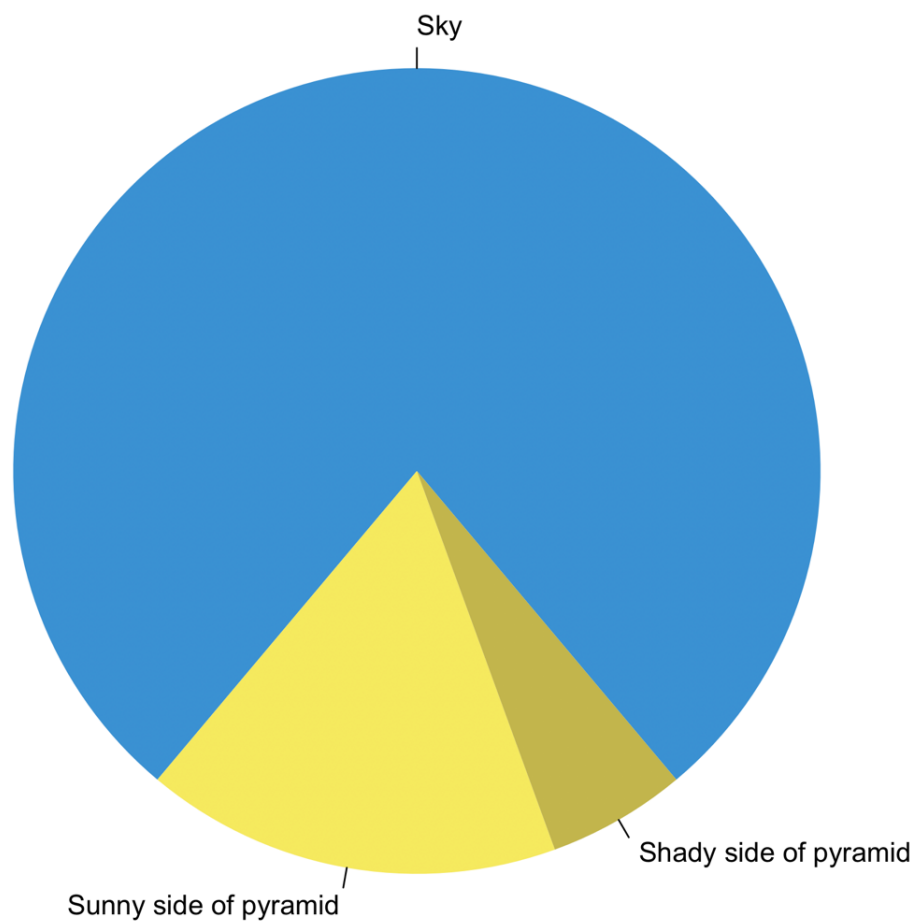


Grundlagen für das Research Project

*Anleitung zum Erstellen von Graphen und zum
Berechnen von Korrelation und Signifikanz*

Christoph.benkler@gymnasium-immensee.ch

Flavio.carrera@gymnasium-immensee.ch



Version: 2.6.20

Inhaltsverzeichnis

<i>Anleitung zum Erstellen von Graphen und zum Berechnen von Korrelation und Signifikanz</i>	1
Inhaltsverzeichnis	2
Welche Faktoren möchtest du erforschen?	4
Welche Arten von Faktoren gibt es?.....	5
Möglichkeit 1: Beide Faktoren sind numerische Faktoren.....	5
Möglichkeit 2: Ein Faktor ist numerisch, der zweite kategorischer	5
Möglichkeit 3: Beide Faktoren sind kategorisch	5
Wie kann ich zwei numerische Faktoren erforschen?	6
Ausgangslage	6
Grafisch	6
Rechnerisch.....	7
r-Wert berechnen (Korrelationskoeffizient)	7
p-Wert berechnen (Signifikanz)	7
Wie kann ich numerische und kategorische Faktoren erforschen?.....	9
Ausgangslage	9
Grafisch	9
Rechnerisch.....	11
r-Wert berechnen (Korrelationskoeffizient)	11
Vorbereitung um Signifikanz (p-Wert) zu bestimmen	12
p-Wert berechnen (Korrelationskoeffizient).....	13
Welchen der drei t-Tests führe ich aus?	13
Welchen P-Wert wähle ich?.....	13
Wie kann ich zwei kategorische Faktoren erforschen?	15
Ausgangslage	15
Grafisch	15
Beobachtete Werte auflisten.....	15
Grafik erstellen.....	16
Rechnerisch.....	Fehler! Textmarke nicht definiert.
r-Wert berechnen	Fehler! Textmarke nicht definiert.

p-Wert berechnen.....	Fehler! Textmarke nicht definiert.
Flussdiagramme	17
Tools / Hilfe.....	19
Was ist schon wieder der r-Wert?	19
Was ist schon wieder der p-Wert?.....	19
Was ist schon wieder der Median?	19
Schöne Datenquellen zum Experimentieren	20
Wo finde ich das Add-In in Excel für Microsoft? 🤖	20
in Excel für Mac? 🐼	20
in Excel online? 😭	21
Weiterführende Quellen.....	21
Noch weiter weiterführende Statistik.....	22
Was ist Varianz?	22
Was ist Varianz?	22
Hilfe, ich will ganz viele Kategorien miteinander vergleichen, z.B. Sternzeichen 🤖	22
2 kategorielle Variablen & Mosaikplots	23

Welche Faktoren möchtest du erforschen?

Am Anfang war eine Frage. Du solltest dir im Klaren darüber sein, was genau du erforschen möchtest, oder noch konkreter: Welche Faktoren du erforschen möchtest. Vielleicht hast du eine kausale Hypothese im Hinterkopf, der du nun statistisch auf den Grund gehen möchtest?

Beispiel 1:

Fragestellung: Haben Schüler*innen mit mehr Taschengeld auch teurere Kleidung im Schrank?

