

Name: _____

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Studierende der Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Probeklausur Bio Data Science

für Pflichtmodule

im 1. & 2. Semester B.Sc./M.Sc.

Prüfer: Prof. Dr. Jochen Kruppa-Scheetz
Fakultät für Agrarwissenschaften und Landschaftsarchitektur
j.kruppa@hs-osnabrueck.de

12. Mai 2025

Erlaubte Hilfsmittel für die Klausur

- Normaler Taschenrechner ohne Möglichkeit der Kommunikation mit anderen Geräten - also ausdrücklich kein Handy!
- Eine DIN A4-Seite als beidseitig, selbstgeschriebene, handschriftliche Formelsammlung - keine digitalen Ausdrucke.
- **You can answer the questions in English without any consequences.**

Ergebnis der Klausur

_____ von 20 Punkten sind aus dem Multiple Choice Teil erreicht.

_____ von 69 Punkten sind aus dem Rechen- und Textteil erreicht.

_____ von 89 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
85.0 - 89.0	1,0
80.5 - 84.5	1,3
76.5 - 80.0	1,7
72.0 - 76.0	2,0
67.5 - 71.5	2,3
63.0 - 67.0	2,7
58.5 - 62.5	3,0
54.5 - 58.0	3,3
50.0 - 54.0	3,7
44.5 - 49.5	4,0

Es ergibt sich eine Endnote von _____.

Multiple Choice Aufgaben

- Pro Multiple Choice Frage ist *genau* eine Antwort richtig.
- **Übertragen Sie Ihre Kreuze in die Tabelle auf dieser Seite.**
- Es werden nur Antworten berücksichtigt, die in dieser Tabelle angekreuzt sind!

	A	B	C	D	E	✓
1 Aufgabe						
2 Aufgabe						
3 Aufgabe						
4 Aufgabe						
5 Aufgabe						
6 Aufgabe						
7 Aufgabe						
8 Aufgabe						
9 Aufgabe						
10 Aufgabe						

- Es sind ____ von 20 Punkten erreicht worden.

Rechen- und Textaufgaben

- Die Tabelle wird vom Dozenten ausgefüllt.

Aufgabe	11	12	13	14	15	16	17
Punkte	9	9	8	12	11	9	11

- Es sind ____ von 69 Punkten erreicht worden.

1 Aufgabe

(2 Punkte)

Welche Aussage zum mathematische Ausdruck $Pr(D|H_0)$ ist richtig?

- A ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativhypothese und somit $1 - Pr(H_A)$
- B ☐ $Pr(D|H_0)$ stellt die Wahrscheinlichkeit die Teststatistik T zu beobachten dar, wenn die Nullhypothese falsch ist.
- C ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit nicht die Daten D zu beobachten sondern die Nullhypothese, wenn diese wahr ist.
- D ☐ Die Inverse der Wahrscheinlichkeit unter der die Nullhypothese nicht mehr die Alternativhypothese überdeckt.
- E ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit die Daten D zu beobachten, wenn die Nullhypothese wahr ist.

2 Aufgabe

(2 Punkte)

Die einfaktorielle ANOVA ist ein Standardverfahren in der agrawissenschaftlichen Forschung wenn es um den Vergleich von Behandlungsgruppen geht. Welche der folgenden Aussage zu der Berechnung der Teststatistik der einfaktoriellen ANOVA ist richtig?

- A ☐ Die ANOVA berechnet die T-Statistik indem den Mittelwertsunterschied der Gruppen simultan durch die Standardabweichung der Gruppen teilt. Wenn die T-Statistik höher als 1.96 ist, kann die Nullhypothese abgelehnt werden.
- B ☐ Die F-Statistik wird berechnet indem die MS der Behandlung durch die MS des Fehlers geteilt werden. Wenn die F-Statistik sich kaum von der Null unterscheidet kann die Nullhypothese nicht abgelehnt werden.
- C ☐ Die ANOVA berechnet die F-Statistik aus den SS Behandlung geteilt durch die SS Fehler.
- D ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 0 annähert kann die Nullhypothese abgelehnt werden.
- E ☐ Die ANOVA berechnet die T-Statistik aus der Multiplikation der MS Behandlung mit der MS der Fehler. Wenn die F-Statistik genau 0 ist, kann die Nullhypothese nicht abgelehnt werden.

