

Daten- und
Methodenbericht
Oktober 2021

Ute Hoffstätter | Johanna Niebuhr | Sandra Vietgen

DZHW- Absolventenpanel 2013

Daten- und Methodenbericht zur Absol-
vent*innenkohorte 2013 (1. und 2. Befragungswelle)



fdz.DZHW.
Forschungsdatenzentrum
für Hochschul- und Wissenschaftsforschung

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

Ute Hoffstätter

Telefon +49 (0)511 450670-404

E-Mail: hoffstaetter@dzhw.eu

Johanna Niebuhr

Telefon +49 (0)511 450670-418

E-Mail: niebuhr@dzhw.eu

Sandra Vietgen

Telefon +49 (0)511 450670-405

E-Mail: vietgen@dzhw.eu

Unter Mitarbeit von:

Andreas Daniel (Kapitel 3 und 5)

Impressum

Herausgeber

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

Geschäftsführerinnen:

Prof. Dr. Monika Jungbauer-Gans

Karen Schlüter

Vorsitzender des Aufsichtsrats:

Ministerialdirigent Peter Greisler

Registergericht:

Amtsgericht Hannover | B 210251

Umsatzsteuer-Identifikationsnummer:

DE291239300

Oktober 2021

Inhaltsverzeichnis

Abbildungsverzeichnis	I
Tabellenverzeichnis	I
I. Einleitung	2
II. Datennutzungshinweise	3
1 Inhalt und Anlage der Studie	6
2 Sekundär genutzte Erhebungsinstrumente	9
3 Grundgesamtheit, Stichprobenverfahren und Durchführung	10
4 Datenaufbereitung	13
4.1 Vergabe von Variablenamen, Variablenlabels und Wertelabels	13
4.2 Codierung fehlender Werte	15
5 Gewichtung	16
6 Anonymisierung	17
7 Literaturverzeichnis	23

Abbildungsverzeichnis

Abb. 1: Datenzugangswege und Analysepotential.....	4
Abb. 2: Kohorten-Panel-Design der DZHW-Absolventenstudienreihe	6
Abb. 3: Datenzugangswege, statistischer Anonymisierungsgrad und Analysepotential der Daten des DZHW-Absolventenpanels 2013	18

Tabellenverzeichnis

Tab. 1: Anlage der Hauptuntersuchungen der DZHW-Absolventenstudienreihe von 1989 bis 2013	7
Tab. 2: Themengebetskürzel mit Wellennummerierung sowie die Themengebiete für Variablenamen des DZHW-Absolventenpanels 2013	13
Tab. 3: Systematik für fehlende Werte.....	15
Tab. 4: Maßnahmen der statistischen Anonymisierung der Daten des DZHW-Absolventenpanels 2013 nach Zugangsweg	19

I. Einleitung

Die DZHW-Absolventenpanels sind eine Untersuchungsreihe zu den beruflichen Werdegängen von Hochschulabsolvent*innen.¹ Sie werden durch das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung GmbH (DZHW) durchgeführt, vom Bundesministerium für Bildung und Forschung (BMBF) gefördert und dienen – in Ergänzung zur amtlichen Hochschulstatistik – dem nationalen Bildungsmonitoring. Seit 1989 wird jeder vierte Absolvent*innenjahrgang (Kohorte) mehrfach befragt.

Im Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung am DZHW (FDZ-DZHW) werden die Daten einiger Absolvent*innenkohorten nachträglich zum Zweck der Datennachnutzung aufbereitet und dokumentiert. Sie werden über verschiedene Zugangswege als *Scientific Use Files* (SUF) für die wissenschaftliche Sekundärnutzung und als *Campus Use Files* (CUF) für Lehr- und Übungszwecke zur Verfügung gestellt. Neben den Datensätzen der Erhebungen werden auch Dokumentationsmaterialien zu den Datensätzen und zur Durchführung der Studien bereitgestellt.

Der vorliegende Daten- und Methodenbericht ist Teil der Dokumentation zur ersten und zweiten Befragungswelle des Absolventenpanels 2013 (doi: 10.21249/DZHW:gra2013:3.0.0). Weitere Dokumentationsmaterialien zur Studie (Datensatzreports, Fragebögen, Filterführungsdiagramme etc.) können frei im Metadatensuchsystem des FDZ-DZHW (<https://metadata.fdz.dzhw.eu>) heruntergeladen werden.

Abschnitt II dieses Berichts stellt die zentralen Informationen zur Nutzung der Daten dieser Studie dar. Kapitel 1 stellt Inhalt und Anlage der Absolventenstudienreihe bis 2013² im Allgemeinen und des Absolventenpanels 2013 im Speziellen vor. In Kapitel 2 werden die sekundär eingesetzten Erhebungsinstrumente und in den Kapiteln 3 bis 4 der Erhebungsprozess beschrieben (Stichprobenziehung, Datenaufbereitung). In den Kapiteln 5 und 6 folgt die Darstellung der vorgenommenen Gewichtung und Anonymisierung.

¹ Aktuelle Informationen zur Absolventenstudienreihe können über die Website des Projektes (www.dzhw.eu/absolventen) abgerufen werden.

² Es werden nur die Kohorten bis zur hier dokumentierten Kohorte 2013 berücksichtigt.

II. Datennutzungshinweise

[Voraussetzungen der Datennutzung] Die Daten des Absolventenpanels 2013 werden durch das FDZ des DZHW entsprechend der europäischen Datenschutzgrundverordnung (EU-DSGVO) anonymisiert bereitgestellt und ausschließlich zur wissenschaftlichen Nutzung freigegeben.³ Das FDZ bietet ein *Scientific Use File* (SUF) für die wissenschaftliche Sekundärnutzung und ein *Campus Use File* (CUF) für Lehr- und Übungszwecke an.

Voraussetzungen für die Nutzung des SUF sind die Anstellung der Datennutzenden an einer wissenschaftlichen Einrichtung und der Abschluss eines Datennutzungsvertrags mit dem FDZ. Studierende oder Promovierende ohne eine Anstellung an einer wissenschaftlichen Einrichtung müssen gemeinsam mit einer betreuenden Person einen Datennutzungsvertrag abschließen. Im Zuge des Vertragsabschlusses wird durch das FDZ auch das Vorliegen eines wissenschaftlichen Nutzungsinteresses geprüft. Für die Nutzung des CUF ist eine Registrierung beim FDZ nötig. Danach wird das CUF durch das FDZ übermittelt. Ein Datennutzungsvertrag muss nicht abgeschlossen werden.

