

Name: _____

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

**B.Sc. Landwirtschaft, B.Eng. Wirtschaftsingenieurwesen im Agri- und Hortibusiness, B.Sc.
Angewandte Pflanzenbiologie - Gartenbau, Pflanzentechnologie**

Klausur Angewandte Statistik und Versuchswesen

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

25. Juni 2025

Erlaubte Hilfsmittel

- Normaler Taschenrechner ohne Möglichkeit der Kommunikation mit anderen Geräten! Ausdrücklich kein Handy!
- Eine DIN A4-Seite als beidseitig, selbstgeschriebene, handschriftliche Formelsammlung. Keine digitalen Ausdrücke!
- **Die Verwendung eines roten Farbstiftes ist nicht gestattet! Korrekturfarbe!**
- *You can answer the questions in English without any consequences.*

Endnote

_____ von 20 Punkten sind aus den Multiple Choice Aufgaben erreicht.
_____ von 73 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 93 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
89.0 - 93.0	1,0
84.5 - 88.5	1,3
79.5 - 84.0	1,7
75.0 - 79.0	2,0
70.5 - 74.5	2,3
66.0 - 70.0	2,7
61.5 - 65.5	3,0
56.5 - 61.0	3,3
52.0 - 56.0	3,7
46.5 - 51.5	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.

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

- 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	8	10	10	12	12	12

- Es sind ____ von 73 Punkten erreicht worden.

1 Aufgabe

(2 Punkte)

Die Testtheorie hat mehrere Säulen. Einer der Säulen ist das Falsifikationsprinzip. Das Falsifikationsprinzip besagt,

- A** ☐ ... dass in der Wissenschaft immer etwas falsch sein muss. Sonst gebe es keinen Fortschritt.
- B** ☐ ... dass ein schlechtes Modell durch das Falsifikationsprinzip durch ein noch schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- C** ☐ ... dass Modelle meist falsch sind und selten richtig.
- D** ☐ ... dass Fehlerterme in statistischen Modellen nicht verifiziert werden können.
- E** ☐ ... dass ein minderwertiges Modell durch ein weniger minderwertiges Modell ersetzt wird. Es gilt das Falsifikationsprinzip nach Karl Popper.

2 Aufgabe

(2 Punkte)

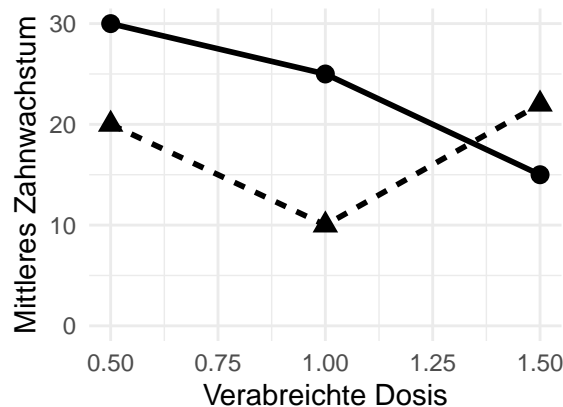
Sie haben folgende unadjustierten p-Werte gegeben: 0.03, 0.21, 0.001 und 0.02. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0075, 0.0525, 3e-04 und 0.005. Die adjustierten p-Werte werden zu einem α -Niveau von 1.25% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.12, 0.84, 0.004 und 0.08. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.12, 0.84, 0.004 und 0.08. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.12, 0.84, 0.004 und 0.08. Die adjustierten p-Werte werden zu einem α -Niveau von 1.25% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0075, 0.0525, 3e-04 und 0.005. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.

3 Aufgabe

(2 Punkte)

Ein Versuch wurde an 52 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung von Vitamin D auf das Zahnwachstum bei Kanarienvögel. Welche Aussage ist richtig, wenn Sie eine zweifaktorielle ANOVA rechnen?



- A** ☐ Das Bestimmtheitsmaß R^2 ist groß.
- B** ☐ Mit ($p \leq 0.05$) liegt eine mittlere bis starke Interaktion vor.
- C** ☐ Es liegt keine Interaktion vor ($p > 0.05$).
- D** ☐ Eine positive Interaktion liegt vor ($\rho \leq -0.5$)
- E** ☐ Keine Korrelation liegt vor ($\rho \geq 0.05$).

