

Daten- und Methodenbericht
Juli 2024

Karsten Becker | Ulrike Schwabe | Daniel Völk | Jonas Koopmann | Frederike Gerdes |
Theresa Schommer | Thorsten Euler | Tina Oestreich | Daniel Klein

Die Studierenden- befragung in Deutschland 2021

Daten- und Methodenbericht zur Erhebung

Dieses Werk steht unter der Creative Commons Namensnennung – Nicht kommerziell – Weitergabe unter gleichen Bedingungen 3.0 Deutschland Lizenz (CC-BY-NC-SA)

<https://creativecommons.org/licenses/by-nc-sa/3.0/de/>



Autor*innen:

Karsten Becker
Ulrike Schwabe
Daniel Völk
Jonas Koopmann
Frederike Gerdes
Theresa Schommer
Thorsten Euler
Tina Oestreich
Daniel Klein

Unter Mitarbeit von:

Mareike Beuße
Lennart Damaske
Marie-Christin Ehrhardt
Martina Kroher
Julia Steinkühler

Impressum

Herausgegeben von

Deutsches Zentrum für Hochschul- und
Wissenschaftsforschung GmbH (DZHW)
Lange Laube 12 | 30159 Hannover | www.dzhw.eu
Postfach 2920 | 30029 Hannover
Tel.: +49 511 450670-0 | Fax: +49 511 450670-960

Geschäftsführung

Prof. Dr. Monika Jungbauer-Gans
Axel Tscherniak

Vorsitzender des Aufsichtsrats

Ministerialdirigent Peter Greisler

Registergericht

Amtsgericht Hannover | B 210251
Umsatzsteuer-Identifikationsnummer:
DE291239300

Juli 2024



Inhaltsverzeichnis

Tabellen-/Abbildungsverzeichnis	II
1 Anlage und Inhalt der Studie	3
2 Erhebungsinstrument	5
2.1 Befragungsdesign und Befragungsinhalte	5
2.2 Pretests	7
3 Grundgesamtheit, Feldzugang und Stichprobenverfahren	9
4 Durchführung der Erhebung	11
5 Rücklauf	14
6 Datenaufbereitung	17
6.1 Datenübertragung, Datenprüfung und Datenbereinigung	17
6.2 Vergabe von Variablennamen, Variablenlabel und Wertelabel	21
6.3 Codierung fehlender Werte	21
6.4 Codierung offener Angaben	23
6.5 Generierung von Variablen	24
6.6 Erstellung des Datensatzes	24
7 Gewichtung	25
7.1 Konzeptionelles Vorgehen	25
7.2 Gewichtungsverfahren und Anwendungshinweise	28
8 Anonymisierung	29
9 Literatur	31
Anhang	36
Sekundär genutzte Instrumente	36
Übersicht generierte Variablen	44

Tabellen-/Abbildungsverzeichnis

Abbildung 1: „Die Studierendenbefragung in Deutschland“ (2021) - Übersicht über die Themenfelder und den modularisierten Befragungsaufbau	6
Tabelle 1: Hochschulbeteiligung an „Die Studierendenbefragung in Deutschland“ (2021)	10
Tabelle 2: Kontaktierung der Studierenden für „Die Studierendenbefragung in Deutschland“ (2021)	10
Tabelle 3: Versandzeitraum der Erhebungsunterlagen.....	12
Tabelle 4: Befragungsdauer „Die Studierendenbefragung in Deutschland“ (2021)	13
Abbildung 2: „Die Studierendenbefragung in Deutschland“ (2021) - Entwicklung des Bruttorücklaufs über die gesamte Feldphase (in absoluten Zahlen)	14
Tabelle 5: Teilnehmende Hochschulen, Stichprobengröße und Rücklauf	16
Tabelle 6: Als bedeutend definierte Variablen des A-Moduls	18
Tabelle 7: Übersicht der straightlining-relevanten Fragebogenseiten in „Die Studierendenbefragung in Deutschland“ (2021)	19
Tabelle 8: Systematik des FDZ-DZHW für fehlende Werte.....	21
Tabelle 9: Verteilung der GewichtungsvARIABLEN, Daten der amtlichen Statistik, ungewichteter und gewichteter Analysedatensatz (in %)	26
Fortsetzung Tabelle 9: Verteilung der GewichtungsvARIABLEN, Daten der amtlichen Statistik, ungewichteter und gewichteter Analysedatensatz (in %).....	27
Tabelle 10: Bereitgestellte Gewichte der Studie „Die Studierendenbefragung in Deutschland“ (2021)	28
Tabelle 11: Zugangswege im FDZ-DZHW	29
Tabelle 12: Quellennachweise zu eingesetzten Erhebungsinstrumenten	36
Tabelle 13: Bedeutende, vom Primärforschungsprojekt generierte Variablen - im Datensatz verfügbar	44
Tabelle 14: Bedeutende, vom Primärforschungsprojekt generierte Hilfs-/Berichtsvariablen - nicht im Datensatz verfügbar.....	52
Tabelle 15: Vom Forschungsdatenzentrum (FDZ-DZHW) im Rahmen der Anonymisierung generierte Variablen - im Datensatz verfügbar.....	58

1 Anlage und Inhalt der Studie

[Studiendesign] „Die Studierendenbefragung in Deutschland“ (2021) ist eine neu konzipierte Studie, die drei bisher separat durchgeführte Studierendenbefragungen integriert (Beuße et al., 2022): die Sozialerhebung (Middendorff et al., 2017; Middendorff & Wallis, 2023), den Studierendensurvey (Multrus et al., 2017; Multrus, 2021) und die Befragung „best - Studieren mit einer gesundheitlichen Beeinträchtigung“ (Poskowsky et al., 2018; Unger et al., 2012).¹ Über einen eigenen Befragungsstrang werden zudem die deutschen Daten für das europäische Bildungsmonitoring EUROSTUDENT erhoben (DZHW, 2018; Cuppen et al., 2023).²

Die Studie „Die Studierendenbefragung in Deutschland“ (2021) ist als Querschnitterhebung mit Langzeitcharakter konzipiert. Konkret heißt dies, dass die Querschnittsbefragungen in einem Abstand von vier Jahren regelmäßig durchgeführt werden. Mit der dadurch entstehenden Studienreihe lassen sich perspektivisch zeitliche Veränderungen der Situation von Studierenden in Deutschland abbilden. Damit setzt die neue, integrierte Studierendenbefragung die lange Tradition der Sozialerhebung und des Studierendensurveys als wichtige querschnittliche Dauerbeobachtungsstudien zur Beschreibung und Analyse der hochschulischen Bildung und der Studierendenpopulation in Deutschland sowie die deutlich kürzere Tradition der bisherigen „best-Studien“ fort.

[Inhaltliches Programm] In der Erhebung wurde eine Vielzahl von Themen zur Studiensituation und den unterschiedlichen Lebenslagen von Studierenden in Deutschland erfasst. Das inhaltliche Programm ist sehr vielfältig, da die thematischen Schwerpunkte der bisher separat durchgeführten Befragungen integriert wurden. In der Tradition der Sozialerhebung wird die wirtschaftliche und soziale Lage der Studierenden in den Blick genommen (Middendorff, 2022). Dieser Fokus wird um Themenbereiche des bisherigen Studierendensurveys ergänzt (Multrus et al., 2017). Zudem ermöglicht das neue Studiendesign erstmals Vergleiche zwischen Studierenden mit und ohne gesundheitliche studienerschwerende Beeinträchtigung (Steinkühler et al., 2023). Durch eine solche Zusammenführung von Befragungsinhalten können Querverbindungen zwischen unterschiedlichen Themenfeldern hergestellt werden. Daraus ergibt sich ein hohes Analysepotenzial. Nähere Erläuterungen zum Befragungsprogramm im Sommersemester 2021 sind Kapitel 2 zu entnehmen.

[Repräsentativität] Nicht alle Studierenden, die im Sommersemester 2021 an einer staatlichen, kirchlichen oder privaten Hochschule in Deutschland immatrikuliert waren, haben eine Einladung zur Befragung erhalten bzw. sind dieser gefolgt. Ein Grund hierfür ist unter anderem, dass an einigen Hochschulen vor dem Einladungsversand eine Stichprobe gezogen wurde. Um dennoch repräsentative Aussagen über Studierende in Deutschland im Sommersemester 2021 zu ermöglichen, wurden die Befragungsdaten in einem zweistufigen Verfahren gewichtet. Um die unterschiedliche Inklusionswahrscheinlichkeit von Studierenden in der Befragung durch die Stichprobenziehung zu berücksichtigen, wurde zuerst ein Design-Gewicht erstellt. Im zweiten Schritt wurde mit einem Redressment-Gewicht zusätzlich eine Adjustierung des Design-Gewichts an die Daten der amtlichen Statistik

¹ Aktuelle Informationen zu „Die Studierendenbefragung in Deutschland“ (2021) können über die Projektwebsite am DZHW abgerufen werden: https://www.dzhw.eu/forschung/projekt?pr_id=650.

² Jede*r 18. Befragungsteilnehmer*in wurde in einen separaten Befragungsstrang zur Erhebung der deutschen EUROSTUDENT-Daten geleitet. Die EUROSTUDENT-Daten werden ebenfalls über das FDZ-DZHW verfügbar gemacht.

vorgenommen. Grundlage für die Redressment-Gewichtung sind Merkmale, die in der amtlichen Hochschulstatistik verfügbar sind. Nähere Erläuterungen zur Gewichtung sind Kapitel 7 zu entnehmen.

[Nutzungspotenzial] Mit der Studie „Die Studierendenbefragung in Deutschland“ (2021) wird eine umfassende und für Deutschland einmalige Datenbasis für Forschung sowie Bildungs- und Hochschulpolitik geschaffen (Beuße et al., 2022). Mit den Daten lässt sich eine Vielzahl wissenschaftlich sowie politisch relevanter Fragestellungen beantworten. Bereits entstanden sind auf Basis der Daten von „Die Studierendenbefragung in Deutschland“ (2021) der Bericht zur 22. Sozialerhebung (Kroher et al., 2023), ein Bericht zur Studiensituation unter Onlinebedingungen (Multrus et al., 2023) sowie der Bericht „best3 – Studieren mit einer gesundheitlichen Beeinträchtigung“ (Steinkühler et al., 2023). Diese sind für das hochschulische Bildungsmonitoring hochrelevant. In ergänzenden Policy-Paper wurden zudem aktuelle hochschulpolitische Fragen zu populistischen Tendenzen unter Studierenden (Multrus et al., 2022), zu Diskriminierungserfahrungen an Hochschulen (Meyer et al., 2022), zu den Auswirkungen der Inflation auf die wirtschaftliche Lage der Studierenden (Meier et al., 2023), zur Attraktivität von Masterstudiengängen (Stefani et al., 2023) und zu den Folgen der Coronapandemie für studentische Auslandsmobilität (Netz & Völk, 2023) adressiert.

2 Erhebungsinstrument

In der Studie „Die Studierendenbefragung in Deutschland“ (2021) wurde als Erhebungsinstrument ein standardisierter Online-Fragebogen in deutscher und englischer Sprache eingesetzt. In Abschnitt 2.1 werden das Befragungsdesign, die zentralen Themenfelder sowie die sekundär genutzten Bestandteile des Erhebungsinstruments dargestellt. In Abschnitt 2.2 werden die zur Prüfung und Optimierung der Befragung durchgeführten kognitiven und quantitativen Pretests beschrieben.

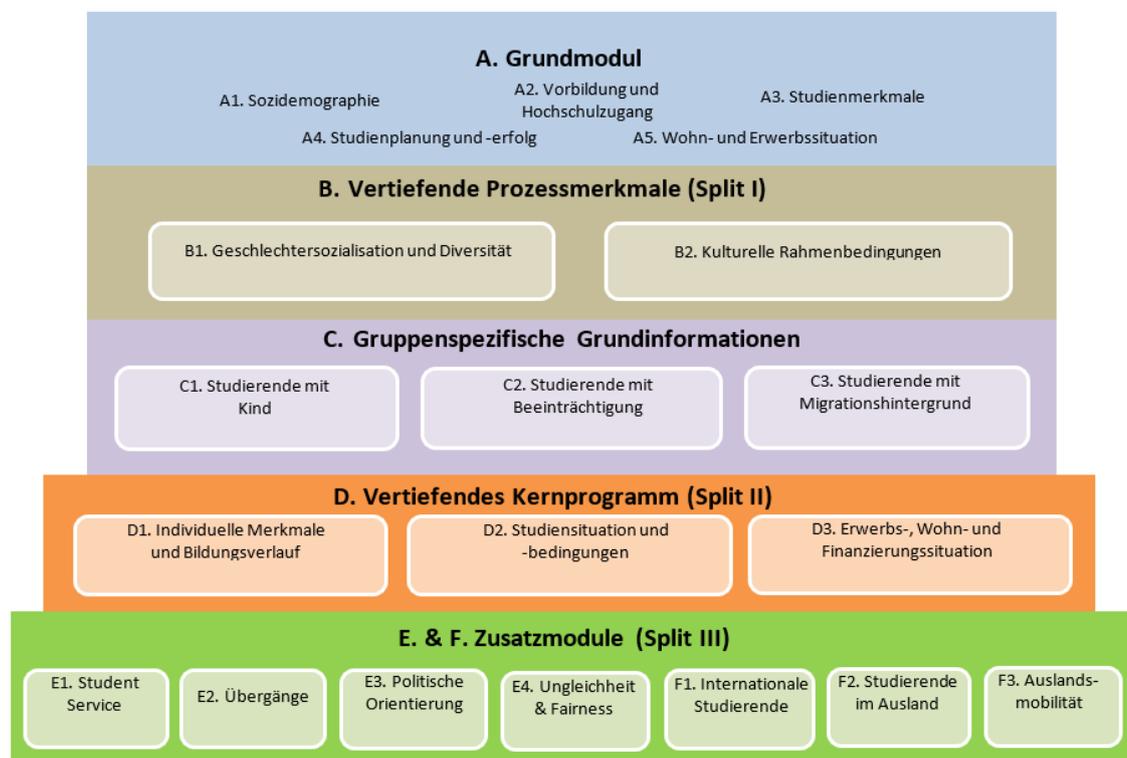
2.1 Befragungsdesign und Befragungsinhalte

[Modularisiertes Befragungs- und Splitdesign] Aus der Integration der bisher separat durchgeführten Studien (siehe Kapitel 1) ergibt sich eine Vielfalt an Befragungsinhalten. Dem gegenüber steht der zeitliche Aufwand für die Befragungsteilnehmer*innen zum Beantworten aller Fragen. Um ein ausgewogenes Verhältnis zwischen Themenvielfalt und Befragungslast zu realisieren, wurde ein modularisiertes Befragungs- und Splitdesign implementiert (siehe Abbildung 1).

[Inhalte und Aufbau des Befragungsprogramms] Abbildung 1 gibt einen Überblick über das Befragungsprogramm von „Die Studierendenbefragung in Deutschland“ (2021). Es setzt sich zusammen aus einem umfassenden Grundmodul A, dessen Fragen an alle Studierenden gingen, ergänzt um gruppenspezifische Zusatzmodule (z. B. für Studierende mit Kind oder Studierende mit einer studienerschwerenden gesundheitlichen Beeinträchtigung), sowie randomisiert zugeteilte thematische Kern- und Zusatzmodule. Diese Kombination aus einem gemeinsamen Grundstock an Fragen, Split-Design und spezifischen Zusatzmodulen ermöglichte es, dass sowohl für das Bildungsmonitoring als auch für die wissenschaftliche Forschung ein breit angelegtes Themenspektrum abgedeckt wurden. Konkret wurden in den Modulen folgende Informationen erhoben: In dem A-Modul wurden zentrale Basis- und Kontextinformationen zu den Studierenden und ihrem Studium wie soziodemografische Merkmale, Studienmerkmale, Vorbildung und Hochschulzugang der Studierenden, Studienplanung und -erfolg, aber auch Basisinformationen zur Wohn- und Erwerbssituation sowie individuelle Merkmale und die Studienverläufe der Studierenden im Grundprogramm erfasst. Dieses Modul wurde durch themenspezifische Kernprogramme zu (a) der Studiensituation und den Studienbedingungen und (b) zur Erwerbs-, Wohn- und Finanzierungssituation der Studierenden sowie Module für die Gruppen (a) Studierende mit Kind(ern), (b) Studierende mit gesundheitlicher Beeinträchtigung, (c) Studierende mit Migrationshintergrund ergänzt. In zwei weiteren Zusatzmodulen standen zum einen Themen wie Serviceangebote für Studierende, Übergänge ins (weiterführende) Studium, politische Orientierungen und Einstellungen zu Ungleichheit und Fairness im Mittelpunkt. Zum anderen wurden die Studierenden entsprechend ihrer jeweiligen Auslandserfahrung bzw. -pläne und des Orts des Erwerbs der Hochschulzugangsberechtigung (in Deutschland oder im Ausland) einem der drei Teilmodule „Internationale Studierende“, „Studierende im Ausland“ oder „Auslandsmobilität“ zugeordnet und zu bereits absolvierten und geplanten Auslandsaufenthalten befragt.

[Randomisierte Zuteilung im Splitdesign] Um die Selbstselektion der Befragten in verschiedene Themenfelder zu vermeiden und gleichzeitig hinreichend große Fallzahlen für spätere Analysen zu realisieren, wurde ein Befragungsdesign entwickelt, auf dessen Basis die Befragten zufällig den verschiedenen Fragebogensplits (B-Module, D-Module und E- & F-Module) nach vorab definierten Quoten zugewiesen wurden.³ Die Zuweisung der Befragten erfolgte über die beiden Zufallsvariablen „mastersplit“ und „zusatzsplit“. Das Design sieht vor, dass im B-Modul des Grundprogramms (Split I) eine Hälfte der Befragten in das Modul B1 „Geschlechtersozialisation und Diversität“ und die andere Hälfte in das Modul B2 „Kulturelle Rahmenbedingungen“ geleitet wird. Im vertiefenden Kernmodul (Split II) wurden die Befragten jeweils zwei der drei D-Module zugewiesen, das heißt, dass *entweder* die Fragen aus D1 und D2, D1 und D3 *oder* D2 und D3 beantwortet wurden. In den Zusatzmodulen E und F (Split III) erfolgte die Zuteilung noch feingliederiger. Zum Teil wurde dort die zufallsbasierte Zuordnung in ein Modul mit einer filterbedingten Frageführung innerhalb des Moduls kombiniert. Befragte des Moduls E2 „Übergänge“ erhielten beispielsweise Folgefragen zu „Masterstudium“ (E2M), „Promotion“ (E2P) oder „Studienanfang“ (E2S) in Abhängigkeit ihrer jeweiligen Studiensituation (siehe Erhebungsinstrument). Die Zuweisung in das F-Modul erfolgte ebenfalls zufallsbasiert, wobei die Befragten in diesem Fall zuerst den gemeinsamen Fragenkatalog der Fragen F1_1 bis F1_5 (siehe Erhebungsinstrument) durchliefen. Anschließend wurden diese filterbedingt jeweils einem der drei Module F1 „Internationale Studierende“, F2 „Studierende im Ausland“ und F3 „Auslandsmobilität“ zugeteilt. Ausgenommen des Grundmoduls A unterscheidet sich somit designbedingt die Anzahl der auswertbaren Fälle zwischen den einzelnen Modulen in Abhängigkeit der zugewiesenen Befragten. Dies sollte bei Analysen mit Variablen aus unterschiedlichen Modulen berücksichtigt werden.

Abbildung 1: „Die Studierendenbefragung in Deutschland“ (2021) - Übersicht über die Themenfelder und den modularisierten Befragungsaufbau



Quelle: „Die Studierendenbefragung in Deutschland“ (2021).

³ Um für einzelne Subgruppen wie zum Beispiel Studierende im Ausland eine ausreichend hohe Fallzahl generieren zu können, wurden deren Anteile in früheren Befragungen der Sozialerhebung bei der Quotierung berücksichtigt.

[Vignettenstudien] Die Studie „Die Studierendenbefragung in Deutschland“ (2021) zeichnet sich auch dadurch aus, dass neben dem modularisierten Befragungs- und Splitdesign im Modul E2 „Übergänge“ zwei Vignettenstudien implementiert wurden: die erste zum Bachelor-Master-Übergang (Stefani et al., 2023) und die zweite zum Übergang vom Master- ins Promotionsstudium.⁴ Bei Vignettenstudien handelt es sich um eine Forschungsmethode, bei der den Befragten hypothetische Szenarien, z. B. zum Übertrittsverhalten ins Masterstudium, vorgelegt werden (Auspurg & Hinz, 2015; Petzold & Netz, 2022). Dabei werden Merkmale der skizzierten Szenarien systematisch variiert, den Befragten zufällig zugewiesen und das Antwortverhalten in Abhängigkeit von der vorgelegten Merkmalsausprägungen analysiert (Stefani et al., 2023). Ein solches Survey-Experiment vereint die Vorteile von Befragungen mit denen von Zufallsexperimenten (Petzold & Netz, 2022). Den Befragten wurden 432 unterschiedliche Vignetten zum Bachelor-Master-Übergang (jeweils 108 Abfolgen von 4 verknüpften Vignetten) bzw. 324 (162 Abfolgen von 2 Vignetten) zum Übergang vom Master- ins Promotionsstudium vorgelegt.

[Sekundär genutzte Bestandteile des Erhebungsinstruments] Bei der Konzeption des Erhebungsinstruments wurde überwiegend auf etablierte Skalen und Fragen zurückgegriffen. In Tabelle 12 im Anhang findet sich eine Übersicht der sekundär genutzten Bestandteile des Erhebungsinstruments der Studie „Die Studierendenbefragung in Deutschland“ (2021).

2.2 Pretests

[Ziel und Verfahren] Das Erhebungsinstrument von „Die Studierendenbefragung in Deutschland“ (2021) wurde vor seinem Einsatz im Feld umfassenden Pretests unterzogen. Es wurden mit kognitiven Pretests, Expert*innen-Pretests und einem quantitativen Pretest verschiedene Pretest-Verfahren (Weichbold, 2022) angewendet.

[Kognitive Pretests] Kognitive Pretests zielen darauf ab, zu testen, wie Befragte Fragen und Begriffe interpretieren, welche Informationen sie abrufen, wie sie zu einer Entscheidung kommen und wie sie diese letztlich einer der vorgegebenen Antwortkategorien zuordnen (Lenzner et al., 2015; Prüfer & Rexroth, 2000). Mit diesem aktiven Pretest-Verfahren werden Einzelfragen geprüft. Indem untersucht wird, welche kognitiven Prozesse bei den Befragten aktiviert werden, können Informationen zur Verständlichkeit von Fragen und Begriffen gewonnen bzw. Probleme identifiziert werden. Für die kognitiven Pretests konnten Studierende an der Universität Hannover rekrutiert werden.

[Expert*innenpretests] Zusätzlich wurde die Expertise interner und externer Expert*innen aus dem Bereich der Bildungs- und Hochschulforschung zu ausgewählten Instrumenten wie beispielsweise „Akademische und soziale Integration“ nach Tinto (1975) eingeholt.

⁴ Die Daten aus den Vignettenstudien werden nicht im Personendatensatz, sondern in separaten Datensätzen für Sekundäranalysen zur Verfügung gestellt. Das Zusammenführen der Daten ist über einen Personen-Identifizier (pid) in beiden Datensätzen möglich.

[Quantitativer Pretest] Ein anschließender quantitativer Pretest zielte darauf ab, das vorläufige Erhebungsinstrument im Feld unter ähnlichen Bedingungen⁵ wie die geplante Haupterhebung im Sommersemester 2021 zu testen. Einerseits wurde die Implementierung des modularisierten Splitdesigns, die Dauer der Befragung und der generelle Aufbau der Online-Befragung getestet. Für den quantitativen Pretest wurde der Fragebogen mit der DZHW-eigenen Onlinebefragungssoftware *Zofar* programmiert (Jungbauer-Gans, 2018). Da diese Befragungssoftware auch für die Haupterhebung im Sommersemester 2021 genutzt wurde, konnte so bereits vor der Durchführung der Befragung ein Ablauf- und Funktionspretest erfolgen. Zusätzlich wurde das responsive Design der Befragungssoftware, d. h. die Darstellung der Fragen auf mobilen Endgeräten mit kleinen Bildschirmen getestet und optimiert.

