



FORSCHUNGSDATENZENTRUM  
der Bundesagentur für Arbeit im Institut für  
Arbeitsmarkt- und Berufsforschung

# FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

---

## **14|2025 DE** Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023

Sandra Dummert, Katrin Hohmeyer, Torsten Lietzmann, Irakli Sauer, Martina Oertel



**Bundesagentur für Arbeit**

# Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023

Sandra Dummert (IAB), Katrin Hohmeyer (IAB), Torsten Lietzmann (IAB), Irakli Sauer (IAB),  
Martina Oertel (IAB)

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other, the data can be used to prepare evaluations.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung und Kurzbeschreibung</b>	<b>8</b>
1.1	Einleitung	8
1.2	Grundgesamtheit der SIG	10
1.3	Abgrenzung zu anderen administrativen Datenprodukten des IAB	11
1.4	Datennutzung	11
1.4.1	Datenzugang und Datenhaltung	11
1.4.2	Dateinamen in der SIG 0723	13
1.5	Veränderung zur SIG 0720	13
1.5.1	Merkmalskanon	13
1.5.2	Datendokumentation	14
1.6	Steckbrief	14
1.7	Mengengerüst	16
<b>2</b>	<b>Datenquellen</b>	<b>17</b>
2.1	Leistungshistorik Grundsicherung (LHG)	18
2.2	Leistungsempfängerhistorik (LeH)	19
2.3	Arbeitssuchendenhistoriken (ASU/XASU)	19
2.4	Maßnahmeteilnahmehistoriken (MTH/XMTH)	20
2.5	Beschäftigtenhistorik (BeH)	21
2.6	Leistungsstatistik nach dem SGB II (LST-S)	22
<b>3</b>	<b>Datenaufbereitung und Stichprobenziehung</b>	<b>23</b>
3.1	Stichprobenziehung	23
3.2	Struktur und Aufbereitung der SIG	23
3.2.1	SGB-II-Anonymisierung	24
3.2.2	Leistungshistorik Grundsicherung (LHG)	25
3.2.3	Leistungsempfängerhistorik (LeH)	25
3.2.4	Arbeitssuchendenhistoriken (ASU und XASU)	26
3.2.5	Maßnahmeteilnahmehistorik (MTH) und Maßnahmeteilnahmehistorik aus XSozial (XMTH)	26
3.2.6	Beschäftigtenhistorik (BeH)	27
3.2.7	Leistungsstatistik nach dem SGB II (LST-S)	27
3.3	Episodensplitting	28
3.4	Bereinigung von Inkonsistenzen	29
3.5	Fehlende Werte	29
<b>4</b>	<b>Datenqualität und Probleme</b>	<b>30</b>

4.1	Einführung des SGB II und spätere Trägerwechsel.....	30
4.2	Leistungshistorik Grundsicherung (LHG) .....	30
4.3	IEB.....	32
4.3.1	Lücken in den Erwerbsbiografien .....	32
4.3.2	Arbeitsuchendenhistoriken (ASU/XASU) .....	32
4.3.3	Maßnahmeteilnahmehistoriken (MTH/XMTH) .....	33
4.3.4	Beschäftigtenhistorik (Beh).....	34
4.4	Leistungsstatistik nach dem SGB II (LST-S) .....	36
<b>5</b>	<b>Variablenbeschreibung.....</b>	<b>38</b>
5.1	Identifikatoren .....	38
5.1.1	Systemfreie Personennummer (persnr).....	38
5.1.2	Systemfreie Betriebsnummer (betnr_sig) .....	38
5.1.3	Systemfreie Bedarfsgemeinschaftsnummer (bg_id) .....	39
5.2	Generierte technische Merkmale .....	39
5.2.1	Quelle (quelle_lhg, quelle_asu, quelle_xasu, quelle_leh, quelle_mth, quelle_xmth, quelle_beh).....	39
5.2.2	Vorhergehende Meldungslücke (lck_links) .....	39
5.2.3	Anschließende Meldungslücke (lck_rechts).....	40
5.2.4	Spell linksenziert 1.1.2007 (zensierung_links_lhg, zensierung_links_x_asu, zensierung_links_leh, zensierung_links_x_mth, zensierung_links_beh).....	40
5.2.5	Spelllänge (spell_length) .....	40
5.3	Gültigkeitszeitraum .....	40
5.3.1	Beginndatum Episode (begepi) .....	40
5.3.2	Enddatum Episode (endepe) .....	41
5.3.3	Erster Monat Episode (beg_mon) .....	41
5.3.4	Letzter Monat Episode (end_mon) .....	41
5.4	Informationen zur Person.....	42
5.4.1	Geschlecht (frau) .....	42
5.4.2	Geburtsjahr (gebjahr).....	42
5.4.3	Geburtsmonat (gebmon) .....	42
5.4.4	Staatsangehörigkeit (nation_lhg, nation_x_asu, nation_leh, nation_x_mth, nation_beh) .....	43
5.4.5	Staatsangehörigkeit vergrößert (nation_gr_lhg, nation_gr_x_asu, nation_ gr_leh, nation_gr_x_mth, nation_gr_beh) .....	43
5.4.6	Familienstand (famst_lhg, famst_x_asu, famst_leh).....	43
5.4.7	Alter des jüngsten Kindes - kategorisiert (alter_juk_kat) .....	44
5.4.8	Ausbildung (ausbildung_x_asu, ausbildung_beh).....	44
5.4.9	Ausbildung (imputiert) (ausbildung_imp_beh) .....	46
5.4.10	Schulausbildung (schule_x_asu, schule_beh) .....	46
5.4.11	Max. Ausbildung bis 2006 (ausbildung_bis2006) .....	47
5.4.12	Max. Schulausbildung bis 2006 (schule_bis2006) .....	47

5.5	Informationen zu Leistungsbezug.....	47
5.5.1	Täglicher Leistungssatz Arbeitslosengeld I (leistung_tag) .....	47
5.5.2	Rolle in der Bedarfsgemeinschaft (bg_rolle).....	48
5.5.3	Erwerbsfähigkeit (erwf).....	48
5.5.4	Bedarfsgemeinschaftstyp (bg_typ) .....	48
5.5.5	Beendigungsgrund (grund_end) .....	49
5.5.6	SGB II Ausschlussgrund (sgbii_ausschl) .....	49
5.5.7	Zu berücksichtigendes Einkommen aus Beschäftigung (zbeink_besch).....	49
5.5.8	Zu berücksichtigendes Einkommen aus Selbständigkeit (zbeink_selbst).....	50
5.5.9	Verfügbares Einkommen gesamt (veink_sum) .....	50
5.5.10	Verfügbares Einkommen aus Beschäftigung (veink_besch) .....	51
5.5.11	Verfügbares Einkommen aus Selbständigkeit (veink_selbst) .....	51
5.5.12	Verfügbares Einkommen aus Kindergeld (veink_kinderg) .....	51
5.5.13	Verfügbares Einkommen aus Unterhalt (veink_unterh).....	51
5.5.14	Verfügbares Einkommen aus Krankengeld (veink_krankeng) .....	52
5.5.15	Verfügbares Einkommen aus Arbeitslosengeld I (veink_algl) .....	52
5.5.16	Verfügbares Einkommen aus Rente (veink_rente) .....	52
5.5.17	Verfügbares Einkommen aus Kapital und Besitz (veink_kub).....	52
5.5.18	Bedarf gesamt ohne BuT (bedarf_sum) .....	53
5.5.19	Bedarf laufende Kosten der Unterkunft (bedarf_lkdu).....	53
5.5.20	Bedarf einmalige Kosten der Unterkunft (bedarf_ekdu) .....	53
5.5.21	Mehrbedarf Schwangerschaft (d_schwanger) .....	54
5.5.22	Mehrbedarf Behinderung (d_behind).....	54
5.5.23	Mehrbedarf Ernährung (d_ernaehr) .....	54
5.5.24	Bedarf Darlehen (d_darleh) .....	55
5.5.25	Zahlungsanspruch gesamt ohne BuT (zanspr_sum) .....	55
5.5.26	Zahlungsanspruch ALG II (zanspr_algII) .....	55
5.5.27	Zahlungsanspruch Sozialgeld (zanspr_sozg).....	56
5.5.28	Zahlungsanspruch Regelbedarf (zanspr_regel) .....	56
5.5.29	Zahlungsanspruch ALG-II-Zuschlag (zanspr_algIIzu) .....	56
5.5.30	Zahlungsanspruch Mehrbedarfe (zanspr_mbedarf) .....	56
5.5.31	Zahlungsanspruch laufende KdU (zanspr_lkdu).....	57
5.5.32	Zahlungsanspruch einmalige KdU (zanspr_ekdu).....	57
5.5.33	Zahlungsanspruch Sozialversicherung (zanspr_sozvers).....	57
5.5.34	Sanktionsbetrag gesamt (sanktionen_sum).....	58
5.6	Informationen zu Arbeitsuche und Maßnahmenteilnahme .....	58
5.6.1	Berufsgruppe – letzte abgeschlossene Ausbildung (KldB 2010) (beruf2010_3_ausb) .....	58
5.6.2	Berufsuntergruppe – letzte abgeschlossene Ausbildung (KldB 2010) (beruf2010_4_ausb) .....	58
5.6.3	Berufsgruppe – gesuchte Tätigkeit (KldB 2010) (beruf2010_3_gesucht) .....	58
5.6.4	Berufsuntergruppe – gesuchte Tätigkeit (KldB 2010) (beruf2010_4_gesucht) .....	59
5.6.5	Arbeitsuchestatus (asustat) .....	59
5.6.6	Erwerbsstatus vor Arbeitsuche (estatvor) .....	59

5.6.7	Status nach Arbeitsuche (estatnach).....	60
5.6.8	Integrationsprognose (ipo) .....	60
5.6.9	Art der Kündigung der letzten Tätigkeit (art_kuend).....	61
5.6.10	Arbeitszeit des Stellengesuchs (arbzeit) .....	61
5.6.11	Abmeldegrund (grund_abm) .....	61
5.6.12	Beginndatum linkszensierter Arbeitslosigkeits-Spell (begin_alo) .....	62
5.6.13	Beginndatum linkszensierter Arbeitsuche-Spell (begin_asu) .....	62
5.6.14	Maßnahmeart - Kategorie (mass_gr).....	63
5.6.15	Maßnahmeart – Gruppe (mass) .....	63
5.7	Informationen zu Beschäftigung.....	63
5.7.1	Tagesentgelt (entgelt_tag) .....	63
5.7.2	Tagesentgelt (inkl. Einmalzahlungen) (entgelt_tag_bonus).....	64
5.7.3	Tagesentgelt (imputiert) (entgelt_tag_imp) .....	65
5.7.4	Beruf – ausgeübte Tätigkeit (KldB 1988) (beruf_aktT).....	65
5.7.5	Berufsgruppe – ausgeübte Tätigkeit (KldB 2010) (beruf2010_3_aktT) .....	65
5.7.6	Berufsuntergruppe – ausgeübte Tätigkeit (KldB 2010) (beruf2010_4_aktT) .....	66
5.7.7	Anforderungsniveau – ausgeübte Tätigkeit (KldB 2010) (niveau_aktT).....	66
5.7.8	Teilzeit (teilzeit).....	67
5.7.9	Beschäftigungsstatus (beschstat) .....	67
5.7.10	Gleitzone (gleitz) .....	68
5.7.11	Leiharbeit (leih) .....	68
5.7.12	Befristung (befrist) .....	69
5.7.13	Abgabegrund (grund_abg).....	69
5.7.14	Mehrfachbeschäftigung (mehrfbesch) .....	69
5.7.15	Tagesentgelt – zweite Beschäftigung (mehrfbesch_tentg) .....	70
5.7.16	Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung (mehrfbesch_bonus_tentg) .....	70
5.7.17	Beschäftigungsstatus – zweite Beschäftigung (mehrfbesch_beschstat).....	70
5.7.18	Teilzeit – zweite Beschäftigung (mehrfbesch_teilzeit) .....	70
5.7.19	Erwerbserfahrung vor 2007 (exp_vor2007) .....	71
5.7.20	Zeit seit letzter Erwerbstätigkeit (time_vor2007) .....	71
5.7.21	Tagesentgelt vor 2007 (tentgelt_vor2007) .....	71
5.7.22	Beruf – letzte ausgeübte Tätigkeit (KldB 1988) vor 2007 (beruf1988_vor2007) ....	71
5.7.23	Berufsgruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 3-Steller (beruf2010_3_vor2007) .....	71
5.7.24	Berufsuntergruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 4-Steller (beruf2010_4_vor2007) .....	72
5.7.25	Teilzeit vor 2007 (teilzeit_vor2007).....	72
5.7.26	Beschäftigungsstatus vor 2007 (beschstat_vor2007) .....	72
5.8	Ortsangaben.....	73
5.8.1	Wohnort - Kreis (wo_kreis_lhg, wo_kreis_x_asu, wo_kreis_leh, wo_kreis_beh) .	73
5.8.2	Wohnort - Bundesland (wo_bula_lhg, wo_bula_x_asu, wo_bula_leh, wo_bula_beh).....	73

5.8.3	Arbeitsort Kreis (ao_kreis)	74
5.8.4	Arbeitsort Bundesland (ao_bula)	74
5.9	Betriebsmerkmale	74
5.9.1	WZ08 5-Steller (w08_5_kons)	74
5.9.2	WZ08 3-Steller (w08_3_kons)	75

## Abbildungsverzeichnis

Abbildung 1:	Datenhaltung der SIG -Daten	12
Abbildung 2:	Quellen der SIG	17
Abbildung 3:	Episodensplitting	29

## Tabellenverzeichnis

Tabelle 1:	Neue, entfallene, umbenannte und veränderte Merkmale in der SIG 0723	14
Tabelle 2:	Steckbrief	14
Tabelle 3:	Mengengerüst	16
Tabelle 4:	Fehlende Werte	29

# Zusammenfassung

Dieser Datenreport beschreibt die Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023.

## Abstract

This data report describes the Sample of Integrated Welfare Benefit Biographies (SIG) 2007 - 2023.

## Keywords

Data manual, German administrative micro data, welfare benefit data.

## Danksagung

Für ihre Mitarbeit und Unterstützung danken wir den Kolleginnen und Kollegen des Forschungsdatenzentrums (FDZ) sowie des Geschäftsbereiches Daten- und IT-Management (DIM) des Instituts für Arbeitsmarkt- und Berufsforschung (IAB), hier insbesondere Alexandra Schmucker (FDZ), Andreas Ganzer (FDZ) und Tobias Graf (DIM). Zur Beschreibung der Daten wurden Textteile aus Datenbeschreibungen von DIM und aus früheren FDZ-Datenreporten übernommen. Zudem danken wir den (ehemaligen) Kolleginnen und Kollegen in der Forschungsgruppe Grundsicherungsbezug und Arbeitsmarkt (FG GAMA), die an der Konzeption früherer Versionen der SIG mitgewirkt haben: Lena Bösel, Kerstin Bruckmeier, Helmut Rudolph und Anna-Theresa Saile.

# 1 Einleitung und Kurzbeschreibung

## 1.1 Einleitung

Mit der Stichprobe der Integrierten Grundsicherungsbiografien (SIG) des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) werden über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit (BA) Daten für wissenschaftliche Auswertungen zu Struktur und Verlauf des Leistungsbezugs in der Grundsicherung nach Sozialgesetzbuch (SGB) II zur Verfügung gestellt. Die SIG besteht aus einer 4 %-Zufallsstichprobe aller Personen, die zwischen 2007 und 2023 in einer nach § 7 SGB II definierten Bedarfsgemeinschaft gelebt haben.

Die Ziehung basiert auf den Daten der Leistungshistorik Grundsicherung (LHG). Die LHG enthält konsolidierte, tagesgenaue, personenbezogene Informationen zu erwerbsfähigen Leistungsberechtigten und ggf. ihren Bedarfsgemeinschaftsmitgliedern aus den Verwaltungsverfahren der BA zur Grundsicherung sowie den von kommunalen Grundsicherungsträgern übermittelten Daten.

Die SIG liefert damit prozessproduzierte Längsschnittdaten, mit denen Forschungsfragen zur Dynamik von Bedürftigkeit im Grundsicherungsbezug analysiert werden können. Um die Prozesse von Zu- und Abgängen bzw. Verringerung von Hilfebedürftigkeit im Zeitverlauf untersuchen zu können, wurden die Daten mit Informationen zu Phasen der Aktivierung, Förderung, Arbeitsuche und Erwerbstätigkeit angereichert. Die Stichprobenpersonen und ihre Erwerbsbiografien bilden den sogenannten SIG-Kern.

Bedürftigkeit im Grundsicherungsbezug definiert sich im Haushalts- bzw. Bedarfsgemeinschaftskontext. Zentrales Anliegen der SIG ist es, die Erforschung von Hilfebedürftigkeit unter Berücksichtigung von besonderen Problemlagen im Haushaltskontext, z. B. der Zugehörigkeit von Kindern zur Bedarfsgemeinschaft, zu ermöglichen. Daher wurden zu den Stichprobenpersonen Informationen über deren Bedarfsgemeinschaftsmitglieder sowie deren Erwerbsverläufe gezogen und dem SIG-Kern hinzugefügt. Dabei können in einer Bedarfsgemeinschaft auch mehrere Stichprobenpersonen leben.

Die Erwerbsbiografien der Stichprobenpersonen wie auch der Bedarfsgemeinschaftsmitglieder stammen zum Großteil aus den Integrierten Erwerbsbiografien (IEB) des IAB. Die IEB enthalten Daten zu allen Personen, die im Beobachtungszeitraum mindestens einmal einen der folgenden Zustände aufweisen:

- sozialversicherungspflichtig beschäftigt in Deutschland (erfasst ab 1975 in West-, ab 1991 in Ostdeutschland),
- geringfügig beschäftigt in Deutschland (erfasst ab 1999),
- Bezug von Leistungen nach dem Rechtskreis SGB III (erfasst ab 1975) oder SGB II bei Erwerbsfähigkeit (erfasst ab 2005),
- bei der Bundesagentur für Arbeit (BA) bzw. den Grundsicherungsträgern als arbeitssuchend gemeldet (erfasst ab 1997),
- Teilnahme an arbeitsmarktpolitischer Maßnahme der BA bzw. eines Grundsicherungsträgers (erfasst ab 2000).

Diese aus unterschiedlichen Datenquellen stammenden Informationen werden in den IEB zusammengeführt und tagesgenau abgebildet aber nicht auf inhaltliche Plausibilität geprüft.

Während die Beschäftigtenhistorik (BeH) Ursprung der Informationen zu sozialversicherungspflichtiger und geringfügiger Beschäftigung ist, wird der Bezug von Leistungen nach den Rechtskreisen des SGB III und des SGB II in der Leistungsempfängerhistorik (LeH) bzw. in der Leistungshistorik Grundsicherung (LHG) erfasst. Die Arbeitssuchendenhistoriken (ASU und XASU) stellen die Datenquelle für die durch die BA bzw. die durch die kommunalen Grundsicherungsträger erfassten Zeiten der Arbeitsuche bzw. der Nichtverfügbarkeit für eine im SGB II verpflichtenden Arbeitsuche dar, während die Teilnahme an arbeitsmarktpolitischen Maßnahmen in den Maßnahmeteilnehmergehistoriken (MTH und XMTH) erfasst wird.

Anders als bei anderen administrativen Datenprodukten des FDZ der BA im IAB, wie beispielsweise bei der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB), werden in der SIG alle Informationen zum Bezug von Leistungen nach dem SGB II nicht aus der IEB, sondern direkt aus der LHG gezogen. Dies ermöglicht ein breiteres Merkmalspektrum insbesondere zur Struktur der Bedarfsgemeinschaft. Diese sind einerseits als Merkmale verfügbar, andererseits sind auch alle Bedarfsgemeinschaftsmitglieder in den Daten enthalten. Zudem wurden einzelne, nicht in der IEB verfügbare Merkmale, direkt aus den weiteren Datenquellen gewonnen.

Die Daten zu den Stichprobenpersonen aus der LHG sowie zu deren Bedarfsgemeinschaftsmitgliedern können zusätzlich – erstmals in einem Datenprodukt des FDZ – um kalendermonatsbezogene, umfassende Finanzdaten aus der Leistungsstatistik SGB II (LST-S) angereichert werden, welche Informationen zu Bedarfen, Einkommen, Zahlungsansprüchen und Sanktionen auf Ebene der Personen beinhalten.

Die SIG-Daten wurden durch das FDZ und die Forschungsgruppe Grundsicherungsbezug und Arbeitsmarkt (FG GAMA) des IAB erstellt. Der SIG-Kern umfasst insgesamt 905.135 Stichprobenpersonen, deren Erwerbsbiografien in insgesamt 24.966.840 Datenzeilen dokumentiert sind. Der Datensatz der Bedarfsgemeinschaftsmitglieder enthält insgesamt 1.784.760 Personen, deren Erwerbsverläufe in insgesamt 50.179.541 Datenzeilen erfasst sind. Der vorliegende Datenreport beschreibt die Merkmale der schwach anonymisierten Version der SIG, welche weitgehend die Originaldaten, d.h. nicht vergrößerte Daten, beinhaltet. Einige Merkmale sind allerdings als besonders sensibel eingestuft und werden nur auf gesonderten Antrag (siehe Kapitel Datenzugang 1.4.1) weitergegeben.

Dieser Datenreport gliedert sich wie folgt. Neben der Einleitung enthält Kapitel 1 unter anderem Informationen zur Abgrenzung zu anderen administrativen Datenprodukten des IAB, zu Datenzugang und -haltung, sowie eine Kurzbeschreibung der Daten, das Mengengerüst und eine Merkmalsliste. Eine Beschreibung der einzelnen Datenquellen findet sich in Kapitel 2. Die Stichprobenziehung, Datenstruktur und -aufbereitung sowie die Datenqualität werden in den Kapiteln 3 und 4 diskutiert, während die einzelnen Merkmale in Kapitel 5 dargestellt werden.

## 1.2 Grundgesamtheit der SIG

Grundgesamtheit der SIG sind alle Personen, die in Bedarfsgemeinschaften mit Leistungsbezug der Grundsicherung für Arbeitsuchende nach SGB II leben.

Leistungsberechtigt sind sog. erwerbsfähige Leistungsberechtigte und die mit ihnen zusammenlebenden Familien- und Haushaltsmitglieder, die eine gemeinsame Bedarfsgemeinschaft bilden.

Leistungsberechtigt sind nach § 7 Abs. 1 SGB II „Personen, die

1. das 15. Lebensjahr vollendet und die Altersgrenze nach § 7a noch nicht erreicht haben,
2. erwerbsfähig sind,
3. hilfebedürftig sind und
4. ihren gewöhnlichen Aufenthalt in der Bundesrepublik Deutschland haben (erwerbsfähige Leistungsberechtigte).“

Nach § 7 Abs. 3 SGB II gehören zur Bedarfsgemeinschaft

- „1. die erwerbsfähigen Leistungsberechtigten,
2. die im Haushalt lebenden Eltern oder der im Haushalt lebende Elternteil eines unverheirateten erwerbsfähigen Kindes, welches das 25. Lebensjahr noch nicht vollendet hat, und die im Haushalt lebende Partnerin oder der im Haushalt lebende Partner dieses Elternteils,
3. als Partnerin oder Partner der erwerbsfähigen Leistungsberechtigten
  - a) die nicht dauernd getrennt lebende Ehegattin oder der nicht dauernd getrennt lebende Ehegatte,
  - b) die nicht dauernd getrennt lebende Lebenspartnerin oder der nicht dauernd getrennt lebende Lebenspartner,
  - c) eine Person, die mit der erwerbsfähigen leistungsberechtigten Person in einem gemeinsamen Haushalt so zusammenlebt, dass nach verständiger Würdigung der wechselseitige Wille anzunehmen ist, Verantwortung füreinander zu tragen und füreinander einzustehen,
4. die dem Haushalt angehörenden unverheirateten Kinder der in den Nummern 1 bis 3 genannten Personen, wenn sie das 25. Lebensjahr noch nicht vollendet haben, soweit sie die Leistungen zur Sicherung ihres Lebensunterhalts nicht aus eigenem Einkommen oder Vermögen beschaffen können.“

Dabei sind unverheiratete Kinder unter 25 Jahren mit eigenem bedarfsdeckendem Einkommen rein rechtlich gesehen keine Mitglieder der Bedarfsgemeinschaft ihrer Eltern (nach § 7 Absatz 3 SGB II). Die Statistik der BA zählt minderjährige Kinder "im Sinne einer umfassenden Sozialberichterstattung" dennoch als nicht leistungsberechtigte Personen in einer Bedarfsgemeinschaft. Auch in diesem Datenreport wird diese Personengruppe unter die "Bedarfsgemeinschaftsmitglieder" subsummiert<sup>1</sup>.

---

<sup>1</sup> Einen Sonderstatus haben außerdem Kinder, deren Bedarfsgemeinschaft ausschließlich Leistungen für Bildung und Teilhabe (BuT) erhält. Deren Eltern(teile) sind in diesem Fall keine Mitglieder der Bedarfsgemeinschaft.

## 1.3 Abgrenzung zu anderen administrativen Datenprodukten des IAB

Die SIG ist der erste rein administrative Längsschnittdatensatz zum SGB-II-Leistungsbezug, der externen Forschenden vom FDZ zur Verfügung gestellt wird. Zweck der SIG ist die Analyse von Leistungsbezugsverläufen auf der Personenebene sowie die Analyse von Struktur und Dynamik im Leistungsbezug auf Aggregatebene. Die SIG bietet dafür drei Alleinstellungsmerkmale, die sie von anderen Datenprodukten des IAB abgrenzen und dem Forschungszweck dienen.

Erstens umfasst die SIG gegenüber der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) und dem PASS-ADIAB (PASS-Befragungsdaten verknüpft mit den administrativen Daten des IAB) eine größere Fallzahl von Personen im SGB-II-Bezug. Damit können Analysen für einzelne Gruppen innerhalb der sehr heterogenen Gruppe von SGB-II-Leistungsbeziehenden, wie z. B. Alleinerziehende oder erwerbstätige Leistungsbeziehende, sowie regionale Analysen durchgeführt werden. Ausdrücklich nicht geeignet ist die SIG aufgrund des dafür nicht vorhandenen notwendigen hohen Detailgrades bei der Erfassung der Maßnahmenteilnahme jedoch für die Maßnahmenevaluation.

Zweitens beinhaltet die SIG als einziger administrativer Datensatz des FDZ einen Bedarfsgemeinschaftskontext. In den SIG-Daten sind für externe Forschende erstmals Informationen zu allen Mitgliedern der Bedarfsgemeinschaft und deren Erwerbsverläufe verfügbar. Dies erlaubt eine für die Analysen von Bedürftigkeit notwendige Differenzierung von Problemlagen im Haushaltskontext wie z. B. die Haushaltsgröße, die Erwerbsbeteiligung der Bedarfsgemeinschaftsmitglieder oder die Betreuung von Kindern. Darüber hinaus kann durch die Abbildung umfassender Erwerbsbiografien aller Bedarfsgemeinschaftsmitglieder der Leistungsbezug im Kontext arbeitsmarktbezogener Lebenslagen, wie z. B. der Arbeitsmarktübertritt von Kindern im Leistungsbezug, untersucht werden.