4 Aufgabe

(2 Punkte)

Um zu überprüfen, ob die Daten die Annahme einer Normalverteilung genügen, können wir folgende Visualisierung nutzen. Dabei kommt dann auch die entsprechende Regel zur Abschätzung der Annahme einer Normalverteilung zur Anwendung.

- A** ☐ Einen Boxplot. Das IQR muss über alle Behandlungen zusammen mit den Whiskers ungefähr gleich aussehen.
- B** ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur audreihen. Eine Abweichung führt zur Ablehnung der Annahme einer Normalverteilung.
- C** ☐ In einer explorativen Datenanalyse nutzen wir den Boxplot. Dabei sollte der Median als dicke Linie in der Mitte der Box liegen. Dann können wir von einer Normalverteilung ausgehen.
- D** ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Normalverteilung angenommen werden.
- E** ☐ Nach dem Einlesen der Daten nutzen wir einen Barplot um zu schauen, ob alle Mittelwerte über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch die Varianz in allen Behandlungen in etwa gleich.

5 Aufgabe

(2 Punkte)

Viele statistische Verfahren nutzen eine Teststatistik um eine Aussage über den Zusammenhang zwischen der Grundgesamtheit und der Stichprobe abzubilden. Ein statistisches Testwerkzeug ist hierbei die ANOVA. Die ANOVA rechnet dabei...

- A** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen oder der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss sich zwischen einem der beiden Varianzquellen entschieden werden.
- B** ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsgruppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- C** ☐ ... den Unterschied zwischen der Varianz ausgelöst durch alle Behandlungsgruppen und der Varianz aus globalen Behandlungsgruppen der Kontrollen. Wenn die ANOVA nicht signifikant ist, muss ein Posthoc-Test ausgeschlossen werden.
- D** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss ein Posthoc-Test angeschlossen werden.
- E** ☐ ... den Unterschied zwischen der Varianz in den verschiedenen Behandlungsgruppen und der Varianz in einer der Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in einer der Gruppen exakt zu bestimmen.

6 Aufgabe

(2 Punkte)

Welche Aussage über den Effekt eines statistischen Tests ist richtig?

- A** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Zum Beispiel den mittleren Unterschied zwischen zwei Gruppen aus einem t-Test. Damit ist der Effekt direkt mit dem Begriff der Relevanz verbunden. Die Entscheidung über die Relevanz trifft der Forschende unabhängig von der Signifikanz eines statistischen Tests.
- B** ☐ Durch den Effekt erfahren wir die statistische interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das η^2 aus einer ANOVA. Damit können wir die Signifikanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutend zu bewerten ist.
- C** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Modernen Algorithmen liefern keine Effekte mehr sondern nur noch bedingte Wahrscheinlichkeiten. Der Effekt spielt in der modernen Statistik keine Rollen mehr.
- D** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.
- E** ☐ Der Effekt eines statistischen Tests beschreibt die mathematisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.

7 Aufgabe

(2 Punkte)

Betrachten wir die Teststatistik aus einem abstrakteren Blickwinkel. Beim statistischen Testen wird das extitsignal mit dem extitnoise aus den Daten D zu einer Teststatistik T_D verrechnet. Welche der Formel berechnet korrekt die Teststatistik T_D ?

- A** ☐ Bei der Berechnung der Teststatistik T_D gewichten wir den Effekt *signal* mit der Varianz *noise*. Wir können verallgemeinert $T_D = \text{signal}/\text{noise}$ schreiben.
- B** ☐ Es gilt $T_D = \text{signal} \cdot \text{noise}$. Der Effekt *signal* wird mit der Varianz *noise* gewichtet.
- C** ☐ Es gilt $T_D = \text{noise}/\text{signal}$. Der Effekt *noise* wird mit der Varianz *signal* gewichtet.
- D** ☐ Wir gewichten den Effekt *noise* mit der Varianz *signal* und erhalten $T_D = \text{signal}/\text{noise}$.
- E** ☐ Wir gewichten den Effekt *noise* mit der Varianz *signal* und erhalten $\text{signal}/\text{noise}^2$.

