

Statistik – Übungen WS 2025

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	58	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 und das dazugehörige Skalenniveau an und bestimmen Sie, ob das Merkmal stetig oder diskret ist:

Merkmal	Ausprägung	Skalenniveau	diskret/stetig
Programmiersprache			
Kreditkartennummer			
Pulsfrequenz in Schlägen pro Minute			
Status des Emirates Skywards Treueprogramms			
täglicher Kalorienverbrauch			
Schmerzskala von 1–10			
Sprachniveau (GER)			
Entfernung zwischen zwei Städten (in km)			

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

- a) Geben Sie die relativen und die kumulierten relativen Häufigkeiten an.
 - b) Stellen Sie die Daten durch ein Diagramm dar.
 - c) Zeichnen Sie eine geeignete Verteilungsfunktion.
 - d) Bestimmen und interpretieren Sie den Wert $F(3)$.
4. P 2 In einer Untersuchung zur kulturellen Aktivität wurde bei 40 Befragten die Anzahl der besuchten Kulturveranstaltungen im letzten Jahr erhoben. Die Antworten sind in der folgenden Tabelle zusammengefasst:

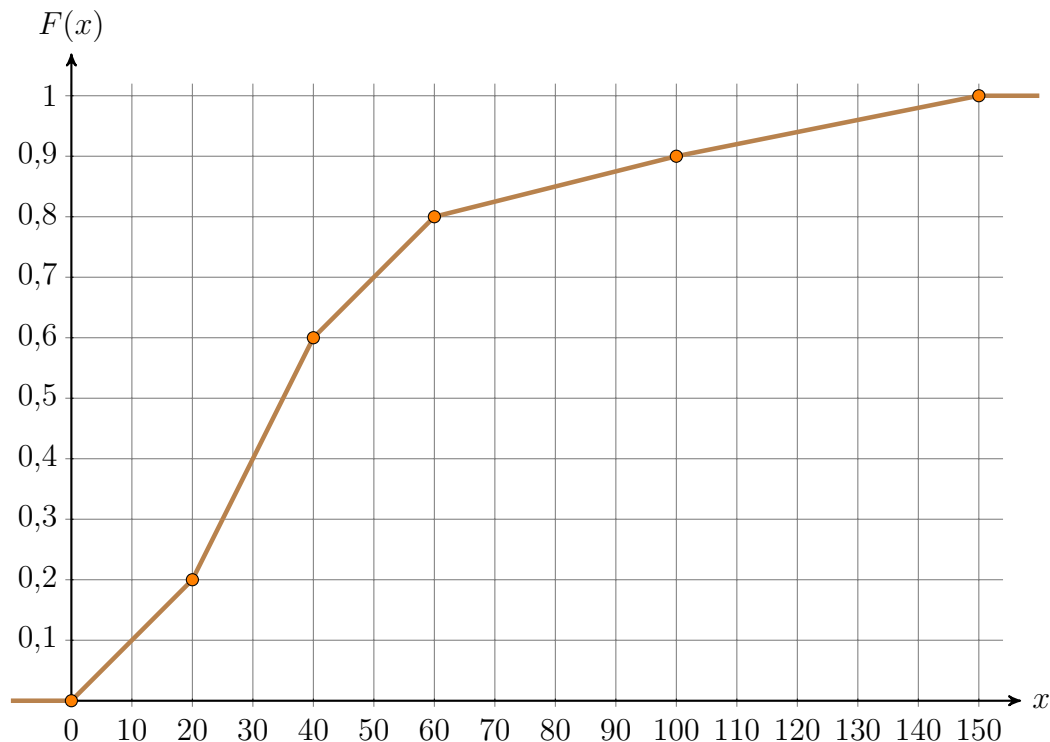
1	0	1	0	2	0	2	3	3	1
1	0	2	3	0	4	2	3	4	0
1	0	3	2	0	2	1	1	0	0
0	1	3	1	0	0	1	1	0	1

- a) Erstellen Sie aus diesen Werten eine Häufigkeitstabelle. In dieser soll die absolute Häufigkeit, die relative Häufigkeit und die relative kumulierte Häufigkeit dargestellt werden.
- b) Zeichnen Sie eine geeignete Verteilungsfunktion.
- c) Bestimmen und interpretieren Sie den Wert $F(2)$!

