



Soziodemographische Variablen in Umfragen: Forschungspotentiale durch Output- Harmonisierung steigern

Meet the Experts – GESIS-Online-Vorträge

Augmenting survey data by linking and harmonization

Dr. Silke Schneider, Lennart Palm, 11.05.2023

Vortragende



Dr. Silke Schneider

- Silke Schneider ist Soziologin und beschäftigt sich mit der Messung und Harmonisierung soziodemographischer Merkmale in nationalen und internationalen Umfragen.
- Sie verantwortet die Messung von Bildung in ESS und PIAAC C2. In KonsortSWD ko-leitet sie ein Teilprojekt zur Harmonisierung. Außerdem ist sie Vertreterin der ASI in der AG „Demographische Standards“.
- Kontakt: silke.schneider@gesis.org



Lennart Palm

- Lennart Palm ist nach seinem Berliner Master-Studium in Soziologie und Stationen in der Privatwirtschaft wissenschaftlicher Mitarbeiter bei GESIS und KonsortSWD.
- Er beschäftigt sich mit Datenharmonisierung und interessiert sich besonders für die Einkommensmessung und kostengünstige push to web Ansätze.
- Kontakt: lennart.palm@gesis.org

Gliederung

1. Einleitung

- Soziodemographische Merkmale
- Vergleichbarkeit und Harmonisierung
- Kontext: NFDI und KonsortSWD

2. Harmonisierung durch soziodemographische Standardvariablen

- Ausgangspunkt: Fragebogenitems verschiedener Studien
- Entwicklung soziodemographischer Standardvariablen
- Validierung der Standardvariablen

... am Beispiel von Hauptstatus und Erwerbstätigkeit

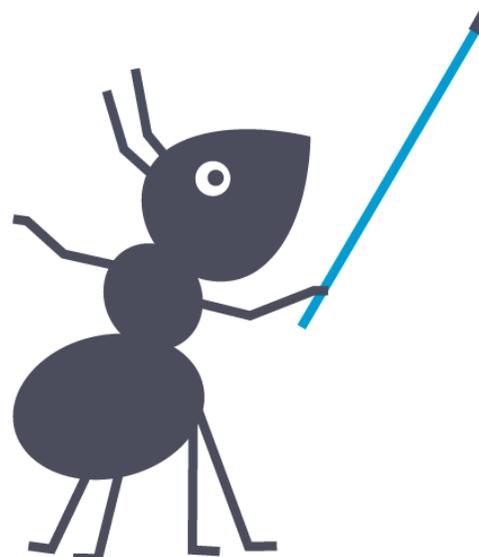
3. Zusammenfassung und Ausblick

Einleitung: Soziodemographische Merkmale, Vergleichbarkeit und Harmonisierung

Relevanz soziodemographischer Merkmale

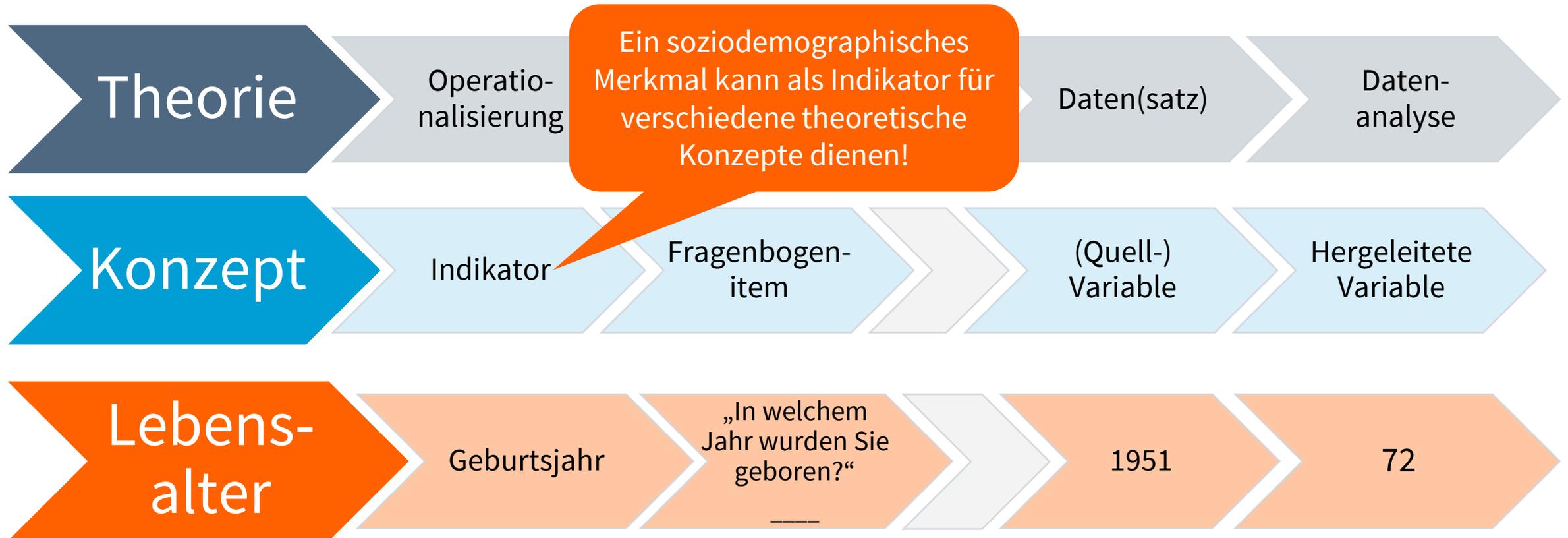
- Merkmale, die Individuen und Haushalte in der Gesellschaft verorten und damit Informationen zum Lebenskontext liefern
 - Langfristig: Geschlecht, Alter, Migrationshintergrund, Religion, Bildung
 - Aktuell: Tätigkeit, Haushalt, Familie, Einkommen, Wohnort
- Soziodemographische Variablen werden in fast allen statistischen Analysen von Mikrodaten eingesetzt
 - Wichtige Kontrollvariablen da sie mit vielen abhängigen und anderen unabhängigen Variablen zusammenhängen (Drittvariablenkontrolle)
 - auch als „Hintergrundvariablen“ bezeichnet
 - In Sozialstrukturanalysen: Hauptvariablen

Bereiche und dazugehörige Merkmale



Basismerkmale	Geburtsmonat und -jahr	Geschlecht	Bildung	Wohnort/ Geocode	
Arbeitsmarkt	Hauptstatus	Erwerbstätigkeit	Beruf	Stellung im Beruf	HH-Nettoeinkommen
Kultur, Migration	Staatsangehörigkeit	Religionszugehörigkeit	In Deutschland geboren	Geburtsland	Sprache
Familie, Haushalt	Familienstand, Kohabitation	Partnerschaft	Haushaltsgröße	HH-Zusammensetzung	Anzahl Kinder

Der Messprozess bei Befragungen



Warum ist Vergleichbarkeit von Variablen wichtig?