Drittens können die Daten der Leistungsstatistik nach dem SGB II (LST-S), die bisher noch nicht als Mikrodaten für die Forschung außerhalb des IAB verfügbar waren, als Erweiterungsmodul zur SIG beantragt werden. Erstmals wird es damit möglich sein, auch Forschung zu Bedarfen, Ansprüchen und Einkommen auf Bedarfsgemeinschaftsebene durchzuführen. Dies erweitert die Analysemöglichkeiten um die Untersuchung von Bedarfslagen und Analysen zur Einkommensstruktur.

## 1.4 Datennutzung

### 1.4.1 Datenzugang und Datenhaltung

Die SIG-Daten in der schwach anonymisierten Version können im Rahmen von Gastaufenthalten am FDZ, über den Remote Desktop Zugang und anschließender Datenfernverarbeitung ausgewertet werden.

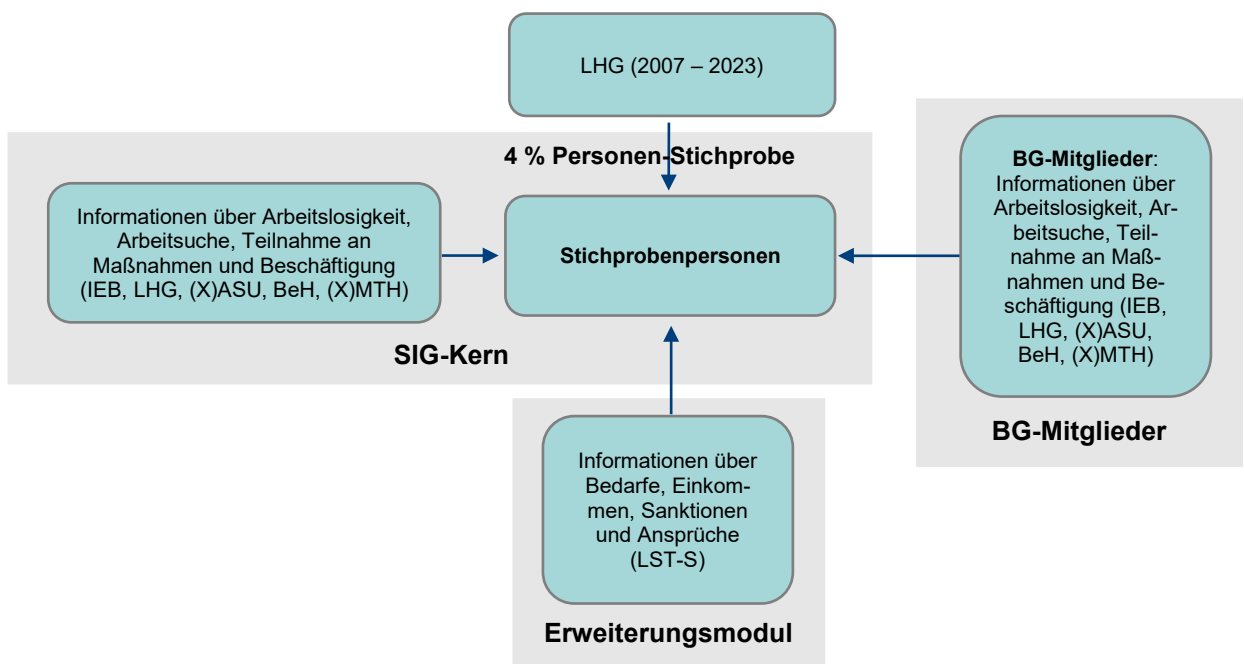
Für die Datennutzung ist zunächst ein Antrag beim FDZ einzureichen. Über die Genehmigung des Forschungsprojekts entscheidet das FDZ im Auftrag und ggf. unter Abstimmung mit dem Bundesministerium für Arbeit und Soziales (BMAS). Nach erfolgter Genehmigung wird ein Nutzungsvertrag mit der wissenschaftlichen Einrichtung des oder der Forschenden geschlossen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können der FDZ-Website entnommen werden.

Die SIG-Daten, die sowohl deutsche als auch englische Labels enthalten,<sup>2</sup> sind modular gegliedert und in mehreren Dateien abgelegt. Das Modul „SIG-Kern“ enthält Identifikatoren (systemfreie Personen-, Bedarfsgemeinschafts- und Betriebsnummer), persönliche Merkmale, Informationen zu Leistungsbezug, Arbeitsuche, Teilnahme an arbeitsmarktpolitischen Maßnahmen und Beschäftigung, Merkmale zum Wohnort der Stichprobenpersonen sowie technische Merkmale. Die persönlichen Merkmale sowie die Informationen zum Erwerbsverlauf der Bedarfsgemeinschaftsmitglieder sind in einer getrennten Datei (BG-Mitglieder) abgelegt.

Auf begründeten Antrag hin kann zudem die Leistungsstatistik nach dem SGB II (Grundsicherungsstatistik) (LST-S) als Erweiterungsmodul bereitgestellt werden. Die LST-S enthält Informationen über Bedarfe, Einkommen, Sanktionen und Ansprüche von Leistungsbeziehenden.

Die Module sind organisatorisch voneinander getrennt, was die Struktur der Daten deutlich macht und Speicherplatz spart (vgl. Abbildung 1). Die Verknüpfung der LST-S mit dem SIG-Kern ist auf Monatsebene mittels des Monats und der Bedarfsgemeinschaftsnummer möglich. Eine Beschreibung und eine Anleitung sind in Kapitel 2.6 dargestellt.

Abbildung 1: Datenhaltung der SIG -Daten



Quelle: Eigene Darstellung

Bestimmte Merkmale, die eine Deanonymisierung von Personen oder Betrieben erleichtern, werden nur dann im Original weitergegeben, wenn es für das Analyseziel notwendig ist und dies im Antrag auf den Datenzugang explizit begründet wird. Diese aus datenschutzrechtlicher Sicht besonders sensiblen Merkmale sind:

<sup>2</sup> Mit den Stata-Befehlen `label language en` oder `label language de` kann auf englisch- bzw. deutschsprachige Labels umgestellt werden.

### **SIG-Kern und BG-Mitglieder Modul:**

- Geburtsmonat (gebmon)
- Staatsangehörigkeit (nation\_lhg, nation\_beh, nation\_leh, nation\_x\_asu, nation\_x\_mth)
- Berufsuntergruppe aktuelle Tätigkeit (beruf2010\_4\_aktT)
- Berufsuntergruppe letzte abgeschlossene Ausbildung (beruf2010\_4\_ausb)
- Berufsuntergruppe gesuchte Tätigkeit (beruf2010\_4\_gesucht)
- Wohnort Kreis (wo\_kreis\_lhg, wo\_kreis\_beh, wo\_kreis\_leh, wo\_kreis\_x\_asu)
- Berufsuntergruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 4-Steller (beruf2010\_4\_vor2007)
- Maßnahmeart – Gruppe (mass)
- Arbeitsort Kreis (ao\_kreis)
- Wirtschaftszweig - Wirtschaftsunterklasse (w08\_5\_kons)

Beachten Sie, dass die sensiblen Merkmale in der Regel in vergrößerter Version bereits in den Daten enthalten sind. Die sensiblen Merkmale werden nur weitergegeben, wenn die Information im vergrößerten Merkmal zur Erreichung des Forschungsziels nicht ausreichend ist.

Die Merkmale der LST-S sind aufgrund ihrer Detailliertheit und des damit einhergehenden Deanonymisierungspotenzials in ihrer Gänze sensibel und können daher nur auf begründeten Antrag zur Verfügung gestellt werden.

#### **1.4.2 Dateinamen in der SIG 0723**

<u>SIG-Kern:</u>	sig_kern_0723_v1.dta
<u>BG-Modul:</u>	sig_bg_0723_v1.dta
<u>LST-S:</u>	sig_0723_v1_lsts_v1.dta

## **1.5 Veränderung zur SIG 0720**

### **1.5.1 Merkmalskanon**

In der SIG 0723 wurden gegenüber der Version SIG 0720 neue Merkmale zum Tagesentgelt aufgenommen, die bei der Anrechnung von Einmalzahlungen unterstützen und eine Imputation der zensierten Werte rechts der Beitragsbemessungsgrenze anbieten<sup>3</sup>. Die zugrundeliegenden Datenarbeiten werden in Drechsler et al. (2023) beschrieben. Weitere Informationen zu den neuen Merkmalen finden sich unter „Tagesentgelt (inkl. Einmalzahlungen) (entgelt\_tag\_bonus)“, „Tagesentgelt (imputiert) (entgelt\_tag\_imp)“ und „Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung (mehrfbesch\_bonus\_tentg)“.

In der SIG 0723 wurde gegenüber der Version SIG 0720 zudem das Merkmal „Integrationsprognose (ipo)“ neu aufgenommen. Die Integrationsprognose wird ab Ende 2016 im Vermittlungsprozess zur Gruppierung der BA-Kunden verwendet.

---

<sup>3</sup> Das in den Skripten von Stüber et al. (2023) enthaltene Verfahren zur Imputation der Löhne kann nun durch die neu angebotenen Variablen ersetzt werden.

Des Weiteren ist das sensible Merkmal „Maßnahmeart – Gruppe (mass)“ neu enthalten. Das Merkmal bietet detailliertere Informationen zu spezifischen Maßnahmeteilnahmen. Das Merkmal ist im Gegensatz zur vergrößerten Version (Maßnahmeart - Kategorie (mass\_gr)) nur auf begründeten Antrag hin erhältlich.

Darüber hinaus wurde auch das Merkmal „Ausbildung (imputiert) (ausbildung\_imp\_beh)“ neu aufgenommen. Das Merkmal ist für die BeH gefüllt.

Tabelle 1 gibt einen Überblick über die Änderungen im Merkmalskanon über alle Quellen hinweg.

**Tabelle 1: Neue, entfallene, umbenannte und veränderte Merkmale in der SIG 0723**

<b>Merkmal</b>	<b>Erläuterung</b>
Tagesentgelt (inkl. Einmalzahlungen) (entgelt_tag_bonus)	Neues Merkmal
Tagesentgelt (imputiert) (entgelt_tag_imp)	Neues Merkmal
Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung (mehrfbesch_bonus_tentg)	Neues Merkmal
Integrationsprognose (ipo)	Neues Merkmal
Maßnahmeart - Gruppe (mass)	Neues Merkmal, detaillierte Maßnahmeart-Differenzierung des IAB
Maßnahmeart - Kategorie (mass_gr)	Umbenannt, alter Variablenname „massn“
Ausbildung imputiert (ausbildung_imp_beh)	Neues Merkmal

### 1.5.2 Datendokumentation

Im vorliegenden Datenreport wurden die Beschreibungen für folgende Merkmale überarbeitet bzw. ergänzt:

- Ausbildung (ausbildung\_beh, ausbildung\_x\_asu)
- Ausbildung imputiert (ausbildung\_imp\_beh)
- Tagesentgelt (entgelt\_tag)
- Tagesentgelt (inkl. Einmalzahlungen) (entgelt\_tag\_bonus)
- Tagesentgelt (imputiert) (entgelt\_tag\_imp)
- Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung (mehrfbesch\_bonus\_tentg)
- Maßnahmeart - Kategorie (mass\_gr)
- Maßnahmeart - Gruppe (mass)
- Integrationsprognose (ipo)

Der Abschnitt „1.7 Merkmalsliste“ wurde gelöscht. Als Ersatz dienen die detaillierten Übersichten zu den einzelnen Merkmalen (Label, Auszählungen, fehlende Werte) im Online-Anhang zum Datenreport.

## 1.6 Steckbrief

**Tabelle 2: Steckbrief**

Themen/Merkmalsgruppen	<p><b>Leistungshistorik Grundsicherung (LHG):</b> Daten zu Personen, die Leistungen zur Sicherung des Lebensunterhalts nach SGB II beziehen (Trägerarten: ARGE/gemeinsame Einrichtungen, getrennte Trägerschaft/getrennte Aufgabenwahrnehmung, zugelassene kommunale Träger)</p> <p><b>Leistungsempfängerhistorik (LeH):</b></p>
------------------------	--

	<p>Informationen zu Leistungsbezug nach SGB III für Bezieher von Arbeitslosengeld</p> <p><b>Arbeitsuchendenhistorik (ASU):</b> Informationen zur Arbeitsuche, die in BA-Verfahren erfasst werden</p> <p><b>Arbeitsuchendenhistorik aus XSozial-BA-SGB II (XASU):</b> Informationen zur Arbeitsuche, die über den Übermittlungsstandard XSozial-BA-SGB-II von zugelassenen kommunalen Trägern an die BA gemeldet werden</p> <p><b>Maßnahmeteilnahnehistorik (MTH):</b> Informationen zu Teilnahmen an Maßnahmen der aktiven Arbeitsmarktförderung (keine Maßnahmen der zugelassenen kommunalen Träger im SGB II)</p> <p><b>Maßnahmeteilnahnehistorik aus XSozial-BA-SGB II (XMTH):</b> Informationen zu Teilnahmen an Maßnahmen zur Eingliederung, die über den Übermittlungsstandard XSozial-BA-SGB-II von zugelassenen kommunalen Trägern an die BA gemeldet werden.</p> <p><b>Beschäftigtenhistorik (BeH):</b> Jahres- und Abmeldungen zur Sozialversicherung für sozialversicherungspflichtig und geringfügig Beschäftigte</p> <p><b>Leistungsstatistik nach dem SGB II (LST-S)</b> Informationen über Bedarfe, Einkommen, Sanktionen und Ansprüche von leistungsberechtigten Personen und weiteren Personengruppen im Umfeld einer Bedarfsgemeinschaft</p>
Untersuchungseinheit(en)	Leistungsempfänger und deren Bedarfsgemeinschaftsmitglieder, sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Arbeitsuchende, Maßnahmeteilnehmende, Beschäftigungsbetriebe
Fallzahlen	905.135 Personen (SIG-Kern) 24.966.840 überschneidungsfreie Sätze (nach Episodensplitting) (SIG-Kern) 1.784.760 Personen (BG-Mitglieder) 50.179.541 überschneidungsfreie Sätze (nach Episodensplitting) (BG-Mitglieder)
Zeitraum	01.01.2007 – 31.12.2023
Zeitbezug	Tagesgenau
Regionale Gliederung	Bundesländer, Kreise
Gebietsstand	Gebietsstandsaktualisiert zum 31.12.2023
Erhebungsdesign	4 % Zufallsstichprobe von Personen aus der LHG
Beteiligte Institutionen	Sozialversicherungsträger, Bundesagentur für Arbeit, kommunale Grundsicherungsträger
Frequenz der Aktualisierung	Laufend
Dateiformat/ -größe	Stata; SIG-Kern: 4,7 GB; BG-Mitglieder: 9,4 GB; LST-S: 6,5 GB
Dateiorganisation	Die Daten werden in drei Dateien abgelegt. Die erste Datei enthält personenbezogene Informationen der Stichprobenpersonen und die zweite Datei personenbezogenen Informationen der Bedarfsgemeinschaftsmitglieder. Auf begründeten Antrag wird eine weitere Datei mit der Leistungsstatistik LST-S bereitgestellt.
Datenzugang	Remote Desktop, Gastaufenthalt, Datenfernverarbeitung
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	<p><b>SIG-Kern und BG-Mitglieder:</b> Geburtsmonat (gebmon), Staatsangehörigkeit (nation_lhg, nation_beh, nation_leh, nation_x_asu, nation_x_mth), Maßnahmeart – Gruppe (mass), Berufsuntergruppe aktuelle Tätigkeit (beruf2010_4_aktT), Berufsuntergruppe letzte abgeschlossene Ausbildung (beruf2010_4_ausb), Berufsuntergruppe gesuchte Tätigkeit (beruf2010_4_gesucht), Wohnort Kreis (wo_kreis_lhg, wo_kreis_beh, wo_kreis_leh, wo_kreis_x_asu), Berufsuntergruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 4-Steller (beruf2010_4_vor2007), Arbeitsort Kreis (ao_kreis), Wirtschaftszweig - Wirtschaftsunterklasse (w08_5_kons)</p> <p><b>LST-S:</b> komplett sensibel</p>
Zitierung der Daten und Datendokumentation	<u>Daten:</u>

	<p>„Die Datengrundlage dieses Beitrags bildet die schwach anonymisierte Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023. Der Datenzugang erfolgte über Remote Desktop/Gastaufenthalt am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (FDZ) und mittels kontrollierter Datenfernverarbeitung beim FDZ.“</p> <p>DOI: 10.5164/IAB.SIG0723.de.en.v1</p> <p><u>Datendokumentation:</u></p> <p>Dummert, Sandra; Hohmeyer, Katrin; Lietzmann, Torsten; Sauer, Irakli; Oertel, Martina (2025): Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023. FDZ-Datenreport, 14/2025 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2514.de.v1</p>
Datensatzversion	Stichprobe der Integrierten Grundsicherungsbiografien (SIG) 2007 - 2023 (SIG 0723); DOI: 10.5164/IAB.SIG0723.de.en.v1

## 1.7 Mengengerüst

**Tabelle 3: Mengengerüst**

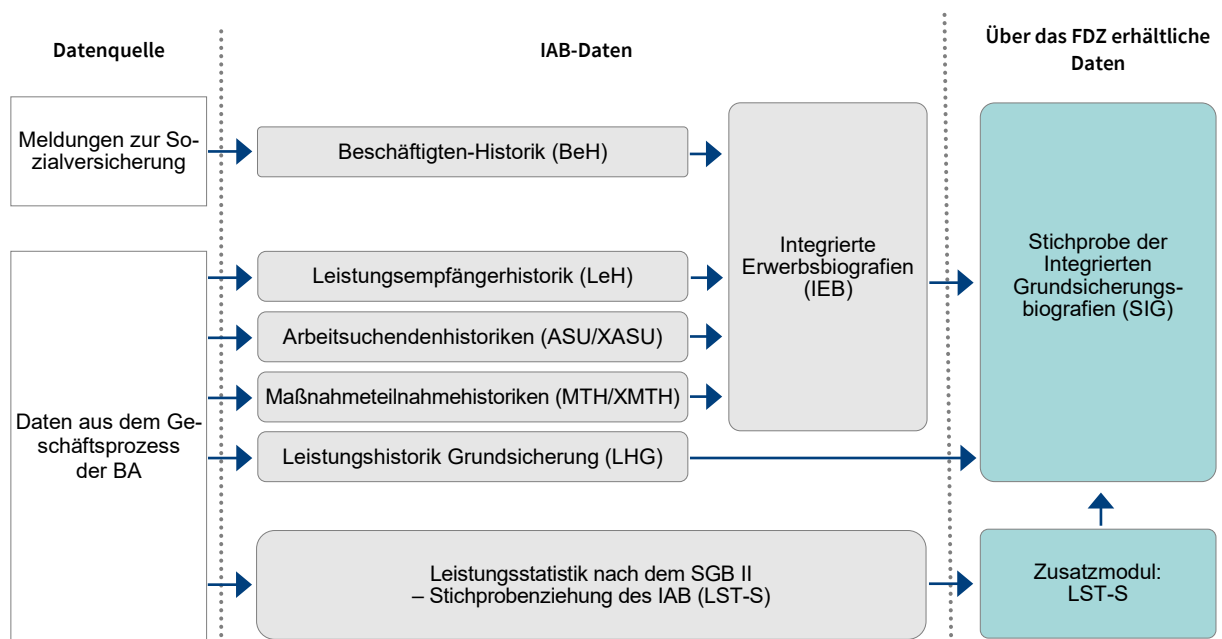
Fallzahlen	SIG-Kern	BG-Mitglieder
LHG	13.577.419	31.887.733
LeH	1.224.743	1.887.418
ASU	13.233.142	25.652.579
XASU	2.209.447	5.086.523
MTH	1.764.780	3.329.212
XMTH	451.848	960.224
BeH	12.267.868	21.599.699
<b>Sätze insgesamt*</b>	<b>24.966.840</b>	<b>50.179.541</b>
<b>Personen</b>	<b>905.135</b>	<b>1.784.760</b>

\* Die Anzahl der Sätze insgesamt entspricht aufgrund der Datenstruktur mit Spells aus potentiell mehreren Quellen nicht der Summe der Sätze über allen Quellen.

## 2 Datenquellen

Die Stichprobenziehung für die SIG erfolgt aus der Leistungshistorik Grundsicherung (LHG). Informationen zum Bezug von Leistungen nach dem SGB II entstammen ebenfalls aus der LHG, während die weiteren administrativen Personendaten zum Großteil aus den Integrierten Erwerbsbiografien (IEB) des IAB gezogen wurden. Diese vereinen Daten aus verschiedenen Datenquellen, die jeweils Informationen aus unterschiedlichen EDV-Verfahren enthalten können. Zusätzlich werden einige ergänzende Merkmale aus diesen Datenquellen in die administrativen Personendaten aufgenommen, die nicht Teil der IEB sind. Abbildung 2 veranschaulicht die Datenflüsse, die zur SIG führen. Im Folgenden werden diese Datenquellen kurz beschrieben.

Abbildung 2: Quellen der SIG



**Hinweis:** Die IAB-eigene Stichprobenziehung LST-S basiert auf der Leistungsstatistik nach dem SGB II der Statistik der Bundesagentur für Arbeit.

Bevor die Daten aus den oben genannten Quellen zur IEB zusammengefügt werden, durchlaufen sie quellenspezifische Bereinigungsverfahren (siehe auch Folgeabschnitte). Die Gesamt-IEB erfahren die folgenden Bereinigungen:

- Sätze, bei denen das Alter an ihrem Ende kleiner 13 oder zu Beginn größer 75 ist, werden gelöscht.
- Datensätze, bei denen das Enddatum kleiner als das Beginndatum ist, werden gelöscht.
- Widersprüchliche Angaben zu Geschlecht oder Geburtsdatum innerhalb eines Kontos werden bereinigt.
- Datensätze ohne Angabe von Geburtsdatum und Geschlecht nach Bereinigung werden gelöscht.

Weitergehende Bereinigungen, wie z.B. das Ergänzen von vermutlich fehlenden Meldungen, Streikbereinigungen, etc. erfolgen nicht.

## 2.1 Leistungshistorik Grundsicherung (LHG)

Die Leistungshistorik Grundsicherung (LHG) enthält Informationen zu erwerbsfähigen Leistungsberechtigten und Personen in ihren Bedarfsgemeinschaften (siehe Kapitel 1.2). Mit den SIG-Daten ist erstmals für Forschende außerhalb des IAB eine Verknüpfung von Individuen mit SGB-II-Leistungsbezug auf Personenebene möglich. Leistungsbezug nach SGB II umfasst sowohl Leistungen zur Sicherung des Lebensunterhalts (z.B. ALG II), wie auch Zuschläge zum ALG II oder auch zusätzliche Leistungen. Die LHG enthält jedoch keine Angaben zu den jeweiligen Leistungssätzen, diese sind in der Leistungsstatistik aus dem SGB II (LST-S) enthalten und können auf Antrag hinzugespielt werden.

Im Gegensatz zu den Leistungen aus dem Rechtskreis des SGB III ist die Bundesagentur für Arbeit nicht alleiniger Träger der Leistungen. Es wird demnach unterschieden nach den drei möglichen Trägerarten:

- Arbeitsgemeinschaften (ARGEn, bis 2010)/gemeinsame Einrichtungen (ab 2011) aus gemeinsamer Aufgabenwahrnehmung von BA und Kommune,
- Getrennte Trägerschaft (gT)/getrennte Aufgabenwahrnehmung (bis 2011) – hier bestand eine Aufgabenteilung zwischen BA und Kommune<sup>4</sup>,
- Zugelassene kommunale Träger, die nach der anfänglichen Experimentierklausel des § 6a auch als optierende Kommunen oder Optionskommunen bezeichnet werden – hier übernimmt die Kommune alle Aufgaben aus dem Bereich des SGB II.

Die Daten der LHG stammen aus verschiedenen Meldeverfahren. In der Regel findet das IT-Verfahren A2LL bis 2010 bei ARGEn und ab 2011 bei gemeinsamen Einrichtungen Verwendung. 2014 wurde A2LL sukzessive durch ALLEGRO als neues IT-Verfahren für die Grundsicherung im SGB II in gemeinsamen Einrichtungen ersetzt. Diese Systemumstellung führte zu einer Neuvergabe von Bedarfsgemeinschaftsnummern<sup>5</sup>. Zugelassene kommunale Träger verwenden verschiedene eigene IT-Verfahren und übermitteln ihre Daten mittels des Standards XSozial-BA-SGB II einmal monatlich an die Bundesagentur für Arbeit. Bei Kommunen mit getrennter Trägerschaft kommen beide Verfahren zum Einsatz. Die unterschiedlichen Datenstandards haben Auswirkungen auf Umfang und Qualität der gelieferten Daten.

Der linke Datenrand der LHG ist der 1. Januar 2005. Die Datenquelle ist jedoch bis Anfang 2007 lückenhaft. Aus diesem Grund wurde für die Stichprobenziehung der SIG das Jahr 2007 als linker Rand gewählt.

---

<sup>4</sup> Die Kommune übernimmt die Kosten der Unterkunft und Heizung (§ 22 SGB II) sowie die abweichend zu erbringenden Leistungen (§ 23 (3) SGB II) und die flankierenden Leistungen nach § 16 (2) Satz 2 Nr. 1 - 4 SGB II. Von der BA werden dagegen Regelleistungen, Sozialversicherungsbeiträge und Eingliederungsleistungen (SGB III und SGB II) sowie spezifische Leistungen ohne die oben genannten flankierenden Leistungen erbracht.

<sup>5</sup> Zusammengehörende LHG-Spells können somit nur anhand der Personennummer, aber nicht anhand der Bedarfsgemeinschaftsnummer identifiziert werden.

## 2.2 Leistungsempfängerhistorik (LeH)

Die Leistungsempfängerhistorik (LeH) erfasst Zeiträume, in denen Personen Lohnersatzleistungen der Bundesagentur für Arbeit beziehen (Rechtskreis SGB III). Die Leistungen beinhalten Arbeitslosengeld, Arbeitslosenhilfe, Unterhaltsgeld und von der Bundesagentur für Arbeit übernommene Beiträge zur privaten Kranken- bzw. Pflegeversicherung während des Leistungsbezugs. Darüber hinaus sind auch Informationen über Sperr- und Unterbrechungszeiten enthalten:

- Eine Sperrzeit ist ein Zeitraum, in dem eine leistungsberechtigte Person aus Gründen, die sie selbst zu vertreten hat, keine Leistung erhält. Typisch ist z.B. eine kurze Sperrzeit zu Beginn eines Leistungsanspruchs, weil die Person sich verspätet arbeitsuchend gemeldet hat.
- Eine Unterbrechungszeit ist hingegen ein Zeitraum ohne Leistungserhalt, den die leistungsberechtigte Person nicht selbst zu vertreten hat. Das kann z.B. eine zwingende Ortsabwesenheit sein. Eine Unterbrechungszeit verringert nicht die Anspruchsdauer, d.h. die Person erhält zwar während der Unterbrechungszeit keine Leistung, dafür verschiebt sich aber das reguläre Ende des Leistungsanspruchs entsprechend nach hinten.

