

BSc PHY – Angewandte Statistik

WS1 Deskriptive Statistik

Patric Eichelberger & Aglaja Busch
aF&E Physiotherapie



patric.eichelberger@bfh.ch | aglaja.busch@bfh.ch



Moodlekurs

9. Mai 2025



M&Ms Stichproben

Fülle die Befragung M&Ms Stichproben auf Moodle aus.

ID	Gewicht	Handspanne
P01	70.0	20.0
P02	66.8	14.0
P03	59.5	15.0
P04	58.0	21.5
P05	64.0	20.0

- ▶ Handelt es sich um eine Population oder eine Stichprobe? Begründe!
Stichprobe aus der Population der Studierenden der PHY an der BFH.
- ▶ Wo in der Tabelle befinden sich die Beobachtungseinheiten? In den Zeilen
- ▶ Wo in der Tabelle befinden sich die Merkmale? In den Spalten
- ▶ Wo in der Tabelle befinden sich die Merkmalsausprägungen? In den Zellen

Skalenniveaus

Ordne zu und begründe.

Merkmal	Qualitativ Kategorien		Quantitativ Kontinuierlich Metrisch	
	Nominal	Ordinal	Intervall	Rational
Fusslänge				x
Schuhgrösse		x		
Temperatur in °C			x	
Nationalität	x			
Alter				x
Kaffees pro Tag		x		

Kennzahlen, Boxplot, Histogramm

In der Datei `MMs.csv` befinden sich die Daten der M&Ms Datenerhebung. Wir betrachten in dieser Aufgabe das M&Ms Gewicht.

- ▶ Importiere die Daten in jamovi.
- ▶ Wie gross ist der Stichprobenumfang?
- ▶ Welcher Wert liegt auf der 25%-Perzentile? Welcher auf der 75%-Perzentile?
- ▶ Wie gross ist der Interquartilsabstand?
- ▶ Wie viele Werte liegen innerhalb des Interquartilsabstands?
- ▶ Wie viele Werte liegen über der 90%-Perzentile? (Tipp: Erstelle einen Filter)
- ▶ Zeichne einen Boxplot für das M&Ms Gewicht von Hand.
- ▶ Erstelle mit jamovi ein Histogramm für das M&Ms Gewicht und interpretiere. Was stellst du fest im Vergleich mit dem Boxplot (Stichwort Symmetrie)?
- ▶ Untersuche den Einfluss der Farbe auf das M&Ms Gewicht.

Robustheit

In der Exceldatei **Alter.xlsx** findest du die Variablen Alter1, Alter2 und Alter3. Die Variable Alter2 enthält einen besonders hohen Wert, die Variable Alter3 einen besonders tiefen Wert (sogenannte Ausreisser).

- ▶ Berechne den Mittelwert und den Median der drei Variablen.
 - ▶ Welche der beiden Lagekennzahlen wird durch die Ausreisser mehr beeinflusst und wieso? **Der Mittelwert. Weil er auf den Zahlenwerten und nicht den Rängen basiert.**
- ▶ Berechne die Standardabweichung und den Interquartilsabstand der drei Variablen.
 - ▶ Welche der beiden Streuungskennzahlen wird durch die Ausreisser mehr beeinflusst und wieso? **Die Standardabweichung. Weil sie auf den Zahlenwerten und nicht den Rängen basiert.**
- ▶ Schaue dir die Variable Alter4 im Boxplot an.
 - ▶ Wieso hat der Boxplot unten keinen Whisker? **Weil das Minimum dem 1. Quartil entspricht.**
 - ▶ Wieso ist der Wert von 50 kein Ausreisser? **Weil er innerhalb $Q_3 + 1.5 \cdot IQR$ liegt.**

Häufigkeitstabellen

In der Datei `MMs.csv` befinden sich die Daten der M&Ms Datenerhebung. Wir betrachten in dieser Aufgabe die Farben der M&Ms. Tipp für diese Aufgabe: Slide Nr. 12 aus dem Screencast.

- ▶ Erstelle eine Tabelle mit den absoluten und relativen Häufigkeiten der M&Ms Farben.

Weitere Übungen auf RPubS

Auf **RPubs** befinden sich weitere Übungen zum Thema, welche uns von unserem Kollegen Lukas Stammer, Dozent im BSc PHY am Standort Basel, zur Verfügung gestellt wurden.

- ▶ <https://rpubs.com/stammer/861165>
- ▶ <https://rpubs.com/stammer/861136> (ausser Übungen 5 und 6)