



Populationsbeschreibung

Thema 03

Wozu Hypothesen?

Drei Arten epistemologischer Ziele von Studien

Beschreibung

Vorhersage

Erklärung

Beispiele für populationsbeschreibende Fragen

Beschreibung

- ▶ Wie viel Geld hat ein Student in Deutschland aktuell (typischerweise) zur Verfügung im Monat?
- ▶ Haben weibliche Studentinnen mehr oder weniger Geld zur Verfügung als männliche Studentis?
- ▶ Wie unterscheidet sich das Budget der Studentis nach Bundesland und nach Studienrichtung?

Beschreibung mit 1 oder 2 (oder mehr) Variablen

1 Variable

- ▶ Wie viel Geld hat ein Studenti in Deutschland aktuell (typischerweise) zur Verfügung im Monat?
- ▶ Wie viele Parties besucht ei Studenti im Schnitt pro Semester (in Deutschland, aktuell, in einem wirtschaftlichen Studiengang)?
- ▶ Wie groß ist der Anteil an Arbeitnehmern, die im Home-Office arbeiten (mind. 1 Tag pro Woche, aktuell, in De)?
- ▶ Wie groß ist der Anteil an Studentis, die heimlich nachts Statistik-Bücher lesen (mind. 30 Minuten, mind. 3 Nächte pro Woche, aktuell, in Deutschland ...)?

2 oder mehr Variablen

- ▶ Haben weibliche Studentinnen mehr oder weniger Geld zur Verfügung als männliche Studentis?
- ▶ Wie unterscheidet sich das Budget der Studentis nach Bundesland und nach Studienrichtung?

