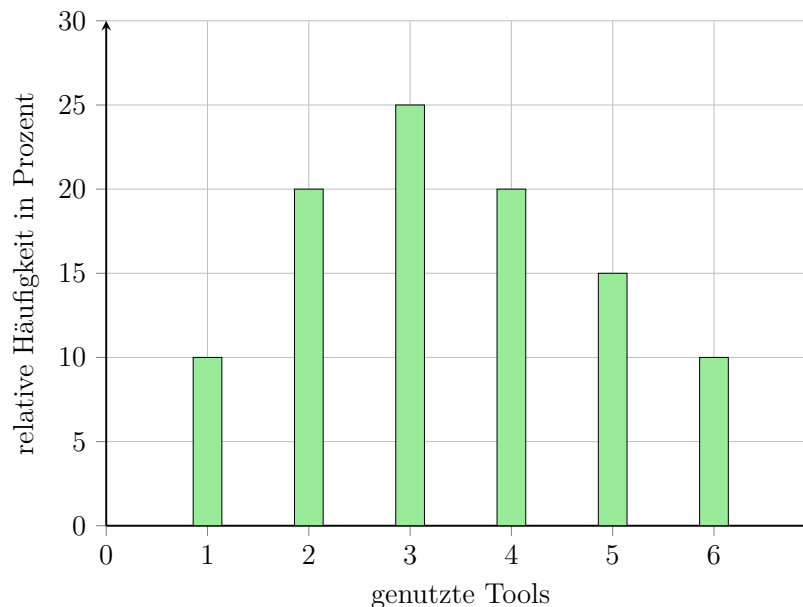
1	1	1	5	2	4	2	5	3	2	3	2
1	5	2	3	4	4	2	3	4	5	5	2
2	4	3	2	4	2	1	1	1	1	3	2
1	4	3	1	2	5	2	2	3	4	1	1
2	2	3	1	2	3	2	3	3	1	4	1

Da viele gleiche Werte auftreten und das Merkmal diskret ist, bietet es sich an, die Werte zu zählen und die Ergebnisse in einer Häufigkeitstabelle zusammenzufassen:

Anzahl der Antritte	1	2	3	4	5
absolute Häufigkeit H_i	15	18	12	9	6

- Geben Sie die relativen und die kumulierten relativen Häufigkeiten an.
- Stellen Sie die Daten durch ein Diagramm dar.
- Zeichnen Sie eine geeignete Verteilungsfunktion.
- Bestimmen und interpretieren Sie den Wert $F(3)$.

4. **P 2** In einem Hochschulprojekt wird erhoben, wie viele verschiedene digitale Werkzeuge Lehrende in einem Semester nutzen. In die Erhebung werden ausschließlich Lehrende einbezogen, die mindestens ein digitales Werkzeug verwenden. Die LV-Leitung wird Ihnen konkrete Prozentwerte vorgeben; das nachstehende Diagramm zeigt beispielhaft die relative Häufigkeiten (in %) für 1 bis 6 genutzte Tools bei insgesamt $n = 80$ Lehrenden:



- Ergänzen Sie in nachstehender Tabelle die absoluten, relativen und kumulierten relativen Häufigkeiten.
- Zeichnen Sie die Verteilungsfunktion $F(x)$.
- Bestimmen und interpretieren Sie $F(4)$ im Kontext der digitalen Lehrpraxis.