- Vergleichbarkeit von Variablen...
 - Stärkt den kumulativen Charakter von Forschung
 - Erleichtert Metanalysen
 - Messunterschiede häufige Begründung für widersprüchliche Forschungsergebnisse!
 - Vereinfacht Kumulation von Daten verschiedener Studien
 - Ermöglicht Analyse kleiner Gruppen
 - Erleichtert data linkage
- Dadurch steigt also Interoperabilität (inhaltlich) und Nachnutzbarkeit von Daten im Sinne von „**FAIR** data“
 - Findable, Accessible, **Interoperable und Reusable**

**Forschungspotentiale
steigern!**

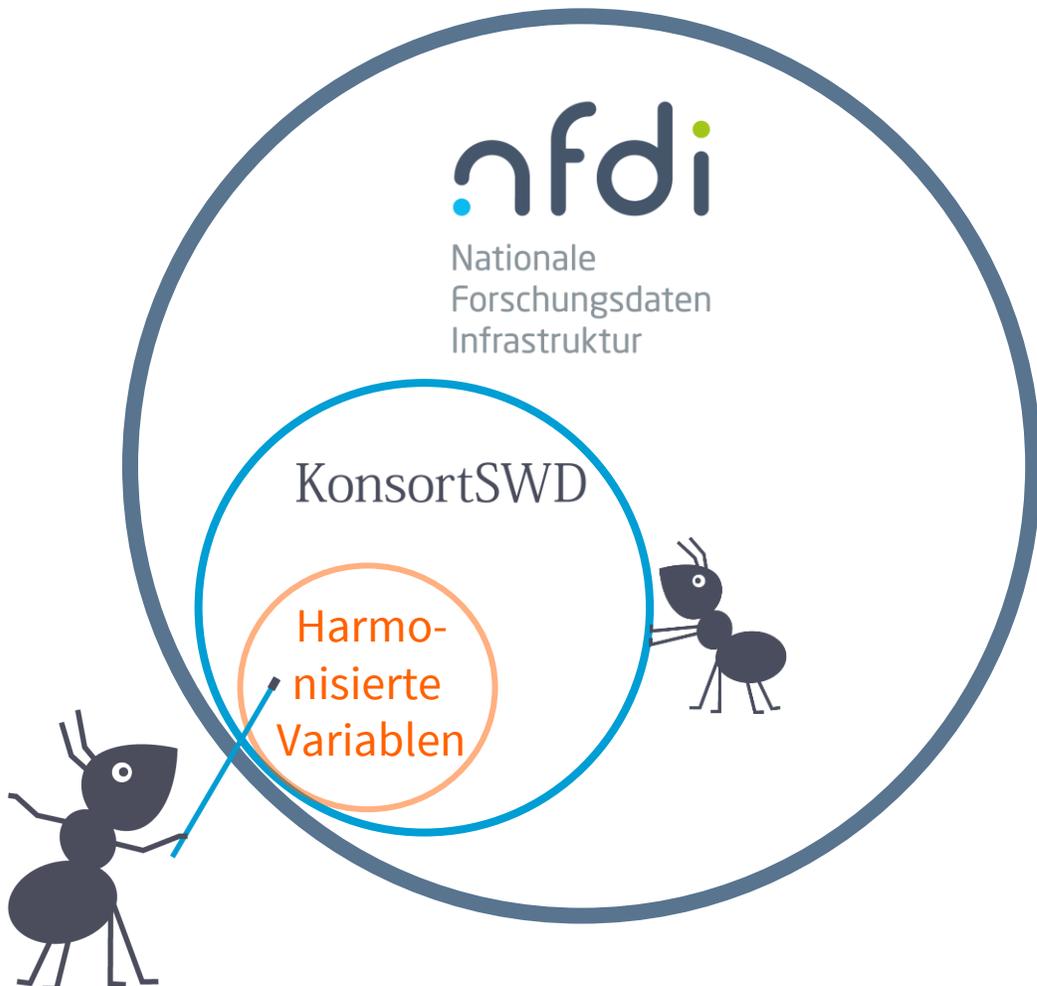


Wie werden soziodemographische Variablen vergleichbar gemacht? Durch Harmonisierung.

- **Input-Harmonisierung: Standardisierung der Erhebungsinstrumente**
 - = Einigung auf soziodemographische Standarditem(s)
 - ✓ Ziel der „Demographischen Standards“ und des Standardfragenkatalogs von Best_FDM
- **ABER: viele Studien wenden diese Standards nicht an!**
- **Output-Harmonisierung: Standardisierung der Variablen im Datensatz**
 - Ex-post: Variablen spezifischer Datensätze werden bestmöglich harmonisiert
 - = keine Einigung auf gemeinsame Standards (aber man kann gleich loslegen!)
 - Ex-ante: fließt ggf. schon in Fragebogendesign ein (wie in international vergleichenden Studien)
 - = Einigung auf soziodemographische Standardvariable(n)
 - ✓ (Ein) Ziel des KonsortSWD-Teilprojekts „Harmonisierte Variablen“

Fokus heute!

NFDI, KonsortSWD und harmonisierte Variablen



- Mit der **Nationalen Forschungsdateninfrastruktur (NFDI)** sollen heute oft dezentral, projektförmig und temporär gelagerte Datenbestände von Wissenschaft und Forschung für das deutsche Wissenschaftssystem systematisch erschlossen, nutzbar gemacht und nach den **FAIR-Prinzipien** zur Verfügung gestellt werden.
- Umsetzung in derzeit 27 Konsortien, die Forschungsfelder repräsentieren.
- **KonsortSWD** ist eines der ersten **NFDI-Konsortien** (2020-2025) und vertritt die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften
- KonsortSWD Measure **TA3.M1 „Harmonisierte Variablen“** zielt auf bessere Vergleichbarkeit von Variablen über (deutsche) Studien hinweg ab
 - Latente Konstrukte (Skalen): QuestionLink
 - Soziodemographische Variablen

Harmonisierung durch soziodemographische Standardvariablen

Ausgangspunkt: Fragebogenitems verschiedener Studien
am Beispiel von Erwerbstätigkeit und Hauptstatus