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

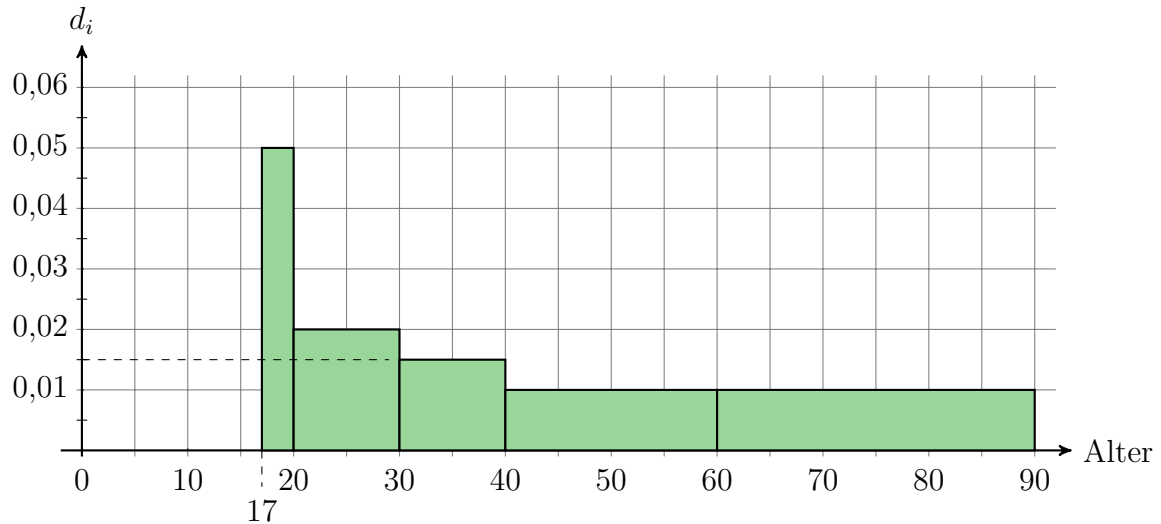
erreichte Punkte	[0; 50]	]50; 60]	]60; 70]	]70; 80]	]80; 100]
absolute Häufigkeit H_i	70	50	40	30	10

- 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** An der Selbstbedienungskasse eines Supermarktes werden die Rechnungsbeträge (x) von 120 Kunden erfasst. Es ergibt sich folgende Häufigkeitsverteilung:



- Stellen Sie die Häufigkeitsverteilung der Rechnungsbeträge als Histogramm dar.
- wie groß ist der Anteil der Rechnungsbeträge, die größer als € 60 sind?
 - Wie groß ist näherungsweise der Anteil der Rechnungsbeträge, die zwischen € 30 und € 50 liegen?

7. **P 4** Um den Einfluss des Alters von Kraftfahrzeuglenkern auf die Häufigkeit der Beteiligung an Unfällen mit Todesfolge zu untersuchen, wurde eine Erhebung durchgeführt. Dabei wurde an 1.000 zufällig ausgewählten Verkehrsunfällen mit tödlichem Ausgang das Alter des Fahrers festgestellt. In der nachstehenden Abbildung sind die Ergebnisse dieser Erhebung dargestellt:



- Geben Sie tabellarisch die absoluten, relativen und kumulierten relativen Häufigkeiten an.
- Zeichnen Sie die zugehörige Verteilungsfunktion.
- Bestimmen und interpretieren Sie den Wert $F(40)$!