3 Aufgabe

(2 Punkte)

Aus einem Feldversuch ergibt sich die Notwendigkeit der Berechnung einer einfaktoriellen ANOVA. Es ergibt sich ein $\eta^2 = 0.12$. Welche Aussage ist richtig?

- A ☐ Das η^2 beschreibt den Anteil der globalen Varianz, der von den Umweltbedingungen erklärt wird.
- B ☐ Das η^2 ist die Korrelation der ANOVA. Mit der Ausnahme, dass $\eta^2 = 0$ der beste Wert ist.
- C ☐ Die Berechnung von η^2 ist ein Wert für die Interaktion in der einfaktoriellen ANOVA.
- D ☐ Der Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird, wird durch das η^2 beschrieben.
- E ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen nicht erklärt wird. Somit der Rest an nicht erklärbarer Varianz.

4 Aufgabe

(2 Punkte)

Die Abkürzung CLD steht für welches statistische Verfahren? Welche folgende Beschreibung der Interpretation ist korrekt?

- A ☐ Compact letter display. Das CLD ist umstritten, da es die Gleichheit der Behandlungen durch gleiche Buchstaben darstellt. Dadurch ist das CLD nicht mehr sauber auf einer Linie mit dem statistischen Testen. Wir lehnen die Nullhypothese ab und zeigen keine Gleichheit im statistischen Testen.

- B** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.
- C** ☐ Compound letter display. Gleichheit in dem Outcomes wird durch den gleichen Buchstaben oder Symbol dargestellt. Teilweise ist die Interpretation des Verbunds (eng. compound) herausfordernd, da wir ja nach dem Unterschied suchen.
- D** ☐ Compact letter display. Gleiche Buchstaben bedeuten, dass sich die Behandlungen unterscheiden. Daher ist das CLD sehr unintuitiv. Es wäre besser, wenn gleiche Buchstaben Gleichheit anzeigen würden. Dies ist aber leider in der statistischen Testtheorie nicht möglich.
- E** ☐ Compact letter display. Gleiche Buchstaben zeigen Gleichheit in den Behandlungen. Die Interpretation ist deshalb sehr intuitiv und einfach. Darüber hinaus ist damit das CLD auch auf einer Linie mit der Testtheorie, da wir ja auch dort die Gültigkeit der Nullhypothese nachweisen. Wir suchen ja Gleichheit.

5 Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit rechnen Sie einen Student t-Test. Welche Aussage ist auch für den Welch t-Test richtig?

- A** ☐ Der t-Test vergleicht zwei oder mehr Gruppen indem die Mittelwerte miteinander verglichen werden.
- B** ☐ Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- C** ☐ Der t-Test vergleicht die Mittelwerte von zwei Gruppen unter der strikten Annahme von Varianzhomogenität. Sollte keine Varianzhomogenität vorliegen, so gibt es keine Möglichkeit den t-Test in einer Variante anzuwenden.
- D** ☐ Der t-Test berechnet die Differenz von zwei Mittelwerten als Effekt und gibt eine Entscheidung, ob sich die beiden Mittelwerte *jeweils* von Null unterscheiden.
- E** ☐ Der t-Test vergleicht zwei Gruppen indem die Mittelwerte miteinander verglichen werden.

6 Aufgabe

(2 Punkte)

Die Standardabweichung ist eine bedeutende deskriptive Statistik für die Analyse von Daten. Wie müssen Sie vorgehen um die Standardabweichung zu berechnen?