Operationalisierung von Erwerbstätigkeit

Amtliche Statistik: ILO-Konzept

Erwerbstätige sind alle über 15 Jahre alten Personen, die in einem festgelegten Zeitraum

- Mindestens eine Stunde
- Gegen Entgelt oder zur Erzielung eines Gewinns in Geld oder Sachleistungen arbeiteten,

oder die vorübergehend

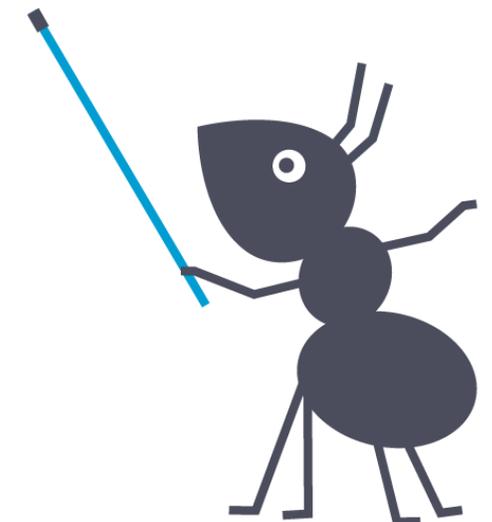
- von einem Arbeitsplatz, an dem sie bereits gearbeitet hatten und mit dem sie weiterhin formell verbunden waren,
- oder von einer selbständigen Beschäftigung abwesend waren (z.B. Urlaub, Krankheit oder Elternzeit).

➤ Ziel: international vergleichbare Erwerbsstatistik (insbes. Arbeitslosenquoten)

Wissenschaftliche Umfragen

Die große Mehrzahl begnügt sich mit der subjektiven Einschätzung der eigenen Erwerbstätigkeit (und des Hauptstatus).

Dazu nun ein paar Beispiele:



Die Fragebogenitems der Demographischen Standards

Nr.	Demographische Standards	weiter mit
8 ³	<p>Welche Erwerbssituation passt für Sie? Was in dieser Liste trifft auf Sie zu? Bitte beachten Sie, dass unter Erwerbstätigkeit jede bezahlte bzw. mit einem Einkommen verbundene Tätigkeit verstanden wird. <i>(Nur eine Nennung möglich)</i> Liste „8“ vorlegen!</p> <p>A Vollzeitwerbstätig () 8 A</p> <p>B Teilzeiterwerbstätig () 8 A</p> <p>C Altersteilzeit (unabhängig davon, ob in der Arbeits- oder Freistellungsphase befindlich) () 8 A</p> <p>D Geringfügig erwerbstätig, 450-Euro-Job⁴, Minijob () 8 A</p> <p>E „Ein-Euro-Job“ (bei Bezug von Arbeitslosengeld II) () 8 A</p> <p>F Gelegentlich oder unregelmäßig beschäftigt () 8 A</p> <p>G In einer beruflichen Ausbildung/Lehre () 11</p> <p>H In Umschulung () 10</p> <p>I Freiwilliger Wehrdienst () 10</p> <p>J Bundesfreiwilligendienst oder Freiwilliges Soziales Jahr () 10</p> <p>K Mutterschafts-, Erziehungsurlaub, Elternzeit oder sonstige Beurlaubung (Altersteilzeit unter C angeben) () 9</p> <p>L Nicht erwerbstätig (einschließlich: Schüler/-innen oder Studierende, die nicht gegen Geld arbeiten, Arbeitslose, Vorruheständler/-innen, Rentner/-innen ohne Nebenverdienst) () 9</p>	

12 Kategorien

9	<p>Wenn Sie nicht vollzeit- oder teilzeiterwerbstätig sind: Sagen Sie bitte, zu welcher Gruppe auf dieser Liste Sie gehören. Liste „9“ vorlegen!</p> <p>A Schüler/-innen an einer allgemeinbildenden Schule () 13 A</p> <p>B Studenten/-innen () 10</p> <p>C Rentner/-innen, Pensionäre/-innen, im Vorruhestand () 10</p> <p>D Arbeitslose () 10</p> <p>E Dauerhaft Erwerbsunfähige () 10</p> <p>F Hausfrauen/Hausmänner () 10</p> <p>G Sonstiges, () 10</p> <p>und zwar: _____</p>	
---	--	--

7 Kategorien

- C-F werden nach weiteren Fragen zu ihrer Erwerbstätigkeit (8A-D) auch zu Frage 9 geführt, um den Hauptstatus zu erheben, da sie nur nebenher erwerbstätig sind
- Viele Antwortkategorien, dafür keine Interviewerhinweise nötig!

Die Fragebogenitems im ALLBUS

Fragebogensdokumentation ALLBUS 2018

F058

work

⇒ *Liste 58 vorlegen!*

Nun weiter mit der Erwerbstätigkeit und Ihrem Beruf.
Was von dieser Liste trifft auf Sie zu?

⇒ *Nur eine Nennung möglich!*

- A Hauptberufliche Erwerbstätigkeit, ganztags
- B Hauptberufliche Erwerbstätigkeit, halbtags
- C Nebenher erwerbstätig
- D Nicht erwerbstätig
- KA

4 Kategorien

Weniger Antwortkategorien, dafür sehr ausführliche Interviewerhinweise!

Da dies eine wichtige Frage ist, versuchen Sie bitte eine Antwort zu erhalten.
Falls es Schwierigkeiten bezüglich der Einstufung gibt, hier noch einige Hinweise:
Lehrlinge / Auszubildende gelten als HAUPTBERUFLICH Erwerbstätige.

F068 (Falls Befragter nicht oder nebenher erwerbstätig ist („C“ oder „D“ in F058))

dw03

⇒ *Liste 68 vorlegen!*

Bitte sehen Sie diese Liste einmal durch, und sagen Sie mir, was davon auf Sie zutrifft.
Nennen Sie mir bitte den entsprechenden Kennbuchstaben.

⇒ *Nur eine Nennung möglich!*

- A Ich bin Schüler / Student
- B Ich bin Rentner / Pensionär
- C Ich bin zur Zeit arbeitslos
- D Ich bin Hausfrau / Hausmann
- E Ich leiste freiwilligen Wehrdienst/Bundesfreiwilligendienst/FSJ/FÖJ
- F Ich bin aus anderen Gründen nicht (hauptberuflich) erwerbstätig
- KA

erwerbstätig.

