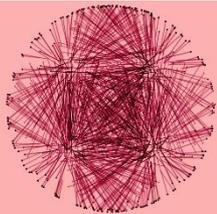


Wie empirische Daten entstehen: Von der Befragung zu den Daten

Dr. Stefan Stuth

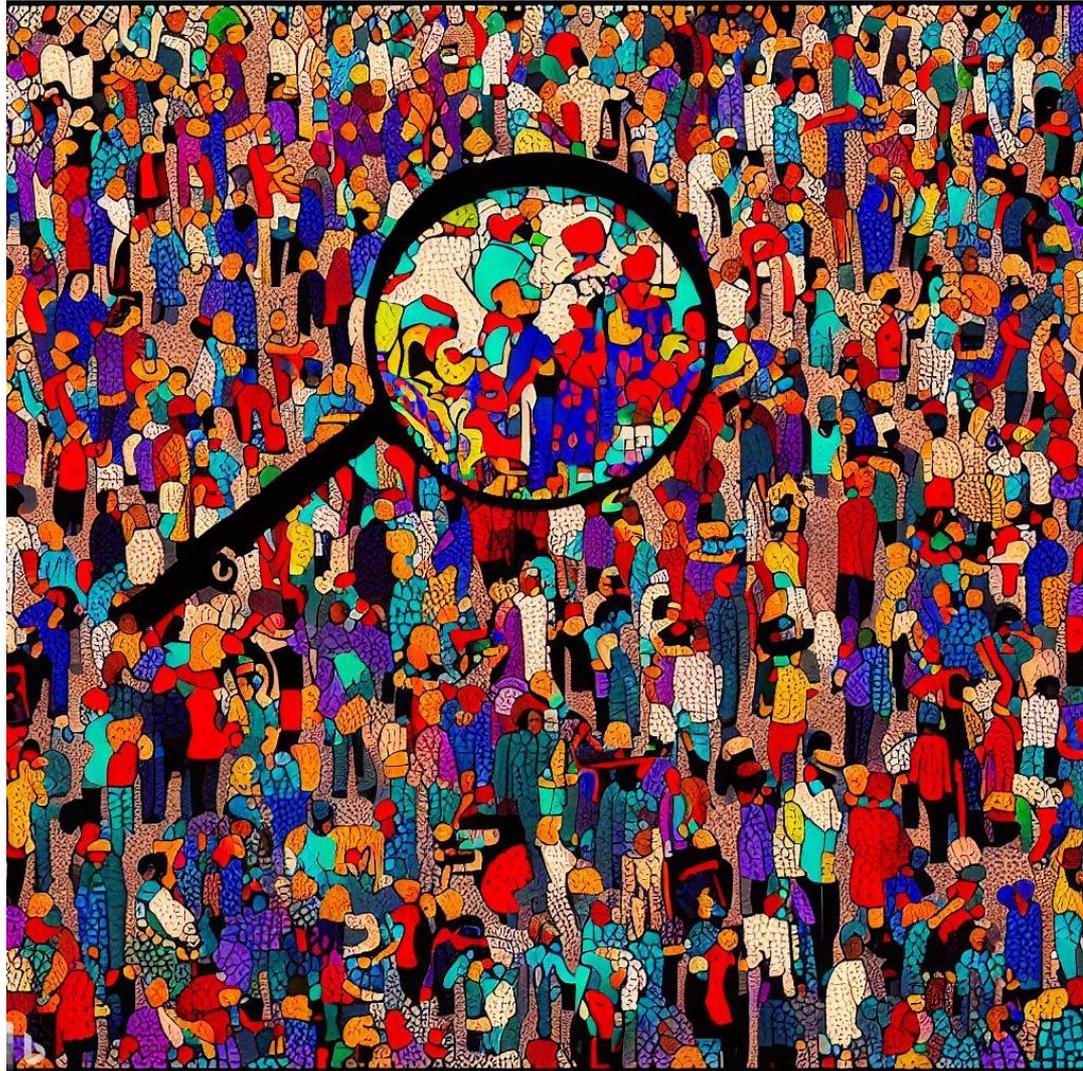


Vortrag im Rahmen der
Langen Nacht der Wissenschaften
28. Juni 2025, Berlin

Befragungen

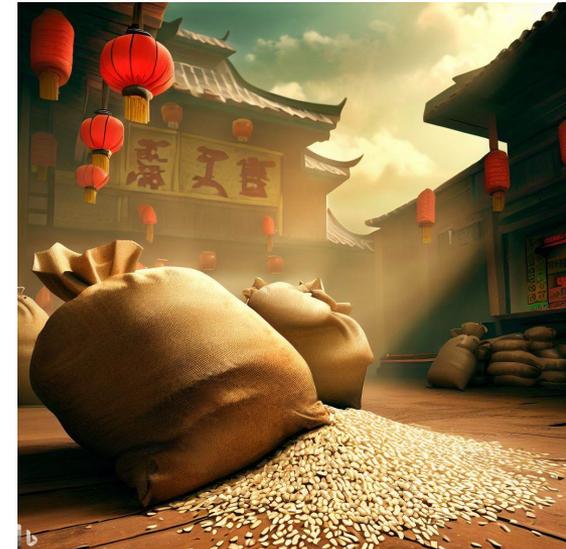
1. Erfassung von Meinungen, Einstellungen, Präferenzen, Wünsche
2. Informationsquelle für objektive Sachverhalte (z.B. Wohlstand und Armut)
3. Grundlage für fundierte Entscheidungen von Politik und Organisationen

Umfragen sind eine Wissenschaft



Design von Fragebögen ist eine Wissenschaft

