

Übungsblatt 1

Beschreibende Statistik

Stochastik@AIN2

Prof. Dr. Barbara Staehle

Wintersemester 2025/2026

HTWG Konstanz

Einfache und mittelschwere Aufgaben

AUFGABE 1.1 EIGENSCHAFTEN VON MERKMALEN, 3 PUNKTE

+--- Qualitativ ---

| Ordinal = anordbar (Ränge)

| nominal = nicht anordbar (Gruppen)

+--- Quantitativ ---

| diskret = endlich viele oder abzählbar unendlich viele (1,2,3..)

| stetig = überabzählbar unendlich viele (1.1, 2.98, 1/5, etc.)

Geben Sie für jedes Merkmal an, ob es quantitativ diskret, quantitativ stetig, qualitativ ordinal oder qualitativ nominal ist.

- a) Anzahl der Insassen in einem PKW bei der Verkehrszählung **Quantitativ diskret**
- b) Höchstgeschwindigkeit eines PKWs **Quantitativ stetig**
- c) Schultypen **Qualitativ Ordinal**
- d) Temperaturangaben in °C **Quantitativ stetig**
- e) Fassungsvermögen einer Festplatte in MB **Quantitativ diskret**
- f) Zugriffszeiten auf die Daten einer Festplatte in sec **Quantitativ stetig**
- g) Energieeffizienz von Elektrogeräten **Qualitativ Ordinal**
- h) Trinkmenge eines Menschen pro Tag in Liter **Quantitativ stetig**
- i) Anzahl Likes eines Instagram-Posts **Quantitativ diskret**
- j) Nationalität eines Menschen **Qualitativ nominal**
- k) Anfangsbuchstabe des Vornamens eines Menschen **Qualitativ nominal**
- l) Wohnort eines Menschen **Qualitativ nominal**

AUFGABE 1.2 TIKTOK-VIDEOS

Sie interessieren sich für die Frage, wie ein typisches TikTok-Video aussieht und führen eine statistische Stichprobenerhebung durch.

TEILAUFGABE 1.2.1 3 PUNKTE

Geben Sie **Beispiele** an, was für diese Untersuchung die folgenden Kenngrößen sein könnten:


- Statistische Einheiten [Kurz Video \(TikTok Video\)](#)
- Grundgesamtheit [Alle Videos der Plattform](#)
- Stichprobe [10000 Videos](#)
- Merkmale (mindestens 3) [Likes, Kommentare, Aufrufe](#)
[Likes -> 10.000](#)
[Kommentare -> 3.000](#)
[Aufrufe -> 70.000](#)
- **Beispielhafte** Merkmalsausprägungen (jeweils für Ihre Beispiele)
- Eigenschaften der Merkmale (qualitativ ordinal/nominal, quantitativ diskret/kontinuierlich) (jeweils für Ihre Beispiele)
[Likes = Quantitativ diskret](#)
[Kommentare = Quantitativ diskret](#)
[Aufrufe = Quantitativ diskret](#)

TEILAUFGABE 1.2.2 1 PUNKT

Welche Eigenschaften von TikTok-Videos finden Sie besonders spannend und einer näheren Untersuchung wert? Begründen Sie Ihre Meinung. [Besonders die Watchtime \(also wie lange ein benutzer das Video geschaut hat\) und das engagement \(Likes, Kommentare etc.\) sind spannend, da diese den Algorithmus beeinflussen, welche Videos anderen Usern vorgeschlagen werden. Diese Eigenschaften bestimmen auch, welche Videos Viral gehen und welche unbekannt bleiben.](#)

AUFGABE 1.3 GELDSCHEINE, 2 PUNKTE

Angenommen, es wären von allen vorhandenen Werten (5 €, 10 €, 20 €, 50 €, 100 €, 200 €, 500 €) gleich viele Geldscheine im Umlauf.

- a) Bestimmen Sie den Median des Werts der Geldscheine im Umlauf. $\bar{x} = 50€$ 
- b) Berechnen Sie den Median, der sich nach der Einführung eines 1000 €-Scheins (und davon ebenfalls so viele in den Umlauf gebracht wie alle anderen) ergeben würde.

Werte = [5€, 10€, 20€, 50€, 100€, 200€, 500€, 1000€]



$$\bar{x} = \frac{50 + 100}{2} = 75€$$

AUFGABE 1.4 AN DER FAHRRADBRÜCKE (KLAUSURAUFGABE WS 16/17), 4 PUNKTE

Alice steht an der Konstanzer Fahrradbrücke und zählt die Anzahl der Räder der Fahrzeuge (Einräder, Fahrräder, Dreiräder, Kinderwagen, ...) welche an ihr vorbei rollen.