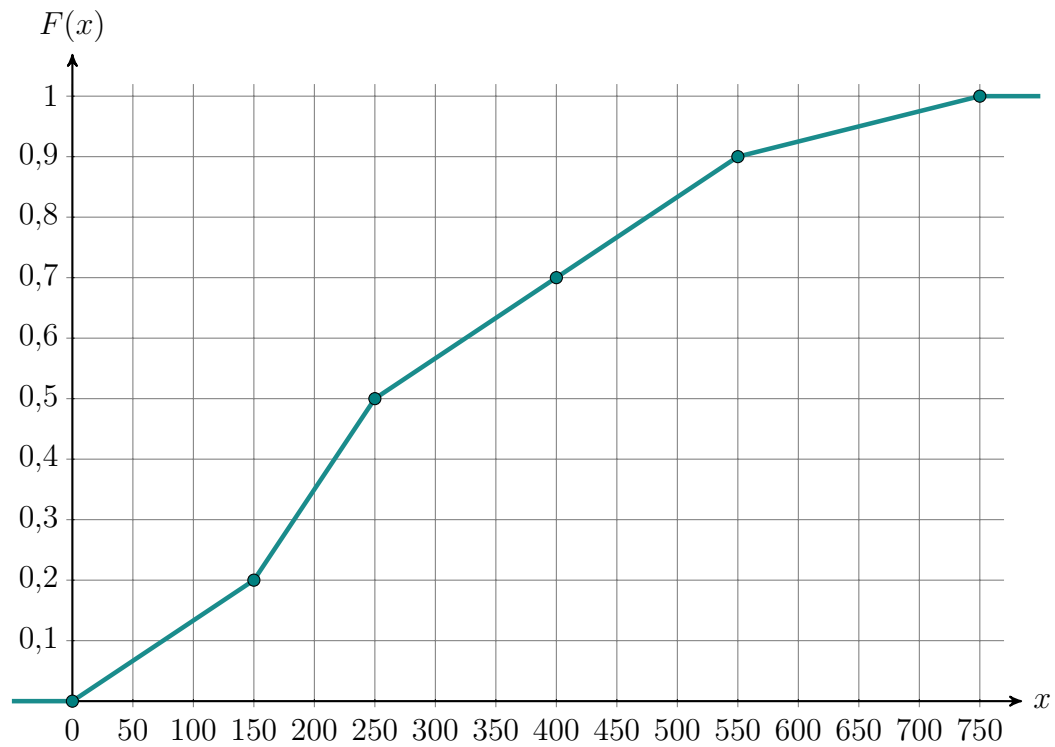
Anzahl genutzter Tools	1	2	3	4	5	6
absolute Häufigkeit H_i						
relative Häufigkeit h_i						
$F(x)$						

5. An einer Statistik-Klausur haben 200 Studierende teilgenommen. Die erreichten Punkte wurden erfasst und in folgender Tabelle dargestellt:

erreichte Punkte	$[0; 50]$	$]50; 60]$	$]60; 70]$	$]70; 80]$	$]80; 100]$
absolute Häufigkeit H_i	70	50	40	30	10

- Ergänzen Sie in der Häufigkeitstabelle die relativen Häufigkeiten sowie die absolute und die relative Besetzungsdichte.
- Stellen Sie die Häufigkeitsverteilung der Punkte als Histogramm dar. Berücksichtigen Sie dabei insbesondere die unterschiedliche Klassenbreite des klassierten Merkmals „erreichte Punkte“.
- Zeichnen Sie die geeignete Verteilungsfunktion.

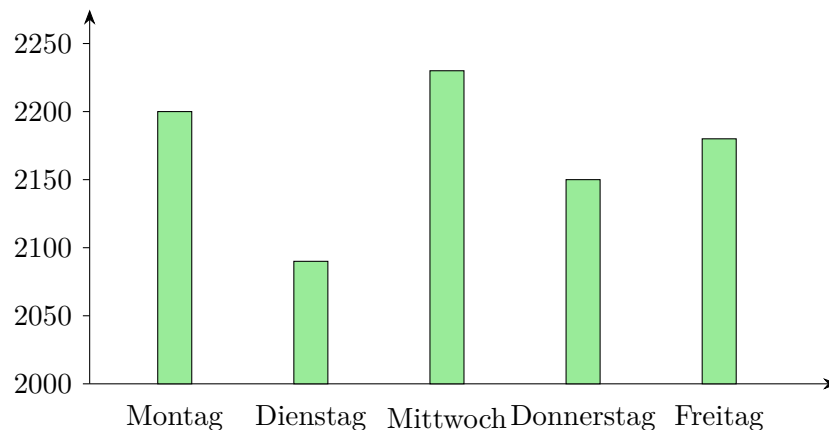
6. **P 3** Eine Bank beobachtet die monatlichen Sparraten (in 100 €) von 1.500 Kundinnen und Kunden eines neuen Sparprodukts. Die folgende Verteilungsfunktion wurde aus den Daten abgeleitet:



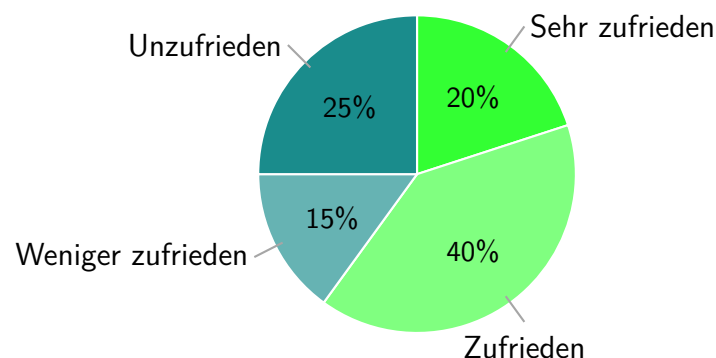
- Stellen Sie die zugrunde liegende Häufigkeitsverteilung der Sparraten in Form eines Histogramms dar.
- Wie groß ist der Anteil der Sparraten, die größer als 400 € sind?
- Bestimmen Sie näherungsweise den Anteil der Sparraten, der zwischen 150 € und 650 € liegt.

