

### 01: Klassik auf dem Klo

Ein Team von Psychologinnen möchte die Hypothese überprüfen, dass Besucher\*innen eines großen Einkaufszentrums die Toiletten als sauberer empfinden, wenn in den Räumlichkeiten klassische Musik im Hintergrund gespielt wird (obwohl nicht häufiger geputzt wird). Es werden zwei Toilettenräume mit identischer Ausstattung und vergleichbarer Lage gewählt. In einem der Räume wird eine Woche lang klassische Musik gespielt, im anderen gibt es keine Hintergrundmusik. Beim Verlassen der Toilettenräume werden die Kund\*innen gefragt, wie zufrieden sie heute mit der Sauberkeit der sanitären Anlagen waren.

Das Rating erfolgt auf folgender Skala (wobei später für die Auswertung der rote Smiley mit einer 1 kodiert wird, dann weiter aufsteigend bis 5 für den grünen Smiley):



Am Ende der Woche werden die Daten für die beiden Räume verglichen.

- Welcher Test sollte verwendet werden, um die Hypothese zu prüfen?
- Wie lautet die statistische Hypothese?
- Zusatzfrage: Was für Probleme (die sich auch auf die Wahl des statistischen Tests auswirken) könnten bei diesem Versuchsaufbau auftreten?

### 02: Studium und Allgemeinwissen

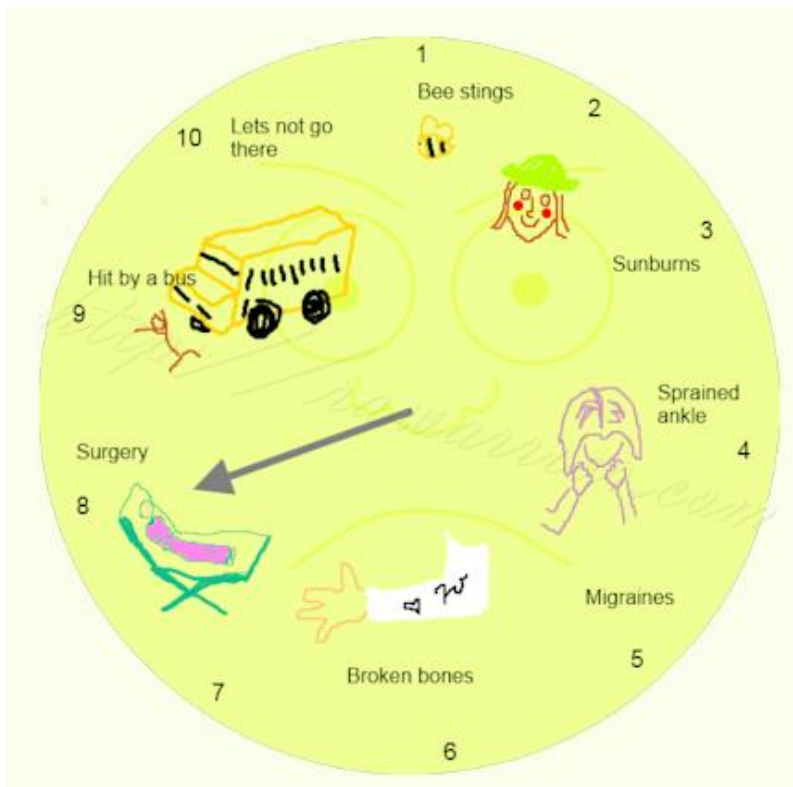
Ein Team von Forscherinnen möchte prüfen, ob Studierende am Ende des Studiums ein besseres oder schlechteres Allgemeinwissen aufweisen als zu Beginn des Studiums. Die Forscherinnen rekrutieren 50 Studierende im ersten Semester und legen ihnen einen Wissenstest mit 100 Fragen aus vielen verschiedenen Bereichen vor. Am Ende ihres Bachelorstudiums werden dieselben 50 Studierenden erneut demselben Test unterzogen. Gehen Sie davon aus, dass die Testergebnisse auf Intervallskalenniveau vorliegen.

- Welcher Test sollte verwendet werden, um die Hypothese zu prüfen?
- Wie lautet die statistische Hypothese?
- Zusatzfrage: Nehmen Sie an, die Studie zeigt, dass die Studierenden am Ende des Studiums besser im Test abschneiden als zu Beginn. Welche Aussage ließe sich aufgrund dieser Ergebnisse treffen und welche nicht? Was könnte man am Versuchsaufbau ändern, um aussagekräftigere Ergebnisse zu erzielen?