Lagemaße, Streuungsmaße, Konzentrationsmaße

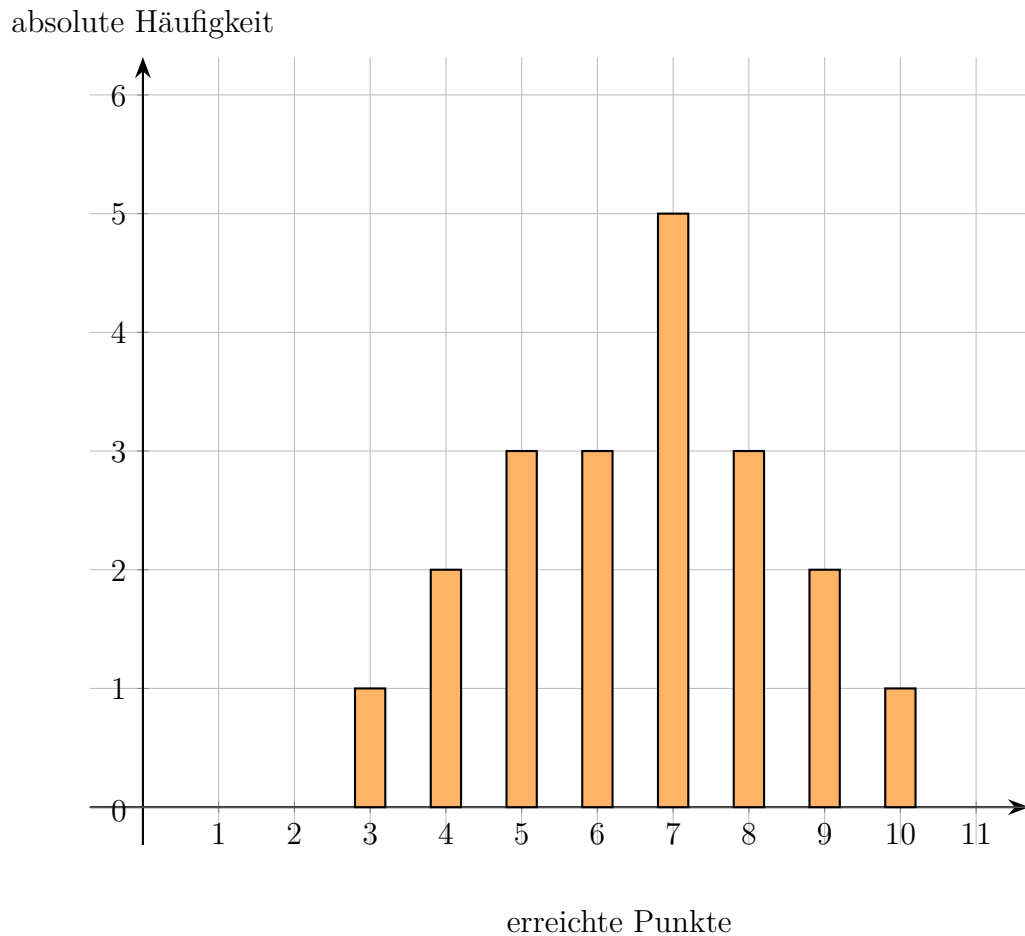
8. Die 14 Filialen eines Kaufhauskonzerns erzielten 2024 folgende Umsätze (in Mio. Euro):

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

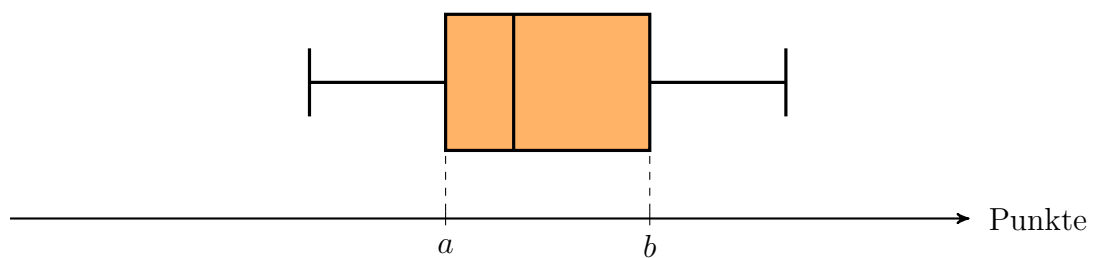
Berechnen Sie für die angegebenen Umsatzwerte:

- den Median, das 1. und das 3. Quartil, sowie Spannweite und den Quartilsabstand 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 20 Studierende nahmen an einem Aufnahmetest teil, bei dem maximal 10 Punkte erreicht werden konnten. Das nachfolgende Stabdiagramm zeigt die absoluten Häufigkeiten der erreichten Punkte:



Die bei diesem Aufnahmetest erreichten Punkte sind im nachstehenden Boxplot dargestellt:



- Berechnen Sie die Werte a und b .
- Berechnen Sie das arithmetische Mittel und die Standardabweichung der erreichten Punkte.
- Bestimmen Sie die Spannweite und den Interquartilsabstand.