6 Kategorien

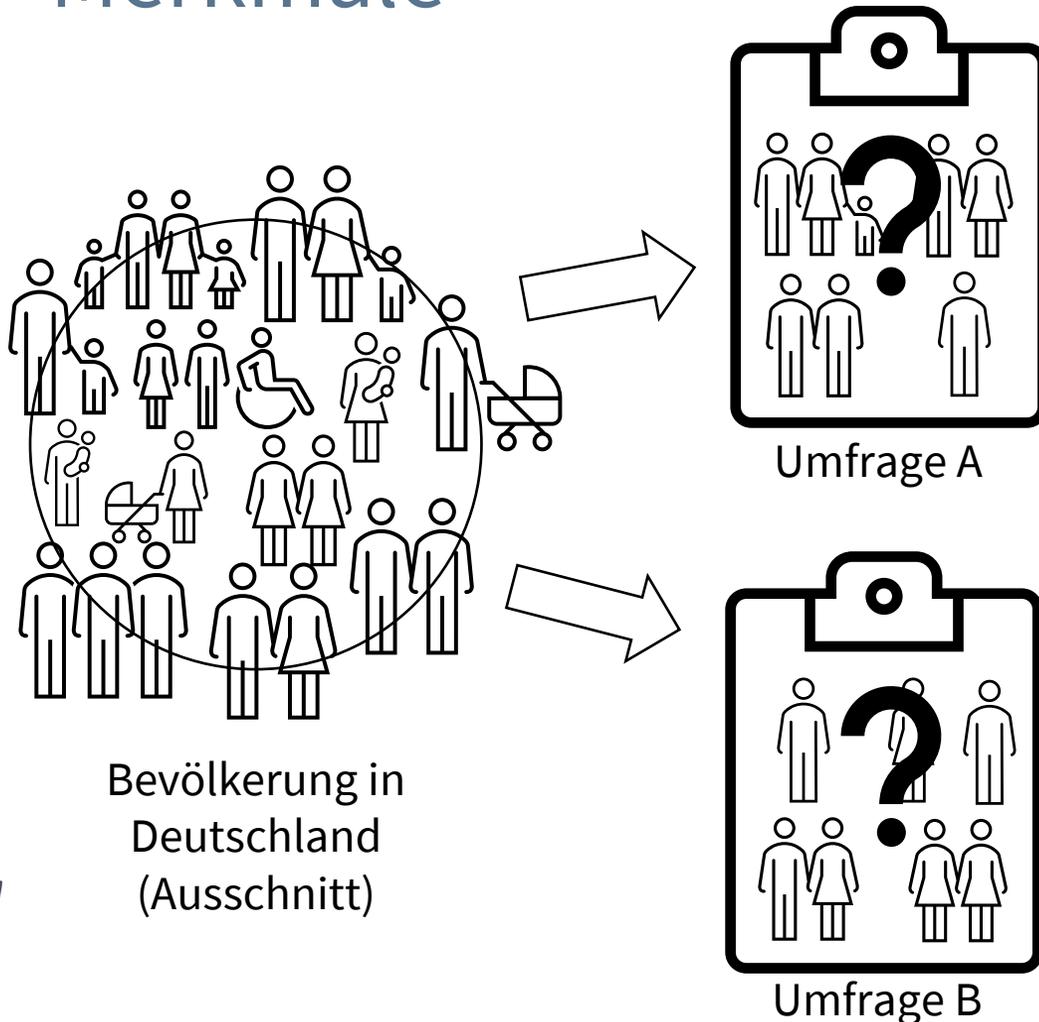
Das Fragebogenitem in der GLES

- Hier wird nur der Hauptstatus erfragt
- Erwerbstätigkeit *nebenher* wird nicht erfasst

vn138	Erwerbstätigkeit
Vw q138	Vw T06.6
Nw q137	Nw T06.6
Gefiltert von Fragennummer(n) Vw/Nw:	
Programmieranweisung Vw/Nw:	
Intervieweranweisung Vw/Nw: [Liste vorlegen]	
Fragetext Vw/Nw: Nun weiter mit der Erwerbstätigkeit und Ihrem Beruf. Was von dieser Liste trifft auf Sie zu?	
Items Vw/Nw:	
Scale Vw/Nw:	
(1) 1 - Vollzeit berufstätig (mehr als 30 Stunden/Woche)	
(2) 2 - Teilzeit berufstätig (bis 30 Stunden/Woche)	
(3) 3 - Lehrling/Azubi	
(4) 4 - Schüler	
(5) 5 - Student	
(6) 6 - in Umschulung	
(7) 7 - zurzeit arbeitslos	
(8) 8 - zurzeit in Kurzarbeit	
(9) 9 - Bundesfreiwilligendienst, Freiwilliges Soziales Jahr (FSJ), Freiwilliges Ökologisches Jahr (FÖJ)	
(10) 10 - Pensionär/Rentner (früher berufstätig)	
(11) 11 - in Mutterschutz, Elternzeit	
(12) 12 - nicht berufstätig (Hausfrau/Hausmann)	