Ihre Beobachtungen resultieren in folgender Urliste:

2, 4, 3, 1, 2, 4, 2, 2, 2, 3.

$1 = 1$
 $2 = 5$
 $3 = 2$
 $4 = 2$

Berechnen Sie für die Beobachtungen von Alice: $[1, 2, 2, 2, 2, 2, 3, 3, 4, 4]$
 $n = 10$

- a) das arithmetische Mittel $= (1+10+6+8)/10 = 2,5$
- b) den Median $= (2+2)/2 = 2$
- c) den Modalwert $mw = 2$
- d) das 10%-Quantil $Position = 10 * 0,1 = 1 \quad 1$
- e) das 25%-Quantil $= 2$
- f) das 75%-Quantil $= 3$
- g) den Interquartilabstand $iqr = Q3 - Q1 = 3 - 2 = 1$
- h) die Spannweite $= \text{Maximalerwert} - \text{minimalerwert} = 4 - 1 = 3$

Geben Sie jeweils an, wie Sie die geforderten Charakteristika berechnet haben!

AUFGABE 1.5 TASCHENGELD (KLAUSURAUFGABE SS 18)

Frank schreibt für seine Schülerzeitung einen Artikel zum Thema „Taschengeld eines Achtklässlers“. Hierfür befragt er 6 zufällig ausgewählte Schüler der Klasse 8a nach ihrem monatlichen Taschengeld in Euro. Seine Umfrage ergibt die folgende Urliste:

25, 10, 30, 25, 35, 25

TEILAUFGABE 1.5.1 2 PUNKTE

Berechnen Sie für Frank's Urliste folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar!

- a) Arithmetisches Mittel
- b) Median
- c) 75%-Quantil
- d) empirische Standardabweichung

TEILAUFGABE 1.5.2 3 PUNKTE

Beantworten Sie folgende Fragen möglichst allgemeinverständlich!

- a) Frank berechnet außerdem noch die Spannweite seiner Stichprobe zu $R = 5$ und den Interquartilabstand zu $I = 25$. Machen diese Berechnungen Sinn? Begründen Sie Ihre Meinung!
- b) Mittelwert und Median von Franks Daten sind gleich. Ist das immer so, oder ist das Zufall?
- c) Was hat der Wert des 75%-Quantils zu bedeuten? Fassen Sie Ihr Ergebnis in Worte.

- d) Frank führt seine Umfragen nach dem monatlichen Taschengeld noch in zwei anderen Klassen, den Klassen 8b und 8c, durch. In beiden Fällen erhält er das selbe arithmetische Mittel wie bei Klasse 8a. Im Fall von Klasse 8b erhält er allerdings eine Standardabweichung von 0, im Fall von Klasse 8c erhält er eine Standardabweichung, die deutlich größer ist als die der Klasse 8a. Was schließen Sie hieraus hinsichtlich der Stichproben des monatlichen Taschengelds aus den anderen beiden Klassen?
- e) Frank interessiert sich weiterhin für die Schuhgröße der Befragungsteilnehmer, sowie für ihre Mathenote. Er berechnet hierfür die Korrelationskoeffizienten von Schuhgröße und Taschengeld zu $r_{S,T} = 0.875$ und von Mathenote und Taschengeld zu $r_{M,T} = -0.038$. Was schließen Sie hieraus über den Zusammenhang der betrachteten Größen?
- f) Grace befragt 3 Schüler nach ihrem Taschengeld und berechnet den arithmetische Mittelwert dieser Stichprobe zu 300 €. Beschreibt diese Zahl die Stichprobe ausreichend gut, oder sollte sie besser noch andere Werte berechnen?

AUFGABE 1.6 SPORTWAGEN (KLAUSURAUFGABE WS 17/18)

Charlie besucht die IAA (Internationale Automobilausstellung). Er interessiert sich für Sportwagen und sucht daher Informationen zu den Höchstgeschwindigkeiten der vorgestellten Autos.

Am Stand von Hersteller 1 analysiert er die Datenblätter von 10 verschiedenen Fahrzeugen, denen er folgende Höchstgeschwindigkeiten entnimmt:

250, 330, 250, 350, 270, 300, 260, 330, 200, 400

TEILAUFGABE 1.6.1 2 PUNKTE

Berechnen Sie für die Stichprobe der von Charlie ermittelten Höchstgeschwindigkeiten folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar!