8 Aufgabe

(2 Punkte)

Sie sollen in Ihrer Abschlussarbeit die Relevanz und die Signifikanz in einer statistischen Maßzahl vereinen. Welche Aussage ist richtig?

- A** ☐ Der p-Wert. Durch den Vergleich mit α lässt sich über die Signifikanz entscheiden und der β -Fehler erlaubt über die Power eine Einschätzung der Relevanz.
- B** ☐ Das OR. Als Chancenverhältnis gibt es das Verhältnis von Relevanz und Signifikanz wieder.
- C** ☐ Über das Konfidenzintervall. Das Konfidenzintervall inkludiert eine Entscheidung über die Relevanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Signifikanzschwelle vom Forschenden definiert werden.
- D** ☐ Das Konfidenzintervall. Durch die Visualisierung des Konfidenzintervalls kann eine Relevanzschwelle vom Anwender definiert werden. Zusätzlich erlaubt das Konfidenzintervall auch eine Entscheidung über die Signifikanz.
- E** ☐ Die Teststatistik. Durch den Vergleich von T_c zu T_k ist es möglich die H_0 abzulehnen. Die Relevanz ergibt sich aus der Fläche rechts vom dem T_c -Wert.

9 Aufgabe

(2 Punkte)

Auf wissenschaftlichen Postern finden Sie unter Abbildungen häufig die Abkürzung *CLD*. Für welchen statistischen Fachbegriff steht die Abkürzung und wie interpretieren Sie ein *CLD*?

- A** ☐ Compact line display. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt. Früher wurden keine Buchstaben sondern eine durchgezogene Linie verwendet. Bei mehr als drei Gruppen funktioniert die Linie aber graphisch nicht mehr.
- B** ☐ 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.
- C** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.
- D** ☐ Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.
- E** ☐ 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.

10 Aufgabe

(2 Punkte)

Sie wollen nach einem Feldversuch die Varianz berechnen. Welche der folgenden Rechenoperationen müssen durchgeführt werden?

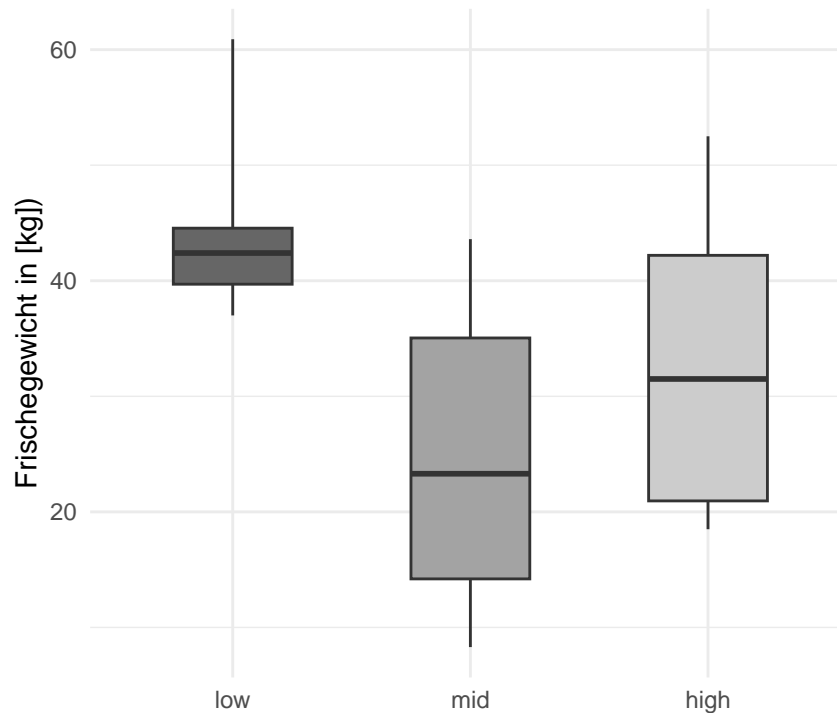
- A** ☐ Den Mittelwert berechnen, dann die absoluten Abstände zum Mittelwert aufsummieren. Die Fallzahl $(n - 1)$ entsprechend gewichten.
- B** ☐ Den Mittelwert berechnen, dann die quadratischen Abstände zum Mittelwert aufsummieren und durch die Fallzahl $(n - 1)$ teilen.
- C** ☐ Wir berechnen erst den Mittelwert und dann die quadratischen Abstände zu dem Mittelwert. Diese quadratischen Abstände summieren wir auf und teilen am Ende durch die Fallzahl $(n - 1)$. Als letzten Schritt ziehen wir die quadratische Wurzel.
- D** ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl $(n - 1)$ multiplizieren.
- E** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl $(n - 1)$.