(-98) weiß nicht	
(-99) keine Angabe	

Problem: heterogene Messung soziodemographischer Merkmale

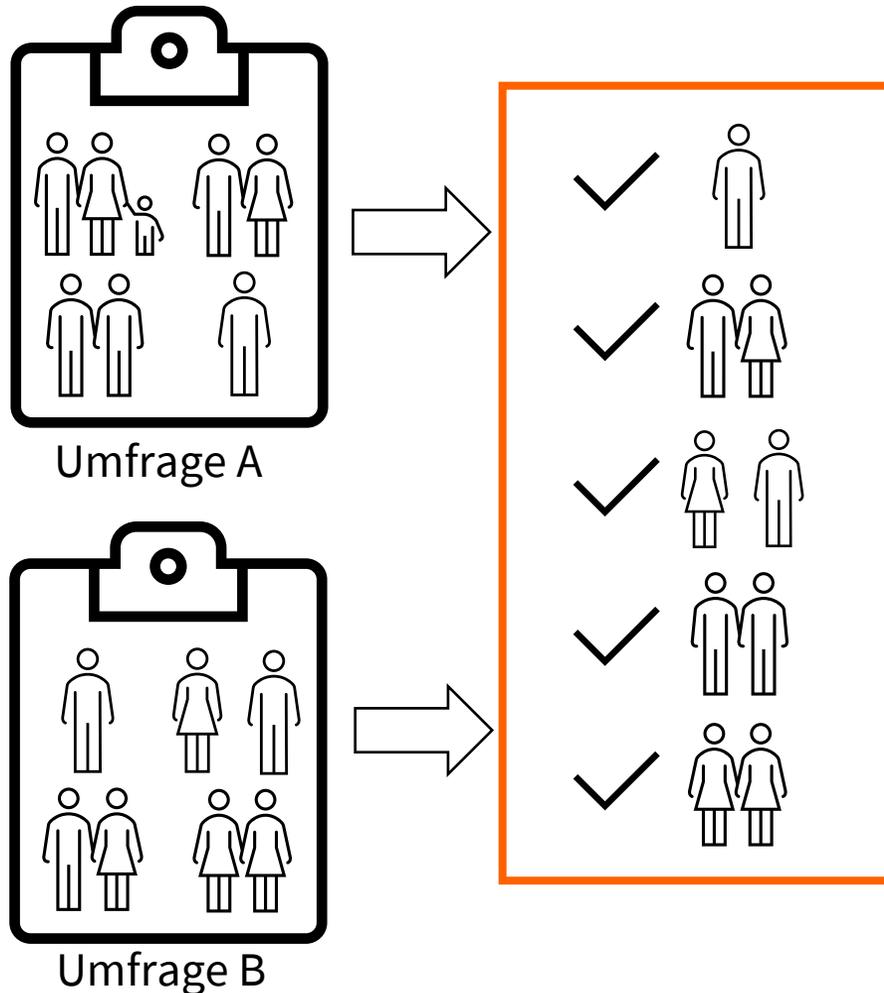


- Umfragen in Deutschland nutzen unterschiedliche Erhebungsinstrumente für soziodemographische Merkmale*
 - Trotz Vorliegen der „Demographischen Standards“
- Die erhobenen Variablen sind daher nicht unmittelbar vergleichbar, nicht „*interoparabel*“.

*Review von Erhebungsinstrumenten in großen deutschen Umfragen: [Schneider et al. 2022](#) ¹⁶



Lösung: soziodemographische Standardvariablen



- Vergleichbarkeit lässt sich durch die Rekodierungen der Variablen der unterschiedlichen Umfragen häufig herstellen
- **Aber:** Hierzu liegen bislang keine allgemeinen Standards vor, an denen sich Studien, FDZen oder Datennutzende orientieren könnten

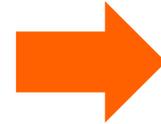
➤ *Wir entwickeln diese Standards*



Entwicklung soziodemographischer Standardvariablen

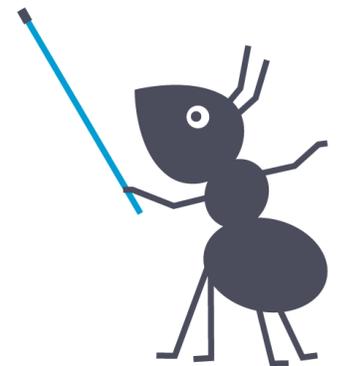
Komponenten von Standardvariablen

- Ebene der Variablen:
 - Variablenname (EN)
 - Präfix „x_“ für leichte Auffindbarkeit
 - Variablenlabel (EN/DE)



	Hauptstatus	Erwerbstätigkeit
Variablenname	x_mainstatus	x_empl
Variablenlabel EN	Main status (subjective)	Employment (subjective)
Variablenlabel DE	Hauptstatus (subjektiv)	Erwerbstätigkeit (subjektiv)

- Ebene der Ausprägungen/Werte:
 - Werte bzw. Codes
 - Wertelabels bzw. Kategorienbezeichnungen (EN/DE)
 - Für alle Merkmale anzuwenden: Fehlende Werte



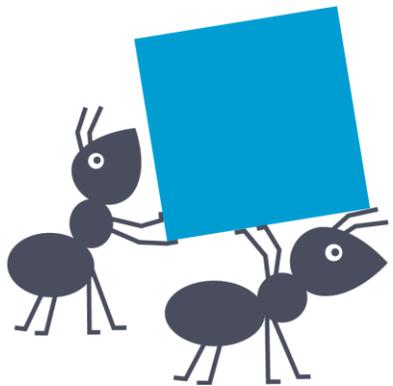
Standardspezifikation für fehlende Werte

Werte	Label	Erklärung
-1	No answer	Zielperson hat keine Angabe getätigt, Angabe verweigert
-2	Don't know	Zielperson hat „weiß nicht“ geantwortet
-3	Filter missing	Merkmal trifft nicht zu , d.h. Zielperson bekommt Frage durch Filterführung nicht gestellt (computer-assisted modes)/hätte Frage nicht beantworten sollen (schriftliche Befragungen).
-4	Not included	Merkmal nicht erhoben (entweder generell oder nur in dieser Erhebungswelle)
-5	Cannot be generated	Variable ist nicht generierbar. Für generierte Variablen: mind. in einer Quellvariable nicht inhaltlich geantwortet und/oder entsprechende Frage nicht gestellt, d.h. wenn mind. eine der Quellvariablen entweder -1 oder -2 oder -3 oder -4 ist.



Unsere Vorschläge für Kategorienschemata beruhen auf:

- Etablierten Klassifikationen: z.B.
 - International Standard Classification of Education (ISCED) und das CASMIN-Schema (aus Projekt „Comparative Analysis of Social Mobility in Industrial Nations“) für Bildung
 - International Standard Classification of Occupations (ISCO) und Klassifikation der Berufe (KldB) für Berufe
- Quellen für einfachere Kategorienschemata:
 - Demographische Standards
 - EU Key Social Variables (EC 2020)
 - Vorhandene Umfragen, Harmonisierungsprojekte
- Empirischen Validierungsanalysen
- Feedback aus der Community



Ausgangspunkte für Standardvariablen für Erwerbstätigkeit und Hauptstatus

Erwerbstätigkeit

1. Fragebogenitem des ALLBUS 2018

Fragebogendokumentation ALLBUS 2018

F058

work

⇒ Liste 58 vorlegen!

Nun weiter mit der Erwerbstätigkeit und Ihrem Beruf.
Was von dieser Liste trifft auf Sie zu?

⇒ Nur eine Nennung möglich!