7. **P 4** Das folgende Säulendiagramm zeigt die Gesamtzahl der Besuche in einem Jugendzentrum im gesamten vergangenen Jahr, aufgeteilt nach den Wochentagen Montag bis Freitag. In einem Artikel einer Lokalzeitung wird das Diagramm verwendet, um zu behaupten: „Man kann aus dem Diagramm ablesen, dass die Besucherzahl am Dienstag weniger als die Hälfte der Besucherzahl am Mittwoch beträgt.“

Anzahl der Besuche



- Argumentieren Sie anhand der Grafik, warum die Aussage der Zeitung nicht stimmt.
- Im vergangenen Jahr wurden im Jugendzentrum insgesamt 10.840 Besuche gezählt. Berechnen Sie die jährlichen Einnahmen des Jugendzentrums aus Eintrittsgeldern, wenn angenommen wird, dass 80 % der Besucher einen regulären Eintrittspreis von 3,00 € zahlen und 20 % einen ermäßigten Preis von 1,50 € zahlen.
- Um die Zufriedenheit der Jugendlichen mit dem Angebot des Jugendzentrums zu erfassen, führt das Sozialforschungsbüro der Stadt eine Umfrage durch. Die Jugendlichen können das Angebot mit den Kategorien „Sehr zufrieden“, „Zufrieden“, „Weniger zufrieden“ und „Unzufrieden“ bewerten. Das folgende Kreisdiagramm gibt eine Zusammenfassung der Ergebnisse wieder.



- Berechnen Sie, wie viele Personen an der Umfrage teilgenommen haben, wenn 48 Jugendliche die Kategorie „Unzufrieden“ angekreuzt haben.
- Berechnen Sie, um wie viel Prozent die Zahl der Jugendlichen, die mit „Weniger zufrieden“ abgestimmt haben, kleiner ist als die Zahl derjenigen, die „Zufrieden“ angegeben haben.

Lagemaße, Streuungsmaße, Konzentrationsmaße

8. Die 14 Filialen eines Kaufhauskonzerns erzielten 2025 folgende Umsätze (in Mio. Euro):

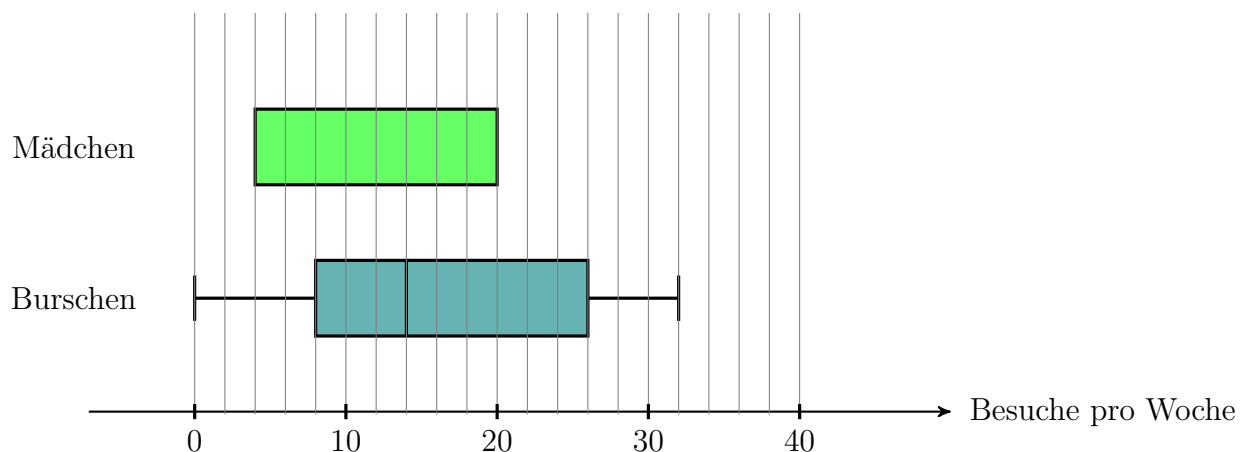
Filiale i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Umsatz x_i	4	2	7	6	7	3	3	13	4	2	9	18	5	2

Berechnen Sie für die angegebenen Umsatzwerte:

- den Median, das 1. und das 3. Quartil, sowie Spannweite und den Interquartilsabstand und zeichnen Sie einen Boxplot.
 - das arithmetische Mittel und die Standardabweichung.
 - Beurteilen Sie anhand des Boxplot, ob die Verteilung linksschief/rechtsschief/symmetrisch ist.
9. P 5 Die Nutzung einer bestimmten Internetplattform durch Jugendliche wird getrennt nach Mädchen und Burschen untersucht. Dabei wird erhoben, wie häufig die befragten Jugendlichen diese Plattform pro Woche besuchen. Die folgenden Boxplots stellen die Ergebnisse der Untersuchung für beide Gruppen dar. Für die Gruppe der Mädchen sind

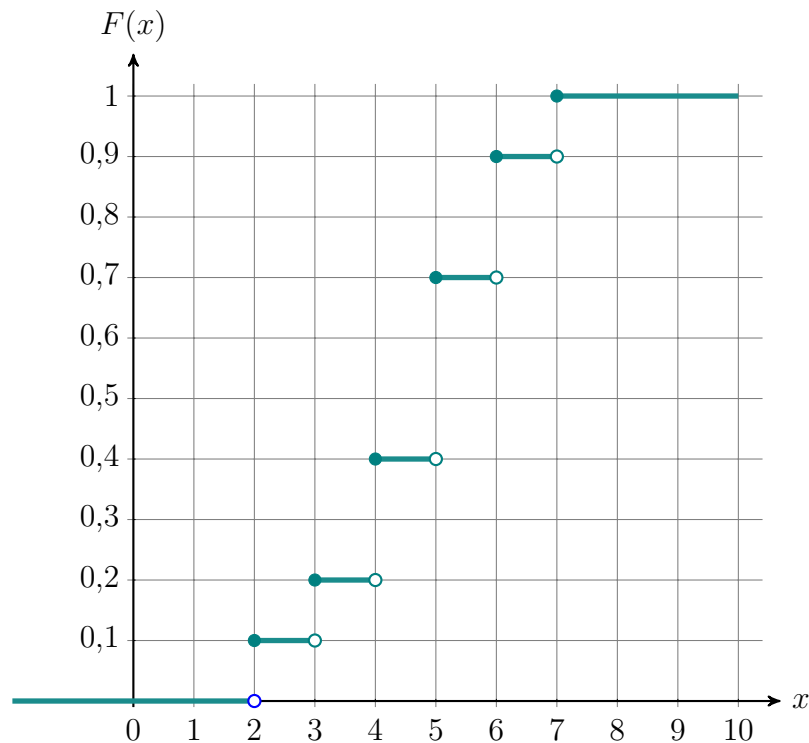
folgende Daten bekannt:

Maximum	34
Spannweite	32
Median	12



- Vervollständigen Sie den Boxplot der Mädchen.
- Schätzen Sie anhand der Form des Boxplots der Mädchen ab, ob das arithmetische Mittel vermutlich größer, kleiner oder ungefähr gleich dem Median ist. Begründen Sie Ihre Entscheidung.
- Aus dem Boxplot für die Burschen kann Folgendes abgelesen werden: Mindestens ein Viertel der Burschen hat eine Anzahl an Besuchen pro Woche kleiner oder gleich einem Wert a und zugleich haben mindestens drei Viertel der Burschen eine Anzahl an Besuchen größer als oder gleich diesem Wert a . Geben Sie diesen Wert a an.

10. **P 6** In einem Weiterbildungsseminar wurden die wöchentlichen Lernzeiten (in Stunden) von 200 Teilnehmenden erhoben und in folgender empirischen Verteilungsfunktion dargestellt:



Berechnen Sie

- die durchschnittliche Lernzeit.
 - den Median sowie
 - die Standardabweichung des Merkmals „wöchentliche Lernzeit“.
11. In der nachstehenden Tabelle ist die Anzahl der Kleinstunternehmen (Umsatz zwischen 0 und 10 Mio. Euro), Kleinunternehmen (Umsatz zwischen 10 und 20 Mio. Euro) und mittleren Unternehmen (Umsatz zwischen 20 und 60 Mio. Euro) in einer bestimmten Region angegeben:

Umsatz (in Mio. €) [von; bis[	[0; 10[	[10; 20[	[20; 60]
Anzahl der Unternehmen	40	80	80

Berechnen sie näherungsweise das arithmetische Mittel sowie die Standardabweichung des Merkmals Umsatz.