- A** ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl ($n - 1$) multiplizieren.
- B** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl ($n - 1$).
- C** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl ($n - 1$). Nicht zu vergessen, am Ende dann noch die Wurzel zu ziehen.
- D** ☐ Den Median berechnen, dann die quadratischen Abstände zum Median aufsummieren, dann die Wurzel ziehen. Am Ende durch die Fallzahl ($n - 1$) teilen
- E** ☐ Den Mittelwert berechnen, dann die quadratischen Abstände zum Mittelwert aufsummieren und durch die Fallzahl ($n - 1$) teilen.

7 Aufgabe

(2 Punkte)

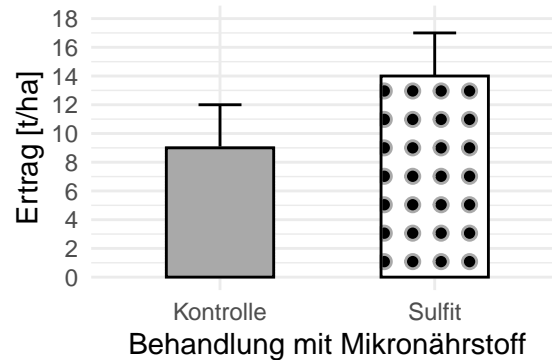
Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile von y mit 20, 10, 12, 3, 27, 26, 16, 33, 5, 23 und 51.

- A** ☐ Es berechnet sich 21 [11; 28]
- B** ☐ Sie erhalten 20 [8; 25]
- C** ☐ Es ergibt sich 20 +/- 10
- D** ☐ Es ergibt sich 20 [10; 27]
- E** ☐ Es berechnet sich 21 [11; 26]

8 Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Sulfit auf den Ertrag in t/ha von Mango im Vergleich zu einer Kontrolle. Der Versuch wurde in 6 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- A** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei 5 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- B** ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei 5.
- C** ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt liegt bei 5.
- D** ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei 5. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.
- E** ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei 0.5.

9 Aufgabe

(2 Punkte)

Ein statistischer Test produziert für einen Gruppenvergleich einen p -Wert. Welche Aussage zusammen mit dem Signifikanzniveau α gleich 5% stimmt?

- A** ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α absolute Werte auf einem Zahlenstrahl und damit den Unterschied der Teststatistiken, wenn die H_0 gilt.
- B** ☐ Wir machen ein Aussage über die Flächen und der Kurve der Teststatistik, wenn die H_0 gilt. Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve repräsentiert werden.
- C** ☐ Wir vergleichen die Effekte des p -Wertes mit den Effekten der Signifikanzschwelle unter der Annahme der Nullhypothese. Dabei gilt, dass wir die Nullhypothese nur ablehnen können anhand des Falsifikationsprinzips.
- D** ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die absoluten Werte auf einem Zahlenstrahl, wenn die H_0 gilt.
- E** ☐ Wir schauen, ob der p -Wert größer ist als das Signifikanzniveau α und vergleichen somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststatistik dargestellt, wenn die H_A gilt.

10 Aufgabe

(2 Punkte)

Berechnen Sie den Mittelwert und Standardabweichung von y mit 12, 14, 19, 15 und 19.

- A** ☐ Es ergibt sich 14.8 +/- 4.85
- B** ☐ Sie erhalten 15.8 +/- 1.76
- C** ☐ Sie erhalten 15.8 +/- 1.555
- D** ☐ Es berechnet sich 15.8 +/- 3.11
- E** ☐ Es berechnet sich 16.8 +/- 9.7

11 Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Yuki und Alex schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten mit einer einfaktoriellen ANOVA auswerten damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse in den multiplen Gruppenvergleichen zu erwarten sind. Da hilft die Katze von Alex auch nur bedingt. Die beiden waren im Emsland um ein Gewächshausexperiment mit Brokkoli durchzuführen. Dabei haben Yuki und Alex den Messwert Trockengewicht [kg/ha] unter der Behandlung Düngestufen (*ctrl*, *low*, *mid* und *high*) ermittelt. Nachher wollen sich beide noch mit dem Hobby Starcraft von Alex beschäftigen. Kennt Yuki noch nicht, klingt aber interessant.