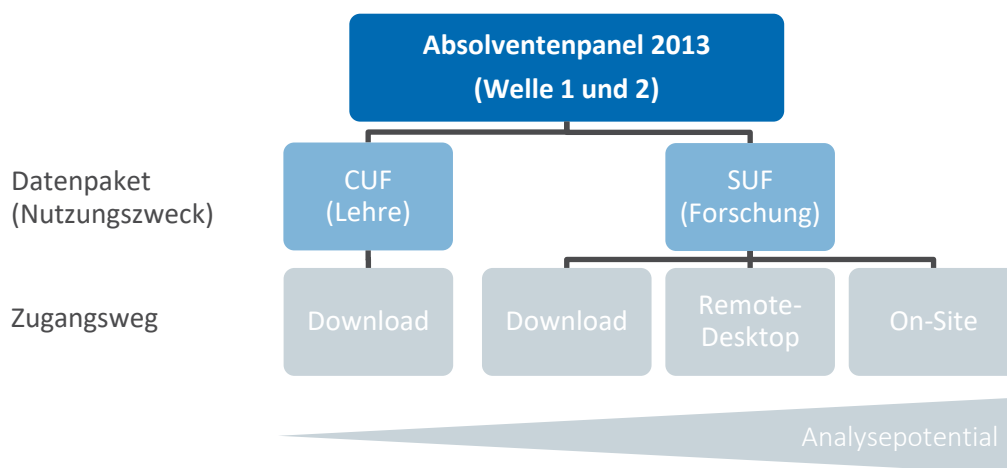
[Datenzugang] Das CUF des Absolventenpanels 2013 kann nach Bereitstellung am lokalen Computer genutzt werden. Das SUF wird über drei Zugangswege angeboten, die hinsichtlich des Speicherortes, der Möglichkeit der eigenständigen Verknüpfung mit externen Daten und der Kontrollmöglichkeiten des FDZ unterschiedlich restriktiv sind.

- **Download:** Die Daten werden verschlüsselt auf der Website des FDZ zum Download bereitgestellt. Datennutzer*innen können die Daten auf ihrem lokalen Computer speichern, falls gewünscht selbst mit Daten aus externen Quellen verknüpfen und die Daten mit eigener Software analysieren.
- **Remote-Desktop:** Die Daten werden auf einem Terminal-Server des FDZ bereitgestellt. Über eine besonders gesicherte Verbindung zwischen dem lokalen Computer der nutzenden Person und dem Terminal-Server des FDZ können die Daten mit der auf dem Terminal-Server vorhandenen Software analysiert werden. Das Übertragen der Daten auf den lokalen Computer ist nicht möglich. Analyseergebnisse werden erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.
- **On-Site:** Die Daten werden in den Räumlichkeiten des FDZ in einer kontrollierten Umgebung an einem speziell gesicherten Computer zur Analyse bereitgestellt. Wie beim Remote-Desktop-Zugang werden Analyseergebnisse erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.

³ 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 (IfBi) (vgl. Koberg, 2016, 699ff.) und das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (vgl. Hochfellner, Müller, Schmucker & Roß, 2012, 9f.) orientieren. Das FDZ des DZHW hat diesen Ansatz an die Anforderungen der eigenen Datenbestände angepasst und nutzt vier Kategorien von Maßnahmen zur Sicherstellung des Datenschutzes, die in unterschiedlicher Weise kombiniert werden: Rechtlich-institutionelle Maßnahmen, informationelle Maßnahmen, technische Maßnahmen und statistische Maßnahmen.

Die bereitgestellten Daten weisen je nach Zugangsweg einen unterschiedlich hohen Informationsgehalt auf und unterscheiden sich damit hinsichtlich ihres Analysepotentials (vgl. Abbildung 1). Dabei gilt: Je stärker der Datenzugriff der Nutzer*innen durch technische und organisatorische Maßnahmen kontrolliert wird, desto mehr Informationen können für die Datennutzer*innen bereitgestellt werden.⁴ Mit diesem Vorgehen wird ein Höchstmaß an Nutzbarkeit und gleichzeitig ein bestmöglicher Schutz der bereitgestellten Daten sichergestellt.

Abb. 1: Datenzugangswege und Analysepotential



[Datenpakete] Über den *Digital Object Identifier* (DOI) 10.21249/DZHW:gra2013:3.0.0 ist eine Webseite mit zentralen Informationen zur Studie, weiteren Dokumentationsmaterialien sowie einer Übersicht der zur Verfügung stehenden Datenpakete zur Studie erreichbar.

Die bereitgestellten Daten des Absolventenpanels 2013 sind in zwei Datensätzen abgelegt. Es liegen ein Personendatensatz im wide-Format und ein Episodendatensatz im long-Format vor. Für SUF und CUF⁵ werden für jeden im FDZ-DZHW angebotenen Zugangsweg beide Datensätze – jeweils mit zugangswegspezifischem Analysepotential (vgl. Abbildung 1) – bereitgestellt.

[Gebühren der Datenbereitstellung] CUF und SUF werden derzeit (Stand: Oktober 2021) kostenfrei zur Verfügung gestellt. Änderungen bzw. die aktuelle Gebührenordnung können auf der Website des FDZ (<https://fdz.dzhw.eu>) eingesehen werden.

[Pflichten der Datennutzer*innen] Die Datennutzer*innen sind verpflichtet, folgende Regeln⁶ einzuhalten:

- **Wissenschaftliche Nutzung:** Die Daten dürfen ausschließlich für wissenschaftliche Zwecke verwendet werden. Eine kommerzielle Nutzung ist untersagt.
- **De-Anonymisierungsverbot:** Jeder Versuch der Re-Identifikation von Analyseeinheiten (z. B. Personen, Haushalten, Institutionen) ist verboten.

⁴ Zu den verschiedenen Anonymisierungsgraden und Analysepotentialen des CUF und der verschiedenen SUF-Varianten vgl. Kapitel 6.

⁵ Aus Anonymisierungsgründen werden dabei jedoch nur die Daten einer Substichprobe bereitgestellt (vgl. Kapitel 6).

⁶ Der Datennutzungsvertrag regelt die Nutzungsbedingungen im Detail.

- **Gebot zur Mitteilung von Sicherheitslücken:** Falls Datennutzer*innen Kenntnis von Sicherheitslücken hinsichtlich Datenschutz bzw. Datensicherheit erlangen, müssen diese dem FDZ-DZHW unverzüglich angezeigt werden.
- **Keine Weitergabe der Daten:** SUF dürfen nur durch die Person genutzt werden, die den Datennutzungsvertrag abgeschlossen hat. CUF dürfen ausschließlich im Rahmen der angegebenen Lehrveranstaltung weitergegeben werden.
- **Löschungsgebot:** Download-SUF sind nach Ablauf der vereinbarten Nutzungsdauer (in der Regel 1,5 Jahre) von jeglichen Rechnern, Servern und Datenträgern zu löschen. Ebenso müssen alle Sicherungskopien, modifizierten Datensätze (z. B. Arbeits-, Auszugs- oder Hilfsdateien) sowie Ausdrucke vernichtet werden.
- **Bereitstellung/Meldung von Publikationen:** Jede Art von Publikation, die aus der Arbeit mit Daten des FDZ hervorgeht, ist dem FDZ unmittelbar nach Veröffentlichung anzuzeigen. Dabei ist dem FDZ eine elektronische Version der Druckfassung zur Verfügung zu stellen.
- **Zitationspflicht:** Die verwendeten Daten müssen in Veröffentlichungen, anderen Arbeiten (z. B. Abschlussarbeiten) und Vorträgen laut der Vorgaben des FDZ zitiert werden (vgl. Zitationsanleitung unter 10.21249/DZHW:gra2013:3.0.0).

