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

FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

09|2026 DE Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell (LIAB LM) 1975-2024

Andreas Filser, Alexandra Schmucker, Matthias Umkehrer



Bundesagentur für Arbeit

Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell (LIAB LM) 1975-2024

Andreas Filser (IAB), Alexandra Schmucker (IAB), Matthias Umkehrer (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

1 Einleitung und Kurzbeschreibung 6

1.1 Einleitung 6

1.2 Datennutzung 8

1.2.1 Datenzugang 8

1.2.2 Datenhaltung 8

1.2.3 Sensible Merkmale und Zusatzvariablen 10

1.3 Veränderungen zum LIAB LM 7521 12

1.3.1 Stichprobe und Verknüpfung 12

1.3.2 Datenquellen und Merkmalskanon 12

1.3.3 Ziehung der Daten aus dem BHP 13

1.3.4 Datendokumentation 14

1.4 Steckbrief 14

1.5 Mengengerüst 16

2 Datenquellen 17

3 Datenaufbereitung und Stichprobenziehung 19

3.1 Bereinigungen und Aufbereitungen 19

3.2 Episodensplitting 20

3.3 Stichprobenziehung 21

3.4 Qualität der Verknüpfung 22

3.5 Gewichtung 23

3.6 Fehlende Werte 23

4 Datenqualität und Probleme 24

4.1 Gesamte IEB 24

4.2 Beschäftigtenhistorik 25

5 Variablenbeschreibung 29

5.1 Identifikatoren 29

5.1.1 Systemfreie Personennummer (persnr_liab_lm) 29

5.1.2 Systemfreie Betriebsnummer (betnr_liab_lm) 29

5.2 Generierte technische Merkmale 30

5.2.1 Status der Betriebsnummer (betr_st) 30

5.2.2 Beobachtung wurde gezogen (smpl_st) 31

5.2.3 Satzzähler pro Konto (spell) 31

5.2.4 Jahr (jahr) 32

5.3 Gültigkeitszeitraum 32

5.3.1 Beginndatum Originalbeobachtung (begorig) ... 32
5.3.2 Enddatum Originalbeobachtung (endorig) ... 32
5.3.3 Beginndatum Episode (begepi) ... 33
5.3.4 Enddatum Episode (endepi) ... 33
5.4 Informationen zur Person ... 33
5.4.1 Geschlecht (frau) ... 33
5.4.2 Geburtsjahr (gebjahr) ... 34
5.4.3 Geburtsmonat (gebmon) ... 34
5.4.4 Staatsangehörigkeit (nation) ... 34
5.4.5 Staatsangehörigkeit vergröbert (nation_gr) ... 34
5.4.6 Ausbildung (ausbildung) ... 35
5.4.7 Ausbildung (imputiert) (ausbildung_imp) ... 36
5.4.8 Schulausbildung (schule) ... 36
5.5 Informationen zu Beschäftigung ... 37
5.5.1 Tagesentgelt/ täglicher Leistungssatz (tentgelt) ... 37
5.5.2 Tagesentgelt (inkl. Einmalzahlungen) (tentgelt_bonus) ... 38
5.5.3 Tagesentgelt (imputiert) (tentgelt_imp) ... 38
5.5.4 Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf) ... 39
5.5.5 Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3) ... 39
5.5.6 Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4) ... 40
5.5.7 Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau) ... 41
5.5.8 Teilzeit (teilzeit) ... 42
5.5.9 Stellung im Beruf und Arbeitszeit (stib) ... 42
5.5.10 Erwerbsstatus (erwstat) ... 43
5.5.11 Gleitzone (gleitz) ... 44
5.5.12 Leiharbeit (leih) ... 44
5.5.13 Befristung (befrist) ... 45
5.5.14 Abmeldegrund / Abgabegrund / Beendigungsgrund (grund) ... 45
5.5.15 Dauer der Arbeitslosigkeit (alo_dau) ... 46
5.6 Ortsangaben ... 46
5.6.1 Wohnort - Kreis (wo_kreis) ... 46
5.6.2 Wohnort - Bundesland (wo_bula) ... 47
5.6.3 Wohnort - Arbeitsagentur (wo_aa) ... 47
5.6.4 Wohnort - Regionaldirektion (wo_rd) ... 48
5.7 Betriebsmerkmale ... 48
5.7.1 WZ73 3-Steller (w73_3) ... 48
5.7.2 WZ93 5-Steller (w93_5) ... 49
5.7.3 WZ93 3-Steller (w93_3) ... 49
5.7.4 WZ03 5-Steller (w03_5) ... 50
5.7.5 WZ03 3-Steller (w03_3) ... 50
5.7.6 WZ08 5-Steller (w08_5) ... 51
5.7.7 WZ08 3-Steller (w08_3) ... 51
5.7.8 w73_3 vervollständigt durch Extrapolation/Imputation (w73_3_gen) ... 52