Darüber hinaus wurden die im Rahmen des quantitativen Pretests erhobenen Daten verwendet, um die eingesetzten Instrumente hinsichtlich der Gütekriterien Reliabilität und Validität zu überprüfen. So wurde beispielsweise das Instrument zur Erfassung der „Lernumwelt Hochschule“ (Schaeper & Weiß, 2016) auf seine faktorielle Struktur und die Subskalen auf ihre interne Konsistenz geprüft. Für den quantitativen Pretest wurde das HISBUS-Studierendenpanel am DZHW genutzt.⁶ Die Feldphase erstreckte sich über drei Wochen im Wintersemester 2019/2020 (07.11.2019–28.11.2019). An der Befragung nahmen insgesamt 1.779 Studierende teil.

Zudem wurden die im Rahmen des quantitativen Pretests erhobenen Paradata ausgewertet. Es handelt sich bei Paradata um prozessgenerierte Daten in Onlinebefragungen, beispielsweise Angaben zum Zeitpunkt der Befragungsteilnahme, Informationen zum genutzten Endgerät oder der Seitenverweildauer (Kreuter, 2013). Die mit der Onlinebefragungssoftware *Zofar* erfassten serverseitig und clientseitig generierten Paradata geben Hinweise dazu, *wie* an der Befragung teilgenommen wurde und damit auch zur Datenqualität (Couper, 1998; Felderer et al., 2019). Die Auswertung dieser Informationen gibt beispielsweise Auskunft über den Bearbeitungsprozess, aber auch über Muster der Teilnahme und den Rekrutierungsprozess. Im Rahmen des im Projekt durchgeführten quantitativen Pretests wurden Zeitstempel und Seitenverweildauern ausgewertet, um die Dauer der Beantwortung des Fragebogens insgesamt und für einzelne Module zu schätzen.

⁵ Die gleiche Befragtengruppe (Studierende an deutschen Hochschulen) wird mit dem gleichen Erhebungsmodus (standardisierte Onlinebefragung, CAWI) befragt. Um den Beantwortungsprozess nicht zu beeinflussen, wurden die Befragten nicht darüber informiert, dass es sich um einen Pretest handelt.

⁶ Das HISBUS-Studierendenpanel war ein Online-Access-Panel, welches vom 01.08.2000 bis zum 31.12.2020 am DZHW im Einsatz war (https://www.dzhw.eu/forschung/projekt?pr_id=273). Dafür wurden aus den regelmäßigen Studien des DZHW Studierende rekrutiert, die mehrmals im Jahr zu verschiedenen Themen befragt wurden. Das HISBUS-Studierendenpanel wurde zum Zweck der Politikberatung und Studierendenforschung mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und unter methodischer Expertise des ZUMA Mannheim aufgebaut (nähere Informationen: http://www.hisbus.de/about/index_html).

3 Grundgesamtheit, Feldzugang und Stichprobenverfahren

[Grundgesamtheit] Die Grundgesamtheit der Studie „Die Studierendenbefragung in Deutschland“ (2021) bilden alle Studierenden, die im Sommersemester 2021 an einer Hochschule in Deutschland immatrikuliert waren, mit Ausnahme Studierender an Verwaltungsfachhochschulen. Zur Grundgesamtheit zählen folglich Studierende an staatlichen, kirchlichen und privaten Hochschulen. Anders als in vorherigen Befragungen im Rahmen der Sozialerhebung (Becker et al., 2019; Middendorff et al., 2017) sind Studierende an Fernhochschulen ebenfalls Teil der Grundgesamtheit. Dies trifft auch auf Studierende an dualen Hochschulen zu. Auf die Einbeziehung deutscher und bildungsinländischer Promotionsstudierender wurde, mit Ausnahme von promovierenden Studierenden, die zusätzlich in einem regulären Studiengang immatrikuliert sind (z. B. Medizin), verzichtet. Bildungsausländische Promotionsstudierende hingegen zählen weiterhin zur Grundgesamtheit, da sie eine relevante Gruppe unter den internationalen Studierenden bilden (Beuße et al., 2022).

[Feldzugang] Da es in Deutschland kein zentrales, bundesweites Studierendenregister mit Kontaktdaten gibt, erfolgte der Feldzugang über die einzelnen Hochschulen (Beuße et al., 2022). In einem ersten Schritt wurden deshalb alle Hochschulen der definierten Grundgesamtheit zur Teilnahme an der Studie „Die Studierendenbefragung in Deutschland“ (2021) eingeladen. Auf Ebene der Hochschulen handelt es sich demzufolge um eine Vollerhebung. Basierend auf der amtlichen Statistik wurden insgesamt 396 deutsche Hochschulen gebeten, sich an der Studie „Die Studierendenbefragung in Deutschland“ (2021) zu beteiligen (Beuße et al., 2022). An diesen Hochschulen waren im Sommersemester 2021 rund 2,7 Millionen Studierende immatrikuliert (Statistisches Bundesamt, 2022). Es handelt sich hierbei sowohl um Studierende, die ihre Hochschulzugangsberechtigung in Deutschland als auch um Studierende, die ihre Hochschulzugangsberechtigung außerhalb Deutschlands erworben haben.

Insgesamt haben sich 250 der 396 eingeladenen Hochschulen an der Studie „Die Studierendenbefragung in Deutschland“ (2021) beteiligt (siehe Tabelle 1). Dies entspricht einer Beteiligungsquote auf Ebene der Hochschulen von gut 63 Prozent. Unter den teilnehmenden Hochschulen befinden sich alle 30 staatlichen Universitäten und Hochschulen für angewandte Wissenschaften im Land Nordrhein-Westfalen. Diese Hochschulen haben mit der Teilnahme an dem Projekt „Die Studierendenbefragung in Deutschland“ (2021) gleichzeitig auch das landeseigene Projekt „Landesweite Studierendenbefragung an den Hochschulen in Nordrhein-Westfalen“ umgesetzt und sich deshalb vollständig an der Erhebung beteiligt (Gerdes & Völk, 2022). Mit etwa 2,3 Millionen immatrikulierten Studierenden repräsentieren alle teilnehmenden Hochschulen 84 Prozent der Studierenden der definierten Grundgesamtheit (Beuße et al., 2022; siehe Tabelle 2). Diese bilden den Auswahlrahmen für die anschließende Stichprobenziehung auf Ebene der Studierenden.

Tabelle 1: Hochschulbeteiligung an „Die Studierendenbefragung in Deutschland“ (2021)

Hochschulen der definierten Grundgesamtheit	Anzahl	Beteiligungsquote
Eingeladene Hochschulen	396	
Teilnehmende Hochschulen	250	
Teilnehmende auf Hochschulebene		63,1 %

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Berechnungen.

[Stichprobe] Für die Stichprobenziehung auf Ebene der Studierenden wurde in einem zweiten Schritt auf die aktuellen, semestergenauen Studierendenlisten der teilnehmenden Hochschulen zurückgegriffen. Mit einer Ziehungsquote von knapp 61 Prozent wurden ca. 1,4 Millionen Studierende zur Teilnahme an der Studie „Die Studierendenbefragung in Deutschland“ (2021) eingeladen (siehe Tabelle 2). Die Ziehungsquote variierte jedoch zwischen den Hochschulen je nach gewähltem Vorgehen: An allen teilnehmenden Hochschulen wurde zufallsbasiert mindestens jede*r dritte Studierende zur Befragungsteilnahme eingeladen. An den meisten teilnehmenden Hochschulen wurde jedoch eine Vollerhebung umgesetzt. Dies hatte zwei Gründe: Zum einen wurde angestrebt, auch an eher kleinen Hochschulen eine ausreichend hohe Fallzahl zu generieren. Zum anderen war es für einige Hochschulen aus organisatorischen Gründen einfacher, eine Vollerhebung zu realisieren, d. h. alle im Sommersemester 2021 immatrikulierten Studierenden zu kontaktieren.

Tabelle 2: Kontaktierung der Studierenden für „Die Studierendenbefragung in Deutschland“ (2021)

Studierende an Hochschulen der definierten Grundgesamtheit	Anzahl	Anteil
Studierende an eingeladenen Hochschulen (n = 396)	2.733.151	
Studierende an teilnehmenden Hochschulen (n = 250)	2.295.075	
Auswahlrahmen für die Stichprobenziehung auf Studierendenebene		84,0 %
Studierende nach Stichprobenziehung durch die teilnehmenden Hochschulen	1.397.326	
Kontaktierung auf Studierendenebene		60,9 %

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Berechnungen.

[Stichprobenziehung und Hochschulkooperation] Ein bundesweites (Kontaktdaten-)Register für Studierende ist aktuell nicht vorhanden. Aus diesem Grund nahmen die teilnehmenden Hochschulen als Adressmittlungsstelle zwischen dem Erhebungsprojekt und der Zielgruppe der ausgewählten Studierenden eine wichtige Gatekeeper-Funktion bei der Kontaktierung der Studierenden ein. Kurz vor dem Beginn der Feldphase zogen die Kontaktpersonen bzw. Ansprechpartner*innen in den Hochschulverwaltungen nach Anleitung die zufallsbasierte Stichprobe aus den ihnen vorliegenden Studierendenlisten— sofern an der entsprechenden Hochschule keine Vollerhebung realisiert wurde. Ein solches Vorgehen zur Kontaktierung der Studierenden setzt die Kooperationsbereitschaft und Unterstützung der Hochschulen voraus. Die generelle Maßgabe bei der Stichprobenziehung bestand darin, ein Drittel der Studierenden aus der definierten Grundgesamtheit einzuladen bzw. eine Vollerhebung an der jeweiligen Hochschule zu realisieren. Außerdem wurde ein seitens der Hochschulen vorab zu leistender Ausschluss bestimmter Studierendengruppen aus der zu ziehenden Stichprobe vermieden, indem z. B. deutsche Promotionsstudierende, die darüber hinaus *nicht* in einem regulären Studiengang eingeschrieben sind, über das Erhebungsinstrument zu Beginn der Befragung identifiziert und mittels entsprechender Filterführung aus der Befragung herausgeführt wurden.

4 Durchführung der Erhebung

[Erhebungsmodus] „Die Studierendenbefragung in Deutschland“ (2021) wurde als standardisierte Onlinebefragung (CAWI: Computer-Assisted Web Interview) implementiert (Wagner-Schelewsky & Hering, 2022). Dieser Erhebungsmodus ermöglichte die technische Implementierung des komplexen Befragungs- und Splitdesigns (siehe Kapitel 2) und die Realisierung eines großen Stichprobenumfangs (siehe Kapitel 3).

Das Erhebungsinstrument (siehe Kapitel 2) wurde mit dem DZHW-eigenen Onlinebefragungssystem *Zofar* programmiert und für *mixed-devices* optimiert (Jungbauer-Gans, 2018; Weiß et al., 2022). Durch dieses Vorgehen konnten die Befragten die Fragen sowohl am Desktop-PC als auch mit mobilen Endgeräten wie Tablets oder Smartphones beantworten. Letztere zeichnen sich vor allem dadurch aus, dass die Bildschirme im Vergleich zum Desktop-PC wesentlich kleiner und berührungssensitiv sind (Weiß et al., 2022).

[Hochschulkommunikation] Im Januar 2021 wurden die Leitungen und zum Teil weitere Ansprechpersonen von allen der Grundgesamtheit zugehörigen Hochschulen in Deutschland per E-Mail kontaktiert und ausführlich über das Projekt „Die Studierendenbefragung in Deutschland“ (2021) informiert. Alle Hochschulen wurden darum gebeten, sich an der Befragung zu beteiligen und die Organisation der Feldphase zu unterstützen (siehe Kapitel 3). In der Zeit zwischen Januar und Mai 2021 gab es anschließend eine umfangreiche Kommunikation mit den Hochschulen, die keine direkte Teilnahmebereitschaft signalisiert hatten. Das Projektteam am DZHW versuchte, auch diese zunächst unentschlossenen Hochschulen zu einer Teilnahme zu bewegen. Insgesamt hatten zum Befragungsbeginn 250 Hochschulen eine Teilnahmeerklärung unterzeichnet. Mit dieser Teilnahmeerklärung teilten die Hochschulen den Projektmitarbeiter*innen am DZHW zudem feste Ansprechpersonen in der Hochschuladministration mit, die die Befragung an den Hochschulen jeweils organisatorisch umsetzten.

[Erhebungsunterlagen] Um sicherzustellen, dass die Befragung so standardisiert wie möglich durchgeführt wird, wurde den teilnehmenden Hochschulen ein „Handbuch zur Durchführung der Befragung an den Hochschulen“ per E-Mail zugesandt. Dieses umfasste eine detaillierte Beschreibung der gesamten Befragungsorganisation, z. B. wie an den Hochschulen die Zufallsstichprobe zu ziehen sei und wie der Versand der Befragungslinks an die Studierenden technisch umgesetzt werden sollte. Die Hochschulen erhielten mit dem Handbuch außerdem das Formular „Kontaktdaten und Stichprobenhöhe“, das sie vor Beginn der Befragung ausfüllten und an das DZHW übermittelten. Mit dem Formular wurden aktualisierte Daten zu den Ansprechpersonen an den Hochschulen sowie die Anzahl der Studierenden an der Hochschule insgesamt und die Anzahl der zu befragenden Studierenden nach Stichprobenplan erfasst. Das Handbuch enthielt ein weiteres Formular, das sog. „Feldprotokoll“, mit dem die Hochschulen Angaben zur Anzahl der versendeten Befragungs-E-Mails, inkl. der Erinnerungsschreiben, dokumentierten. Das Feldprotokoll wurde am Ende der Feldphase an das Projektteam übermittelt und ist eine wichtige Voraussetzung für die Berechnung der Rücklaufquoten auf Ebene der Studierenden (siehe Kapitel 5).

[Technische Umsetzung] Um die Serverlast, die eine Befragung dieser Größenordnung über die Feldphase erzeugt, zu verteilen, wurde, in Absprache mit den teilnehmenden Hochschulen, gestaffelt zur Befragung eingeladen. Folglich erstreckte sich der Einladungsversand über einen Zeitraum von etwa zwei Wochen (siehe Tabelle 3).

Unter Anleitung der Projektmitarbeiter*innen am DZHW verschickten die Hochschulen an die ausgewählten Studierenden in einem ersten Schritt die Einladungsschreiben zur Teilnahme an der

Befragung. Ein personalisierter Zugangslink zur Onlinebefragung, sog. „token“, stellte sicher, dass die gleichen Studierenden nicht mehrfach an der Befragung teilnehmen konnten. Durch diesen personalisierten Zugangslink war es auch möglich, die Befragung zu unterbrechen und sie zu einem späteren Zeitpunkt fortzuführen.

[Feldphase] Die Feldphase der Befragung begann am 04. Mai 2021 und endete für alle Hochschulen am 19. September 2021 (siehe Tabelle 3). Ab dem 04. Mai 2021 verschickten die Hochschulen, gestaffelt nach Bundesland, über einen Zeitraum von zwei Wochen die Einladung per E-Mail zur Befragung an die Studierenden.

Nach dieser initialen Einladung zur Befragungsteilnahme wurden die Hochschulen außerdem dazu aufgefordert, Erinnerungsschreiben an diejenigen Studierenden zu senden, die sich bis zu diesem Zeitpunkt noch nicht an der Befragung beteiligt bzw. diese noch nicht beendet hatten. Dies diente der Motivation der Studierenden, an der Befragung teilzunehmen und damit den Rücklauf zu erhöhen (Göritz & Crutzen, (2012)). Zu insgesamt drei Zeitpunkten wurden den Hochschulen Textvorlagen für Erinnerungsschreiben sowie eine tagesaktuelle Liste derjenigen Befragungslinks zugeschickt, deren Empfänger*innen nicht erneut angeschrieben werden sollten, da diese den Fragebogen bereits vollständig ausgefüllt hatten (siehe Tabelle 3).

Einige Hochschulen baten aus internen organisatorischen Gründen darum, die Feldphase abweichend von den vorgesehenen Zeiträumen durchzuführen. Diesen wurden die Unterlagen zu einem abweichenden Zeitpunkt zugeschickt.

Tabelle 3: Versandzeitraum der Erhebungsunterlagen

Versandzeiträume ^a	
Beginn Feldphase: Versand Einladungsschreiben	04.05.2021–18.05.2021
Versand 1. Erinnerung	ca. 2 Wochen nach Feldstart
Versand 2. Erinnerung	ca. 4 Wochen nach Feldstart
Versand 3. Erinnerung	29.06.2021 (für alle Hochschulen)
Ende Feldphase: Befragung wird offline genommen	19.09.2021

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Darstellung.

Anmerkung: ^a = Teilweise wurde von den Zeiträumen abgewichen, um auf besondere Bedarfe der Hochschulen Rücksicht zu nehmen.

[Rücklaufsteigernde Maßnahmen] Die Rücklaufquoten von Onlinebefragungen fallen geringer aus als bei anderen Erhebungsmodi (Lozar Manfreda et al., 2008). Deshalb wurde die Befragung von einer intensiven Kommunikationsarbeit begleitet. Während der Feldphase wurden die Hochschulen vom DZHW, aber auch vom Deutschen Studierendenwerk (DSW) und den regionalen Studierendenwerken dazu ermuntert, über verschiedene Kommunikationskanäle auf das Projekt „Die Studierendenbefragung in Deutschland“ (2021) aufmerksam zu machen und dieses aktiv zu bewerben. Dazu wurden z. B. Logos für Postings zur Verfügung gestellt, die die Hochschulen in ihren Social-Media-Kanälen (Twitter, Facebook und Instagram) veröffentlichen konnten. Da das Sommersemester 2021 immer noch von der Coronapandemie geprägt war, wurden ausschließlich digitale Werbeformate angeboten und keine Printprodukte verbreitet. Mit dem Podcast-Format „Der Podcast der Studierendenbefragung in Deutschland“ wurden Themen der Studierendenbefragung mit Vertreter*innen aus Wissenschaft und Politik diskutiert und die Befragung und ihre Relevanz öffentlich bekannt gemacht.⁷

Die Teilnahme an der Studie „Die Studierendenbefragung in Deutschland“ (2021) wurde nicht incentiviert. Allerdings wurden für besonders kleine, aber wichtige Teilgruppen (z. B. Studierende ohne schulische Hochschulzugangsberechtigung) 10-Euro-Gutscheine angeboten, wenn diese bereit waren, ihre Kontaktdaten (E-Mail-Adressen) am Ende der Befragung anzugeben. Ziel dieser Postpaid-

⁷ Die Podcast-Folgen sind unter <https://einfueralle.podigee.io/episodes> abrufbar.

Incentivierung war es, diese Teilgruppen für nachfolgende DZHW-Befragungen erneut anschreiben zu können.

[Befragungsdauer] Für Onlinebefragungen gibt es die Empfehlung, eine maximale Befragungsdauer von etwa 20 Minuten nicht zu überschreiten, da sich die Länge negativ auf die Teilnahme und das Antwortverhalten auswirken kann (Galesic & Bosnjak, 2009). Aufgrund des Splitdesigns der Befragung erfolgt die Berechnung der durchschnittlichen Befragungsdauer auf Ebene der einzelnen Module, die dann addiert werden (siehe Tabelle 4). Im Mittel haben die Befragten 32,3 Minuten (Median: 30,9 Minuten) benötigt, um die Befragung auszufüllen. Werden nur diejenigen Befragten betrachtet, die den Fragebogen vollständig beantwortet haben, beträgt der Mittelwert für diese Gruppe 43,6 Minuten, der Median liegt bei 39,2 Minuten

Tabelle 4: Befragungsdauer „Die Studierendenbefragung in Deutschland“ (2021)

Modul	Mittelwert (in Minuten)		Median (in Minuten)	
	insgesamt ^a	beendet ^b	insgesamt ^a	beendet ^b
A. Grundmodul	9,1	11,6	8,3	10,0
B. Vertiefende Prozessmerkmale				
B1. Geschlechtersozialisation und Diversität	4,2	5,0	3,7	4,3
B2. Kulturelle Rahmenbedingungen	5,0	5,5	4,3	4,7
C. Gruppenspezifische Grundinformationen				
C1. Studierende mit Kind	2,2	2,2	1,8	1,9
C2. Studierende mit Beeinträchtigung	3,4	3,5	2,8	3,0
C3. Studierende mit Migrationshintergrund	0,4	0,4	0,3	0,3
D. Vertiefendes Kernprogramm				
D1. Individuelle Merkmale und Bildungsweg	5,0	6,1	4,2	6,0
D2. Studiensituation und -bedingungen	9,9	11,6	8,6	9,8
D3. Erwerbs-, Wohn- und Finanzierungssituation	11,0	11,7	9,3	9,9
E. Zusatzmodul				
E1. Student Services	5,0	5,4	4,6	4,8
E2. Übergänge	3,9	4,0	3,2	3,3
E3. Politische Orientierung	4,4	4,6	3,8	3,9
E4. Ungleichheit & Fairness	6,0	6,3	5,2	5,4
F. Zusatzmodul				
F1. Internationale Studierende	6,2	6,3	5,5	5,6
F2. Studierende im Ausland	4,7	4,8	4,2	4,3
F3. Auslandsmobilität	2,3	2,3	1,9	1,9
Befragung insgesamt	32,3	43,6	30,9	39,2

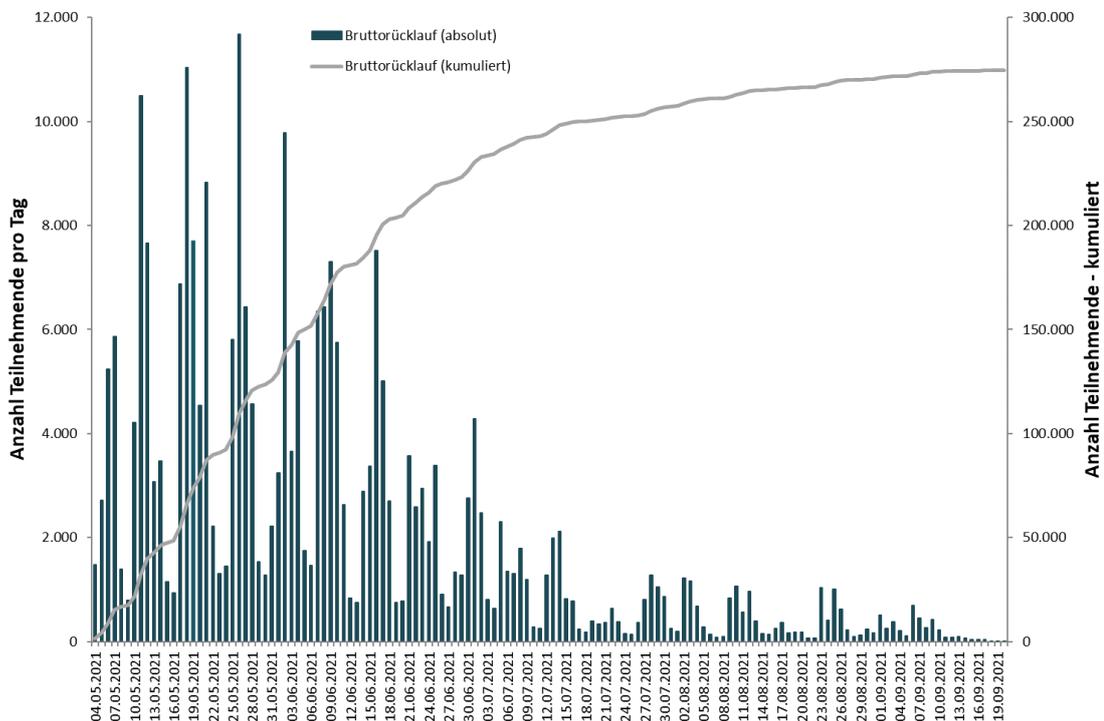
Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Berechnungen.

Anmerkung: ^a = Alle Befragten werden berücksichtigt; ^b = Es werden nur die Befragten berücksichtigt, die auf alle Fragen des jeweiligen Moduls geantwortet haben.