Hypothese: Je mehr Taschengeld, desto mehr Geld steht für Kleidung zur Verfügung, deshalb leisten sich Personen mit hohem Taschengeld die teureren Kleidungsstücke.

- ➔ Faktor 1: Taschengeld in Schweizer Franken
- ➔ Faktor 2: Teuerstes Kleid im Kleiderschrank in Schweizer Franken

Beispiel 2:

Fragestellung: Haben Schüler*innen, die rauchen einen tieferen Blutdruck als Nichtraucher?

Hypothese: Rauchen verengt die Blutgefäße, darum steigt bei ihnen der Blutdruck.

- ➔ Faktor 1: Raucher / Nichtraucher
- ➔ Faktor 2: Blutdruck

Beispiel 3:

Fragestellung: Sind Schüler*innen, die am häufigsten TikTok verwenden häufiger Gamer?

Hypothese: Auf TikTok werden ...

- ➔ Faktor 1: Lieblings Social Media-App
- ➔ Faktor 2: Gamer (Ja/Nein)

Welche Arten von Faktoren gibt es?

Jetzt, wo du weisst, was du erforschen willst, müssen wir uns die Faktoren etwas genauer anschauen, die du untersuchen willst. Es existieren zwei Arten von Faktoren, die wir numerische und die kategorischen Faktoren nennen wollen:

1. **Numerische Faktoren:** Sind graduelle Werte, die sich über ein Spektrum erstrecken.
 - Taschengeld in Schweizer Franken
 - Körpergrösse in Zentimeter
2. **Kategorische Faktoren:** Sind Faktoren, die in Gruppen aufgeteilt sind.
 - Geschlecht: Weiblich, männlich, divers
 - Sternzeichen: Widder, Wassermann, Skorpion, Archaeopteryx, ...

Wenn du in deinen Untersuchungen nun Faktoren auf mögliche Zusammenhänge, Unterschiede und kausale Beziehungen hin untersuchen willst, hängt die Vorgehensweise davon ab, welche Arten von Faktoren du untersuchst. Darum ist es ganz wichtig, dass du zuerst herausfindest, welche Art von Faktoren es sind.

? Welche Möglichkeit trifft auf dein Forschungsvorhaben zu?

Möglichkeit 1: Beide Faktoren sind numerische Faktoren

(Numerischer Faktor | Numerischer Faktor)

- ➔ Beispiel 1: «Wie gross bist du?» | «Was hast du für eine Schuhgrösse?»
- ➔ Beispiel 2: «Wie viel Ozon befindet sich in der Luft?» | «Wie viel Sonneneinstrahlung haben wir?»

Möglichkeit 2: Ein Faktor ist numerisch, der zweite kategorischer

(Numerischer Faktor | Kategorischer Faktor)

- ➔ Beispiel 3: «Welches Geschlecht hast du?» | «Welches ist deine Lieblingsfarbe?»
- ➔ Beispiel 4: «Wie viele Einwohner besitzt dein Dorf?» | «Welche Religion dominiert im Dorf.»

Möglichkeit 3: Beide Faktoren sind kategorisch

(Numerischer Faktor | Numerischer Faktor)

- ➔ Beispiel 5: Welches Sternzeichen hast du? | Welches ist deine Lieblingsmusik?
- ➔ Beispiel 6: Wie

Wie kann ich zwei numerische Faktoren erforschen?

Schwierigkeitsgrad



Ausgangslage

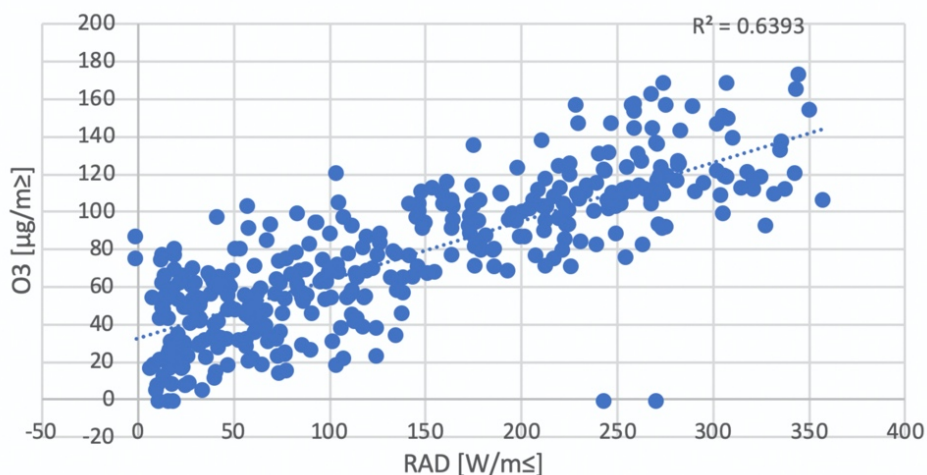
Du hast dir also beispielsweise die Frage gestellt, ob Schüler*innen mit mehr Taschengeld auch teurere Kleidung im Schrank haben. Deine Vermutung ist, dass dies der Fall ist. Wie kannst du diese zwei numerischen Faktoren nun auswerten?

Grafisch

Für deine grafische Darstellung zweier numerischer Faktoren eignet sich der so genannte Scatterplot (Punktediagramm). Du findest diesen in Excel unter «Diagrammen». Wenn du alle Daten auswählst und «Scatterplot» anwählst, kreiert Excel für dich ein leicht interpretierbares Punktediagramm.

Scatterplot (X/Y)

Ozon vs. Sonneneinstrahlung



Wichtig: Achte auf folgende Punkte:

- Füge eine Trendlinie hinzu («Diagrammelement hinzufügen» / «Trendlinie» / «linear»)
- Beschrifte beide Achsen
- Beschneide gegebenenfalls die Achsen
- Gib die Stichprobengröße an (z.Bsp.: N=100)

Rechnerisch

r-Wert berechnen (Korrelationskoeffizient)

Der Korrelationskoeffizient «r» lässt sich in Excel einfach berechnen. Nutze hierzu die folgende Formel:

=KORREL(...;...)