Leistungen nach dem Rechtskreis SGB II (z.B. Arbeitslosengeld II) sind nicht enthalten. Da der Leistungsanspruch von der Erfüllung gesetzlicher Voraussetzungen abhängt, werden Arbeitslosigkeitsperioden, in denen die Voraussetzungen nicht vorliegen (z.B. fehlende Bedürftigkeit im Falle der Arbeitslosenhilfe oder Nichterfüllung der Anwartschaftszeit beim Arbeitslosengeld), nicht in der Leistungsempfängerhistorik erfasst.

Es werden folgende für das Datenprodukt relevante Bereinigungen der LeH vorgenommen:

- Sätze ohne gültiges Beginndatum wurden ausgeschlossen.
- Sätze mit leerem oder ungültigem Enddatum wurden ausgeschlossen, da hier nicht davon ausgegangen werden kann, dass überhaupt eine Zahlung geleistet worden ist.
- Für die SIG werden Informationen zu Arbeitslosenhilfe und Unterhaltsgeld entfernt.

## 2.3 Arbeitssuchendenhistoriken (ASU/XASU)

Daten über Arbeitssuchende werden in den Arbeitssuchendenhistoriken (ASU und XASU) gespeichert. Die Quelle ASU enthält Informationen zu Arbeitssuchenden, welche bei Arbeitsagenturen verzeichnet sind, ab 2005 auch solche, die bei Arbeitsgemeinschaften und getrennten Trägerschaften gemeldet sind. In der Quelle XASU sind ab 2005 die Daten von arbeitssuchenden Personen im ALG-II-Bezug von zugelassenen kommunalen Trägern enthalten, welche nach dem XSozial-Standard konsolidiert gemeldet werden. Neben arbeitssuchenden und nichtarbeitssuchenden Arbeitssuchenden sind auch Informationen zu anderweitig gemeldeten Erwerbsfähigen enthalten, insbesondere der Personenkreis von Bedarfsgemeinschaftsmitgliedern, die nach § 10 SGB II für die Arbeitssuche nicht verfügbar sind (ab 2011 in beiden Datenquellen valide erfasst).

Es werden folgende für das Datenprodukt relevante Bereinigungen der (X)ASU vorgenommen:

- Es erfolgt keine Konsolidierung der ASU-Sätze einer einzelnen Person. Es kann daher vorkommen, dass Überschneidungen zwischen ASU-Sätzen existieren.
- Die personenbezogenen Merkmale, die für die Quellen (X)ASU gefüllt sind, beziehen sich immer auf den Beginn des Satzes.

- Ein neuer ASU-Satz wird angelegt, sobald ein Statuswechsel (z.B. von arbeitsuchend zu arbeitslos) erfolgt. Dies gilt auch, wenn sich die Trägerart (Agentur für Arbeit, ARGE, gemeinsame Einrichtung, zugelassener kommunaler Träger, getrennte Aufgabenwahrnehmung) ändert.
- In der XASU werden überschneidungsfreie Zeiträume bezogen auf eine einzelne Person abgebildet. Bei Änderungen folgender Merkmale wird für die XASU jeweils ein neuer Datensatz gebildet:
  - Wechsel des Arbeitsuche-Status
  - Wechsel der Verfügbarkeit nach § 10 SGB II
  - Wechsel des SGB II-Trägers
  - Wechsel des Wohnortes

## 2.4 Maßnahmeteilnahmehistoriken (MTH/XMTH)

Die Maßnahmeteilnahmehistoriken (MTH und XMTH) enthalten Informationen, die verschiedenen Rechtskreisen zugeordnet werden können. Zum einen sind Maßnahmen und Maßnahmeteilnahmen der aktiven Arbeitsmarktförderung nach dem SGB III enthalten. Zum anderen beinhalten die MTH Maßnahmen aus dem Rechtskreis SGB II, wenn diese in Verfahren der BA erfasst werden. In der Quelle XMTH sind ab 2005 Maßnahmen der zugelassenen kommunalen Träger geführt.

Es werden folgende für das Datenprodukt relevante Bereinigungen der MTH vorgenommen:

- Sätze, die mehr als ein Jahr nach Maßnahmeende erstellt wurden, werden gelöscht, falls ein anderer Satz existiert, der innerhalb eines Jahres nach Abschluss der Maßnahme erstellt wurde.
- Es wird lediglich der aktuellste Datensatz zu einer einzelnen Maßnahmeteilnahme verwendet.
- In die IEB werden nur Maßnahmeteilnahmen aufgenommen, die als „tatsächlich stattgefunden“ klassifiziert sind. Teilnahmen, die nicht oder noch nicht stattgefunden haben, werden gelöscht. Teilnahmen werden auch dann als nicht stattgefunden gewertet, wenn während der Teilnahme an einer Maßnahme ein Löschdatum gesetzt wird.
- Bestimmte Maßnahmearten werden nicht aufgenommen. Hierzu zählen Unterstützung der Beratung und Vermittlung, Mobilitätshilfen und reine Reha-Maßnahmen.

Es werden folgende Bereinigungen der XMTH vorgenommen:

- Für die XMTH besteht die besondere Herausforderung darin, technisch oder organisatorisch verursachte Mehrfachmeldungen für eine Maßnahmeteilnahme zu identifizieren und zu behandeln. Die Ursachen von diesen Doppelmeldungen sind zum einen eine bis April 2009 fehlende Identifikationsnummer für Maßnahmeteilnahmen bzw. unsachgemäße Handhabung der eingeführten Förder-ID und zum anderen Trägernummernwechsel im Zusammenhang mit Fusionen oder anderen Umorganisationen von Grundsicherungsträgern mit der Neuorganisation der Arbeitsagenturen 2012/2013. Zudem kam es zu trägerspezifischen Problemen bei Versions- oder Produktwechseln der kommunalen Software oder bei Versionsänderungen des Standards XSozial. Daraus resultieren zeitweise in beträchtlichem Umfang Meldungen für Maßnahmeteilnahmen der gleichen Förderungsart, die sich pro Teilnehmendem teilweise oder vollständig zeitlich überlappen. Echte Dubletten werden identifiziert und aussortiert, da

sie mit hoher Wahrscheinlichkeit technische Dubletten sind. Versteckte Dubletten sind ebenfalls identifizierbar und werden aussortiert. Dabei gilt im Prinzip immer die Regel, dass der aktuellste Informationsstand behalten wird.

- Überlappende und unmittelbar aneinander angrenzende Meldungen der gleichen Maßnahmeart (aus XSozial) werden zu einer Beobachtung zusammengefasst. Maßnahmespezifische Merkmale der zusammengefassten Zeiträume werden auf Systemmissings gesetzt.
- Anschließend werden die weiteren Merkmale zusammengestellt und berechnet. Dabei wird auf konsolidierte Personendaten aus XSozial-Historiken und LHG zurückgegriffen.
- Abschließend werden folgende Maßnahmen ausgeschlossen:
  - Einmalleistungen (wie UBV/Mobi-/Vermittlungsbudget/LES) und
  - spezifische Reha-Maßnahmen

## 2.5 Beschäftigtenhistorik (BeH)

Die Quelle für Informationen über Beschäftigung ist die Beschäftigtenhistorik (BeH) des IAB. Die Grundlage der Daten ist das mit Wirkung vom 1. Januar 1973 eingeführte (und seit dem 1. Januar 1991 auf Ostdeutschland ausgedehnte) integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung, das durch das Kürzel DEÜV (früher DEVO/DÜVO) bezeichnet ist (vgl. für weitere Details: Bender et al., 1996, S. 4ff.; Wermter/Cramer, 1988). Es verlangt von den Arbeitgebern, mindestens einmal jährlich, Meldungen für alle sozialversicherungspflichtig beschäftigten Arbeitnehmer an die zuständigen Sozialversicherungsträger abzugeben. Die BeH erfasst alle Arbeiter und Angestellten sowie alle Auszubildenden, soweit sie nicht von der Sozialversicherungspflicht befreit sind. Dies bedeutet, dass verbeamtete und selbstständige Tätigkeiten sowie Studierende<sup>6</sup> (vgl. Cramer, 1985) grundsätzlich nicht in der BeH erfasst werden. Seit der Änderung des Meldeverfahrens zum 1. Januar 1999 werden auch geringfügig Beschäftigte und mithelfende Familienangehörige erfasst (enthalten erst ab 1. April 1999). Die Daten werden von den Krankenkassen aufgenommen, von der Bundesagentur für Arbeit gesammelt und aufbereitet und anschließend vom IAB in dessen Historikdatei integriert.

Es wurden folgende für das Datenprodukt relevante Bereinigungen der BeH vorgenommen:

- Um einen im Zeitablauf möglichst konstanten Personenkreis zu erfassen, werden einige Personengruppen ausgeschlossen, deren Daten nicht für den gesamten Beobachtungszeitraum zur Verfügung stehen. Ab dem Meldejahr 2011 stammen die Daten der BeH aus neu konzipierten Quelldaten. Dadurch wurde eine Reihe von Personengruppen reaktiviert bzw. neu eingeführt, da sie von der BA-Statistik als sozialversicherungspflichtig eingestuft werden. Damit sind fortan die Personengruppen 101 - 107, 111 - 114, 118, 119, 120, 140, 141, 142, 143, 149, 201 und 203 - 205 enthalten sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind unter anderem kurzfristig Beschäftigte, d. h. Personengruppen 110, 202 und 210.
- Die Personengruppen 123, 124 und 127 wurden 2011 neu eingeführt.
- Aus datenschutzrechtlichen Gründen werden die Personengruppen 107, 111, 113, 114, 127 und 204 zur Personengruppe "Sonstige Erwerbstätige" (599) zusammengefasst.

---

<sup>6</sup> Studierende können dennoch in der BeH erscheinen, wenn sie parallel zum Studium z.B. geringfügig beschäftigt waren.

- Ab dem Meldejahr 2012 wurden Auszubildende mit den Personengruppen 121 und 122 neu aufgenommen.
- Sätze mit fehlendem Entgelt sowie der Ausprägung 101 beim Merkmal Erwerbsstatus und dem Wert 50 beim Abgabegrund (Jahresmeldung) werden nicht in die IEB übernommen.
- Geschlecht und Geburtsdatum werden aus dem Data Warehouse (DWH) der Statistik der BA übernommen. Dort liegen diese Angaben quellenübergreifend harmonisiert vor.

## 2.6 Leistungsstatistik nach dem SGB II (LST-S)

Die Leistungsstatistik nach dem SGB II (Grundsicherungsstatistik) – Stichprobenziehung des IAB (LST-S) enthält Informationen über Bedarfe, Einkommen, Sanktionen und Ansprüche von leistungsberechtigten Personen und weiteren Personengruppen im Umfeld einer Bedarfsgemeinschaft<sup>7</sup>. Die Daten werden von der Statistik der Bundesagentur für Arbeit (BA) aus A2LL, ALLEGRO und zPDV / STEP sowie dem Übermittlungsstandard XSozial-BA-SGB-II (XSozial) gewonnen und quellenübergreifend konsolidiert. Das Verfahren "LST-S" erzeugt Datensätze mit monatsgenauen Informationen. Das Merkmalspektrum konzentriert sich dabei auf Geldbeträge. Für vom Leistungsbezug ausgeschlossene BG-Mitglieder sind generell keine Finanzdaten verfügbar.

Die Finanzdaten des LST-S liefern Informationen zu Bedarfen, Einkommen, Ansprüchen und Sanktionen. Vereinfachend lässt sich zusammenfassen, dass zuerst der Gesamtbedarf einer Person ermittelt wird. Dieser wird dann gegebenenfalls um vorhandenes Einkommen aus Erwerbstätigkeit (auch Selbständigkeit), anderen Transferleistungen sowie aus Kapital und Vermögen reduziert. Der sich aus dieser Berechnung ergebende Leistungsanspruch gleicht häufig dem finalen Zahlungsanspruch, kann aber dann abweichen, wenn vom Jobcenter Sanktionen ausgesprochen wurden, um diese der Anspruch reduziert wird.

Für das SIG-Datenprodukt wurden Kinder ohne Leistungsberechtigung, also mit bedarfsdeckendem Einkommen aufgrund unsicherer Datenqualität ausgeschlossen. Die kleinste zeitliche Einheit ist dabei ein Monat. Im Anschluss wurden aus Gründen der Datensparsamkeit die Informationen mehrerer konsekutiver Monate einer Person immer dann aggregiert bzw. aufsummiert, wenn sämtliche ausgewählten Finanzinformationen dieser Monate identisch sind.

Eine Verbindung der LST-S mit dem SIG-Kern ist auf unterschiedliche Weise möglich und hängt von der Fragestellung ab. Die zu präferierende Strategie verknüpft beide Datensätze auf Monatsebene. Dieser Schritt ist allerdings rechen- und speicherplatzintensiv, weshalb ausreichend Zeit eingeplant werden sollte. Außerdem wird hierfür dringend angeraten, beide Datensätze – den SIG-Kern und die LST-S – vorab um Informationen zu bereinigen, die für die Analysen nicht benötigt werden, sprich die Datensätze auf diejenigen Personengruppen, Zeiträume und Variablen einzugrenzen, die tatsächlich für die Analysen benötigt werden. Dies reduziert die Laufzeit und verhindert, dass dieser Prozess an den Kapazitätsgrenzen der FDZ-Nutzerserver (30 GB) scheitert.

Für die Verknüpfung empfehlen wir, sowohl die LHG-Spells des SIG-Kerns als auch die LST-S-Spells zunächst auf Monatsebene herunterzubrechen, anschließend zusammenspielen und zu guter Letzt wieder auf die ursprünglichen Zeiträume der Spells des SIG-Kerns zu aggregieren.

---

<sup>7</sup> Statistik der Bundesagentur für Arbeit: Statistik der Grundsicherung für Arbeitsuchende nach dem SGB II, Nürnberg April 2018.

Nachfolgender Stata-Code (siehe Kasten 1) zeigt eine Variante auf, wie die LST-S auf einfache Weise zurück auf die monatliche Darstellung gebracht werden kann, im Rahmen derer sämtliche über größere Zeiträume ohne inhaltliche Änderungen hinweg zusammengefassten Spells wieder getrennt und ihren jeweiligen Monaten zugewiesen werden.

---

#### Kasten 1: Beispielcode für Stata 18

```
* Ausweitung des LST-S-Datensatzes entsprechend der jeweiligen Spelllänge:
expand spell_length
* Aufteilung der aggregierten Finanzinformationen auf die einzelnen Monate:
qui d bedarf* zanspr* zbeink* veink* sankt*, varl
foreach v in `r(varlist)' {
    dis "`v'"
    gen z_`v'=`v'/spell_length
    drop `v'
    rename z_`v' `v'
}
* Generierung eines Monatsidentifikators:
bys persnr bg_id beg_mon end_mon: gen count = _n-1
by persnr bg_id beg_mon end_mon: gen int mon_id = beg_mon if _n==1
by persnr bg_id beg_mon end_mon: replace mon_id = mon_id[1] + count if mon_id==.
format mon_id %tm
drop count beg_mon end_mon
```

## 3 Datenaufbereitung und Stichprobenziehung

### 3.1 Stichprobenziehung

Die SIG ist eine 4 %-Zufallsstichprobe<sup>8</sup> aller Personen mit einem Eintrag in der Leistungshistorik Grundsicherung (LHG). Dadurch ist die SIG repräsentativ für den Personenkreis der LHG im entsprechenden Zeitraum ab 2007.

### 3.2 Struktur und Aufbereitung der SIG

Die SIG stellt die bislang umfassendste Datengrundlage für den Bezug von Leistungen der Grundsicherung für Arbeitsuchende nach SGB II dar, die über das FDZ zur Verfügung gestellt wird. Dementsprechend stehen im Mittelpunkt der Daten die Informationen zu Zeiträumen des Leistungsbezugs und die entsprechenden Merkmale der Personen und Bedarfsgemeinschaften. Um die Bio-

---

<sup>8</sup> Im Zuge der Datenaufbereitung und Anonymisierung gehen einige wenige Fälle verloren, sodass die SIG streng genommen geringfügig weniger als 4 % aller potentiellen Fälle enthält.

grafien der Leistungsbeziehenden möglichst breit abzubilden, wurden die entsprechenden biographischen Informationen aus anderen administrativen Datenprodukten zu Arbeitslosigkeit und Arbeitssuche, Beschäftigung, der Teilnahme an Maßnahmen der aktiven Arbeitsmarktpolitik und der Bezug von Leistungen des SGB III (Arbeitslosengeld I) hinzugespielt. Diese können auf Grund des institutionellen Aufbaus des Systems der Grundsicherung sowohl parallel zum Leistungsbezug (z.B. erwerbstätige Leistungsbeziehende, „Aufstocker“) als auch außerhalb des Leistungsbezugs auftreten.

Das Grundprinzip des Aufbaus der SIG ist, dass es pro Person und Zeitraum nur eine Zeile in den Daten gibt. Ein solcher Aufbau ist zum einen inhaltlich sinnvoll, da es eine dominante Quelle in den Daten gibt, nämlich die Leistungshistorik Grundsicherung (LHG) mit den Bezugszeiten in der Grundsicherung, die auf Grund der Stichprobe von Leistungsbeziehenden für jede Person mindestens einmal auftritt. Dieser Basisinformation werden dann die Informationen aus den anderen Quellen zugeordnet. Zum anderen erspart dieser Aufbau dem Nutzer die Aufbereitung der Daten über mehrere Zeilen und Quellen hinweg. Des Weiteren ist dieser Aufbau platzsparender als die Bereithaltung aller parallelen Informationen in den verschiedenen Einzelquellen. Die finale Datenstruktur hält somit überschneidungsfreie, tagesgenaue Informationen zu Leistungsbezugs- und Erwerbsverläufen auf einer einheitlichen Zeitschiene bereit, wie sie beispielsweise für Verweildauer- oder Sequenzmusteranalysen benötigt werden.

Da diese Struktur nicht derjenigen in den Rohdaten entspricht, müssen die Daten bereinigt werden. Dies erfolgt in einem ersten Schritt auf Basis der einzelnen Datenquellen und im Anschluss für die zusammengespilten Daten. Dafür müssen Entscheidungen getroffen werden, welche Informationen als relevant oder prioritär eingestuft und damit behalten werden. Dieses Vorgehen wird in den folgenden Abschnitten beschrieben.

### 3.2.1 SGB-II-Anonymisierung

Um das Deanonymisierungsrisiko zu verringern, wird in der SIG standardmäßig nur das Geburtsjahr ausgewiesen. Der Geburtsmonat kann als sensibles Merkmal nur bei begründetem Bedarf beantragt werden. In den Quellen LHG und (X)ASU besteht allerdings die Gefahr, dass durch die zeitliche Struktur der Beobachtungen dennoch das exakte Geburtsdatum sichtbar werden könnte. Sätze könnten systematisch an bestimmten Geburtstagen oder dem Renteneintritt beginnen bzw. am Tag davor enden.

Um eine indirekte Bestimmung des exakten Geburtstags zu verhindern, wird daher das folgende Verfahren angewendet. Sätze, die an im SGB-II-Kontext neuralgischen Geburtstagen geteilt wurden, bei denen der Geburtstag also einen Split auslöst, erhalten den 15. des jeweiligen Monats als neuen Spellbeginn. Dementsprechend wird das Enddatum des vorhergehenden Spells auf den 14. des Monats gelegt. In seltenen Fällen liegt zwischen diesem Satz und dem 15. des Monats ein weiterer, abgeschlossener Spell, der in Folge der Anonymisierung gelöscht wird. Parallel liegende Episoden anderer Bedarfsgemeinschaftsmitglieder werden entsprechend angepasst.

Aufgrund dieser Anonymisierungsschritte geht außer dem exakten Geburtstag keine weitere Information verloren. Bei dieser Korrektur kommt es allerdings zu einer leichten Verzerrung der Satzdauer.

### 3.2.2 Leistungshistorik Grundsicherung (LHG)

Die Informationen der LHG werden wie folgt weiteraufbereitet:

- In der LHG kann es auf Personenebene zu Doppelmeldungen kommen (siehe Kapitel 4.2), die wie folgt bereinigt werden:
  - Doppelmeldungen, die in Übergangszeiten von Meldesystemwechseln bestehen, werden anhand einer Filtertabelle bereinigt.
  - Bei parallelen Spells einer Person, die zeitgleich in zwei Bedarfsgemeinschaften ist, werden folgende Bereinigungen durchgeführt:
    - Es bleibt der Spell mit der "höheren BG-Rolle" erhalten;
    - bei gleicher BG-Rolle bleibt derjenige, der erwerbsfähiger Hilfebedürftiger ist, erhalten;
    - wenn bei beiden Spells die Erwerbsfähigkeitsangabe gleich ist, dann bleibt derjenige, der aus einer Meldung bei der BA oder einer gemeinsamen Einrichtung stammt, erhalten;
    - wenn auch das gleich ist, dann bleibt schließlich derjenige mit der niedrigeren BG-ID erhalten.
- Bei überlappenden Enden zweier Spells wird der Beginn des späteren Spells einen Tag nach Ende des vorangehenden Spells gesetzt.
- Leistungsepisoden, die nur einen einzigen Tag dauern, werden gelöscht. Bei diesen handelt es sich meist um einmalige Leistungen, die gewährt werden. Personen, die nur eine Leistungsepisode haben, die nur einen einzigen Tag dauert, sind daher nicht enthalten.
- Umsetzung Monatskonzept: Beginn und Ende von Leistungsepisoden werden auf Monatsbeginn bzw. -ende gesetzt und Lücken innerhalb eines Kalendermonats geschlossen, da die Bedürftigkeit für den gesamten Monat gilt.
- Aneinandergrenzende Spells werden zusammengehängt, wenn dies ohne Informationsverlust möglich ist, und der linke und der rechte Datenrand des Datensatzes entsprechend gesetzt.

### 3.2.3 Leistungsempfängerhistorik (LeH)

Aus der LeH werden für die SIG nur Informationen zum Bezug von Arbeitslosengeld genutzt und keine zu z. B. Arbeitslosenhilfe und Unterhaltsgeld. Zudem werden nur die Informationen zur Höhe des Leistungsbezugs und zum Abmeldegrund weiter verwertet.

Es werden die Episoden – wie bei den anderen Quellen auch – so bereinigt, dass es zeitgleich nur eine Information aus dieser Quelle gibt. Dafür werden im Einzelnen folgende Schritte vollzogen:

- Es werden Ein-Tages-Episoden gelöscht.
- Parallele Episoden mit gleichen Eigenschaften werden bereinigt.
- Bei parallelen Episoden mit unterschiedlichen Informationen wird diejenige mit einem höheren Entgelt und einer niedrigeren (aber gültigen) Abgangsnummer behalten.
- Anschließend werden LeH-Episoden ohne Lücke zusammengefasst, wenn dies ohne Informationsverlust möglich ist, und der linke und rechte Datenrand entsprechend gesetzt.

### 3.2.4 Arbeituchendehistoriken (ASU und XASU)

Die ASU und XASU werden gemeinsam aufbereitet. Zunächst wird das höchste schulische und berufliche Ausbildungsniveau vor dem linken Datenrand gesichert. Anschließend werden folgende Schritte durchgeführt, um parallele Episoden zu bereinigen:

- Bei parallelen Episoden, die aus unterschiedlichen Quellen stammen, wird diejenige aus der ASU behalten, da hier von einer besseren Datenqualität ausgegangen werden kann.
- Bei parallelen Episoden aus derselben Quelle wird die mit dem jeweils niedrigeren Wert des Erwerbsstatus beibehalten (d.h. Arbeitslosigkeit sticht Arbeitssuche oder Nichtverfügbarkeit).
- Wenn eine der parallelen Episoden keine Angabe zu wichtigen Variablen enthält, die andere aber schon, so wird letztere behalten und die andere verworfen.
- Von den übrigen, parallelen Episoden wird die mit längerer, anonymisierter Dauer im Originaldatensatz verwendet.
- Von den darüber hinaus verbleibenden, parallelen Episoden wird zufällig eine behalten.
- Anschließend werden Episoden, die maximal 7 Tage Lücke haben und bei denen es keine Änderungen in zentralen Variablen gibt, zusammengefasst und die Daten am linken und rechten Rand zensiert.

### 3.2.5 Maßnahmeteilnehnehistorik (MTH) und Maßnahmeteilnehnehistorik aus XSozial (XMTH)

MTH und XMTH werden einzeln, aber analog zueinander aufbereitet. Die Aufbereitung erfolgt in folgenden Schritten:

- Es werden Ein-Tagesepisoden entfernt. Hierbei handelt es sich z.B. um Umzugshilfen oder Bewerbungstrainings.
- Die Maßnahmen werden kategorisiert und anschließend wie folgt nach ihrer angenommenen Intensität priorisiert: (1) Aufnahme einer Erwerbstätigkeit, (2) Beschäftigungsschaffende Maßnahmen, (3) Berufliche Weiterbildung, (4) Berufswahl & Berufsausbildung, (5) Aktivierung & berufliche Eingliederung, (6) Freie und sonstige Förderung. Dabei wird bei parallelen Episoden die mit der Maßnahme aus der Kategorie, die als wichtiger eingestuft wurde, behalten. Bei mehreren parallelen Episoden der gleichen Maßnahmekategorie wird die Episode ohne fehlende Werte bevorzugt behalten. Jetzt gibt es pro Teilquelle und Zeitraum nur noch maximal eine Episode.
- Aufeinanderfolgende Episoden ohne Lücke, die die gleichen Ausprägungen haben, werden zusammengehängt.
- Episoden werden am linken und rechten Datenrand zensiert.

Anschließend werden die beiden Quellen zusammengefügt und gemeinsam bereinigt, sodass es pro Zeitraum nur noch eine Information gibt. Bei parallelen Episoden in MTH und XMTH wird die aus der MTH behalten, da hier eine höhere Datenqualität angenommen wird. Liegt eine Episode in einer anderen, so wird die innenliegende Episode gelöscht und die längere behalten. Bei überlappenden Episoden, werden die Daten der XMTH-Episode so korrigiert, dass die MTH-Teile behalten und die XMTH-Teile gelöscht werden. Anschließend werden aneinandergrenzende Episoden zusammengehängt, sofern das ohne Informationsverlust möglich ist.

### 3.2.6 Beschäftigtenhistorik (BeH)

Es gehen nur BeH-Spells mit einem Beginndatum ab dem 1.1.1993 in die Datenaufbereitung mit ein, da ab 1993 Daten der Rentenversicherungsträger in auswertbarer Qualität auch für Ostdeutschland vorliegen. Zunächst werden Episoden von Altersvollrentnern (`ieb_erw_stat= 119, 149`) und bei vermuteter Beschäftigung (`120`) entfernt.

Die angestrebte Datenstruktur, die nur eine Information pro Zeitraum und Quelle für eine Person erlaubt, macht es notwendig, dass bei mehreren parallelen oder überlappenden Beschäftigungen die Hauptbeschäftigung identifiziert wird. Bei parallelen Beschäftigungsepisoden wird diejenige als Hauptbeschäftigung behalten, die sozialversicherungspflichtig ist bzw. diejenige mit höherem Tagesentgelt. Zudem werden einige zentrale Informationen zu einer möglichen zweiten (parallelen) sozialversicherungspflichtigen und geringfügigen Beschäftigung in separaten Variablen gesichert.