5 Rücklauf

[Rücklaufentwicklung] Während der Feldphase wurden von den 250 teilnehmenden Hochschulen insgesamt 1.397.326 Einladungen zur Teilnahme an der Befragung verschickt (siehe Tabelle 2). Abbildung 2 zeigt den Brutorücklauf in absoluten Zahlen für jeden dritten Tag während der Feldphase und kumuliert über den gesamten Erhebungszeitraum Feldphase. Erwartungsgemäß steigt der Rücklauf sprunghaft an den Tagen an, an denen Einladungen und Erinnerungen zur Befragungsteilnahme von den Hochschulen verschickt wurden und nimmt zum Ende der Feldphase hin ab. Im Zeitraum von Mai bis September 2021 haben sich insgesamt 274.466 Studierende⁸ an der Erhebung „Die Studierendenbefragung in Deutschland“ (2021) beteiligt. Daraus ergibt sich eine Brutorücklaufquote von 19,6 Prozent (siehe Tabelle 5).

Abbildung 2: „Die Studierendenbefragung in Deutschland“ (2021) - Entwicklung des Brutorücklaufs über die gesamte Feldphase (in absoluten Zahlen)



Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Darstellung.

[Rücklaufquoten nach AAPOR-Standard] Die Rücklaufquote wurde in Anlehnung an die Standards der American Association for Public Opinion Research ermittelt (AAPOR 2023; siehe Tabelle 5).

⁸ Die Angabe der Beteiligung in absoluten Zahlen liegt hierbei höher als in Abbildung 2 dargestellt, da dies auch Befragte einschließt, die lediglich die Startseite der Befragung aufgerufen haben oder bei denen keine Zustimmung zum Informed Consent vorlag.

Darüber hinaus wurden, um einen Analysedatensatz von guter Qualität zu erhalten, weitere Fälle entfernt. Von den insgesamt 274.466 Befragungsteilnahmen, wurden

- Befragte, die Angaben auf weniger als der Hälfte der als bedeutend definierten Variablen gemacht haben (sog. nicht-valide Fälle nach AAPOR 2023 (37.089), siehe Tabelle 5)
- Befragte, die nur die Startseite aufgerufen haben (28.667),
- Befragte, die zu Beginn in die EUROSTUDENT-Befragung gefiltert wurden (13.915),
- Befragte, die dem Informed Consent nicht zugestimmt haben (4.392),
- Befragte, bei denen Speeding oder Straightlining im Antwortverhalten identifiziert wurde (1.699),
- Befragte, die eine nachträgliche Datenlöschungsanfrage gestellt haben und Testzugänge (769)

aus dem Analysedatensatz ausgeschlossen. Nach diesen Ausschlüssen verblieben noch Angaben von 187.935 Studierenden, die als valide Fälle definiert wurden. Daraus ergibt sich eine bereinigte Rücklaufquote⁹ von 13,4 Prozent auf Studierendenebene (siehe Tabelle 5). Die Angaben von diesen 187.935 Befragten können für Analysen genutzt werden. Aufgrund des Splitdesigns der Befragung unterscheidet sich die Fallzahlen in den verschiedenen Modulen zum Teil erheblich (siehe Kapitel 2).

[Einschätzung des Rücklaufs] Da „Die Studierendenbefragung in Deutschland“ im Sommersemester 2021 erstmalig durchgeführt wurde, gibt es bezüglich des Rücklaufs keine Vergleichszahlen aus vergangenen Erhebungen. Jedoch können andere Studien für Vergleichszwecke herangezogen werden: In der 21. Sozialerhebung wurde ebenfalls ein standardisierter Online-Fragebogen bei einer ähnlichen Population eingesetzt (Becker et al., 2019). Für deutsche und bildungsinländische Studierende lag die bereinigte Rücklaufquote bei 16,2 Prozent und für bildungsausländische Studierende bei 10,2 Prozent. Im letzten Studierendensurvey aus dem Wintersemester 2015/2016 wurden sowohl Papier- (PAPI) als auch Online- (CAWI) Fragebögen eingesetzt (Multrus, 2021). Die Rücklaufquote für PAPI lag bei 15,2 Prozent, die für CAWI bei 8,8 Prozent (Multrus, 2021).

Für die genannten Beispiele finden sich in den jeweiligen Methodenberichten keine Angaben dazu, ob für die Berechnung der Rücklaufquoten die AAPOR-Standards angewandt wurden, die zusammen mit den weiteren Ausschlusskriterien¹⁰ im vorliegenden Fall zu einer konservativen Berechnung der Rücklaufquote geführt haben. Dennoch sind vorsichtige Vergleiche der Rücklaufquoten möglich. Vor dem Hintergrund dieser Zahlen sowie unter Berücksichtigung, dass die Befragten ausschließlich online befragt, nicht schriftlich erinnert und nicht (pre-paid) incentiviert wurden, ist die bereinigte Rücklaufquote der Studie „Die Studierendenbefragung in Deutschland“ (2021) mit 13,4 Prozent auf Ebene der Studierenden (SSU) als insgesamt sehr zufriedenstellend zu bewerten.

⁹ In die Berechnung der bereinigten Stichprobe fließen auch die „partial completed“ mit ein, sofern diese mindestens die Hälfte der als bedeutend definierten Variablen beantwortet haben (AAPOR, 2016).

¹⁰ Zu berücksichtigen ist zudem der nachträgliche Ausschluss der Fälle, die der EUROSTUDENT-Befragung zugeführt sowie der Fälle, die aufgrund von Speeding oder Straightlining aus dem Analysedatensatz entfernt wurden.

Tabelle 5: Teilnehmende Hochschulen, Stichprobengröße und Rücklauf¹¹

	Anzahl	Anteil (in %)
Grundgesamtheit		
<i>Hochschulen</i>		
Hochschulen insgesamt ^{a)}	396	
Teilnehmende Hochschulen	250	63,1 %
<i>Studierende</i>		
Studierende insgesamt ^{b)}	2.733.151	
Studierende an teilnehmenden Hochschulen	2.295.075	84,0 %
Stichprobe		
<i>Angeschriebene Studierende an den teilnehmenden Hochschulen ^{c)}</i>	1.397.326	60,9 %
Rücklauf		
<i>Brutto-Rücklauf</i>	274.466	19,6 %
davon ...		
... nicht-valide Fälle ^{d)}	37.089	2,7 %
... nur Startseite aufgerufen	28.667	2,1 %
... in EUROSTUDENT-Befragung geleitet ^{e)}	13.915	1,0 %
... keine Zustimmung Inform Consent	4.392	0,3 %
... Speeding & Straightlining ^{f)}	1.699	0,1 %
... Datenlöschungsanfragen & Testzugänge	769	0,1 %
Fälle im Analysedatensatz (bereinigter Rücklauf)	187.935	13,4 %

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Berechnungen.

Anmerkung: ^{a)} Alle Hochschulen in Deutschland im Sommersemester 2021 ohne Verwaltungsfachhochschulen. Eigene Berechnungen mit Daten des Statistischen Bundesamts (Statistisches Bundesamt, 2022); ^{b)} Alle Studierenden in Deutschland im Sommersemester 2021 ohne Studierende an Verwaltungsfachhochschulen. Eigene Berechnungen mit Daten des Statistischen Bundesamts; ^{c)} Die teilnehmenden Hochschulen haben entweder eine Zufallsauswahl von einem Drittel der Studierenden in die Stichprobe aufgenommen oder eine Vollerhebung durchgeführt. Die Stichprobe umfasst die angeschriebenen Studierenden. In dieser sind Studierende ohne E-Mailadresse oder mit ungültiger E-Mailadresse (auch „Mail-delivery-failed“-Meldungen) nicht enthalten; ^{d)} In Anlehnung an die Richtlinie der American Association for Public Opinion Research (AAPOR) gelten Befragte als nicht-valide Fälle, wenn sie Angaben auf weniger als der Hälfte der als bedeutend definierten Variablen gemacht haben (AAPOR, 2023). Für „Die Studierendenbefragung in Deutschland“ sind dies insgesamt 22 Variablen, u. a. Hochschule, Geschlecht, angestrebter Abschluss, Studienfach, Studienform. ^{e)} Ein zufällig ausgewählter Teil der Befragten (5,6 Prozent) wurde in die Befragung der internationalen Vergleichsstudie Eurostudent geleitet; ^{f)} Das unplausible Antwortverhalten derjenigen, die entweder sehr schnell (Speeding) oder mit immer identischen Antworten (Straightlining) den gesamten Fragebogen beantwortet haben, wird von den nachfolgenden Analysen ausgeschlossen.

¹¹ Die Tabelle findet sich auch bei Kroher et al. (2023: 146).

6 Datenaufbereitung

Im Folgenden werden die verschiedenen Schritte der Datenaufbereitung beschrieben. Im Einzelnen handelt es sich um die Datenübertragung, -prüfung und -bereinigung (Kap. 6.1), die Vergabe von Variablennamen, Variablenlabel und Wertelabel (Kap. 6.2), die Codierung fehlender Werte (Kap. 6.3), die Codierung offener Angaben (Kap. 6.4), die Generierung von Variablen (Kap. 6.5) sowie die Erstellung des Datensatzes (Kap. 6.6). Die beschriebenen Tätigkeiten wurden vom Projektteam am DZHW und an der Universität in Konstanz in Zusammenarbeit mit dem FDZ-DZHW durchgeführt. Die Datenaufbereitungsschritte wurden fast ausschließlich mit der Software *Stata* durchgeführt. Die im Rahmen der Datenaufbereitung vorgenommenen Aufbereitungsprozesse der Gewichtung und Anonymisierung werden in den beiden folgenden Kapiteln 7 und 8 gesondert erläutert. Die Gewichtung erfolgte durch das Projektteam, die Anonymisierung durch das Team des FDZ-DZHW.

6.1 Datenübertragung, Datenprüfung und Datenbereinigung

[Datenübertragung] Die Befragung wurde mit der DZHW-eigenen Onlinebefragungssoftware *Zofar* durchgeführt. Die Daten aus der Online-Befragung wurden direkt aus der Befragungssoftware als *.csv*-Datei exportiert und als *Stata*-Datei (*.dta*) weiterverarbeitet.

[Konsistenz- und Plausibilitätsprüfung] Im Anschluss an die Datenübertragung aus der Onlinebefragungssoftware *Zofar* erfolgte eine umfassende Prüfung der Daten. Die jeweiligen Schritte dieser Prüfung wurden syntaxbasiert mittels *Stata*-Do-Files und regelgeleitet durchgeführt. Es wurden folgende Typen von Prüfungen vorgenommen:

[Prüfung des Exportdatensatzes] Während der Datenerhebung wurden zum Zwecke der Feldkontrolle Testzugänge zur Online-Befragung bereitgestellt. Über diese Testzugänge erhielten Projektmitarbeiter*innen am DZHW ebenfalls die von den Hochschulen versandten E-Mails. Dadurch gewann das Projektteam einen Einblick in den (de facto-)Zeitpunkt und die Güte der von den Hochschulen versendeten Einladungs- und Erinnerungs-E-Mails. Diese Testzugänge wurden nach Abschluss der Datenerhebung aus dem Datensatz vollständig gelöscht. Weiterhin wurden beispielsweise Fälle der EU-ROSTUDENT-Befragung oder Fälle, die ihre Teilnahme (nachträglich) zurückgezogen haben gelöscht (siehe Kapitel 5). Ferner wurden etwaige im Zuge der Datenübermittlung entstandenen Exportinkonsistenzen (bspw. Javascript-basierte Fehler) oder durch das Befragungssystem inadäquat vergebene (Missing-)Werte im Rohdatensatz recodiert.¹²

[Prüfung von Wertebereichen] Um von vorneherein das Ausmaß an unplausiblen Werten im Datensatz zu reduzieren, wurden mit der Fragebogenprogrammierung Wertebereiche definiert, die nicht unter- bzw. überschritten werden sollten. Im Falle einer Unter- bzw. Überschreitung dieser Wertebereiche wurden die Befragten mit einem Hinweis in roter Schrift darauf aufmerksam gemacht, dass ihre Eingabe unplausibel erscheint und sie wurden darum gebeten, den Wert zu korrigieren. Erfolgte nach diesem Hinweis keine Korrektur, konnte die Befragung dennoch fortgesetzt werden („*soft-forcing*“). Deshalb wurde im Datensatz geprüft, ob die erfasste Ausprägung einer Variable in dem für diese Variable definierten Wertebereich liegt.

¹² Dies betrifft beispielsweise die Codierung der Kategorie „bitte auswählen“, welche im Online-Fragebogen über ein Drop-Down-Menü erhoben worden ist. Hier wurde der Exportwert „0“ in einen (Item-Nonresponse-)Missing-Wert recodiert.

[Prüfung der Einhaltung der Filterführung] Aufgrund des modularisierten Befragungs- und Splitdesigns ist die Filterführung komplex. Daher wurde einerseits geprüft, ob aufgrund der im Fragebogen vorgesehenen Filterführung für die jeweilige befragte Person Angaben zu erwarten gewesen wären („Vollständigkeitsprüfung“), und andererseits, ob für die jeweilige Person keine Angaben hätten erfolgen dürfen („Filterverstöße“) und die entsprechenden korrekten Missingwerte gesetzt. Aufgrund der Möglichkeit des „Zurückgehens“ im Fragebogen und der möglichen Korrektur von Angaben ist der relevante Pfad für die Prüfung des Fragebogenverlaufs und der Filterführung immer der letztgespeicherte Pfad.

[Konsistenzprüfung] Nach der Filterprüfung wurde eine grundlegende Konsistenzprüfung der Angaben innerhalb des Fragebogens vorgenommen. Hierbei wurden insbesondere inhaltliche Angaben in den jeweiligen Themenfeldern auf mögliche Widersprüche überprüft.

[Definition gültiger Fälle] Die Bestimmung der gültigen Fälle erfolgte anhand mehrerer Kriterien in sukzessiven Schritten. In einem ersten Schritt wurden unvollständig ausgefüllte Fragebögen genauer betrachtet. Einer Richtlinie der American Association for Public Opinion Research (AAPOR 2016) folgend, wurden alle Fälle, bei denen mindestens 50 Prozent der zuvor als bedeutend definierten Fragen beantwortet wurden, als valide bzw. gültig bestimmt. Insgesamt wurden 22 Variablen aus dem A-Modul für die Definition „bedeutend“ herangezogen. Diese sind in Tabelle 6 aufgelistet.

Tabelle 6: Als bedeutend definierte Variablen des A-Moduls

Fragebogenseite	Variablenbeschreibung	Variablenname
A_18	Hochschule	preloadhs_id
A_17	Fachsemester	ssemfs
A_22	Zufriedenheit mit Hochschulbedingungen	sperleiszufr
A_23	Studienfach	sfach1o2
A_25	Abschlussart	sabsan
A_27	Studienform	sformpraes-sformsons
A_1	Ort des HZB-Erwerb (Dtl. vs. Ausland)	vsbdeba
A_5	Geburtsort (Dtl. vs. Ausland)	dgebort
A_9b	Geschlecht	demosex
A_12	Elternschaft	dkinja
A_15	Beeinträchtigung	gbeges
A_16	Derzeitiger Auslandsaufenthalt	imausl
A_31	Erfolgswahrscheinlichkeit Studienabschluss	studerfolg
A_32	Höchster Berufsabschluss Vater	deltberuv
A_32	Höchster Berufsabschluss Mutter	deltberum
A_34	Geburtsland Vater	deltgebv
A_34	Geburtsland Mutter	deltgebm
A_38	Höchster Schulabschluss	vsbart
A_42	Erwerbstätigkeit	eaktsens
A_51b	Studienabbruchintention	ssaaja
A_54	Bisherige studienbez. Auslandsaufenthalte	ainfaus
D1_9	Voll-/Teilzeitstudium	sartzeit

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Darstellung.

[Straightlining] Negativ auf die Datenqualität wirkt sich das sogenannte Straightlining aus, ein Indikator für die mangelnde Motivation der Befragten, valide Antworten zu geben (Roßmann, 2017). In Anlehnung an Roßmann (2017) wird Straightlining über die Nicht-Differenzierung der Antworten in Matrixfragen definiert, beispielsweise, wenn bei Zustimmungsitems, unabhängig vom konkreten Inhalt des Items, immer der gleiche Wert, zum Beispiel „stimme voll und ganz zu“ (Wert 5), angegeben wird. Im Anschluss an die Identifikation der validen Fälle über die bedeutenden Variablen wurde deshalb in einem zweiten Schritt zur Identifikation der Straightliner*innen das Antwortverhalten auf den hierfür relevanten Fragebogenseiten analysiert. Nach obiger Definition liegt Straightlining auf der jeweiligen Fragebogenseite dann vor, wenn für eine Person alle Antworten einer Matrixfrage auf einer Fragebogenseite denselben Wert annehmen.

Nach Roßmann (2017) wurden ex-ante die Fragebogenseiten mit mindestens drei Items mit einer 4er-, 5er- oder 6er-Skala betrachtet. Lediglich Fragebogenseiten, auf die diese Kriterien zwar zutreffen, die jedoch Faktenwissen abfragen (z. B. Hochschulabschluss der Eltern oder Großeltern, siehe Frage B2_10) oder nur bestimmten Befragtengruppen vorgelegt wurden (z.B. Studierenden mit Kind(ern)), wurden nicht zur Straightlining-Identifikation herangezogen. Dieses Vorgehen zielte unter anderem darauf, die Wahrscheinlichkeit, Straightlining zufällig zu beobachten, zu reduzieren (Roßmann, 2017: 87). Insgesamt floßen 65 Fragebogenseiten in die Betrachtung ein (siehe Tabelle 7).

Tabelle 7: Übersicht der straightlining-relevanten Fragebogenseiten in „Die Studierendenbefragung in Deutschland“ (2021)

Nr.	Seite	Skalierung	Nr.	Seite	Skalierung
1	A_57	5er-Skala mit 12 Items	34	E1_3	5er-Skala mit 11 Items
2	B1_2	5er-Skala mit 7 Items	35	E1_3a	5er-Skala mit 10 Items
3	B1_3a	5er-Skala mit 6 Items	36	E1_6	5er-Skala mit 13 Items
4	B1_3b	5er-Skala mit 7 Items	37	E1_11	5er-Skala mit 6 Items
5	B2_6	5er-Skala mit 6 Items	38	E2M_4	5er-Skala mit 4 Items
6	B2_7	5er-Skala mit 7 Items	39	E2M_5	4er-Skala mit 11 Items
7	D1_18	5er-Skala mit 3 Items	40	E2M_6	5er-Skala mit 13 Items
8	D1_7	5er-Skala mit 12 Items	41	E2M_7	5er-Skala mit 4 Items
9	D1_8	5er-Skala mit 12 Items	42	E2P_2	5er-Skala mit 18 Items
10	D1_19	5er-Skala mit 4 Items	43	E2S_2	5er-Skala mit 18 Items
11	D1_20	5er-Skala mit 15 Items	44	E2S_3	5er-Skala mit 12 Items
12	D1_21	5er-Skala mit 7 Items	45	E2S_4	5er-Skala mit 12 Items
13	D1_24	5er-Skala mit 10 Items	46	E2S_6	4er-Skala mit 7 Items und 5er-Skala mit 7 Items
14	D1_25	5er-Skala mit 3 Items	47	E2S_7	6er-Skala mit 7 Items
15	D1_29	5er-Skala mit 18 Items	48	E2S_8	6er-Skala mit 11 Items
16	D2_2	5er-Skala mit 9 Items	49	E3_3	6er-Skala mit 10 Items
17	D2_3	5er-Skala mit 10 Items	50	E3_4	6er-Skala mit 13 Items
18	D2_4	5er-Skala mit 6 Items	51	E3_5	6er-Skala mit 5 Items
19	D2_6a	5er-Skala mit 3 Items	52	E3_6	5er-Skala mit 7 Items
20	D2_7	5er-Skala mit 6 Items	53	E4_3	5er-Skala mit 3 Items
21	D2_9	5er-Skala mit 3 Items	54	E4_4	5er-Skala mit 12 Items
22	D2_11	5er-Skala mit 6 Items	55	E4_5	6er-Skala mit 6 Items
23	D2_12	5er-Skala mit 3 Items	56	E4_8	5er-Skala mit 9 Items
24	D2_13	5er-Skala mit 6 Items	57	F1_2	5er-Skala mit 14 Items

25	D2_16	5er-Skala mit 3 Items	58	N_9	5er-Skala mit 5 Items
26	D2_21	5er-Skala mit 11 Items	59	N_9c	5er-Skala mit 4 Items
27	D3_5	5er-Skala mit 5 Items	60	N_11	5er-Skala mit 8 Items
28	D3_11	5er-Skala mit 4 Items	61	N_12	5er-Skala mit 9 Items
29	D3_22	5er-Skala mit 4 Items	62	CO_10	5er-Skala mit 7 Items
30	D1_27	5er-Skala mit 4 Items	63	N_17	5er-Skala mit 3 Items
31	D3_25	5er-Skala mit 4 Items	64	N_16	5er-Skala mit 8 Items
32	E1_1	5er-Skala mit 5 Items	65	B2_5	5er-Skala mit 5 Items
33	E1_2	5er-Skala mit 11 Items			

Quelle: „Die Studierendenbefragung in Deutschland“ (2021). Eigene Darstellung

Straightlining wurde für jede der in Tabelle 7 aufgelisteten Fragebogenseiten als binäre Indikator-Variable indiziert. Beim Vorliegen von Straightlining nahm die Indikator-Variable den Wert „1“, in allen anderen Fällen den Wert „0“ an.

Daran anknüpfend wurde ein Straightlining-Index pro Person i aus der Summe der mit Straightlining indizierten Seiten j berechnet (siehe Formel 1).

$$\text{Straightlining} - \text{Index}_i = \sum_{j=1}^{n_i} \text{Straightlining}_j \quad (1)$$

Aufgrund des modularisierten Splitdesigns des Fragebogens wurden nicht alle Fragen den Befragten vorgelegt, weshalb der Straightlining-Index Werte zwischen $[0, 32] := \{x \in \mathbb{N} \mid 0 \leq x \leq 32\}$ annehmen konnte. Da diese Indexvariable rechtsschief verteilt war, lag das 99er-Perzentil bei dem Wert 9.

Dies berücksichtigend, wurde darauf aufbauend die Straightlining-Quote aus der Summe der Fragen, bei denen Straightlining durch die Indikator-Variablen angezeigt wurden ($j = 1$), und der Anzahl der im Befragungsverlauf gesehenen bzw. den Befragten vorgelegten Fragebogenseiten n berechnet (siehe Formel 2).

$$\text{Straightlining} - \text{Quote}_i = \frac{\sum_{j=1}^{n_i} \text{Straightlining}_j}{n_i} \quad (2)$$

Abschließend wurden mit Blick auf den gesamten Fragebogen diejenigen Befragten als Straightliner*innen ausgewiesen, die auf mehr als acht Fragebogenseiten (99er-Perzentil, siehe oben) Straightlining im Antwortverhalten zeigten und bei denen gleichzeitig Straightlining auf mindestens 50 Prozent der gesehenen Fragebogenseiten vorlag. Neben den als ungültig definierten Fällen wurden auch die als Straightliner*innen definierten Befragten aus dem Datensatz ausgeschlossen.

[Speeding] Neben Straightlining wirkt sich auch das sogenannte Speeding negativ auf die Datenqualität aus. Bei Speeder*innen handelt es sich um Befragte, die besonders bzw. vergleichsweise schnell den Fragebogen durchlaufen haben. Auch hier wird eine mangelnde Motivation der Befragten, valide Antworten zu geben, unterstellt. Zur Identifikation dieser Gruppe wurde die Gesamtfragebogendauer aller Befragten, die den Fragebogen beendet hatten, herangezogen (siehe Tabelle 4). Als Speeder*innen wurden diejenigen Befragten definiert, die den Fragebogen schneller als 99 Prozent der übrigen Befragten durchlaufen hatten. Analog zu der Berechnung der Straightliner*innen wurde demnach auch hier entsprechend verteilungstheoretischer Annahmen auf das 99er-Perzentil der Verteilung der Speedingvariable zurückgegriffen. Die als Speeder*innen identifizierten Befragten wurden aus dem Datensatz gelöscht (siehe Tabelle 5).

