

„Ab welcher Fallzahl ist eine Umfrage repräsentativ?“

Immer wieder wird uns diese Frage gestellt. Wie oft im Leben: Die Antwort ist nicht einfach und deshalb nicht in einem Satz oder einer Zahl zu geben.

Wir haben deshalb einige Aspekte zusammengefasst.

Wie viele Interviews führen zu einer repräsentativen Befragung?

Eine entsprechende Mindestgröße der Stichprobe führt zu dem, dass eine heterogene Grundgesamtheit in all ihren Merkmalsausprägungen Einzug in die Stichprobe halten kann. Zum anderen wird das Ergebnis mit zunehmender Stichprobengröße genauer.

Ein Beispiel: Werfe ich einen Würfel nur dreimal, können nicht alle sechs Seiten vertreten sein. Werfe ich den Würfel aber 60 oder sogar 300 Mal, ist zum einen davon auszugehen, dass alle Zahlen eins bis sechs repräsentiert sind. Zum anderen wird sich das Ergebnis immer mehr einer Gleichverteilung über alle sechs Zahlen annähern. Somit bilden 300 Würfe bereits sehr gut ab, wie das Ergebnis bei zig Millionen Würfeln ausfallen würde.

So viele sollten es mindestens sein!

Die deutsche Bevölkerung lässt sich nicht oder nur sehr schwer mit der Auswahl von beispielsweise 100 Personen passend über die Merkmale Alter, Bildung, Berufstätigkeit und deren Kombinationen abbilden. Um dies zu erreichen und hier zuverlässige Aussagen machen zu können, sollten mindestens 500, besser 1.000 Personen befragt werden. Bei einer B2B-Studie über verschiedene Branchen, Größenklassen und Regionen ist eine Stichprobe ab 120, besser 200 Befragten empfehlenswert.

Die folgende Tabelle verdeutlicht den Zusammenhang zwischen dem Stichprobenumfang und der Präzision der Ergebnisse.

In dem Beispiel soll der Anteil der Wählerstimmen für eine größere Volkspartei basierend auf einer Zufallsstichprobe vom Umfang n aus der wahlberechtigten Bevölkerung geschätzt werden (ähnlich der „Sonntagsfrage“).

Zur Quantifizierung der Unsicherheit ist neben dem Stichprobenumfang jeweils die Breite eines 95%-Konfidenzintervalls für den Wähleranteil angegeben. Das Konfidenzintervall ist ein Verfahren aus der Inferenzstatistik, welches einen Bereich ausweist, in dem der tatsächliche Wert aus der Grundgesamtheit mit großer Sicherheit liegt. Je schmaler das Intervall ist, desto präziser die Aussage.

Der Berechnung liegt die Annahme zugrunde, dass der Wähleranteil der Partei in der Grundgesamtheit bei 30% liegt:

Stichproben- umfang	Genauigkeit: 95%-Konfidenzintervall als Abw. vom Anteil in der Stichprobe
n=10	-23.3% bis +35.2%
n=20	-18.1% bis +24.3%
n=50	-12.1% bis +14.6%
n=100	-8.8% bis +10.0%
n=200	-6.3% bis +6.9%
n=500	-4.0% bis +4.2%
n=1000	-2.8% bis +2.9%
n=2000	-2.0% bis +2.1%
n=5000	± 1.3%
n=10000	± 0.9%
n=100000	± 0.3%

Lesebeispiel:

Der Wähleranteil liegt „in der Realität“ bei 30%

- Bei **100 Befragten** wird das Ergebnis mit 95prozentiger Sicherheit zwischen 21,2% und 40% liegen.
- Bei **500 Befragten** wird das Ergebnis mit 95prozentiger Sicherheit zwischen 26% und 34,2% liegen.
- Bei **1.000 Befragten** wird das Ergebnis mit 95prozentiger Sicherheit zwischen 27,2% und 32,9% liegen.

Die Tabelle zeigt auch, wie der Stichprobenumfang entsprechend den Anforderungen an die Genauigkeit der Ergebnisse und des Budgets festgelegt werden kann. Zu beachten ist, dass die Kosten i.d.R. fast proportional zum Stichprobenumfang steigen, während der Gewinn an Präzision mit wachsendem n immer geringer ausfällt.

Quellen:

https://maisbergerinsights.maisberger.de/wann_ist_eine_studie_repraesentativ.html

https://www.inwt-statistics.de/blog-artikel-lesen/Repraesentativtaet_Teil3.html

Für diejenigen, die tiefer in die Thematik eintauchen wollen:

<http://von-der-lippe.org/dokumente/Wieviele.pdf>