Im nächsten Schritt werden fehlende Jahresmeldungen (Vorgehen siehe FDZ-Methodenreport 6/2007) ergänzt, wenn die BeH-Meldung am 31.12. endet und ein Jahr später am 1.1. wieder eine Meldung vorhanden ist und bei beiden Beschäftigungsverhältnissen Betriebsnummer, Erwerbsstatus und Stellung im Beruf übereinstimmen.

Es werden Episoden zusammengefügt, die maximal sieben Tage Lücke haben und bei denen sich Beruf, Entgelt, berufliche Stellung, Erwerbsstatus, Wirtschaftszweig und Zweitbeschäftigung nicht ändern. Sehr kurze Episoden von bis zu sieben Tagen werden anschließend entfernt, wenn diese davor und danach keine weiteren, innerhalb von sieben Tagen angrenzenden Spells haben.

Für Analysen kann es wünschenswert sein, Informationen zur Erwerbstätigkeit außerhalb des Datenfensters der SIG zu haben. Daher werden einige Informationen zur Erwerbserfahrung und letzten Beschäftigung zwischen dem 1.1.1993 und dem linken Rand gesichert. Die Episoden werden zusammengehängt, wenn dies ohne Informationsverlust möglich ist.

### 3.2.7 Leistungsstatistik nach dem SGB II (LST-S)

Die Finanzdaten der LST-S mehrerer konsekutiver Monate werden für das SIG-Datenprodukt immer dann aggregiert bzw. aufsummiert, wenn sämtliche ausgewählte Finanzinformationen dieser Monate identisch sind. Die kleinste zeitliche Einheit im Endprodukt ist daher ein Monat und jede Information gilt auch immer für mindestens einen Monat, nicht für kürzere Zeiträume.

Um das Deanonymisierungsrisiko, insbesondere von exakten Geburtstagen, zu verringern, werden in der LST-S verschiedene Schritte unternommen: Zum einen werden manche Merkmale nur in dichotomer Form bereitgestellt, um das der Kombination aus vergleichsweise seltenen Finanzflüssen und deren Höhe anhängende Risiko zu verringern. Diese Informationen gehen aber nicht verloren, sondern erlauben die Identifikation bestimmter Mehrbedarfsfälle in einem gegebenen Zeitraum. Hierbei handelt es sich um Fälle von Schwangerschaft, Behinderung oder besonderer, kostenintensiver Ernährung aufgrund medizinischer Notwendigkeit (siehe § 21 SGB II). Zudem kann über einen solchen dichotomen Indikator die Gewährung von Darlehen anstelle von Zuschüssen – nicht aber deren Höhe – identifiziert werden.

Des Weiteren werden sämtliche Leistungsdaten zu Einkommen, Bedarfen, Ansprüchen und Sanktionen durch Addition ganzzahliger Zufallswerte zwischen -10 und +10 verfremdet und zu guter Letzt auf das nächstgelegene Vielfache von 10 gerundet.

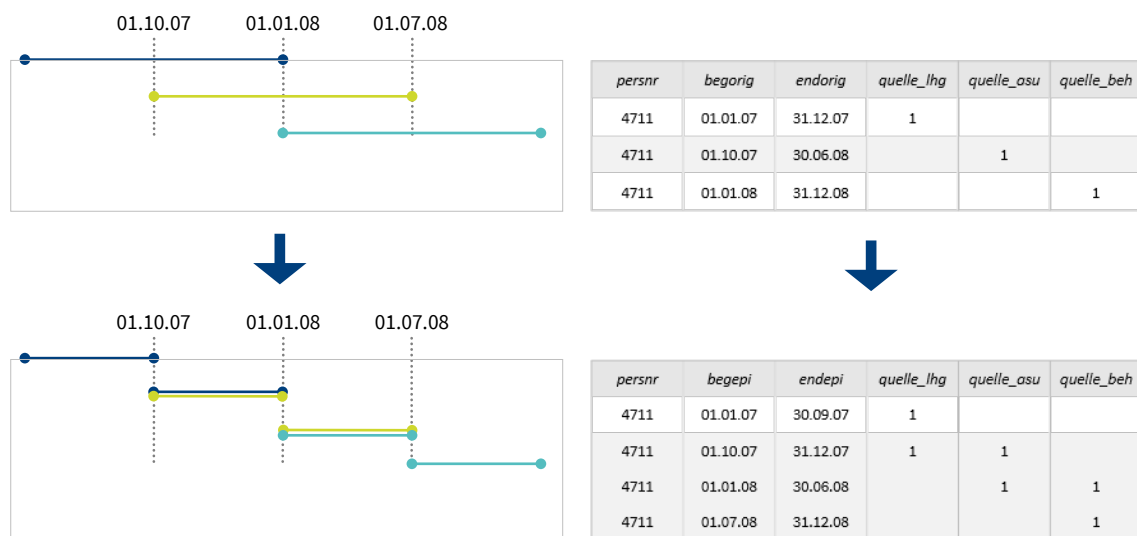
### 3.3 Episodensplitting

Die administrativen Personendaten liegen „gesplittet“ vor. Falls sich innerhalb eines Kontos Sätze überlappen, werden diese durch künstliche Sätze mit neuen Datumsangaben derart ersetzt, dass vollständig parallele oder überschneidungsfreie Zeiträume entstehen (Episodensplitting). Die Anzahl der Sätze erhöht sich dadurch (siehe Abbildung 3). In der SIG wird dieses Episodensplitting folgendermaßen umgesetzt: zum einen liegen die Einzelquellen ursprünglich in gesplitteter Form vor. Das heißt, wenn sich einzelne Merkmale ändern, sind die Spells in der jeweiligen Quelle geteilt an dem Datum ab dem die neue Merkmalsausprägung auftritt. In der Aufbereitung der Einzelquellen (siehe oben) werden dann Spells wieder zusammengefasst, wenn der Split auf einem Merkmal beruhte, das in der SIG nicht enthalten ist. Zudem werden kurze Unterbrechungen von max. sieben Tagen (BeH, (X)ASU) geschlossen. In der LHG werden für den Grundsicherungsbezug Spells, die während eines Monats beginnen oder enden, jeweils auf den jeweiligen Monatsbeginn bzw. das jeweilige Monatsende gesetzt.

In einem zweiten Schritt werden die einzelnen Quellen zusammengespielt und gesplittet um die angestrebte Datenstruktur von einem Spell pro Person und Zeitraum zu erreichen. Finden also Veränderungen innerhalb einer Quelle statt, werden auch die Informationen zu den anderen Quellen gesplittet. Ändert sich beispielsweise für eine Person während eines zusammenhängenden Leistungsbezug in der Grundsicherung (Quelle LHG) der Arbeitslosigkeitsstatus der Person von arbeitslos zu arbeitsuchend (Quelle (X)ASU), wird der Spell dieser Person an dieser Stelle gesplittet. Es entstehen dadurch zwei parallele Spells, bei denen anschließend alle relevanten Informationen auf einen übertragen werden und der zweite dann gelöscht wird. Eine inhaltliche Plausibilitätsprüfung oder Konsolidierung der Informationen zwischen den verschiedenen Quellen findet nicht statt. Ein solches, durch Splitting erzeugtes Ende (oder Beginn) bedeutet dementsprechend nicht zwangsläufig, dass eine zusammenhängende Episode hinsichtlich jeder der Quellen tatsächlich endet (oder beginnt). Für die Identifikation von neu beginnenden (oder endenden) Leistungsbezugs-, Arbeitslosigkeits-, Maßnahmeteilnahme- oder Beschäftigungsepisoden muss jeweils geprüft werden, ob der jeweilige Zustand auch davor (danach) besteht.

Für welche Quellen der jeweilige Spell Informationen enthält, kann aus den „quelle-\*“-Variablen abgelesen werden.

Abbildung 3: Episodensplitting



### 3.4 Bereinigung von Inkonsistenzen

Verschiedene – notwendige – Eingriffe in die den Zeitraum einer jeden Episode bestimmenden Datumsangaben können in manchen Konstellationen dazu führen, dass bestimmte Variablen innerhalb einer Bedarfsgemeinschaft von leichten Inkonsistenzen betroffen sind. Dies betrifft insbesondere den Bedarfsgemeinschaftstyp sowie die Zuordnung von Neugeborenen.

In beiden Fällen werden die Inkonsistenzen im Zuge einer Bereinigung behoben. Dabei wird für den 1. Tag einer jeden Episode (einer Person innerhalb einer Bedarfsgemeinschaft) eine neue Bedarfsgemeinschaftstyp-Variable gebildet, die in den wenigen Fällen einer Abweichung der Originalvariable zum Einsatz kommt. Zuvor werden – getreu dem 1. Tag als Stichtag – bereits Episoden von Neugeborenen entfernt, falls ihr Geburtsdatum nach dem 1. Tag oder durch die Datenaufbereitung bzw. Anonymisierung gar erst nach dem Ende einer Episode liegt.

### 3.5 Fehlende Werte

In den Daten gilt die folgende Codierung von fehlenden Werten:

Tabelle 4: Fehlende Werte

Bezeichnung	Wert	Beschreibung
Keine (valide) Angabe	.z	Nicht systematisch fehlender Wert. Das Merkmal ist prinzipiell für die Quelle gefüllt, jedoch ist für den betrachteten Satz ein Wert nicht vorhanden bzw. nicht sinnvoll zu interpretieren.
Fehlt systematisch	.n	Ein Merkmal ist grundsätzlich für eine Quelle oder für einen gewissen Zeitraum nicht gefüllt.
Wert unplausibel (nur LST-S)	.p	Die Plausibilität der übermittelten Leistungsdaten aus der LST-S wird als Ergebnis einer Prüfung der BA-Statistik in Abstimmung mit den liefernden Grundsicherungsträgern infrage gestellt und somit für die SIG auf den Missing-Wert .p gesetzt.

## 4 Datenqualität und Probleme

### 4.1 Einführung des SGB II und spätere Trägerwechsel

Mit der Einführung des SGB II zum 1.1.2005 wurden auch die Zuständigkeiten für die Betreuung von Arbeitsuchenden neugestaltet, sodass die Aufgabenwahrnehmung im SGB II von unterschiedlichen Trägertypen übernommen werden kann:

- In der Regel übernahmen die Arbeitsgemeinschaften (ARGE) aus BA und Kreis die Aufgaben im Bereich des SGB II. Diese wurde 2011 von den gemeinsamen Einrichtungen (gE) – auch „Jobcenter“ genannt – abgelöst.
- Zudem war es bis Ende 2011 möglich, dass BA und Kreis die ihnen übertragenen Aufgaben in eigener (getrennter) Zuständigkeit – in sogenannten getrennten Trägerschaften (gT) oder auch getrennten Aufgabenwahrnehmungen (gAw) erbrachten.
- Seit 2005 besteht zudem die Möglichkeit, dass zugelassene kommunale Träger (zkT; auch: optierende Kommunen) die Aufgaben übernehmen. Zunächst erhielten 69 Landkreise oder kreisfreie Städte die alleinige Trägerschaft der Grundsicherung für Arbeitsuchende bis 31.12.2010. Mit der Verfassungsänderung wurde die Optionsmöglichkeit ab 2012 auf 110 Träger ausgedehnt.

Während die ARGE/gE das Leistungs- und Fallmanagement über die BA-Verfahren pflegen und die erfassten Daten an die Statistik der BA übermitteln, verwenden die kommunalen Grundsicherungsträger jeweils eigene Softwaresysteme. Die Übermittlung an die Statistik der BA erfolgt über XSozial. Für getrennte Träger lag eine Aufteilung der Zuständigkeiten vor. Der jeweilige Trägerteil war nur für die Erfassung und Lieferung der Daten verantwortlich, die in seinen Zuständigkeitsbereich fallen. Daher gab es reduzierte Datenanforderungen für die kommunale Seite der getrennten Träger. Die Übermittlung über den Standard XSozial erfolgt einmal monatlich zu einem festgelegten Zeitfenster. In der Woche darauf besteht an einem zweiten Meldetag die Möglichkeit, missglückte Lieferungen zu wiederholen. Im Lauf der Jahre wurden zahlreiche Qualitätssicherungsinstrumente entwickelt. Dennoch erhöhen Versionswechsel des Standards oder der vor Ort genutzten Software regelmäßig das Risiko für Datenqualitätsprobleme.

Im Rahmen der Neuorganisation von SGB-II-Trägern jeweils zum 01. Januar der Jahre 2011 bis 2014 gingen diverse ARGE/gE und getrennte Träger in kommunale Trägerschaft über bzw. wurden von zkT zu gE umgewandelt. Diese Trägerwechsel und die damit einhergehende – zum Teil auch ohne Trägerwechsel – erfolgte Umstellung der vor Ort verwendeten Software führen teilweise auch zu Brüchen in den Daten der SGB-II-Quellen. Detaillierte Informationen zu den daraus resultierenden Qualitätsproblemen in den Daten finden sich in den einzelnen Abschnitten zu den Datenquellen.

### 4.2 Leistungshistorik Grundsicherung (LHG)

- Längsschnittuntersuchungen für Personen sind mit Unsicherheiten behaftet, da es zu monatlichen Untererfassungen des Leistungsbezugs auf Grund von Datenlieferproblemen oder –ausfällen auf Trägerebene kommen kann. Das heißt, dass bei Personen, die in solchen Gebieten

wohnen oder dorthin umziehen, ein nicht zu beobachtender Leistungsbezug auf Lieferprobleme zurückzuführen ist oder tatsächlich kein Bezug vorlag.

- Es kommt bei zugelassenen kommunalen Trägern immer wieder zu meist einmonatiger trägerspezifischer Untererfassung.
- Die Lückenvariable kann identifizieren, ob vor oder nach einem Spell eine solche Lücke vorgelegen haben kann.
- Im Zusammenhang mit Trägerartwechseln kommt es zu Unter-/Übererfassungen:
  - Im Rahmen der Trägergebietsreform vom 1. Januar 2011 kommt es zu Untererfassungen im Gebiet der Arbeitsagenturen Dessau-Roßlau, Halberstadt, Halle und Sangerhausen.
  - Doppelmeldungen aufgrund der Gebietsreformen 2009/2011 und der Trägerformwechsel zum 1. Januar 2012 und zum 1. Januar 2013 werden in der IEB bereits soweit möglich bereinigt. Dennoch können Doppelmeldungen auftreten.
- Bei folgenden Jobcentern bestehen Unschärfen bei der regionalen Zuordnung von Leistungsfällen:
  - zwischen Emden und Norden zwischen September und Dezember 2009
  - zwischen Döbeln und Mittelsachsen von Oktober bis Dezember 2012
  - zwischen Tirschenreuth und Wunsiedel von November 2012 bis März 2013
- Manche Personen, für die eine LHG-Beobachtung vorliegt, sind vom Leistungsbezug nach SGB II ganz oder teilweise ausgeschlossen, da sie z. B. eine geförderte Ausbildung machen, Altersrente beziehen, in einer stationären Einrichtung untergebracht sind oder Versicherungsleistungen zur Vermeidung von Hilfebedürftigkeit bekommen. Dies betrifft pro Jahr ca. 3 bis 5 Prozent der Fälle. In XSozial kommt es zu trägerspezifischen Untererfassungen dieser Personengruppe.
- In der offiziellen Leistungsstatistik der BA werden Meldelücken durch ein statistisches Schätzverfahren auf aggregierter regionaler Ebene ergänzt. In der LHG werden keine ergänzenden Datensätze bereitgestellt. Ein Vergleich der SIG mit der Leistungsstatistik ist daher z.T. nur eingeschränkt möglich.
- Eine Besonderheit der Grundsicherungsdaten besteht darin, dass es sich zwar um Individualdaten handelt, Art und Höhe des Leistungsbezugs jedoch nicht immer individuell bestimmt sind. Vielmehr sieht das Gesetz vor, dass Personen in einer Bedarfsgemeinschaft (BG) gewillt sind „Verantwortung füreinander zu tragen und füreinander einzustehen" (§ 7, Abs. 3 SGB II). Dies bedeutet, dass Veränderungen in der Bedarfsgemeinschaft Auswirkungen auf die Leistungsberechtigung aller Personen des jeweiligen Grundsicherungsfalls haben können. Einer fachlichen Bedarfsgemeinschaft können mehrere BG-Nummern zugeordnet sein, was im Rahmen der Datenaufbereitung (siehe Kapitel 3.2.2) bereinigt wurde. Ursachen können sein:
  - Ein Umzug in ein Gebiet eines anderen Grundsicherungsträgers. Da die Träger rechtlich unabhängig voneinander sind, werden die Fälle zwar teilweise in dem gemeinsamen IT-Verfahren A2LL verwaltet, fachlich sind sie jedoch getrennt zu halten. Daher ist die Trägernummer Bestandteil der BG-Nummer (die ersten fünf Stellen).

- Ein Trägernummernwechsel wegen einer Neuorganisation des Trägers bzw. der zugeordneten Arbeitsagentur erzeugt ohne Umzug neue Kunden- und Bedarfsgemeinschafts-IDs.
- Technischer Grund im IT-Verfahren, z. B. bei einer Datenmigration mit Vergabe einer neuen Nummer oder Korrektur einer Falschmeldung. Auch der Wechsel von A2LL zu ALLEGRO führt zum Wechsel der BG-Nummer.
- Zu- oder Abgänge von BGs im Zeitverlauf oder BG-Wechsel von Personen sind nicht zuverlässig identifizierbar. Die Betrachtung oder Stichprobenziehung von Bedarfsgemeinschaften zu einem Zeitpunkt ist mit der BG-ID weitgehend unproblematisch.
- Wenn Kinder nur temporär im Haushalt leben, z.B. weil sie regelmäßig abwechselnd bei jeweils einem der getrennt lebenden Elternteile wohnen, können sie Mitglieder einer solchen BG sein und daher theoretisch zwei BGs gleichzeitig angehören.

## 4.3 IEB

### 4.3.1 Lücken in den Erwerbsbiografien

Die IEB beinhalten Informationen zu unterschiedlichen Erwerbszuständen. Allerdings ist in den administrativen Daten nicht jede Art von Erwerbstätigkeit erfasst. Zudem gibt es keine umfassenden Informationen zu anderen biografischen Zuständen (z.B. bzgl. Bildungs- oder Erziehungszeiten). Dies führt dazu, dass es Zeiträume gibt, für die für eine Person kein Satz in den IEB enthalten ist (individuelle Lücken). Für Auswertungszwecke kann es jedoch durchaus interessant sein, Informationen zu diesen Lücken in den IEB zu haben (z.B. für Kontrollgruppenbildung, Lebensverlaufsanalysen etc.). Für einige dieser Lücken kann das Merkmal „Abgabegrund/Beendigungsgrund/SGB-II-Einstellungsgrund/Abmeldegrund“ des direkten Vorgängersatzes (sofern ein entsprechender Satz vorhanden ist) Hinweise auf die Art der Lücke liefern (z. B. wenn dort Informationen enthalten sind, dass die Beendigung einer Arbeitslosigkeitsepisode wegen Aufnahme einer selbständigen Tätigkeit erfolgt ist).

### 4.3.2 Arbeitssuchendenhistoriken (ASU/XASU)

#### 4.3.2.1 ASU

- Mit der Einführung des SGB II zum 1.1.2005 werden arbeitssuchende Personen nicht mehr vollständig mit BA-Verfahren erfasst. Die Quelle ASU berücksichtigt ab diesem Zeitpunkt nur noch Personen, die von der BA im Rechtskreis SGB III (Arbeitsförderung), oder von ARGE n, gE bzw. gT im Rechtskreis SGB II (Grundsicherung) betreut werden.
- Für die Vermittlungskräfte ist es nicht möglich, die Zuordnung zum Rechtskreis immer unmittelbar zu erfassen, da vielfach erst zeitverzögert feststeht, welcher Träger aufgrund einer möglichen SGB-II-Leistungsberechtigung vorrangig zuständig ist.
- Für manche Personen, für die ab 2005 eine Optionskommune zuständig war, wurden von der BA parallel "künstliche" ASU-Sätze angelegt. Diese sind über das Merkmal „estatvor“ (Wechsel zu einem zugelassenen kommunalen Träger) zu identifizieren.
- Das bis Juni 2006 genutzte Verfahren coArb unterstützte ausschließlich die Vermittlung von Arbeitslosen und Arbeitssuchenden. Lückenhaft erfasst wurden auch Daten zu Personen, die

sich lediglich beraten lassen wollten. Die Daten der Berufsberatung wurden in einem separaten System gehalten. Mit VerBIS erweiterten sich die Ausprägungen des Arbeitsuche-Status um „Ratsuchende“ und Personen „ohne Status“. Zu Letzteren gehören ALG-II-Leistungsberechtigte, die eingeschränkt der Arbeitsvermittlung verfügbar sind. Ihre Erfassung in VerBIS gilt erst seit Januar 2008 als weitgehend vollständig.

- Ein Wechsel der Grundsicherungsträger oder des Wohnorts führen nicht zu einem neuen ASU-Satz, die Merkmalsausprägung jeweils zu Beginn einer Episode wird fortgeführt. Mit zunehmender Dauer des Spells wird also das Risiko größer, dass Träger oder Wohnort nicht mehr zutreffend sind.
- Seit Frühjahr 2011 werden die Arbeitsuchendendaten aus BA-Verfahren und XSozial in der integrierten Arbeitslosenstatistik konsolidiert verarbeitet. Hier kann es zu größeren Abweichungen mit der BA-Statistik kommen. Am aktuellen Rand sind die Bestandsdaten jedoch identisch.

#### 4.3.2.2 XASU

- Da in die XASU auch LHG-Episoden als Grundlage für die Meldungen zu Nichtarbeitsuchenden nach § 10 SGB II einfließen, sind einzelne Monate von Lieferausfällen betroffen. Daneben bestehen Untererfassungen aufgrund von technisch induzierten fälschlichen Stornierungen, die mit den angewendeten Entstornierungsverfahren nicht vollständig rekonstruiert werden können.
- Punktuell kann es auch zu Übererfassungen kommen. Hierfür kann eine Vielzahl technischer Probleme verantwortlich sein, u.a. systematisch nicht erfolgte Abmeldungen, falsche Entstornierungen und parallele Arbeitsuchendmeldungen in mehreren Kreisen.
- Verschiedene Merkmale weisen für die XASU zum Teil nur einen sehr geringen Füllgrad auf. Hiervon betroffen sind unter anderem die Merkmale Schulabschluss, Abgabegrund sowie Erwerbsstatus vor Arbeitsuche. Die Füllgrade dieser Merkmale verbessern sich zwar im Zeitablauf, bleiben aber dennoch zum Teil trägerspezifisch unbefriedigend. Das Merkmal Beruf ist nahezu für den gesamten zur Verfügung stehenden Zeitraum in der XASU nicht gefüllt.
- Bei etlichen Trägern (Kreisen) erscheint der Anteil der gemeldeten Arbeitsuchenden an den gemeldeten erwerbsfähigen Hilfebedürftigen in den IEB durchgängig oder zeitweise unplausibel hoch. Gründe dafür könnten eine möglicherweise fehlerhafte Bestimmung des Status „Nichtarbeitslos, arbeitssuchend“ oder eine technisch fehlerhafte Übermittlung mit hohem Anteil an Nachmeldungen bei diesen Trägern sein.
- Die träger- und zeitraumabhängige Plausibilität der XASU-Angaben sollte jeweils unter Berücksichtigung der Untersuchungsfrage vor der Verwendung geprüft werden.
- Durch Unterschiede in Konsolidierungsregeln, Zeitbezügen und regionalen Zuordnungen kommt es zu Abweichungen von den veröffentlichten BA-Statistiken.

#### 4.3.3 Maßnahmeteilnahmehistoriken (MTH/XMTH)

##### 4.3.3.1 MTH

- Die MTH enthält nur Meldungen, die in BA-Verfahren gebucht sind. Deren Nutzung nimmt in Arbeitsgemeinschaften / Getrennten Trägerschaften / getrennten Aufgabenwahrnehmungen

zwischen 2005 und 2007 kontinuierlich zu. Ab März 2007 sind die Meldungen für diese Träger komplett. Meldungen, die zugelassenen kommunalen Träger über XSozial übermitteln, finden sich in der XMTH.

- Durch die Neuorganisation von SGB-II-Trägern 2011-2014 (siehe Abschnitt 4.1) kann es zur Aufspaltung der Dokumentation von Maßnahmeteilnahmen in der MTH und XMTH kommen, wenn es zu einer Umstellung des Meldeverfahrens kommt<sup>9</sup>. Dies kann sich in einem Split des Maßnahmespells, aber auch in Doppelungen niederschlagen (siehe Abschnitt 2.4).
- Bei Meldungen zum Überbrückungsgeld wird zum Teil die maximal zulässige Dauer von einem halben Jahr überschritten. Dies kann in den meisten Fällen durch eine Voreinstellung der Eingabemaske im Erfassungssystem erklärt werden.
- Die MTH wird um Bewerbermerkmale (z. B. Ausbildung) aus anderen BA-Verfahren ergänzt. Für diese Merkmale erfolgte im Jahr 2006 die Umstellung des operativen Verfahrens von coArb auf VerBIS. Es gelten hier die gleichen Qualitätseinschränkungen wie für die ASU.

#### 4.3.3.2 XMTH

- Ca. 13 % aller Grundsicherungsträger melden von 2005 bis Anfang 2017 fast keine kommunalen Eingliederungsleistungen (vormals Flankierende Leistungen), die u.a. in der zusammengefassten Kategorie „sonstige Förderung“ enthalten sind. Viele weitere Jobcenter melden nur zeitweise und / oder nur ausgewählte Arten. Das führt zu einer Untererfassung der „sonstigen Förderung“.
- Übererfassung von Teilnahmen: Die Gesamtbestände wirken ab Ende 2008 stabil. Bei der Mehrzahl der Träger sind aber auch danach zeitweise auffällige Lieferprofile zu registrieren, die Zweifel an der Qualität wecken. Erst die Einführung der XSozial-Förder-ID im Jahr 2009 und die mittlerweile mehrjährige Erfahrung aller Beteiligten stabilisieren das Meldegeschehen und die daran anschließende Datenverarbeitung. Auch die ab 2012 neu hinzukommenden zkT melden weit überwiegend unauffällig, so dass Umfang und Dauer mutmaßlicher Übererfassung merklich zurückgehen.
- Durch die Neuorganisation von SGB-II-Trägern 2011-2014 (siehe Abschnitt 4.1) kann es zur Aufspaltung der Dokumentation von Maßnahmeteilnahmen in der MTH und XMTH kommen, wenn es zu einer Umstellung des Meldeverfahrens kommt. Dies kann sich in einem Split des Maßnahmespells, aber auch in Dopplungen niederschlagen (siehe Abschnitt 2.4).
- Die über monatliche Stichtagsauszählungen in der XMTH erhaltenen Zahlen bzw. Personen unterscheiden sich durch die umfangreichen Konsolidierungsarbeiten in mehrfacher Hinsicht von den veröffentlichten BA-Statistiken.