1 Inhalt und Anlage der Studie

[Studienreihe] Das DZHW-Absolventenpanel 2013 ist Teil der DZHW-Absolventenstudienreihe, in der anhand von standardisierten Befragungen Informationen zu Studium, Berufseintritt, Berufsverlauf und Weiterqualifizierung von Hochschulabsolvent*innen erfasst werden. Das erste Absolventenpanel wurde 1989⁷ durchgeführt, seitdem wird jeder vierte Absolvent*innenjahrgang (Kohorte) untersucht. Die Grundgesamtheit einer Kohorte sind Hochschulabsolvent*innen, die im Winter- oder Sommersemester des betreffenden Prüfungsjahrs ihren Studienabschluss an einer Hochschule in Deutschland erworben haben.⁸

Für jede Absolvent*innenkohorte werden mehrere Befragungswellen durchgeführt, wobei jede Welle in unterschiedlichem zeitlichen Abstand zum Studienabschluss stattfindet. Es handelt sich somit um ein kombiniertes Kohorten-Panel-Design (vgl. Abbildung 2).

Abb. 2: Kohorten-Panel-Design der DZHW-Absolventenstudienreihe

Jahr	Absolvent*innenkohorte						
	1989	1993	1997	2001	2005	2009	2013
1989	Abschluss						
1990	1. Welle						
1991							
1992							
1993	2. Welle	Abschluss					
1994		1. Welle					
1995							
1996							
1997			Abschluss				
1998			1. Welle				
1999		2. Welle					
2000							
2001				Abschluss			
2002				1. Welle			
2003			2. Welle				
2004							
2005					Abschluss		
2006					1. Welle		
2007				2. Welle			
2008			3. Welle*				
2009						Abschluss	
2010						1. Welle	
2011					2. Welle		
2012				3. Welle*			

⁷ Schon seit 1974 werden Hochschulabsolvent*innen – neben den Studienabbrecher*innen und Hochschulwechsler*innen – im Rahmen der Exmatrikuliertenbefragung des DZHW befragt. Diese Untersuchungsreihe wird seit Anfang der 2000er Jahre unter dem Namen „Studienabbruch – Umfang und Motive“ durchgeführt.

⁸ Für die Kohorte 1989 wurden ausschließlich Absolvent*innen aus den Bundesländern der damaligen Bundesrepublik Deutschland ausgewählt.



*Hauptbefragung + Vertiefungsbefragungen

Die Befragungen der Absolvent*innenkohorten von 1989 und 1993 umfassten zwei Wellen, seit 1997 wird eine zusätzliche dritte Befragungswelle durchgeführt. Die erste Welle einer Kohorte findet im Mittel ein Jahr nach dem jeweiligen Studienabschluss statt. Die zweite Befragungswelle folgt etwa fünf Jahre nach dem Studienabschluss. Etwa zehn Jahre nach dem Abschluss schließt sich eine dritte Befragungswelle an. Teilweise setzen sich die zweiten bzw. dritten Wellen aus einer Hauptbefragung und gesonderten Vertiefungsbefragungen zu spezifischen Themen zusammen.

Die verschiedenen Erhebungen wurden als schriftlich-postalische Paper-and-Pencil-Befragung (Paper and Pencil Interview; PAPI) durchgeführt, mittlerweile aber zunehmend auch als Online-Befragung (Computer Assisted Web Interview; CAWI) (vgl. Tabelle 1).

Tab. 1: Anlage der Hauptuntersuchungen der DZHW-Absolventenstudienreihe von 1989 bis 2013

Welle	Befragungszeitpunkt	Thematischer Fokus	Absolvent*innenkohorte				
			1989 u. 1993	1997 u. 2001	2005	2009	2013
1	ca. 1 Jahr nach Studienabschluss	Studienverlauf und -erfahrung, akademische Weiterqualifizierung, Übergang in den Beruf	Paper & Pencil	Paper & Pencil	Paper & Pencil	Paper & Pencil	Paper & Pencil
2	ca. 5 Jahre nach Studienabschluss	Aktuelle Tätigkeit, Erwerbstätigkeit, akademische und berufliche Weiterbildung	Paper & Pencil	Paper & Pencil	Paper & Pencil	Online ^b	Online ^b
3	ca. 10 Jahre nach Studienabschluss	Aktuelle Tätigkeit, Erwerbstätigkeit, akademische und berufliche Weiterbildung, familiäre Situation	---	Paper & Pencil ^a	Online ^b	Online ^b	Online (geplant)

^a Die Hauptbefragung wurde als Paper & Pencil- und die Vertiefungsbefragungen als Online-Befragung durchgeführt.

^b Sowohl die Hauptbefragung als auch die Vertiefungsbefragungen wurden als Online-Befragung durchgeführt.

Die Erhebungsinstrumente aller Kohorten enthalten Fragen zum Studium, zum Übergang in den Beruf, zur akademischen und beruflichen Weiterbildung, zur Erwerbstätigkeit und zu soziodemographischen und bildungsbiographischen Merkmalen. Der thematische Fokus einer Befragungswelle orientiert sich an der jeweils typischen Bildungs-, Berufs- und Lebensphase der Befragten zum Befragungszeitpunkt.

[Analysepotential] In allen Kohorten wird je Befragungswelle ein identischer Kern an Informationen erhoben. Auf dieser Grundlage können mit Zeitreihen- bzw. Kohortenvergleichen langfristige Trends der Hochschulbildung und Arbeitsmarktentwicklung in den Blick genommen werden. Zudem wird innerhalb einer Kohorte ein Teil der Fragen in den verschiedenen Befragungswellen wiederholt gestellt. Dies ermöglicht die Betrachtung intra-individueller Veränderungen zwischen den Wellen (z. B.

kausale Panelanalysen). Besonders hervorzuheben ist, dass in allen Kohorten über die Wellen hinweg monatsgenaue kontinuierliche Verlaufsdaten zum individuellen Tätigkeitsverlauf seit dem Studienabschluss erfasst werden, die sich für Ereignisdaten- und Sequenzmusteranalysen eignen. Darüber hinaus werden in Abhängigkeit von aktuellen Entwicklungen und Forschungsinteressen in einzelnen Kohorten bestimmte Aspekte vertiefend oder ergänzend abgefragt.

[Einordnung ins Forschungsfeld] Das Stichproben- und Erhebungsdesign sowie die damit verbundenen Analysemöglichkeiten unterscheiden die DZHW-Absolventenstudienreihe von anderen in Deutschland durchgeführten Absolventenstudien. So ist beispielsweise das Bayerische Absolventenpanel (BAP) des Bayerischen Staatsinstituts für Hochschulforschung und Hochschulplanung (IHF) auf Absolvent*innen bayrischer Hochschulen begrenzt.⁹ Das Kooperationsprojekt Absolventenstudie (KOAB) des International Centre for Higher Education Research (INCHER) befragt Absolvent*innen seiner Kooperationshochschulen und ermöglicht individuelle Analysen auf Hochschul- und Studienebene, die zur Evaluation und Weiterentwicklung genutzt werden können.