12. **P 7** In einer Studie zur Arbeitszufriedenheit in einem Callcenter werden für 14 Mitarbeitende folgende Merkmale erfasst: **Alter** (in Jahren), **Team** (Inbound, Outbound, Support), **Arbeitszufriedenheit** (Skala 1 (sehr niedrig) bis 5 (sehr hoch)), **Beschäftigungsmaß** (Teilzeit/Vollzeit), **Schulungsstatus** (0: keine, 1: geplant, 2: absolviert), **Anzahl der Pausen pro Schicht** (in ganzen Zahlen).

ID	Alter	Team	Zufriedenheit	Beschäftigung	Schulung	Pausen
1	24	Inbound	3	Vollzeit	2	3
2	31	Outbound	2	Teilzeit	1	2
3	29	Support	4	Vollzeit	0	4
4	42	Inbound	5	Vollzeit	2	3
5	35	Inbound	2	Teilzeit	0	1
6	27	Support	4	Vollzeit	1	3
7	39	Inbound	3	Vollzeit	2	2
8	30	Outbound	1	Teilzeit	1	2
9	33	Support	4	Vollzeit	2	3
10	26	Inbound	3	Teilzeit	0	2
11	41	Outbound	2	Vollzeit	2	1
12	28	Support	5	Vollzeit	1	4
13	37	Inbound	3	Vollzeit	0	3
14	32	Outbound	1	Teilzeit	1	2

- a) Berechnen Sie für die von der LV-Leitung genannten Merkmale jeweils ein geeignetes Lagemaß und begründen Sie Ihre Wahl.
- b) Berechnen Sie für das Merkmal „Alter“ ein geeignetes Streuungsmaß und begründen Sie Ihre Wahl.
13. In einem Betrieb mit 18 Beschäftigten wurden Informationen über das jährliche Einkommen erhoben. Dazu wurde von den 7 weiblichen und den 11 männlichen Beschäftigten des Betriebes das Jahreseinkommen erfasst. Bei den Frauen ergaben sich folgende Werte (in 1.000 €):

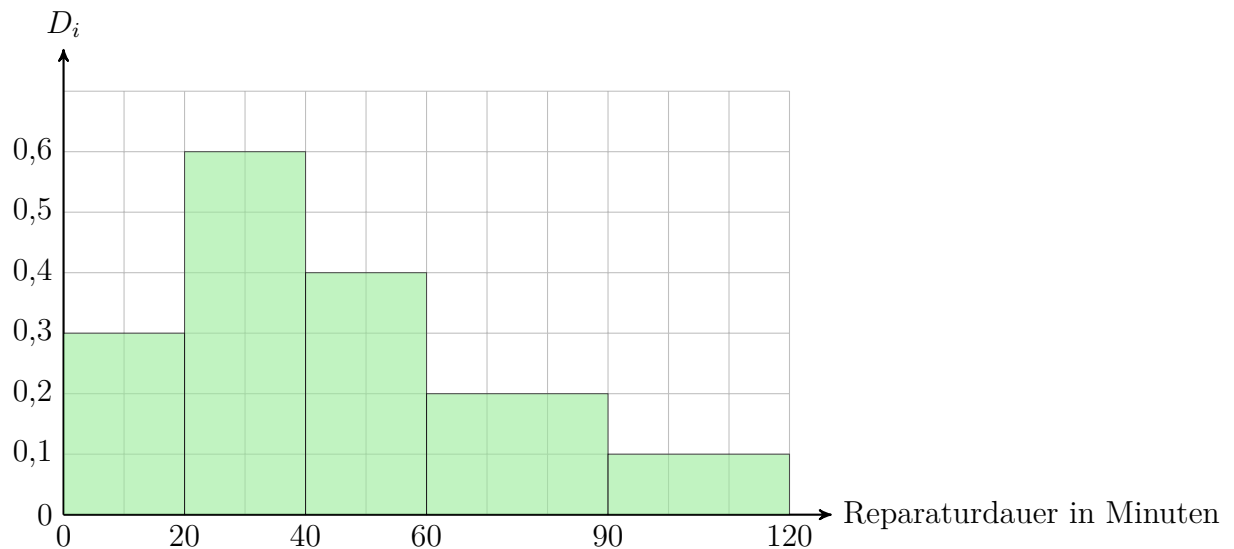
22,4	26,0	14,9	35,1	50,2	20,5	30,8
------	------	------	------	------	------	------

Für die Gruppe der Männer sind folgende Werte bekannt:

Arithmetische Mittel der Jahreseinkommen (in € 1.000)	Standardabweichung der Jahreseinkommen	Anzahl an befragten Männern
30,4	11,7	11

- a) Berechnen Sie für die Frauen Mittelwert und Standardabweichung der Jahreseinkommen.
- b) Vergleichen Sie die Streuung der Jahreseinkommen zwischen Frauen und Männern. Bei welcher Gruppe ist die Streuung höher?
- c) Bestimmen Sie unter Verwendung von a) Mittelwert und Standardabweichung der Jahreseinkommen für die Gesamtheit aller 18 Beschäftigten.