#### 4.3.4 Beschäftigtenhistorik (BeH)

- Die Übermittlung von Informationen zur Ausbildung, ausgeübten Tätigkeit und Stellung im Beruf erfolgt durch Meldungen der Arbeitgeber nach DEÜV (siehe Abschnitt 2.5) über einen sogenannten Tätigkeitsschlüssel. Für Zeiträume mit einem Enddatum nach dem 30.11.2011 er-

---

<sup>9</sup> Weitere Informationen zur Gebietsstruktur der SGB-II-Träger und relevanten Änderungen finden sich unter: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikationen/Regionale-Gliederungen/Gebietsstruktur-Traeger-Grundsicherung-Nav.html>

folgte der Wechsel zum neuen Tätigkeitsschlüssel 2010 (siehe Bertat et al., 2013, für mehr Details). Dieser Wechsel wurde von den Spitzenorganisationen der Sozialversicherungsträger beschlossen, da der Tätigkeitsschlüssel 2003 eine Reihe von Sachverhalten nicht mehr zeitgemäß und nicht mehr realitätsnah erfassen konnte. Da die Meldungen der Arbeitgeber nach DEÜV in die BeH einfließen, hat die Umstellung des Tätigkeitsschlüssels ausschließlich Auswirkungen für Beobachtungen aus dieser Quelle. Die Messung folgender bisher enthaltener Angaben ist von dieser Umstellung betroffen: Arbeitszeit, Beruf, Stellung im Beruf sowie schulisches und berufliches Qualifikationsniveau. Darüber hinaus gibt es seit der Umstellung auch Angaben zur Befristung eines Beschäftigungsverhältnisses sowie dazu, ob eine Person als Verleihpersonal in einem Zeitarbeitsunternehmen angestellt ist. Die wichtigste Auswirkung ist der Umstieg auf eine neue Berufsklassifikation. Statt der bisherigen Klassifikation der Berufe 1988 (KldB 1988) wird mit dem neuen Tätigkeitsschlüssel die differenziertere KldB 2010 gemeldet.<sup>10</sup>

- Die Einführung des neuen Tätigkeitsschlüssels hat zu einer Reihe von Problemen geführt. So kam es während des Übergangszeitraums, der Arbeitgebern im Meldeverfahren zur Sozialversicherung eingeräumt wurde, zu einer temporären Zunahme fehlender Angaben. Analysen der Statistik der BA (Bertat et al., 2013, S. 10) zeigen, dass die neuen bzw. umgestellten Merkmale „ausgeübte Tätigkeit“, „Arbeitszeit“ und „Ausbildungsabschluss“ nach der Umstellung zu etwa 20 bis 30 Prozent keine Angaben enthielten. Eine deutliche Verbesserung stellt sich im ersten Halbjahr 2013 ein. Um die Qualität der Arbeitszeitangabe im Übergangszeitraum zu verbessern, haben Ludsteck/Thomsen (2016) ein Imputationsverfahren entwickelt, das die fehlenden durch imputierte Werte ersetzt. Die imputierten Daten werden ab der SIG 0720 verwendet. Für die Lücken in den anderen Merkmalen werden keine Imputationen vorgenommen.
- Innerhalb des Meldeverfahrens ist eine gewisse Zeitverzögerung nicht vermeidbar. Zwar müssen Änderungen von Beschäftigungsverhältnissen sofort gemeldet und bestehende Beschäftigungsverhältnisse jährlich bis zum April (bis Jahresmeldung 2012) bzw. Mitte Februar (ab Jahresmeldung 2013) des folgenden Jahres bestätigt werden. Faktisch treffen Meldungen jedoch zum Teil um Jahre verspätet ein. Die Historikdatei des IAB wird aber nicht kontinuierlich fortgeschrieben, sondern in bestimmten Abständen aktualisiert. Dies geschieht unter Verwendung von Dateien mit den Beschäftigten-Meldungen jeweils eines Jahres, die 36, 18, 12 oder 6 Monate nach Ablauf des Berichtsjahres eingetroffen sind (frühestens im Juli 2015 z.B. kann die 18-Monats-Datei für das Jahr 2013 erstellt werden). Mehr als 3 Jahre verspätete Meldungen werden im IAB nicht berücksichtigt, so dass eine 36-Monats-Datei per definitionem einen Füllgrad von 100% aufweist.
- In der zugrundeliegenden Version der IEB ist 2014 das letzte Jahr mit einem Füllgrad der BeH-Meldungen von 100%<sup>11</sup>. Für die Jahre 2015 und 2016 wurden 18-Monats-Dateien verwendet, für 2017 eine 6-Monats-Datei. Es ist daher davon auszugehen, dass Beschäftigungsmeldungen für die Jahre 2015 bis 2016 leicht und für 2017 etwas stärker untererfasst sind. Die Auswertbarkeit der Daten auf Personenebene sollte dadurch aber nicht eingeschränkt sein. Allerdings häufen sich die fehlenden Meldungen in wenigen Betrieben. Dies bedeutet, dass in Einzelfällen

---

<sup>10</sup> Siehe Paulus/Matthes, 2013, für Details zur Klassifikation der Berufe 2010.

<sup>11</sup> Durch ein Redesign der Datengrundlage der BeH stand zum Zeitpunkt der Erstellung der BeH ausnahmsweise nur das 30-statt des 36-Monatsmaterials zur Verfügung. Analysen mit früheren Datenversionen haben jedoch gezeigt, dass i.d.R. bereits die 18-Monats-Datei einen Füllgrad von rund 99% aufweist. Demnach kommt es in den weiteren 18 Monaten nur noch zu extrem wenigen Nach- oder Korrekturmeldungen.

die Betriebsangaben, z.B. die Betriebsgröße, grob falsch sind und sich in Folgeversionen stark verändern werden.

#### 4.4 Leistungsstatistik nach dem SGB II (LST-S)

Auswertungen auf Träger- oder Kreisebene sind aufgrund von temporären Lieferausfällen teilweise eingeschränkt. Hierzu sind zum einen die beiden Kennzeichen "Vorhergehende Meldungslücke" und "Nachfolgende Meldungslücke" zu berücksichtigen, die eine generelle Untererfassung anzeigen. Daneben beobachtet die Statistik der BA die Aggregatwerte eines meldenden Trägers. Erweisen sich diese themenspezifisch als unplausibel, sind alle Datensätze des Trägers entsprechend gekennzeichnet. In der SIG werden diese Zeiträume auf den Missingwert .p gesetzt und bieten dafür die Voraussetzung für projektspezifische Imputationen.

Die Aussagekraft für regionale und zeitliche Vergleiche der Kosten der Unterkunft ist ab Mitte 2016 deutlich eingeschränkt, da Leistungen für die Unterbringung von Geflüchteten in Sammelunterkünften teilweise für einen ggf. länger zurückreichenden Zeitraum in einer Summe eingefordert werden und vom meldenden Träger nicht immer auf die entsprechenden Anspruchsmonate verteilt, sondern nur einem Monat zugeordnet werden. Das wirkt sich dann auch auf die Summen von Zahlungsansprüchen und Gesamtregelleistungen aus (Bundesagentur für Arbeit, 2017).

Zudem stimmen die Zeiträume aus LHG und LST-S nicht immer vollständig überein. So kann es vorkommen, dass aufgrund von notwendigen Eingriffen in die Daten, bspw. zu Zwecken der Anonymisierung, eine LHG-Episode bereits vor dem dazugehörigen LST-S-Spell auftritt oder erst später endet. Zur näherungsweisen Quantifizierung des Phänomens eine einfache Auszählung: Zu knapp 9 Prozent Beobachtungen (auf Monatsebene) mit Informationen aus der LST-S gibt es in der aktuellen SIG-Version keine LHG-Entsprechung, also keinen LHG-Spell, der mindestens einen Tag des jeweiligen Monats einschließt. Umgekehrt existieren für 3 Prozent der Beobachtungen mit LHG-Informationen (auf Monatsebene) keine LST-S-Entsprechung für den jeweiligen Monat.

Beträge von exakt 0,00 Euro können nicht nur dann vermerkt sein, wenn tatsächlich kein Bedarf vorliegt, sondern auch aus anderen Gründen. Dies ist der Fall, wenn die Art des Bedarfs zum Zeitpunkt der Meldung noch nicht oder nicht mehr gültig ist oder (technische) Restriktionen bzw. mangelnde Datenqualität einer Eintragung im Wege stehen. Fälle mit einem Gesamtbedarf von 0 Euro wurden aus der LST-S-Zuspieldatei zur SIG ausgeschlossen.

Die Gültigkeit von Merkmalen wird zunächst von der gesetzlichen Gültigkeit bestimmt, die Sachverhalte sind rechtlich nur innerhalb bestimmter Zeitgrenzen möglich. Die Daten sind jedoch nicht anhand der gesetzlichen Gültigkeit konsolidiert. Beispielsweise werden auch nach 2010 in Einzelfällen ALG-II-Zuschläge gemeldet. Die Finanzdaten der LST-S können zudem unabhängig von der trägerspezifischen Plausibilität im Einzelfall unterschiedliche Güte haben. Vor allem nach Einführung neuer Leistungsarten ist damit zu rechnen, dass sich die Qualität erst im Lauf der Zeit stabilisiert. Die Merkmale sind in dieser Hinsicht nicht konsolidiert, sondern entsprechen dem Meldefluss. Deren Qualität beeinflusst damit mehr oder weniger die Qualität der in der SIG veröffentlichten Summenwerte.

Die Betragsmerkmale zu "Bildung und Teilhabe" (BuT) haben besondere Qualitätseinschränkungen und sind daher lediglich als Bedarfskennzeichen in der SIG enthalten. Sie werden erst seit Mitte

2015 rückwirkend ab 2013 veröffentlicht. In diesem Zeitraum stieg die Anzahl der meldenden Träger kontinuierlich, aber bei der Erstveröffentlichung gab es immer noch einzelne Träger, die keinerlei BuT-Fälle melden. Für einen regionalen Vergleich sind die Angaben nicht geeignet, da dazu flächendeckend weitere Informationen zu regionalen Angeboten vorliegen müssten (z.B. Fahrkostenzuschüsse durch Bundesländer, kommunale Beteiligung an der Mittagsverpflegung usw.) (weitere Details zu der generellen Qualität siehe Bundesagentur für Arbeit (2015)).

Auszubildende, deren Ausbildung nach dem Bundesausbildungsförderungsgesetz (BAföG) oder im Rahmen der Berufsausbildungsbeihilfe (BAB) förderfähig ist, waren eigentlich bis einschließlich Berichtsmonat August 2016 vom Leistungsbezug nach dem SGB II ausgeschlossen. Sie können aber dennoch unter bestimmten Voraussetzungen Leistungen nach dem SGB II beanspruchen, wenn deren Bedarf nicht anderweitig gedeckt werden kann. Auszubildende können danach Mehrbedarfe, Zuschüsse zu den Kosten der Unterkunft, Darlehen für Regelbedarf und Beiträge für Kranken- und Pflegeversicherung sowie Leistungen für Mietschulden erhalten. Alle Leistungen, die Auszubildenden gewährt werden, gelten nicht als Arbeitslosengeld II. In der statistischen Abbildung wird nicht nach den einzelnen Leistungsarten unterschieden, sondern ein Gesamtwert für Leistungen für Auszubildende dargestellt.

Für vom Leistungsanspruch voll ausgeschlossene Personen sind keine Finanzdaten enthalten. Bei einer Berechnung von Durchschnittsbeträgen pro Bedarfsgemeinschaftsmitglied ist dies entsprechend zu berücksichtigen.

Die Werte der Zahlungsansprüche werden je nach Quellverfahren unterschiedlich ermittelt: während aus ALLEGRO konsolidierte Meldungen von bedarfsspezifischen Zahlungsansprüchen übermittelt werden, wird der Wert bei XSozial-Daten bei der Statistik-Aufbereitung aus Bedarfs-, Einkommens- und Sanktionsmeldungen berechnet. Bei A2LL-Daten handelt es sich um Berechnungen direkt aus dem Meldefluss der Zahlungsanweisungen. Daher sind dort negative Werte möglich.

In der LST-S sind keine imputierten Datensätze bei Lieferausfall verfügbar, sondern ausschließlich die Originalmeldungen. Daher unterscheiden sich auf die Grundgesamtheit hochgerechnete Aggregatdaten von den veröffentlichten Daten der Statistik der BA in Monaten mit Lieferausfällen.

## 5 Variablenbeschreibung

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in den Arbeitshilfen zur SIG unter <https://fdz.iab.de>. Weitere detaillierte Übersichten zu den einzelnen Merkmalen (Label, Auszählungen, fehlende Werte) sind dem Online-Anhang zum Datenreport zu entnehmen.

### 5.1 Identifikatoren

#### 5.1.1 Systemfreie Personennummer (persnr)

Kategorie	Beschreibung
Variablenlabel	Systemfreie Personennummer
Variablenname	persnr
Kategorie	Identifikatoren
Herkunft	BeH, LeH, LHG, ASU, XASU, MTH, XMTH, LST-S
Datentyp	Numerisch
Detailbeschreibung	<p>Die systemfreie Personennummer zeigt an, welche Sätze zur gleichen Person gehören. Systemfrei bedeutet, dass von dieser Personennummer nicht auf Eigenschaften der Person oder Original-Identifikatoren geschlossen werden kann.</p> <p>Da es in den verschiedenen Datenquellen keinen einheitlichen Personen-Identifikator gibt, ist die Zuordnung der Informationen aus den unterschiedlichen Datenquellen (z.B. Beschäftigung und Leistung) zu Personen(-Konten) nicht immer eindeutig. In solchen Fällen kann es zu unplausiblen Erwerbshistorien kommen.</p> <p>Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.</p>

#### 5.1.2 Systemfreie Betriebsnummer (betnr\_sig)

Kategorie	Beschreibung
Variablenlabel	Systemfreie Betriebsnummer
Variablenname	betnr_sig
Kategorie	Identifikatoren
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Die systemfreie Betriebsnummer zeigt an, welche Sätze zum gleichen Betrieb gehören. Um die Daten zu anonymisieren wurde die von der BA vergebene Betriebsnummer durch eine systemfreie, aber eindeutige Nummer ersetzt. (weitere Informationen zur Betriebsnummernvergabe der BA finden sich in Bender et al., (1996, S. 15f. und S. 27-30) sowie direkt bei dem Betriebsnummern-Service der BA unter <a href="https://www.arbeitsagentur.de/betriebsnummern-service/alles-wichtige">https://www.arbeitsagentur.de/betriebsnummern-service/alles-wichtige</a>). Die Betriebsnummer kann zusammen mit der Jahresangabe benutzt werden, um Personen- und Betriebsinformationen zusammenzuspielen.</p> <p>Für die Betriebsnummer sollte Folgendes grundsätzlich beachtet werden:</p> <p>Besteht das Unternehmen aus nur einer Niederlassung oder hat das Unternehmen in einer Gemeinde nur eine Niederlassung, so ist diese Niederlassung der Betrieb und erhält eine Betriebsnummer.</p> <p>Hat das Unternehmen <b>mehrere</b> Niederlassungen in <b>einer</b> Gemeinde, so sind diese Betriebs- und Arbeitsstätten zu einem Betrieb unter einer Betriebsnummer zusammenzufassen, wenn sie derselben Wirtschaftsklasse angehören bzw. ist jede Niederlassung ein Betrieb und erhält je eine Betriebsnummer, wenn sie <b>nicht</b> derselben Wirtschaftsklasse angehören.</p> <p>Hat das Unternehmen <b>mehrere</b> Niederlassungen in <b>mehreren</b> Gemeinden, so ist jede dieser Niederlassungen ein Betrieb und erhält je eine Betriebsnummer.</p> <p>In diesem Zusammenhang sind auch folgende Definitionen im Sinne der Betriebsnummernvergabe im Rahmen des Meldeverfahrens zur Sozialversicherung zu beachten:</p>

Kategorie	Beschreibung
	<p>Ein <b>Betrieb</b> ist eine regional und wirtschaftsfachlich abgegrenzte Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer zu vergeben wird.</p> <p>Eine <b>Arbeitsstätte</b> ist eine Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer nicht zu vergeben ist.</p> <p>Ein <b>Unternehmen</b> als Begriff fasst Betriebs- und Arbeitsstätten desselben Arbeitgebers zusammen.</p> <p>Ein <b>Arbeitgeber</b> ist jede natürliche oder juristische Person, für die mindestens ein sozialversicherungspflichtig bzw. geringfügig Beschäftigter tätig ist.</p> <p>Betrieb und Betriebsstätten sind synonyme Begriffe; Niederlassung ist dann ein synonyme Begriff für Filialen, Zweigstellen, Auslagerungen, Dependancen, Arbeitsstätten u.ä., wenn sie kein Betrieb ist.</p>
Qualitätshinweis	Die Betriebsnummer fehlt in nur sehr wenigen BeH-Meldungen. Bei diesen Meldungen handelt es sich in den meisten Fällen um Meldungen der Personengruppe „205“ (Entgeltmeldungen für unständig Beschäftigte). Da Betriebsmerkmale (Arbeitsort, Wirtschaftszweig usw.) über die Betriebsnummer zugespielt werden, fehlen sie in diesen Meldungen.

### 5.1.3 Systemfreie Bedarfsgemeinschaftsnummer (bg\_id)

Kategorie	Beschreibung
Variablenlabel	Systemfreie Bedarfsgemeinschaftsnummer
Variablenname	bg_id
Kategorie	Identifikatoren
Herkunft	LHG, LST-S
Datentyp	Numerisch
Detailbeschreibung	Die systemfreie Bedarfsgemeinschaftsnummer zeigt an, welche Sätze zur gleichen Bedarfsgemeinschaft gehören. Systemfrei bedeutet, dass von dieser Bedarfsgemeinschaftsnummer nicht auf Eigenschaften der Personen/BG oder Original-Identifikatoren geschlossen werden kann.
Qualitätshinweis	Die Analyse von Bedarfsgemeinschaften zu einem Zeitpunkt ist mit der Bedarfsgemeinschaftsnummer weitgehend unproblematisch. Zu- oder Abgänge von Bedarfsgemeinschaften im Zeitverlauf oder BG-Wechsel von Personen sind nicht zuverlässig identifizierbar.

## 5.2 Generierte technische Merkmale

### 5.2.1 Quelle (quelle\_lhg, quelle\_asu, quelle\_xasu, quelle\_leh, quelle\_mth, quelle\_xmth, quelle\_beh)

Kategorie	Beschreibung
Variablenlabel	Quelle
Variablenname	quelle_lhg, quelle_asu, quelle_xasu, quelle_beh, quelle_leh, quelle_mth, quelle_xmth
Kategorie	Generierte technische Merkmale
Herkunft	LHG, ASU, XASU, LeH, MTH, XMTH, BeH
Datentyp	Numerisch
Detailbeschreibung	Die Merkmale geben die Datenquellen an, aus denen Informationen in dem Spell enthalten sind.

### 5.2.2 Vorhergehende Meldungslücke (lck\_links)

Kategorie	Beschreibung
Variablenlabel	Vorhergehende Meldungslücke
Variablenname	lck_links
Kategorie	Generierte technische Merkmale
Herkunft	LHG; SIG-Berechnung

Kategorie	Beschreibung
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal gibt an, ob der meldende Grundsicherungsträger vor dem Beginndatum des Spells Datenübermittlungsprobleme hatte. Dadurch entstand eine Meldungslücke im Datenbestand. Das Beginndatum kann daher falsch sein.

### 5.2.3 Anschließende Meldungslücke (lck\_rechts)

Kategorie	Beschreibung
Variablenlabel	Anschließende Meldungslücke
Variablenname	lck_rechts
Kategorie	Generierte technische Merkmale
Herkunft	LHG; SIG-Berechnung
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal gibt an, ob der meldende Grundsicherungsträger nach dem Enddatum des Spells Datenübermittlungsprobleme hatte. Dadurch entstand eine Meldungslücke im Datenbestand. Das Enddatum kann daher falsch sein.

### 5.2.4 Spell linksenziert 1.1.2007 (zensierung\_links\_lhg, zensierung\_links\_x\_asu, zensierung\_links\_leh, zensierung\_links\_x\_mth, zensierung\_links\_beh)

Kategorie	Beschreibung
Variablenlabel	Spell linksenziert 1.1.2007
Variablenname	zensierung_links_lhg, zensierung_links_x_asu, zensierung_links_leh, zensierung_links_x_mth, zensierung_links_beh
Kategorie	Generierte technische Merkmale
Herkunft	SIG-Berechnung
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal gibt an, ob ein Spell zum 1.1.2007 linksenziert wurde (nach Quelle).

### 5.2.5 Spelllänge (spell\_length)

Kategorie	Beschreibung
Variablenlabel	Spelllänge
Variablenname	spell_length
Kategorie	Generierte technische Merkmale
Herkunft	LST-S
Datentyp	Numerisch
Detailbeschreibung	Die Variable gibt die Länge eines jeden LST-S-Spells in Monaten an. Sie kann dazu verwendet werden die ursprüngliche Datenstruktur wiederherzustellen (wie in Kapitel 2.6 vorgeschlagen). Im Zuge dessen werden alle Episoden, die aufgrund identischer Angaben zusammengefasst wurden, wieder in Ein-Monats-Episoden getrennt.

## 5.3 Gültigkeitszeitraum

### 5.3.1 Beginndatum Episode (begepi)

Kategorie	Beschreibung
Variablenlabel	Beginndatum Episode

Kategorie	Beschreibung
Variablenname	begepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	SIG-Berechnung
Datentyp	Datum
Detailbeschreibung	Das Beginndatum der Episode ist immer gleich dem oder größer als das Beginndatum der Originalbeobachtung (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.3).

### 5.3.2 Enddatum Episode (endepi)

Kategorie	Beschreibung
Variablenlabel	Enddatum Episode
Variablenname	endepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	SIG-Berechnung
Datentyp	Datum
Detailbeschreibung	Das Enddatum der Episode ist immer gleich dem oder kleiner als das Enddatum der Originalbeobachtung (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.3).

### 5.3.3 Erster Monat Episode (beg\_mon)

Kategorie	Beschreibung
Variablenlabel	Erster Monat Episode
Variablenname	beg_mon
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	LST-S; SIG-Berechnung
Datentyp	Datum (Monat)
Detailbeschreibung	Das Merkmal bezeichnet den Beginn einer Episode aus der LST-S gemäß der zusammenfassenden SIG-eigenen Berechnung bei Monaten mit unveränderter Leistungsstatistik auf Bedarfsgemeinschaftsebene.

### 5.3.4 Letzter Monat Episode (end\_mon)

Kategorie	Beschreibung
Variablenlabel	Letzter Monat Episode
Variablenname	end_mon
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	LST-S; SIG-Berechnung
Datentyp	Datum (Monat)
Detailbeschreibung	Das Merkmal bezeichnet das Ende einer Episode aus der LST-S gemäß der zusammenfassenden SIG-eigenen Berechnung bei Monaten mit unveränderter Leistungsstatistik auf Bedarfsgemeinschaftsebene

## 5.4 Informationen zur Person

### 5.4.1 Geschlecht (frau)

Kategorie	Beschreibung
Variablenlabel	Geschlecht
Variablenname	frau
Kategorie	Informationen zur Person
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	Geschlechter-Dummy (0 - Mann, 1 - Frau). Die Geschlechtsangabe ist innerhalb eines Personenkontos konstant.
Qualitätshinweis	Die Angabe zum Geschlecht wurde aus den Angaben zum Leistungsbezug (LHG) auf alle Spells einer Person überschrieben.

### 5.4.2 Geburtsjahr (gebjahr)

Kategorie	Beschreibung
Variablenlabel	Geburtsjahr
Variablenname	gebjahr
Kategorie	Informationen zur Person
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	Das Geburtsjahr ist innerhalb eines Personenkontos konstant.
Qualitätshinweis	Die Angabe zum Geburtsjahr wurde aus den Angaben zum Leistungsbezug (LHG) auf alle Spells einer Person überschrieben. Waren in der LHG unterschiedliche Angaben für eine Person vorhanden, wurde die jüngste (aktuellste) Angabe verwendet.

### 5.4.3 Geburtsmonat (gebmon)

Kategorie	Beschreibung
Variablenlabel	Geburtsmonat
Variablenname	gebmon
Kategorie	Informationen zur Person
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	Der Geburtsmonat ist innerhalb eines Personenkontos konstant. Um aus den Angaben zum Geburtsjahr (gebjahr) und -monat (gebmon) eine Variable im Datumsformat JJJJmM (z.B. 1984m6) zu generieren, können Sie folgenden Befehl verwenden: <pre>gen int gebdat = ym(gebjahr, gebmon) format gebdat %tm</pre>
Qualitätshinweis	Die Angabe zum Geburtsmonat wurde aus den Angaben zum Leistungsbezug (LHG) auf alle Spells einer Person überschrieben. Waren in der LHG unterschiedliche Angaben für eine Person vorhanden, wurde die jüngste (aktuellste) Angabe verwendet.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt. Ansonsten wird lediglich das Geburtsjahr (gebjahr) bereitgestellt.

#### 5.4.4 Staatsangehörigkeit (nation\_lhg, nation\_x\_asu, nation\_leh, nation\_x\_mth, nation\_beh)

Kategorie	Beschreibung
Variablenlabel	Staatsangehörigkeit
Variablenname	nation_lhg, nation_x_asu, nation_leh, nation_x_mth, nation_beh
Kategorie	Informationen zur Person
Herkunft	LHG, ASU, XASU, LeH, MTH, XMTH, BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die Staatenschlüssel des Statistischen Bundesamtes (Statistisches Bundesamt, 2019).
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Staatsangehörigkeitsmerkmal lediglich die Staatsangehörigkeit vergrößert (nation_gr) bereitgestellt.

#### 5.4.5 Staatsangehörigkeit vergrößert (nation\_gr\_lhg, nation\_gr\_x\_asu, nation\_gr\_leh, nation\_gr\_x\_mth, nation\_gr\_beh)

Kategorie	Beschreibung
Variablenlabel	Staatsangehörigkeit vergrößert
Variablenname	nation_gr_lhg, nation_gr_x_asu, nation_gr_leh, nation_gr_x_mth, nation_gr_beh
Kategorie	Informationen zur Person
Herkunft	LHG, ASU, XASU, LeH, MTH, XMTH, BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die zu Gruppen zusammengefassten Staatenschlüssel des Statistischen Bundesamtes.

#### 5.4.6 Familienstand (famst\_lhg, famst\_x\_asu, famst\_leh)

Kategorie	Beschreibung
Variablenlabel	Familienstand
Variablenname	famst_lhg, famst_x_asu, famst_leh
Kategorie	Informationen zur Person
Herkunft	LHG, ASU, XASU, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal beschreibt den Familienstand. Diese Variable ist für Beschäftigungsmeldungen nicht gefüllt. <b>1) LeH</b> In der Quelle LeH umfasst das Merkmal nur zwei Ausprägungen (0 - nicht verheiratet, 1 - verheiratet) <b>2) LHG, ASU, XASU</b> In den Quellen LHG/ASU/XASU wird zwischen sechs Ausprägungen (Werte 11-16) unterschieden. Ein Abgleich der Informationen aus den Quellen wurde nicht vorgenommen.
Qualitätshinweis	<b>1) LeH</b> Ab Beginnjahr 2014 ist die Qualität als fraglich einzustufen. So sinkt der Anteil der Verheirateten z.B. zwischen 2013 und 2017 erheblich. <b>2) LHG, XASU, XMTH</b> Aufgrund abweichender Meldestandards in XSozial sind die Angaben aus den zKT bis Dezember 2009 als weniger valide einzustufen.