Das Strichpunkt (Semikolon) trennt die beiden Faktoren. Wenn du nicht mehr weisst, was Korrelation überhaupt bedeutet, navigiere zur Hilfe, oder klicke hier:

➔ [Hilfe, was ist Korrelation schon wieder?](#)

p-Wert berechnen (Signifikanz)

Um den p-Wert (d.h. die Signifikanz) zweier numerischer Faktoren zu berechnen, musst du die Datenanalyse starten. Damit die Datenanalyse auch erscheint, muss ein «Ad-In» installiert sein. Findest du die Datenanalyse nicht unter «Daten» (ganz rechts), klicke unten:

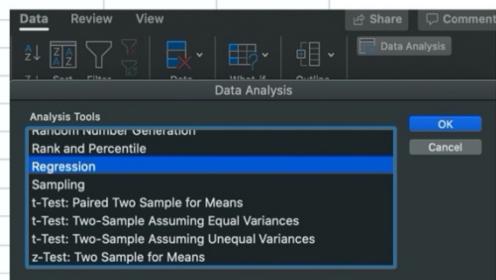
➔ [Hilfe, ich finde die Datenanalyse nicht auf meinem Mac?!?](#)

➔ [Hilfe, ich finde die Datenanalyse nicht auf meinem PC?!?](#)

Hast du das Ad-In installiert, startest du die Datenanalyse und wählst «Regression» an. Excel spuckt dir dann eine komplizierte Tabelle aus. Für dich relevant ist das Feld «Significance F». Damit ist die «Signifikanz» oder der p-Wert gemeint.

Regression

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.79956685							
R Square	0.63930716							
Adjusted R Square	0.63831624							
Standard Error	23.1209632							
Observations	366							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	344893.704	344893.704	645.168896	1.308E-82			
Residual	364	194586.734	534.578939					
Total	365	539480.437						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	32.6879764	2.11056214	15.4878057	6.4307E-42	28.5375505	36.8384022	28.5375505	36.8384022
RAD [W/ms]	0.31275095	0.01231294	25.4001751	1.308E-82	0.28853751	0.33696439	0.28853751	0.33696439



Weisst du nicht, was diese Zahl bedeutet, oder ganz hast du ganz allgemein vergessen, was wir mit dem p-Wert meinen, dann klicke hier:

→ [Hilfe, was ist der p-Wert schon wieder?](#)

Wie kann ich numerische und kategoriale Faktoren erforschen?

Schwierigkeitsgrad:

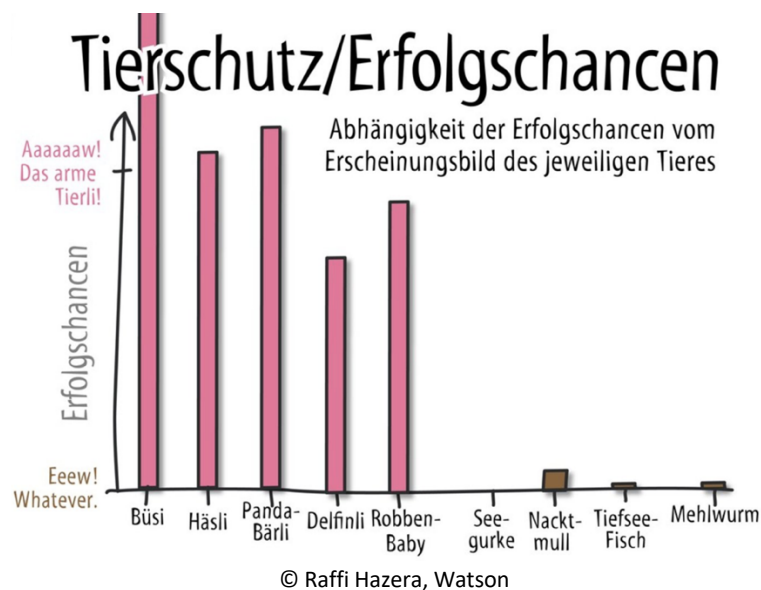


Ausgangslage

Du hast dir also beispielsweise die Frage gestellt, ob Schüler*innen, die rauchen einen höheren Blutdruck haben als Nichtraucher*innen. Deine Vermutung ist, dass dies der Fall ist, weil Rauchen die Blutgefäße verengt, was wiederum den Druck in den Gefäßen erhöht. Nun hast du einen numerischen Faktor (Blutdruck) und einen kategorischen Faktor (Raucher: Ja, Nein). Wie kannst du diese Faktoren nun auswerten?

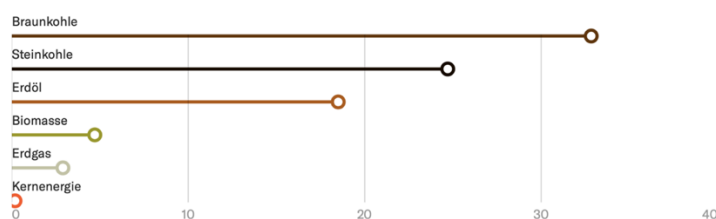
Grafisch

Es existieren verschiedene Möglichkeiten, wie du numerische und kategoriale Faktoren auswerten kannst. Zwei Beispiele haben wir dir unten aufgeführt.