Notation für Forschungsfragen

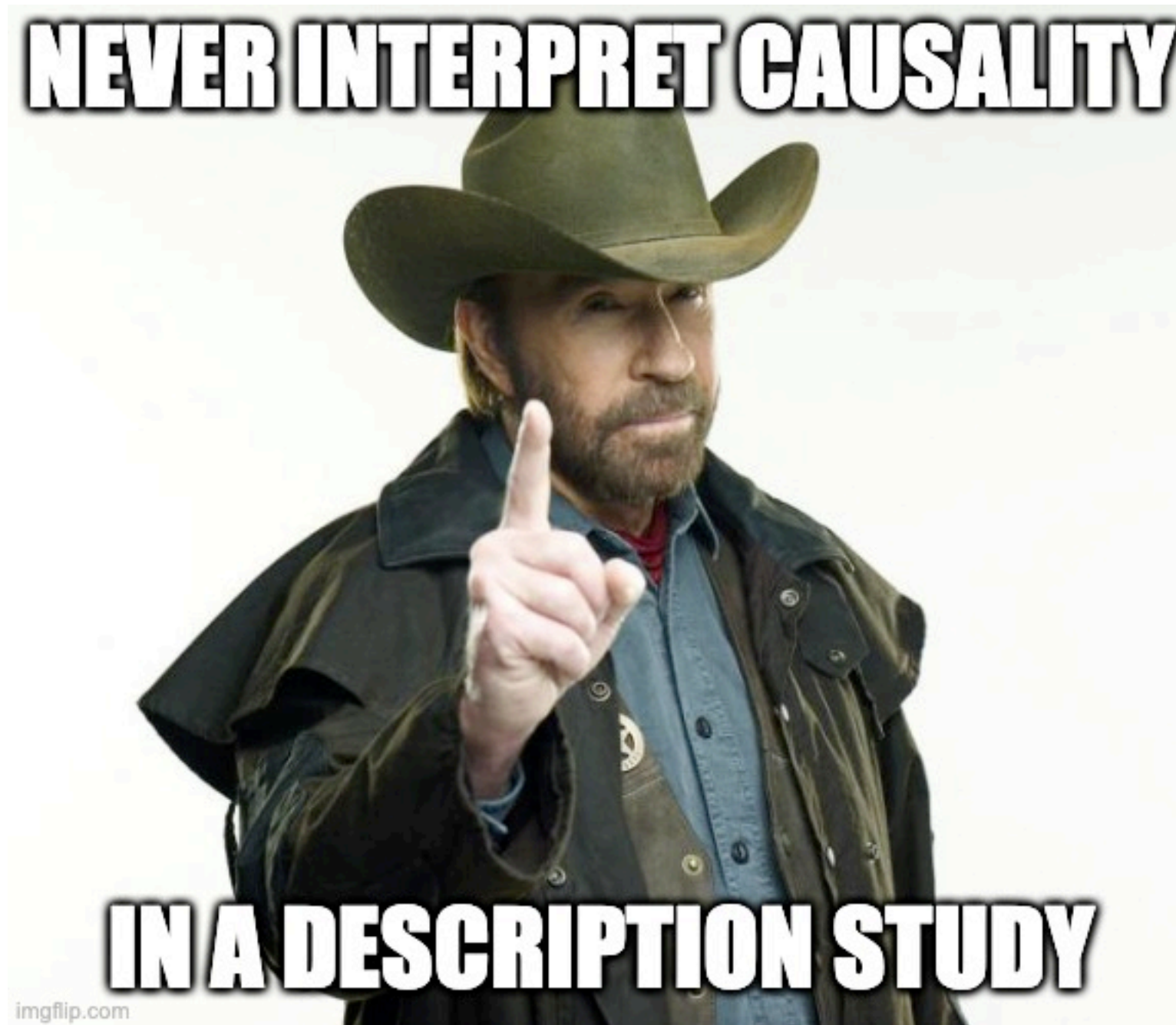
1 AV, keine UV

$$y \sim 1$$

1 AV, 1 oder mehr UV

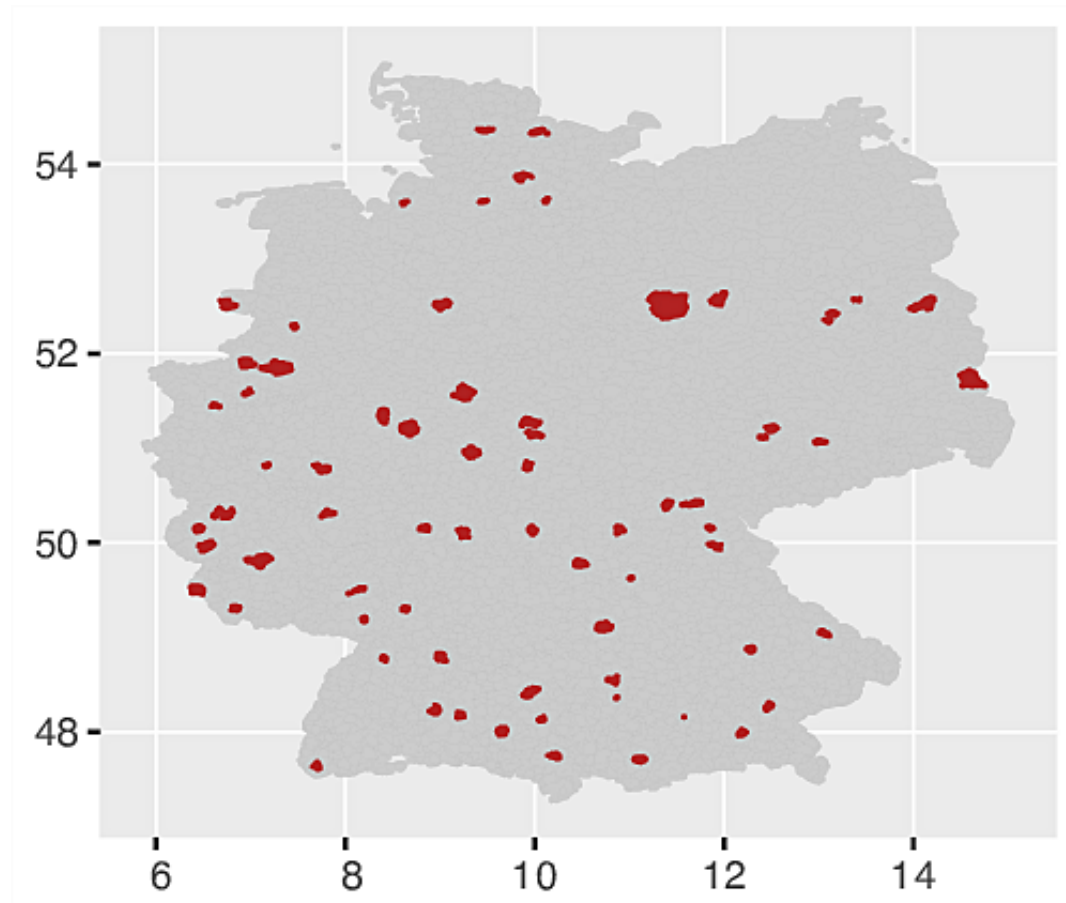
$$y \sim x_1$$
$$y \sim x_1 + x_2$$

Kausalinterpretation sind nur begründet möglich

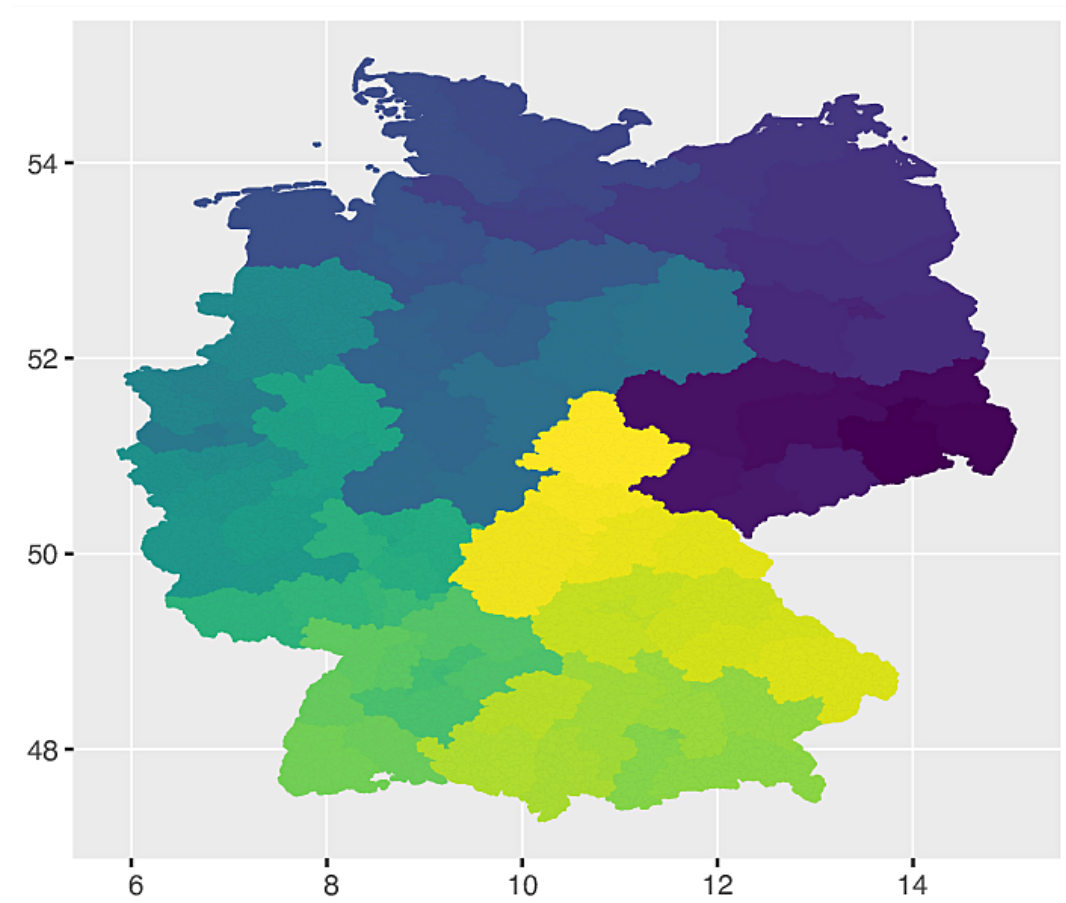


Stichprobe vs. Grundgesamtheit

Stichprobe deutscher Gemeinden

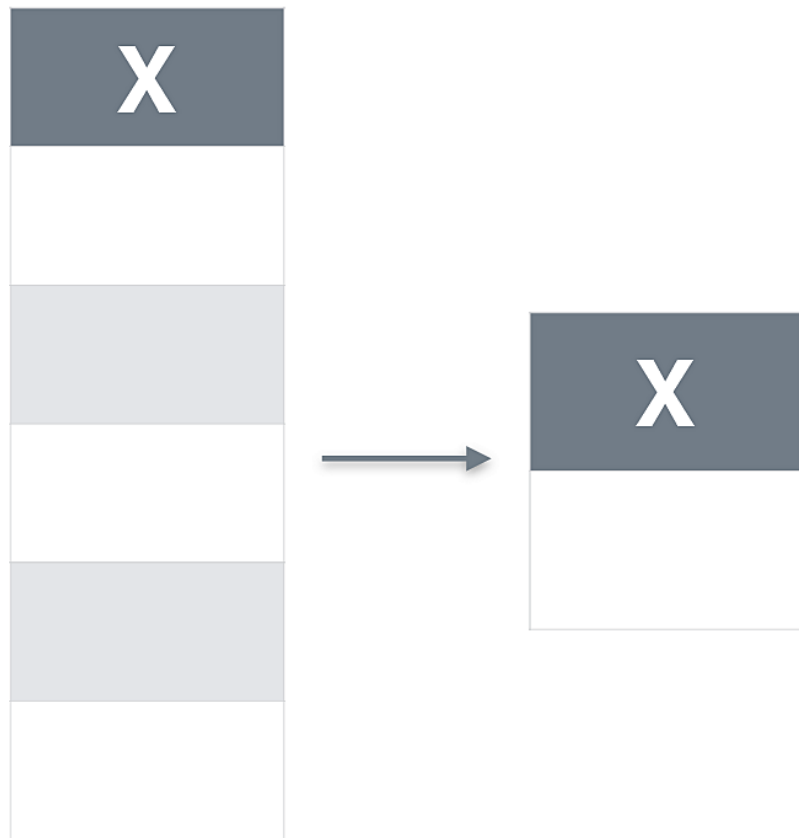


Population deutscher Gemeinden

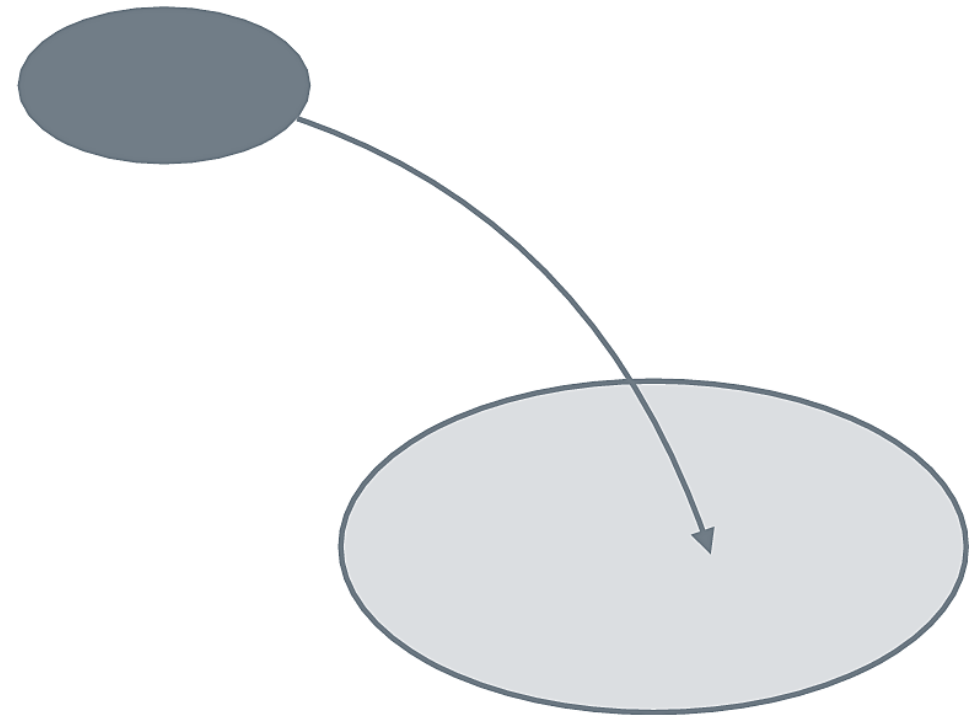


Deskriptivstatistik vs. Inferenzstatistik

Deskriptivstatistik