14. **P 8** Eine Fahrradwerkstatt dokumentiert für eine Woche die Dauer der Reparaturaufträge (nur Arbeitszeit, in Minuten). Zur Auswertung werden die Daten klassiert und in einem Histogramm dargestellt:



- a) Ergänzen Sie die nachstehende Tabelle der absoluten Häufigkeiten:

Reparaturdauer (in Minuten)	[0; 20[	[20; 40[	[40; 60[	[60; 90[	[90; 120]
H_i					

- b) Berechnen Sie näherungsweise die mittlere Reparaturdauer.
c) Berechnen Sie näherungsweise die Standardabweichung der Reparaturdauer.

15. **P 9** In einer Marktstudie eines großen Online-Händlers wurden die durchschnittlichen Ausgaben pro Einkauf (Warenkorbwert in Euro) für das Jahr 2025 in drei verschiedenen Altersgruppen untersucht. Man erhielt folgende Daten:

Altersgruppe	Mittelwert des Warenkorbwerts (in €)	Standardabweichung des Warenkorbwerts (in €)	Anzahl der befragten Kunden
Gen Z (16–26 Jahre)	120	28	500
Millennials (27–42 Jahre)	100	25	600
Gen X (43–58 Jahre)	90	25	450

- a) In welcher Altersgruppe ist die Streuung der Ausgaben am höchsten?
b) Berechnen Sie den Gesamtdurchschnitt sowie die Gesamtstandardabweichung des Warenkorbwerts für diese drei Altersgruppen zusammen.

16. In einem Land beträgt das gesamte Bruttovermögen (Sachvermögen und Finanzvermögen, ohne Schulden) 1.000 Mrd. Euro. Dieses verteilt sich auf vier Haushaltsgruppen („Untere Hälfte“, „Obere Mitte“, „Vermögende“ und „Top 5“). Die Gruppe „Untere Hälfte“ umfasst dabei 50 %, „Obere Mitte“ 30 %, „Vermögende“ 15 % und „Top 5“ 5 % aller Haushalte. Die Aufteilung des gesamten Bruttovermögens auf die vier Haushaltsgruppen ist in nachstehender Tabelle zusammengefasst:

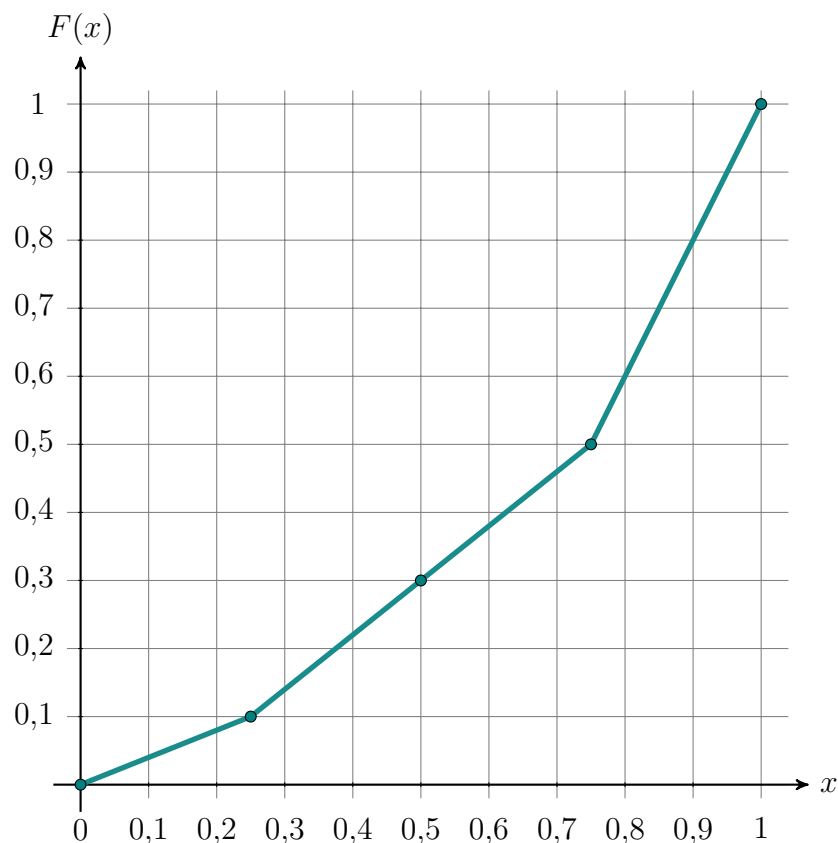
Haushaltsgruppe	Untere Hälfte	Obere Mitte	Vermögende	Top 5
Bruttovermögen (in Mrd. Euro)	50	200	300	450

Es soll die Konzentration des Bruttovermögens untersucht werden.