### 03: Medikamentenstudie

In einer Studie soll die Wirksamkeit eines neuen Schmerzmittels überprüft werden. Die Hypothese ist, dass das Medikament wirkt. 60 Patient\*innen, die mit akuten Schmerzen in der Notaufnahme

eines Krankenhauses vorstellig werden, werden bei der Aufnahme gebeten, die Intensität ihres Schmerzes auf folgender Skala einzuschätzen:



Anschließend erhalten sie das Schmerzmittel. Nach einer Stunde werden sie erneut gebeten, die Intensität ihres Schmerzes anzugeben.

- Welcher statistische Test sollte verwendet werden?
- Wie lautet die statistische Hypothese?
- Zusatzfrage 1: Ist das Studiendesign geeignet, um die Hypothese zu prüfen? Wenn nein, warum nicht?
- Zusatzfrage 2: Angenommen, es lässt sich in dieser Studie keine Wirksamkeit des neuen Medikaments nachweisen. Was könnte dafür sprechen, es in einem anderen Kontext dennoch weiter zu prüfen?

#### 04: Netflix all day every day – all year?

Marktforscherinnen wollen die Hypothese prüfen, dass es einen Zusammenhang zwischen der Jahreszeit und der auf Netflix & Co verbrachten Zeit gibt. Da die meisten Streaming-Services keine Daten zugänglich machen, führen sie eine eigene Studie durch. Zu einem Erhebungszeitpunkt im Winter, einem Erhebungszeitpunkt im Sommer, einem im Frühling und einem im Herbst bitten sie eine Stichprobe von Versuchspersonen (unterschiedliche Personen zu allen Messzeitpunkten), die Zeit einzuschätzen, die sie derzeit pro Woche mit dem Schauen von Filmen und Serien auf Streaming-Plattformen verbringen. Die Angaben werden klassifiziert in: „bis zu zwei Stunden“ und „mehr als zwei Stunden“.

- Welcher Test ist geeignet, um die Hypothese zu prüfen?
- Wie lautet die statistische Hypothese?

- Wie viele Personen sollten insgesamt befragt werden, wenn die Forscherinnen von einem mittleren Effekt ausgehen und 80% Power bei  $\alpha = .05$  erreichen wollen?

### 05: Frustessen

Psychologinnen möchten untersuchen, ob „Frustessen“ die Laune hebt. Sie führen eine Studie durch, in der alle Teilnehmer\*innen einen Allgemeinwissenstest bearbeiten. Anschließend erhalten sie von einem schauspielerisch geschulten Versuchsleiter die schockierte (falsche) Rückmeldung, dass sie im Test derart schlecht abgeschnitten hätten, dass man ihre Daten leider nicht verwenden könne. Sie werden aber gebeten, nach einer kurzen Pause dennoch einen Fragebogen für ein anderes Experiment auszufüllen, für das sie gerade noch geeignet seien. Eine Gruppe darf in der anschließenden Pause drei Tafeln Schokolade pro Person essen, eine Gruppe darf einen kleinen Schokoriegel pro Person essen und die dritte Gruppe gar nichts. Nach der Snackpause wird ein Fragebogen ausgefüllt, der die Stimmung auf Intervallskalenniveau misst (höhere Werte entsprechen besserer Stimmung).

Die Psychologinnen vermuten, dass die Gruppe, die gar nichts essen durfte, am schlechtesten gelaunt sein wird. Sie vermuten zudem, dass die Stimmung bei einem Schokoriegel besser wird, aber bei Verzehr von drei Tafeln Schokolade wieder abfällt.

- Welcher Test ist geeignet, um die Hypothese zu prüfen?
- Wie lautet die zentrale statistische Hypothese?
- Was muss zusätzlich getestet werden?

### 06: Gruppengröße und Konzentrationsfähigkeit bei Schulkindern

In einer Studie soll die Hypothese geprüft werden, dass die Konzentrationsfähigkeit von Schulkindern höher ist, wenn sie in kleinen im Vergleich zu größeren Gruppen am Tisch zusammensitzen. Zur Prüfung dieser Hypothese wurden die Kinder einer Schulklasse ( $N = 30$ ) randomisiert in zwei Gruppen aufgeteilt ( $n_1 = n_2 = 15$ ). Die Kinder der ersten Gruppe durften in drei Fünfer-Gruppen zusammensitzen, die Kinder der zweiten Gruppe wurden in fünf Dreier-Gruppen eingeteilt. Allen Kindern wurde ein Konzentrationstest vorgelegt, der die Konzentrationsfähigkeit auf Intervallskalenniveau misst.

- Welcher Test sollte verwendet werden, um die Hypothese zu prüfen?
- Formulieren Sie die statistische Hypothese.
- Zusatzfrage: Wie groß müsste der Effekt der Gruppengröße auf die Konzentrationsfähigkeit mindestens sein, damit er mit der gegebenen Stichprobengröße, einem alpha-Fehlerniveau von 0.05 und einer Teststärke von 0.80 entdeckt werden kann?