11 Aufgabe

(9 Punkte)

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

Boxplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Steffen nicht gemerkt, welche statistischen Maßzahlen für einen Boxplot erhoben werden müssen. Besser wäre was anderes gewesen. Klemmbausteine. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Steffen denkt gerne über Klemmbausteine nach. Das ist in soweit doof, da nach seinem Betreuer erstmal ein Boxplot nachgebaut werden soll, bevor es mit seiner Abschlussarbeit losgeht. Dann hat er schonmal den R Code vorliegen und nachher geht dann alles schneller. Na dann mal los. Steffen schafft sich die nötige Stimmung. Steffen nickt im Takt von Taylor Swift und bemerkt dabei gar nicht was die Schlange schon wieder anstellt. In der Behandlung für Kartoffeln werden verschiedene Bewässerungstypen (*low*, *mid* und *high*) sein. Erfasst wird als Endpunkt (*Y*) *Frischegewicht*. Steffen soll dann *freshmatter* in seiner Exceldatei eintragen. Aber nur in passender Atmosphäre! Auf seinem Second Screen läuft Harry Potter und Steffen schaufelt Oreos. Nicht effizient, aber gut.



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

1. Erstellen Sie eine Tabelle mit den statistischen Maßzahlen aus der obigen Abbildung der drei Boxplots! *Beachten Sie die korrekte Darstellungsform der statistischen Maßzahlen!* **(3 Punkte)**
2. Beschriften Sie *einen* der Boxplots mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
3. Erstellen Sie einen beispielhaften Datensatz, aus dem die drei Boxplots *möglicherweise* erstellt wurden, im R üblichen Format! **(2 Punkte)**
4. Kann Steffen einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

12 Aufgabe

(8 Punkte)

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

In einem Gespräch mit seinem Betreuer wird Yuki gebeten seine Daten aus einem Stallexperiment mit Lamas in einem Histogramm darzustellen. Yuki schmeißt noch eine Handvoll Reese's Peanut Butter Cups in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von London Grammar. In seinem Experiment hat er die mittleren dunklen Pigmentstörungen erst fotografiert und dann ausgezählt. Laut seinem Betreuer soll das Histogramm helfen, die Verteilung der die mittleren dunklen Pigmentstörungen zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Wenn die Faulheit nicht wäre, ja dann wäre wohl vieles möglich für Yuki! Aber so.. Yuki streichelt liebevoll das Minischwein. Der Kopf ist in seinem Schloß vergraben um den Klang von London Grammar zu dämpfen.

Die mittleren dunklen Pigmentstörungen: 8.2, 8.5, 8.6, 11.6, 9.8, 11.3, 8.7, 8.9, 12.1, 8.2, 9.1, 9.9, 14.8, 9.7, 10.9, 11.5, 9.1, 9.5, 13.4, 8, 8.8, 12, 8.5, 12.5, 10.9, 11.4, 11.2, 6.5

Leider kennt sich Yuki mit der Erstellung von Histogrammen überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Zeichnen Sie ein Histogramm um die Verteilung der Daten zu visualisieren! **(3 Punkte)**
2. Erläutern Sie Ihr Vorgehen um ein Histogramm für kontinuierliche Daten zu zeichnen! **(2 Punkte)**
3. Beschriften Sie die Achsen der Abbildung! **(2 Punkte)**
4. Ergänzen Sie die relativen Häufigkeiten in der Abbildung! **(1 Punkt)**