- Interessante Themen
- Fragen
 - müssen verständlich und kurz sein
 - Dürfen nicht suggestiv sein
 - Müssen neutral sein
 - Dürfen jeweils nur einen Sachverhalt abfragen
 - Dürfen nicht zu allgemein sein
 - Dürfen nicht abschreckend sein
 - Dürfen keine Verneinungen enthalten

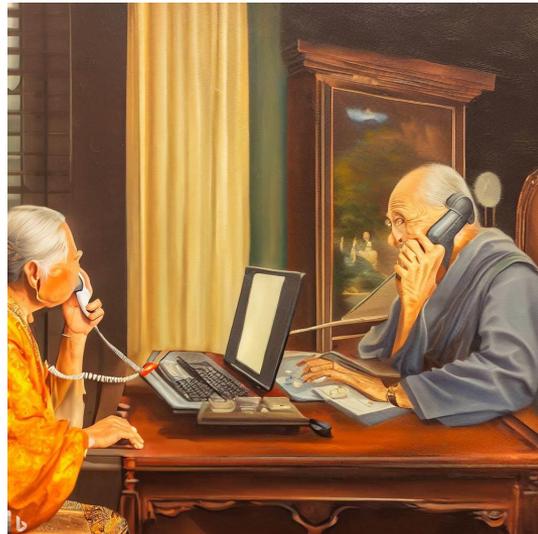


Negativ-Beispiel

Verstehen Sie nicht die absolute Nutzlosigkeit einer solchen exorbitanten Extravaganz in einer sinnentleerten Suggestivfrage?

- 1 – Nein, gar nicht
- 2 – Nein, nicht einmal ein bisschen
- 3 – Sowohl als auch
- 4 – Ja, ein bisschen schon
- 5 – Jawohl, absolut