6.2 Vergabe von Variablennamen, Variablenlabel und Wertelabel

[Vergabe von Variablennamen] Die Vergabe der Variablennamen erfolgte anhand eines internen Schemas. Die Namen geben das jeweilige Merkmal bzw. Konstrukt in Kurzform wieder (siehe hierzu auch den Variablenfragebogen). Insgesamt finden sich im Datensatz vier unterschiedliche Suffixe an den Variablennamen:

- **_p-Variablen:** Variablen, deren Ausprägung plausibilisiert wurde.
- **_g-Variablen:** Variablen, die im Zuge der Datenaufbereitung generiert wurde.¹³¹⁴
- **_pg-Variablen:** Variablen, die sowohl plausibilisiert als auch generiert wurden.
- **_a-Variablen:** Variablen, deren Ausprägungen aggregiert wurden.

[Vergabe von Variablen- und Wertelabel] Für die Variablen- und Wertelabel wurden Formulierungen des Fragebogens übernommen oder prägnante Kurzformen dieser Formulierungen gewählt. Dabei basieren die Variablenlabel in der Regel auf der inhaltlichen Beschreibung der Variablen. Grundlage für die Wertelabel sind je nach Fragetyp die Texte der Antwortoptionen bzw. eine Kombination der Texte von Frage und Antwortoption.

6.3 Codierung fehlender Werte

[Missing-Kategorien] Die Codierung fehlender Werte orientiert sich an der Systematik des FDZ-DZHW. Fehlende Angaben werden diesem Schema folgend durch dreistellige negative Werte codiert. In Tabelle 9 sind die Kategorien, die vergeben wurden, gelistet. Fett hervorgehoben sind die Kategorien, die in der Studie „Die Studierendenbefragung in Deutschland“ (2021) verwendet wurden.

Tabelle 8: Systematik des FDZ-DZHW für fehlende Werte

Wertebereich	Code	Wertelabel
-999 bis -990: Nonresponse	-999	Weiß nicht
	-998	Keine Angabe
	-997	Keine Angabe (Antwortkategorie)
	-996	Interviewabbruch
	-995	Keine Teilnahme (Panel)
	-994	Verweigert
-989 bis -970: nicht zutreffend	-989	Filterbedingt fehlend
	-988	Trifft nicht zu
	-987	Designbedingt fehlend (Fragebo- gensplit)
	-986	Designbedingt fehlend (Welle) ^a
	-985	Designbedingt fehlend (Kohorte) ^b

¹³ Bei Variablen, die vom FDZ-DZHW im Zuge der Anonymisierung gebildet wurden, werden die gebildeten Variablen mit dem Suffix **_g#** versehen (# = Ziffer), da dort ggf. mehrere neue Variablen auf Basis einer Originalvariable gebildet wurden. Die vom Primärforschungsprojekt vergebenen Suffixe wurden vom FDZ-DZHW nur in Ausnahmefällen überarbeitet.

¹⁴ Zudem existieren generierte Zusatzvariablen ohne spezifische Endung (siehe hierzu auch Tabelle 13 und 14 im Anhang).

-969 bis -950: Editierter fehlender Wert	-969	Unbekannter fehlender Wert ^c
	-968	Unplausibler Wert ^d
	-967	Anonymisiert
	-966	Nicht bestimmbar ^e
	-965	Ungültige Mehrfachnennung
-949 bis -930: Item-spezifische fehlende Werte (nicht vergeben)		
-929 bis -920: Andere fehlende Werte	-929	Datenverlust

^a Dieser Wert wird nur für Datensätze im Long-Format vergeben.
^b Dieser Wert wird nur in gepoolten Datensätzen vergeben.
^c Dieser Wert wird vergeben, wenn keinerlei Ursache rekonstruiert werden kann.
^d Angaben, die aufgrund unterschiedlicher Faktoren in der Codierphase als nicht plausibel eingestuft werden, erhalten diesen Wert. Eine exakte Rekonstruktion ist ggf. nicht mehr möglich.
^e Diese Kategorie wird vergeben, wenn eine eindeutige Codierung nicht möglich ist.

[Missings für Befragungsabbruch] Zur Rekonstruktion von Befragungsabbrüchen wurde überprüft, bis zu welcher Seite die Befragung von den Befragten ausgefüllt wurde. Im Falle eines vorzeitigen Abbruchs der Befragung wurde für alle nachfolgenden Seiten ein Missing für „Interviewabbruch“ vergeben (-996).

[Filtermissings] Die Onlinebefragungssoftware *Zofar* arbeitet mit Ausgangsfiltern (*transitions*). Diese Ausgangsfilter aus der Programmiervorlage, dem Flowchart und der QML-Datei von *Zofar* werden seitenweise in Eingangsfiler übersetzt. Für Befragte, die eine Variable aufgrund zuvor gemachter Filterangaben nicht gesehen haben, wurde die entsprechende Missingkategorie für Filtermissings vergeben (-989).

[Designmissings] Mit dem modularisierten Befragungs- und Splitdesign (siehe Kapitel 2) wurde festgelegt, dass nicht alle Befragten alle Fragen vorgelegt bekommen, sondern – je nach den zufällig vergebenen Preloads für die Master- und Zusatzsplits – ein bestimmter Pfad durch den Fragebogen vorgegeben wurde. Fragen, die aufgrund des jeweiligen Mastersplits nicht vorgelegt wurden, wurden mit der Missingkategorie „designbedingt fehlend (Fragebogensplit)“ (-987) versehen.

Des Weiteren wurden bei bestimmten Fragen zusätzliche Splits eingesetzt, wenn Fragen nur von einer Teilstichprobe gesehen werden sollten (z. B. nur 12,5 Prozent aller Befragten) oder wenn zwei Fragen alternativ jeweils 50 Prozent der Befragten gezeigt wurden (*Split-Half-Design*). Auch in diesen Fällen wurde, sofern Fragen aufgrund des jeweiligen Designs nicht vorgelegt wurden, die Missingkategorie für „designbedingt fehlend (Fragebogensplit)“ (-987) vergeben.

[Missings für „keine Angabe“] Nach Bearbeitung der oben genannten Missingkategorien wurden die noch verbleibenden Missings noch einmal geprüft. Im Rahmen der Modul-Plausibilisierungen wurden Missings für Item-Nonresponse (*keine Angabe*) vergeben, beispielsweise bei fehlenden Angaben auf Multiple-Choice-Fragen.

[Missings für Ausweichkategorien] Bei einigen Fragen wurde den Befragten zur Vermeidung einer vorherigen, zusätzlichen Abfrage neben den Antwortkategorien eine zusätzliche Ausweichkategorie angeboten. In dieser konnte u.a. ausgedrückt werden, dass ein zu beantwortender Sachverhalt nicht zutrifft, bspw. weil noch keine angebbare Note vorliegt. Diese Ausweichkategorien wurden im Regelfall als außerhalb des gültigen Antwortbereichs der Fragen definiert und mit einem negativen zweistelligen Missingcode mit eindeutigem Label abgelegt. Die Festlegung, ob es sich in diesem Fällen tatsächlich um *fehlende* Angaben handelt, erfolgte durch das Primärforschungsprojekt.

Nach der Recodierung der Missings sowie nach der Plausibilisierung der einzelnen Module wurden die restlichen noch verbleibenden Missings auf „unbekannter fehlender Wert“ (-969) gesetzt.

6.4 Codierung offener Angaben

[Vielzahl offener Angaben] Der Fragebogen der Studie „Die Studierendenbefragung in Deutschland“ (2021) enthält ca. 130 Variablen, bei denen (halb-)offene Angaben möglich waren. Für die Sekundärdatennutzung im Scientific Use File (SUF) und Campus Use File (CUF) werden keine offenen Angaben zur Verfügung gestellt. Das Projektteam am DZHW und an der Universität Konstanz hat die (halb-)offenen Angaben kodiert und für die Sekundärdatennutzung entsprechend aufbereitet. Die offenen Angaben wurde entweder bereits bestehenden Antwortkategorien zugeordnet¹⁵ oder es wurden neue Variablen oder Variablenausprägungen generiert. Nachfolgend werden mit der Fächercodierung und der Berufscodierung zwei Codierungsverfahren, die teilautomatisiert erfolgten, skizziert.

[Fächercodierung] Die Codierung der Studienfächer erfolgte teilautomatisiert mittels eines *Python*-Skriptes durch das Projektteam an der Universität Konstanz (Weeber et al., 2024). Grundlage für die Klassifizierung war die Systematik für Studienfächer, Studienbereiche und Fächergruppen des Statistischen Bundesamtes für das Prüfungsjahr 2021¹⁶ (Destatis, 2020). Die Besonderheit der hier verwendeten regelbasierten Klassifizierungsmethode war die Vorgabe, dass keine falsch-positiven Zuweisungen vorgenommen werden sollten. Daher wurden nur als sicher geltende Zuweisungen direkt vorgenommen. Für die verbleibenden Angaben wurden – sofern möglich – Vorschläge erstellt, aus denen Coder*innen den richtigen auswählen oder neu zuweisen konnten. Die übrigen Angaben, für die weder ein Code direkt zugeordnet noch vorgeschlagen werden konnte, mussten per Einzelfallprüfung codiert werden. Die finalen Klassifikationen konnten dann als neue Variablen an den Datensatz angefügt und anschließend exportiert werden (für weitere Details siehe Weeber et al., 2024). Für die Sekundärdatennutzer*innen sind die Studienfächer und die zugehörige Klassifikation der Studienbereiche bzw. Fächergruppen nach dem Statistischen Bundesamt (bspw. Variable *SB1_STB_p* bzw. *FG1_STB_p*) sowie die im Bericht der 22. Sozialerhebung verwendete projektspezifische Klassifikation der Studienbereiche (Variable *studfach*) im Datensatz verfügbar. Im Bericht zur Studiensituation unter Onlinebedingungen wurde die Kategorie Lehramt noch zusätzlich ausgewiesen (Multrus et al. 2023).

[Berufscodierung] Die Angaben zu den Berufen der Eltern, dem erlernten Ausbildungsberuf der Befragten sowie dem aktuellen und angestrebten Beruf der Befragten wurden in der Studie „Die Studierendenbefragung in Deutschland“ (2021) ebenfalls offen erfasst. Im Zuge der Datenaufbereitung wurden diese Angaben ebenfalls teilautomatisiert vercodet. Die automatisierte Vercodung erfolgte auf der Grundlage der Klassifikation der Berufe des Statistischen Bundesamtes in der Fassung von 2010 (KldB, 2010) und nach einem Algorithmus von Safikhani et al. (2023). Alle nicht eindeutig zuordenbaren Angaben wurden durch Coder*innen nachvercodet.

¹⁵ Sofern Angaben zu bestehenden Antwortkategorien zugeordnet wurden, kann die Konsistenz der Filterführung (siehe Kapitel 6.1) davon berührt worden sein – falls die betroffene Antwortkategorie als Filtervariable diente.

¹⁶ Diese bezieht sich auf das Wintersemester 2020/2021 und Sommersemester 2021.

6.5 Generierung von Variablen

[Übersicht zu generierten Variablen] Die für die Sekundärdatennutzung verfügbaren Datensätze enthalten (teilweise) Variablen, die vom Projektteam am DZHW und der Universität Konstanz erstellt wurden. Eine Übersicht der wichtigsten dieser Variablen ist in Tabelle 13 und Tabelle 14 im Anhang hinterlegt.

Im Zuge der Anonymisierung für die jeweiligen Zugangswege der Datensätze wurden vom FDZ-DZHW zusätzliche Variablen mit Aggregationen bestehender Informationen generiert. Eine Übersicht zu diesen findet sich in Tabelle 15 im Anhang.

6.6 Erstellung des Datensatzes

[Datenstruktur und Dateiformat] Die verfügbaren Datensätze als SUF und CUF enthalten die Befragungsdaten sowie die zusätzlich generierten und aggregierten Variablen. Die Reihenfolge der Variablen im Datensatz orientiert sich an der Reihenfolge der entsprechenden Fragen im Fragebogen. Für die Sekundärdatennutzung werden die Datensätze im *Stata*- und *SPSS*-Format bereitgestellt.

7 Gewichtung

In diesem Kapitel erfolgt in Kapitel 7.1 eine allgemeine Einführung in das konzeptionelle Vorgehen und eine Darstellung der erstellten Gewichte. Im Anschluss wird in Kapitel 7.2 die Gewichtungsprozedur im Detail beschrieben.

7.1 Konzeptionelles Vorgehen

[Ziel von Gewichtung] Die Studie „Die Studierendenbefragung in Deutschland“ (2021) erhebt den Anspruch, repräsentative Aussagen zur Studierendenschaft in Deutschland im Sommersemester 2021 zu ermöglichen (Beuße et al., 2022). Hierfür war es notwendig, die erhobenen Daten wissenschaftlich zu gewichten (Kiesl, 2022). Gewichtungsverfahren ermöglichen es, Stichprobenverzerrungen, beispielsweise nach Geschlecht, Alter oder Hochschulträgerschaft, auszugleichen, indem in den Befragungsdaten unterrepräsentierte Gruppen ein höheres Gewicht erhalten. Grundlage für die Gewichtung bilden die Daten der amtlichen (Hochschul-)Statistik zu den genannten Merkmalen (HStatG, siehe Kapitel 3). Darüber hinaus gleichen Gewichtungsverfahren die unterschiedlichen Auswahlwahrscheinlichkeiten von Studierenden an den teilnehmenden Hochschulen aus (siehe Kapitel 3). Tabelle 9 gibt einen Überblick zur Verteilung der Gewichtungsvariablen in der amtlichen Statistik sowie im ungewichteten und im gewichteten Datensatz von „Die Studierendenbefragung in Deutschland“ (2021).

[Konzeptionelles Vorgehen] Es wurde ein mehrstufiges Gewichtungsverfahren durchgeführt, welches eine Design- und eine Redressement-Gewichtung der Befragungsdaten kombiniert:

(1) Mit dem Design-Gewicht wird zunächst der unterschiedlichen Inklusionswahrscheinlichkeit von Studierenden Rechnung getragen.

(2) Da auf individueller Ebene keine Informationen zu Nicht-Teilnehmenden vorliegen, wurde in einem nächsten Schritt ein Redressement-Gewicht basierend auf dem „Iterative Proportional Fitting-Ansatz“ erstellt (Kolenikov, 2014).

Tabelle 9: Verteilung der GewichtungsvARIABLEN, Daten der amtlichen Statistik, ungewichteter und gewichteter Analysedatensatz (in %)

Gewichtungsvariablen	Amtliche Statistik	Ungewichteter Analysedatensatz	Differenz	Gewichteter Analysedatensatz
Hochschulart				
Universität (inkl. Kunst- und Musikhochschulen)	62,4 %	60,0 %	2,4 %	62,3 %
Hochschule für angewandte Wissenschaften	37,6 %	40,0 %	2,4 %	37,7 %
	100,0 %	100,0 %		100,0 %
Bundesland des Studienorts				
Baden-Württemberg	12,2 %	14,0 %	1,8 %	12,2 %
Bayern	13,7 %	13,1 %	0,6 %	13,7 %
Berlin	6,9 %	4,0 %	2,9 %	6,9 %
Brandenburg	1,7 %	1,6 %	0,1 %	1,7 %
Bremen	1,3 %	0,9 %	0,4 %	1,3 %
Hamburg	4,1 %	2,0 %	2,1 %	4,1 %
Hessen	9,0 %	7,2 %	1,8 %	9,0 %
Mecklenburg-Vorpommern	1,3 %	1,8 %	0,5 %	1,3 %
Niedersachsen	7,0 %	9,6 %	2,6 %	7,0 %
Nordrhein-Westfalen	26,2 %	30,7 %	4,5 %	26,2 %
Rheinland-Pfalz	4,2 %	3,2 %	1,0 %	4,2 %
Saarland	1,0 %	1,5 %	0,5 %	1,0 %
Sachsen	3,5 %	3,1 %	0,4 %	3,5 %
Sachsen-Anhalt	1,8 %	2,7 %	0,9 %	1,9 %
Schleswig-Holstein	2,2 %	2,0 %	0,2 %	2,2 %
Thüringen	4,0 %	2,6 %	1,4 %	4,0 %
	100,0 %	100,0 %		100,0 %
Geschlechtsidentität				
Männlich	50,1 %	39,1 %	11,0 %	49,6 %
Weiblich	49,9 %	60,0 %	10,1 %	49,4 %
Divers	n. v.	0,7 %	n. v.	0,7 %
Andere Geschlechtsidentität	n. v.	0,3 %	n. v.	0,2 %
	100,0 %	100,0 %		100,0 %
Alter				
Bis 19 Jahre	6,1 %	7,3 %	1,2 %	6,1 %
20-22 Jahre	27,1 %	35,5 %	8,4 %	27,1 %
23-25 Jahre	27,8 %	30,4 %	2,6 %	27,8 %
26-30 Jahre	23,0 %	17,3 %	5,7 %	23,0 %
31 Jahre und älter	16,0 %	9,5 %	6,5 %	16,0 %
	100,0 %	100,0 %		100,0 %

Fortsetzung auf der nächsten Seite

Fortsetzung Tabelle 9: Verteilung der Gewichtungsvariablen, Daten der amtlichen Statistik, ungewichteter und gewichteter Analysedatensatz (in %)

Gewichtungsvariablen	Amtliche Statistik	Ungewichteter Analysedatensatz	Differenz	Gewichteter Analysedatensatz
Studienbereich^a				
Sprachwissenschaften	6,3 %	6,0 %	0,3 %	6,0 %
Geisteswissenschaften (ohne Sprachwissenschaften)	4,9 %	4,6 %	0,3 %	5,2 %
Erziehungswissenschaften, Sozialwesen	7,4 %	11,7 %	4,3 %	8,0 %
Sozial-, Politik-, Regional-, Verwaltungswissenschaften, Psychologie	8,1 %	8,7 %	0,6 %	7,4 %
Rechtswissenschaften	4,9 %	4,4 %	0,5 %	4,9 %
(Internationale) Betriebswirtschaftslehre	10,3 %	9,5 %	0,8 %	10,0 %
Wirtschaftswissenschaften (ohne Betriebswirtschaftslehre)	6,8 %	6,7 %	0,1 %	7,2 %
Mathematik, Naturwissenschaften	10,9 %	11,3 %	0,4 %	10,9 %
Medizin (Human- und Zahnmedizin)	4,2 %	4,8 %	0,6 %	4,2 %
Gesundheitswissenschaften (inkl. Sport)	3,6 %	3,9 %	0,3 %	3,6 %
Agrar-, Forst-, Ernährungswissenschaften, Veterinärmedizin	2,2 %	2,6 %	0,4 %	2,2 %
Informatik	8,6 %	7,2 %	1,4 %	9,4 %
Maschinenbau und Verfahrens-, Elektro-, Informationstechnik	8,6 %	7,3 %	1,3 %	8,6 %
Ingenieurwissenschaften (ohne Maschinenbau und Verfahrens-, Elektro-, Informationstechnik)	9,8 %	8,3 %	1,5 %	9,0 %
Kunst, Kunstwissenschaften	3,4 %	3,0 %	0,4 %	3,4 %
	100,0 %	100,0 %		100,0 %
Ort des Erwerbs der Hochschulzugangsberechtigung				
Studierende aus Deutschland	85,2 %	94,1 %	8,9 %	85,2 %
Internationale Studierende	14,8 %	5,9 %	8,9 %	14,8 %
	100,0 %	100,0 %		100,0 %

Quelle: Eigene Berechnungen mit Daten des Statistischen Bundesamts (Statistisches Bundesamt, 2022) sowie mit ungewichteten und gewichteten Daten aus „Die Studierendenbefragung in Deutschland“ (2021). Die Analysen basieren auf Angaben von 179.908 Studierenden.

Anmerkung: Durch Rundungsdifferenzen können aufsummierte Werte von 100,0 Prozent abweichen. n. v. = nicht in der amtlichen Statistik (Fachserie 11 Reihe 4.1) vorhanden. ^a = die hier dargestellten Studienbereiche stellen nicht die Studienbereiche der Hochschulstatistik des Statistischen Bundesamts dar, sondern eine Aggregationsvariante von Studienfächern, die in den Berichten von „Die Studierendenbefragung in Deutschland“ (2021) Anwendung fand. Sie wurde nicht zur Gewichtung genutzt und dient lediglich der Visualisierung der Verteilung der gewichteten und ungewichteten Daten.

7.2 Gewichtungsverfahren und Anwendungshinweise

[Gewichtungsverfahren] Für die Daten der Studie „Die Studierendenbefragung in Deutschland“ (2021) wurden in einem mehrstufigen Verfahren mittels Design- und Redressementgewichtung die folgenden zwei Gewichte erstellt, die sowohl im SUF als auch im CUF bereitgestellt werden (siehe Tabelle 10):

Tabelle 10: Bereitgestellte Gewichte der Studie „Die Studierendenbefragung in Deutschland“ (2021)

Variablenname	Beschreibung
wgt_design	Design-Gewicht
wgt_cal_n	Redressement-Gewicht

[Design-Gewichtung] Wie in Kapitel 3 beschrieben, sind Studierende einiger Hochschulen aufgrund des Stichprobendesigns überrepräsentiert. Folglich weisen diese Individuen eine höhere Auswahlwahrscheinlichkeit auf, in die Stichprobe zu gelangen, als andere Studierende. Deshalb wurde zunächst ein Designgewicht erstellt, um die unterschiedlichen Inklusionswahrscheinlichkeiten von Studierenden auszugleichen. Diese Korrektur kann durch die Verwendung der inversen Auswahlwahrscheinlichkeit (π_i) vorgenommen werden (*Horvitz-Thompson-Schätzer*). Das Designgewicht wgt_d_i ist damit:

$$wgt_d_i = \frac{1}{\pi_i}$$

Elemente, die mit höherer Wahrscheinlichkeit als andere in die Stichprobe eingehen, wie zum Beispiel Studierende, deren Hochschulen eine Vollerhebung durchgeführt haben, erhalten somit ein niedrigeres Gewicht und umgekehrt.

[Kalibrierung der Design-Gewichtung] Da auf individueller Ebene keine Informationen über die Nicht-Teilnehmenden vorliegen, wurde in einem nächsten Schritt ein Redressement-Gewicht, basierend auf dem „Iterative Proportional Fitting-Ansatz“, erstellt (Kolenikov, 2014). Dazu wurde das Designgewicht anhand der aus der amtlichen Statistik bekannten Randverteilungen demografischer Merkmale der Studierendenschaft (Alter, Geschlecht, internationale Studierende) sowie der Merkmale Fächergruppe,¹⁷ Bundesland und Hochschulart für das Sommersemester 2021 kalibriert.

[Hinweise zur Anwendung der Gewichte] Das erstellte Gewicht kann in *Stata* mit Hilfe *ado-spezifischer* Optionen berücksichtigt werden.¹⁸ Das Gewicht „wgt_design“ berücksichtigt lediglich die ungleichen Inklusionswahrscheinlichkeiten der Studierenden. Das kalibrierte Designgewicht „wgt_cal_n“ beachtet darüber hinaus auch die Randverteilung der Gewichtsmerkmale im Abgleich mit der ausgewiesenen Verteilung der amtlichen (Hochschul-)Statistik.¹⁹

Grundlegend ist zu beachten, dass Gewichte nur dann sinnvolle Korrekturgrößen darstellen, wenn das verwendete Analysemodell die zur Gewichtung herangezogenen Variablen (siehe Tabelle 9) enthält oder mit diesen in einem Zusammenhang steht. Aus diesem Grund müssen Gewichte immer mit Fokus auf die zu analysierende Fragestellung verwendet werden. Im Bedarfsfall sind die Nutzer*innen angehalten, eigene, analysespezifische Gewichte zu erstellen.

¹⁷ Als Fächergruppen wurden nicht die Fächergruppe des Statistischen Bundesamts genutzt, sondern eine projektspezifische Variante, die stark an die Aufteilung der amtlichen Statistik angelehnt ist.

¹⁸ Ausführliche Erläuterungen finden sich im *Stata-Manual*, wenn der Befehl `help weights` eingegeben wird.

¹⁹ Aufgrund fehlender Informationen für das Redressement, konnte für 8.027 Befragte kein kalibriertes Gewicht „wgt_cal_n“ berechnet werden.