- A Hauptberufliche Erwerbstätigkeit, ganztags
- B Hauptberufliche Erwerbstätigkeit, halbtags
- C Nebenher erwerbstätig
- D Nicht erwerbstätig
- KA

Hauptstatus

Für den Hauptstatus bieten sich die Kategorien der EU key social variables an:

Employed
Unemployed
Retired
Unable to work due to long-standing health problems
Student, pupil
Fulfilling domestic tasks
Compulsory military or civilian service
Other

Maximal- und Minimalversionen

**Umsetzung: hierarchische
Codeschemata**

Maximalversion:

- ✓ Soll eine möglichst hohe Validität aufweisen
- ✓ Erlaubt weitere Vereinfachungen für spezifische Forschungsfragen durch Datennutzende
- Lässt sich jedoch nicht für alle Studien herleiten

Minimalversion:

- ✓ Soll von einer möglichst großen Zahl von Studien herleitbar sein
- Lässt jedoch größere Validitätsverluste erwarten, die je nach Forschungskontext unterschiedlich sein können
- Mglw. für spezifische Forschungsfragen zu undifferenziert

Die Vorschläge für Erwerbstätigkeit

Maximalversion 2	Maximalversion 1=ALLBUS	Minimalversion 2	Minimalversion1
Vollzeit-erwerbstätig	Vollzeit-erwerbstätig	Erwerbstätig in Vollzeit	Erwerbstätig
Teilzeit-erwerbstätig	Teilzeit-erwerbstätig	Erwerbstätig, aber nicht in	
Nebenher erwerbstätig	Nebenher erwerbstätig	Vollzeit (Teilzeit, nebenher oder unregelmäßig)	
Arbeitslos	Nicht erwerbstätig	Nicht erwerbstätig	Nicht erwerbstätig
Nicht erwerbstätig und nicht arbeitslos			

- Unterschiedlich detaillierte Version(en) entworfen
- Unklar wie gut welche Version in der Datenanalyse „funktioniert“
- Validierung als Entscheidungshilfe: vergleicht Vorhersagekraft der Versionen
- Nach Auswertung ggf. Erstellung eines hierarchischen Codeschemas

Standardvariable „Hauptstatus“

- 1. Stelle im hierarchischen Codeschema entspricht Kategorie der EU key social variables
- 2. Stelle gibt zusätzliche für Deutschland relevante Details

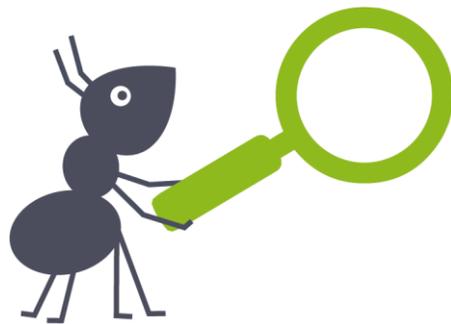
Standardvariable x_mainstatus	EU key social variable
Wert Wertelabels	
10 Erwerbstätig, nicht näher spezifiziert	
11 Vollzeiterwerbstätig	Employed
12 Teilzeiterwerbstätig	
20 Erwerbslos	Unemployed
30 Im Ruhestand, pensioniert	Retired
40 Dauerhaft erwerbsunfähig	Unable to work due to long-standing health problems
50 Studierende, SchülerIn	Student, pupil
60 Hausarbeit leistend	Fulfilling domestic tasks
70 In verpflichtendem Wehr- oder Zivildienst	Compulsory military or civilian service
80 Andere, nicht näher spezifiziert	
81 In freiwilligem Wehrdienst, Bundesfreiwilligendienst u.ä.	Other

Da etabliertes Schema und wenig zusammenfassbare Kategorien: keine Validierung!

Validierung der Standardvariablen Erwerbstätigkeit

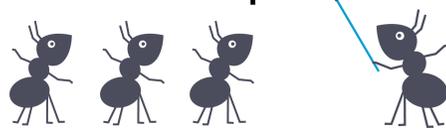
Validierung zur Qualitätssicherung

- Wie schneiden die unterschiedlichen Versionen ab?
- Können die Minimalversionen überhaupt empfohlen werden?
- Geht durch die Output-Harmonisierung Vorhersagekraft verloren?