10. **P 6** Für die 25 Beschäftigten einer GmbH wurde die Anzahl an Überstunden für die letzte Woche erhoben:

Anzahl Überstunden	0	1	2	3	4	12
Anzahl Beschäftigte	3	10	5	4	2	1

Berechnen Sie

- die durchschnittliche Überstundenzahl
 - den Median sowie
 - die Standardabweichung des Merkmals „Überstunden“.
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 soll der Einfluss sozioökonomischer Faktoren, Lebensstile und individuelle Gesundheitspraktiken auf das Wohlbefinden untersucht werden. Die folgende Tabelle zeigt die Ergebnisse einer Umfrage von 10 zufällig ausgewählten Personen, die mit nachstehendem Fragebogen durchgeführt wurde:

ID	F1	F2	F3	F4	F5	F6	F7
1	1	38	178	2,5	1	1	4
2	2	28	170	4	2	2	2
3	2	36	154	1,2	2	3	1
4	3	47	151	0,5	3	4	3
5	1	51	175	4,2	5	3	2
6	1	61	178	5	1	1	1
7	2	22	171	1,5	5	2	3
8	1	48	185	0,7	1	2	1
9	2	19	163	1,8	2	4	3
10	2	22	157	2,8	3	3	4

Fragebogen

Sehr geehrte Teilnehmerinnen und Teilnehmer, wir möchten Sie herzlich einladen, an unserer Umfrage teilzunehmen.

Zweck dieser Umfrage ist es, ein besseres Verständnis dafür zu entwickeln, wie sozioökonomische Faktoren, Lebensstile und individuelle Gesundheitspraktiken unser Wohlbefinden beeinflussen. Ihre Teilnahme liefert wertvolle Einblicke in Zusammenhänge zwischen dem sozialen Umfeld, der persönlichen Lebensweise und der Gesundheit.

Alle Ihre Angaben werden streng vertraulich behandelt werden. Die Ergebnisse dieser Befragung werden anonymisiert und dienen ausschließlich wissenschaftlichen Zwecken.

F1	Welches Geschlecht haben Sie?	männlich	☐	(1)
		weiblich	☐	(2)
		divers	☐	(3)
F2	Wie alt sind Sie?	(Jahre)		
F3	Wie groß sind Sie?	(cm)		
F4	Wie viele Stunden nutzen Sie soziale Medien pro Tag?	(h)		
F5	Wie würden Sie Ihre Ernährungsgewohnheiten beschreiben? Wählen Sie bitte die Option, die am besten zutrifft.	Omnivor	☐	(1)
		Vegetarisch	☐	(2)
		Vegan	☐	(3)
		Pescetarisch	☐	(4)
		Flexitarisch	☐	(5)
F6	Wie oft betreiben Sie pro Woche Sport?	Weniger als 1 Mal	☐	(1)
		1-2 Mal	☐	(2)
		3-4 Mal	☐	(3)
		5 Mal oder mehr	☐	(4)
F7	Wie beurteilen Sie Ihren derzeitigen Gesundheitszustand?	sehr gut	☐	(1)
		gut	☐	(2)
		mittel	☐	(3)
		weniger gut	☐	(4)
		sehr schlecht	☐	(5)

- a) Berechnen Sie für die Merkmale, die Sie von Ihrer LV-Leitung genannt bekommen, jeweils ein geeignetes Lagemaß und begründen Sie Ihre Wahl.
- b) Berechnen Sie für das Merkmal F4 ein geeignetes Streuungsmaß und begründen Sie Ihre Wahl.

Omnivor (Ich konsumiere sowohl Pflanzen als auch Tiere.)
Vegetarisch (Ich meide Fleisch und Fisch, konsumiere jedoch Milchprodukte und Eier.)
Vegan (Ich meide alle tierischen Produkte.)
Pescetarisch (Meine Ernährung ist hauptsächlich pflanzlich, inkludiert aber Fisch.)
Flexitarisch (Ich esse meist pflanzliche Kost, gelegentlich jedoch auch Fleisch/Fisch.)

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):