8 Anonymisierung

[Datenschutzrechtlicher Rahmen] Für personenbezogene Daten²⁰, die in freiwilligen Befragungen durch das DZHW erhoben werden, gelten die EU-Datenschutz-Grundverordnung (EU-DSGVO) und das Bundesdatenschutzgesetz in seiner Neufassung vom 30. Juni 2017.²¹ Danach sind personenbezogene Daten für die Weitergabe zur wissenschaftlichen Sekundärnutzung (ohne Vorliegen einer Einverständniserklärung zur Sekundärnutzung der personenbezogenen Daten) in der Regel derart aufzubereiten, dass „die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden können“ (Art. 4 Abs. 5 DSGVO; s. auch Art. 89 DSGVO sowie Erwägungsgrund 26 DSGVO).

[Sicherstellung des Datenschutzes im FDZ-DZHW] Im FDZ-DZHW wird der Datenschutz der Befragten über rechtlich-organisatorische Maßnahmen, technische Maßnahmen und statistische Maßnahmen sichergestellt²², die in unterschiedlicher Weise kombiniert werden können. Aus den Kombinationen ergeben sich vier potenzielle Zugangswege, die in Tabelle 11 im Überblick dargestellt sind.²³

Tabelle 11: Zugangswege im FDZ-DZHW

	Scientific Use File (SUF): On-Site	Scientific Use File (SUF): Remote-Desktop	Scientific Use File (SUF): Download	Campus Use File (CUF): Download
rechtlich-organisatorische Maßnahmen	hoch	moderat	moderat	gering
technische Maßnahmen	hoch	moderat	gering	sehr gering
statistische Maßnahmen	gering	moderat	hoch	sehr hoch

Je stärker der Datenzugang rechtlich-organisatorisch und technisch kontrolliert und beschränkt wird, desto geringer ist das Risiko einer De-Anonymisierung der Daten, desto weniger müssen die Daten mittels statistischer Maßnahmen anonymisiert, also um Informationen reduziert, werden, d.h. desto größer bleibt ihr Analysepotenzial.

²⁰ „Personenbezogene Daten (sind) alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden „betroffene Person“) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen identifiziert werden kann, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind“ (Art. 4 DSGVO: 1).

²¹ Die DSGVO gilt grundsätzlich innerhalb der EU und somit ebenfalls für das DZHW. Das BDSG in seiner Neufassung vom 30. Juni 2017 (Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und Umsetzungsgesetz EU DSAnpUG-EU)) kommt teils zusätzlich zur Anwendung, da die DZHW GmbH juristisch als öffentliche Stelle des Bundes betrachtet wird (§ 2 Abs. 3 BDSG). Der Bund hält die absolute Mehrheit der Anteile der DZHW GmbH und das Institut erfüllt Aufgaben der öffentlichen Verwaltung des Bundes im weitesten Sinn.

²² Das Datenschutzkonzept des FDZ ist angelehnt an den Portfolio-Ansatz von Lane, Heus und Mulcahy (2008: 6ff.), an dem sich bereits das Leibniz-Institut für Bildungsverläufe (LifBi) (Koberg, 2016: 699ff.) und das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (Hochfellner, Müller, Schmucker & Roß, 2012: 9f.) orientieren.

²³ Für genauere Informationen . <https://www.fdz.dzhw.eu/de/datennutzung>.

Für die Daten von „Die Studierendenbefragung in Deutschland“ (2021) gibt es alle Zugangswege. Im Folgenden werden die konkreten durchgeführten statistischen Anonymisierungsmaßnahmen genauer erläutert.

[Statistische Anonymisierungsmaßnahmen] Im Rahmen der statistischen Anonymisierungsmaßnahmen wurden zunächst alle Informationen darauf geprüft, ob sich über sie Personen direkt identifizieren lassen. Diese direkten Identifikatoren, wie Namen, Adressen und E-Mail-Adressen, wurden bereits während der Feldphase in einem separaten Datensatz erfasst und sind somit per se nicht in den Daten enthalten. Die Original-Identifikationsnummer wurde entfernt und durch eine neue zufällig vergebene Identifikationsnummer ersetzt.

Darüber hinaus wurde der Empfehlung von Ebel und Meyermann gefolgt, offene Angaben zu löschen „selbst wenn die jeweiligen Fragestellungen an sich unproblematisch sind. Denn es besteht die Gefahr, dass Studienteilnehmer*innen bei eigentlich unbedenklichen Fragen mit offener Antwortmöglichkeit kritische Informationen preisgegeben haben, die zu einer Identifikation führen könnten“ (Ebel & Meyermann, 2015).

Anschließend wurden die Quasi-Identifikatoren bestimmt, also Informationen, die in Kombination oder durch die Anspielung externer Informationen geeignet sind, eine Person indirekt zu identifizieren. Um eine eindeutige Zuordnung der Daten zu unterbinden, werden diese Schlüsselmerkmale – je nach Zugangsweg – nicht bzw. nur in aggregierter Form freigegeben. Die betrifft sowohl Informationen über die befragte Person selbst als auch über Dritte, mit deren Kontextualisierung eine Person leichter zu de-anonymisieren wäre. Bei der Abwägung der Quasi-Identifikatoren ist bei „Die Studierendenbefragung in Deutschland“ (2021) zu berücksichtigen, dass die Kombination bestimmter Schlüsselmerkmale bei einer befragten Person im Datensatz einmalig sein kann, dies aber in der Grundgesamtheit mehrere Personen betreffen könnte. Dies ergibt sich aus der Teilnahmebereitschaft bzw. der Nichtteilnahme von Hochschulen, der Ziehung von einzuladenden Studierenden aus der Auswahlgesamtheit, dem Ausfall von potenziellen Befragten durch Nichtteilnahme sowie den späteren Ausschluss von Personen im Rahmen der Datenaufbereitung und Rücklaufkontrolle (siehe Kapitel 5 und 6).

Zuletzt wurde geprüft, ob in den Daten sensible Informationen, z. B. zur Gesundheit, sexuellen Orientierung und zu politischen Einstellungen, enthalten waren. Diese eignen sich zwar nicht notwendig zur Re-Identifikation von Individuen oder Institutionen, jedoch können die Informationen im Falle einer De-Anonymisierung nutzbringend sein (Koberg, 2016) und sind daher besonders schützenswert (Art. 9 DSGVO, Erwägungsgrund 51 DSGVO). Die sensiblen Informationen werden – je nach Zugangsweg – nicht bzw. nur in aggregierter Form freigegeben.

Ein Überblick, welche Variablen in den bereitgestellten Daten enthalten sind und über welchen Zugangsweg sie jeweils nutzbar sind, findet sich im jeweiligen Datensatzreport der Datensätze des Datenpakets.

Für den Zugangsweg „Campus Use File (CUF): Download“ wurde der höchste Grad an Informationsreduktion vorgenommen. Abweichend von anderen Datenpaketen am FDZ-DZHW wurde auf das Beschränken des CUF auf eine per Zufallsauswahl gewonnene Substichprobe der Daten verzichtet, um der komplexen modularisierten Struktur der Befragung Rechnung zu tragen und ausreichende Analysemöglichkeiten für Lehr- und Übungszwecke zu ermöglichen.

9 Literatur

- AAPOR (2016). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR.
<https://aapor.org/wp-content/uploads/2022/11/Standard-Definitions20169theditionfinal.pdf>
(Abruf: 26.03.2024)
- AAPOR (2023). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. American Association for Public Opinion Research. 10th Edition. AAPOR.
<https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf>
(Abruf: 26.03.2024)
- Augspurg, K. & Hinz, T. (2015). Factorial survey experiments. Sage
- Becker, K., Baillet, F. & Weber, A. (2019). 21. Sozialerhebung. Daten- und Methodenbericht zu der Erhebung der wirtschaftlichen und sozialen Lage der Studierenden 2016. FDZ-DZHW.
[https://metadata.fdz.dzhw.eu/public/files/studies/stu-ssy21\\$/attachments/ssy21_MethodReport_de.pdf](https://metadata.fdz.dzhw.eu/public/files/studies/stu-ssy21$/attachments/ssy21_MethodReport_de.pdf) (Abruf: 32.01.2024)
- Beuße, M., Kroher, M., Becker, K., Ehrhardt, M., Isleib, S., Koopmann, J., Steinkühler, J., Völk, D., Buchholz, S., Meyer, J., Multrus, F., Hinz, T., Marczuk, A., & Strauß, S. (2022). Die Studierendenbefragung in Deutschland: Eine neue, integrierte Datenbasis für Forschung, Bildungs- und Hochschulpolitik. (DZHW Brief 06| 2022). DZHW.
https://doi.org/10.34878/2022.06.dzhw_brief
- Cuppen, J., Muja, A., Hauschildt, K., Buck, D. & Daniel, A. (2023). Eurostudent VII. Micro Data Data and methods report SUF version 3.0. DZHW.
[https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-es7\\$-3.1.0/attachments/es7_methodreport_en_3_0.pdf](https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-es7$-3.1.0/attachments/es7_methodreport_en_3_0.pdf)
(Abruf: 05.02.2024)
- Couper, M. P. (1998): Measuring Survey Quality in a CASIC Environment. In American Statistical Association. *Proceedings of the Survey Research Methods Section* (S. 41-49). American Statistical Association.
- Datenschutz-Grundverordnung (DSGVO).
<https://dsgvo-gesetz.de/> (Abruf: 15.02.2024)
- DZHW (Eds.) (2018). Social and economic conditions of student life in Europe: Synopsis of Indicators. EUROSTUDENT VI 2016-2018. W. Bertelsmann Verlag.
https://www.eurostudent.eu/download_files/documents/EUROSTUDENT_VI_Synopsis_of_Indicators.pdf/
(Abruf: 22.01.2024)
- Destatis (2020). Schlüsselerzeichnisse für die Studenten- und Prüfungsstatistik, Promovierendenstatistik und Gasthörerstatistik. WS 2020/2021 und SS 2021. Statistisches Bundesamt.
https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1_Studenten_2_Pr%C3%BCfungen/Schluesselverzeichnis_WS_2020.pdf
(Abruf: 26.03.2024)

- Ebel, T., Meyermann, A. (2015). Hinweise zur Anonymisierung von quantitativen Daten. forschungsdaten bildung informiert 3. Forschungsdatenzentrum (FDZ) Bildung am DIPF
- Felderer, B., Birg, A. & Kreuter, F. (2019). Paradata. In Baur, N., Blasius, J. (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung* (S.413-422). Springer VS.
https://doi.org/10.1007/978-3-658-21308-4_29
- Gerdes, F., & Völk, D. (2022). „Landesweite Studierendenbefragung an den Hochschulen in Nordrhein-Westfalen“ im Kontext der „Studierendenbefragung in Deutschland“. Ergebnisbericht. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (nicht zur Veröffentlichung vorgesehen).
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349–360.
<https://doi.org/10.1093/poq/nfp031>
- Gesetz über die Statistik für das Hochschulwesen sowie für die Berufsakademien (Hochschulstatistikgesetz- HStatG). https://www.gesetze-im-internet.de/hstatg_1990/index.html
(Abruf: 26.03.2024)
- Görritz, A. S., & Crutzen, R. (2012). Reminders in Web-Based Data Collection: Increasing Response at the Price of Retention?. *American Journal of Evaluation*, 33(2), 240-250.
<https://doi.org/10.1177/1098214011421956>
- Hochfellner, D., Müller, D., Schmucker, A. & Roß, E. (2012). FDZ-Methodenreport. Datenschutz am Forschungsdatenzentrum (Nr. 06). Institut für Arbeitsmarkt- und Berufsforschung (IAB).
- Jungbauer-Gans, M. (2018). Forschungsperspektiven am DZHW. In M. Fuhrmann, J. Güdler, J. Kohler, P. Pohlenz & U. Schmidt (Hrsg.), *Handbuch Qualität in Studium, Lehre und Forschung* (B4.19). DUZ Verlags- und Medienhaus.
- Kiesl, H. (2022). Gewichtung. In Baur, N., Blasius, J. (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung* (S.473-481). Springer VS.
https://doi.org/10.1007/978-3-658-37985-8_30
- KldB (2010). Klassifikation der Berufe. Bundesagentur für Arbeit.
<https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/Klassifikation-der-Berufe-Nav.html>
(Abruf: 07.02.2024)
- Koberg, T. (2016). Disclosing the National Educational Panel Study. In H.-P. Blossfeld, J. v. Maurice, M. Bayer & J. Skopek (Hrsg.), *Methodological Issues of Longitudinal Surveys. The example of the National Educational Panel Study* (S. 691–708). Springer VS.
<https://doi.org/10.1007/978-3-658-11994-2>
- Kolenikov, S. (2014). Calibrating survey data using iterative proportional fitting (raking). *The Stata Journal*, 14(1), 22–59.
- Kreuter, F. (2013): Improving Surveys with Paradata. Analytic Uses of Process Information. Wiley
- Kroher, M., Beuße, M., Isleib, S., Becker, K., Ehrhardt, M. C., Gerdes, F., Koopmann, J., Schommer, T., Schwabe, U., Steinkühler, J., Völk, D., Peter, F., & Buchholz, S. (2023). Die Studierendenbefragung in Deutschland: 22. Sozialerhebung: Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2021. Leibniz Universität Hannover.
<https://doi.org/10.15488/15772>

- Lane, J., Heus, P. & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1(1), 2–16.
- Lenzner, T., Neuert, C., & Otto, W. (2015). Kognitives Pretesting. GESIS – Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines).
https://doi.org/10.15465/gesis-sg_010
- Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I. & Vehovar, V. (2008). Web Surveys versus other Survey Modes: A Meta-Analysis Comparing Response Rates. *International Journal of Market Research*, 50(1), 79-104.
- Meier, D. H., Thomsen, S. L., & Kroher, M. (2023). Die Bedeutung der Inflation für die wirtschaftliche Situation von Studierenden in Deutschland im Zeitraum 2021 bis 2024: Eine Abschätzung. (DZHW Brief 01 | 2023). DZHW.
https://doi.org/10.34878/2023.01.dzhw_brief
- Meyer, J., Strauß, S., & Hinz, T. (2022). Die Studierendenbefragung in Deutschland: Fokusanalysen zu Diskriminierungserfahrungen an Hochschulen. (DZHW Brief 08 | 2022). DZHW.
https://doi.org/10.34878/2022.08.dzhw_brief
- Middendorff, E. (2022). Die Sozialerhebungen des Deutschen Studentenwerks 1951 – 2016. Ein historischer Überblick über Akteure, Wellen, Methoden, Themen und projektbezogene Publikationen (Working Paper). Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW).
[https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-ssypool\\$-3.0.0/attachments/SE_History_1-21_Sept22.pdf](https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-ssypool$-3.0.0/attachments/SE_History_1-21_Sept22.pdf)
(Abruf: 26.06.2024)
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017). Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016: 21. Sozialerhebung des Deutschen Studentenwerks. Bundesministerium für Bildung und Forschung.
- Middendorff, E., Wallis, M. (2023). 10.-21-Sozialerhebung 1982-2016. Daten- und Methodenbericht zum gepoolten Datensatz der zwölf Studierendenbefragungen. DZHW.
- Multrus, F. (2021). Daten- und Methodenbericht des Datenpakets „Konstanzer Studierendensurvey 1983- 2016“. AG Hochschulforschung, Universität Konstanz.
- Multrus, F.S., Majer, T., Bargel, M., Schmidt (2017). Studiensituation und studentische Orientierungen. 13. Studierendensurvey an Universitäten und Fachhochschulen. Bundesministerium für Bildung und Forschung.
- Multrus, F., Marczuk, A., Hinz, T. & Strauß, S. (2023). Studiensituation unter Onlinebedingungen. Studieren während der Corona-Pandemie 2021. Bundesministerium für Bildung und Forschung.
<https://www.soziologie.uni-konstanz.de/ag-hochschulforschung/publikationen/publikationsarten/hauptberichte/>
(Abruf: 18.01.2024)
- Multrus, F., Strauß, S., & Hinz, T. (2022). Die Studierendenbefragung in Deutschland: Fokusanalysen zu populistischen Tendenzen unter Studierenden. (DZHW Brief 07 | 2022). DZHW.
https://doi.org/10.34878/2022.07.dzhw_brief
- Netz, N., & Völk, D. (2023). Studienbezogene Auslandsmobilität und soziale Ungleichheiten im Kontext der Coronapandemie. (DZHW Brief 03 | 2023). DZHW.
https://doi.org/10.34878/2023.03.dzhw_brief

- Petzold, K., & Netz, N. (2022). Vignettenexperimente in der Hochschul- und Wissenschaftsforschung: Konstruktion, Potenziale und Fallstricke illustriert am Beispiel einer Professor*innenbefragung. In G. Brandt & S. de Vogel (Hrsg.), *Survey-Methoden in der Hochschulforschung* (S. 153-199). Springer.
https://doi.org/10.1007/978-3-658-36921-7_7
- Poskowsky, J., Heißenberg, S., Zaussinger, S., & Brenner, J. (2018). Beeinträchtigt Studieren – best2: Datenerhebung zur Situation Studierender mit Behinderung und chronischer Krankheit 2016/17. Deutsches Studentenwerk. https://www.studierendenwerke.de/fileadmin/api/files/beeintraehtigt_studieren_2016_barrierefrei.pdf
(Abruf: 16.01.2024)
- Prüfer, P., & Rexroth, M. (2000). Zwei-Phasen-Pretesting. (ZUMA-Arbeitsbericht 2000/08). Zentrum für Umfragen, Methoden und Analysen (ZUMA).
<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200864>
(Zugegriffen am: 23.01.2024).
- Roßmann, J. (2017). Satisficing in Befragungen: Theorie, Messung und Erklärung. Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute. Springer VS
<https://doi.org/10.1007/978-3-658-16668-7>
- Safikhani, P., Avetisyan, H., Föste-Eggers, D., & Broneske, D. (2023). Automated occupation coding with hierarchical features: A data-centric approach to classification with pre-trained language models. *Discover Artificial Intelligence* 3, 2023(6).
<https://doi.org/10.1007/s44163-023-00050-y>
- Schaeper, H. & Weiß, T. (2016). The Conceptualization, Development, and Validation of an Instrument for Measuring the Formal Learning Environment in Higher Education. In H.-P.Blossfeld, J. von Maurice, M. Bayer & J. Skopek, (Hrsg.), *Methodological Issues of Longitudinal Surveys. The Example of the National Educational Panel Study* (S. 267–290). Springer VS.
<https://doi.org/10.1007/978-3-658-11994-2>
- Statistisches Bundesamt (2022). Studierende an Hochschulen – Sommersemester 2021 (Fachserie 11 Reihe 4.1). Statistisches Bundesamt (Destatis).
https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/studierende-hochschulen-ss-2110410217314.pdf?__blob=publicationFile
(Abruf: 10.05.2023)
- Stefani, A., Hinz, T., & Strauß, S. (2023). Die Studierendenbefragung in Deutschland: Fokusanalysen zur Attraktivität von Masterstudiengängen. (DZHW Brief 02|2023). DZHW.
https://doi.org/10.34878/2023.02.dzhw_brief
- Steinkühler, J., Beuße, M., Kroher, M., Gerdes, F., Schwabe, U., Koopmann, J., Becker, B., Völk, D., Schommer, T., Ehrhardt, M.-C., Sören, I., & Buchholz, S. (2023). Die Studierendenbefragung in Deutschland: best3. Studieren mit einer gesundheitlichen Beeinträchtigung. Leibniz Universität Hannover
<https://doi.org/10.15488/15773>
- Tinto, V. (1975). Dropout from Higher Education. A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), 89–125.
<https://doi.org/10.3102/00346543045001089>
- Unger, M., Wejwar, P., Zaussinger, S., & Laimer, A. (2012). Beeinträchtigt Studieren: Datenerhebung zur

Situation Studierender mit Behinderung und chronischer Krankheit 2011. Deutsches Studentenwerk.

Wagner-Schelewsky, P. & Hering, L. (2022). Online-Befragung. In Baur, N., Blasius, J. (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S.1051-1065), Springer VS. https://doi.org/10.1007/978-3-658-37985-8_70

Weeber, F., Hinz, T., Meyer, J., Multrus, F. (2024). Rule-based semi-automated coding procedure to classify fields of study in surveys among university students in Germany. *Hefte zur Bildungs- und Hochschulforschung* 99. Higher Education Research Group (AG Hochschulforschung), University of Konstanz. <https://doi.org/10.48787/kops/352-2-15qrl6ayxtl7u8>

Weiß, B., Silber, H., Struminskaya, B., Durrant, G. (2022). Mobile Befragungen. In Baur, N., Blasius, J. (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung* (S.1067-1080), Springer VS. https://doi.org/10.1007/978-3-658-37985-8_71

Weichbold, M. (2022). Pretests. In: Baur, N. & Blasius, J. (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung*. Springer VS (S.443-451). https://doi.org/10.1007/978-3-658-37985-8_28

Anhang

Sekundär genutzte Instrumente

Tabelle 12: Quellennachweise zu eingesetzten Erhebungsinstrumenten

Fragenummer	Quellenangabe
Modul A	
A_13	Adaptiert nach Institut für Arbeitsmarkt- und Berufsforschung (2018) Arbeitsmarkt und Soziale Sicherung - Welle 11 (2017). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung. (DOI:10.5164/IAB.FDZD.1806.de.v1) Seniorenfragebogen PP1500
A_14	Adaptiert nach Institut für Arbeitsmarkt- und Berufsforschung (2018) Arbeitsmarkt und Soziale Sicherung - Welle 11 (2017). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung. (DOI:10.5164/IAB.FDZD.1806.de.v1) Seniorenfragebogen zur Panelstudie „Arbeitsmarkt und Soziale Sicherung“. Fragenummer PP1300
A_15	Adaptiert nach Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2017). 20. Sozialerhebung – (2012). Hannover: FDZ-DZHW. (https://doi.org/10.21249/DZHW:ssy20:1.0.0). Datenkuratierung: Daniel, A., Sarcletti, A. & Vietgen, S. Die wirtschaftliche und soziale Situation der Studierenden Fragebogen. Fragenummer 41
Modul B1	
B1_2	Adaptiert nach Berger, A. (2011). Entwicklung und Validierung eines Inventars zur Erfassung positiver und negativer Attribute des Geschlechtsrollenelbstkonzepts. Universität Potsdam, S. 158
B1_3a	Adaptiert nach Eckes, T., & Six-Materna, I. (1999). Hostilität und Benevolenz: Eine Skala zur Erfassung des ambivalenten Sexismus. Zeitschrift für Sozialpsychologie, 30(4), 211–228. https://doi.org/10.1024//0044-3514.30.4.211 Adaptiert nach Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. Journal of Personality and Social Psychology, 70(3), 491–512
B1_3b	Adaptiert nach Walter J. G. (2018). Measures of gender role attitudes under revision: The example of the German General Social Survey. Social science research, 72, 170–182. https://doi.org/10.1016/j.ssresearch.2018.02.009 Adaptiert nach GESIS - Leibniz-Institut für Sozialwissenschaften (2013). Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 2012. GESIS Datenarchiv, Köln. ZA4614 Datenfile Version 1.1.1, https://doi.org/10.4232/1.11753 . ALLBUS 2012 Fragebogendokumentation, Fragenummern F020A, F020B.