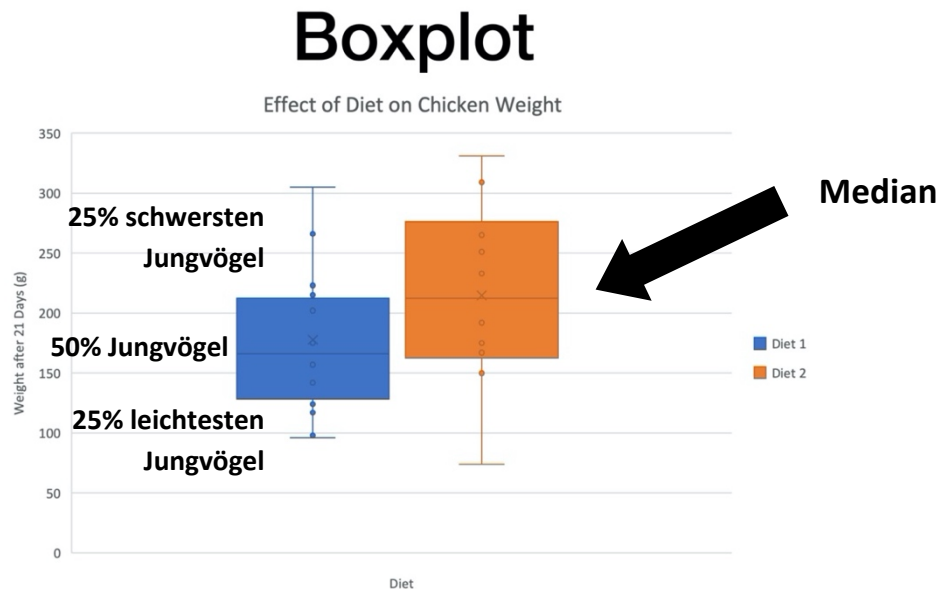
Kohle und Öl sind am gefährlichsten

Todesfälle pro Terawattstunde erzeugten Strom



Todesfälle bei der Produktion und als Folge der Umweltverschmutzung. Quelle: Ourworldindata.org.

Die Möglichkeit, die wir dir allerdings empfehlen möchten, ist der so genannte **Boxplot**. Er ist in den Wissenschaften weit verbreitet, relativ einfach zu interpretieren und kann von Excel erstellt werden.



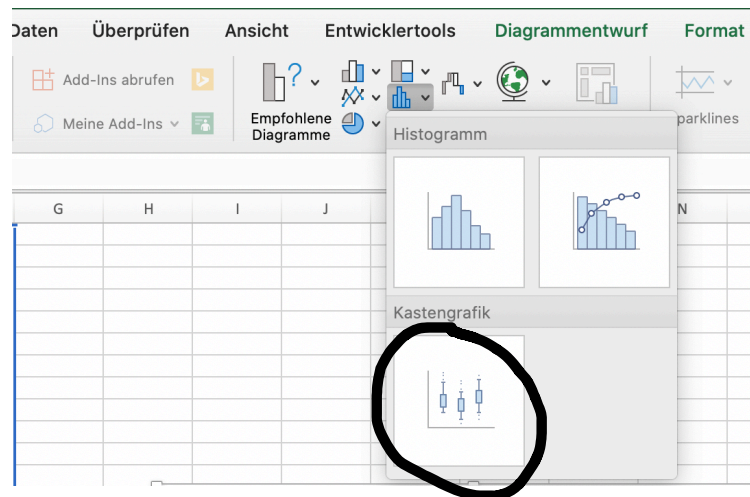
Quelle: <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/ChickWeight.html>

Im obigen Beispiel siehst du eine Studie, bei der Jungvögel unterschiedliche Nahrung erhielten (kategorischer Faktor) und man hat geschaut, wie sich dies auf ihr Gewicht ausgewirkt hat. Die Grafik zeigt den Stand nach 21 Tagen (Y-Achse).

- Der blaue Boxplot entspricht der Nahrung 1
- Der Orange Boxplot der Nahrung 2
- Der Strich entspricht dem Median
- Die grosse Box entspricht genau 50% der Jungvögel.
- Oberhalb der Box sind die 25% schwersten und unterhalb die 25 leichtesten

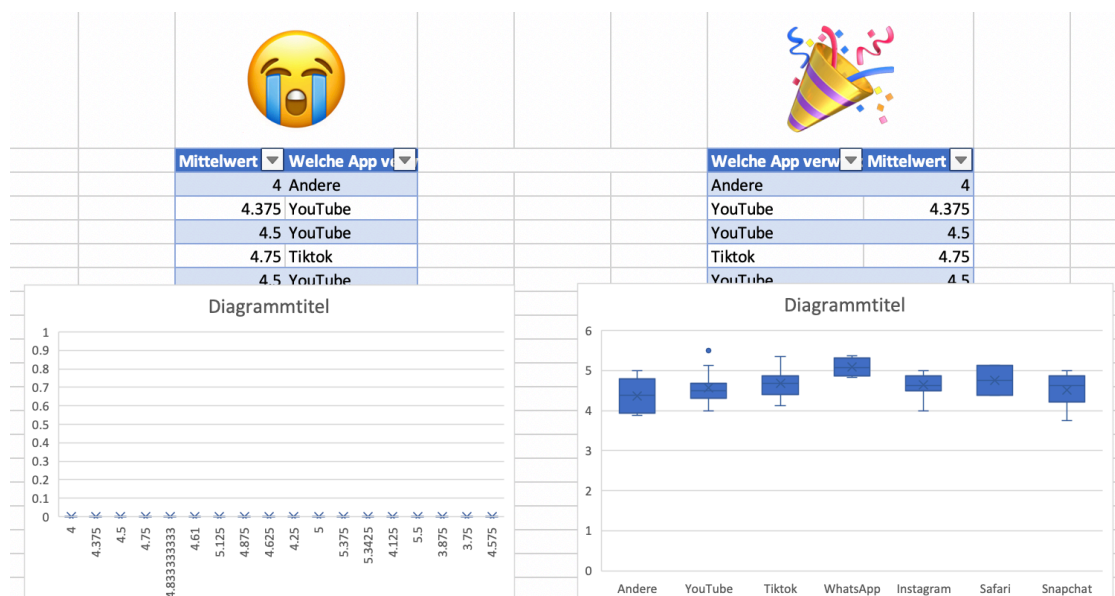
➔ [Hilfe, was ist schon wieder der Median?](#)

Du findest die «Box-Plots» unter «Einfügen/Statistik/Kastengrafik» (siehe Screenshot unten).