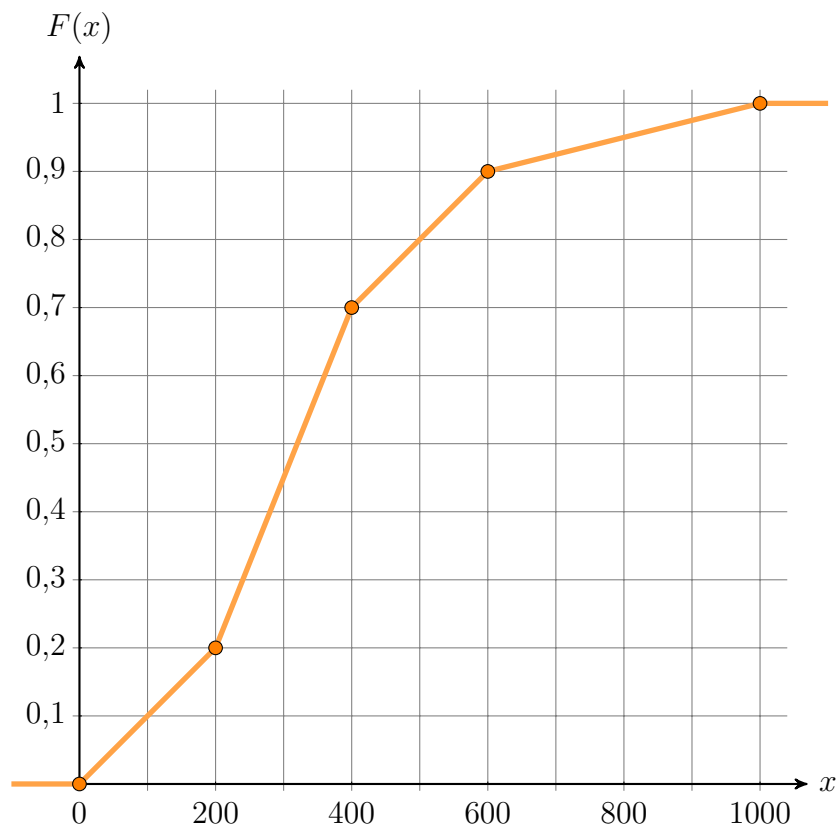
22,4	26,0	14,9	35,1	50,2	20,5	30,8
------	------	------	------	------	------	------

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 Bei einer Untersuchung über den täglichen Wasserverbrauch (in Liter pro Tag) von privaten Haushalten in einer österreichischen Großstadt ergab sich folgende Häufigkeitsverteilung:



- a) Ergänzen Sie die nachstehende Tabelle

Wasserverbrauch (in Liter/Tag)	$[0; 200[$	$[200; 400[$	$[400; 600[$	$[600; 1.000[$
h_i				

- b) Berechnen Sie näherungsweise den mittleren täglichen Wasserbrauch.
c) Berechnen Sie näherungsweise die Standardabweichung des täglichen Wasserverbrauchs.

15. **P 9** An zwei Standorten A und B eines Unternehmens werden Aggregate gleicher Sorte hergestellt. Zur Untersuchung der Qualität der Produktion wurden Lebensdaueranalysen durchgeführt. Im Standort A wurden sechs Aggregate überprüft, im Standort B acht Aggregate. Es ergaben sich folgende Werte:

Standort	Anzahl an geprüften Aggregaten	Durchschnittliche Lebensdauer pro Aggregat (in Jahren)	Standardabweichung der Lebensdauer
A	5	24,8	5,31
B	8	33,1	5,72

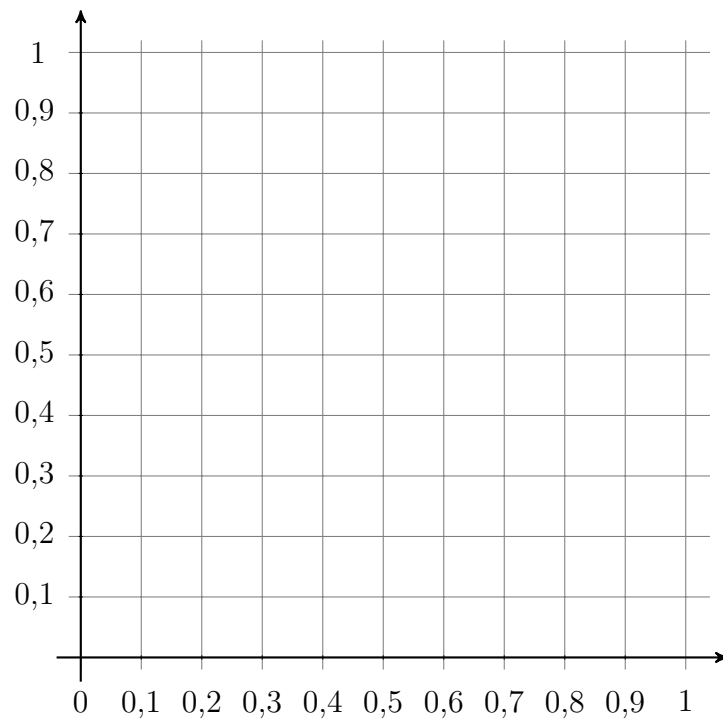
- Bestimmen Sie die Variationskoeffizienten der Lebensdauer der Aggregate für die einzelnen Standorte. An welchem Standort war die Streuung der Lebensdauer höher?
 - Bestimmen Sie unter Verwendung von a) Mittelwert und Standardabweichung der Lebensdauer für die Gesamtheit aller 13 Aggregate.
16. In einem 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** In einer Stadt gibt es zehn Bäckereien, die sich bezüglich ihres Umsatzes in drei Gruppen einteilen lassen: kleine, mittlere und große Bäckereien (wobei innerhalb einer Gruppe der Umsatz jeweils gleich ist). Im Jahr 2024 erzielten alle Bäckereien zusammen einen Gesamtumsatz von 4 Millionen Euro. 40 Prozent davon entfielen auf die zwei großen Bäckereien, während die fünf kleinen Bäckereien einen Umsatz von insgesamt 1.200.000 Euro erzielten.
- a) Bestimmen Sie die Punkte der Lorenzkurve.
 - b) Zeichnen Sie die Lorenzkurve in nachstehendes Koordinatensystem.
 - c) Die Fläche unter der Lorenzkurve aus Aufgabe a) beträgt 0,37. Berechnen Sie den Ginikoeffizienten und interpretieren Sie den erhaltenen Wert.