- a) Arithmetisches Mittel
- b) Median
- c) 90%-Quantil
- d) Standardabweichung

TEILAUFGABE 1.6.2 3 PUNKTE

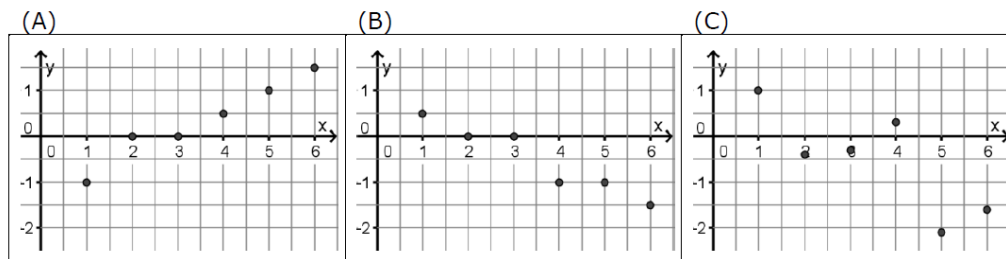
Gefragt sind Beschreibungen in Ihren eigenen Worten (Zielgruppe: nicht-technisches Publikum).

- a) Nennen Sie **zwei** Fakten, wie die von Ihnen berechneten Größen einzeln oder gemeinsam betrachtet die von Charlie erhobene Stichprobe der Höchstgeschwindigkeiten charakterisieren.
- b) Charlie entnimmt aus einem Datenblatt von Hersteller 2 ebenfalls die Höchstgeschwindigkeiten von 10 verschiedenen Fahrzeugen und berechnet den Mittelwert dieser Höchstgeschwindigkeiten zu 294 sowie die Standardabweichung zu 20. Was schließen Sie hieraus hinsichtlich des Unterschieds zwischen den Fahrzeugen der beiden Hersteller?
- c) Charlie betrachtet weiterhin die Preise der Reifen welche für die Fahrzeuge von Hersteller 1 zugelassen sind, sowie die Umweltfreundlichkeit des Fahrzeugs auf einer Skala von 0 (gar nicht) bis 10 (sehr).
Er berechnet hierfür die Korrelationskoeffizienten von Höchstgeschwindigkeit und Reifenpreis zu $r_{G,R} = 0.853$ und von Höchstgeschwindigkeit und Umweltfreundlichkeit zu $r_{G,U} = -1.137$. Was schließen Sie hieraus über den Zusammenhang der betrachteten Größen?

Mittelschwere und schwere Aufgaben

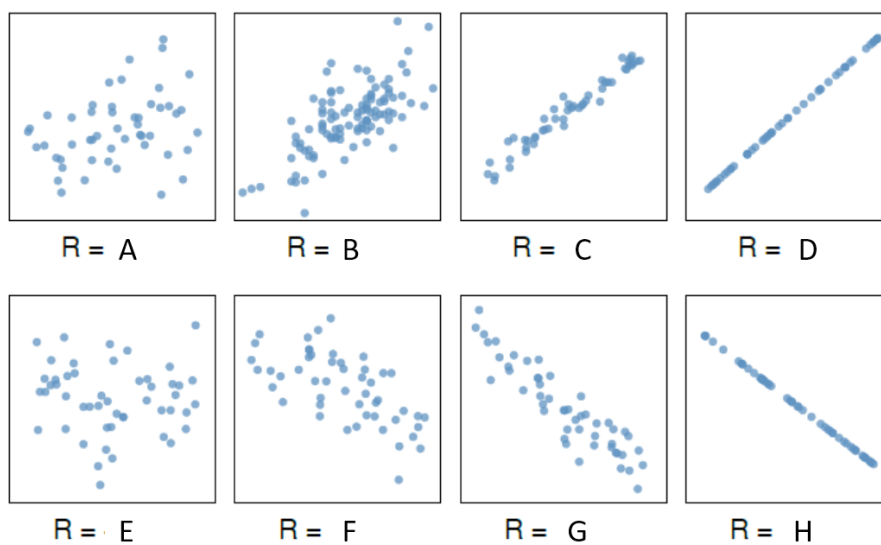
AUFGABE 1.7 PUNKTWOLKEN

TEILAUFGABE 1.7.1 3 PUNKTE



Bei welcher der dargestellten Punktwolken (A), (B) oder (C) erhält man als Korrelationskoeffizienten $r_{x,y} = -0,966$? Begründen Sie Ihre Meinung und nennen Sie für die beiden anderen Punktwolken jeweils einen Ablehnungsgrund.