[Spezifika des Absolventenpanels 2013] Neben den allgemeinen Charakteristika der Studienreihe weist die hier betrachtete Absolvent*innenkohorte 2013 folgende Spezifika auf. Noch mehr als die Absolvent*innenkohorten 2005 und 2009 ist die Studienphase der Kohorte 2013 durch den Hochschulwandel im Rahmen des Bologna-Prozesses und Veränderungen hinsichtlich der Vergleichbarkeit von Hochschulabschlüssen und daraus folgend der veränderten Mobilität geprägt. Die berufliche Einstiegsphase der Absolvent*innenkohorte 2013 ist durch einen Aufschwung nach einer starken Depression als Folge der Wirtschafts- und Finanzkrise im Jahr 2008 gekennzeichnet.

Im Unterschied zu vorangegangenen Kohorten sind in der Kohorte von 2013 Personen mit traditionellen Studienabschlüssen (Diplom- oder Magisterabschluss) nicht mehr Teil der Stichprobe (Ausnahmen bestehen für bestimmte Studiengänge in einzelnen Bundesländern, die weiterhin mit einem Diplom abschließen).

Die Stichprobe der Kohorte 2013 unterscheidet sich zudem von denen vorheriger Kohorten: Da ein großer Teil der Hochschulen aus Nordrhein-Westfalen und Baden-Württemberg die Teilnahme an der Befragung ablehnte, konnten Befragte aus diesen Bundesländern größtenteils nicht vom DZHW befragt werden und sind daher nicht Teil der Stichprobe (Details siehe Kapitel 3).

⁹ vgl. dazu <http://www.bap.ihf.bayern.de>

2 Sekundär genutzte Erhebungsinstrumente

Folgende Teile der Erhebungsinstrumente wurden sekundär genutzt:

Frage 8.4 (Welle 1) basiert auf:

Abele, A. E. & Schradi, M. (2000). *Methodisches Vorgehen und Fragebogen der ersten Erhebungswelle* (Bericht Nr. 1 des Projekts "Frauen in der Mathematik"). Erlangen: FAU Erlangen-Nürnberg.

Frage 9.11 (Welle 1) basiert auf:

Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). *Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen* (GESIS - Leibniz-Institut für Sozialwissenschaften, Hrsg.). doi:10.6102/zis35

3 Grundgesamtheit, Stichprobenverfahren und Durchführung

[Grundgesamtheit] Die Grundgesamtheit des Absolventenpanels 2013 umfasst alle Hochschulabsolvent*innen, die im Wintersemester 2012/2013 oder im Sommersemester 2013 ihren ersten berufsqualifizierenden Studienabschluss an einer staatlich anerkannten Hochschule in der Bundesrepublik Deutschland erworben haben. Hochschulabsolvent*innen von Bundeswehrhochschulen, Verwaltungsfachhochschulen, Berufsakademien und Fernhochschulen sowie weiterer Hochschulen, die sich ausschließlich an Berufstätige richten, waren dabei von vornherein ausgenommen. Ebenfalls ausgeschlossen sind Absolvent*innen von dualen und berufsbegleitenden Studiengängen sowie Online-studiengängen. Einbezogen wurden alle Bachelor-, Master-, Staatsexamens- und künstlerischen Abschlüsse. Ausgenommen von der Grundgesamtheit sind Personen, die ihr Studium mit einem Diplom- oder Magisterabschluss beendet haben.¹⁰

[Stichprobenverfahren] Aufgrund fehlender oder nicht zugänglicher Listen von Hochschulabsolvent*innen mussten die Individuen über die Hochschulen rekrutiert werden. Zu diesem Zweck wurden zunächst auf Basis der vorliegenden Studienanfänger*innenzahlen und der Studiendauern aller Studierenden an Universitäten und Fachhochschulen für jedes Bundesland einzeln diejenigen Studiengänge prognostiziert, die für den Prüfungsjahrgang 2013 voraussichtlich Absolvent*innen verzeichnen würden.¹¹ Daraus wurde eine geschichtete Klumpenstichprobe für Hochschulabsolvent*innen gezogen.

[Geschichtete Klumpenstichprobe] Die „primary sampling units“ (Klumpen) wurden anhand der Hochschule und des Studienbereichs¹² definiert. Die „secondary sampling units“ stellten die Hochschulabsolvent*innen des Prüfungsjahrgangs 2013 innerhalb dieser Klumpen dar. Die Klumpen wurden dabei zunächst nach Abschlussart¹³ und Hochschulart¹⁴ disproportional¹⁵ geschichtet. Innerhalb dieser Schichten wurde dann eine Klumpenstichprobe gezogen.

Beim Ausfall eines Klumpens (z. B. bei Teilnahmeverweigerung auf Hochschul- oder Fakultätsebene) wurde ein – hinsichtlich der Merkmalskombination Studienbereich, Hochschulart, Abschlussart – möglichst ähnlicher Klumpen als Ersatz gesucht. Bei mehreren Klumpen mit ähnlichen Merkmals-

¹⁰ Ausnahmen bestehen für bestimmte Studiengänge in einzelnen Bundesländern, die weiterhin mit einem Diplom abschließen.
¹¹ Bei Erwartung annähernd konstanter Absolvent*innenzahlen hätten die Zahlen des vorherigen Prüfungsjahrgangs herangezogen werden können. Durch die Umstellung auf die neuen Studiengänge unterschieden sich die Statistiken zwischen den Prüfungsjahrgängen jedoch zu stark.

¹² Entsprechend der Aufgliederung nach der amtlichen Statistik (gemäß Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2012/2013 und SoSe 2013).

¹³ Die Ausprägungen waren Bachelor, Master sowie Staatsexamen.

¹⁴ Die Ausprägungen waren Universität und Fachhochschule.

¹⁵ Die exakte Auswahlwahrscheinlichkeit kann für das Datenpaket nicht ermittelt werden, da hier nur der DZHW-Teil der Daten zur Verfügung gestellt wird, die Auswahlwahrscheinlichkeiten bezogen sich jedoch auf die gesamte Stichprobe inkl. der KOAB-Erhebung.

kombinationen wurde der Klumpen ausgewählt, der dem Ausgangsklumpen hinsichtlich der Fallzahlen am ähnlichsten ist.

[Durchführung] Während der Feldphase hat ein großer Teil der Hochschulen aus Nordrhein-Westfalen und Baden-Württemberg die Teilnahme an der Befragung abgelehnt, sodass Befragte aus diesen Bundesländern größtenteils nicht vom DZHW befragt werden konnten (Fabian, Hillmann, Trennt & Briedis, 2016, S. 10; Trennt, 2019, S. 380). Hintergrund war, dass sich ein Großteil dieser Hochschulen an der KOAB-Erhebung¹⁶ desselben Jahrgangs beteiligte. Aufgrund der Ähnlichkeiten beider Befragungen konnten für das Primärforschungsprojekt¹⁷ die Stichprobenausfälle durch die Daten der KOAB-Befragung kompensiert werden (Fabian et al., 2016, S. 10). Das vorliegende Datenpaket enthält aus rechtlichen Gründen nur die Daten, die durch das DZHW erhoben wurden.