Damit die Kastengrafik (der Boxplot) so aussieht, wie im Beispiel mit den Jungvögeln musst du vielleicht noch einige Anpassungen vornehmen. Schau dir hierzu die Möglichkeiten unter «Diagrammelement hinzufügen» und «Schnelllayout» an.

Ganz wichtiger Hinweis: Excel kann Boxplots nur berechnen, wenn die linke Spalte der kategoriale Faktor und die rechte Spalte der numerische Faktor ist. Andersrum geht es leider nicht. (Siehe Bild unten) Achtet also darauf, dass ihr die Daten so sortiert, dass auf der linken Seite der kategoriale Faktor und auf der rechten die numerischen Werte stehen, bevor ihr den Boxplot erstellt.



Rechnerisch

r-Wert berechnen (Korrelationskoeffizient)

Der r-Wert lässt sich nicht berechnen.

Vorbereitung um Signifikanz (p-Wert) zu bestimmen

Die Signifikanz (der p-Wert) lässt sich bestimmen, allerdings muss zunächst eine Vorüberlegung durchgeführt werden. Dies soll an einem konkreten Beispiel aufgezeigt werden. In unserem Beispiel wurde die Bildschirmzeit (in Minuten) von acht Jugendlichen unterschiedlicher Klassen bestimmt.

- Faktor 1: Geschlecht (Kategorisch)
- Faktor 2: Bildschirmzeit in Minuten (Numerisch)
- Faktor 3: Klassen (Kategorisch)

Weiblich	389.00	3. Klasse
Weiblich	400.00	3. Klasse
Weiblich	420.00	4. Klasse
Weiblich	524.00	2. Klasse
Weiblich	540.00	3. Klasse
Männlich	91.00	2. Klasse
Männlich	120.00	4. Klasse
Männlich	120.00	2. Klasse
Männlich	140.00	3. Klasse
Männlich	140.00	4. Klasse

Nun soll in einem ersten Schritt berechnet werden, ob die Schüler*innen signifikant längere Bildschirmzeiten haben, als die Schüler. Dies geht indem der t-Test (siehe unten) mit den grünen und gelben Werten durchgeführt wird. Soll nun aber bestimmt werden, ob die 2., 3. oder 4. Klässler eine signifikant längere Bildschirmzeit haben, dann muss zunächst ausgewählt werden, welche miteinander verglichen werden sollen. Alle drei aufs Mal geht nicht. Wir vergleichen also bspw. zunächst die 2. Klassen und die 3. Klassen und sortieren unsere Tabelle nach der Klasse. Nun wenden wir den t-Test auf die zyanfarbigen und die rosafarbenen Werte an.

Weiblich	389.00	3. Klasse
Weiblich	400.00	3. Klasse
Weiblich	540.00	3. Klasse
Männlich	140.00	3. Klasse
Weiblich	524.00	2. Klasse
Männlich	91.00	2. Klasse
Männlich	120.00	2. Klasse

p-Wert berechnen (Korrelationskoeffizient)

Damit die Datenanalyse auch erscheint, muss ein «Ad-In» installiert sein. Findest du die Datenanalyse nicht unter «Daten» (ganz rechts), klicke unten:

→ [Hilfe, ich finde die Datenanalyse nicht auf meinem Mac?!?](#)

→ [Hilfe, ich finde die Datenanalyse nicht auf meinem PC?!?](#)

Hast du das Ad-In installiert, startest du die Datenanalyse und wählst führst du einen sogenannten t-Test durch. Excel zeigt dir Arten von t-Tests:

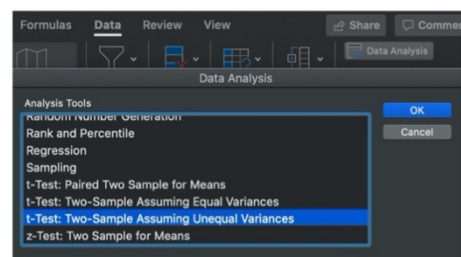
Welchen der drei t-Tests führe ich aus?

- **Paired:** wenn die beiden Stichproben abhängig voneinander sind (z.B. vorher/nachher)
- **Equal Variances:** falls die beiden Stichproben eine ähnliche Varianz aufweisen (praktisch nie der Fall)
- **Unequal Variances:** falls die beiden Stichproben eine unterschiedliche Varianz aufweisen (oft)

In der Regel wählen wir Option 3, also «t-Test: Two-Sampling Assuming Unequal Variances» an (siehe Screenshot unten). Excel spuckt dir dann eine komplizierte Tabelle aus. Für dich relevant sind die Felder, die mit P beschrieben sind.

t-Test

t-Test: Two-Sample Assuming Unequal Variances		
	WeightDiet1	WeightDiet2
Mean	177.75	214.7
Variance	3445.93333	6105.5667
Observations	16	10
Hypothesized Mean D	0	
df	15	
t Stat	-1.2857114	
P(T<=t) one-tail	0.1090194	
t Critical one-tail	1.75305036	
P(T<=t) two-tail	0.21803879	
t Critical two-tail	2.13144955	



Welchen P-Wert wähle ich?

Du siehst, dass dir Excel zwei P-Werte zeigt. Folgendes gilt:

- **One-tail:** wenn ich analysieren will, ob die eine Stichprobe *grösser* (oder *kleiner*) ist als die andere
- **Two-tail:** wenn ich analysieren will, ob die eine Stichprobe *anders* ist als die andere

In unserem Beispiel wäre der p-Wert also «0.218» und somit nicht signifikant, da er die 5% Grenze deutlich übersteigt.