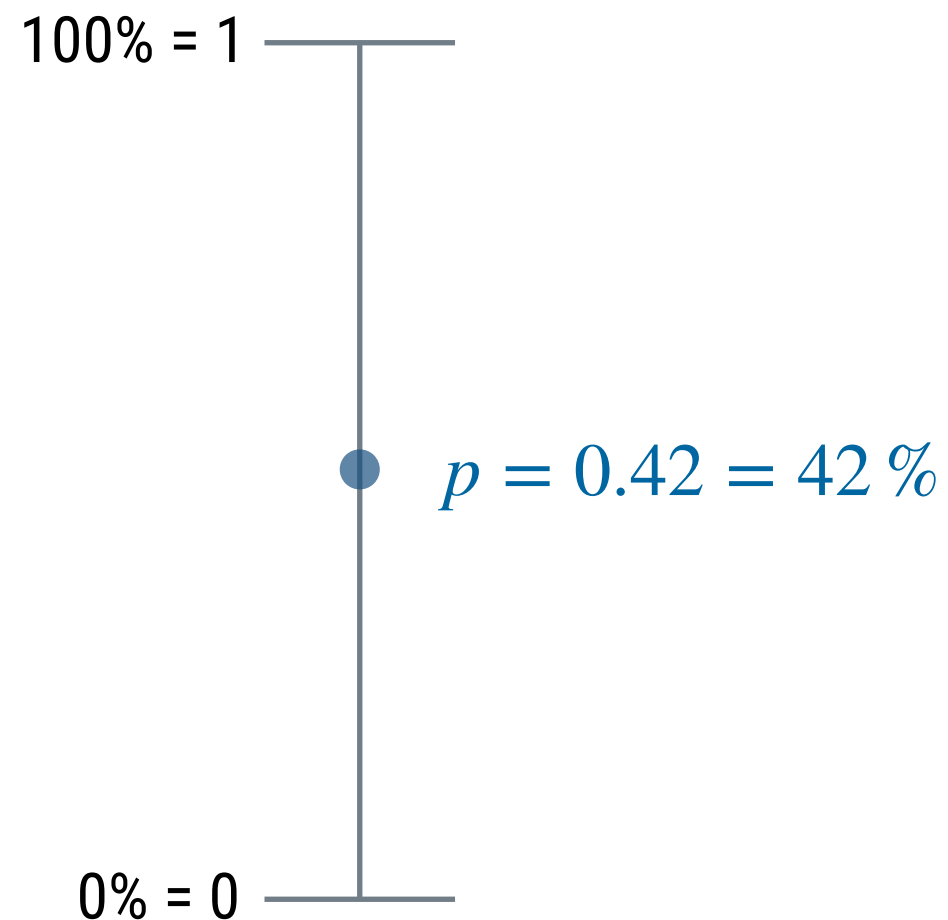


Inferenzstatistik



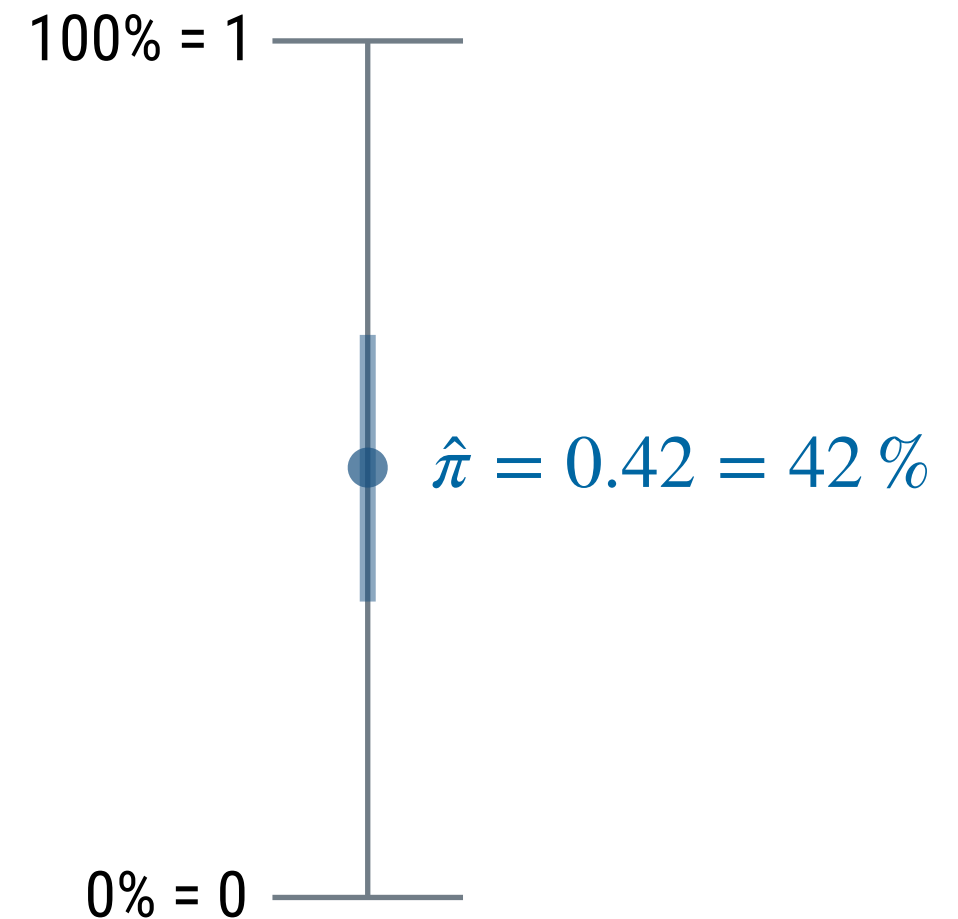
Anteil der heimlichen Statistik-Fans?

Deskriptive Statistik



In der Stichprobe ($n=100$) fanden sich 42 heimliche Statistik-Fans.

Inferenzstatistik



Wir schätzen den Anteil in der Population auf 0.32-0.52 (95%-PI).

Eine gute Stichprobe anstelle einer Vollerhebung



[Quelle: Flickr, Rhonda, 2009](#)

Stichprobenkennwerte und Populationsparameter

	Stichprobe	Population
Anteil	p	π
Arithmetisches Mittel	\bar{X}	μ