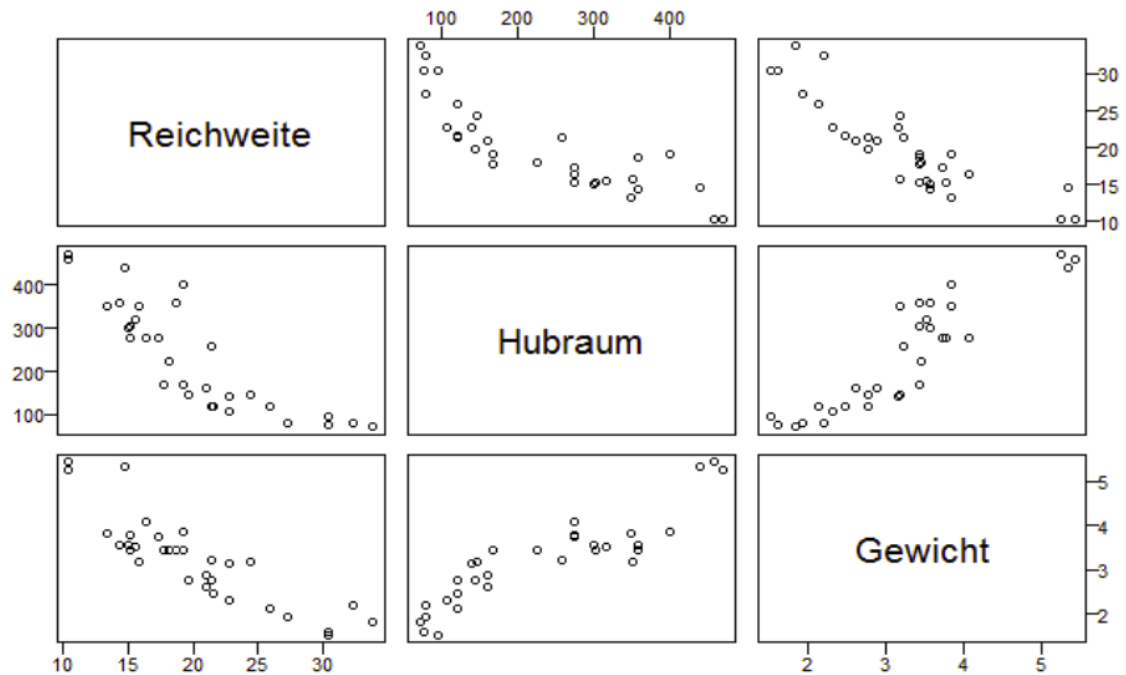
TEILAUFGABE 1.7.2 3 PUNKTE



Wählen Sie für jede der dargestellten Punktwolken aus den folgenden Werten denjenigen aus, der am besten für den Korrelationskoeffizienten passt. Achtung: Nicht alle Werte passen zu einem Bild.

−0.92 0.00 0.69 −0.08 −0.64 −0.20 0.33 1.00 −1.00 0.98

TEILAUFGABE 1.7.3 3 PUNKTE



Die abgebildete Matrix stellt Punktwolken für einige Eigenschaften einer Auswahl verschiedener Autotypen des gleichen Herstellers dar. Der Plot oben rechts z.B. zeigt auf der x-Achse den Faktor Gewicht, auf der y-Achse den Faktor Reichweite. Ein Punkt entspricht einem Auto, geeignete Einheiten fehlen, diese dürfen Sie sich selbst ausdenken.

Welche Aussagen über die Zusammenhänge zwischen diesen drei Faktoren lassen sich aus der Grafik ableiten?

AUFGABE 1.8 KLAUSURNOTEN (KLAUSURAUFGABE SS19)

Prof. Schmidt liest für alle Studierenden des Studiengangs „Bachelor Super-Informatik“ die Vorlesungen „Mathematik 1“ und „Mathematik 2“.

Um die Noten der beiden Vorlesungen vergleichen, analysiert sie die Noten, welche 6 zufällig ausgewählte Studierende (s_1, s_2, \dots, s_6) in den beiden Klausuren jeweils erreichten. Sie erhält die folgende Tabelle:

Studierende(r)	s_1	s_2	s_3	s_4	s_5	s_6
Note Mathematik 1	1.3	5.0	1.3	2.7	4.0	3.7
Note Mathematik 2	2.0	3.0	1.0	2.0	1.0	2.7

TEILAUFGABE 1.8.1 2 PUNKTE

Berechnen Sie **nur für die Noten der Vorlesung Mathematik 1** folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar, bzw. geben Sie die Verwendung des Taschenrechners oder eines anderen elektronischen Tools an!