Modul B2	
B2_3a	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 – Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.759, Fragenummer t34005a. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D1_18	Adaptiert nach Weber, A., Daniel, A., Becker, K., Bornkessel, P. (2018). Proximale Prädiktoren objektiver wie subjektiver Studienerfolgsindikatoren, In: Bornkessel, P. (Hrsg.) Erfolg im Studium: Konzeptionen, Befunde und Desiderate. Bielefeld: wbv Media GmbH & Co. KG. https://doi.org/10.3278/6004654w
B2_11	Adaptiert nach Institut für Arbeitsmarkt- und Berufsforschung (2018) Arbeitsmarkt und Soziale Sicherung - Welle 11 (2017). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung. (DOI:10.5164/IAB.FDZD.1806.de.v1) Personenfragebogen zur Panelstudie Arbeitsmarkt und Soziale Sicherung. Fragenummer PA0800
Modul C1	
C1_14	In Anlehnung an Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2019). 18. Sozialerhebung – (2006). Hannover: FDZ-DZHW. (https://doi.org/10.21249/DZHW:ssy18:1.0.0). Datenkuratierung: Middendorff, E. & Hoffstätter, U. Die wirtschaftliche und soziale Situation der Studierenden Fragebogen. Fragenummer 13
Modul C2	
C2_2	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer BE6
C2_2a	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer BE6
C2_7	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer SB3, SB4
C2_8	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer SB3, SB4, SB7
C2_9	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer SB3, SB4, SB7, SB8
C2_10	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer SB10
C2_12	Adaptiert nach Poskowsky, J., Heißenberg, S., Zaussinger, S., Brenner, J. (2018). <i>beeinträchtigt studieren – best2, Fragebogen</i> . Hannover, Berlin, Wien: DZHW, DSW, IHS. Fragenummer B1
Modul D1	
D1_19	Dickhäuser, O., Schöne, C., Spinath, B., Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept: Konstruktion und Überprüfung eines neuen Instrumentes. Zeitschrift für differentielle und diagnostische Psychologie 23 (4), S. 393 - 405 https://doi.org/10.1024//0170-1789.23.4.393

D1_20	Adaptiert nach Nerdinger, F. W., Blickle, G., Schaper, N., Solga, M. (2018). <i>Arbeits- und Organisationspsychologie</i> . Deutschland: Springer Berlin Heidelberg. DOI https://doi.org/10.1007/978-3-662-56666-4
D1_22	Adaptiert nach Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. <i>British Journal of Educational Psychology</i> , 71(1), S. 133–149. https://doi.org/10.1348/000709901158433
D1_23	Adaptiert nach Boerner, Sabine & Seeber, Günther & Keller, Helmut & Beinborn, Peter. (2005). <i>Lernstrategien und Lernerfolg im Studium</i> . Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie. 37. https://doi.org/10.1026/0049-8637.37.1.17f Adaptiert nach Wild, K.-P., & Schiefele, U. (1994). <i>Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens</i> . Zeitschrift für Differentielle und Diagnostische Psychologie, 15(4), 185–200. Adaptiert nach Pintrich, P.R., Smith, D., García, T., McKeachie, W. (1991). <i>Motivated Strategies for Learning Questionnaire (MSLQ) Manual</i> . https://doi.org/10.13140/RG.2.1.2547.6968 .
D1_24	Rammstedt, B., Kemper, C., Klein, M., Beierlein, C., Kovaleva, A. (2013). Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit: Big-Five-Inventory-10 (BFI-10) Methoden – Daten – Analysen. 7. 233-249. doi: 10.12758/mda.2013.013
D1_25	Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). <i>Allgemeine Selbstwirksamkeit Kurzskala (ASKU)</i> . Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis35
D1_29	Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2017) DZHW-Studienberechtigtenpanel 2008 - Ausbildungswege von Studienberechtigten – (2007-2013). Hannover: FDZ-DZHW. (https://doi.org/10.21249/DZHW:gsl2008:1.0.0). Datenkuratierung: Daniel, A., Hoffstätter, U., Huß, B. & Scheller, P. <i>Schule. Und dann? Erste Befragung von Schülerinnen und Schülern des Abschlussjahrgangs 2007/2008 Fragebogen</i> . Fragenummer 16
Modul D2	
D2_2	Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). <i>Konstanzer Studierenden survey 2016 (2015-2017)</i> . Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuratierung: Daniel, A. <i>Fragebogen Forschungsprojekt Studiensituation 13</i> . Studierenden survey, Fragenummer 53, 56
D2_3	Adaptiert nach FDZ-LifBi (2019). <i>Codebook NEPS Startkohorte 5 – Studierende Hochschulstudium und Übergang in den Beruf</i> . Scientific Use File Version 12.0.0, S.2047ff., Fragenummer t243402, t243401, t243403, t246421, https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024). Adaptiert nach Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). <i>Konstanzer Studierenden survey 2016 (2015-2017)</i> . Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuratierung: Daniel, A.

	Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragennummer 17, 18, 24
D2_6a	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.2392ff., Fragennummer t245414, t245411, t245413 https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D2_7	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.2399ff., Fragennummer t245421, t245424, t245422, t245423, t246404, t243411. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D2_8	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.2405f., Fragennummer t246404, t246412. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D2_9	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.2387ff., Fragennummer t246401, t246413, t246411. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D2_11	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.2407ff., Fragennummer t244403, t244411, t244402, t244413, t244401, t244412. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
D2_12	Adaptiert nach FDZ-LIfBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S. 2396ff., Fragennummer t244421, t244422, t244423. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).
Modul D3	
D3_9	Adaptiert nach Engstler, H. Vogel, C., Böger, A. Franke, J., Klaus, D., Mahne, K., Simonson, J., Spuling, S., Wettstein, M., Wolff, J., Tesch-Römer, C. (2018) Deutscher Alterssurvey (DEAS): Instrumente der DEAS-Erhebung 2017, Berlin: Deutsches Zentrum für Altersfragen. DOI 10.5156/DEAS.2017.D.001, Fragennummer 808
D3_23	Adaptiert nach Ware, J., Jr., & Sherbourne, C.D. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection, in <i>Medical Care</i> , June 1992, Vol. 30, No. 6, S. 473-483
D1_27	Adaptiert nach Thönnissen, C., Wilhelm, B., Alt, P., Greischel, H., Walper, S. (2019): pairfam Scales and Instruments Manual, Waves 1 to 10, Release 10.0. S. 172 ff. GESIS Data Archive, Cologne. ZA5678 Data file Version 14.0.0, https://doi.org/10.4232/pairfam.5678.14.0.0
	Adaptiert nach Fliege, H., Rose, M., Arck, P., Levenstein, S. & Klapp, B. F. (2009). PSQ. Perceived Stress Questionnaire [Verfahrensdokumentation,

	<p>PSQ20-Skalenberechnung, PSQ20-Fragebogen Englisch, Deutsch, Deutsch (letzte 2 Jahre), PSQ30-Skalenberechnung, PSQ30-Fragebogen Englisch, Französisch, Deutsch, Italienisch und Spanisch]. In Leibniz-Institut für Psychologie (ZPID) (Hrsg.), Open Test Archive. Trier: ZPID. https://doi.org/10.23668/psycharchives.12937</p> <p>Adaptiert nach Levenstein, S., Prantera, C., Varvo, V., Scribano, M., Berto, E., Luzzi, C., Andreoli, A. (1993). Development of the Perceived Stress Questionnaire: A new tool for Psychosomatic Research. <i>Journal of psychosomatic research</i>. 37. S. 19-32. https://doi.org/10.1016/0022-3999(93)90120-5.</p>
D3_25	<p>Adaptiert nach Chmitorz, A., Wenzel, M., Stieglitz, R., Kunzler, A., Bagusat, C., Helmreich I, Gerlicher, A., Kampa, M., Kubiak, T., Kalisch, R., Lieb, K., Tüscher, O. (2018) Population-based validation of a German version of the Brief Resilience Scale. <i>PLoS ONE</i> 13(2): e0192761. https://doi.org/10.1371/journal.pone.0192761</p> <p>Adaptiert nach Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. <i>International journal of behavioral medicine</i>, 15(3), 194-200. DOI: 10.1080/10705500802222972</p>
Modul E2	
E2M_2	<p>Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0. Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey. Fragennummer 66</p>
E2M_4	<p>Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0. Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragennummer 67</p>
E2S_1	<p>Adaptiert nach FDZ-LifBi (2019). Codebook NEPS Startkohorte 5 — Studierende Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 12.0.0, S.194, t320410. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/12-0-0/SC5_12-0-0_Codebook_de.pdf (abgerufen am 20.03.2024).</p>
E2S_2	<p>Adaptiert nach Holland, J.L. (1997) <i>Making Vocational Choices: A Theory of Vocational Personalities and Work Environments</i>. 3rd Edition, Psychological Assessment Resources, Odessa.</p>
E2S_3	<p>Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0. Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragennummer 8</p>
E2S_6	<p>Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0. Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragennummer 9</p>

E2S_7	Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Sime- aner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015- 2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragenummer 59, 60
E2S_8	Adaptiert nach Jänsch, V. K., & Bosse, E. (2018). Messinstrument für die Wahrnehmung von Studienanforderungen (MWS). Zusammenstellung sozial- wissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis263 Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Sime- aner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015- 2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragenummer 60
Modul E3	
E3_1	Adaptiert nach GESIS Leibniz-Institut für Sozialwissenschaften (2019) ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (2018) Köln: GE- SIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.13250 ALLBUS 2018 Fragebogendokumentation Material zu den Datensätzen der Studennummern ZA5270 und ZA5271. Fragenummer F028
E3_2	Adaptiert nach GESIS Leibniz-Institut für Sozialwissenschaften (2019) ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (2018) Köln: GE- SIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.13250 ALLBUS 2018 Fragebogendokumentation Material zu den Datensätzen der Studennummern ZA5270 und ZA5271. Fragenummer F023 Adaptiert nach Presse- und Informationsamt der Bundesregierung, Berlin (2016). Vertrauen der Bevölkerung in die Politik (2016) Köln: GESIS Datenar- chiv für Sozialwissenschaften. https://doi.org/10.4232/1.12710 Vertrauen der Bevölkerung in die Politik. Endgültiger Fragebogen. Fragenummer q7
E3_3	Adaptiert nach GESIS Leibniz-Institut für Sozialwissenschaften (2019) ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (2018) Köln: GE- SIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.13250 ALLBUS 2018 Fragebogendokumentation Material zu den Datensätzen der Studennummern ZA5270 und ZA5271. Fragenummer F033 Adaptiert nach Vehrkamp, R., Merkel, W. (2019) Populismusbarometer 2018. Populistische Einstellungen bei Wählern und Nichtwählern in Deutschland 2018. Bertelsmann Stiftung, Gütersloh, S. 26 DOI 10.11586/2018059 Adaptiert nach Mannheimer Zentrum für Europäische Sozialforschung (2015). Demokratie Audit Mannheim (2015). Köln: GESIS Datenarchiv für So- zialwissenschaften. https://doi.org/10.4232/1.12205 Mannheimer Demokra- tie Audit Fragebogen. Fragenummer 31
E3_4	Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Sime- aner, H. & Lang, D. (2022). Konstanzer Studierendensurvey 2016 (2015- 2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierendensurvey, Fragenummer 77

E3_6	Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierenden survey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierenden survey, Fragenummer 64
E3_7	Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H., Lang, D., Ramm, M., Wuttke, K., Wiehn, E. R., Sandberger, J., Gawatz, R., Lensing, G., Framhein-Peisert, G., Peisert, H., Lenke, W., Behringer, F. & Gleich, M. (2021). Konstanzer Studierenden survey 1983 - 2016. (1982-2017). Hannover: FDZ-DZHW. Datenkuratierung: Daniel, A. & Buck, D. https://doi.org/10.21249/DZHW:stsu8316:1.0.0 , Codebuch, Seite 48
Modul E4	
E4_1	Adaptiert nach Georg, W., Multrus, F., Bargel, T., Majer, S., Schmidt, M., Simeaner, H. & Lang, D. (2022). Konstanzer Studierenden survey 2016 (2015-2017). Hannover: FDZ-DZHW. https://doi.org/10.21249/DZHW:stsu16:1.0.0 . Datenkuration: Daniel, A. Fragebogen Forschungsprojekt Studiensituation 13. Studierenden survey, Fragenummer 73
E4_3	Adaptiert nach GESIS Leibniz-Institut für Sozialwissenschaften (2019) ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (2018) Köln: GESIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.13250 ALLBUS 2018 Fragebogendokumentation Material zu den Datensätzen der Studiennummern ZA5270 und ZA5271. Fragenummer F025 Adaptiert nach GESIS Leibniz-Institut für Sozialwissenschaften (2018) ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (2014) Köln: GESIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.13141 ALLBUS 2014 Fragebogendokumentation Material zu den Datensätzen der Studiennummern ZA5240 und ZA5241. Fragenummer F056
E4_4	Adaptiert nach Sandberger, J. U. (1997). Aufstiegsmobilität. Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis213 Adaptiert nach SSP Research Group (2017). International Social Survey Programme: Social Inequality IV – ISSP 2009 (2008-2012) Köln: GESIS Datenarchiv für Sozialwissenschaften. https://doi.org/10.4232/1.12777 International Social Survey Programme 2009 Social Inequality IV Final questionnaire. Fragenummer Q1
E4_5	Revers, M., Traunmüller, R. Is Free Speech in Danger on University Campus? Some Preliminary Evidence from a Most Likely Case. Köln Z Soziol 72, 471–497 (2020). https://doi.org/10.1007/s11577-020-00713-z
E4_9	Adaptiert nach Müller, U. & Kellmer, A. (2011) Diskriminierungserfahrungen von Studierenden Ergebnisse der großen UDE-Studierendenbefragung. Universität Duisburg-Essen
Modul F1	
F1_1	Adaptiert nach Petzold, K., Peter, T. The social norm to study abroad: determinants and effects. High Educ 69, 885–900 (2015). https://doi.org/10.1007/s10734-014-9811-4
F1_9	Adaptiert nach Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2019). 21. Sozialerhebung – (2016). Hannover: FDZ-DZHW. (https://doi.org/10.21249/DZHW:ssy21:2.0.0). Datenkuratierung: Baillet, F.

& Weber, A. 21. Sozialerhebung Fragebogen-Screenshots zu der Erhebung der wirtschaftlichen und sozialen Lage der Studierenden 2016: Bildungsausländer(innen). Fragenummer B51

Modul F2

F2_8 Adaptiert nach Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2019). 21. Sozialerhebung – (2016). Hannover: FDZ-DZHW. (<https://doi.org/10.21249/DZHW:ssy21:2.0.0>). Datenkuratierung: Baillet, F. & Weber, A. 21. Sozialerhebung Fragebogen-Screenshots zu der Erhebung der wirtschaftlichen und sozialen Lage der Studierenden 2016: Bildungsausländer(innen). Fragenummer B51

Modul F3

F3_6 Adaptiert nach Petzold, K., Peter, T. The social norm to study abroad: determinants and effects. High Educ 69, 885–900 (2015). <https://doi.org/10.1007/s10734-014-9811-4>

Zusätzliche Fragen

N_16 Otte, G. (2019). Weiterentwicklung der Lebensführungstypologie, Version 2019. Johannes Gutenberg-Universität Mainz

Übersicht generierte Variablen

Tabelle 13: Bedeutende, vom Primärforschungsprojekt generierte Variablen - im Datensatz verfügbar

Beschreibung der Variablen 1: Variablenname 2: Variablenlabel	Beschreibung der Generierung
migraerf „Migrationshintergrund und internationale Studierende“	<p>Grundprogramm</p> <p><i>Wurde mittels zweier Hilfsvariable (migra und migra_a) gebildet, die nicht im Datensatz zur Verfügung stehen.</i></p> <p><i>migra wurde wie folgt gebildet:</i></p> <pre>gen migra = . replace migra = 1 if vsbdeba==1 & dgebort == 1 replace migra = 3 if vsbdeba==1 & deltegebv == 2 replace migra = 3 if vsbdeba==1 & deltegebv == 2 replace migra = 2 if vsbdeba==1 & dgebort == 2</pre> <p><i>migra_a wurde wie folgt gebildet:</i></p> <pre>gen migra_a = migra recode migra_a (3=2) replace migra_a =.a if vsbdeba==.a replace migra_a =.a if dgebort==.a & deltegebv==.a & deltegebv ==.a replace migra_a =.a if vsbdeba==1 & dgebort==.a & deltegebv!=2 & deltegebv !=2</pre> <p><i>migraerf wurde wie folgt gebildet:</i></p> <pre>gen migraerf = . replace migraerf = 1 if migra_a == 1 replace migraerf = 2 if migra_a == 2 replace migraerf = 3 if intstud == 1</pre>
dnatsta_p „ausländische Staatsangehörigkeit: Land“	<pre>gen dnatsta_p = dnatsta replace dnatsta_p = dnatstao_g1 if dnatsta_p ==.a replace dnatsta_p = dnatstao_g2 if dnatsta_p ==.a</pre>
beeintrart „Art der Beeinträchtigung“	<pre>gen beeintrart = beeintr_art recode beeintrart (4=8) (5=4) (6=5) (7=6) (8=7) (10=8)</pre>
<i>Finale Berichtsvariable</i>	
hs_p_g1 „Hochschule akt. Studium: Hochschule“	<p><i>Für die Variable hs_p wurde ein Abgleich zwischen der Hochschule, über die die Studierenden zur Befragung eingeladen wurden und der Hochschule, die die Studierenden im Fragebogen angegeben haben, vorgenommen.</i></p>
blandhs_preload_p „Hochschule akt. Studium: Bundesland“	<p><i>Aggregation des Hochschulstandorts zu Bundesländern gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i></p>
FG1_STB_p „1. Studienfach (Fächergruppe - StaBu)“	<p><i>In dieser Variable wurden die Fächerinformationen aus Frage A_23 zusammengeführt. Die Fächergruppen basieren auf der Variable sfach1 – verkodet gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i></p>

SB1_STB_p „1. Studienfach (Studienbereich - StaBu)“	<i>Die Studienbereiche basieren auf der Variable sfach1 – verkodet gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
FG2_STB_p „2. Studienfach (Fächergruppe - StaBu)“	<i>In dieser Variable wurden die Fächerinformationen aus Frage A_23 zusammengeführt. Die Fächergruppen basieren auf der Variable sfach1 – verkodet gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
SB2_STB_p „2. Studienfach (Studienbereich - StaBu)“	<i>Die Studienbereiche basieren auf der Variable sfach2 – verkodet gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
studfach „1. Studienfach (Fächergruppen der SiD)“	<i>Den Ausgangspunkt für die Variable studfach bildet die Variable sfach1 – verkodet als Studienfächer gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021. Für die einzelnen Fachcodes wurde anschließend eine Feineinteilung zu 18 Fächergruppen vorgenommen. Aus diesen 18 Fächergruppen wurde die Variable studfach mit 15 Fächergruppen generiert. Dazu wurden die Fächergruppe „3. Sportwissenschaften“ in die Gruppe „12. Gesundheitswissenschaften“ inkludiert und die Fächergruppe „4. Psychologie“ in die „6. Sozial-, Politik-, Regional- und Verwaltungswissenschaften“. Die Fächergruppe „16. Elektrotechnik und Informationstechnik“ wurde in „15. Maschinenbau/Verfahrenstechnik (inkl. Verkehrstechnik und Nautik)“ inkludiert.</i>
sabsan_a „angestrebter Abschluss: Art (aggregiert)“	gen sabsan_a =. replace sabsan_a = 1 if sabsan==1 & inlist(sabslaja, 2, 3, 4, 5, 6) replace sabsan_a = 2 if sabsan==1 & sabslaja ==1 replace sabsan_a = 3 if sabsan==2 & inlist(sabslaja, 2, 3, 4, 5, 6) replace sabsan_a = 4 if sabsan==2 & sabslaja ==1 replace sabsan_a = 5 if sabsan==3 & inlist(sabslaja, 2, 3, 4, 5, 6) replace sabsan_a = 6 if sabsan==3 & sabslaja ==1 replace sabsan_a = 7 if inlist(sabsan, 4, 5, 7)
hsart_preload_p „Hochschule akt. Studium: Hochschulart“	<i>Aggregation des Hochschulstandortes zu drei Hochschularten gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
hs_groesse „HS. akt. Studium: Größe Hochschule“	<i>Aggregation der Hochschulstandorte zu drei Hochschulgrößen.</i>
akad_sid „Bildungsherkunft: Eltern (nicht vs. akademisch)“	gen akad_sid=. replace akad_sid=0 if (inlist(deltberuv,4,3,2,1) inlist(deltberum,4,3,2,1)) replace akad_sid=1 if (inlist(deltberuv,7,6,5) inlist(deltberum,7,6,5))
hszug „Art der HZB: Hochschulreife und berufliche HZB“	gen hszug = vsbart if inlist(vsbart,1,2,3) replace hszug = 4 if vsbtypone_a == 3
vsbtypone „Art der HZB: Zugangsweg“	<i>In der Variable wurden die Einfach- und Mehrfachantworten aus den Variablen vsbtyp1 – vsbtyp15 zusammengeführt und um die offenen Angaben aus vsbtyp9o und vsbtyp15o ergänzt.</i>
vsbtypone_a „Art der HZB: aggregierter Zugangsweg“	gen vsbtypone_a = vsbtypone replace vsbtypone_a=1 if inlist(vsbtypone, 1, 2, 3, 5, 6, 7, 8, 9) replace vsbtypone_a=2 if inlist(vsbtypone, 4) replace vsbtypone_a=3 if inlist(vsbtypone, 10, 11) replace vsbtypone_a=4 if inlist(vsbtypone, 16,17,18,19)
berqual „Berufliche Qualifizierung vor aktuellem Studium“	<i>Wurde mittels zweier Hilfsvariablen (ausbild, fortbild) gebildet, die nicht im Datensatz zur Verfügung stehen und fasst zusammen, ob</i>

	<p><i>vor dem Studium eine Ausbildung und/oder eine Fortbildung absolviert wurde.</i></p> <p><i>ausbild wurde wie folgt gebildet:</i> gen ausbild = vsbtyp11_p replace ausbild = 1 if vausbja1_p == 1</p> <p><i>fortbild wurde wie folgt gebildet:</i> gen fortbild = vsbtyp10_p replace fortbild = 1 if vausbja2_p == 1</p> <p><i>berqual wurde wie folgt gebildet:</i> gen berqual = ausbild replace berqual = 1 if fortbild ==1</p> <p>replace ausbild=0 if vausbnein ==1 replace fortbild=0 if vausbnein ==1 replace berqual=0 if vausbnein ==1</p>
<p>Städte_Kat „HS. akt. Studium: Größe Hochschulort“</p>	<p>gen Städte_Kat=0 replace Städte_Kat=1 if Einwohnerzahl <100000 replace Städte_Kat=2 if Einwohnerzahl >=100000 & Einwohnerzahl <500000 replace Städte_Kat=3 if Einwohnerzahl >=500000 & Einwohnerzahl <1000000 replace Städte_Kat=4 if Einwohnerzahl >=1000000 & Einwohnerzahl < 10000000</p>
	<p>B-Module</p>
<p>sexorient_a „Sexuelle Orientierung (aggregiert)“</p>	<p><i>Die Ausprägungen „Schwul“ und „Lesbisch“ wurden zu „Homosexuell“ zusammengefasst und „Andere Orientierung“ und „Unklar“ wurden zu „Andere Orientierung/ Unklar“ zusammengefasst</i></p> <p>gen sexorient_a = sexorient recode sexorient_a (2 3 = 2) recode sexorient_a (4 = 3) recode sexorient_a (5 6 = 4)</p>
	<p>C-Module</p>
<p>stud_erschw „Studienschwerende Beeinträchtigung Ja/Nein“</p>	<p><i>Wurde mittels zweier Hilfsvariablen (beeintr1 und studbeeintr1) gebildet, die nicht im Datensatz zur Verfügung stehen.</i></p> <p><i>beeintr1 wurde wie folgt gebildet:</i> gen beeintr1 = 0 replace beeintr1 = 1 if gartmob_p==1 gartseh_p==1 gartohr_p==1 gartspr_p==1 gartpsy_p==1 gartsom_p==1 garttls_p==1 gartson_p==1 gartka==1 gartneuro_g==1 gartass_g==1 gartads_g==1 gartmig_g==1 gartcan_g==1 tab beeintr1</p> <p><i>studbeeintr1 wurde wie folgt gebildet:</i> gen studbeeintr1 =. replace studbeeintr1 = 0 if beeintr1==1 & beschwer_p==1 replace studbeeintr1 = 1 if beeintr1==1 & beschwer_p==2 tab studbeeintr1</p> <p><i>stud_erschw wurde wie folgt gebildet:</i> gen stud_erschw = 0 replace stud_erschw = 1 if studbeeintr1==1 tab stud_erschw</p>
<p>grad_se „Ausmaß der Studienschwernis“</p>	<p>gen grad_se =.</p>