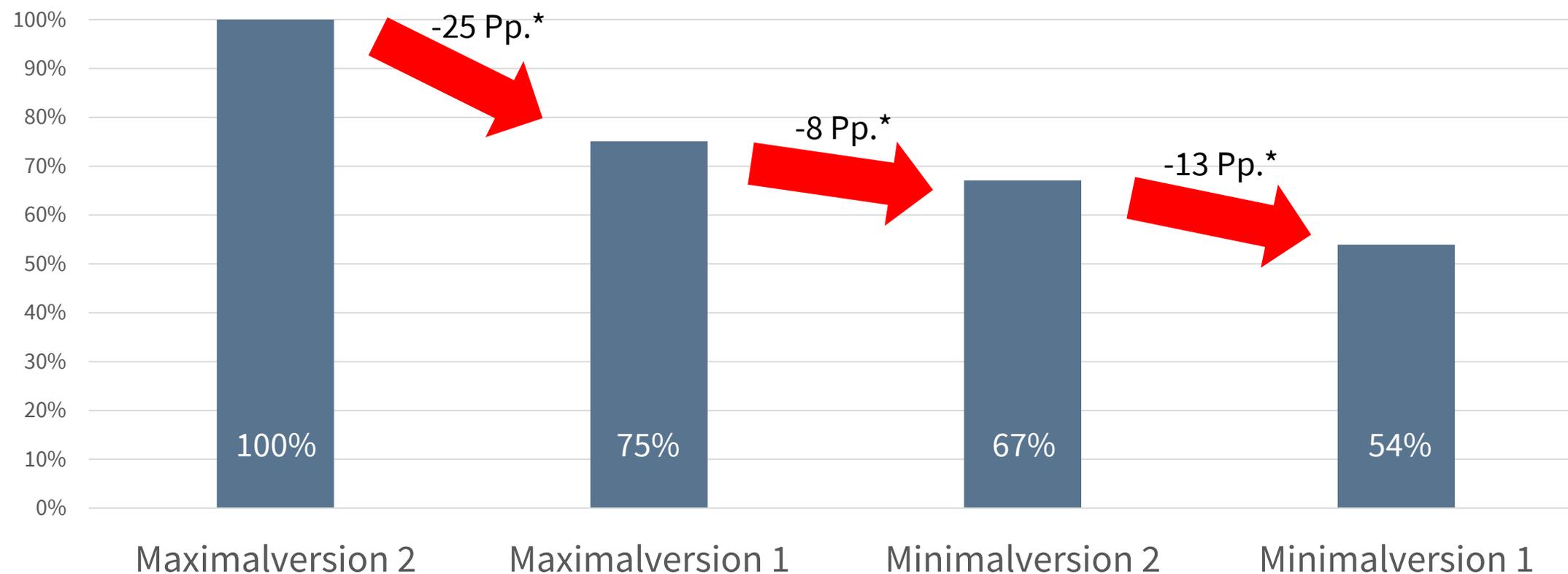
Vorgehen in der Validierung

1. Standardvariablen werden als Prädiktoren in Regressionsmodellen eingesetzt
 2. Vergleich der Vorhersagekraft der unterschiedlichen Versionen der Standardvariablen mit der Vorhersagekraft der detailliertesten Version
 - Vorhersagekraft = partielles adjustiertes R^2
- A - Theoriegeleitetes Verfahren (Ergebnisse hier nicht gezeigt):
- Theoretische begründete Auswahl von 2 bis 4 Validierungsvariablen pro Merkmal
 - inhaltliche Prüfung der Koeffizienten
- B - Datengeleitetes Verfahren:
- 189 mindestens 5-stufig ordinal- oder kontinuierlich skalierten Variablen aus ALLBUS 2018 kommen als abhängige Variablen in Frage
 - Auswahl von Validierungsvariablen pro Merkmal: detaillierteste Version zeigt partielles adj. $R^2 \geq .01$ (bei Erwerbstätigkeit sind es 24 Variablen)
 - Berechnung der durchschnittlichen partiellen adj. R^2 über Validierungsvariablen hinweg



Abnahme der Vorhersagekraft unterschiedlicher Versionen von Erwerbstätigkeit

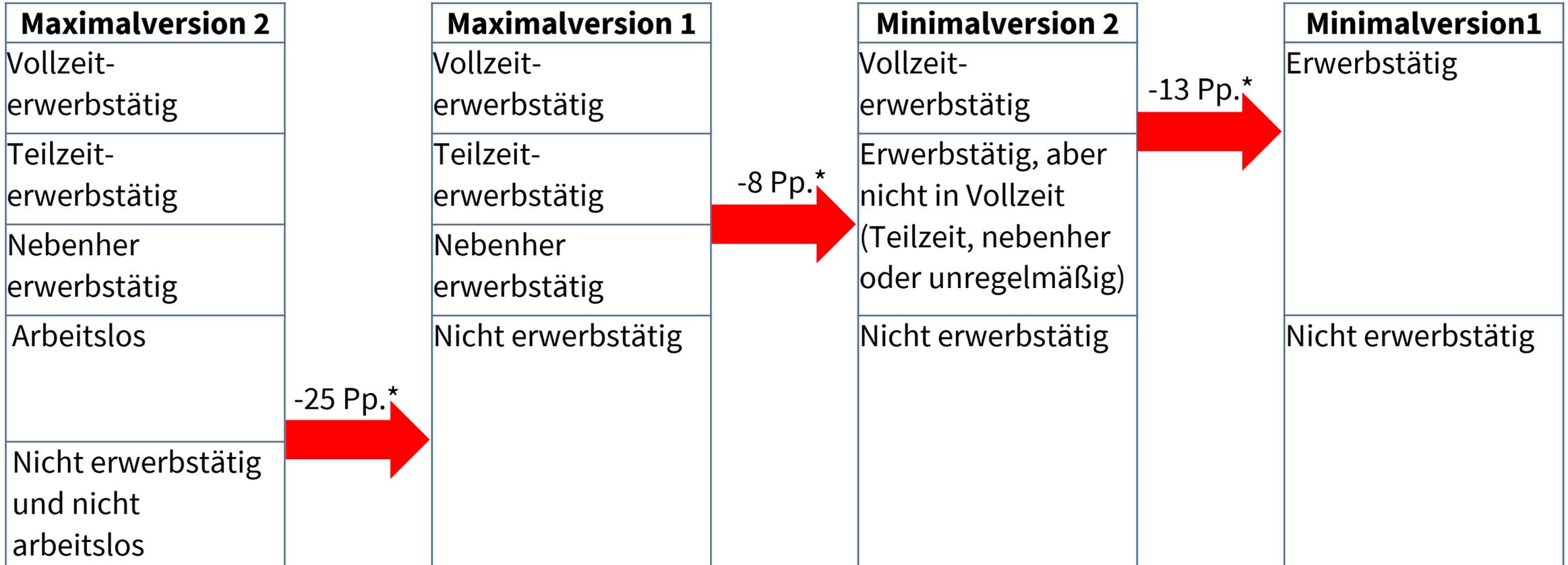
Ergebnisse der datengeleiteten Validierung für Erwerbstätigkeit



*Pp.: Prozentpunkte

Ergebnisse: in 24 der 189 Variablen des Validierungspools erklärt Maximalversion2 jeweils mehr als 1% der Varianz.

Die Auswirkungen der Zusammenfassung unterschiedlicher Kategorien bei Erwerbstätigkeit



*Pp.: Prozentpunkte

Nur eine Version kann vorgeschlagen werden



Maximalversion 2
Vollzeit- erwerbstätig
Teilzeit- erwerbstätig
Nebenher erwerbstätig
Arbeitslos
Nicht erwerbstätig und nicht arbeitslos



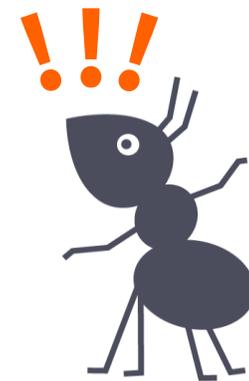
x_empl	
Wert	Wertelabels
1	Vollzeit- erwerbstätig
2	Teilzeit- erwerbstätig
3	Nebenher erwerbstätig
4	Arbeitslos
5	Nicht erwerbstätig und nicht arbeitslos

- Die Ergebnisse der datengeleiteten, aber auch der theoretisch geleiteten Validierung zeigen Maximalversion 2 als beste Version.
- Weitere Versionen konnten sich nicht bewähren und werden daher auch nicht vorgeschlagen.
- Studien, die nebenher Erwerbstätige nicht identifizieren können, können nur die Standardvariable Hauptstatus anbieten

Zusammenfassung und Ausblick

Output Harmonisierung ist kein Allheilmittel!