- a) Arithmetisches Mittel
- b) Alle Quartile
- c) 90%-Quantil
- d) empirische Standardabweichung

TEILAUFGABE 1.8.2 3 PUNKTE

Beantworten Sie folgende Fragen in ganzen Sätzen und möglichst allgemeinverständlich!

- a) Was bedeutet der Wert des 90%-Quantils (Ihr Ergebnis aus Aufgabe 1.10.2c) für die Noten von Mathematik 1?
- b) Prof. Schmidt berechnet für die Noten der Vorlesung Mathematik 2 folgende Größen:

$$\bar{x} = 2.0, \quad (Q1, Q2, Q3) = (1.0, 2.0, 2.7), \quad \tilde{x}_{0.9} = 3.0, \quad s = 0.8462$$

Verwenden Sie diese Zahlen, sowie Ihre eigenen Berechnungen aus Aufgabe 1.10.2 um **zwei** Gemeinsamkeiten oder Unterschiede für die Noten in Mathematik 1 und 2 abzuleiten. Begründen Sie Ihre Meinung!

- c) Welcher der folgenden Werte ist der korrekte empirische Korrelationskoeffizienten für die Noten in Mathematik 1 und Mathematik 2 $r_{1,2}$ (für die obige Stichprobe)? Begründen Sie Ihre Meinung.

$$r_{1,2} = \quad 0 \quad -0.571 \quad 0.423 \quad 1.1 \quad 0.952$$

AUFGABE 1.9 JOBSUCHE (KLAUSURAUFGABE WS20/21)

Sie suchen nach Ihrem Studium einen neuen Job, der bestimmte Anforderungen erfüllen soll.

- Der neue Job soll in einer attraktiven Stadt sein: Sie bewerten die Stadt in dieser Hinsicht mit Punkten von 1 (furchtbar) bis 10 (total toll).
- Natürlich ist Ihnen auch das Gehalt wichtig: Sie notieren das (geschätzte) Brutto-Gehalt in Tausend Euro ($55 \triangleq 55\,000$).
- Und die Tätigkeit soll interessant sein: Auch hier vergeben Sie Punkte von 1 (langweilig) bis 10 (genau Ihr Ding).

Nach zwei Wochen Suche haben Sie sich auf folgende 8 Jobs eingeschränkt und diese wie folgt bewertet:

Job	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8
Attraktivität der Stadt	7	4	10	10	9	3	9	6
Gehalt	65	60	40	70	45	40	65	55
interessant	3	10	9	1	5	10	2	3

TEILAUFGABE 1.9.1 3 PUNKTE

Berechnen Sie **nur für Ihre Punktevergabe der Attraktivität der Städte** (Zeile 1 der Tabelle) folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar, bzw. geben Sie die Verwendung des Taschenrechners (TR), bzw. des genutzten Tools an!:

- Modalwert
- Arithmetisches Mittel
- empirische Standardabweichung
- Alle Quartile
- 60%-Quantil

TEILAUFGABE 1.9.2 6 PUNKTE

- Für welchen der 8 Jobs würden Sie sich auf Basis dieser Daten entscheiden? Begründen Sie Ihre Meinung (es gibt keine richtige oder falsche Antwort).
- Ihre Freunde Alice und Bob haben sich jeweils auch 8 Jobs ausgesucht und die Attraktivität der Städte ebenso mit 1-10 bewertet. Alice berechnet den Mittelwert der Städte als $\overline{a_B} = 7.25$ und die entsprechende Standardabweichung zu $s_{a_B} = 5$. Bob berechnet den Mittelwert der Städte als $\overline{a_C} = 5$ und die entsprechende Standardabweichung zu $s_{a_C} = 0$. Was können Sie aus diesen Werten über die Attraktivität der Städte von Alice und Bob im Vergleich zu „ihren“ Städten sagen?
- Schätzen** Sie die empirischen Korrelationskoeffizienten $r_{A,G}$, $r_{A,I}$, $r_{G,I}$ zwischen „Attraktivität der Stadt“ und „Gehalt“, „Attraktivität der Stadt“ und „interessant“ sowie „Gehalt“ und „interessant“ für die ausgewählten Jobs.
Rechnen Sie diese nicht aus, sondern wählen Sie unter den folgenden Werten. Begründen Sie Ihre Meinung!