Nach einem Abgleich mit den Zahlen der amtlichen Statistik des Prüfungsjahrgangs 2013 als Annäherung an die Grundgesamtheit zeigt sich eine Disproportionalität bezüglich der Schichten. So liegt der Anteil der Fachhochschulabsolvent*innen unter den Absolvent*innen mit Bachelorabschluss in der Stichprobe deutlich unter dem entsprechenden Anteil in der Grundgesamtheit (Fabian et al., 2016, S. 48). „Für die Verteilungsunterschiede können im Wesentlichen drei Gründe angeführt werden: Erstens ist bei Zufallsauswahlen immer ein Stichprobenfehler möglich [...]. Zweitens kommen Designeffekte durch disproportionale Ziehung zum Tragen, wobei die disproportionale Auswahl [...] geplant sein kann und [...] [es ist] Ergebnis der Ungenauigkeiten bei der Prognostizierung der Absolventenzahlen nach Abschlussart und Hochschulart. [...] Drittens können Verteilungsunterschiede auch durch gruppenspezifische Ausschöpfungsquoten bedingt sein. So ist auch aus anderen Untersuchungen bekannt, dass sich Studierende der Wirtschaftswissenschaften seltener an Befragungen beteiligen als Studierende anderer Fachrichtungen.“ (Fabian et al., 2016, S. 48).

[Erste Befragungswelle] Da die Hochschulen die Kontaktdaten ihrer Absolvent*innen aus Datenschutzgründen nicht herausgeben durften, teilten sie dem DZHW lediglich die jeweilige Absolvent*innenanzahl mit. Anschließend sendete das DZHW die passende Anzahl an Erhebungsunterlagen für die erste Befragungswelle postalisch an die jeweiligen Prüfungsämter, die diese an die Zielpersonen weiterverschickten.

Die Erhebungsunterlagen der ersten Befragungswelle bestanden aus einem Anschreiben (inkl. Datenschutzinformationen), einem Papierfragebogen und einem an das DZHW adressierten portofreien Umschlag zur Rücksendung des ausgefüllten Fragebogens. Außerdem wurde ein Flyer mit ausgewählten Ergebnissen bisheriger Absolvent*innenbefragungen beigelegt. Der Erhebungszeitraum der ersten Befragungswelle erstreckte sich vom 1. März 2014 bis zum 31. März 2015. Aufgrund des angewendeten Kontaktverfahrens über die Prüfungsämter konnte das DZHW keinen direkten Einfluss auf den genauen Versandzeitpunkt der Erhebungsunterlagen nehmen.

Die Nettostichprobe der ersten Befragungswelle umfasst 8.477 Fälle. Da nicht eindeutig klar ist, wie viele Absolvent*innen durch die Prüfungsämter erreicht wurden, lässt sich die Netto-Rücklaufquote nur näherungsweise bestimmen: „Die Erfahrungswerte vergleichbarer Untersuchungen sowie stichpunktartige Rückfragen bei den Prüfungsämtern lassen auf eine Quote von nicht zustellbaren Unterlagen von 15 bis 20 Prozent schließen. Unter Berücksichtigung dieser Erfahrungswerte beträgt die geschätzte Rücklaufquote 25 Prozent.“ (Fabian et al., 2016, S. 46). Um die Erhebungsunterlagen in der zweiten Befragungswelle direkt durch das DZHW verschicken zu können, wurden im Fragebogen

¹⁶ Das Kooperationsprojekt Absolventenstudien (KOAB) ist ein vom International Centre for Higher Education Research (INCHER-Kassel) koordiniertes Projekt, in dem Absolvent*innen der teilnehmenden Hochschulen zum Studium und zu Berufswegen befragt werden.

¹⁷ https://www.dzhw.eu/forschung/projekt?pr_id=304 (letzter Zugriff: 31.08.2021)

der ersten Befragungswelle die Kontaktdaten (Post- und E-Mail Adresse) der Befragungsteilnehmer*innen erfasst.

[Zweite Befragungswelle] Die zweite Befragungswelle besteht aus drei Teilerhebungen. Im vorliegenden SUF wurde lediglich die erste Teilerhebung aufbereitet. Dabei handelt es sich um die Hauptbefragung der zweiten Welle, der anschließend noch zwei kürzere Online-Vertiefungsbefragungen zu speziellen Themen folgten. Die Erhebungsunterlagen der Hauptbefragung der zweiten Welle wurden je nach vorliegenden Kontaktdaten sowohl postalisch als auch per Mail direkt an die Absolvent*innen versandt. Sowohl das postalische als auch das digitale Anschreiben enthielten einen Link zur Online-Befragung, das zugehörige Passwort und Informationen zum Datenschutz. Der Erhebungszeitraum der Hauptbefragung der zweiten Befragungswelle erstreckte sich vom 05. Oktober 2018 bis zum 08. Januar 2019. Die Nettostichprobe der Hauptbefragung der zweiten Befragungswelle umfasst 4.485 Fälle. Bezogen auf die 8.477 Befragten in der ersten Befragungswelle konnte in der zweiten Befragungswelle eine Rücklaufquote von 52,9 Prozent erzielt werden.

4 Datenaufbereitung

Im Folgenden werden die verschiedenen Schritte der Datenaufbereitung beschrieben. Die im Rahmen der Datenedition vorgenommenen Aufbereitungsprozesse der Gewichtung und Anonymisierung werden in den beiden folgenden Kapiteln 5 und 6 gesondert erläutert.

4.1 Vergabe von Variablennamen, Variablenlabels und Wertelabels

[Variablen- und Wertelabelvergabe] Für Variablen- und Wertelabel wurden Formulierungen des Fragebogens übernommen oder prägnante Kurzformen dieser Formulierungen gewählt. Dabei basieren die Variablenlabel in der Regel auf dem entsprechenden Fragetext. Grundlage für die Wertelabel sind je nach Fragetyp die Texte der Antwortoptionen bzw. eine Kombination der Texte von Frage und Antwortoption. Die Variablen- und Wertelabel liegen auf Deutsch vor.

[Variablenbenennung im Personendatensatz] Für die Variablenbenennung wurde das Variablennamenschema des Primärforschungsprojektes übernommen. Tabelle 2 stellt die verschiedenen Themengebetskürzel mit Wellennummerierung sowie die Themengebiete für Variablennamen des Absolventenpanels 2013 im Überblick dar.