Wie kann ich zwei kategoriale Faktoren erforschen?

Schwierigkeitsgrad:



Ausgangslage

Du hast dir also beispielsweise die Frage gestellt, ob Schüler*innen, die ein bestimmtes Sternzeichen haben (kategorischer Faktor) auch gehäuft einen bestimmten Persönlichkeitstyp besitzen (kategorischer Faktor).

🌟 Wir möchten dich darauf aufmerksam machen, dass dieser dritte Typ mit Abstand der schwierigste ist. Bestimmt nicht unmöglich, aber vielleicht eher für jene geeignet, die eine Herausforderung suchen.

Grafisch

Beobachtete Werte auflisten

Um die Berechnungen müssen ein paar Vorarbeiten geleistet werden: Angenommen, wir haben die folgenden Ergebnisse von 13 Personen aus der Untersuchung:

Persönlichkeit	Sternzeichen
Wachen	Wassermann
Wachen	Wassermann
Forscher	Steinbock
Wachen	Steinbock
Forscher	Steinbock
Diplomaten	Skorpion
Diplomaten	Krebs
Forscher	Krebs
Diplomaten	Krebs
Diplomaten	Jungfrau
Forscher	Jungfrau
Analysten	Jungfrau

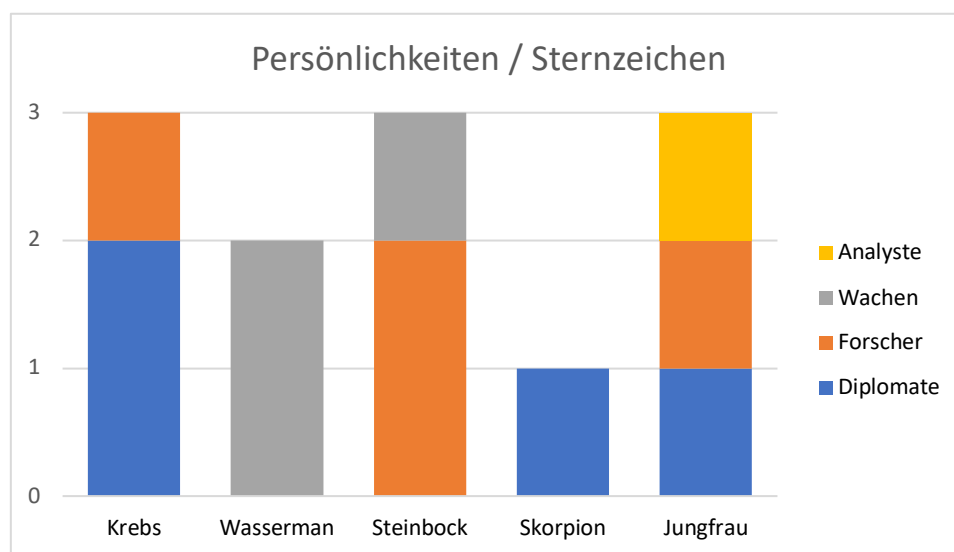
Jetzt musst du zunächst alle beobachteten Werte auflisten und gruppieren, damit du auf einen Blick sehen kannst, wie viele Persönlichkeiten mit welchem Sternzeichen beobachtet wurden. Das machst du wieder in Excel und könnte in etwa so aussehen:

Beobachtet	Krebs	Wasserman	Steinbock	Skorpion	Jungfrau	SUMME
Diplomate	2	0	0	1	1	4
Forscher	1	0	2	0	1	4
Wachen	0	2	1	0	0	3
Analyste	0	0	0	0	1	1
SUMME	3	2	3	1	3	

Die hellgelb eingefärbten Zellen entsprechen jeweils der Summe, da wir ja nicht von jedem Sternzeichen und Persönlichkeitstyp gleich viele beobachtet haben. Die «12» unten rechts zeigt die Summe aller Personen an, muss also der Gesamtzahl befragter Personen entsprechen, bei uns 12.

Grafik erstellen

Jetzt kannst du deine Werte anwählen (ohne die Summen) und daraus ein gestapeltes Säulendiagramm erstellen. Excel zeigt dir in diesem Diagramm schön, wie die Persönlichkeiten auf die verschiedenen Sternzeichen verteilt sind.



Rechnerisch

r-Wert berechnen

Einen r-Wert kann man hier nicht berechnen.

p-Wert berechnen

Zum Schluss stellen wir uns noch der grössten Herausforderung, indem wir nämlich den p-Wert (also die Signifikanz) berechnen. Hierzu müssen wir den so genannten Chi-Quadrat-Test durchführen. Wie geht das genau? Es ist auch keine Hexerei, aber es braucht etwas Vorarbeit.

Du hast nach wie vor deine «beobachteten Werte», wir brauchen nun zusätzlich auch noch die erwarteten Werte, das heisst, was würde man eigentlich erwarten?

Wir möchten dir nun zeigen, wie du die erwarteten Werte aus den beobachteten Werten berechnen kannst. Sollte dir ein Schritt unklar sein, kannst du dir jederzeit das nachfolgende Lernvideo anschauen:

➔ <https://www.youtube.com/watch?v=M0K4av6Wnac>

Du erhältst die korrekten erwarteten Zellenwerte, indem du jeweils die Spaltensumme mit der Zeilensumme multiplizierst und durch das Total (hier 12) teilst.