$r_{A,G}$	a) 0.1983	b) 1.983	c) -0.9308	d) 0.0093
$r_{A,I}$	a) 5.599	b) -0.5599	c) 0.5599	d) 0.0055
$r_{G,I}$	a) 0.7032	b) 0.0073	c) -0.7032	d) -7.0320

AUFGABE 1.10 AUTOTESTS (KLAUSURAUFGABE WS19/20)

Alice liest auf der Webseite eines Automagazins einen Testbericht, der 5 verschiedene Autos (a_1, a_2, \dots, a_5) hinsichtlich ihres Spritverbrauchs vergleicht. Konkret wurde überprüft, ob der Benzinverbrauch in Liter pro 100 km Fahrleistung der Herstellerangabe entspricht:

Auto	a_1	a_2	a_3	a_4	a_5
Verbrauch [l/100km] laut Hersteller	4.2	3.9	4.3	4.1	4.1
Verbrauch [l/100km] laut Test	4.8	4.1	4.8	4.5	4.4

TEILAUFGABE 1.10.1 1 PUNKT

Geben Sie an, welche der Eigenschaften **qualitativ**, **quantitativ**, **diskret**, **stetig** auf das ebenfalls untersuchte Merkmal „Anzahl der Räder des Fahrzeugs“ zutrifft bzw. zutreffen. Begründen Sie Ihre Wahl (kurz).

TEILAUFGABE 1.10.2 3 PUNKTE

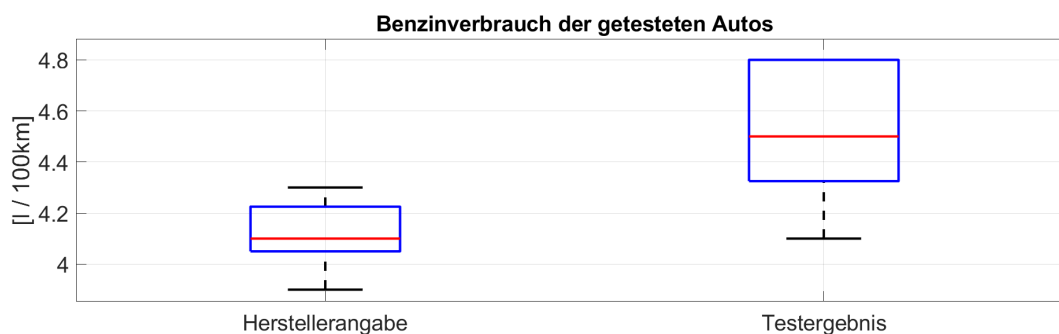
Berechnen Sie **nur für die Verbrauchsangaben des Herstellers** folgende Größen.

- a) Arithmetisches Mittel b) Alle Quartile c) 10%-Quantil
- d) Spannweite e) empirische Standardabweichung

TEILAUFGABE 1.10.3 5 PUNKTE

Beantworten Sie folgende Fragen in ganzen Sätzen und möglichst allgemeinverständlich!

- a) Was bedeutet der Wert der empirischen Standardabweichung (Ihr Ergebnis aus Aufgabe 1.10.2??) für den Benzinverbrauch eines Autos (laut Hersteller)?
- b) Alice interessiert sich dafür, inwieweit die Verbrauchsangaben der Hersteller und der im Test erzielte Benzinverbrauch vergleichbar sind. Daher erstellt Sie folgende Grafik:



Beschreiben Sie anhand dieser Grafik (und ggf. unter Verwendung Ihrer eigenen Berechnungen aus Aufgabe 1.10.2) **zwei** Gemeinsamkeiten oder Unterschiede für den Benzinverbrauch der untersuchten Autos laut Hersteller und laut Test. Können diese Sachverhalte Sinn machen? Begründen Sie Ihre Meinung!

- c) Welcher der folgenden Werte ist der korrekte empirische Korrelationskoeffizienten für den Benzinverbrauch eines Autos laut Hersteller und laut Test $r_{H,T}$ (für die vorliegende Stichprobe)? Kreisen Sie den richtigen Wert bzw. streichen Sie falsche Werte durch. Begründen Sie Ihre Meinung.

$$r_{H,T} = \quad 0.0001 \quad 0.961 \quad 0.235 \quad 1.1 \quad -0.521$$

- d) Um abschätzen zu können, ob man aus der Herstellerangabe des Benzinverbrauchs das Testergebnis ausrechnen kann, berechnet Alice die Regressionsgerade für die beiden Datensätze und erhält hierfür das Bestimmtheitsmaß $R^2 = 0.92$. Was bedeutet dieses? Ist die Anwendung der linearen Regression hier sinnvoll und zulässig? Begründen Sie Ihre Meinung.