Tab. 2: Themengebetskürzel mit Wellennummerierung sowie die Themengebiete für Variablennamen des DZHW-Absolventenpanels 2013

Themengebetskürzel mit Wellennummerierung	Themengebiet
b1	1. Studienverlauf und Studienerfahrungen
g1	2. Vor dem Masterstudium
a1, d1	3. Nach dem Bachelorstudium
d1	4. Promotion
h1, f1, c1	5. Auf dem Weg in den Beruf
c1	6. Erwerbstätigkeit
c1 ¹⁸	7. Werdegang nach dem Studienabschluss
c1	8. Berufs- und Lebensziele
k1	9. Angaben zur Person
k1	10. Vorbildung
k1	11. Mobilität
k1	12. Bildungsherkunft
c1	13. Zu guter Letzt

¹⁸ Diese Variablen wurden in einen Episodendatensatz umgewandelt, wodurch sich die Variablennamen verändert haben (s. Abschnitt Variablenbenennung im Episodendatensatz).

i2, h2, g2, mo ¹⁸	1. Aktuelle Situation
a2, c2	2. Berufliche Situation
d2, e2	3. Hochschulische und außerhochschulische Weiterbildung
k2, c2	4. Fragen zur Person

Bestimmte Variablen sind aus Anonymisierungsgründen nicht über alle potentiellen Zugangswege (Download-CUF, Download-SUF, Remote-Desktop-SUF, On-Site-SUF; vgl. Kapitel II) einsehbar. In diesen Fällen wird bei den Personendaten im Variablenlabel in Klammern derjenige Zugangsweg angegeben, *ab* dem die Variable nutzbar¹⁹ ist:

- **c:** Variable ist im CUF, im Download-SUF, im Remote-Desktop-SUF und im On-Site-SUF nutzbar.
- **d:** Variable ist nicht im CUF, aber im Download-SUF, im Remote-Desktop-SUF und im On-Site-SUF nutzbar.
- **r:** Variable ist nicht im CUF und im Download-SUF, aber im Remote-Desktop-SUF und im On-Site-SUF nutzbar.
- **o:** Variable ist nicht im CUF, im Download-SUF und im Remote-Desktop-SUF, aber im On-Site-SUF nutzbar.
- **a:** Variable ist über keinen Zugangsweg nutzbar. Sie wird aber dokumentiert, da es zugehörige Fragen im Fragebogen gibt.

[Variablenbenennung im Episodendatensatz] Die Variablen im Episodendatensatz sind die Identifikationsnummer der befragten Person (pid), die Identifikationsnummer der jeweiligen Episode (eid), die ausgeübte Tätigkeitsart (status) sowie Beginn und Ende des Episodenzeitraums, der über vier Variablen (Monat: begin_m und end_m; Jahr: begin_j und end_j) codiert wird.

[Besondere Hinweise zu den Variablen c1job*taet und c2job*taet „Art der Tätigkeit“] Die Variablen „Art der Tätigkeit“ (c1job*taet; c2job*taet) wurden vom Primärforschungsprojekt aus den Variablen „Art des Arbeitsverhältnisses“ (c1job*arve; c2job*arve) und „Berufliche Stellung“ (c1job*best; c2job*best) generiert. Die beiden Ursprungsvariablen wurden in Welle 1 in Frage 6.2 und in Welle 2 in Frage p304 erhoben. Die Angaben zum Arbeitsverhältnis und zur Beruflichen Stellung wurden vom Primärforschungsprojekt kombiniert und mit den Angaben zum beruflichen Werdegang aus dem Kalendarium (Welle 1: Frage 7.1; Welle 2: p103) abgeglichen. Die Variablen zur Art der Tätigkeit sind somit Kunstvariablen, die als Ergänzung zu den restlichen Angaben zur beruflichen Tätigkeit genutzt werden können.

¹⁹ „Nutzbar“ heißt: die Variable enthält nicht ausschließlich Fälle mit dem Missing-Wert „anonymisiert“.

4.2 Codierung fehlender Werte

Zur Codierung fehlender Werte wurde folgende Systematik verwendet:

Tab. 3: Systematik für fehlende Werte

Code	Wertelabel
-1	Keine Angabe
-2	Filter
-3	Nicht erwerbstätig/ Angabe keine Stelle gefunden/ kein BA in PJ20123 ²⁰ / Master erwogen aber (noch) nicht umgesetzt
-4	Kein Code vergeben/ Suche ja, k.a. finden/ kein Master(studium) erwogen
-9	Keine Teilnahme (Panel)
-966	nicht bestimmbar
-967	anonymisiert
9999	Hochschule unbekannt

²⁰ PJ20123 steht für Prüfungsjahr 2012/13.

5 Gewichtung

Es werden keine Gewichte bereitgestellt, da aufgrund der systematischen Stichprobenausfälle (s. Kapitel 3) keine adäquate Datenbasis für eine Kalibrierung herangezogen werden kann.

6 Anonymisierung

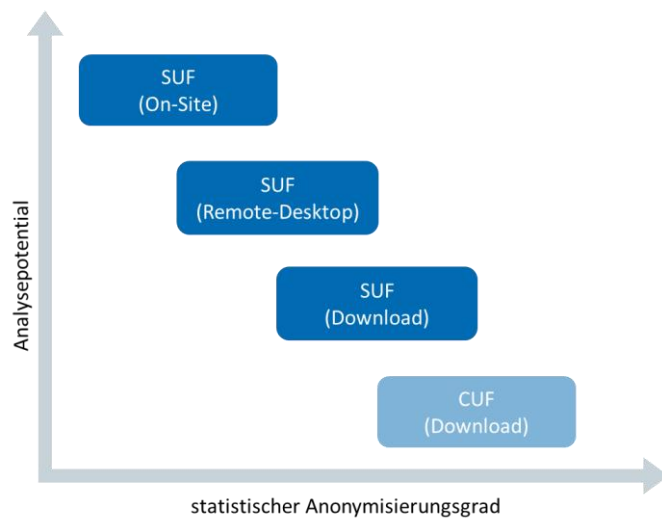
[Datenschutzrechtlicher Rahmen] Für personenbezogene Daten²¹, die in freiwilligen Befragungen durch das DZHW erhoben werden, gelten die EU-Datenschutz-Grundverordnung (EU-DSGVO) und das Bundesdatenschutzgesetz. Danach sind personenbezogene Daten für die Weitergabe zur wissenschaftlichen Sekundärnutzung (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). Das heißt, für die Weitergabe von Daten aus wissenschaftlichen Forschungsprojekten an Dritte sind die Daten derart zu anonymisieren, dass kein Bezug zur Person mehr hergestellt werden kann.

[Datenzugang, Anonymisierungsgrad und Analysepotential] Das FDZ des DZHW stellt für das Absolventenpanel 2013 ein SUF für die wissenschaftliche Sekundärnutzung und ein CUF für Lehr- und Übungszwecke zur Verfügung. Die Anonymität der Befragten wird dabei über eine Kombination aus statistischen Maßnahmen und technischen Zugriffsbeschränkungen sichergestellt. Je stärker der Datenzugang technisch kontrolliert wird, desto geringer ist das Risiko einer De-Anonymisierung der Daten, desto weniger müssen die Daten mittels statistischer Maßnahmen um Informationen reduziert werden und desto größer bleibt ihr Analysepotential.