#### 5.4.7 Alter des jüngsten Kindes - kategorisiert (alter\_juk\_kat)

Kategorie	Beschreibung								
Variablenlabel	Alter des jüngsten Kindes - kategorisiert								
Variablenname	alter_juk_kat								
Kategorie	Informationen zur Person								
Herkunft	LHG								
Datentyp	Numerisch								
Detailbeschreibung	<p>Das Merkmal enthält das Alter des jüngsten Kindes (unter 15 Jahren) in folgenden Kategorien:</p> <table> <tr> <td>0</td> <td>kein Kind unter 15 Jahre</td> </tr> <tr> <td>1</td> <td>0 bis 2 Jahre</td> </tr> <tr> <td>2</td> <td>3 bis 6 Jahre</td> </tr> <tr> <td>3</td> <td>7 bis 14 Jahre</td> </tr> </table>	0	kein Kind unter 15 Jahre	1	0 bis 2 Jahre	2	3 bis 6 Jahre	3	7 bis 14 Jahre
0	kein Kind unter 15 Jahre								
1	0 bis 2 Jahre								
2	3 bis 6 Jahre								
3	7 bis 14 Jahre								

#### 5.4.8 Ausbildung (ausbildung\_x\_asu, ausbildung\_beh)

Kategorie	Beschreibung																																						
Variablenlabel	Ausbildung																																						
Variablenname	ausbildung_x_asu, ausbildung_beh																																						
Kategorie	Informationen zur Person																																						
Herkunft	ASU, XASU, BeH																																						
Datentyp	Numerisch																																						
Detailbeschreibung	<p>Das Merkmal enthält den beruflichen Ausbildungsabschluss. Eine Einordnung der Abschlüsse in die Internationale Standardklassifikation des Bildungswesens (ISCED 2011) ist den zentralen Arbeitshilfen des FDZ zu entnehmen. Zu beachten ist, dass das Merkmal je nach Quelle eine unterschiedliche Bedeutung hat.</p> <p><b>1) BeH</b>  Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Berufsausbildung. Folgende Ausprägungen kommen vor:</p> <table> <tr> <td>1</td> <td>ohne abgeschlossene Berufsausbildung</td> </tr> <tr> <td>2</td> <td>betriebliche/außerbetriebliche Ausbildung</td> </tr> <tr> <td>11</td> <td>FH ohne nähere Spezifikation</td> </tr> <tr> <td>12</td> <td>Hochschulabschluss ohne nähere Spezifikation</td> </tr> </table> <p>Bei Meldungen nach dem neuen Tätigkeitsschlüssel (siehe Abschnitt 4.3.4) ist es nicht mehr möglich, FH-Absolventen wie bisher eindeutig zu identifizieren, da der neue Tätigkeitsschlüssel die separate Erfassung dieses Bildungsabschlusses nicht länger erlaubt. Sie werden der Kategorie 12 zugeordnet.</p> <p><b>2) ASU</b>  Für diese Sätze ist die letzte abgeschlossene Berufsausbildung enthalten. Folgende Ausprägungen kommen bei Spells mit Beginndatum bis zum 30.06.2006 vor:</p> <table> <tr> <td>1</td> <td>ohne abgeschlossene Berufsausbildung</td> </tr> <tr> <td>2</td> <td>betriebliche/außerbetriebliche Ausbildung</td> </tr> <tr> <td>3</td> <td>Berufsfachschule</td> </tr> <tr> <td>4</td> <td>Fachschule</td> </tr> <tr> <td>5</td> <td>Fachhochschule</td> </tr> <tr> <td>6</td> <td>Universität</td> </tr> </table> <p>Im Jahr 2006 wurde das EDV-Verfahren, aus dem die Arbeitsuchendendaten stammen, von coArb auf VerBIS umgestellt. Viele Merkmale, wie auch die Ausbildung, wurden in beiden Systemen in unterschiedlicher Differenziertheit abgefragt. Das bedeutet, dass in den Quellen ASU und MTH bei Spells mit Beginndatum ab dem 01.07.2006 folgende differenzierteren Aufprägungen vorliegen:</p> <table> <tr> <td>7</td> <td>in Deutschland nicht anerkannte Berufsausbildung</td> </tr> <tr> <td>8</td> <td>in Deutschland nicht anerkannter Hochschulabschluss</td> </tr> <tr> <td>9</td> <td>betriebliche/schulische Ausbildung</td> </tr> <tr> <td>10</td> <td>sonstige Examen</td> </tr> <tr> <td>11</td> <td>FH ohne nähere Spezifikation</td> </tr> <tr> <td>12</td> <td>Hochschulabschluss ohne nähere Spezifikation</td> </tr> <tr> <td>13</td> <td>Promotion</td> </tr> <tr> <td>14</td> <td>Bachelor (BA)</td> </tr> <tr> <td>15</td> <td>Bachelor (FH)</td> </tr> </table>	1	ohne abgeschlossene Berufsausbildung	2	betriebliche/außerbetriebliche Ausbildung	11	FH ohne nähere Spezifikation	12	Hochschulabschluss ohne nähere Spezifikation	1	ohne abgeschlossene Berufsausbildung	2	betriebliche/außerbetriebliche Ausbildung	3	Berufsfachschule	4	Fachschule	5	Fachhochschule	6	Universität	7	in Deutschland nicht anerkannte Berufsausbildung	8	in Deutschland nicht anerkannter Hochschulabschluss	9	betriebliche/schulische Ausbildung	10	sonstige Examen	11	FH ohne nähere Spezifikation	12	Hochschulabschluss ohne nähere Spezifikation	13	Promotion	14	Bachelor (BA)	15	Bachelor (FH)
1	ohne abgeschlossene Berufsausbildung																																						
2	betriebliche/außerbetriebliche Ausbildung																																						
11	FH ohne nähere Spezifikation																																						
12	Hochschulabschluss ohne nähere Spezifikation																																						
1	ohne abgeschlossene Berufsausbildung																																						
2	betriebliche/außerbetriebliche Ausbildung																																						
3	Berufsfachschule																																						
4	Fachschule																																						
5	Fachhochschule																																						
6	Universität																																						
7	in Deutschland nicht anerkannte Berufsausbildung																																						
8	in Deutschland nicht anerkannter Hochschulabschluss																																						
9	betriebliche/schulische Ausbildung																																						
10	sonstige Examen																																						
11	FH ohne nähere Spezifikation																																						
12	Hochschulabschluss ohne nähere Spezifikation																																						
13	Promotion																																						
14	Bachelor (BA)																																						
15	Bachelor (FH)																																						

Kategorie	Beschreibung																					
	<p>16 Bachelor (Uni)  17 Master (FH)  18 Master (Uni)  19 Diplom (BA)  20 Diplom (FH)  21 Diplom (Uni)  22 ohne anerkannte abgeschlossene Berufsausbildung  23 grundständige Studienfächer /-gänge  24 weiterführende Studienfächer /-gänge  25 sonstige (Aufstiegs-)Weiterbildungen für Hochschulabsolventen  26 Duales Studium, ausbildungsintegrierend  27 Duales Studium, praxisintegrierend</p> <p><b>3)XASU</b>  Für diese Sätze ist die letzte abgeschlossene Berufsausbildung enthalten. Folgende Ausprägungen kommen vor:</p> <p>1 ohne abgeschlossene Berufsausbildung  2 betriebliche/außerbetriebliche Ausbildung  3 Berufsfachschule  4 Fachschule  5 Fachhochschule  6 Universität  7 in Deutschland nicht anerkannte Berufsausbildung  8 in Deutschland nicht anerkannter Hochschulabschluss</p> <p>Für Auswertungen über einen längeren Zeitraum könnten die Ausprägungen anhand folgender Tabelle aggregiert werden:</p> <table> <tr> <td>1, 22</td> <td>auf</td> <td>1</td> </tr> <tr> <td>7</td> <td>auf</td> <td>2</td> </tr> <tr> <td>8</td> <td>auf</td> <td>3</td> </tr> <tr> <td>2, 3, 4, 9</td> <td>auf</td> <td>4</td> </tr> <tr> <td>5, 11, 14, 15, 17, 19, 20</td> <td>auf</td> <td>5</td> </tr> <tr> <td>6, 10, 12, 13, 16, 18, 21, 26, 27</td> <td>auf</td> <td>6</td> </tr> <tr> <td>23, 24, 25</td> <td>auf</td> <td>7</td> </tr> </table> <p>Die aggregierten Werte haben folgende Labels:</p> <p>1 ohne (anerkannte) abgeschlossene Berufsausbildung  2 in Deutschland nicht anerkannte Berufsausbildung  3 in Deutschland nicht anerkannter Hochschulabschluss  4 betriebliche / schulische Ausbildung  5 Fachhochschule  6 Universität  6 sonstige Examen  7 Grundständige Studienfächer/-gänge  7 Weiterführende Studienfächer/-gänge  7 Sonstige (Aufstiegs-)Weiterbildungen für Hochschulabsolventen</p>	1, 22	auf	1	7	auf	2	8	auf	3	2, 3, 4, 9	auf	4	5, 11, 14, 15, 17, 19, 20	auf	5	6, 10, 12, 13, 16, 18, 21, 26, 27	auf	6	23, 24, 25	auf	7
1, 22	auf	1																				
7	auf	2																				
8	auf	3																				
2, 3, 4, 9	auf	4																				
5, 11, 14, 15, 17, 19, 20	auf	5																				
6, 10, 12, 13, 16, 18, 21, 26, 27	auf	6																				
23, 24, 25	auf	7																				
Qualitätshinweis	<p><b>1) BeH</b>  „Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammengestellt werden. Wenn ein Beschäftigter z.B. durch berufs begleitende Weiterbildung einen höheren Abschluss erlangt hat, dann wird dieser Statuswechsel wahrscheinlich erst von einem neuen Beschäftigungsbetrieb erfasst. Es ist in der Regel davon auszugehen, dass bei längeren Beschäftigungszeiten innerhalb eines Betriebs die einmal erhobenen personenbezogenen Daten fortgeschrieben werden.“ (Meinken/Koch, 2004, S. 63)  Der Anteil der fehlenden Werte steigt in der BeH im Zeitverlauf fast kontinuierlich an. Durch die Einführung des neuen Tätigkeitsschlüssels gab es in 2011 einen stark erhöhten Anteil von rund 51%. Seit 2014 pendelt sich Anteil der fehlenden Werte in der BeH bei rund 40% ein. Fehlende Werte treten besonders häufig bei folgenden Personengruppen auf: geringfügig Beschäftigte, Personen in Teilzeit, ausländische Beschäftigte und ostdeutsche Arbeitnehmer. Hintergrund ist, dass das Merkmal versicherungsrechtlich keine besondere Bedeutung hat (siehe Meinken/Koch, 2004, S. 63).  Mit der Umstellung des Tätigkeitsschlüssels ist ein Bruch in den Daten verbunden, der über die reine Umstellung des Schlüssels hinausgeht, und evtl. durch Aktualisierungseffekte bei der Umstellung der Lohnbuchhaltungssoftware bei den meldenden Betrieben zu erklären ist.  Für das Merkmal Ausbildung (imputiert) ((ausbildung_imp_beh), siehe Abschnitt 5.4.9) wurde eine Methode angewendet, um fehlende Werte oder inkonsistente zeitliche Verläufe der Bildungsangaben zu bereinigen. Diese Variable ist jedoch nur in der Quelle BeH gefüllt und hat andere Ausprägungen als die Variable ausbildung.</p> <p><b>2) ASU</b></p>																					

Kategorie	Beschreibung
	<p>Durch die Umstellung von coArb auf VerBIS kann in der Quelle ASU zwischen 2006 und 2008 nicht korrekt zwischen „ohne abgeschlossene Berufsausbildung“ und „Keine Angabe“ unterschieden werden. Ein fehlender Wert in diesem Zeitraum besagt also nicht unbedingt, dass keine (Information zur) Ausbildung vorliegt, sondern kann auch bedeuten, dass das einschlägige Generierungsverfahren nicht angewandt werden konnte. In der Quelle MTH sind auch nach dem 01.07.2006 zum Teil noch die älteren Aufprägungen zu finden.</p> <p><b>3) XASU</b> Der Füllgrad in der Quelle XASU ist generell niedrig.</p>

#### 5.4.9 Ausbildung (imputiert) (ausbildung\_imp\_beh)

Kategorie	Beschreibung
Variablenlabel	Ausbildung (imputiert)
Variablenname	ausbildung_imp_beh
Kategorie	Persönliche Merkmale
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal Ausbildung (imputiert) ist ein Zusatzmerkmal zum Merkmal Ausbildung (imputiert) (ausbildung_imp_beh) und enthält für BeH-Meldungen ergänzte und harmonisierte Angaben zur Berufsausbildung von Beschäftigten. Das Merkmal bietet damit einen Lösungsansatz für die in Abschnitt 5.4.8 beschriebenen Probleme der Variable Ausbildung (ausbildung_beh) bei BeH-Meldungen. Das Imputationsverfahren wird in Thomsen et al. (2018) beschrieben, welches auf der Arbeit von Fitzenberger et al. (2006) basiert.</p> <p>Da das Merkmal nur Bildungsangaben der BeH-Meldungen nutzt, und weil die Bildungskategorien des alten und des neuen Tätigkeitsschlüssels für das Merkmal Ausbildung (ausbildung_beh) harmonisiert werden mussten, bietet das Merkmal Ausbildung (imputiert) (ausbildung_imp_beh) andere Kategorien als das Merkmal Ausbildung (ausbildung_beh).</p> <p>Eine Einordnung der Abschlüsse in die Internationale Standardklassifikation des Bildungswesens (ISCED 2011) ist den zentralen Arbeitshilfen des FDZ (<a href="https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/">https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/</a>) zu entnehmen.</p>
Qualitätshinweis	Trotz des Imputationsverfahrens nimmt der Anteil der fehlenden Angaben im Zeitverlauf zu und erreichen 2023 ca. 11%. Mit der Umstellung des Tätigkeitsschlüssels ist ein Bruch in den Daten verbunden, der über die reine Umstellung des Schlüssels hinausgeht, und evtl. durch Aktualisierungseffekte bei der Umstellung der Lohnbuchhaltungssoftware bei den meldenden Betrieben zu erklären ist.

#### 5.4.10 Schulausbildung (schule\_x\_asu, schule\_beh)

Kategorie	Beschreibung
Variablenlabel	Schulausbildung
Variablenname	schule_x_asu, schule_beh
Kategorie	Informationen zur Person
Herkunft	ASU, XASU, BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal enthält den Schulabschluss. Eine Einordnung der Abschlüsse in die Internationale Standardklassifikation des Bildungswesens (ISCED 2011) ist den zentralen Arbeitshilfen des FDZ zu entnehmen. Je nach Quelle sind unterschiedliche Ausprägungen möglich.</p> <p><b>1) BeH</b> Mit der Umstellung des Tätigkeitsschlüssels ändern sich die möglichen Ausprägungen (siehe Abschnitt 4.3.4). Davor gelten:</p> <p>5 Volks-/Hauptschule o./m. Abschluss, Mittlere Reife oder gleichwertige Schulbildung 8 Fachhochschulreife oder Abitur/Hochschulreife 9 Abitur/Hochschulreife</p> <p>Danach gelten:</p> <p>1 Kein Hauptschulabschluss 4 Hauptschulabschluss/Volksschulabschluss 6 Mittlere Reife</p>

Kategorie	Beschreibung
	<p>8 Fachhochschulreife oder Abitur/Hochschulreife</p> <p>Die Unterscheidungsmöglichkeiten im Sekundarbereich I wurden damit erweitert (1, 4, 6 statt vorher 5), die im Sekundarbereich II dagegen verringert (8 und 9 zusammengefasst unter 8).</p> <p><b>2) ASU, XASU</b></p> <p>Für diese Quellen sind die folgenden Ausprägungen möglich:</p> <p>1 Kein Hauptschulabschluss  4 Hauptschulabschluss/Volksschulabschluss  6 Mittlere Reife  7 Fachhochschulreife  9 Abitur/Hochschulreife</p> <p>Sie sind zu Beginn der Arbeitssuche/Maßnahmeteilnahme gültig. In der Quelle XASU kann das Merkmal bei Ausbildungsplatzsuchenden auch den angestrebten Abschluss enthalten.</p>
Qualitätshinweis	<p><b>1) BeH</b></p> <p>Der Füllgrad nimmt in der BeH über die Jahre kontinuierlich ab und liegt in jüngeren Jahren unter 2/3. Mit der Umstellung des Tätigkeitsschlüssels ist ein Bruch in den Daten verbunden, der über die reine Umstellung des Schlüssels hinausgeht, und evtl. durch Aktualisierungseffekte bei der Umstellung der Lohnbuchhaltungssoftware bei den meldenden Betrieben zu erklären ist.</p> <p><b>2) ASU, XASU</b></p> <p>In der XASU nimmt der Füllgrad kontinuierlich zu und liegt seit 2018 über 75%. In ASU ist der Füllgrad insgesamt gut.</p>

#### 5.4.11 Max. Ausbildung bis 2006 (ausbildung\_bis2006)

Kategorie	Beschreibung
Variablenlabel	Max. Ausbildung bis 2006
Variablenname	ausbildung_bis2006
Kategorie	Informationen zur Person
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt das maximale Ausbildungsniveau einer Person bis 2006 an.

#### 5.4.12 Max. Schulausbildung bis 2006 (schule\_bis2006)

Kategorie	Beschreibung
Variablenlabel	Max. Schulausbildung bis 2006
Variablenname	schule_bis2006
Kategorie	Informationen zur Person
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt das maximale Schulausbildungsniveau einer Person bis 2006 an.

### 5.5 Informationen zu Leistungsbezug

#### 5.5.1 Täglicher Leistungssatz Arbeitslosengeld I (leistung\_tag)

Kategorie	Beschreibung
Variablenlabel	Täglicher Leistungssatz
Variablenname	leistung_tag
Kategorie	Informationen zu Leistungsbezug
Herkunft	LeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	Das Merkmal enthält den täglichen Leistungssatz, jeweils umgerechnet in Euro. Es ist zu berücksichtigen, dass der tägliche Leistungssatz für Kalendertage gilt. Die Angabe von 0 Euro ist seit 1.1.2005 auf Sperrzeiten oder Leistungsunterbrechungen zurückzuführen. Ist für einen Satz mit Leistungssatz gleich 0 ein Beendigungsgrund angegeben, so handelt es sich um eine Unterbrechungsmeldung. Bei Sätzen, die eine Sperrzeit widerspiegeln, ist der Anspruch derselbe wie vor Beginn der Sperrzeit.

### 5.5.2 Rolle in der Bedarfsgemeinschaft (bg\_rolle)

Kategorie	Beschreibung
Variablenlabel	Rolle in der Bedarfsgemeinschaft
Variablenname	bg_rolle
Kategorie	Informationen zu Leistungsbezug
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal beschreibt, in welcher Beziehung die Personen in einer Bedarfsgemeinschaft zueinander stehen (Partner/in, Kinder). Es sind folgende Ausprägungen möglich: 1        Hauptperson / Partner/in 2        Minderjähriges unverheiratetes Kind 3        Volljähriges, unverheiratetes Kind unter 25 Jahren

### 5.5.3 Erwerbsfähigkeit (erwf)

Kategorie	Beschreibung
Variablenlabel	Erwerbsfähigkeit
Variablenname	erwf
Kategorie	Informationen zu Leistungsbezug
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal wurde aus den Angaben zum SGB-II-Status berechnet. Es sind folgende Ausprägungen möglich: 0        Nicht erwerbsfähige Person 1        Erwerbsfähige Person
Qualitätshinweise	Im SGB-II-Status wird nicht berücksichtigt, ob ein Ausschlussgrund vorliegt. Zur Identifikation von ALG-II- bzw. Sozialgeld-Leistungsberechtigten ist daher zusätzlich das Merkmal SGB II Ausschlussgrund (sgbii_ausschl) zu berücksichtigen. Zudem zeigt der SGB-II-Status der LHG nicht an, ob ein individueller Leistungsanspruch besteht. Minderjährige Kinder ohne Leistungsanspruch sind in diesem Merkmal ebenfalls nicht zu identifizieren. Unplausibel überdurchschnittlich viele Datensätze zeigen 15- bis 17-jährige als nicht erwerbsfähig an.

### 5.5.4 Bedarfsgemeinschaftstyp (bg\_typ)

Kategorie	Beschreibung
Variablenlabel	Bedarfsgemeinschaftstyp
Variablenname	bg_typ
Kategorie	Informationen zu Leistungsbezug
Herkunft	LHG
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>Dieses Merkmal enthält den berechneten BG-Typ des BG-Kunden nach Definition des IAB. Es sind folgende Ausprägungen möglich:</p> <ol style="list-style-type: none"> <li>1 Alleinstehend</li> <li>2 Alleinstehend mit Volljährigem/n unter 25 Jahre</li> <li>3 Alleinerziehend mit Kind(ern) unter 18 Jahre</li> <li>4 Alleinerziehend mit Kind(ern) unter 18 Jahre und Volljährigem/n unter 25 Jahre</li> <li>5 Paar ohne Kind</li> <li>6 Paar ohne Kind und mit Volljährigem/n unter 25 Jahre</li> <li>7 Paar mit Kind(ern) unter 18 Jahre</li> <li>8 Paar mit Kind(ern) unter 18 Jahre und Volljährigem/n unter 25 Jahre</li> <li>9 Sonstige</li> </ol>

### 5.5.5 Beendigungsgrund (grund\_end)

Kategorie	Beschreibung
Variablenlabel	Beendigungsgrund
Variablenname	grund_end
Kategorie	Informationen zu Leistungsbezug
Herkunft	LeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal gibt den Grund für das Ende des Bezugs von Arbeitslosengeld an. Über die Gründe für den Beginn des Leistungsbezugs finden sich in der LeH keine Informationen, da die LeH mit den Meldungen der Arbeitsagenturen an die Krankenkassen über abgeschlossene Leistungsbezugsdauern gefüllt wird. Die Ausprägungen dieses Merkmals sind nur für den fachlich bzw. gesetzlich vorgesehenen Gültigkeitszeitraum gefüllt. Kommen Ausprägungen außerhalb ihres Gültigkeitszeitraums vor, werden sie auf Ersatzwerte umgeschlüsselt {Wertebereich 1173-1181}, weil ihre Bedeutung dann nicht sicher geklärt werden kann.</p>

### 5.5.6 SGB II Ausschlussgrund (sgbii\_ausschl)

Kategorie	Beschreibung
Variablenlabel	SGB II Ausschlussgrund
Variablenname	sgbii_ausschl
Kategorie	Informationen zu Leistungsbezug
Herkunft	LHG
Datentyp	Numerisch
Detailbeschreibung	<p>Der SGB-II-Ausschlussgrund gibt an, aus welchem personenbezogenen Grund ein BG-Mitglied vom Leistungsbezug ganz oder teilweise ausgeschlossen wird.</p>
Besondere Hinweis	<p>Für vom Leistungsanspruch voll ausgeschlossenen Personen (Werte 1-10) sind keine Finanzdaten enthalten. Bei einer Berechnung von Durchschnittsbeträgen pro Bedarfsgemeinschaftsmitglied ist dies entsprechend zu berücksichtigen.)</p>

### 5.5.7 Zu berücksichtigendes Einkommen aus Beschäftigung (zbeink\_besch)

Kategorie	Beschreibung
Variablenlabel	Zu berücksichtigendes Einkommen aus Beschäftigung
Variablenname	zbeink_besch
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal gibt die Summe des zu berücksichtigenden Einkommens einer Person aus Beschäftigung an.</p>

Kategorie	Beschreibung
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Einkommensangaben verfügbar, d.h. das Gesamteinkommen der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.8 Zu berücksichtigendes Einkommen aus Selbständigkeit (zbeink\_selbst)

Kategorie	Beschreibung
Variablenlabel	Zu berücksichtigendes Einkommen aus Selbständigkeit
Variablenname	zbeink_selbst
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Summe des zu berücksichtigenden Einkommens einer Person aus Selbständigkeit an.
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Einkommensangaben verfügbar, d.h. das Gesamteinkommen der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.9 Verfügbares Einkommen gesamt (veink\_sum)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen gesamt
Variablenname	veink_sum
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal gibt die Summe des verfügbaren Einkommens einer Person an. Dabei gilt – ebenso wie für die nachfolgenden Merkmale zu verfügbarem Einkommen:</p> <p>Summe aller Einkommen einer Person</p> <ul style="list-style-type: none"> <li>- privilegiertes Einkommen</li> <li>= zu berücksichtigendes Einkommen (Bruttoeinkommen bzw. Einnahmen bei Selbständigen)</li> <li>- Steuern und Sozialversicherungsbeiträge (bzw. Betriebsausgaben bei Selbständigen)</li> <li>= verfügbares Einkommen (Nettoeinkommen bzw. Gewinn bei Selbständigen)</li> <li>- gesetzliche Freibeträge</li> <li>= anrechenbares Einkommen</li> </ul> <p>Die anrechenbaren Einkommen aller Mitglieder einer Bedarfsgemeinschaft werden aufsummiert und anschließend gemäß der Bedarfsanteilmethode wieder auf die Mitglieder verteilt. Somit ergibt sich das angerechnete Einkommen einer Person. Dieses angerechnete Einkommen reduziert in der Folge die Bedarfe für Leistungen zur Sicherung des Lebensunterhalts für die Leistungsberechtigten. Für die SIG werden die Summen dann wieder über alle Mitglieder aufsummiert.</p> <p>Für das SIG-Datenprodukt wird nur das tatsächlich verfügbare Einkommen berücksichtigt, da dieses für die Minderung der Bedürftigkeit entscheidend ist und im Falle von Selbständigkeit den tatsächlichen Gewinn des Betriebs angibt.</p> <p>Auch wenn die nachfolgenden Merkmale Unterkategorien zum gesamten verfügbaren Einkommen darstellen und somit in dessen Summe eingehen, sind diese nicht erschöpfend. Daher kann die eigens gebildete Summe aus diesen Unterkategorien von der Variable „Verfügbares Einkommen gesamt“ nicht nur aufgrund der Rundungen abweichen.</p>
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Einkommensangaben verfügbar, d.h. das Gesamteinkommen der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.10 Verfügbares Einkommen aus Beschäftigung (veink\_besch)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Beschäftigung
Variablenname	veink_besch
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Beschäftigung an. Hierzu zählen nicht nur Einkommen aus SV-pflichtiger sowie geringfügiger Beschäftigung, sondern auch weitere Einkommen aus nichtselbständiger Arbeit wie Taschengeld bei Freiwilligendiensten, Einkommen aus ehrenamtlicher Tätigkeit usw.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.11 Verfügbares Einkommen aus Selbständigkeit (veink\_selbst)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Selbständigkeit
Variablenname	veink_selbst
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Selbständigkeit an. Damit können nicht alle Selbstständigen im SGB II identifiziert werden, da im verfügbaren Einkommen nur ein positiver Saldo nach Abzug der Betriebsausgaben von den Einnahmen dargestellt wird.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.12 Verfügbares Einkommen aus Kindergeld (veink\_kinderg)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Kindergeld
Variablenname	veink_kinderg
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Kindergeldzahlungen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.13 Verfügbares Einkommen aus Unterhalt (veink\_unterh)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Unterhalt
Variablenname	veink_unterh
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Unterhaltszahlungen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.14 Verfügbares Einkommen aus Krankengeld (veink\_krankeng)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Krankengeld
Variablenname	veink_krankeng
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Krankengeldzahlungen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.15 Verfügbares Einkommen aus Arbeitslosengeld I (veink\_algl)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus ALG I
Variablenname	veink_algl
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Leistungen aus Arbeitslosengeld I (SGB III) an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.16 Verfügbares Einkommen aus Rente (veink\_rente)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Rente
Variablenname	veink_rente
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Leistungen aus Rentenversicherungsansparungen an. Hierzu zählen alle Rentenarten (auch z.B. Waisen- und Witwenrente).
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.17 Verfügbares Einkommen aus Kapital und Besitz (veink\_kub)

Kategorie	Beschreibung
Variablenlabel	Verfügbares Einkommen aus Kapital und Besitz
Variablenname	veink_kub

Kategorie	Beschreibung
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des der Person verfügbaren Einkommens aus Kapitaleinkünften sowie aus Vermietung und Verpachtung an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.18 Bedarf gesamt ohne BuT (bedarf\_sum)

Kategorie	Beschreibung
Variablenlabel	Bedarf gesamt ohne BuT
Variablenname	bedarf_sum
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Gesamtbedarfs einer Person ohne Bedarf aus Bildung und Teilhabe an.
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.19 Bedarf laufende Kosten der Unterkunft (bedarf\_lkdu)

Kategorie	Beschreibung
Variablenlabel	Bedarf laufende KdU
Variablenname	bedarf_lkdu
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Gesamtbedarfs einer Person aus laufenden Kosten der Unterkunft an.
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.20 Bedarf einmalige Kosten der Unterkunft (bedarf\_ekdu)

Kategorie	Beschreibung
Variablenlabel	Bedarf einmalige KdU
Variablenname	bedarf_ekdu
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Gesamtbedarfs einer Person aus einmaligen Kosten der Unterkunft an.