AUFGABE 1.11 RESPOSITORIES (KLAUSURAUFGABE SS21)

Alice arbeitet für den Webhoster Host123 und ist dafür zuständig, den automatischen Build-Prozess für Kunden-Webapplikationen zu optimieren. Hierfür analysiert sie in 7 zufällig ausgewählten Kundenrepositories die Anzahl der in den am häufigsten genutzten Programmiersprachen Java, Python, Ruby gespeicherten Dateien.

Ihre Ergebnisse fasst sie in der folgenden Tabelle zusammen:

Repository	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Java	25	17	25	29	20	15	11
Python	11	11	13	9	10	7	3
Ruby	0	4	0	0	0	2	2

TEILAUFGABE 1.11.1 8 PUNKTE

Berechnen Sie **für die Anzahl der Java-Dateien** (Zeile 1 der Tabelle) die im Folgenden gesuchten Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar, bzw. geben Sie die Verwendung des Taschenrechners (TR), bzw. des genutzten Tools an!

- a) Modalwert
- b) Arithmetisches Mittel
- c) Median
- d) 90%-Quantil
- e) empirische Standardabweichung
- f) Alle Quartile
- g) Interquartilsabstand
- h) Spannweite (Range)

TEILAUFGABE 1.11.2 9 PUNKTE

- a) (2 Punkte) Formulieren Sie mit Hilfe der Kenngrößen, die Sie in Aufgabe 1.11.1 ausgerechnet haben, **zwei** Aussagen über die Anzahl der Java-Dateien in den untersuchten Repositories. Schreiben Sie also zwei Sätze auf, die Sie mit Hilfe der Zahlen begründen können. Begründen Sie Ihre Meinung (es gibt keine richtige oder falsche Antwort).
- b) (2 Punkte) Finden Sie jeweils einen Unterschied, bzw. eine Gemeinsamkeit zwischen der Anzahl der Python- und der Python-Dateien und der Java und der Ruby- Dateien in den Repositories, indem Sie Ihre Berechnungen aus Aufgabe 1.11.1 mit den folgende Kennwerten für Java und Ruby vergleichen. Schreiben Sie hierzu zwei Sätze auf.

Sprache	Mittelwert	Median	Std	IQR	Range
Python	9.1	10.0	3.288	3.0	10.0
Ruby	1.1	0.0	1.574	2.0	4.0

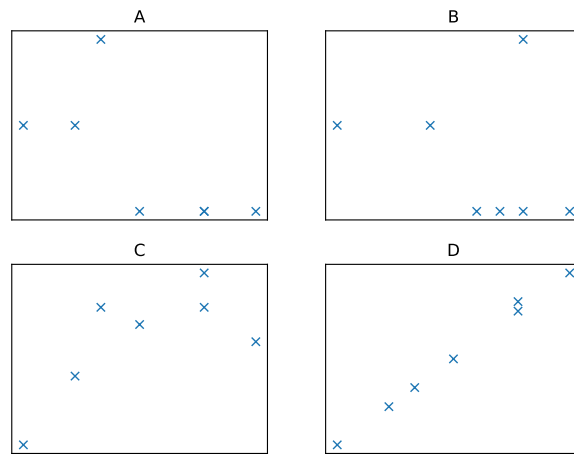


Abbildung 1: Welcher Scatterplot entspricht welcher Sprachkombination?

- b) (3 Punkte) Die Scatter-Plots in Abbildung 1 illustrieren den Zusammenhang zwischen der Anzahl der Dateien in

(1) Java und Python	(2) Java und Ruby	(3) Python und Ruby
---------------------	-------------------	---------------------

Geben Sie an, welches der Bilder (A,B,C,D) welchen Zusammenhang darstellt. Begründen Sie Ihre Meinung.

- c) (3 Punkte) **Schätzen** Sie die empirischen Korrelationskoeffizienten $r_{J,P}$, $r_{J,R}$, $r_{P,R}$ zwischen der Anzahl der Dateien in Java und Python, Java und Ruby, sowie Python und Ruby.