13 Aufgabe

(10 Punkte)

'Okay, für was war jetzt nochmal das 95% Konfidenzintervall gut?', fragt Alex und schaut in das leere Gesicht von Tina. 'Keine Ahnung. Irgendwas mit Relevanz und Effekt oder Signifikanz. Da kannst du irgendwie was verbinden. Keine Ahnung warum', entgegnet Tina. 'Wir haben doch als Messwert *Wasserverbrauch der Bewässerung* erhoben.', stellt Alex fest. Jetzt haben beide das Problem, die möglichen 95% Konfidenzintervalle zu interpretieren.

Leider kennen sich Alex und Tina mit der Visualisierung des 95% Konfidenzintervall überhaupt nicht aus.


1. Beschriften Sie die untenstehende Abbildung mit der Signifikanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
2. Ergänzen Sie eine *in den Kontext passende* Relevanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Skizzieren Sie in die untenstehende Abbildung sechs einzelne Konfidenzintervalle (a-f) mit den jeweiligen Eigenschaften! **(6 Punkte)**
 - (a) Ein signifikantes, relevantes 90% Konfidenzintervall.
 - (b) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall
 - (c) Ein signifikantes, relevantes 95% Konfidenzintervall
 - (d) Ein 95% Konfidenzintervall mit höherer Varianz s_p in der Stichprobe als der Rest der 95% Konfidenzintervalle
 - (e) Ein 95% Konfidenzintervall mit niedriger Varianz s_p in der Stichprobe als der Rest 95% der Konfidenzintervalle
 - (f) Ein signifikantes, nicht relevantes 95% Konfidenzintervall



14 Aufgabe

(10 Punkte)

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

'Wir waren in der Uckermark um Maiss in einem Gewächshausexperiment zu messen.', Jessica legt das Dokument auf den Tisch und schaut Alex und Mark fragend an. Beide schauen fragend zurück. Gäbe es die Unsicherheit nicht, dann wäre es für Mark irgendwie einfacher hier zu helfen. Echt unangenehm. Die beiden sind zu Jessica gekommen, da sie sich nicht mit  auskennen und daher Hilfe bei der Interpretation des t-Tests brauchen. Im Hintergrund wummert David Bowie und leere Schokobons Packungen stapeln sich auf dem Boden. 'Kein Problem', sagt Jessica und streichelt langsam die Hündin. 'Aber worum es in dem Versuch geht, lässt sich nur aus dem Text in seiner Hand erahnen.' merkt sie an. Vielleicht hilft da ja die Ausgabe des t-Tests in R weiter. Draußen geht blutrot die Sonne unter.

```
##  
## Two Sample t-test  
##  
## data: Trockengewicht by Lüftungssystemen  
## t = 0.74576, df = 13, p-value = 0.4691  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -4.935199 10.138771  
## sample estimates:  
## mean in group ctrl mean in group tornado  
## 37.31429 34.71250
```

Helfen Sie Jessica bei der Interpretation des t-Tests! Sonst geht es auch für Alex und Mark 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. Skizzieren Sie das sich ergebende 95% Konfidenzintervall! **(2 Punkte)**
5. Beschriften Sie die Abbildung und das 95% Konfidenzintervall entsprechend! **(2 Punkte)**
6. Interpretieren Sie den Effekt des 95% Konfidenzintervalls! **(2 Punkte)**

15 Aufgabe

(12 Punkte)

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

Tina und Jessica 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. Deshalb erstmal Schokobons mampfen, die Jessica mitgebracht hat. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA-Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Tina schaut Jessica sehen erstmal gar nichts. Die beiden waren in der Uckermark um ein Kreuzungsexperiment mit Schweinen durchzuführen. Dabei haben Tina und Jessica den Messwert Schlachtgewicht [kg] unter der Behandlung Genotypen (AA, AB und BB) ermittelt. Später wollen die beiden dann noch raus um Rad zu fahren.