Leider kennen sich Yuki und Alex mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe, die Katze reicht als Hilfe nicht aus!

1. Formulieren Sie die wissenschaftliche Fragestellung! (1 Punkt)
2. Formulieren Sie das statistische Hypothesenpaar! (1 Punkt)
3. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Düngestufen	3				
error	23	897.01			
Total	26	6063.19			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.03$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
5. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? (2 Punkte)

12 Aufgabe

(9 Punkte)

An einem schwülen Sommernachmittag sitzen Alex und Mark in einem Eiskaffee und wollen sich auf die Klausur vorbereiten. In fast allen Fragen geht es ja um die Interpretation eines statistischen Tests. Daher wollen die beiden jetzt nochmal nacharbeiten, was die Grundlagen der Stichprobe (eng. *sample*) und der Grundgesamtheit (eng. *population* oder *ground truth*) sind. Alex hat sich Gummibärchen Eisbecher bestellt und Mark bleibt lieber bei einem Marzipankugeln Eis. 'Irre, was die Lebensmittelindustrie alles auf die Beine kriegt', merk Mark an und Alex schüttelt anerkennend den Kopf.

Leider kennen sich Alex und Mark mit der Grundgesamtheit und der Stichprobe überhaupt nicht aus. Daher sind Sie gefragt!

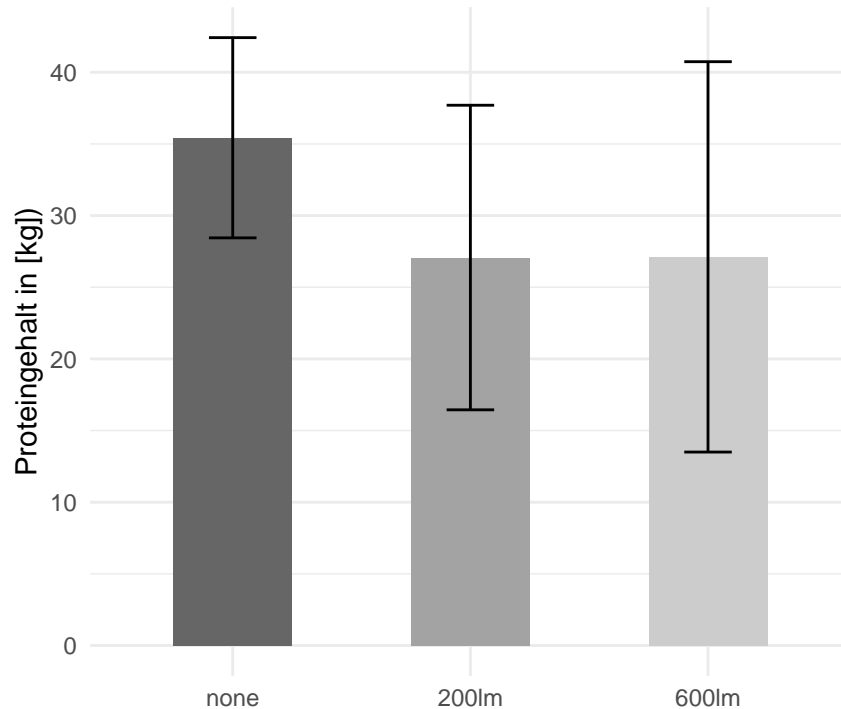
1. Erklären Sie den Zusammenhang zwischen Stichprobe und Grundgesamtheit an einem Schaubild! Beschriften Sie das Schaubild entsprechend! *Nutzen Sie hierfür als Veranschaulichung ein aussagekräftiges Beispiel!* **(3 Punkte)**
 - a) Nennen Sie das statistische Verfahren um von einer Grundgesamtheit auf eine Stichprobe zu gelangen! **(1 Punkt)**
 - b) Nennen Sie ein konkretes Beispiel zur Durchführung um von einer Grundgesamtheit auf eine Stichprobe zu gelangen! **(1 Punkt)**
 - c) Benennen Sie die Eigenschaft, die zwischen Grundgesamtheit und Stichprobe vorliegen muss! **(1 Punkt)**
2. Erweitern Sie das Schaubild um die Entstehung von $Pr(D|H_0)$! *Nutzen Sie hierfür als Veranschaulichung ein aussagekräftiges Beispiel!* **(3 Punkte)**