5.7.9 Art der Vervollständigung w73_3 (group_w73_3) 52
5.7.10 w93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen) 52
5.7.11 Art der Vervollständigung w93_3 (group_w93_3) 53
5.7.12 w08_3 vervollständigt durch Extrapolation/Imputation (w08_3_gen) 53
5.7.13 Art der Vervollständigung w08_3 (group_w08_3) 54
5.7.14 Jahr erstes Auftreten Betriebsnummer (grd_jahr) 54
5.7.15 Jahr letztes Auftreten Betriebsnummer (lzt_jahr) 54
5.7.16 Anzahl Beschäftigte gesamt (az_ges) 55
5.7.17 Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz) 55
5.7.18 Anzahl geringfügig Beschäftigte (az_gf) 55
5.7.19 Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw) 56
5.7.20 Arbeitsort - Kreis (ao_kreis) 56
5.7.21 Arbeitsort - Bundesland (ao_bula) 57

6 Literatur 58

7 Anhang 61

7.1 Labels, Auszählungen und fehlende Werte 61

7.2 Abkürzungsverzeichnis 61

Zusammenfassung

Dieser Datenreport beschreibt das LIAB-Längsschnittmodell (LIAB LM) 1975 - 2024.

Abstract

This data report describes the LIAB Longitudinal Model (LIAB LM) 1975 - 2024.

Keywords

German administrative micro data, labour market data, linked employer-employee data, data manual

Danksagung

Für ihre Mitarbeit und Unterstützung danken wir den Kolleginnen und Kollegen des Forschungsdatenzentrums (FDZ), der Forschungseinheit Betriebe und Beschäftigung (BB) sowie des Geschäftsbereichs Daten- und IT-Management (DIM) des Instituts für Arbeitsmarkt- und Berufsforschung. Der vorliegende Datenreport ist eine Überarbeitung des FDZ-Datenreports 04/2024 (Panahian Fard, Schmucker, Seth, Umkehrer, Zimmermann, DOI: 10.5164/IAB.FDZD.2404.de.v1). Darüber hinaus wurden einzelne Passagen aus IAB-internen Datendokumentationen von IAB-DIM übernommen.

Datenverfügbarkeit

Der in diesem Beitrag beschriebene Datensatz ist für die Fachöffentlichkeit zugänglich. Nähere Informationen dazu auf der Internetseite https://fdz.iab.de/.

1 Einleitung und Kurzbeschreibung

1.1 Einleitung

Das LIAB-Längsschnittmodell (LIAB LM) 7524 ist einer der Linked-Employer-Employee-Datensätze des Instituts für Arbeitsmarkt- und Berufsforschung (IAB), die über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit (BA) für wissenschaftliche Auswertungen zur Verfügung stehen. Es verknüpft Informationen zu Betrieben aus dem IAB-Betriebspanel, einer jährlich stattfindenden Betriebsbefragung, mit Information zu deren Beschäftigten aus den administrativen Daten der BA.