- Zeichnen Sie die Lorenzkurve.
 - Bestimmen Sie den Gini-Koeffizienten für das Merkmal „Bruttovermögen“ und interpretieren Sie Ihr Ergebnis.
17. **P 10** Ein privates Gründer-Duo hat zwei erfolgreiche Start-ups aufgebaut und möchte nun die interne Anteilsstruktur der beiden Unternehmen analysieren, bevor eine neue große Finanzierungsrunde startet. Für die Gründer ist es wichtig zu wissen, wie stark die Stimmrechte konzentriert sind. Bei Start-up A (FinTech) teilen sich die drei Hauptgesellschafter die Firmenanteile (bewertet in Millionen Euro) aktuell wie folgt auf:

Gesellschafter	G_1	G_2	G_3
Firmenanteil (in Mio. €)	20	40	140

Bei Start-up B (HealthTech) teilen sich vier Investoren das Unternehmen. Die Konzentration der Firmenanteile ist hier durch folgende Lorenzkurve gegeben:



Für strategische Verhandlungen soll die Konzentration der Anteile (Stimmrechte) als Entscheidungsgrundlage dienen.

- a) Erstellen Sie die Lorenzkurve und berechnen Sie den Gini-Koeffizienten für die Anteilsverteilung bei Start-up A.
- b) Bei welchem der beiden Start-ups ist die Gefahr einer „Vormachtstellung“ (Dominanz eines einzelnen Anteilseigners) geringer, weil die Anteile gleichmäßiger verteilt sind (d. h. eine möglichst geringe Konzentration vorliegt)? (Sie können auch anhand der grafischen Darstellung der Lorenzkurven argumentieren.)

Zusammenhangsmaße, Regressionsgerade

18. In einem Experiment zur Wirkung von Musik auf die Ausgabebereitschaft beim Einkauf in einem Supermarkt wurden 100 Kunden zufällig ausgewählt. Ein Teil der Kunden tätigte seine Einkäufe an Tagen, an denen im Supermarkt keine „Hintergrundmusik“ eingespielt wurde. Der andere Teil kaufte an einem Tag, an dem der Einkauf durch Musik und Werbeansagen begleitet wurde, ein. Jeder Kunde wurde hinsichtlich der Gesamtsumme der getätigten Einkäufe in eine der drei Gruppen der Ausgabebereitschaft (hoch, mittel und gering) eingeordnet. Es ergab sich folgende (bivariate) Verteilung:

	Ausgabebereitschaft		
Ambiente	gering	mittel	hoch
Mit Musik	8	16	21
Ohne Musik	24	20	11

Berechnen Sie eine geeignete Kennzahl, um eine Aussage über eine mögliche Abhängigkeit der beiden Merkmale treffen zu können!

19. P 11 Eine Online-Lernplattform möchte wissen, ob es einen Zusammenhang zwischen dem Lernerfolg (gering oder hoch) und dem Nutzungstyp (Gelegenheitsnutzer oder Intensivnutzer) gibt. Zu diesem Zweck wurden 140 Nutzer befragt.

35 der befragten Gelegenheitsnutzer gaben einen geringen Lernerfolg an. Von den 85 Intensivnutzern berichteten 63 einen hohen Lernerfolg.

- a) Welche Skalenniveaus haben hier die Variablen „Lernerfolg“ und „Nutzungstyp“?
- b) Tragen Sie die entsprechenden absoluten Häufigkeiten in die nachstehende Tabelle ein und bilden Sie alle Zeilen- und Spaltensummen.
- c) Geben Sie die Tabelle mit den unter Unabhängigkeit erwarteten Häufigkeiten an.
- d) Berechnen Sie eine geeignete Kennzahl, um eine Aussage über eine mögliche Abhängigkeit der beiden Merkmale treffen zu können und interpretieren Sie diese im Kontext zur Aufgabenstellung.

H_{ij}	Lernerfolg gering	Lernerfolg hoch	Σ
Gelegenheitsnutzer			
Intensivnutzer			
Σ			

20. Für acht Studierende einer Statistikübung wurde die Selbsteinschätzung ihrer Statistikkenntnisse (A: sehr gut, B: gut, C: teils gut/teils schlecht, D: schlecht, E: sehr schlecht) vor Beginn der Übung, sowie die erreichte Punktezahl am Ende der Übung erhoben. Gibt es einen Zusammenhang zwischen diesen beiden Merkmalen? Berechnen Sie eine geeignete Kennzahl.