13 Aufgabe


(8 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Jessica ist es eine Möglichkeit schneller ans Ziel zu gelangen. Jessica soll in ihrer Hausarbeit Kartoffeln untersuchen. Die Behandlung in ihrer Hausarbeit werden verschiedene Lichtstufen (*none*, *200lm* und *600lm*) sein. Erheben wird Jessica als Messwert (Y) *Proteingehalt* benannt als *protein* in ihrer Exceldatei. Von ihrer Betreuer erhält sie nun folgende Abbildung von Barplots, die sie erstmal zur Übung nachbauen soll, bevor sie mit dem eigentlichen Versuch beginnt.



Leider kennt sich Jessica mit der Erstellung von Barplots in  nicht aus. Deshalb braucht sie bei der Visualisierung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Erstellen Sie eine Tabelle mit den statistischen Maßzahlen der drei Barplots! *Beachten Sie die korrekte Darstellungsform der statistischen Maßzahlen!* **(3 Punkte)**
3. Erstellen Sie einen beispielhaften Datensatz im  üblichen Format, aus dem die drei Barplots *möglicherweise* erstellt wurden! **(2 Punkte)**
4. Kann Jessica einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

14 Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Der t-Test. Alex erschauert. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der einen normalverteilten Messwert (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Alex überhaupt aus? Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Alex hat einen Leistungssteigerungsversuch mit Lamas durchgeführt um eine neue technische Versuchsanlage zu testen. Bei dem Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) wurde die Behandlung Ernährungszusatz (*ctrl* und *fedX*) an den Lamas getestet und dabei wurde geschaut, ob der Versuch überhaupt technisch klappen könnte. Gemessen hat Alex dann als Messwert Gewichtszuwachs in der 1LW [%/kg]. Warum der Versuch im Oldenburger Land für seinen Projektbericht stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Gewichtszuwachs in der 1LW [%/kg]?

Ernährungszusatz	Gewichtszuwachs
ctrl	13.0
fedX	25.0
ctrl	11.7
ctrl	15.4
fedX	23.0
fedX	21.5

Leider kennt sich Alex mit der Berechnung eines Student t-Tests überhaupt nicht aus. Deshalb braucht er bei der Berechnung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
5. Formulieren Sie eine Antwort an Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

15 Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Alex und Yuki schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten visualisieren damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse zu erwarten sind. Die beiden waren in der Uckermark um einen Versuch in einer Klimakammer mit Maiss durchzuführen. Dabei haben Alex und Yuki den Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] unter der Behandlung Lichtstufen (*none*, 200lm und 600lm) ermittelt. Kennengelernt haben sich die beiden auf einem Konzert von London Grammar. Später wird noch Matrix geguckt. Yuki befürwortet das!

Lichtstufen	Chlorophyllgehalt
200lm	41
200lm	39
none	38
none	37
600lm	41
200lm	40
200lm	40
200lm	40
none	40
600lm	42
600lm	39
600lm	40
600lm	40
200lm	40
none	40
none	41
none	40
none	39

Leider kennen sich Alex und Yuki mit Darstellung einer einfaktoriellen ANOVA überhaupt nicht aus.