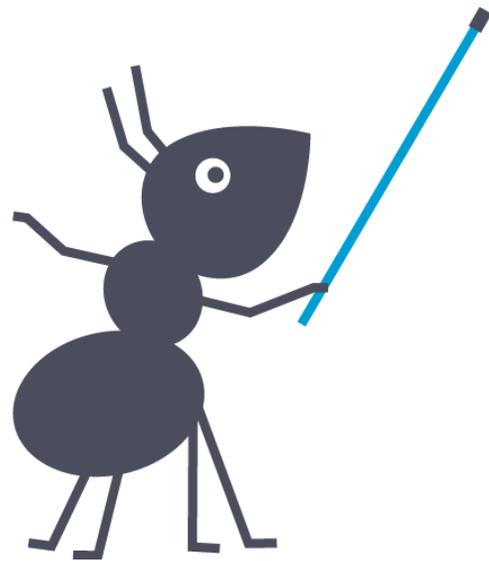
- Es kann nur vergleichbar gemacht werden, was in unterschiedlichen Umfragen vergleichbar gemessen wurde
 - Voraussetzung: Gleiches Konzept, gleicher Indikator, ähnliche Fragebogenitems
- Unterschiede in Fragebogenitems (Formulierung, Instruktionen, Antwortkategorien...) führen zu Messunterschieden, die durch Harmonisierung nicht verschwinden
 - Bei soziodemographischen Merkmalen weniger problematisch als bei Einstellungen und anderen stärker subjektiven Konzepten



Fazit und Ausblick

- Soziodemographische Standardvariablen könnten bei Sekundärdatenanalysen unterstützen, kumulative Erkenntnisse fördern und das Zusammenführen von Daten vereinfachen
- Vereinfachungen können Validität massiv reduzieren
- Implementierung der Standardvariablen durch Datennutzende oder FDZen möglich
 - Erster use case: NEPS SC6 am LIfBi
- Standardvariablen werden im Herbst veröffentlicht bei <https://zis.gesis.org/>
- Wir freuen uns über Anmerkungen und Feedback!

Merkmale, für die Standardvariablen entwickelt werden



Basismerkmale	Geburtsmonat und -jahr	Geschlecht	Bildung	Wohnort/ Geocode	
Arbeitsmarkt	Hauptstatus	Erwerbstätigkeit	Beruf	Stellung im Beruf	HH-Nettoeinkommen
Kultur, Migration	Staatsangehörigkeit	Religionszugehörigkeit	In Deutschland geboren	Geburtsland	Sprache
Familie, Haushalt	Familienstand, Kohabitation	Partnerschaft	Haushaltsgröße	HH-Zusammensetzung	Anzahl Kinder

Literatur

- European Commission. (2020). Standardised key social variables-Implementing guidelines. <https://ec.europa.eu/eurostat/documents/54431/1966394/Standardised-key-social-variables.pdf>
- GESIS-Leibniz-Institut Für Sozialwissenschaften. (2019). ALLBUS/GGSS 2018 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2018) (2.0.0) [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.13250>
- GLES. (2019). Vor- und Nachwahl-Querschnitt (Kumulation) (GLES 2017) (3.0.1) [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.13236>
- Hoffmeyer-Zlotnik, J. H. P., Beckmann, K., Glemser, A., Heckel, C., von der Heyde, C., Schneider, S. L., Hanefeld, U., Herter-Eschweiler, R., & Kühnen, C. (2016). Demographische Standards, Ausgabe 2016: Eine gemeinsame Empfehlung des ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e. V. (ASI) und des Statistischen Bundesamtes. Statistisches Bundesamt. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf
- Rat Für Sozial- und Wirtschaftsdaten. (2023). Standardfragenkatalog zur Erhebung soziodemographischer und krisenbezogener Variablen: Erstellt von der Arbeitsgruppe des RatSWD zur Vernetzung sozialwissenschaftlicher Forschung in gesellschaftlichen Krisen. RatSWD Output Series. <https://doi.org/10.17620/02671.76>
- Schneider, S. L., Ortmanns, V., Diaçó, A., & Müller, S. (2022). Die Erhebung sozio-demographischer Variablen in großen deutschen Umfragen: Ein Überblick über Möglichkeiten und Herausforderungen der Harmonisierung (Nr. 2/2022; KonsortSWD Working paper). GESIS - Leibniz-Institut für Sozialwissenschaften. <https://doi.org/10.5281/zenodo.6810973>

Kontakt und GESIS-Beratung



Kontakt: Sie können die Vortragenden über E-Mail kontaktieren:

Silke.Schneider@gesis.org

Lennart.Palm@gesis.org

GESIS-Beratung: GESIS bietet individuelle Beratungsleistungen zu einer Vielzahl von Themen – hierunter Umfragemethodik, Datenarchivierung oder digitale Verhaltensdaten & Computational Social Sciences – und in allen Phasen des Forschungsdatenzyklus.

Besuchen Sie uns unter www.gesis.org; weitere Informationen zu unseren Angeboten und den jeweiligen Nutzungsbedingungen finden Sie [hier](#).

Weitere Vorträge in der Reihe

- Informationen zu weiteren Vorträgen erhalten Sie hier:
 - [Meet the Experts - Expert*innen geben Einblicke in aktuelle Forschungsthemen – GESIS](#)

Das hierarchische Codeschema am Beispiel Familienstand

Die vollständige Standardvariable Familienstand

Im hierarchischen Codeschema werden Unterscheidungen je nach Wichtigkeit und Verfügbarkeit in Umfragedaten auf drei Stellen verteilt:

- Die erste Stelle bildet die vier grundlegenden Unterscheidungen (ledig, verheiratet, geschieden, verwitwet) ab.
- Die zweite Stelle ergänzt für Verheiratete Personen, ob sie zusammenleben oder getrennt (d.h. im Trennungsjahr) sind.
- Die dritte Stelle bildet ab, ob es sich rechtlich um eine Ehe handelt oder um eine eingetragene Lebenspartnerschaft.

Wert	Wertelabels
100	Ledig
200	Verheiratet (auch eingetragene Lebenspartnerschaft) - nfs
201	Verheiratet, nfs
202	Eingetragene Lebenspartnerschaft, nfs
210	Verheiratet (auch eingetragene Lebenspartnerschaft), zusammenlebend, nfs
211	Verheiratet, zusammenlebend
212	Eingetragene Lebenspartnerschaft, zusammenlebend
220	Verheiratet (auch eingetragene Lebenspartnerschaft), getrennt lebend, nfs
221	Verheiratet, getrennt lebend
222	Eingetragene Lebenspartnerschaft, getrennt lebend
300	Geschieden (auch eingetragene Lebenspartnerschaft aufgehoben), nfs
301	Ehe geschieden
302	Eingetragene Lebenspartnerschaft aufgehoben
400	Verwitwet (auch eingetragene/r /r Lebenspartner/in verstorben) , nfs
401	Ehepartner/-in verstorben
402	Eingetragene/r Lebenspartner/-in verstorben