Während das CUF nach einer Registrierung direkt durch das FDZ des DZHW übermittelt wird, wird das SUF über drei verschiedene Zugangswege angeboten: Download, Remote-Desktop und On-Site (für weiterführende Informationen vgl. Abschnitt II). Für jeden Zugangsweg wird eine andere SUF-Variante bereitgestellt, die unterschiedlich stark anonymisiert worden ist und entsprechend weniger oder mehr Informationen umfasst. Abbildung 3 gibt einen Überblick über den jeweiligen Grad der statistischen Anonymisierung und dem damit verbundenen Analysepotential. Im Folgenden werden die durchgeführten statistischen Anonymisierungsmaßnahmen in Abhängigkeit von Datenpaket (SUF/CUF) und Zugangsweg erläutert.

²¹ „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, S. 1).

Abb. 3: Datenzugangswege, statistischer Anonymisierungsgrad und Analysepotential der Daten des DZHW-Absolventenpanels 2013



[Statistische Anonymisierungsmaßnahmen] Im Rahmen der Anonymisierung sind zunächst alle Informationen, mit denen sich Personen oder Institutionen direkt identifizieren lassen, zu löschen. Diese sogenannten *direkten Identifikatoren*, wie Namen, Adressen und E-Mail Adressen, wurden im Absolventenpanel 2013 bereits während der Feldphase in einem separaten Datensatz erfasst und sind somit weder im CUF noch in den verschiedenen SUF-Varianten enthalten. Um einen Rückbezug auf diesen Datensatz zu unterbinden, wurde zudem die Original-Identifikationsnummer entfernt und durch eine neue zufällig vergebene Identifikationsnummer ersetzt.

Anschließend wurden die *Quasi-Identifikatoren* bestimmt, also Informationen, die in Kombination oder durch die Anspielung externer Informationen geeignet sind, eine Person indirekt zu identifizieren.²² Für das Absolventenpanel 2013 wurden unter anderem die folgenden Quasi-Identifikatoren identifiziert: Hochschule, Studienfach, Abschlussart, Berufsangaben, regionale Informationen (zur Hochschule, zum Ort des Erwerbs der Studienberechtigung oder Arbeitsort), Staatsangehörigkeit und Geburtsland. Um eine eindeutige Zuordnung der Absolvent*innendaten zu unterbinden, wurden diese Schlüsselmerkmale – je nach Datenpaket bzw. Zugangsweg – aggregiert oder gelöscht (vgl. Tabelle 4). Beispielsweise wird das Merkmal „Hochschule“ in dem SUF für die On-Site Nutzung zu NUTS-2-Regionen, im Remote-Desktop-SUF zu Bundesländern, im Download-SUF und im Download-CUF zu den zwei Kategorien alte vs. neue Bundesländern aggregiert.

Darüber hinaus empfehlen Ebel und Meyermann, 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, S. 5). Die offenen Angaben wurden größtenteils bereits im Rahmen der Datenaufbereitung durch das Primärforschungsprojekt vercodet und werden in dieser Form (teilweise

²² Dabei ist darauf hinzuweisen, dass die Identifikation einer Person bereits durch die Stichprobenauswahl erschwert wird, da eine Ungewissheit darüber besteht, ob eine befragte Person eine einzigartige Merkmalskombination in der Population aufweist.

aggregiert) zur Verfügung gestellt. Nicht codierte offene Angaben wurden im CUF und in allen SUF-Varianten gelöscht.

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 (vgl. Koberg, 2016, S. 694) und sind daher besonders schützenswert (Art. 9 DSGVO, Erwägungsgrund 51 DSGVO). Im Absolventenpanel 2013 wurden Gesundheitsinformationen (längerfristige Krankheitsepisoden im Kalendarium und gesundheitliche Gründe für Kinderlosigkeit) erhoben, für die bei den Befragten kein zusätzliches Einverständnis für die Sekundärnutzung eingeholt wurde. Daher wurden diese Antworten im CUF- und allen SUF-Varianten mit der Kategorie „Sonstiges“ zusammengefasst oder gelöscht/anonymisiert.

Zur Realisierung der Anonymität der Befragten im CUF wurden zum einen im Vergleich zu den SUF-Varianten restriktivere statistische Anonymisierungsmaßnahmen auf Variablenebene vorgenommen. Zum anderen wurde eine per Zufallsauswahl gewonnene Substichprobe der Daten (2/3 der befragten Absolvent*innen) gezogen.

Die nachfolgende Tabelle 4 stellt in Kurzform die durchgeführten statistischen Anonymisierungsmaßnahmen je nach Datenform bzw. Zugangsweg dar.

Tab. 4: Maßnahmen der statistischen Anonymisierung der Daten des DZHW-Absolventenpanels 2013 nach Zugangsweg²³

Merkmal	On-Site-SUF	Remote-Desktop-SUF	Download-SUF	Download-CUF (Substichprobe)
Direkte Identifikatoren	Löschung und Vergabe einer zufälligen ID	Löschung und Vergabe einer zufälligen ID	Löschung und Vergabe einer zufälligen ID	Löschung und Vergabe einer zufälligen ID
Studien-, Promotions- und Fortbildungsfach	Freigabe	Aggregation zu Studienbereichen ^a	Aggregation zu Studienbereichen ^a	Aggregation zu Fächergruppen ^a
Angestrebte Abschlussart	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45
Schulform Lehramtsausrichtung	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45	Aggregation nach cl-dzhw-45
(Promotions-/ Fortbildungs-) Hochschule	Aggregation zu Hochschulart nach cl-destatis-hochschule-2013 und Hochschulort zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b nach cl-eurostat-	Aggregation zu Hochschulart nach cl-destatis-hochschule-2013 und Hochschulort zu Bundesländern	Aggregation zu Hochschulart nach cl-destatis-hochschule-2013 und Hochschulort zu neuen bzw. alten Bundesländern	Aggregation zu Hochschulart nach cl-destatis-hochschule-2013 und Hochschulort zu neuen bzw. alten Bundesländern

²³ Detaillierte Informationen zu den anonymisierten Variablen sind dem Datensatzreport sowie dem Metadaten suchsystem (<https://metadata.fdz.dzhw.eu>) zu entnehmen. Die verwendeten Codierlisten (z.B. cl-dzhw-45 oder cl-destatis-hochschule-2013) sind einsehbar unter <https://metadata.fdz.dzhw.eu/#!/de/data-packages/stu-gra2013>, Abschnitt „Instrumente“.