Kategorie	Beschreibung
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.21 Mehrbedarf Schwangerschaft (d\_schwanger)

Kategorie	Beschreibung
Variablenlabel	Mehrbedarf Schwangerschaft
Variablenname	d_schwanger
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das dichotome Merkmal zeigt an, ob eine Person Mehrbedarf aufgrund von Schwangerschaft hat.
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.22 Mehrbedarf Behinderung (d\_behind)

Kategorie	Beschreibung
Variablenlabel	Mehrbedarf Behinderung
Variablenname	d_behind
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das dichotome Merkmal gibt an, ob eine Person Mehrbedarf aufgrund von Behinderung hat.
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.23 Mehrbedarf Ernährung (d\_ernaehr)

Kategorie	Beschreibung
Variablenlabel	Mehrbedarf Ernährung
Variablenname	d_ernaehr
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das dichotome Merkmal gibt an, ob eine Person Mehrbedarf aufgrund von besonderer, kostenaufwändiger Ernährung aus medizinischer Notwendigkeit hat.

Kategorie	Beschreibung
Besonderheiten	Für vom Leistungsanspruch voll ausgeschlossenen Personen sind keine Bedarfsangaben verfügbar, d.h. der Gesamtbedarf der Bedarfsgemeinschaft kann untererfasst sein. Meldungen aus A2LL enthalten aufgrund einer Umgehungslösung auch Mehrbedarfe nach § 21 (6) (Härtefall). Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.24 Bedarf Darlehen (d\_darleh)

Kategorie	Beschreibung
Variablenlabel	Bedarf Darlehen
Variablenname	d_darleh
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das dichotome Merkmal gibt an, ob (einzelne oder die Gesamtheit der) Bedarfe einer Person als Darlehen gewährt wurden. Der Verwendungszweck kann dabei vielfältig sein, von Darlehen für die Instandhaltung selbst bewohnten Wohneigentums über Mietschulden bis hin zu mehrtägigen Klassenfahrten und orthopädischen Schuhen.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.25 Zahlungsanspruch gesamt ohne BuT (zanspr\_sum)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch gesamt ohne BuT
Variablenname	zanspr_sum
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Gesamtzahlungsanspruchs einer Person ohne Ansprüche auf Sozialversicherungsbeiträge und Leistungen zur Bildung und Teilhabe und ggf. nach Abzug von Sanktionsbeträgen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

#### 5.5.26 Zahlungsanspruch ALG II (zanspr\_algII)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch ALG II
Variablenname	zanspr_algII
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus Arbeitslosengeld II an. Es entspricht der Summe von Regelleistungen, Mehrbedarfen, Kosten der Unterkunft und ALG-II-Zuschlägen ggf. nach Abzug von Sanktionsbeträgen für erwerbsfähige Bedarfsgemeinschaftsmitglieder.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.27 Zahlungsanspruch Sozialgeld (zanspr\_sozg)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch Sozialgeld
Variablenname	zanspr_sozg
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus Sozialgeld an. Es entspricht der Summe von Regelleistungen, Mehrbedarfen und Kosten der Unterkunft ggf. nach Abzug von Sanktionsbeträgen für nichterwerbsfähige Bedarfsgemeinschaftsmitglieder.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.28 Zahlungsanspruch Regelbedarf (zanspr\_regel)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch Regelbedarf
Variablenname	zanspr_regel
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus Regelbedarf ggf. nach Abzug von Sanktionsbeträgen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.29 Zahlungsanspruch ALG-II-Zuschlag (zanspr\_algllzu)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch ALG-II-Zuschlag
Variablenname	zanspr_algllzu
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus ALG-II-Zuschlägen ggf. nach Abzug von Sanktionsbeträgen an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.30 Zahlungsanspruch Mehrbedarfe (zanspr\_mbedarf)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch Mehrbedarfe
Variablenname	zanspr_mbedarf
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus Mehrbedarfen ggf. nach Abzug von Sanktionsbeträgen an. Eine Differenzierung nach Mehrbedarfsart erfolgt hier nicht, diese wird nur für die Bedarfe vorgenommen.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.31 Zahlungsanspruch laufende KdU (zanspr\_lkdu)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch laufende KdU
Variablenname	zanspr_lkdu
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus anerkannten laufenden Kosten der Unterkunft und ggf. nach Abzug von Sanktionsbeträgen an. Hierunter zählen neben direkten Mietkosten auch Betriebs- und Heizkosten. Nachzahlungen zu diesen fallen ebenfalls unter die laufenden Kosten der Unterkunft.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.32 Zahlungsanspruch einmalige KdU (zanspr\_ekdu)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch einmalige KdU
Variablenname	zanspr_ekdu
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Höhe des Zahlungsanspruchs aus anerkannten einmaligen Kosten der Unterkunft ggf. nach Abzug von Sanktionsbeträgen an. Hierunter zählen beispielsweise Mietschulden, Wohnbeschaffungskosten sowie Instandhaltungskosten selbst bewohnten Wohneigentums. Nachzahlungen zu Betriebs- und Heizkosten fallen hingegen unter die laufenden Kosten der Unterkunft.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.33 Zahlungsanspruch Sozialversicherung (zanspr\_sozvers)

Kategorie	Beschreibung
Variablenlabel	Zahlungsanspruch Sozialversicherung
Variablenname	zanspr_sozvers
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Summe der Beiträge zur gesetzlichen Kranken-/Pflege-/Rentenversicherung an. Diese Beiträge sind durch Sanktionen nicht kürzbar. Zuschüsse zu Versicherungsbeiträgen gemäß § 26 SGB II sind in der Gesamtsumme enthalten.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

### 5.5.34 Sanktionsbetrag gesamt (sanktionen\_sum)

Kategorie	Beschreibung
Variablenlabel	Sanktionsbetrag gesamt
Variablenname	sanktionen_sum
Kategorie	Informationen zu Leistungsbezug
Herkunft	LST
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt Summe der Abzüge an Regelleistung, ALG-Zuschlag, Mehrbedarf und Kosten der Unterkunft an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt.

## 5.6 Informationen zu Arbeitsuche und Maßnahmenteilnahme

### 5.6.1 Berufsgruppe – letzte abgeschlossene Ausbildung (KldB 2010) (beruf2010\_3\_ausb)

Kategorie	Beschreibung
Variablenlabel	Berufsgruppe – letzte abgeschlossene Ausbildung (KldB 2010), 3-Steller
Variablenname	beruf2010_3_ausb
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal enthält die letzte erfolgreich abgeschlossene Berufsausbildung des Bewerbers (seit dem Jahr 2000) auf der Ebene der Berufsgruppe.

### 5.6.2 Berufsuntergruppe – letzte abgeschlossene Ausbildung (KldB 2010) (beruf2010\_4\_ausb)

Kategorie	Beschreibung
Variablenlabel	Berufsuntergruppe – letzte abgeschlossene Ausbildung (KldB 2010), 4-Steller
Variablenname	beruf2010_4_ausb
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal enthält die letzte erfolgreich abgeschlossene Berufsausbildung des Bewerbers (seit dem Jahr 2000) auf der Ebene der Berufsuntergruppe.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die Berufsgruppe (beruf2010_3_ausb) bereitgestellt.

### 5.6.3 Berufsgruppe – gesuchte Tätigkeit (KldB 2010) (beruf2010\_3\_gesucht)

Kategorie	Beschreibung
Variablenlabel	Berufsgruppe – gesuchte Tätigkeit (KldB 2010), 3-Steller
Variablenname	beruf2010_3_gesucht
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme

Kategorie	Beschreibung
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält die vom Kunden als Hauptziel angestrebte Tätigkeit zum Episodenbeginn auf Ebene der Berufsgruppe.

#### 5.6.4 Berufsuntergruppe – gesuchte Tätigkeit (KldB 2010) (beruf2010\_4\_gesucht)

Kategorie	Beschreibung
Variablenlabel	Berufsuntergruppe – gesuchte Tätigkeit (KldB 2010), 4-Steller
Variablenname	beruf2010_4_gesucht
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält die vom Kunden als Hauptziel angestrebte Tätigkeit zum Episodenbeginn auf Ebene der Berufsuntergruppe
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die Berufsgruppe (beruf2010_3_gesucht) bereitgestellt.

#### 5.6.5 Arbeitsuchestatus (asustat)

Kategorie	Beschreibung
Variablenlabel	Arbeitsuche-Status
Variablenname	asustat
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal zeigt den Arbeitsvermittlungsstatus an. Es wird zwischen „arbeitslos arbeitssuchend“, „nicht arbeitslos arbeitssuchend“, „arbeitslos – arbeitsunfähig bis 42 Tage“, „ratsuchend“ und „nicht arbeitssuchend“ unterschieden. Als Ratsuchende gelten Bewerber, die von der BA nur eine Beratung wünschen. Dazu kommen Rehabilitanden und vor 2008 Personen ab dem 58. Lebensjahr, die einer Vermittlung nicht uneingeschränkt zur Verfügung stehen. Ab 01.08.2016 können auch Personen dazu zählen, deren Hilfebedürftigkeit nicht mehr besteht, die aber dennoch weiter vom Jobcenter betreut werden.</p> <p>„Nicht arbeitssuchend“ impliziert hauptsächlich Personen, denen nach § 10 SGB II eine Aktivierung bzw. Vermittlung nicht zuzumuten ist. Ebenso werden Personen mit einer Arbeitsunfähigkeit von mehr als 42 Tagen, die weiterhin ALG II beziehen, unter diesem Status im System geführt.</p> <p>Daneben existiert bei XASU-Fällen ab 1.1.2017 der Status „ALG-I-aufstockend“ als Qualitätshinweis. Demnach wurde zwar vom zkt ein Arbeitsuchestatus gemeldet, die Zuständigkeit für die Vermittlung liegt jedoch nach der geltenden Rechtslage bei der Arbeitsagentur. Daher wird eine Statusmeldung aus der ASU erwartet, die dann auch maßgeblich ist. Da die Meldelogik in vielen Fällen eine Reidentifikation des exakten Geburtsdatums ermöglichen würde, wurden die ursprünglichen Datumsangaben durch das in Abschnitt 3.2.1 beschriebene Anonymisierungsverfahren verändert.</p>
Qualitätshinweis	Die Kategorien „ratsuchend“ und „nicht arbeitssuchend“ sind erst seit der Einführung von Verbis (siehe Abschnitt 4.3.2) vorhanden. Aufgrund von Nachmeldungen sind aber auch Eintragungen vor 2006 zu finden. „Nicht arbeitssuchende“ gelten vor 2008 als untererfasst. In der XASU ist die Qualität der Statusbestimmung bei etlichen Trägern zum Teil eingeschränkt, insbesondere in den ersten Jahren.

#### 5.6.6 Erwerbsstatus vor Arbeitsuche (estatvor)

Kategorie	Beschreibung
Variablenlabel	Erwerbsstatus vor Arbeitsuche
Variablenname	estatvor

Kategorie	Beschreibung
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	Die Variable zeigt für ASU- und XASU-Sätze den Erwerbsstatus vor der Arbeitsuche an. Wird ein Datensatz innerhalb der ASU/XASU künstlich gesplittet (etwa bei einem Rechtskreiswechsel zwischen SGB III und SGB II), wird der Zugangsgrund „DS durch Splitting erzeugt“ vergeben. Dies gilt jedoch nicht für das Episodensplitting innerhalb der SIG. Die Angabe bezieht sich daher nicht notwendigerweise auf das Beginnndatum der Episode, sondern auf den Beginn des Originalzeitraums.

### 5.6.7 Status nach Arbeitsuche (estatnach)

Kategorie	Beschreibung
Variablenlabel	Status nach Arbeitsuche
Variablenname	estatnach
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält den Status der Person nach Abgang aus der Arbeitslosigkeit. Längere Krankheitsepisoden können über dieses Merkmal identifiziert werden. Die Ausprägungen sind folgendermaßen unterteilt: 1000er: Maßnahme (geförderte Beschäftigung) 2000er: ungeförderte Beschäftigung 3000er: in Ausbildung etc. 4000er: in Selbständigkeit 5000er: in Sperre 6000er: in Sonstige
Qualitätshinweis	Es zeigen sich je nach Merkmalsausprägung im Zeitverlauf Auffälligkeiten in den relativen Häufigkeiten. Bei Auswertungen ist daher besondere Vorsicht geboten.

### 5.6.8 Integrationsprognose (ipo)

Kategorie	Beschreibung
Variablenlabel	Integrationsprognose
Variablenname	ipo
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	ASU
Datentyp	numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>Das Merkmal enthält die im Erstkontakt von der Vermittlungs-, Beratungs- und Integrationsfachkraft (bzw. dem Lotsen im SGB III) eingeschätzte Integrationsprognose. Die Integrationsprognose ist das Ergebnis der Einschätzung zur Marktnähe der Kundinnen und Kunden. "Marktnah" gilt dann, wenn eine Integration innerhalb von sechs Monaten zu erwarten ist, "nicht marktnah", wenn die Integration voraussichtlich erst nach mehr als sechs Monaten gelingt. Erkennbare Handlungsbedarfe, die innerhalb von sechs Monaten behoben werden können, schließen Marktnähe nicht aus.</p> <p>Die Integrationsprognose "integriert, aber hilfebedürftig" ist nur gültig für Bewerberinnen und Bewerber aus dem Rechtskreis SGB II.</p> <p>Besteht ausschließlich ein Vermittlungswunsch in Ausbildung, ist keine Integrationsprognose festzulegen.</p> <p>Dieses Merkmal steht ab Ende 2016 zur Verfügung und ist Nachfolger des Merkmals Profillage (profil). Für die Zeit davor werden die Angaben aus der Profillage in die Integrationsprognose wie folgt umgeschlüsselt:</p> <p>Die alten Ausprägungen 2 (Marktprofil), 3 (Aktivierungsprofil) und 4 (Förderprofil) werden zur neuen Ausprägung 1 (Marktnah), die alten Ausprägungen 5 (Entwicklungsprofil), 6 (Stabilisierungsprofil) und 7 (Unterstützungsprofil) zur neuen Ausprägung 2 (nicht marktnah), die alte Ausprägung 8 (Zuordnung nicht erforderlich) zu 3 (Zuordnung nicht erforderlich) und 9 (Integriert, aber hilfebedürftig) zu 4 (Integriert, aber hilfebedürftig) sowie 1 (Noch nicht festgelegt) zu .z (Keine Angabe).</p>
Qualitätshinweis	Das Merkmal wurde 2016 eingeführt, aber bis 2006 aus anderen Merkmalen umgeschlüsselt. Die Qualität wird von der BA-Statistik seit 2010 als verlässlich eingestuft.

### 5.6.9 Art der Kündigung der letzten Tätigkeit (art\_kuend)

Kategorie	Beschreibung
Variablenlabel	Art der Kündigung der letzten Tätigkeit
Variablenname	art_kuend
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU
Datentyp	Numerisch
Detailbeschreibung	<p>Diese Variable ist für Beschäftigungsmeldungen nicht gefüllt!</p> <p>Dieses Merkmal beschreibt die Art der Beendigung des letzten Arbeits- oder Ausbildungsverhältnisses vor einer erneuten Arbeitsuche. Es kann daher zur Identifikation von Job-to-Job-Vermittlungen verwendet werden.</p>
Qualitätshinweis	Der Anteil der fehlenden Werte ist in dieser Variable besonders hoch und liegt durchschnittlich bei etwa 65%. Dies liegt vorwiegend aber daran, dass die Variable nur für Personen gefüllt sein kann, die arbeitssuchend sind und zuvor sozialversicherungspflichtig oder geringfügig beschäftigt waren.

### 5.6.10 Arbeitszeit des Stellengesuchs (arbzeit)

Kategorie	Beschreibung
Variablenlabel	Arbeitszeit des Stellengesuchs
Variablenname	arbzeit
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU
Datentyp	Numerisch
Detailbeschreibung	Arbeitsuchende geben beim Vermittlungsvorgang an, welchem zeitlichen Umfang die gesuchte Stelle entsprechen sollte.

### 5.6.11 Abmeldegrund (grund\_abm)

Kategorie	Beschreibung
Variablenlabel	Abmeldegrund
Variablenname	grund_abm
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme

Kategorie	Beschreibung
Herkunft	ASU, XASU
Datentyp	Numerisch
Detailbeschreibung	<p><b>1) ASU</b> Bei ASU-Sätzen ist der Abmeldegrund enthalten. Im Fall eines Rechtskreiswechsels wird der Satz künstlich gesplittet und als Abmeldegrund "DS durch Splittung erzeugt" eingetragen. Zur korrekten Abbildung der Abgangsgründe ist außerdem das Merkmal „Status nach Arbeitsuche“ zu beachten. Ab dem 26.04.2003 erfolgte eine Reduktion der Merkmalsausprägungen. Für Auswertungen über längere Zeiträume können die alten Werte auf die aktuell gültigen anhand folgender Regeln umgeschlüsselt werden:</p> <p>229 auf 260, 230 auf 260, 231 auf 261, 232 auf 260, 233 auf 260, 234 auf 260, 235 auf 260, 236 auf 261, 237 auf 266, 238 auf 266, 239 auf 271, 240 auf 269, 242 auf 265, 243 auf 270, 244 auf 274, 245 auf 277, 246 auf 267, 247 auf 267, 248 auf 278, 249 auf 269, 250 auf 275, 251 auf 274, 252 auf 276, 253 auf 268, 254 auf 278.</p> <p><b>2) XASU</b> In der XASU werden überschneidungsfreie Zeiträume bezogen auf einen zKT-Kunden abgebildet. Bei Änderungen einer Reihe von Merkmalen wird für die XASU jeweils ein neuer Datensatz gebildet, davon sind folgende Bestandteil der SIG:</p> <ul style="list-style-type: none"> <li>• Wechsel des Arbeitsuche-Status</li> <li>• Wechsel der Verfügbarkeit</li> <li>• Wechsel des SGB-II-Trägers</li> <li>• Wechsel des Wohnortes</li> </ul> <p>Wird die vom Grundsicherungsträger gemeldete Episode aus anderen Gründen künstlich gesplittet (siehe Beginndatum Episode), wird als Abmeldegrund "DS durch Splittung erzeugt" eingetragen.</p>

### 5.6.12 Beginndatum linksenzierter Arbeitslosigkeits-Spell (begin\_alo)

Kategorie	Beschreibung
Variablenlabel	Beginndatum linksenzierter Arbeitslosigkeits-Spell
Variablenname	begin_alo
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Datum
Detailbeschreibung	<p>Das Beginndatum des linksenzierten Arbeitslosigkeits-Spells gibt an, zu welchem Datum die Arbeitslosigkeitsepisode vor dem 1.1.2007 begonnen hat. Der linke Rand ist der 1.1.1997. Dieses technische Datum ist nicht gleichzusetzen mit dem Beginn eines tatsächlich zusammenhängenden Zeitraums von Arbeitslosigkeit, da u.a. kurze Krankheitszeiten die Arbeitslosigkeit rechtlich nicht unterbrechen oder ALO-Episoden aus technischen Gründen aneinandergrenzen können.</p>

### 5.6.13 Beginndatum linksenzierter Arbeitsuche-Spell (begin\_asu)

Kategorie	Beschreibung
Variablenlabel	Beginndatum linksenzierter Arbeitsuche-Spell
Variablenname	begin_asu
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	ASU, XASU
Datentyp	Datum
Detailbeschreibung	<p>Das Beginndatum des linksenzierten Arbeitsuche-Spells gibt an, zu welchem Datum die Arbeitsucheepisode vor dem 1.1.2007 begonnen hat. Der linke Rand ist der 1.1.1997. Dieses technische Datum ist nicht gleichzusetzen mit dem Beginn eines tatsächlich zusammenhängenden Zeitraums von Arbeitsuche, da (X)ASU-Episoden aus technischen Gründen aneinandergrenzen können.</p>

#### 5.6.14 Maßnahmeart - Kategorie (mass\_gr)

Kategorie	Beschreibung
Variablenlabel	Maßnahmeart - Kategorie
Variablenname	mass_gr
Kategorie	Informationen zu Arbeitsuche und Maßnahmenteilnahme
Herkunft	MTH, XMTH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal zeigt die Maßnahmeart-Kategorie an. Dies ist die höchste Hierarchiestufe der Maßnahmeart-Systematiken der Statistik der Bundesagentur für Arbeit.

#### 5.6.15 Maßnahmeart - Gruppe (mass)

Kategorie	Beschreibung
Variablenlabel	Maßnahmeart - Gruppe
Variablenname	mass
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	MTH, XMTH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal zeigt die Maßnahmeart-Gruppe an. Dies ist die zweithöchste Hierarchiestufe der Maßnahmeart-Systematiken der Bundesagentur für Arbeit.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt. Ansonsten wird lediglich die Maßnahmeart-Kategorie bereitgestellt (mass_gr).

### 5.7 Informationen zu Beschäftigung

#### 5.7.1 Tagesentgelt (entgelt\_tag)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt
Variablenname	entgelt_tag
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>In BeH-Sätzen weist das Merkmal jeweils das Bruttotagesentgelt eines Beschäftigten aus. Es wird berechnet aus dem vom Arbeitgeber gemeldeten Zeitraumentgelt sowie der Dauer des Zeitraums der (ungesplitteten) Originalmeldung in Kalendertagen. Das Tagesentgelt ist in Euro angegeben.</p> <p>Bis 1998 meldeten Arbeitgeber grundsätzlich nur das sozialversicherungspflichtige Arbeitsentgelt. Entgelte, die unterhalb der Geringfügigkeitsgrenze lagen, wurden nicht gemeldet. Entgelte, welche die Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung überschritten, werden nur bis zu dieser Grenze gemeldet. In der gesetzlichen Rentenversicherung gibt es vier Beitragsbemessungsgrenzen. Die Grenze der Knappschaftlichen Rentenversicherung liegt meist höher als die Grenze der Deutschen Rentenversicherung Bund (bis 2005 die Arbeiter- und Angestelltenrentenversicherungen). Jedoch ist eine Unterscheidung der beiden Versicherungsträger in den Daten nicht möglich.</p> <p>Seit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren zum 1.4.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze erfasst. Die Beitragsbemessungsgrenze als obere Kapplungsgrenze gilt jedoch weiterhin. In einigen Fällen übersteigen die gemeldeten Entgelte die Beitragsbemessungsgrenze allerdings trotzdem. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnen kann. Dabei ist es dann unerheblich, ob durch diese Addition die für den Meldezeitraum maßgebliche Beitragsbemessungsgrenze in der Rentenversicherung überschritten wird. Es könnte sich allerdings auch um fehlerhafte Angaben bei der Beschäftigungszeit handeln. (Die Entgeltinformationen hingegen können auf Grund ihrer Versicherungsrelevanz als weniger fehleranfällig gesehen werden.)</p> <p>Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist jeweils der Standort des Betriebs). Eine Übersicht über diese Grenzen findet sich unter (<a href="https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/">https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/</a>).</p> <p>Eine Angabe von 0 Euro ist auf "Unterbrechungsmeldungen" zurückzuführen. In diesen Zeiten besteht das Beschäftigungsverhältnis zwar rechtlich fort, es wird aber kein Entgelt gezahlt. Dies ist der Fall bei Krankheit nach Ende der Lohnfortzahlung, Mutterschaftszeiten und Sabbaticals.</p> <p>Das Tagesentgelt wird mit zwei Nachkommastellen ausgewiesen. Alle Werte größer als 0 und kleiner als 0,01 wurden auf 0,01 aufgerundet. Somit können die o.a. Unterbrechungsmeldungen mit der Bedingung Tagesentgelt = 0 identifiziert werden.</p>
Qualitätshinweis	<p>Ab 2013 nimmt die Anzahl der Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) sprunghaft zu (siehe Variable Abgabegrund (grund_abg)). Es ist davon auszugehen, dass Sonderzahlungen, die vor 2013 mit den Jahresmeldungen angegeben wurden, nun extra gemeldet werden. Daher empfiehlt es sich bei Analysen der Entgelte im Zeitverlauf, die Sonderzahlungen zu den entsprechenden Entgelten von zeitgleichen Beschäftigungsepisoden im gleichen Betrieb dazu zu addieren (siehe Merkmal Tagesentgelt (inkl. Einmalzahlungen)).</p>

### 5.7.2 Tagesentgelt (inkl. Einmalzahlungen) (entgelt\_tag\_bonus)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt (inkl. Einmalzahlungen)
Variablenname	entgelt_tag_bonus
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal Tagesentgelt (inkl. Einmalzahlungen) ist ein Zusatzmerkmal zum Merkmal Tagesentgelt (entgelt_tag) und enthält für BeH-Meldungen um Einmalzahlungen/Sonderzahlungen ergänzte und harmonisierte Angaben zum Tagesentgelt. Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) müssen damit nicht mehr selbst integriert werden. Werte oberhalb von 97,5% des Modus der auf ganze Eurobeträge gerundeten Tagesentgelte vollzeitbeschäftigter Männer (pro Jahr und separat für Ost- und Westdeutschland) sind trunziert. Dieser Wert ist niedriger als die offizielle Beitragsbemessungsgrenze. Das Verfahren wird in Abschnitt 4.1. von Drechsler et al. (2023) beschrieben.</p>
Qualitätshinweis	<p>Das Merkmal ist nicht für alle Beobachtungen der Quelle BeH gefüllt. Zum einen werden Episoden mit Abmeldegrund 154 nicht berücksichtigt, zum anderen werden von Drechsler et al. (2023) Mehrfachbeschäftigungen zum gleichen Zeitraum im gleichen Betrieb vorab bereinigt und für die Generierung des Merkmals nicht verwendet. Im vorliegenden Datenprodukt sind diese Episoden noch enthalten.</p>

### 5.7.3 Tagesentgelt (imputiert) (entgelt\_tag\_imp)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt (imputiert)
Variablenname	entgelt_tag_imp
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal Tagesentgelt (imputiert) ist ein Zusatzmerkmal zum Merkmal Tagesentgelt (entgelt_tag) und enthält für BeH-Meldungen von in Vollzeit Beschäftigten im Alter zwischen 14 und 64 Jahren (ohne Auszubildende) imputierte Tagesentgelte, wenn diese wegen Überschreitens der Beitragsbemessungsgrenze in den Meldedaten zensiert sind. Das Merkmal baut auf dem Merkmal Tagesentgelt (inkl. Einmalzahlungen) (entgelt_tag_bonus) auf. Tagesentgelte dieser Subpopulation, die die Beitragsbemessungsgrenze nicht überschreiten, werden aus „entgelt_tag_bonus“ übernommen. Details zum Imputationsverfahren sind in Drechsler et al. (2023) bzw. Drechsler/Ludsteck (2025) beschrieben.
Qualitätshinweis	Drechsler et al. (2023, Kapitel 5) warnen davor, das Merkmal ohne weitere Prüfschritte in Analysen zu verwenden. Insbesondere sollte immer geprüft werden, wie hoch der Anteil zensierter Entgeltangaben im jeweiligen Zielmodell ist, und ob alle Regressoren des Zielmodells entweder direkt oder durch ausreichend gute Proxies auch Eingang in das Imputationsmodell gefunden haben. Es wird daher empfohlen, sich vor einer Nutzung mit den Annahmen und Grenzen des Verfahrens vertraut zu machen.