	<pre> replace grad_se = 1 if gbegrmob_g==1 gbegrseh_g==1 gbegrrohr_g==1 gbegrspr_g==1 gbegrpsy_g==1 gbegrchron_g==1 gbegrtls_g==1 gbegrson_g==1 gbegrka_g == 1 replace grad_se = 2 if gbegrmob_g==2 gbegrseh_g==2 gbegrrohr_g==2 gbegrspr_g==2 gbegrpsy_g==2 gbegrchron_g==2 gbegrtls_g==2 gbegrson_g==2 gbegrka_g == 2 replace grad_se = 3 if gbegrmob_g==3 gbegrseh_g==3 gbegrrohr_g==3 gbegrspr_g==3 gbegrpsy_g==3 gbegrchron_g==3 gbegrtls_g==3 gbegrson_g==3 gbegrka_g == 3 replace grad_se = 4 if gbegrmob_g==4 gbegrseh_g==4 gbegrrohr_g==4 gbegrspr_g==4 gbegrpsy_g==4 gbegrchron_g==4 gbegrtls_g==4 gbegrson_g==4 gbegrka_g == 4 replace grad_se = 5 if gbegrmob_g==5 gbegrseh_g==5 gbegrrohr_g==5 gbegrspr_g==5 gbegrpsy_g==5 gbegrchron_g==5 gbegrtls_g==5 gbegrson_g==5 gbegrka_g == 5 </pre>
beausw_g „Amtlich festgestellte Behinderung – Schwerbehindertenausweis“	<p><i>Gibt an, ob eine amtlich festgestellte Behinderung bzw. ein Schwerbehindertenausweis bei den Studierenden vorliegt, bereinigt um unplausible Angaben.</i></p> <pre> gen beausw_g = replace beausw_g = 1 if beausw_p==1 replace beausw_g = 2 if beausw_p==2 replace beausw_g = 3 if beausw_p==3 beauswo_p<=40 replace beausw_g = 4 if beauswo_p>=50 & beauswo_p<=100 </pre>
stfwein_g „Grund Fachwahl: Einkommens- und Karrierechancen“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p> <pre> gen stfwein_g = replace stfwein_g = 1 if stfwein_p == 1 stfweinb_p == 1 replace stfwein_g = 2 if stfwein_p == 2 stfweinb_p == 2 replace stfwein_g = 3 if stfwein_p == 3 stfweinb_p == 3 replace stfwein_g = 4 if stfwein_p == 4 stfweinb_p == 4 replace stfwein_g = 5 if stfwein_p == 5 stfweinb_p == 5 </pre>
stfwart_g „Grund Fachwahl: gute Aussichten auf sicheren Arbeitsplatz“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p> <pre> gen stfwart_g = replace stfwart_g = 1 if stfwart_p == 1 stfwartb_p == 1 replace stfwart_g = 2 if stfwart_p == 2 stfwartb_p == 2 replace stfwart_g = 3 if stfwart_p == 3 stfwartb_p == 3 replace stfwart_g = 4 if stfwart_p == 4 stfwartb_p == 4 replace stfwart_g = 5 if stfwart_p == 5 stfwartb_p == 5 </pre>
stfwalt_g „Grund Fachwahl: keine Zulassung im gewünschten Studienfach“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p> <pre> gen stfwalt_g = replace stfwalt_g = 1 if stfwalt_p == 1 stfwaltb == 1 replace stfwalt_g = 2 if stfwalt_p == 2 stfwaltb == 2 replace stfwalt_g = 3 if stfwalt_p == 3 stfwaltb == 3 replace stfwalt_g = 4 if stfwalt_p == 4 stfwaltb == 4 replace stfwalt_g = 5 if stfwalt_p == 5 stfwaltb == 5 </pre>
stfwzeit_g „Grund Fachwahl: kurze Studienzeit“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p>

	<pre>gen stfwzeit_g = replace stfwzeit_g = 1 if stfwzeit_p == 1 stfwzeitb_p == 1 replace stfwzeit_g = 2 if stfwzeit_p == 2 stfwzeitb_p == 2 replace stfwzeit_g = 3 if stfwzeit_p == 3 stfwzeitb_p == 3 replace stfwzeit_g = 4 if stfwzeit_p == 4 stfwzeitb_p == 4 replace stfwzeit_g = 5 if stfwzeit_p == 5 stfwzeitb_p == 5</pre>
stfwfam_g „Grund Fachwahl: Eltern/Familie haben mir dazu geraten“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwfam_g = replace stfwfam_g = 1 if stfwfam_p == 1 stfwfamb_p == 1 replace stfwfam_g = 2 if stfwfam_p == 2 stfwfamb_p == 2 replace stfwfam_g = 3 if stfwfam_p == 3 stfwfamb_p == 3 replace stfwfam_g = 4 if stfwfam_p == 4 stfwfamb_p == 4 replace stfwfam_g = 5 if stfwfam_p == 5 stfwfamb_p == 5</pre>
stfwber_g „Grund Fachwahl: fester Berufswunsch“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwber_g = replace stfwber_g = 1 if stfwber_p == 1 stfwberb_p == 1 replace stfwber_g = 2 if stfwber_p == 2 stfwberb_p == 2 replace stfwber_g = 3 if stfwber_p == 3 stfwberb_p == 3 replace stfwber_g = 4 if stfwber_p == 4 stfwberb_p == 4 replace stfwber_g = 5 if stfwber_p == 5 stfwberb_p == 5</pre>
stfwbega_g „Grund Fachwahl: eigene Begabung“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwbega_g = replace stfwbega_g = 1 if stfwbega_p == 1 stfwbegab_p == 1 replace stfwbega_g = 2 if stfwbega_p == 2 stfwbegab_p == 2 replace stfwbega_g = 3 if stfwbega_p == 3 stfwbegab_p == 3 replace stfwbega_g = 4 if stfwbega_p == 4 stfwbegab_p == 4 replace stfwbega_g = 5 if stfwbega_p == 5 stfwbegab_p == 5</pre>
stfwint_g „Grund Fachwahl: spezielles Fachinteresse“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwint_g = replace stfwint_g = 1 if stfwint_p == 1 stfwintb_p == 1 replace stfwint_g = 2 if stfwint_p == 2 stfwintb_p == 2 replace stfwint_g = 3 if stfwint_p == 3 stfwintb_p == 3 replace stfwint_g = 4 if stfwint_p == 4 stfwintb_p == 4 replace stfwint_g = 5 if stfwint_p == 5 stfwintb_p == 5</pre>
stfwwiss_g „Grund Fachwahl: eine gute wissenschaftliche Ausbildung“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwwiss_g = replace stfwwiss_g = 1 if stfwwiss_p == 1 stfwwissb_p == 1 replace stfwwiss_g = 2 if stfwwiss_p == 2 stfwwissb_p == 2 replace stfwwiss_g = 3 if stfwwiss_p == 3 stfwwissb_p == 3 replace stfwwiss_g = 4 if stfwwiss_p == 4 stfwwissb_p == 4 replace stfwwiss_g = 5 if stfwwiss_p == 5 stfwwissb_p == 5</pre>
stfwverf_g „Grund Fachwahl: Vereinbarkeit Familie mit späteren Beschäftigungsmögl.“	<i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i> <pre>gen stfwverf_g = replace stfwverf_g = 1 if stfwverf_p == 1 stfwverfb_p == 1 replace stfwverf_g = 2 if stfwverf_p == 2 stfwverfb_p == 2</pre>

	<pre>replace stfwverf_g = 3 if stfwverf_p == 3 stfwverfb_p == 3 replace stfwverf_g = 4 if stfwverf_p == 4 stfwverfb_p == 4 replace stfwverf_g = 5 if stfwverf_p == 5 stfwverfb_p == 5</pre>
stfwverb_g „Grund Fachwahl: Vereinbarkeit Beeinträchtigung mit späteren Beschäftigungsmögl.“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p> <pre>gen stfwverb_g =. replace stfwverb_g = 1 if stfwverb_p == 1 stfwverbb == 1 replace stfwverb_g = 2 if stfwverb_p == 2 stfwverbb == 2 replace stfwverb_g = 3 if stfwverb_p == 3 stfwverbb == 3 replace stfwverb_g = 4 if stfwverb_p == 4 stfwverbb == 4 replace stfwverb_g = 5 if stfwverb_p == 5 stfwverbb == 5</pre>
stfwand_g „Grund Fachwahl: Anderes“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_7 und E2S_3 zusammengeführt.</i></p> <pre>gen stfwand_g =. replace stfwand_g = 1 if stfwand_p == 1 stfwandb == 1 replace stfwand_g = 2 if stfwand_p == 2 stfwandb == 2 replace stfwand_g = 3 if stfwand_p == 3 stfwandb == 3 replace stfwand_g = 4 if stfwand_p == 4 stfwandb == 4 replace stfwand_g = 5 if stfwand_p == 5 stfwandb == 5</pre>
hswatt_g „Grund Hochschulwahl: Attraktivität von Stadt und Umgebung“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i></p> <pre>gen hswatt_g =. replace hswatt_g = 1 if hswatt_p==1 hswattb_p==1 replace hswatt_g = 2 if hswatt_p==2 hswattb_p==2 replace hswatt_g = 3 if hswatt_p==3 hswattb_p==3 replace hswatt_g = 4 if hswatt_p==4 hswattb_p==4 replace hswatt_g = 5 if hswatt_p==5 hswattb_p==5</pre>
hswbarr_g „Grund Hochschulwahl: Barrierefreiheit der Hochschule“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i></p> <pre>gen hswbarr_g =. replace hswbarr_g = 1 if hswbarr_p==1 hswbarrb==1 replace hswbarr_g = 2 if hswbarr_p==2 hswbarrb==2 replace hswbarr_g = 3 if hswbarr_p==3 hswbarrb==3 replace hswbarr_g = 4 if hswbarr_p==4 hswbarrb==4 replace hswbarr_g = 5 if hswbarr_p==5 hswbarrb==5</pre>
hswbed_g „Grund Hochschulwahl: günstige Lebensbedingungen am Hochschulort“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i></p> <pre>gen hswbed_g =. replace hswbed_g = 1 if hswbed_p==1 hswbedb_p==1 replace hswbed_g = 2 if hswbed_p==2 hswbedb_p==2 replace hswbed_g = 3 if hswbed_p==3 hswbedb_p==3 replace hswbed_g = 4 if hswbed_p==4 hswbedb_p==4 replace hswbed_g = 5 if hswbed_p==5 hswbedb_p==5</pre>
hswfach_g „Grund Hochschulwahl: gewünschte Fachrichtung“	<p><i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i></p> <pre>gen hswfach_g =. replace hswfach_g = 1 if hswfach_p==1 hswfachb_p==1 replace hswfach_g = 2 if hswfach_p==2 hswfachb_p==2 replace hswfach_g = 3 if hswfach_p==3 hswfachb_p==3 replace hswfach_g = 4 if hswfach_p==4 hswfachb_p==4 replace hswfach_g = 5 if hswfach_p==5 hswfachb_p==5</pre>

hswint_g „Grund Hochschulwahl: internationale Ausrichtung der Hochschule“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswint_g = replace hswint_g = 1 if hswint_p==1 hswintb_p==1 replace hswint_g = 2 if hswint_p==2 hswintb_p==2 replace hswint_g = 3 if hswint_p==3 hswintb_p==3 replace hswint_g = 4 if hswint_p==4 hswintb_p==4 replace hswint_g = 5 if hswint_p==5 hswintb_p==5</pre>
hswkont_g „Grund Hochschulwahl: Freund*innen, Familie vor Ort“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswkont_g = replace hswkont_g = 1 if hswkont_p==1 hswkontb_p==1 replace hswkont_g = 2 if hswkont_p==2 hswkontb_p==2 replace hswkont_g = 3 if hswkont_p==3 hswkontb_p==3 replace hswkont_g = 4 if hswkont_p==4 hswkontb_p==4 replace hswkont_g = 5 if hswkont_p==5 hswkontb_p==5</pre>
hswrank_g „Grund Hochschulwahl: gute Platzierung meines Fachs in Rankings“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswrank_g = replace hswrank_g = 1 if hswrank_p==1 hswrankb==1 replace hswrank_g = 2 if hswrank_p==2 hswrankb==2 replace hswrank_g = 3 if hswrank_p==3 hswrankb==3 replace hswrank_g = 4 if hswrank_p==4 hswrankb==4 replace hswrank_g = 5 if hswrank_p==5 hswrankb==5</pre>
hswsupp_g „Grund Hochschulwahl: hochschulspez. Beratungs- und Unterstützungsangeb.“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswsupp_g = replace hswsupp_g = 1 if hswsupp_p==1 hswsuppb_p==1 replace hswsupp_g = 2 if hswsupp_p==2 hswsuppb_p==2 replace hswsupp_g = 3 if hswsupp_p==3 hswsuppb_p==3 replace hswsupp_g = 4 if hswsupp_p==4 hswsuppb_p==4 replace hswsupp_g = 5 if hswsupp_p==5 hswsuppb_p==5</pre>
hswtrad_g „Grund Hochschulwahl: Tradition und Ruf der Hochschule“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswzul_g = replace hswzul_g = 1 if hswzul_p==1 hswzulb==1 replace hswzul_g = 2 if hswzul_p==2 hswzulb==2 replace hswzul_g = 3 if hswzul_p==3 hswzulb==3 replace hswzul_g = 4 if hswzul_p==4 hswzulb==4 replace hswzul_g = 5 if hswzul_p==5 hswzulb==5</pre>
hswtz_g „Grund Hochschulwahl: Möglichkeit, in Teilzeit studieren zu können“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswtz_g = replace hswtz_g = 1 if hswtz_p==1 hswtzb==1 replace hswtz_g = 2 if hswtz_p==2 hswtzb==2 replace hswtz_g = 3 if hswtz_p==3 hswtzb==3 replace hswtz_g = 4 if hswtz_p==4 hswtzb==4 replace hswtz_g = 5 if hswtz_p==5 hswtzb==5</pre>
hswzul_g „Grund Hochschulwahl: keine Zulassung an Wunschhochschule“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i>
	<pre>gen hswzul_g =</pre>

	<pre>replace hswzul_g = 1 if hswzul_p==1 hswzulb==1 replace hswzul_g = 2 if hswzul_p==2 hswzulb==2 replace hswzul_g = 3 if hswzul_p==3 hswzulb==3 replace hswzul_g = 4 if hswzul_p==4 hswzulb==4 replace hswzul_g = 5 if hswzul_p==5 hswzulb==5</pre>
hswand_g „Grund Hochschulwahl: Anderes“	<i>In dieser Variable wurden die Informationen aus Frage D1_8 und E2S_4 zusammengeführt.</i> <pre>replace hswand_g = 1 if hswand==1 hswandb==1 replace hswand_g = 2 if hswand==2 hswandb==2 replace hswand_g = 3 if hswand==3 hswandb==3 replace hswand_g = 4 if hswand==4 hswandb==4 replace hswand_g = 5 if hswand==5 hswandb==5</pre>
vfach1 „1. Fach, Verlauf/Wechsel“	<i>Den Ausgangspunkt für die Variable vfach1 bildet die Variable fach01, verkodet als Studienfächer gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
vfach2 „2. Fach, Verlauf/Wechsel“	<i>Den Ausgangspunkt für die Variable vfach2 bildet die Variable fach02, verkodet als Studienfächer gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
vfach3 „3. Fach, Verlauf/Wechsel“	<i>Den Ausgangspunkt für die Variable vfach3 bildet die Variable fach03, verkodet als Studienfächer gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
vfach4 „4. Fach, Verlauf/Wechsel“	<i>Den Ausgangspunkt für die Variable vfach4 bildet die Variable fach04, verkodet als Studienfächer gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
E-Module	
Keine Variablen aus diesen Modulen vorhanden	
F-Module	
Keine Variablen aus diesen Modulen vorhanden	

Tabelle 14: Bedeutende, vom Primärforschungsprojekt generierte Hilfs-/Berichtsvariablen - nicht im Datensatz verfügbar

Beschreibung der Variablen 1: Variablenname 2: Variablenlabel	Beschreibung der Generierung
demofam_a „Familienstand – 3er Kategorie“	Grundprogramm egen demofam_count = anycount(demofamsin demofampar demofamehe demofamaus demofamtod), val(1)
<i>Variable nicht im Datensatz verfügbar.</i>	gen demofam_a =. replace demofam_a = 4 if (demofam_count>0 & demofam_count<6) replace demofam_a = 2 if demofampar==1 & demofam_count==1 replace demofam_a = 1 if demofamehe==1 & demofam_count==1 replace demofam_a = 3 if demofamsin==1 & demofam_count==1
pfleg_heads „Pflegeausgaben“	gen pfleg_heads = pflegang1 replace pfleg_heads = 1 if pflegang1 == 1 pflegang2 == 1 replace pfleg_heads = 0 if pflegangno == 1
<i>Variable nicht im Datensatz verfügbar.</i>	
pflegangno12 „Pflegetaufgaben – Single-Choice“	gen pflegangno12 = pfleg_heads replace pflegangno12 = 1 if pflegang1 == 1 replace pflegangno12 = 2 if pflegang2 == 1 replace pflegangno12 = 3 if pflegang1==1 & pflegang2 == 1
<i>Variable nicht im Datensatz verfügbar</i>	
semob, seseh, seohr, sespr, sepsy, sesom, setls, seson, seneuro, seass, awads, semig, secan „Hilfsvariablen Beeinträchtigungsart“	gen semob =. replace semob = 0 if gartmob_p!=1 & beschwer_p==2 replace semob = 1 if gartmob_p==1 & beschwer_p==2 gen seseh =. replace seseh = 0 if gartseh_p!=1 & beschwer_p==2 replace seseh = 1 if gartseh_p==1 & beschwer_p==2 gen seohr =. replace seohr = 0 if gartohr_p!=1 & beschwer_p==2 replace seohr = 1 if gartohr_p==1 & beschwer_p==2 gen sespr =. replace sespr = 0 if gartspr_p!=1 & beschwer_p==2 replace sespr = 1 if gartspr_p==1 & beschwer_p==2 gen sepsy =. replace sepsy = 0 if gartpsy_p!=1 & beschwer_p==2 replace sepsy = 1 if gartpsy_p==1 & beschwer_p==2 gen sesom =. replace sesom = 0 if gartsom_p!=1 & beschwer_p==2 replace sesom = 1 if gartsom_p==1 & beschwer_p==2 gen setls =. replace setls = 0 if garttls_p!=1 & beschwer_p==2 replace setls = 1 if garttls_p==1 & beschwer_p==2 gen seson =. replace seson = 0 if gartson_p!=1 & beschwer_p==2 replace seson = 1 if gartson_p==1 & beschwer_p==2 gen seneuro =. replace seneuro = 0 if gartneuro_g!=1 & beschwer_p==2 replace seneuro = 1 if gartneuro_g==1 & beschwer_p==2 gen seass =. replace seass = 0 if gartass_g!=1 & beschwer_p==2 replace seass = 1 if gartass_g==1 & beschwer_p==2 gen seads =. replace seads = 0 if gartads_g!=1 & beschwer_p==2 replace seads = 1 if gartads_g==1 & beschwer_p==2
<i>Diese Variablen wurden als Hilfsvariablen erstellt und sind nicht im Datensatz verfügbar.</i>	

	<pre> gen semig = replace semig = 0 if gartmig_g!=1 & beschwer_p==2 replace semig = 1 if gartmig_g==1 & beschwer_p==2 gen secan = replace secan = 0 if gartcan_g!=1 & beschwer_p==2 replace secan = 1 if gartcan_g==1 & beschwer_p==2 </pre>
<p>se_count „Zählvariable für die Mehrfachbeeinträchtigungen“</p>	<pre> egen se_count = anycount(semob seseh seohr sespr sepsy sesom setls seson seneuro seass seads semig secan), values(1) </pre>
<p><i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i></p>	
<p>semob_ seseh_ seohr_ sespr_ sepsy_ sesom_ setls_ seson_ seneuro_ seass_ semig_ secan_ „Hilfsvariablen für die Einfachbeeinträchtigungen“</p>	<pre> gen semob_ = replace semob_ = 0 if semob==0 replace semob_ = 1 if semob==1 & se_count==1 replace semob_ = 0 if semob==1 & se_count>=2 gen seseh_ = replace seseh_ = 0 if seseh==0 replace seseh_ = 1 if seseh==1 & se_count==1 replace seseh_ = 0 if seseh==1 & se_count>=2 gen seohr_ = replace seohr_ = 0 if seohr==0 replace seohr_ = 1 if seohr==1 & se_count==1 replace seohr_ = 0 if seohr==1 & se_count>=2 gen sespr_ = replace sespr_ = 0 if sespr==0 replace sespr_ = 1 if sespr==1 & se_count==1 replace sespr_ = 0 if sespr==1 & se_count>=2 gen sepsy_ = replace sepsy_ = 0 if sepsy==0 replace sepsy_ = 1 if sepsy==1 & se_count==1 replace sepsy_ = 0 if sepsy==1 & se_count>=2 gen sesom_ = replace sesom_ = 0 if sesom==0 replace sesom_ = 1 if sesom==1 & se_count==1 replace sesom_ = 0 if sesom==1 & se_count>=2 gen setls_ = replace setls_ = 0 if setls==0 replace setls_ = 1 if setls==1 & se_count==1 replace setls_ = 0 if setls==1 & se_count>=2 gen seson_ = replace seson_ = 0 if seson==0 replace seson_ = 1 if seson==1 & se_count==1 replace seson_ = 0 if seson==1 & se_count>=2 gen seneuro_ = replace seneuro_ = 0 if seneuro==0 replace seneuro_ = 1 if seneuro==1 & se_count==1 replace seneuro_ = 0 if seneuro==1 & se_count>=2 gen seass_ = replace seass_ = 0 if seass==0 replace seass_ = 1 if seass==1 & se_count==1 replace seass_ = 0 if seass==1 & se_count>=2 gen seads_ = replace seads_ = 0 if seads==0 replace seads_ = 1 if seads==1 & se_count==1 replace seads_ = 0 if seads==1 & se_count>=2 gen semig_ = replace semig_ = 0 if semig==0 replace semig_ = 1 if semig==1 & se_count==1 </pre>
<p><i>Diese Variablen wurden als Hilfsvariablen erstellt und sind nicht im Datensatz verfügbar.</i></p>	

	<pre>replace semig_ = 0 if semig==1 & se_count>=2 gen secan_ = replace secan_ = 0 if secan==0 replace secan_ = 1 if secan==1 & se_count==1 replace secan_ = 0 if secan==1 & se_count>=2</pre>
<p>grad1_count „Zählvariable Grad der Beeinträchtigung“</p> <p><i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i></p>	<pre>egen grad1_count = anycount(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p), values(1) egen grad2_count = anycount(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p), values(2) egen grad3_count = anycount(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p), values(3) egen grad4_count = anycount(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p), values(4) egen grad5_count = anycount(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p), values(5)</pre>
<p>grad_countmax „Zählvariable stärkste Beeinträchtigung“</p> <p><i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i></p>	<pre>egen grad_countmax = rmax(gbegrmob_p gbegrseh_p gbeg- rohr_p gbegrspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbe- grson_p)</pre>
<p>semehr_ „Bildung der Mehrfachbeeinträchtigungen“</p> <p><i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i></p>	<pre>gen semehr_ = replace semehr_ = 1 if grad_countmax==5 & grad5_count>1 replace semehr_ = 1 if grad_countmax==4 & grad4_count>1 replace semehr_ = 1 if grad_countmax==3 & grad3_count>1 replace semehr_ = 1 if grad_countmax==3 & grad2_count>1 replace semehr_ = 1 if grad_countmax==1 & grad1_count>1 replace semehr_ = 0 if grad_countmax==5 & grad5_count==1 replace semehr_ = 0 if grad_countmax==4 & grad4_count==1 replace semehr_ = 0 if grad_countmax==3 & grad3_count==1 replace semehr_ = 0 if grad_countmax==2 & grad2_count==1 replace semehr_ = 0 if grad_countmax==1 & grad1_count==1</pre>
<p>beart_einf „Art der Beeinträchtigung (nur Einfachbeeinträchtigungen)“</p> <p><i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i></p>	<pre>gen beart_einf = replace beart_einf = 1 if semob==1 replace beart_einf = 2 if seseh==1 replace beart_einf = 3 if seohr==1 replace beart_einf = 4 if sespr==1 replace beart_einf = 5 if sepsy==1 replace beart_einf = 6 if sesom==1 replace beart_einf = 7 if setls==1 replace beart_einf = 8 if seson==1 replace beart_einf = 9 if seneuro==1 replace beart_einf = 10 if seass==1 replace beart_einf = 11 if seads==1 replace beart_einf = 12 if semig==1 replace beart_einf = 13 if secan==1 gen beart_einf1 = beart_einf recode beart_einf1 (9/13=8)</pre>
<p>semehr_1 „Mehrfachbeeinträchtigung anhand des Ausmaßes der Beeinträchtigung“</p>	<pre>gen semehr_1 = semehr_ replace semehr_1 = 2 if beart_einf1<=8</pre>

Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.

gbegrmob_p_h, gbegrseh_p_h, gbegrohr_p_h, gbe-grspr_p_h, gbegrpsy_p_h, gbegrchron_p_h, gbegr-rtls_p_h, gbegrson_p_h „Hilfsvariablen für die Beeinträchtigungsarten“	foreach var of varlist gbegrmob_p gbegrseh_p gbegrohr_p gbe-grspr_p gbegrpsy_p gbegrchron_p gbegrtls_p gbegrson_p { gen `var'_h = `var' replace `var'_h = -999 if mi(`var'_h)}
---	--

Diese Variablen wurden als Hilfsvariablen erstellt und sind nicht im Datensatz verfügbar.

beeintr_art „Art der Beeinträchtigung mit Mehrfachbeeinträch-tigung“ <i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i>	gen beeintr_art = beart_einf1 replace beeintr_art=1 if semob==1 & semehr_1==0 & (gbe-grmob_p_h > gbegrseh_p_h & gbegrmob_p_h > gbegrohr_p_h & gbegrmob_p_h > gbegrspr_p_h & gbegrmob_p_h > gbegrpsy_p_h & gbegrmob_p_h > gbegrchron_p_h & gbegrmob_p_h > gbegr-rtls_p_h & gbegrmob_p_h > gbegrson_p_h) replace beeintr_art=2 if seseh==1 & semehr_1==0 & (gbegr-seh_p_h > gbegrmob_p_h & gbegrseh_p_h > gbegrohr_p_h & gbegrseh_p_h > gbegrspr_p_h & gbegrseh_p_h > gbegrpsy_p_h & gbegrseh_p_h > gbegrchron_p_h & gbegrseh_p_h > gbegr-rtls_p_h & gbegrseh_p_h > gbegrson_p_h) replace beeintr_art=3 if seohr==1 & semehr_1==0 & (gbeg-rohr_p_h > gbegrmob_p_h & gbegrohr_p_h > gbegrseh_p_h & gbegrohr_p_h > gbegrspr_p_h & gbegrohr_p_h > gbegrpsy_p_h & gbegrohr_p_h > gbegrchron_p_h & gbegrohr_p_h > gbegr-rtls_p_h & gbegrohr_p_h > gbegrson_p_h) replace beeintr_art=5 if sepsy==1 & semehr_1==0 & (gbe-grpsy_p_h > gbegrmob_p_h & gbegrpsy_p_h > gbegrseh_p_h & gbegrpsy_p_h > gbegrohr_p_h & gbegrpsy_p_h > gbegrspr_p_h & gbegrpsy_p_h > gbegrchron_p_h & gbegrpsy_p_h > gbegr-rtls_p_h & gbegrpsy_p_h > gbegrson_p_h) replace beeintr_art=6 if sesom==1 & semehr_1==0 & (gbe-grchron_p_h > gbegrmob_p_h & gbegrchron_p_h > gbegrseh_p_h & gbegrchron_p_h > gbegrohr_p_h & gbegrchron_p_h > gbe-grspr_p_h & gbegrchron_p_h > gbegrpsy_p_h & gbegrchron_p_h > gbegr-rtls_p_h & gbegrchron_p_h > gbegrson_p_h) replace beeintr_art=7 if setls==1 & semehr_1==0 & (gbegr-rtls_p_h > gbegrmob_p_h & gbegrtls_p_h > gbegrseh_p_h & gbegr-rtls_p_h > gbegrohr_p_h & gbegrtls_p_h > gbegrspr_p_h & gbegr-rtls_p_h > gbegrpsy_p_h & gbegr-rtls_p_h > gbegrchron_p_h & gbegr-rtls_p_h > gbegrson_p_h) replace beeintr_art=8 if seson==1 & semehr_1==0 & (gbe-grson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h & gbegrson_p_h > gbegrohr_p_h & gbegrson_p_h > gbegrspr_p_h & gbegrson_p_h > gbegrpsy_p_h & gbegrson_p_h > gbe-grchron_p_h & gbegrson_p_h > gbegr-rtls_p_h) replace beeintr_art=8 if seneuro==1 & semehr_1==0 & (gbe-grson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h & gbegrson_p_h > gbegrohr_p_h & gbegrson_p_h > gbegrspr_p_h & gbegrson_p_h > gbegrpsy_p_h & gbegrson_p_h > gbe-grchron_p_h & gbegrson_p_h > gbegr-rtls_p_h) replace beeintr_art=8 if seass==1 & semehr_1==0 & (gbe-grson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h & gbegrson_p_h > gbegrohr_p_h & gbegrson_p_h > gbegrspr_p_h & gbegrson_p_h > gbegrpsy_p_h & gbegrson_p_h > gbe-grchron_p_h & gbegrson_p_h > gbegr-rtls_p_h) replace beeintr_art=8 if seads==1 & semehr_1==0 & (gbe-grson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h &
--	---

	<pre> gbgrson_p_h > gbegrroh_p_h & gbgrson_p_h > gbegrspr_p_h & gbgrson_p_h > gbegrpsy_p_h & gbgrson_p_h > gbegrchro_n_p_h & gbegrson_p_h > gbegrtps_p_h) replace beeintr_art=8 if semig==1 & semehr_1==0 & (gbgrson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h & gbgrson_p_h > gbegrroh_p_h & gbegrson_p_h > gbegrspr_p_h & gbgrson_p_h > gbegrpsy_p_h & gbegrson_p_h > gbegrchro_n_p_h & gbegrson_p_h > gbegrtps_p_h) replace beeintr_art=8 if secan==1 & semehr_1==0 & (gbgrson_p_h > gbegrmob_p_h & gbegrson_p_h > gbegrseh_p_h & gbgrson_p_h > gbegrroh_p_h & gbegrson_p_h > gbegrspr_p_h & gbgrson_p_h > gbegrpsy_p_h & gbegrson_p_h > gbegrchro_n_p_h & gbegrson_p_h > gbegrtps_p_h) </pre>
	<pre> replace beeintr_art=10 if semehr_1==1 </pre>
studfach_best „Studienbereiche für Best-Bericht“ <i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i>	<i>Die Variable studfach_best wurde auf der Basis von studfach gebildet. Dabei wurde die Gruppe 12 „Gesundheitswissenschaften (inkl. Sport)“ in die Gruppe 11 „Medizin (Human- und Zahnmedizin)“ inkludiert. Die Variable beinhaltet dementsprechend 14 Fächergruppen.</i>
sabsan_int „Abschlussart international Studierende“ <i>Variable nicht im Datensatz verfügbar.</i>	<pre> gen sabsan_int = sabsan if intstud==1 recode sabsan_int (4 5 7 8 = 99) recode sabsan_int (6 = 4) replace sabsan_int = 5 if inlist(sabsabsba,2,3) recode sabsan_int (99 = 6) </pre>
vsbtypbin „Art der HZB (binär)“ <i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i>	<pre> recode vsbtypone (1 2 3 4 5 6 7 8 16 18 19 = 1) (10 11 17 = 2) (12 13 14 = .a), into (vsbtypbin) </pre>
Einwohnerzahl „Einwohnerzahl Hochschulstadt“ <i>Diese Variable wurde als Hilfsvariable erstellt und ist nicht im Datensatz verfügbar.</i>	<i>Offizielle Einwohnerzahl mittels der ID den Hochschulstandorten zugeordnet.</i>
B-Module	
Keine Variablen aus diesen Modulen vorhanden	
C-Module	
grad_se1 „Ausmaß der Studienschwernis (3er Ausprägung)“ <i>Variable nicht im Datensatz verfügbar.</i>	<i>Aggregiert die Ausprägungen 1 und 2 bzw. 4 und 5 der Variable grad_se.</i>
bzeitj_g „Zeitpunkt des Auftretens der Beeinträchtigung“ <i>Variable nicht im Datensatz verfügbar.</i>	<i>Der Zeitpunkt des Auftretens der Beeinträchtigung wurde mittels einer Kombination aus dem Alter und dem Hochschulsemester erstellt.</i> <pre> gen alter_imma = (demoage-(ssemhs_p/2)) if demoage>14 & demoage<=99 & ssemhs_p<. recode alter_imma (-25.5/14.5=.) gen bezeitj_g =. replace bezeitj_g = 1 if bezeit_p==1 replace bezeitj_g = 2 if alter_imma>bezeitj_p & alter_imma<. & bezeitj_p<. replace bezeitj_g = 3 if alter_imma<bezeitj_p & alter_imma<. & bezeitj_p<. </pre>

<p>einallse20 „Gesamteinnahmen“</p> <p>Variable nicht im Datensatz verfügbar.</p>	<p>Die Variable ist nicht im SUF vorhanden und wurde für die Berichte von SiD erstellt. Sie umfasst die baren und unbaren Einnahmen der Studierenden.</p> <p>Die Variable setzt sich aus der Variable feininsg und feininsg2 zusammen.</p>
<p>ges_insg „Gesamtausgaben“</p> <p>Variable nicht im Datensatz verfügbar.</p>	<p>Die Variable ist nicht im SUF vorhanden und wurde für die Berichte von SiD erstellt. Sie umfasst die eigenen Ausgaben der Studierenden, sowie die Ausgaben von Dritten.</p> <p>Die Variable setzt sich aus der Variable fausinsg und feininsg2 zusammen.</p>
<p>erw_quote1 „Erwerbstätigenquote wie in der 21. SE“</p> <p>Variable nicht im Datensatz verfügbar.</p>	<pre>gen erw_quote1 = replace erw_quote1 = 0 if eaktsens==1 replace erw_quote1 = 1 if (erw_quote==1 & sform_a !=2) (erw_quote==1 & sform_a==2 & etat_h==1)</pre>
<p>Fokustyp „Fokustyp SE“</p> <p>Variable nicht im Datensatz verfügbar.</p>	<p>Die Variable ist nicht im SUF vorhanden und wurde für die Berichte von SiD erstellt. Sie umfasst eine spezifische Studierendengruppe. Es handelt sich hierbei um unverheiratete, nicht bei den Eltern oder anderen Familienangehörigen wohnende, Erststudierende in einem Vollzeitstudium in Präsenz, die ihre Hochschulzugangsberechtigung in Deutschland erworben haben.</p> <p>Für die Generierung der Variable Fokustyp werden folgende Variablen benötigt: demofam_a, wohnel bzw. wohnfam, sabsja, sartzzeit, sformpraes und vsbdeba</p>
E-Module	
Keine Variablen aus diesen Modulen vorhanden	
F-Module	
Keine Variablen aus diesen Modulen vorhanden	

Tabelle 15: Vom Forschungsdatenzentrum (FDZ-DZHW) im Rahmen der Anonymisierung generierte Variablen - im Datensatz verfügbar

Beschreibung der Variablen 1: Variablenname 2: Variablenlabel	Beschreibung der Generierung
	Grundprogramm
pid „Identifikationsnummer Befragte (randomisiert)“	Randomisierte Identifikationsnummer auf Basis der projektinternen Identifikationsnummer zur Anonymisierung der Befragten
hs_cluster_id „Identifikationsnummer Hochschule für Clustering (randomisiert)“	Randomisierte Identifikationsnummer der Hochschulen auf Basis der Hochschulcodes aus <i>hs_p_g1</i> zur Bildung von Clustern von Befragten der gleichen Hochschule.
dnatsta_p_g2 „ausländische Staatsangehörigkeit: Land (NEPS-Regionen)“	Aggregation der Länder aus <i>dnatsta_p</i> bzw. <i>dnatsta_p_g1</i> auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
dnatsta_p_g3 „ausländische Staatsangehörigkeit: Land (Weltregionen)“	Aggregation der Länder aus <i>dnatsta_p</i> bzw. <i>dnatsta_p_g1</i> auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.
dgebsta_p_g2 „Geburtsland: Studierende*r (NEPS-Regionen)“	Aggregation der Länder aus <i>dgebsta_p</i> bzw. <i>dgebsta_p_g1</i> auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
dgebsta_p_g3 „Geburtsland: Studierende*r (Weltregionen)“	Aggregation der Länder aus <i>dgebsta_p</i> bzw. <i>dgebsta_p_g1</i> auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.
demosex_p_g „Geschlecht (männlich/weiblich/missing)“	Zusammenfassung der Angaben „divers“, „keine der genannten Kategorien“ und fehlender Angabe aus <i>demosex_p</i> zu einer Kategorie „anonymisiert“.
demoage_g1 „Alter (in Jahren) (Top-Bottom-Codierung)“	Zusammenfassung der Altersangaben aus <i>demoage_p</i> für Befragte jünger als 19 Jahre bzw. älter als 30 Jahre.
demoage_g2 „Alter (in Jahren) (kategorisiert)“	Aggregation der Altersangaben aus <i>demoage_p</i> in fünf Altersgruppen. Oberste und unterste Altersgruppe entsprechen den zusammengefassten Kategorien aus <i>demoage_g1</i> .
dsibanz_g „Anzahl Geschwister (Top-Codierung)“	Aggregation der Anzahl der Geschwister aus <i>dsibanz</i> ; Zusammenfassung von mehr als drei Geschwistern in einer Kategorie.
dsibageo1_p_g1 „1. Geschwisterteil: Alter (Top-Bottom-Codierung)“	Zusammenfassung der Altersangaben aus <i>dsibageo1_g1</i> für Geschwister jünger als 19 Jahre bzw. älter als 30 Jahre.
dsibageo2_p_g1 „2. Geschwisterteil: Alter (Top-Bottom-Codierung)“	Zusammenfassung der Altersangaben aus <i>dsibageo2_g1</i> für Geschwister jünger als 19 Jahre bzw. älter als 30 Jahre.
dsibageo3_p_g1 „3. Geschwisterteil: Alter (Top-Bottom-Codierung)“	Zusammenfassung der Altersangaben aus <i>dsibageo3_g1</i> für Geschwister jünger als 19 Jahre bzw. älter als 30 Jahre.
ssemfs_p_g „Studienphase: Anzahl Fachsemester (Top-Codierung)“	Zusammenfassung der Anzahl der Fachsemester aus <i>ssemfs_p</i> für Befragte mit 25 oder mehr Fachsemestern.
ssemhs_p_g „Studienphase: Anzahl Hochschulsemester (Top-Codierung)“	Zusammenfassung der Anzahl der Hochschulsemester aus <i>ssemhs_p</i> für Befragte mit 35 oder mehr Hochschulsemestern.
hs_p_g2 „Hochschule akt. Studium: Hochschule (Bundesland)“	Aggregation der Ortsangaben aus <i>hs_p_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
hs_p_g3 „Hochschule akt. Studium: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus <i>hs_p_g1</i> auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs_p_g4 „Hochschule akt. Studium: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus <i>hs_p_g1</i> in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“

	<i>(einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.</i>
deltjobvo_g2 „Beruf Vater (KldB2010 3-Steller)“	<i>Aggregation der vercodeten Berufsangaben aus deltjobvo_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufsgruppen nach KldB2010 (3-Steller).</i>
deltjobvo_g3 „Beruf Vater (KldB2010 2-Steller)“	<i>Aggregation der vercodeten Berufsangaben aus deltjobvo_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufshauptgruppen nach KldB2010 (2-Steller).</i>
deltjobmo_g2 „Beruf Mutter (KldB2010 3-Steller)“	<i>Aggregation der vercodeten Berufsangaben aus deltjobmo_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufsgruppen nach KldB2010 (3-Steller).</i>
deltjobmo_g3 „Beruf Mutter (KldB2010 2-Steller)“	<i>Aggregation der vercodeten Berufsangaben aus deltjobmo_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufshauptgruppen nach KldB2010 (2-Steller).</i>
deltgstav_p_g2 „Geburtsland: Vater (NEPS-Regionen)“	<i>Aggregation der Länder aus deltgstav_p_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
deltgstav_p_g3 „Geburtsland: Vater (Weltregionen)“	<i>Aggregation der Länder aus deltgstv_p_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
deltgstam_p_g2 „Geburtsland: Mutter (NEPS-Regionen)“	<i>Aggregation der Länder aus deltgstam_p_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
deltgstam_p_g3 „Geburtsland: Mutter (Weltregionen)“	<i>Aggregation der Länder aus deltgstam_p_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
B-Module	
Keine Variablen aus diesen Modulen vorhanden	
C-Module	
dkinagej_p_g „Alter: jüngstes Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinagej_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
dkinage1_p_g „Alter: einziges Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinage1_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
dkinage2_p_g „Alter: 1. Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinage2_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
dkinage3_p_g „Alter: 2. Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinage3_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
dkinage4_p_g „Alter: 3. Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinage4_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
dkinage5_p_g „Alter: 4. Kind (Altersgruppen)“	<i>Aggregation der Altersangaben aus dkinage5_p in zehn Altersgruppen, orientiert an gängigen Altersgrenzen (Kita, Schule, Jugendliche, Heranwachsende etc.).</i>
D-Module	
vausberufo_p_g2 „Berufsausbildung (KldB2010 3-Steller)“	<i>Aggregation der vercodeten Berufsangaben des Ausbildungsberufs aus vausberufo_p_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufsgruppen nach KldB2010 (3-Steller).</i>
vausberufo_p_g3 „Berufsausbildung (KldB2010 3-Steller)“	<i>Aggregation der vercodeten Berufsangaben des Ausbildungsberufs aus vausberufo_p_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufshauptgruppen nach KldB2010 (2-Steller).</i>

sabserhs1_p_g2 „letzter Hochschulabschluss: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus <i>sabserhs1_p_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
sabserhs1_p_g3 „letzter Hochschulabschluss: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus <i>sabserhs1_p_g1</i> auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
sabserhs1_p_g4 „letzter Hochschulabschluss: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus <i>sabserhs1_p_g1</i> in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
sabserhs3_g2 „letzter Hochschulabschluss: Land (NEPS-Regionen)“	Aggregation der Länder aus <i>sabserhs3_g1</i> auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
sabserhs3_g3 „letzter Hochschulabschluss: Land (Weltregionen)“	Aggregation der Länder aus <i>sabserhs3_g1</i> auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.
hs01_g2 „1. Studienepisode: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus <i>hs01_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
hs01_g3 „1. Studienepisode: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus <i>hs01_g1</i> auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs01_g4 „1. Studienepisode: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus <i>hs01_g1</i> in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
hs02_g2 „2. Studienepisode: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus <i>hs02_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
hs02_g3 „2. Studienepisode: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus <i>hs02_g1</i> auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs02_g4 „2. Studienepisode: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus <i>hs02_g1</i> in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
hs03_g2 „3. Studienepisode: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus <i>hs03_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
hs03_g3 „3. Studienepisode: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus <i>hs03_g1</i> auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs03_g4 „3. Studienepisode: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus <i>hs03_g1</i> in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
hs04_g2 „4. Studienepisode: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus <i>hs04_g1</i> auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).

hs04_g3 „4. Studienepisode: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus hs04_g1 auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs04_g4 „4. Studienepisode: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus hs04_g1 in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
hs05_g2 „5. Studienepisode: Hochschule (Bundesländer)“	Aggregation der Ortsangaben aus hs05_g1 auf 16 deutsche Bundesländern (ggf. zusätzlich Ausland).
hs05_g3 „5. Studienepisode: Hochschule (alte/neue Bundesländer)“	Aggregation der Ortsangaben aus hs05_g1 auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).
hs05_g4 „5. Studienepisode: Hochschule (Hochschulart – 2 Kategorien)“	Aggregation der Hochschulangaben aus hs05_g1 in „Universitäten“ (einschl. Pädagogischen Hochschulen, Theologischen Hochschulen, Kunst- und Musikhochschulen) und „Fachhochschulen“ (einschl. Verwaltungsfachhochschulen, Berufs-, Verwaltungs- und Wirtschaftsakademien) gemäß Schlüsselverzeichnis des Statistischen Bundesamts für das Wintersemester 2020/21 und Sommersemester 2021.
jobbbdso_p_g2 „Aktueller Beruf: Berufsbezeichnung (KldB2010 – 3-Steller)“	Aggregation der vercodeten Berufsangaben des Berufs der aktuellen Erwerbstätigkeit aus jobbbdso_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufsgruppen nach KldB2010 (3-Steller).
jobbbdso_p_g3 „Aktueller Beruf: Berufsbezeichnung (KldB2010 – 3-Steller)“	Aggregation der vercodeten Berufsangaben des Berufs der aktuellen Erwerbstätigkeit aus jobbbdso_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufshauptgruppen nach KldB2010 (2-Steller).
E-Module	
berufswahl_g2 „Angestrebter Beruf: Berufsbezeichnung (KldB2010 – 3-Steller)“	Aggregation der vercodeten Berufsangaben des angestrebten Berufs aus berufswahl_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufsgruppen nach KldB2010 (3-Steller).
berufswahl_g3 „Angestrebter Beruf: Berufsbezeichnung (KldB2010 – 2-Steller)“	Aggregation der vercodeten Berufsangaben des angestrebten Berufs aus berufswahl_g1 (Berufsgattungen nach KldB2010 – 5-Steller) auf Berufshauptgruppen nach KldB2010 (2-Steller).
F-Module	
ainflosia_p_g2 „akt. Auslandsaufenthalt: Land (NEPS-Regionen)“	Aggregation der Länder aus ainflosia_p_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
ainflosia_p_g3 „akt. Auslandsaufenthalt: Land (Weltregionen)“	Aggregation der Länder aus ainflosia_p_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.
ainflo1_g2 „letzter Auslandsaufenthalt: Land (NEPS-Regionen)“	Aggregation der Länder aus ainflo1_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
ainflo1_g3 „letzter Auslandsaufenthalt: Land (Weltregionen)“	Aggregation der Länder aus ainflo1_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.
ainflo2_g2 „vorletzter Auslandsaufenthalt: Land (NEPS-Regionen)“	Aggregation der Länder aus ainflo2_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.
ainflo2_g3 „vorletzter Auslandsaufenthalt: Land (Weltregionen)“	Aggregation der Länder aus ainflo2_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.

ainflo3_g2 „drittletzter Auslandsaufenthalt: Land (NEPS-Regionen)“	<i>Aggregation der Länder aus ainflo3_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
ainflo3_g3 „drittletzter Auslandsaufenthalt: Land (Weltregionen)“	<i>Aggregation der Länder aus ainflo3_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando1_g2 „geplanter Auslandsaufenthalt: Land 1 (NEPS-Regionen)“	<i>Aggregation der Länder aus aplanlando1_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando1_g3 „geplanter Auslandsaufenthalt: Land 1 (Weltregionen)“	<i>Aggregation der Länder aus aplanlando1_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando2_g2 „geplanter Auslandsaufenthalt: Land 2 (NEPS-Regionen)“	<i>Aggregation der Länder aus aplanlando2_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando2_g3 „geplanter Auslandsaufenthalt: Land 2 (Weltregionen)“	<i>Aggregation der Länder aus aplanlando2_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando3_g2 „geplanter Auslandsaufenthalt: Land 3 (NEPS-Regionen)“	<i>Aggregation der Länder aus aplanlando3_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
aplanlando3_g3 „geplanter Auslandsaufenthalt: Land 3 (Weltregionen)“	<i>Aggregation der Länder aus aplanlando3_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
hzb_bl_g2 „Hochschulzugangsberechtigung: Bundesland (alte/neue Bundesländer)“	<i>Aggregation der Ortsangaben aus hzb_bl_g1 auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).</i>
wohnland_g2 „Wohnort in Vorlesungszeit: Land (NEPS-Regionen)“	<i>Aggregation der Länder aus wohnland_g1 auf Basis der Anonymisierungs-Klassifikation des NEPS. Einschließlich ehemaliges Jugoslawien und ehemalige Sowjetunion. Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
wohnland_g3 „Wohnort in Vorlesungszeit: Land (Weltregionen)“	<i>Aggregation der Länder aus wohnland_g1 auf Basis von Kontinenten. Nord- und Südamerika zusammengefasst, Europa zusätzlich differenziert nach EU/Nicht-EU.</i>
wohn_bl_g2 „Wohnort in Vorlesungszeit: Bundesland (alte/neue Bundesländer)“	<i>Aggregation der Ortsangaben aus wohn_bl_g1 auf „alte Bundesländer“ (ehem. BRD ohne Berlin), „neue Bundesländer“ (ehem. DDR inkl. Berlin) und „Ausland“ (falls gegeben).</i>