Merkmal	On-Site-SUF	Remote-Desktop-SUF	Download-SUF	Download-CUF (Substichprobe)
deplznuts-2010				
Auslandsaufenthalt, Praktikum (Ort/Land)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)
Arbeitsort (Bundesland/Land)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation zu Bundesländern und Ausland (ja/nein)	Aggregation zu alte, neue Bundesländer und Ausland (ja/nein)
Arbeitsort (PLZ 5-Steller)	Freigabe und Aggregation zu NUTS-3 ^b nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b und Bundesländer (NUTS 1) nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b und Bundesländer (NUTS 1) nach cl-eurostat-deplznuts-2010	Aggregation zu alte, neue Bundesländer
Ort der Studienberechtigung (Bundesland/ Ausland)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)	Löschung	Löschung
Ort der Studienberechtigung (PLZ)	Aggregation zu NUTS-3 ^b nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b nach cl-eurostat-deplznuts-2010	Aggregation zu alte, neue Bundesländer
Wohnsitz (Ausland)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation zu Ausland (ja/nein)	Aggregation zu Ausland (ja/nein)
Wohnsitz (PLZ-3-Steller)	Freigabe	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b und Bundesländer (NUTS 1) nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b und Bundesländer (NUTS 1) nach cl-eurostat-deplznuts-2010	Aggregation zu alte, neue Bundesländer
Beruf	Vercodung zu Berufsgattungen (KldB 5-Steller) ^c	Aggregation zu Berufsordnungen (KldB 3-Steller) ^c nach cl-destatis-kldb-2010	Aggregation zu Berufsordnungen (KldB 3-Steller) ^c nach cl-destatis-kldb-2010	Aggregation zu Berufshauptgruppen (KldB 2-Steller) ^c nach cl-destatis-kldb-2010
Ausbildungsberuf	Löschung (da offene Angabe)	Löschung (da offene Angabe)	Löschung (da offene Angabe)	Löschung (da offene Angabe)
Dauer Arbeitslosigkeit	Freigabe	Top-Codierung ab 25 Monate	Top-Codierung ab 25 Monate	Top-Codierung ab 25 Monate

Merkmal	On-Site-SUF	Remote-Desktop-SUF	Download-SUF	Download-CUF (Substichprobe)
Staatsangehörigkeit (Ausland)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach NEPS Klassifikation ^d nach cl-dzhw-44 (Code_NEPS)	Aggregation zu Weltregionen nach cl-dzhw-44 (Code_Weltregionen)	Löschung
Deutsche Staatsangehörigkeit seit (seit Geburt bzw. Jahreszahl)	Freigabe	Freigabe	Freigabe	Löschung
Geburtsort (PLZ 5-Steller)	Freigabe und Aggregation zu NUTS-3 ^b nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b nach cl-eurostat-deplznuts-2010	Aggregation zu NUTS 2: Basisregionen für regionalpolitische Maßnahmen ^b nach cl-eurostat-deplznuts-2010	Löschung
Geburtsland (Ausland)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach NEPS Klassifikation ^d nach cl-dzhw-44 (Code_NEPS)	Aggregation zu Weltregionen nach cl-dzhw-44 (Code_Weltregionen)	Löschung
Geburtsjahr und -monat	Freigabe	Freigabe	Freigabe	Aggregation: 1992 und jünger; 1963 und älter
Geburtsort Eltern (Bundesland)	Freigabe	Freigabe	Freigabe	Aggregation zu alte, neue Bundesländer
Geburtsland Eltern (Ausland)	Aggregation nach cl-dzhw-44 (Code_Länder)	Aggregation nach NEPS Klassifikation ^d nach cl-dzhw-44 (Code_NEPS)	Aggregation zu Weltregionen nach cl-dzhw-44 (Code_Weltregionen)	Aggregation zu Weltregionen nach cl-dzhw-44 (Code_Weltregionen)
Beruf Eltern	Löschung (da offene Angabe)	Löschung (da offene Angabe)	Löschung (da offene Angabe)	Löschung (da offene Angabe)
Alter und Anzahl Kinder	Freigabe	Geburtsjahr: Aggregation: bis 1997; 1998-2003; 2004-2009; 2010-2012; ab 2013 Geburtsmonat: Löschung	Geburtsjahr (nur für die vier jüngsten Kinder): Aggregation: bis 1997; 1998-2009; ab 2010 Geburtsmonat: Löschung	Geburtsjahr (nur für die vier jüngsten Kinder): Aggregation: bis 1997; 1998-2009; ab 2010 Geburtsmonat: Löschung
Gründe für Kinderlosigkeit: - Kann keine Kinder bekommen - Gesundheitliche Gründe	Löschung	Löschung	Löschung	Löschung
Krankheitsepisoden (Kalendarium)	Zusammenfassung mit der Kategorie „Sonstiges“	Zusammenfassung mit der Kategorie „Sonstiges“	Zusammenfassung mit der Kategorie „Sonstiges“	Zusammenfassung mit der Kategorie „Sonstiges“

Merkmal	On-Site-SUF	Remote-Desktop-SUF	Download-SUF	Download-CUF (Substichprobe)
Art der berufsqualifizierenden Weiterbildung	Löschung	Löschung	Löschung	Löschung
Sonstige offene Angaben	Löschung oder Aggregation zu genannt/ nicht genannt	Löschung oder Aggregation zu genannt/ nicht genannt	Löschung oder Aggregation zu genannt/ nicht genannt	Löschung oder Aggregation zu genannt/ nicht genannt

^a Nach dem Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2012/2013 und SoSe 2013 von Destatis.

^b Statistisches Amt der Europäischen Union (Eurostat): Nomenclature of territorial units for statistics (NUTS)
<http://ec.europa.eu/eurostat/web/nuts/overview>.

^c Nach Klassifikation der Berufe (KldB) von 2010.

^d Die Aggregation der Staaten zu Weltregionen ist angelehnt an die Klassifikation des NEPS mit Anpassungen bei europäischen Ländern https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5_6-0-0_Anonymisation.pdf.

7 Literaturverzeichnis

- Abele, A. E. & Schradi, M. (2000). *Methodisches Vorgehen und Fragebogen der ersten Erhebungswelle* (Bericht Nr. 1 des Projekts "Frauen in der Mathematik"). Erlangen: FAU Erlangen-Nürnberg.
- Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). *Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen* (GESIS - Leibniz-Institut für Sozialwissenschaften, Hrsg.). doi:10.6102/zis35
- Ebel, T. & Meyermann, A. (2015). *Hinweise zur Anonymisierung von quantitativen Daten* (forschungsdaten bildung informiert Nr. 3). Verbund Forschungsdaten Bildung. Verfügbar unter https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert_nr-7.pdf
- Fabian, G., Hillmann, J., Trennt, F. & Briedis, K. (2016). *Hochschulabschlüsse nach Bologna. Werdegänge der Bachelor- und Masterabsolvent(inn)en des Prüfungsjahrgangs 2013* (Forum Hochschule 1|2016). Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW).
- Hochfellner, D., Müller, D., Schmucker, A. & Roß, E. (2012). *Datenschutz am Forschungsdatenzentrum* (FDZ-Methodenreport Nr. 6). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).
- Koberg, T. (2016). Disclosing the National Educational Panel Study. 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. 691–708). Wiesbaden: Springer VS. doi:10.1007/978-3-658-11994-2
- Lane, J., Heus, P. & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1(1), pp. 2–16.
- Trennt, F. (2019). Zahlt sich ein Master aus? Einkommensunterschiede zwischen den neuen Bachelor- und Masterabschlüssen. In M. Lörz & H. Quast (Hrsg.), *Bildungs- und Berufsverläufe mit Bachelor und Master. Determinanten, Herausforderungen und Konsequenzen* (1. Auflage 2019, S. 371–397). Wiesbaden: Springer Fachmedien Wiesbaden GmbH. doi:10.1007/978-3-658-22394-6_12