1. Erstellen Sie eine Visualisierung der Datentabelle! Beschriften Sie die Abbildung! **(2 Punkte)**
2. Benennen Sie die Visualisierung mit dem korrekten, statistischen Fachbegriff! **(1 Punkt)**
3. Zeichnen Sie folgende statistischen Maßzahlen passend ein!
 - Den globalen Mittelwert β_0 **(1 Punkt)**
 - Die Mittelwerte der einzelnen Behandlungsstufen **(1 Punkt)**
 - Die Mittelwertsdifferenz der einzelnen Behandlungsstufen mit β_{none} , β_{200lm} und β_{600lm} **(1 Punkt)**
 - Die Residuen oder Fehler mit ϵ **(1 Punkt)**
4. Liegt ein *vermutlicher* signifikanter Unterschied vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Schätzen Sie die Effekte der Behandlungsstufen! **(2 Punkte)**

16 Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Mark steht vor einem ersten Problem, denn wenn es nach seinem Betreuer geht, soll er in einem Feldexperiment Maiss auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Geocaching. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Mark denkt gerne über Geocaching nach. Die Behandlung waren verschiedene Lichtstufen (*none* und *600lm*). In seiner Excel-datei hat er den Messwert (*Y*) *Trockengewicht* als *drymatter* aufgenommen. Nun soll Mark die Daten einmal als Boxplots in einer Präsentation visualisieren, damit seinem Betreuer wieder klar wird, was er eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Anhand von Boxplots lässt sich eine Aussage über die Varianzhomogenität über die Behandlungsgruppen treffen. Wäre da nicht noch etwas. Eine echte Herausforderung für ihn war schon immer die Unsicherheit gewesen. Ein leidiges Lied. Aber egal. Mark will später nochmal raus um zu Reiten. Druck ablassen, dass muss er auch.

treatment	drymatter
600lm	28.2
600lm	13.1
none	31.5
none	29.2
none	20.0
600lm	38.7
none	30.1
600lm	24.4
none	31.7
600lm	36.3
none	34.9
600lm	20.3
none	15.9
600lm	10.8
none	34.1
none	31.9




Leider kennt sich Mark mit der Erstellung von Boxplots nicht aus. Deshalb braucht er bei der Visualisierung Ihre Hilfe!

1. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Behandlungen von Maiss! Beschriften Sie die Achsen entsprechend! **(5 Punkte)**
2. Wie ist Ihr Vorgehen, wenn Sie eine *gerade* Anzahl an Beobachtungen pro Gruppe haben? **(1 Punkt)**
3. Beschriften Sie *einen* der beiden Boxplots mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
4. Wenn Sie keinen Effekt zwischen den Behandlungen von Maiss erwarten würden, wie sehen dann die beiden Boxplots aus? Antworten Sie mit einer Skizze der Boxplots! **(1 Punkt)**

17 Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

‘Wir sind uns relativ sicher, dass unser Messwert Fettgehalt [%/kg] ist!’, ruft Steffen wild gestikulierend. Steffen wäre mehr präsent, wenn es die Romantik nicht gäbe. Als würde sowas die Ausgabe von  interessieren. Steffen und Tina sind in einem Café mit Nilufar um sich Hilfe von ihr in  zu holen. Während Nilufar Kirschstreuselkuchen und Takis Blue Heat mampft, versuchen die Steffen und Tina ihren Versuch in der Uckermark mit Schweinen in einem Stallexperiment zu erklären. Nilufar hofft insgeheim, dass die  Ausgabe des t-Tests ihr mehr Informationen liefert. Eigentlich würde sie dann doch lieber raus um zu Kicken vielleicht mit Tina?

```
##
## Two Sample t-test
##
## data: Fettgehalt by Bestandsdichte
## t = -0.74325, df = 15, p-value = 0.4688
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -14.520079 7.011746
## sample estimates:
## mean in group Verordnung mean in group Gesteigert
##                25.31250                29.06667
```

Helfen Sie Nilufar bei der Interpretation des t-Tests! Sonst geht es auch für Steffen und Tina nicht weiter.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**
5. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.13|$ einzeichnen! **(4 Punkte)**
6. Beschriften Sie die Abbildung! **(1 Punkt)**