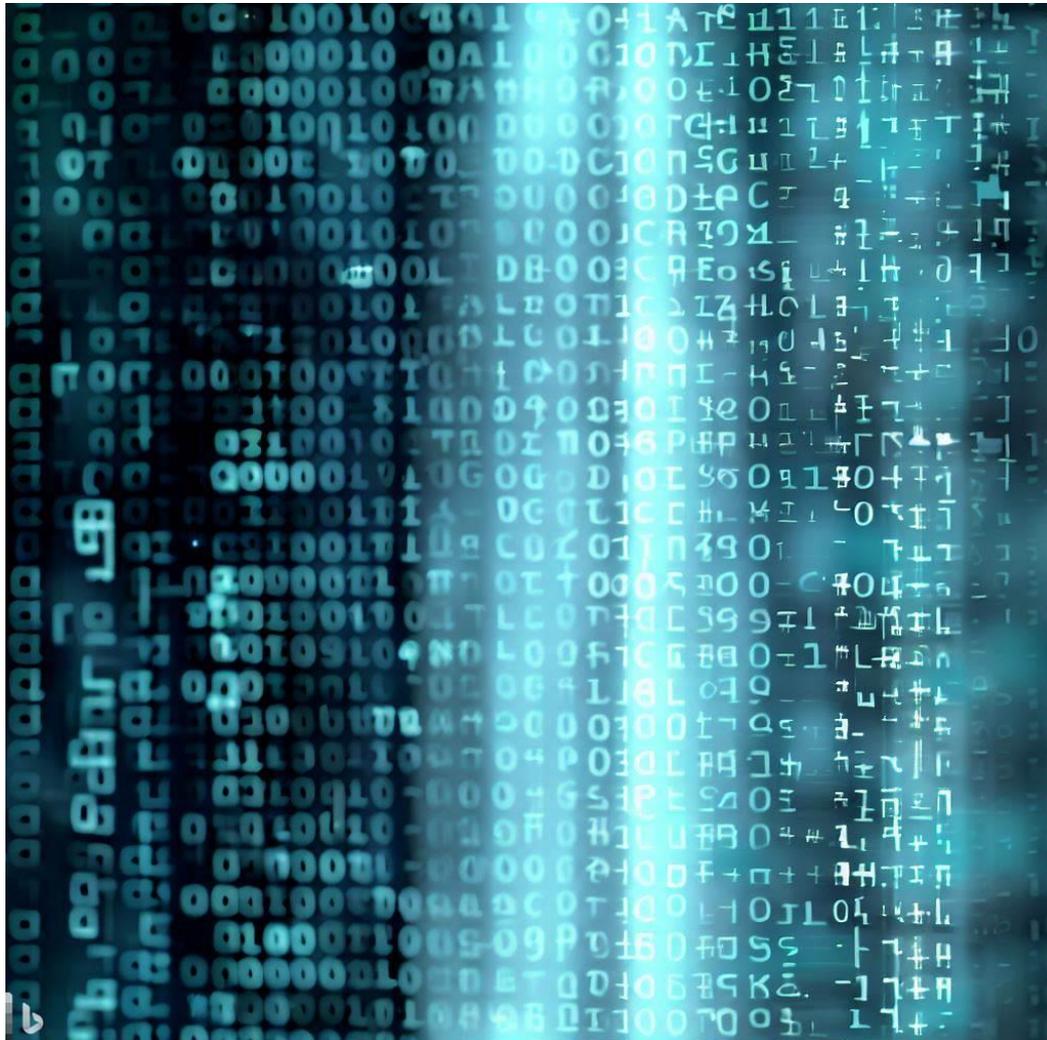
Die Durchführung der Umfrage



Trennung von Personen- und Befragungsdaten

- Befragungsdaten werden von Wissenschaftlern anonym ausgewertet
 - Befragungsdaten werden von Adressdaten getrennt
- Aber:
 - Warum haben schriftliche Fragebögen häufig einen Barcode oder Lfd-Nummer?
 - Selektivität und Repräsentativität
 - Wer hat mitgemacht und wer nicht?
 - Wichtig für Gewichtungsfaktoren
 - Erinnerungsschreiben vs Dankschreiben