Erwartet	Krebs	Wasserman	Steinbock	Skorpion	Jungfrau	SUMME
Diplomate	1	0.666666667	1	0.333333333	0.333333333	4
Forscher	1	0.666666667	1	0.333333333	0.333333333	4
Wachen	0.75	0.5	0.75	0.25	0.25	3
Analyste	0.25	0.166666667	0.25	0.083333333	0.083333333	1
SUMME	3	2	3	1	3	12

Beispiel: Die erwartete Anzahl Analysten, die das Sternzeichen «Wassermann» haben ergibt sich aus der Summe aller Wassermänner (hier 2) multipliziert mit der Summe aller Analysten (hier 1) dividiert durch die Gesamtzahl Personen (hier 12), d.h. es ergibt sich die Rechnung:

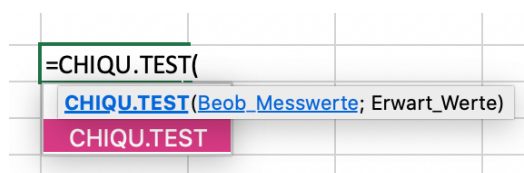
$$2 \cdot 1 / 12 = 0.16667$$

CHI-Test ausführen

Sobald wir alle beobachteten und alle erwarteten Werte kennen, können wir den Chi-Quadrat-Test starten. Dies ist mit der folgenden Formel möglich:

=CHISQ.TEST(...;...)

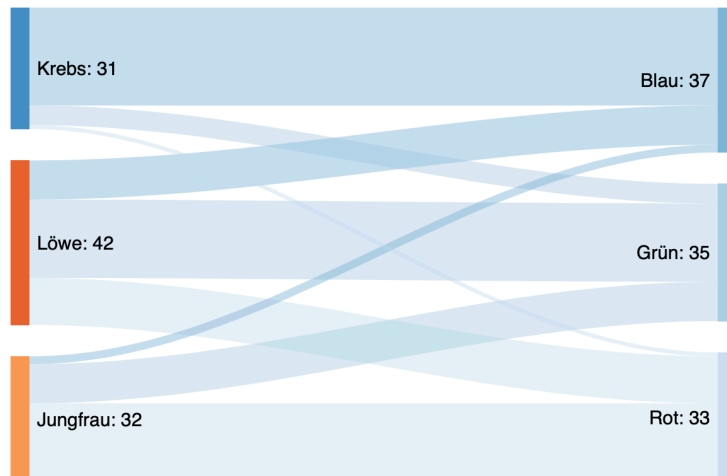
Excel verlangt nun von uns die beobachteten und die erwarteten Messwerte. Wir wählen diese an:



Excel gibt dann einen Wert aus, der dem p-Wert entspricht, in unserem Fall also 0.01. Dieser Wert liegt deutlich unter der 5% Schwelle, wir haben also ein signifikantes Ergebnis.

Flussdiagramme

Eine hübsche Möglichkeit um zwei kategoriale Faktoren grafisch darzustellen, sind so genannte Flussdiagramme (Sankeys). Nachfolgend siehst du so ein Diagramm.



Sie sind einfach zu interpretieren. In unserem Beispiel wird gezeigt, wie viele Leute eines bestimmten Sternzeichens welche Farbe als Lieblingsfarbe gewählt haben. Leider kann Excel diese nicht darstellen, es gibt aber eine kostenlose Webseite, mithilfe der dies gemacht werden kann.

[Tool zum Erstellen von Flussdiagrammen: Link](#)

Das Tool funktioniert so, dass du die Werte, die du hast, einträgst und jeweils mit «Preview» überprüfst, wie die Daten in ein Flussdiagramm übersetzt werden. Du kannst die Farben, die Breite der Flussdiagramme, usw. anpassen. Am Schluss kannst du sie als PNG exportieren und deiner Studie hinzufügen.

Tools / Hilfe

Was ist schon wieder der r-Wert?

r-Wert: Korrelationskoeffizient

➔ Besteht ein statistischer Zusammenhang zwischen den Faktoren?

Der r-Wert kann zwischen -1 und +1 betragen. Je näher der Wert an 0 liegt, desto weniger stark korrelieren die Faktoren, desto schwächer ist also der statistische Zusammenhang. Folgende Konventionen gelten:

Sozialwissenschaften:¹

$r = 0.1$ für eine geringe Korrelation

$r = 0.3$ für eine mittlere Korrelation

$r = 0.5$ für eine hohe Korrelation

Naturwissenschaften

$r = ?$ für eine geringe Korrelation

$r = ?$ für eine mittlere Korrelation

$r = ?$ für eine hohe Korrelation

Was ist schon wieder der p-Wert?

Kurz und knapp: Der **p-Wert** zeigt die Signifikanz an, wobei man die Signifikanz wie folgt definieren kann:

➔ Wie wahrscheinlich ist es, dass das Resultat eine Folge von Zufall ist?

Am besten schaust du dir das entsprechende Kapitel im «Skript» noch einmal gründlich an.

Was ist schon wieder der Median?

Am besten schaust du dir das entsprechende Kapitel im «Skript» noch einmal gründlich an.

¹ Cohen (1988)

Schöne Datenquellen zum Experimentieren

<https://opendata.swiss>

<https://www.bfs.admin.ch>

Bei abenteuerlichen Datenformaten gibt's converter:

<https://jsonformatter.curiousconcept.com>

<https://www.convertjson.com>

Wo finde ich das Add-In in Excel für Microsoft? 🤔

Laden und Aktivieren der Analyse-Funktionen

1. Klicken Sie auf der Registerkarte **Datei** auf **Optionen**, und klicken Sie dann auf die Kategorie **Add-Ins**.

Wenn Sie Excel 2007 verwenden, klicken Sie auf die **Microsoft Office-Schaltfläche** , und klicken Sie dann auf **Excel-Optionen**.

2. Wählen Sie im Feld **Verwalten** die Option **Excel-Add-Ins** aus, und klicken Sie dann auf **Gehe zu**.

Wenn Sie Excel für Mac verwenden, wechseln Sie im Menü "Ablage" zu **Extras** > **Excel Add-Ins**.