Leider kennen sich Tina und Jessica mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

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)
Genotypen	2	182.86			
Error	19	142.96			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.52$ ab. Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA? **(1 Punkt)**
6. Berechnen Sie *einen* Student t-Test für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_{\alpha=5\%} = 2.03$. Begründen Sie Ihre Auswahl! **(3 Punkte)**

Genotypen	Fallzahl (n)	Mittelwert	Standardabweichung
AA	7	11.86	2.85
AB	5	4.00	2.35
BB	10	9.30	2.83

7. Gegebenen der ANOVA-Tabelle war das Ergebnis des Student t-Tests zu erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

16 Aufgabe

(12 Punkte)

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

Nilufar steht in der Uckermark. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei ihren Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Hühnern. Das Huhn durchbohrt sie mit leeren Blick. 'Woher zum Teufel!', entfährt es ihr. Aber da ist es schon weg. Ja, darum geht es in ihrer Hausarbeit. Und wäre das nicht noch alles genug, ist ihr Experiment auch noch als einen Leistungssteigerungsversuch komplex geraten. Es wurde der Messwert Gewichtszuwachs in der 1LW mit dem Behandlung Genotypen (AA, AB und BB) sowie der Behandlung Elterlinie (*ctrl*, und *Xray*) untersucht. 'Hmpf', denkt Nilufar und ruft 'Und jetzt!?' in die Leere. Und eigentlich wollte Nilufar doch noch ihrem Hobby nachgehen! Am Ende dann doch besser Hip Hop. Wunderbar. Eine echte Ablenkung für Nilufar.

Leider kennen sich Nilufar mit Berechnung einer zweifaktoriellen ANOVA überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung für beide Faktoren separat! (2 Punkte)
2. Formulieren Sie die statistischen Hypothesenpaare für beide Faktoren separat! (2 Punkte)
3. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Genotypen	3	483.71			
Elterlinie	1	124.07			
Genotypen:Elterlinie	3	20.99			
Error	18	297.84			

4. Schätzen Sie den p-Wert der Tabelle ab. Begründen Sie Ihre Antwort! (3 Punkte)


	$F_{\alpha=5\%}$
Genotypen	4.26
Elterlinie	3.40
Genotypen:Elterlinie	5.23

5. Was sagt der Term *Genotypen:Elterlinie* aus? Interpretieren Sie das Ergebnis! (2 Punkte)


17 Aufgabe

(12 Punkte)

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

'Hä? Was ist denn das? Hatten wir das als Aufgabe eine lineare Regression zu rechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?', fragt Tina. Steffen schaut fragend zurück. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern Zahlen in einer Tabelle...', antwortet Steffen leicht angespannt. Die beiden hatten ein Stallexperiment im Teutoburgerwald mit Fleischrindern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Bewegungsscore [Movement/h] und Schlachtgewicht [kg]. Das Bestimmtheitsmaß R^2 hatten die beiden mit 0.3 bestimmt. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der  Ausgabe haben beide jetzt ein Problem.

term	estimate	std.error	t statistic	p-value
(Intercept)	0.6815233	2.1934600		
Durchschnittlicher Bewegungsscore	1.6428104	0.2252332		

Leider kennen sich Tina und Steffen mit der linearen Regression für kontinuierliche Daten in  überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die Regressionsgleichung! **(1 Punkt)**
3. Erstellen Sie eine Visualisierung der `lm()`-Ausgabe. *Beachten Sie die Informationen zum Bestimmtheitsmaß R^2 aus dem Aufgabentext!* Beschriften Sie die Achsen! **(2 Punkte)**
4. Beschriften Sie die Visualisierung mit den statistischen Maßzahlen der `lm()`-Ausgabe! **(2 Punkte)**
5. Ergänzen Sie die t-Statistik in der `lm()`-Ausgabe! **(2 Punkte)**
6. Ergänzen Sie den p-Wert in der `lm()`-Ausgabe mit $T_{\alpha=5\%} = 1.96$! **(2 Punkte)**
7. Interpretieren Sie den p-Wert im Kontext der wissenschaftlichen Fragestellung! **(1 Punkt)**
8. Wie groß ist der Effekt im Kontext der wissenschaftlichen Fragestellung? **(1 Punkt)**