Überführung der Befragung in eine Matrix



	ID	GESCHL	GEBJAHR	BERUFSTAETIG	ARBEITSSSTD	ARZTBES	RAUCH	GRO	GEW	SCHULABSCHLUSS	SCHULABSCHLUSS_V	SCHULABSCHLUSS_M
1	1	WEIBLICH	1980	HAUPTBERUFL.GANZTAGS	40.0	0	NEIN	170	68	HOCHSCHULREIFE	VOLKS-,HAUPTSCHULE	FACHHOCHSCHULREIFE
2	2	WEIBLICH	1964	HAUPTBERUFL.GANZTAGS	43.0	1	NEIN	163	75	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
3	3	MAENNLICH	1957	HAUPTBERUFL.GANZTAGS	70.0	0	JA	174	75	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
4	4	MAENNLICH	1952	NEBENHER BERUFSTAE.	NA	2	NEIN	172	115	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
5	5	WEIBLICH	1954	HAUPTBERUFL.HALBTAGS	20.0	1	NEIN	168	87	MITTLERE REIFE	FACHHOCHSCHULREIFE	FACHHOCHSCHULREIFE
6	6	MAENNLICH	1958	HAUPTBERUFL.GANZTAGS	60.0	0	NEIN	182	90	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	MITTLERE REIFE
7	7	MAENNLICH	1947	NICHT ERWERBSTAETIG	NA	1	NEIN	178	101	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
8	8	WEIBLICH	1988	HAUPTBERUFL.GANZTAGS	40.0	6	NEIN	187	86	HOCHSCHULREIFE	HOCHSCHULREIFE	FACHHOCHSCHULREIFE
9	9	MAENNLICH	1955	NICHT ERWERBSTAETIG	NA	0	NEIN	180	100	FACHHOCHSCHULREIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
10	10	MAENNLICH	1961	HAUPTBERUFL.GANZTAGS	35.0	8	NEIN	190	115	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
11	11	MAENNLICH	1947	NICHT ERWERBSTAETIG	NA	1	NEIN	174	82	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
12	12	WEIBLICH	1968	HAUPTBERUFL.GANZTAGS	70.0	1	NEIN	168	56	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
13	13	WEIBLICH	1962	HAUPTBERUFL.GANZTAGS	40.0	1	NEIN	162	54	MITTLERE REIFE	NA	NA
14	14	WEIBLICH	1956	HAUPTBERUFL.GANZTAGS	39.0	0	NEIN	167	80	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
15	15	WEIBLICH	1964	NEBENHER BERUFSTAE.	NA	3	NEIN	162	63	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
16	16	MAENNLICH	1950	HAUPTBERUFL.GANZTAGS	38.0	3	NEIN	173	82	HOCHSCHULREIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
17	17	WEIBLICH	1939	NICHT ERWERBSTAETIG	NA	15	NEIN	155	60	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
18	18	MAENNLICH	1953	HAUPTBERUFL.GANZTAGS	45.0	8	NEIN	175	105	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE
19	19	WEIBLICH	1989	HAUPTBERUFL.GANZTAGS	42.0	2	JA	160	58	FACHHOCHSCHULREIFE	VOLKS-,HAUPTSCHULE	MITTLERE REIFE
20	20	WEIBLICH	1946	NICHT ERWERBSTAETIG	NA	10	NEIN	158	58	MITTLERE REIFE	NA	VOLKS-,HAUPTSCHULE
21	21	MAENNLICH	1955	HAUPTBERUFL.GANZTAGS	40.0	2	JA	179	83	MITTLERE REIFE	VOLKS-,HAUPTSCHULE	VOLKS-,HAUPTSCHULE

Daten prüfen

- Konsistenz
 - Angaben zur Körpergröße: z.B. 16 cm
 - Gefühltes Alter: 6 Jahre, 175 Jahre, 556 Jahre
- Filterwege
 - Befragte ohne Kinder, sollten keine Angaben zu Enkeln aufweisen
- Antwort durch Zielperson?
 - Selektivität und Repräsentativität

Datendokumentation



10 WELL-BEING/PsYCHOLOGICAL RESOURCES (WOHLBEFINDEN / PSYCHOLOGISCHE RESSOURCEN)

10.1 Life satisfaction (Lebenszufriedenheit)

10.1.1 General life satisfaction (Allgemeine Lebenszufriedenheit)

Items	1. In den meisten Dingen ist mein Leben nahezu ideal 2. Meine Lebensbedingungen sind hervorragend 3. Ich bin zufrieden mit meinem Leben 4. Die wichtigsten Dinge, die ich in meinem Leben will, habe ich weitestgehend erreicht 5. Wenn ich mein Leben noch einmal leben könnte, würde ich kaum etwas anders machen	1. In most ways my life is close to my ideal 2. The conditions of my life are excellent 3. I am satisfied with my life 4. So far I have gotten the important things I want in life 5. If I could live my life over, I would change almost nothing
Coding	1. Trifft genau zu 2. Trifft eher zu 3. Weder noch 4. Trifft eher nicht zu 5. Trifft gar nicht zu	1. Strongly agree 2. Agree 3. Neither agree nor disagree 4. Disagree 5. Strongly disagree

DEAS-Instrument / Drop-off 1996-2021

variable names

	1996	2002	2008	2011	2014	2017	2020	2021
1	d1_11	e3_1	f3_1	gd3_1	hd3_1	id3_1	jp1_1	kd3_1
2	d1_12	e3_2	f3_2	gd3_2	hd3_2	id3_2	jp1_2	kd3_2
3	d1_13	e3_3	f3_3	gd3_3	hd3_3	id3_3	jp1_3	kd3_3
4	d1_14	e3_4	f3_4	gd3_4	hd3_4	id3_4	jp1_4	kd3_4
5	d1_15	e3_5	f3_5	gd3_5	hd3_5	id3_5	jp1_5	kd3_5
scale	lz_96	lz_02	lz_08	lz_11	lz_14	lz_17	jp_20	lz_21