3. Aktivieren Sie im Feld **Add-Ins** das Kontrollkästchen **Analyse-Funktionen**, und klicken Sie auf **OK**.

- Wenn **Analyse-Funktionen** nicht im Feld **Verfügbare Add-Ins** angezeigt wird, klicken Sie auf **Durchsuchen**, um danach zu suchen.
- Wenn angezeigt wird, dass die Analyse-Funktionen derzeit nicht auf dem Computer installiert sind, klicken Sie auf **Ja**, um das Add-In zu installieren.

in Excel für Mac? 🤔

➔ In diesem Video: <https://www.youtube.com/watch?v=BWZeBtIQItY>

in Excel online? 😞

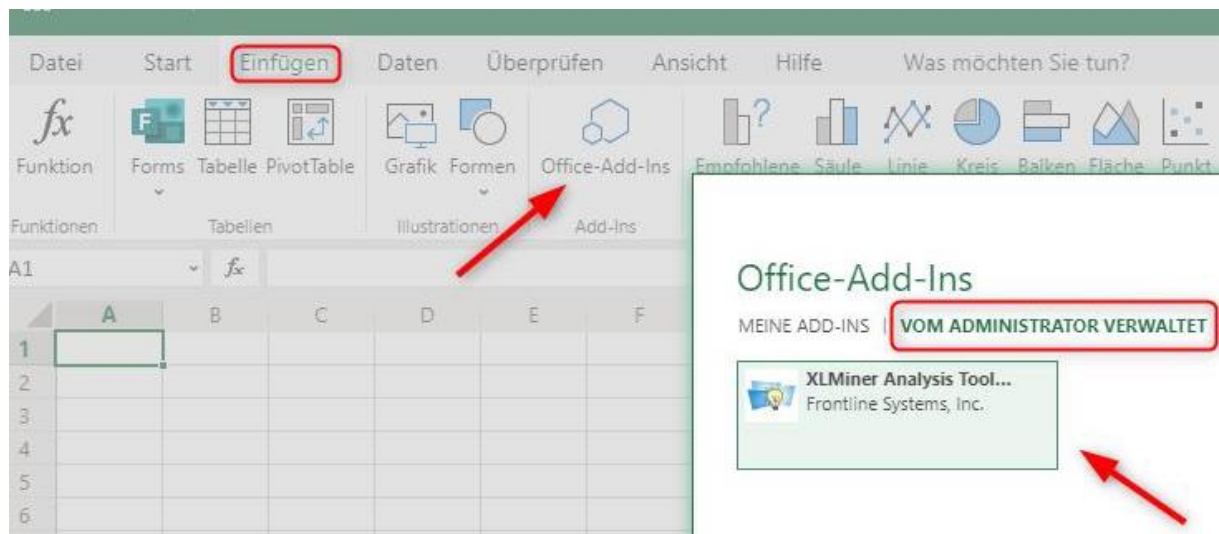


Diagramm-Typen

<https://support.office.com/de-de/article/verfuegbare-diagrammtypen-in-office-a6187218-807e-4103-9e0a-27cdb19afb90>

Pivot-Tabellen

<https://support.office.com/de-de/article/erstellen-einer-pivottable-zum-analysieren-von-arbeitsblatt-daten-a9a84538-bfe9-40a9-a8e9-f99134456576>

Weiterführende Quellen

<https://de.statista.com/statistik/lexikon/definition/123/skalenniveau>

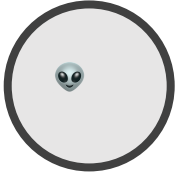
<https://de.wikipedia.org/wiki/T-Test>

Viele Darstellungen von Daten der ganzen Welt: <https://ourworldindata.org>

Ist dir Excel zu Mainstream? Versuchs mal mit <https://d3js.org>

Noch weiter weiterführende Statistik

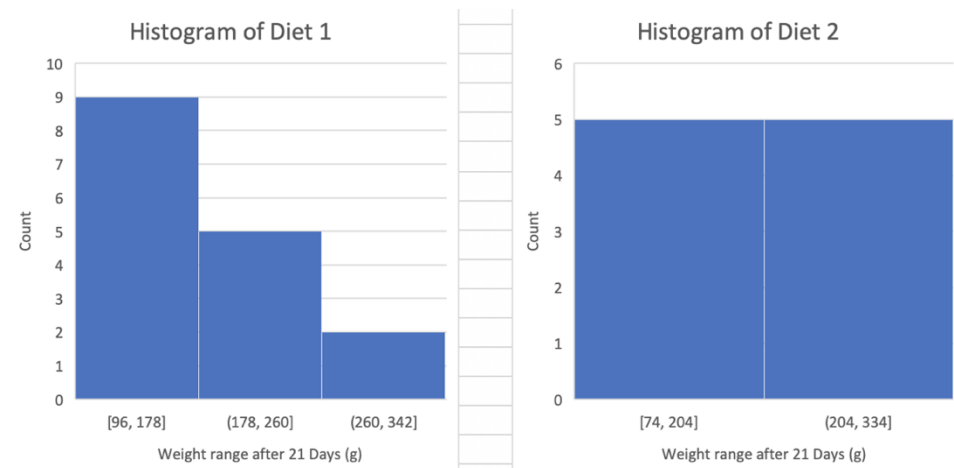
Schwierigkeitsgrad:



Was ist Varianz?

Was ist Varianz?

Die Varianz lässt sich graphisch mit «Histogrammen» darstellen; im folgenden Beispiel sind die Werte sehr unterschiedlich verteilt, weshalb “unequal variances” gewählt wurde:



Hilfe, ich will ganz viele Kategorien miteinander vergleichen, z.B. Sternzeichen 🤔

Dann kann man entweder auf einen [Bonferroni-korrigierten Kruskal-Wallis-Test](#) zurückgreifen, **oder aber** man vergleicht die verschiedenen Kategorien erstmal visuell mit Boxplots und wählt dann zwei-drei Highlights aus, die man im Anschluss jeweils mit einem klassischen t-Test vergleichen kann.

2 kategoriale Variablen & Mosaikplots