Studierender	1	2	3	4	5	6	7	8
Statistikkenntnisse	D	B	C	D	E	A	C	D
erreichte Punkte	21	13	26	19	14	23	24	12

21. **P 12** Ein Fitnessstudio möchte wissen, ob zwischen **Trainingshäufigkeit** pro Woche und **subjektiver Fitnessbewertung** (Skala von 1 = sehr schlecht bis 10 = sehr gut) ein Zusammenhang besteht. Es wurden von 12 Personen Daten erhoben:

Person	1	2	3	4	5	6	7	8	9	10	11	12
Trainingshäufigkeit (Einheiten/Woche)	0	4	1	2	6	2	3	7	3	4	5	5
Fitnessbewertung (1–10)	3	7	4	5	6	6	6	10	7	7	7	8

Gibt es einen Zusammenhang zwischen Trainingshäufigkeit und subjektiver Fitnessbewertung? Berechnen Sie eine geeignete Kennzahl und begründen Sie die Wahl dieser Kennzahl. Interpretieren Sie Ihren berechneten Kennwert hinsichtlich der Richtung und Stärke des Effekts. Welche inhaltliche Schlussfolgerung kann das Fitnessstudio aus diesem Ergebnis ziehen?

22. Die Marketingabteilung einer Handelskette möchte wissen, ob ihre Werbemaßnahmen wirken. Die Buchhaltung liefert Informationen über die monatlichen Umsätze. Die Umsätze von 6 aufeinanderfolgenden Monaten mit den entsprechenden Marketingausgaben liefern folgende Daten (Beträge in 1.000 Euro):

Monat	Jänner	Februar	März	April	Mai	Juni
Marketingausgaben (in Tsd. €)	24	16	19	26	12	18
Umsatz (in Tsd. €)	195	190	220	250	150	185

- Erstellen Sie ein Streudiagramm des Datensatzes.
- Bestimmen Sie den Pearson-Korrelationskoeffizienten.
- Geben Sie die Gleichung der Regressionsgeraden zur Beschreibung der Abhängigkeit des Umsatzes von den Marketingausgaben an.
- Welchen Umsatz würde die Handelskette erzielen, wenn Marketingausgaben in Höhe von 10.000 € geplant werden?

23. **P 13** Ein mittelständischer Hersteller von Hautpflegeprodukten untersucht die monatlichen Produktionsmengen und die monatlichen Gesamtkosten bei der Herstellung seiner Feuchtigkeitscreme. Die Controlling-Abteilung möchte anhand der Daten des Jahres 2025 untersuchen, ob zwischen der monatlichen Produktionsmenge x (in Stück) und den monatlichen Gesamtkosten y (in €) ein linearer Zusammenhang besteht. Für die 12 Monate des Jahres 2025 wurden folgende Kennzahlen ermittelt:
- Produktionsmenge: Insgesamt 120.000 Stück im untersuchten Jahr.
 - Die durchschnittlichen Gesamtkosten der Produktion in € betragen das Zehnfache der durchschnittlichen Produktionsmenge pro Monat.
 - Standardabweichung der monatlichen Produktionsmengen im Untersuchungszeitraum: 9.000 Stück.
 - Standardabweichung der monatlichen Gesamtkosten: 50.000 €.
 - Als Maß für den Zusammenhang zwischen Produktionsmenge und Gesamtkosten wurde ein Korrelationskoeffizient von 0,90 ermittelt.
- a) Berechnen Sie die Regressionsgerade für den Fall eines einfachen linearen Zusammenhangs der beiden Merkmale auf Monatsebene in der Form $y = a + b \cdot x$.
- b) Interpretieren Sie den Wert der Steigung im gegebenen ökonomischen Zusammenhang.
- c) In einem bestimmten Monat betrug die Produktionsmenge 12.000 Stück und die tatsächlichen Gesamtkosten lagen bei 120.000 €. Berechnen Sie die Differenz zwischen den tatsächlichen Kosten und den vom Modell vorhergesagten Kosten für diesen Monat. Die Regressionsfunktion wurde aus realen Daten geschätzt. Warum kann der tatsächliche Kostenwert eines Monats dennoch vom geschätzten Wert abweichen? Erläutern Sie kurz die Bedeutung dieser Abweichung.

Die mit **P** gekennzeichneten Beispiele sind von den Studierenden vorzubereiten und nach Aufruf durch die Lehrveranstaltungsleitung zu präsentieren!