Additional information	Skala entspricht dem Mittelwert von mindestens 3 Items, alle Items wurden rekodiert Hohe Werte entsprechen hoher Lebenszufriedenheit	Scale represents the mean of at least 3 items, all items have been recoded High values represent high life satisfaction
-------------------------------	---	--

Reference	<i>Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. Psychological Assessment, 5, 164-172.</i> <i>Schumacher, J. (2003). SWLS. Satisfaction with Life Scale. In J. Schumacher, A. Klaiberg, & E. Brähler (Eds.), Diagnostische Verfahren zu Lebensqualität und Wohlbefinden (pp.305-308). Göttingen: Hogrefe.</i>	
------------------	--	--

Auswertungen

Stata/SE 17.0 - G:\Forsch\FDZ\Daten_fuer_DZA-Mitarbeitende\DEAS\Daten\Stata\SUF_DEAS_Meta_11_2021_de.dta

File Edit Data Graphics Statistics User Window Help

History

Filter commands here

#	Command	_rc
1	do "C:\Users\STEFAN~1\STU\AppData\Local...	

```
(label LABW already defined)
(label LABX already defined)
(label LABY already defined)
(label LABZ already defined)
(label NWGROESS already defined)
(label SCHICHT already defined)
(label SIOPS_KA already defined)
(label ZELLE_96 already defined)
(label STICH already defined)
(label WESTOST already defined)
(label LABE already defined)
(label LABN already defined)
(label LABV already defined)
(label LABAD already defined)
(label LABAM already defined)
(label LABAP already defined)
(label BILDUNG4 already defined)
(label PARTNER already defined)
(label ANZKIND already defined)

Result
```

Result	Number of obs	
Not matched	15,876	
from master	15,876	(_merge==1)
from using	0	(_merge==2)
Matched	4,838	(_merge==3)

```
. drop _merge
.
. *(Nicht-)Teilnehmer in 2017 die in 2014 Basisbefragte waren
. tab part_17 if entry==2014

Command
```

Variables

Filter variables here

Name	Label
fallnum	Fallnummer mit vorangestellter Ziffer der...
stich	Stichprobenzugehoerigkeit - Befragter st...
gemeindenkenner	Anonymisierte Gemeindekennung basier...
pointkenner	Kennung der Sampling Points innerhalb d...
part_96	Teilnahme in 1996
part_02	Teilnahme in 2002

Do-file Editor - Bias-Bildung in 2021.do

File Edit View Language Project Tools

```
Bias-Bildung in 2021.do
(gewichtet vs ungewichtet)
*Schritt 1: Anteile der Personen mit niedriger, mittlerer, hoher Bildung nach Zellen
Geschlecht, Region)
*Jahr 2021
local i=1
while `i' <=12 {
    proportion isced_21 if entry==2014 & part_21==2 & zelle_14=="i" [pw=qsp5_21]
    local o=1
    while `o' <=3 {
        count if entry==2014 & part_21==2 & zelle_14=="i" & isced_21==1
        if r(N)==0 {
            local z`i'`b1_21`=0
            display "z`i'`b1_21" `z`i'`b1_21'
            local z`i'`b2_21`=round(e(b)[1,1], .0001)
            display "z`i'`b2_21" `z`i'`b2_21'
            local z`i'`b3_21`=round(e(b)[1,2], .0001)
            display "z`i'`b3_21" `z`i'`b3_21'
            local o = `o'+3
        }
        else {
            local z`i'`b`o'_21`=round(e(b)[1,`o'], .0001)
            display "z`i'`b`o'_21" `z`i'`b`o'_21'
            local o = `o'+1
        }
    }
    local i = `i'+1
}
```

C:\Users\Stefan.Stuth\Documents

Suchen

14:25 16.06.2023

Auswertung

- Ergebnisse basieren auf Aggregaten:
 - Mittelwerte
 - Anteile
 - Stärke von Zusammenhängen
 - Wahrscheinlichkeiten
 - Linearkombinationen
 - ...

Vielen Dank