Rechnen Sie diese nicht aus, sondern wählen Sie unter den folgenden Werten. Interpretieren Sie den Wert (was sagt dieser über den Zusammenhang der beiden Metriken aus?) Begründen Sie Ihre Meinung.

$r_{J,P}$	a) -0.7	b) 0.998	c) 0.695	d) -0.295
$r_{J,R}$	a) -0.7	b) 0.998	c) 0.695	d) -0.295
$r_{P,R}$	a) -0.7	b) 0.998	c) 0.695	d) -0.295

Digitalaufgaben

Lösen Sie alle Aufgaben in diesem Abschnitt digital mit einem Werkzeug Ihrer Wahl (MATLAB, Python, Excel, ...) und laden Sie den verwendeten Code bzw. Ihre Bilder und Ergebnisse nach Moodle hoch.

AUFGABE 1.12 AUSBILDUNG UND GEHÄLTER, 3 PUNKTE

Fünf Mitglieder eines Sportclubs gehen der Frage nach, ob zwischen ihrer Schul- und Ausbildungszeit und ihrem Jahreseinkommen ein statistischer Zusammenhang besteht. Das Ergebnis der anonymen Befragung ist in folgender Tabelle wiedergegeben.

Ausbildungsdauer X	9	13	15	18	20
Jahresgehalt Y [1000 €]	18	37	61	125	59

- Berechnen Sie den empirischen Korrelationskoeffizient der beiden untersuchten Merkmale. Interpretieren Sie diesen Wert!
- Bestimmen Sie eine lineare Regressionsfunktion und zeichnen Sie diese. Bewerten Sie die Passgenauigkeit dieser Regressionsgerade optisch und durch das Bestimmtheitsmaß.

AUFGABE 1.13 FERNSEHEN UND AGRESSIONEN, 3 PUNKTE

In einer Studie zur Auswirkung von YouTube-Filmen mit gewalttätigen Szenen auf das Sozialverhalten von Kindern wurden ein Aggressivitätsmaß Y (auf einer Skala von 0-10), die Zeitdauer X (in Minuten), während der das Kind gewöhnlich solche Filme sieht, und das Geschlecht Z des Kindes (0 männlich, 1 weiblich) erfasst.

Die Ergebnisse in einer zufällig gewählten Testgruppe waren wie folgt:

i	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	10	50	30	70	80	60	90	40	10	20	30	50	60
y_i	4	5	2	6	6	8	7	2	7	3	5	1	3
z_i	0	0	0	0	0	0	0	1	1	1	1	1	1

Verwenden Sie diese Testergebnisse, für eine (empirisch begründete) Aussage zum Thema "macht YouTube aggressiv"?

Empfohlene Vorgehensweise:

- Zeichnen Sie ein Streudiagramm für die 13 Kinder, und berechnen Sie den Korrelationskoeffizienten zwischen Dauer des Medienkonsums und Aggressionslevel ohne Berücksichtigung des Geschlechts.
- Zeichnen Sie nun für Jungen und Mädchen getrennt jeweils ein Streudiagramm und berechnen Sie für beide Geschlechter den Korrelationskoeffizienten.
- Vergleichen Sie alle Ergebnisse und ziehen Sie Ihre Folgerungen.

AUFGABE 1.14 3 KEY FACTS, 6 PUNKTE

Suchen Sie sich einen der in Moodle verlinkten Datensätze, oder einen beliebigen Datensatz Ihrer Wahl aus und finden Sie 3 Key Facts (also Sachverhalte, die Sie für spannend halten). Verwenden Sie hierfür Methoden, und Kenngrößen, die Sie in der Vorlesung kennengelernt haben und MATLAB, Python, Excel, Java, Stellen Sie jeden Sachverhalt mit einem Bild, einer Zahl, einer Tabelle, ... dar. Beschreiben Sie die Aussage, die Sie daraus ableiten, sowie wie Sie das Bild, die Zahl, ... erhalten haben, kurz. Bündeln Sie alles (Bild, Code, Zahl, Erklärung) in einer Datei. Keine Punkte ohne (kurze) Vorstellung in der Übung.

Hinweise:

- Teamgröße: 3-5 (ein Team darf identische Lösungen abgeben, Namen der Teammitglieder bitte in der Lösung angeben)
- Bewertung: 1-2 Punkte pro gut erklärtem Fakt guten Code und gute Präsentation
- Geben Sie in Ihrer Lösung an, aus welcher Quelle Sie Ihre Daten haben.
- Beispielhafte Lösungen: siehe Moodle