### 5.7.4 Beruf – ausgeübte Tätigkeit (KldB 1988) (beruf\_aktT)

Kategorie	Beschreibung
Variablenlabel	Beruf – ausgeübte Tätigkeit (KldB 1988)
Variablenname	beruf_aktT
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifizierung der Berufe 1988“ (Bundesanstalt für Arbeit, 1988), in der ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen. Beschäftigungsmeldungen mit einem Enddatum nach dem 30.11.2011 werden nach dem neuen Tätigkeitsschlüssel 2010 (KldB2010) gemeldet (siehe Abschnitt 4.3.4). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB1988 zurückgeschlüsselt. Dabei kommt es zu Unschärfen.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.7.5 Berufsgruppe – ausgeübte Tätigkeit (KldB 2010) (beruf2010\_3\_aktT)

Kategorie	Beschreibung
Variablenlabel	Berufsgruppe – ausgeübte Tätigkeit (KldB 2010), 3-Steller
Variablenname	beruf2010_3_aktT
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V).</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsgruppe wird durch die ersten drei Stellen der Kennziffer erfasst. Das in der KldB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.3.4). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.7.6 Berufsuntergruppe – ausgeübte Tätigkeit (KldB 2010) (beruf2010\_4\_aktT)

Kategorie	Beschreibung
Variablenlabel	Berufsuntergruppe – ausgeübte Tätigkeit (KldB 2010), 4-Steller
Variablenname	beruf2010_4_aktT
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V).</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsuntergruppe wird durch die ersten vier Stellen der Kennziffer erfasst. Das in der KldB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.3.4). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die Berufsgruppe (beruf2010_3_aktT) bereitgestellt.

### 5.7.7 Anforderungsniveau – ausgeübte Tätigkeit (KldB 2010) (niveau\_aktT)

Kategorie	Beschreibung
Variablenlabel	Anforderungsniveau - ausgeübte Tätigkeit (KldB 2010)
Variablenname	niveau_aktT
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V).</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Das an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable Anforderungsniveau (niveau) zur Verfügung gestellt.</p> <p>Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.3.4). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten.</p>
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben.

### 5.7.8 Teilzeit (teilzeit)

Kategorie	Beschreibung
Variablenlabel	Teilzeit
Variablenname	teilzeit
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal Teilzeit (teilzeit) unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Für Teilzeitbeschäftigte wird nur erfasst, ob ihre Stundenzahl eine bestimmte Grenze übersteigt oder nicht. Diese Grenze liegt seit 1988 bei 18 Wochenstunden.
Qualitätshinweis	<p>In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Um dieses Problem abzumildern, wurde am IAB eine Imputation der Arbeitszeit für den fraglichen Zeitraum vorgenommen. Nähere Informationen zum Verfahren können Ludsteck/Thomsen (2016) entnommen werden. Hier wird keine Imputation vorgenommen.</p> <p>Mit der Umstellung des Tätigkeitschlüssels ist ein Bruch in den Daten verbunden, der über die reine Umstellung des Schlüssels hinausgeht, und evtl. durch Aktualisierungseffekte bei der Umstellung der Lohnbuchhaltungssoftware bei den meldenden Betrieben zu erklären ist. Eine weiterführende Diskussion des Problems und ein möglicher Korrekturansatz findet sich in Fitzenberger/Seidlitz (2020).</p>

### 5.7.9 Beschäftigungsstatus (beschstat)

Kategorie	Beschreibung
Variablenlabel	Beschäftigungsstatus
Variablenname	beschstat
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	<p>Das Merkmal Beschäftigungsstatus entspricht der im neuen Meldeverfahren (DEÜV) ab dem 1.1.1999 erfassten Personengruppe. Sie weist beitrags- oder leistungsrechtliche Besonderheiten des Beschäftigungsverhältnisses aus.</p> <p>Treffen für eine Beschäftigungsmeldung mehrere Schlüssel zu, ist vom meldenden Arbeitgeber der kleinste anzugeben. Mehrheitlich handelt sich dabei um sozialversicherungspflichtige Beschäftigungsverhältnisse ohne Besonderheiten, die unter Schlüsselnummer 101 geführt werden. Entsprechend ist es möglich, dass diese Beschäftigungen leicht überschätzt werden.</p> <p>Im Meldeverfahren ist geregelt, dass Wechsel im Beschäftigungsstatus - z.B. die Übernahme eines Auszubildenden nach dem Ende der Berufsausbildung im Ausbildungsbetrieb - durch eine erneute Meldung erfolgen müssen.</p> <p>Seit dem 1.4.1999 werden im Meldeverfahren DEÜV auch geringfügig Beschäftigte erfasst. Diese Personengruppe lässt sich über die Ausprägungen 109 und 209 abgrenzen. Für geringfügig Beschäftigte konnten keine Daten vor dem Zeitpunkt der Einführung der Meldepflicht im Jahr 1999 zurückgeschrieben werden.</p>
Qualitätshinweis	<p>Die Personengruppe kann in Beschäftigungsmeldungen, die sich auf die Jahre vor 1999 beziehen, aber erst 1999 oder später eingegangen sind, enthalten sein. Für Meldungen, die vor 1999 eingegangen sind, wird versucht, anhand bestimmter Regeln mit Hilfe der Merkmale „Ausbildung“, „Stellung im Beruf und Arbeitszeit“ und „Beruf“ sowie weiterer Informationen eine Zuordnung zu den Personengruppen durchzuführen. In vielen Fällen ist das jedoch nicht eindeutig möglich.</p>

### 5.7.10 Gleitzone (gleitz)

Kategorie	Beschreibung
Variablenlabel	Gleitzone
Variablenname	gleitz
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal zeigt an, ob es sich bei der Beschäftigungsmeldung um eine Beschäftigung im Niedriglohnbereich innerhalb der sogenannten „Gleitzone“ bzw. ab Juli 2019 in einem "Übergangsbereich" handelt (sog. Midi-Job), für die der Arbeitnehmer nur einen reduzierten Beitrag zum Gesamtsozialversicherungsbeitrag leisten muss. Da Arbeitnehmer mit einem Entgelt in der Gleitzone freiwillig den „normalen“ Sozialversicherungsbeitrag zahlen können, sind nicht automatisch alle Beschäftigungen mit einem entsprechenden Entgelt als Gleitzone gekennzeichnet.</p> <p>Die Zuordnung erfolgt über das Bruttoarbeitsentgelt und ändern sich im Zeitverlauf:</p> <p>01.04.2003 bis 31.12.2012: 400,01 bis 800,00 Euro  01.01.2013 bis 30.06.2019: 450,01 bis 850,00 Euro  Ab 01.07.2019: 450,01 bis 1.300,00 Euro</p> <p>Abweichende untere Grenzen existieren für Auszubildende und Beschäftigte in Freiwilligendiensten. Bei mehreren Beschäftigungsverhältnissen ist das insgesamt erzielte Arbeitsentgelt maßgebend.</p> <p>Die entsprechende gesetzliche Regelung besteht seit dem 01.04.2003.</p>

### 5.7.11 Leiharbeit (leih)

Kategorie	Beschreibung
Variablenlabel	Leiharbeit
Variablenname	leih
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	<p>Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein Leiharbeitsverhältnis handelt. Die Meldung erfolgt durch den verleihenden Betrieb, somit werden die Beschäftigten dem verleihenden Betrieb zugeordnet. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.</p>

Kategorie	Beschreibung
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Im Jahr 2012 beträgt die Missingquote aber nur noch rund 5 % und sinkt bis 2023 auf unter 1%.

### 5.7.12 Befristung (befrist)

Kategorie	Beschreibung
Variablenlabel	Befristung
Variablenname	befrist
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein befristetes oder unbefristetes Arbeitsverhältnis handelt. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Im Jahr 2012 beträgt die Missingquote nur aber noch rund 5 % und sinkt bis 2023 auf unter 1%.

### 5.7.13 Abgabegrund (grund\_abg)

Kategorie	Beschreibung
Variablenlabel	Abgabegrund
Variablenname	grund_abg
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	In BeH-Meldungen kennzeichnet der Abgabegrund den Anlass, aus dem der Arbeitgeber den Sozialversicherungsträgern die betreffende Beschäftigungsmeldung übermittelt hat. Allerdings werden in die IEB nicht alle im Rahmen des Meldeverfahrens mögliche Abgabegründe integriert. So beinhalten die IEB nur Meldungen mit Entgeltangaben (also Jahres-, Unterbrechungs- und Abmeldungen), während Anmeldungen wegen fehlender Entgeltinformationen nicht enthalten sind. Ein Informationsverlust ist damit jedoch nicht verbunden, da die Angaben aus einer Anmeldung mit der darauffolgenden Jahres-, Unterbrechungs- oder Abmeldung erneut übermittelt werden. Die Abgabegründe für Beschäftigungsmeldungen sind nach den Regeln des seit 01.01.1999 geltenden Meldeverfahrens (gemäß DEÜV) mit einer führenden „1“ kodiert. Um die Originalausprägungen zu erhalten, muss der Wert „100“ von den Ausprägungen abgezogen werden.
Qualitätshinweis	Ab 2013 nimmt die Anzahl der Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) sprunghaft zu. Solange ein Beschäftigungsverhältnis besteht, können Sonderzahlung, die bis März des Folgejahres ausgezahlt werden, mit der normalen Meldung (v.a. Jahresmeldung) für das Vorjahr angegeben werden. Eine gesonderte Meldung mit Abmeldegrund 54 ist dann nicht erforderlich. Bis 2012 konnten die Jahresmeldungen bis Mitte April abgegeben werden, seit 2013 müssen sie nun bis spätestens Mitte Februar erfolgen. Für Sonderzahlungen, die im Februar und März ausgezahlt werden, muss nun eine extra Meldung erfolgen.

### 5.7.14 Mehrfachbeschäftigung (mehrfbesch)

Kategorie	Beschreibung
Variablenlabel	Mehrfachbeschäftigung ja/nein
Variablenname	mehrfbesch
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	Das Merkmal gibt an, ob eine Person eine zweite gleichzeitige Beschäftigung ausübt.

### 5.7.15 Tagesentgelt – zweite Beschäftigung (mehrbesch\_tentg)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt – zweite Beschäftigung
Variablenname	mehrbesch_tentg
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält das Tagesentgelt der zweiten gleichzeitigen Beschäftigung (s. auch Kapitel 5.7.1).

### 5.7.16 Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung (mehrbesch\_bonus\_tentg)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt (inkl. Einmalzahlungen) – zweite Beschäftigung
Variablenname	mehrbesch_bonus_tentg
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält das Tagesentgelt (inkl. Einmalzahlungen) der zweiten gleichzeitigen Beschäftigung (s. auch Kapitel 5.7.2).

### 5.7.17 Beschäftigungsstatus – zweite Beschäftigung (mehrbesch\_beschstat)

Kategorie	Beschreibung
Variablenlabel	Beschäftigungsstatus – zweite Beschäftigung
Variablenname	mehrbesch_beschstat
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält den Beschäftigungsstatus der zweiten gleichzeitigen Beschäftigung (s. auch Kapitel 5.7.9).

### 5.7.18 Teilzeit – zweite Beschäftigung (mehrbesch\_teilzeit)

Kategorie	Beschreibung
Variablenlabel	Teilzeit – zweite Beschäftigung
Variablenname	mehrbesch_teilzeit
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal unterscheidet zwischen Voll- und Teilzeit der zweiten gleichzeitigen Beschäftigung (s. auch Kapitel 5.7.8).

### 5.7.19 Erwerbserfahrung vor 2007 (exp\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Erwerbserfahrung vor 2007 (in Tagen)
Variablenname	exp_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die gemeldete Erwerbserfahrung einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an (in Tagen).

### 5.7.20 Zeit seit letzter Erwerbstätigkeit (time\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Zeit seit letzter Erwerbstätigkeit vor 2007 (in Tagen)
Variablenname	time_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Zeit seit der letzten gemeldeten Erwerbstätigkeit einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an (in Tagen).

### 5.7.21 Tagesentgelt vor 2007 (tentgelt\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt vor 2007
Variablenname	tentgelt_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt das Tagesentgelt der letzten gemeldeten Hauptbeschäftigung einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an.

### 5.7.22 Beruf – letzte ausgeübte Tätigkeit (KldB 1988) vor 2007 (beruf1988\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Beruf – letzte ausgeübte Tätigkeit (KldB 1988) vor 2007
Variablenname	beruf1988_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt den Beruf der letzten ausgeübten und gemeldeten Hauptbeschäftigung (KldB 1988) einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an.

### 5.7.23 Berufsgruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 3-Steller (beruf2010\_3\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Berufsgruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 3-Steller

Kategorie	Beschreibung
Variablenname	beruf2010_3_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Berufsgruppe der letzten und gemeldeten ausgeübten Hauptbeschäftigung (KldB 2010, 3-Steller) einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an.

#### 5.7.24 Berufsuntergruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 4-Steller (beruf2010\_4\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Berufsuntergruppe – letzte ausgeübte Tätigkeit (KldB 2010) vor 2007, 4-Steller
Variablenname	beruf2010_4_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Berufsuntergruppe der letzten ausgeübten und gemeldeten Hauptbeschäftigung (KldB 2010, 4-Steller) einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die Berufsgruppe (beruf2010_3_vor2007) bereitgestellt.

#### 5.7.25 Teilzeit vor 2007 (teilzeit\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Teilzeit vor 2007
Variablenname	teilzeit_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt an, ob für eine Person zwischen dem 1.1.1993 und dem 31.12.2006 ausschließlich eine Teilzeit-Beschäftigung gemeldet wurde hat (letzte Hauptbeschäftigung).

#### 5.7.26 Beschäftigungsstatus vor 2007 (beschstat\_vor2007)

Kategorie	Beschreibung
Variablenlabel	Beschäftigungsstatus vor 2007
Variablenname	beschstat_vor2007
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt den Beschäftigungsstatus der letzten ausgeübten Hauptbeschäftigung einer Person zwischen dem 1.1.1993 und dem 31.12.2006 an.

## 5.8 Ortsangaben

### 5.8.1 Wohnort - Kreis (wo\_kreis\_lhg, wo\_kreis\_x\_asu, wo\_kreis\_leh, wo\_kreis\_beh)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Kreis
Variablenname	wo_kreis_lhg, wo_kreis_x_asu, wo_kreis_leh, wo_kreis_beh
Kategorie	Ortsangaben
Herkunft	LHG, ASU, XASU, LeH, BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	<p>Das Merkmal gibt jeweils den Kreis (kreisfreie Stadt oder Landkreis) an, in dem der Versicherte wohnt. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes (NUTS-1) und in der ersten bis dritten Stelle den Regierungsbezirk (NUTS-2). Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. In der Systematik NUTS (Nomenclature des unités territoriales statistiques) der Europäischen Union entsprechen die Kreise der Hierarchieebene NUTS 3.</p> <p>Für die Quellen LHG und XASU gilt der Wohnort für den gesamten Zeitraum der Originalbeobachtung. Bei BeH, ASU und LeH enthält das Merkmal den Wohnort zum Beginn des Originalzeitraums. Mit der Dauer des Spells steigt also das Risiko, dass der Wohnort veraltet ist und die Angaben zu späteren Stichtagen falsch sind.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben für alle Quellen auf den Gebietsstand vom 31.12.2023 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2023 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandsaktualisierung Fälle auftreten, bei denen die Kreiskennziffer des Wohnortes wechselt, ohne dass der betreffende Versicherte umgezogen ist.</p>
Qualitätshinweis	<p>Hinsichtlich der Wohnortangaben in der BeH bestehen für einzelne Beschäftigte Erhebungsungenauigkeiten. Die Meldevorschrift stellt nicht klar, welcher Wohnsitz – Haupt- oder Nebenwohnsitz mit überwiegendem Aufenthaltsort – vom Arbeitgeber zu melden ist. Dies kann in der Beschäftigungsstatistik zum Nachweis von "Fernpendlern" zwischen gemeldetem Hauptwohnsitz und Arbeitsort führen, obwohl die beschäftigte Person ihrer Beschäftigung am Nebenwohnsitz nachgeht, also faktisch nicht pendelt.</p> <p>Im Jahr 2015 kommt es in den BeH-Daten zu einer Reduktion der Wohnortwechsel von ca. 10-15% gegenüber dem üblichen Niveau der umliegenden Jahre, wobei regionale Unterschiede auftreten. Der genaue Grund für diese Abweichung konnte bisher nicht ermittelt werden.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das Bundesland (wo_bula) bereitgestellt.

### 5.8.2 Wohnort - Bundesland (wo\_bula\_lhg, wo\_bula\_x\_asu, wo\_bula\_leh, wo\_bula\_beh)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Bundesland
Variablenname	wo_bula_lhg, wo_bula_x_asu, wo_bula_leh, wo_bula_beh
Kategorie	Ortsangaben
Herkunft	LHG, ASU, XASU, LeH, BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal Wohnort Kreis auf die 16 deutschen Bundesländer. Weitere Informationen zum Kreis des Wohnortes finden sich unter Wohnort - Kreis (wo_kreis_lhg, wo_kreis_x_asu, wo_kreis_leh, wo_kreis_beh).

### 5.8.3 Arbeitsort Kreis (ao\_kreis)

Kategorie	Beschreibung
Variablenlabel	Arbeitsort Kreis
Variablenname	ao_kreis
Kategorie	Ortsangaben
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	<p>Angegeben ist jeweils der Kreis (kreisfreie Stadt oder Landkreis), in dem der Betrieb des Beschäftigten angesiedelt ist. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes (NUTS-1), in der ersten bis dritten Stelle den Regierungsbezirk (NUTS-2). Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. In der Systematik NUTS (Nomenclature des unités territoriales statistiques) der Europäischen Union entsprechen die Kreise der Hierarchieebene NUTS 3.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben auf den Gebietsstand vom 31.12.2023 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2023 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandsaktualisierung Fälle auftreten, bei denen die Kreiskennziffer des Betriebsortes wechselt, ohne dass dieser Betrieb seinen Standort verlagert hat.</p>
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das Bundesland (ao_bula) bereitgestellt.

### 5.8.4 Arbeitsort Bundesland (ao\_bula)

Kategorie	Beschreibung
Variablenlabel	Arbeitsort Bundesland
Variablenname	ao_bula
Kategorie	Ortsangaben
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Bundesland   Kreis
Detailbeschreibung	Angegeben ist jeweils das Bundesland, in dem der Betrieb angesiedelt ist. Dieses Merkmal wird aus der Kreiskennziffer (ao_kreis) generiert. Die ersten beiden Stellen der Kreiskennziffer bezeichnen das Bundesland (NUTS-1).

## 5.9 Betriebsmerkmale

### 5.9.1 WZ08 5-Steller (w08\_5\_kons)

Kategorie	Beschreibung
Variablenlabel	WZ08 5-Steller
Variablenname	w08_5_kons
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)

Kategorie	Beschreibung
Detailbeschreibung	<p>Dieses Merkmal weist den konsolidierten Wirtschaftszweig gemäß der Klassifikation WZ08 als 5-Steller aus und ist für den kompletten Zeitraum gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.</p>
Besonderheiten	<p>Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w08_3_kons) bereitgestellt.</p>

### 5.9.2 WZ08 3-Steller (w08\_3\_kons)

Kategorie	Beschreibung
Variablenlabel	WZ08 3-Steller
Variablenname	w08_3_kons
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller)   Wirtschaftsabteilung (2-Steller)   Wirtschaftsgruppe (3-Steller)   Wirtschaftsklasse (4-Steller)   Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	<p>Dieses Merkmal weist den konsolidierten Wirtschaftszweig gemäß der Klassifikation WZ08 als 3-Steller aus und ist für den kompletten Zeitraum gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf.</p> <p>Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.</p>

# Literatur

- Bender, Stefan; Hilzendegen, Jürgen; Rohwer, Götz; Rudolph, Helmut (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg.
- Bertat, Thomas; Dundler, Agnes; Grimm, Christopher; Kiewitt, Jochen; Schomaker, Christine; Schridde, Henning; Zemmann, Christian (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: <https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodik-Qualitaet/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erhebungsinhalte-Arbeitszeit-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf> (letzter Zugriff: 27.11.2019).
- Bundesagentur für Arbeit (Hrsg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.
- Bundesagentur für Arbeit (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: <https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB-2010/Generische-Publikationen/KldB2010-Printversion-Band1.pdf> (letzter Zugriff: 27.11.2019).
- Bundesagentur für Arbeit – Statistik (2015): Statistik zu Bildung und Teilhabe nach dem SGB II. Methodenbericht. Nürnberg, September 2015. URL: <https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodik-Qualitaet/Methodenberichte/Grundsicherung-Arbeit-suchende-SGBII/Generische-Publikationen/Methodenbericht-Bildung-und-Teilhabe.pdf> (letzter Zugriff: 22.11.2019).
- Bundesagentur für Arbeit – Statistik (2017): Aktuelle Hinweise zur Interpretation der Kosten der Unterkunft in der Grundsicherungsstatistik. Kurzinfo. Juni 2017 (erweiterte Fassung).
- Bundesanstalt für Arbeit (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: <https://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html> (letzter Zugriff: 27.11.2019).
- Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv, Jg. 69, S. 56-68.
- Drechsler, Jörg; Ludsteck, Johannes (2025): Imputation strategies for rightcensored wages in longitudinal datasets. In: Journal for Labour Market Research, 59(19). DOI: <https://doi.org/10.1186/s12651-025-00410-4>.
- Drechsler, Jörg; Ludsteck, Johannes; Moczall, Andreas (2023): Imputation der rechtszensierten Tagesentgelte für die BeH. FDZ-Methodenreport 05/2023 (de), Nürnberg, DOI:10.5164/IAB.FDZM.2305.de.v1.
- Eberle, Johanna; Jacobebbinghaus, Peter; Ludsteck, Johannes; Witter, Julia (2011): Generation of time-consistent industry codes in the face of classification changes \* Simple heuristic based on the Establishment History Panel (BHP). FDZ Methodenreport, 05/2011 (en), Nürnberg.

- Fitzenberger, Bernd; Seidlitz, Arnim (2020): The 2011 break in the part-time indicator and the evolution of wage inequality in Germany. *Journal for Labour Market Research* 01/2020.
- Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 126(3), pp. 405-436.
- Ludsteck, Johannes; Thomsen, Ulrich (2016): Imputation of the Working Time Information for the Employment Register Data. *FDZ-Methodenreport*, 01/2016 (en), Nürnberg.
- Meinken, Holger; Koch, Iris (2004): BA-Beschäftigtenpanel 1998-2002. *Codebuch*, Nürnberg.
- Paulus, Wiebke; Matthes, Britta (2013): Klassifikation der Berufe \* Struktur, Codierung und Umsteigeschlüssel. *FDZ-Methodenreport*, 08/2013 (de), Nürnberg.
- Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ08.html> (Zugriff: 21.04.2016)
- Statistisches Bundesamt (2019): Staats- und Gebietssystematik, Wiesbaden. URL: [https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietssystematik/Staatsangehoerigkeitsgebietsschluessel\\_pdf.pdf?blob=publicationFile&v=5%20](https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietssystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?blob=publicationFile&v=5%20) (Zugriff: 05.12.2019).
- Stüber, Heiko; Dauth, Wolfgang; Eppelsheimer, Johann (2023): A guide to preparing the sample of integrated labour market biographies (SIAB, version 7519 v1) for scientific analysis. In: *Journal for Labour Market Research*, 57(7). <https://doi.org/10.1186/s12651-023-00335-w>.
- Thomsen, Ulrich; Ludsteck, Johannes; Schmucker, Alexandra (2018): Skilled or unskilled - Improving the information on qualification for employee data in the IAB Employee Biography. *FDZ-Methodenreport*, 09/2018 (en), Nürnberg.
- Wermter, Winfried; Cramer, Ulrich (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? – Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 4/88, S. 468-482.

# Anhang

## **Auszählungen**

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in gesonderten Dateien unter <https://fdz.iab.de>.

# Impressum

## FDZ-Datenreport 14|2025

### Veröffentlichungsdatum

18. Dezember 2025

### Herausgeber

Forschungsdatenzentrum  
der Bundesagentur für Arbeit  
im Institut für Arbeitsmarkt- und Berufsforschung  
Regensburger Straße 104  
90478 Nürnberg

### Nutzungsrechte

Diese Publikation ist unter folgender Creative-Commons-Lizenz veröffentlicht:  
Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-SA 4.0)  
<https://creativecommons.org/licenses/by-sa/4.0/deed.de>

### Bezugsmöglichkeit dieses Dokuments

[https://doku.iab.de/fdz/reporte/2025/DR\\_14-25.pdf](https://doku.iab.de/fdz/reporte/2025/DR_14-25.pdf)

### Dokumentation Version

SIG0723\_DE\_v1\_dok1, DOI: 10.5164/IAB.FDZD.2514.de.v1

### Datensatz Version

SIG 0723 v1, DOI: 10.5164/IAB.SIG0723.de.en.v1

### Bezugsmöglichkeit aller Veröffentlichungen der Reihe „FDZ-Datenreport“

<https://fdz.iab.de/forschung/publikationen/fdz-datenreporte/>

### Website

<https://fdz.iab.de/>

---

### Rückfragen zum Inhalt

Sandra Dummert  
Telefon: 0911 179-7005  
E-Mail: [sandra.dummert@iab.de](mailto:sandra.dummert@iab.de)