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 Ein Unternehmen möchte wissen, ob es einen Zusammenhang zwischen dem Kundentyp (Stamm- oder Neukunde) und der Kundenzufriedenheit (niedrig oder hoch) gibt. Zu diesem Zweck wurden 135 Kunden befragt.

30 der befragten Neukunden bewerteten ihre Zufriedenheit als niedrig. Von den 80 Stammkunden bewerteten 62 ihre Zufriedenheit als hoch.

- Welche Skalenniveaus haben hier die Variablen „Kundenzufriedenheit“ und „Kundentyp“?
- Tragen Sie die entsprechenden absoluten Häufigkeiten in die nachstehende Tabelle ein und bilden Sie alle Zeilen- und Spaltensummen.
- Geben Sie die Tabelle mit den unter Unabhängigkeit erwarteten Häufigkeiten an.
- 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}	Zufriedenheit niedrig	Zufriedenheit hoch	Σ
Stammkunde			
Neukunde			
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** An einem Tanzturnier nahmen 7 Tanzpaare teil. Zwei Wertungsrichter gaben den Paaren folgende Platzziffern

Tanzpaar Nr.	1	2	3	4	5	6	7
Platz von Wertungsrichter A	4	2	6	5	1	7	3
Platz von Wertungsrichter B	4	1	5	7	3	6	2

- a) Berechnen Sie eine geeignete Kennzahl, um den Zusammenhang zwischen den Wertungen der beiden Wertungsrichter zu beschreiben.
- b) Interpretieren Sie die Kennzahl im Kontext der Untersuchung. Gibt es Hinweise auf einen Zusammenhang zwischen der Wertung des Wertungsrichters A und der des Wertungsrichters B?
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 Tausend €)	24	16	19	26	12	18
Umsatz (in Tausend €)	195	190	220	250	150	185

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

23. **P 13** Die in folgender Tabelle angegebenen Daten zeigen die Entwicklung der Arbeitslosenzahlen und der offenen Stellen der Jahre 2016 bis 2024 in Österreich (jeweils als Jahresdurchschnitt zusammengestellt):

Jahr	2016	2017	2018	2019	2020	2021	2022	2023	2024
offene Stellen (in Tsd.)	49	51	57	65	66	72	80	95	91
Arbeitslose (in Tsd.)	357	285	260	240	315	320	221	241	373

- Erstellen Sie ein Streudiagramm des Datensatzes.
- Bestimmen Sie die Gleichung der Regressionsgeraden zur Beschreibung der Abhängigkeit der Arbeitslosenzahl von der Anzahl an offenen Stellen.
- Beurteilen Sie mithilfe des Korrelationskoeffizienten, ob die Regressionsgerade ein geeignetes Modell darstellt, um diesen Zusammenhang zu beschreiben.
- Welche Arbeitslosenzahl prognostiziert das Modell für das Jahr 2025, in dem 78.000 offene Stellen erwartet werden?

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