Standardabweichung	s, sd	σ
Varianz	s^2	σ^2

Repräsentative Stichproben

- ▶ Um mit Hilfe einer Stichprobenerhebung gültige Aussagen über ein Merkmal X in der avisierten Population treffen zu können, muss die Stichprobe repräsentativ sein.
- ▶ „Repräsentativ“ kann zweierlei bedeuten:
 - ▶ Die Stichprobe ähnelt der avisierten Grundgesamtheit stark im
 - ▶ Die Abweichung in X sind schätzbar
- ▶ Ein geeigneter Weg für eine repräsentative Stichprobe ist das Ziehen einer Zufallsstichprobe.

Warum keine Adhoc-Stichprobe?

- ▶ Forschungsfrage: „Wie viel Zeit verbringt ein deutscher Bürger (m/w/d) am Tag im Schnitt am Handy?“ (Nennen wir diese Variable X).
- ▶ Angenommen, wir befragen dazu die nächst besten 100 Menschen am Hauptbahnhof. Basiert man die Stichprobenziehung auf einfache Verfügbarkeit, so spricht man von einer Gelegenheits- oder Adhoc-Stichprobe.
- ▶ Ob wir mit einer Adhoc-Stichprobe (am Hauptbahnhof) die Population für X repräsentativ erfassen?
- ▶ Vermutlich nicht, denn
 - ▶ Berufstätige sind womöglich überrepräsentiert, und ihr Handykonsum unterscheidet sich vielleicht von den Nicht-Berufstätigen
 - ▶ Hiesige Menschen sind vermutlich am hiesigen Hauptbahnhof überrepräsentiert, und ihr Handykonsum unterscheidet sich vielleicht von den Nicht-Berufstätigen
 - ▶ ...
- ▶ Vermutlich gibt es weitere Verzerrungen (Konfundierungen, Kollisionsvariablen) die uns nicht bekannt sind, aber eine Rolle spielen

Aber funktioniert das wirklich?

Alle Welt behauptet, Stichproben zu ziehen, sei eine gute Sache.

Aber stimmt das wirklich? Also wirklich wirklich?

[Probieren wir es aus.](#)