## 07: Länger leben durch Kunst & Kultur?

Folgende Studie machte Schlagzeilen in den Medien:

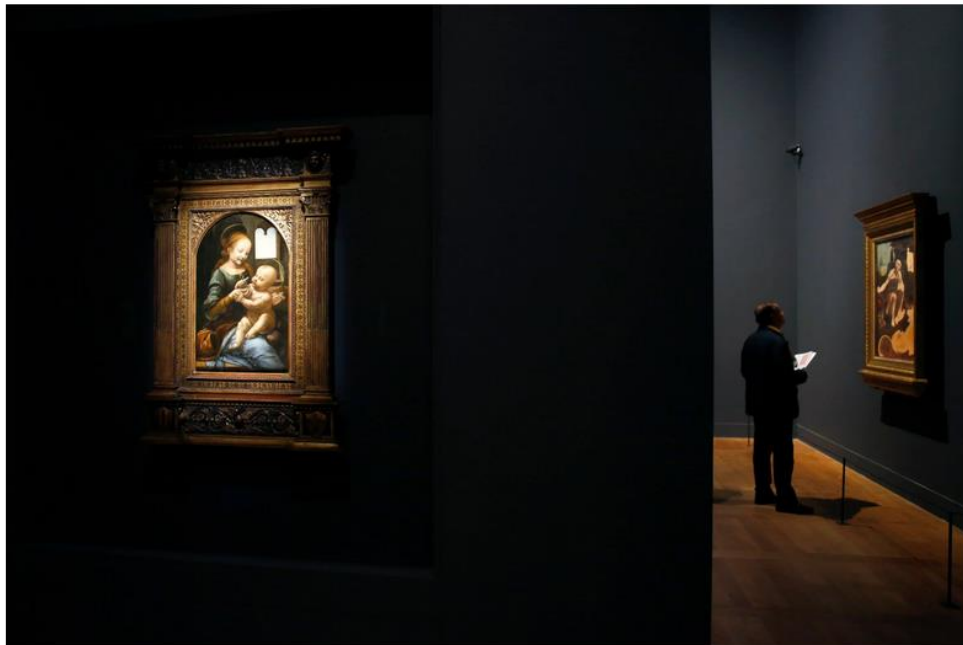
---

*The New York Times*

---

### *Another Benefit to Going to Museums? You May Live Longer*

Researchers in Britain found that people who go to museums, the theater and the opera were less likely to die in the study period than those who didn't.



Leonardo da Vinci's "The Virgin and Child" at the Louvre in Paris. A British study found that people who engaged with the arts had a lower likelihood of dying during the study period than those who did not. Thibault Camus/Associated Press

Nehmen Sie an, die Forscher\*innen wären bei der Erhebung folgendermaßen vorgegangen: 1000 Menschen gaben über einen Zeitraum von 5 Jahren hinweg an, wie häufig sie Museen, Theater oder Oper besuchten. Am Ende des festgelegten Zeitraums wurden die Personen in folgende zwei Kategorien eingeordnet: „keine oder wenige Besuche“ und „einige oder viele Besuche“. Zudem wurde verzeichnet, ob die Person noch lebte oder verstorben war.

Es soll die Hypothese geprüft werden, dass der Anteil der noch lebenden Personen in der Gruppe mit einigen oder vielen Besuchen höher ist als in der Gruppe mit keinen oder wenigen Besuchen.

- Welcher statistische Test ist geeignet, um diese Hypothese zu prüfen?
- Zusatzfrage: Welche Alternativerklärungen könnte es für den Befund geben?

## 08: Persönlichkeit und Social Media

Ein Team von Forscherinnen möchte untersuchen, ob es einen Zusammenhang zwischen Extrovertiertheit und der Aktivität in sozialen Medien gibt. In der geplanten Studie sollen Versuchspersonen einen Fragebogen zu Extrovertiertheit ausfüllen (Intervallskalenniveau, höhere Werte bedeuten stärkere Extrovertiertheit) und Angaben darüber machen, wie viele Freund\*innen sie auf Facebook haben und wie viele Nachrichten sie dort am Tag verschicken. Die Forscherinnen gehen davon aus, dass stärker extrovertierte Personen tendenziell mehr Facebook-Freund\*innen haben. Sie gehen zusätzlich davon aus, dass stärker introvertierte Personen zwar eine geringere Anzahl an Kontakten haben, aber mit diesen intensiver interagieren (und dadurch insgesamt mehr Nachrichten am Tag verschicken als extrovertiertere Personen).

- Welche Tests sind geeignet, um die Hypothesen zu testen?
- Wie lauten die statistischen Hypothesen?
- Wie viele Versuchspersonen müssen getestet werden, wenn für beide Hypothesen von einem mittleren Effekt ausgegangen wird ( $\alpha=.05$ , Teststärke = .80)?