Das IAB-Betriebspanel ist eine repräsentative Arbeitgeberbefragung zu betrieblichen Determinanten der Beschäftigung. Die Erhebung wird seit 1993 in Westdeutschland und seit 1996 in Ostdeutschland durchgeführt. Sie stellt die zentrale Quelle für Analysen zur Arbeitskräftenachfrage auf dem deutschen Arbeitsmarkt dar. Dieser Datenreport konzentriert sich auf die Beschreibung der Personendaten im LIAB LM. Nähere Informationen zum IAB-Betriebspanel finden sich auf den Internetseiten des FDZ (https://fdz.iab.de/unsere-datenprodukte/betriebsdaten/iab-betriebspanel/).

Die administrativen Daten zu den Beschäftigten stammen aus den Integrierten Erwerbsbiografien (IEB) des IAB. Die IEB enthalten Daten zu allen Personen, die im Beobachtungszeitraum mindestens einmal einen der folgenden Zustände aufweisen:

- sozialversicherungspflichtige Beschäftigung in Deutschland (erfasst ab 1975),
- geringfügige Beschäftigung in Deutschland (erfasst ab 1999),
- Bezug von Leistungen nach dem Rechtskreis SGB III (erfasst ab 1975) oder SGB II (erfasst ab 2005),
- bei der Bundesagentur für Arbeit (BA) bzw. den Grundsicherungsträgern als arbeitsuchend gemeldet (erfasst ab 1997),
- Teilnahme an arbeitsmarktpolitischer Maßnahme (erfasst ab 2000).

Diese aus unterschiedlichen Datenquellen stammenden Informationen werden in den IEB zusammengeführt und tagesgenau abgebildet. Das LIAB LM 7524 umfasst davon ausschließlich Daten zu sozialversicherungspflichtiger und geringfügiger Beschäftigung aus der Beschäftigtenhistorik (BeH). Weitere in den IEB einbezogene Datenquellen sind im LIAB LM 7524 nicht enthalten.

Das IAB-Betriebspanel (Version 9324 v1) bildet den Ausgangspunkt für die Ziehung der Personendaten des LIAB LM 7524. In einem ersten Schritt werden alle Betriebe ausgewählt, die in einem bestimmten Zeitraum, hiernach als *Basisperiode* bezeichnet, Teil des IAB-Betriebspanels sind und für die eine Zustimmung zur Verknüpfung der Daten vorliegt.[1] Für die Version LIAB LM 7524 werden folgende Basisperioden[2] angeboten:

[1] Der Anteil der Betriebe, für die eine Zustimmung zur Verknüpfung mit den administrativen Daten vorliegt, liegt bis zum Jahr 2010 bei über 98 Prozent, fällt dann aber auf etwa 95 Prozent im Jahr 2024. Im Jahr 2020 ist der Anteil mit knapp 92 Prozent am geringsten. Dieser Rückgang ist auf eine Änderung der Erhebung der Zuspielbereitschaft ab dem Jahr 2020 zurückzuführen, die in erster Linie Betriebe, die ab dem Jahr 2020 an der Befragung teilgenommen haben, betrifft.

[2] Für die Durchführung eines konkreten Forschungsprojektes ist jeweils die am besten geeignete Basisperiode auszuwählen. Es ist nicht möglich, mehrere Basisperioden zu wählen, zu kombinieren oder zu erweitern.

- 1993-2000 (B9300)
- 2000-2007 (B0007)
- 2007-2017 (B0717)
- 2014-2024 (B1424)

Für diese Betriebe werden dann in einem zweiten Schritt die Beschäftigten, die innerhalb der Basisperiode und ein Jahr davor im Betrieb tätig waren, aus den IEB gezogen. Aufgrund von Größenbeschränkungen können die Erwerbsbiografien der gesamten Belegschaft nur für Betriebe mit bis zu 49 sozialversicherungspflichtig Beschäftigten angeboten werden. Bei größeren Betrieben wird eine Zufallsstichprobe von 20 Prozent der Beschäftigten, die innerhalb der Basisperiode und ein Jahr davor in einem Betrieb tätig waren, gezogen. Für alle so ermittelten Beschäftigten werden schließlich die tagesgenauen Erwerbsbiografien von 1975 bis 2024 bereitgestellt. Je nach Basisperiode umfasst der Datensatz die Erwerbsverläufe von 2.338.291 bis 3.772.719 Personen aus 18.505 bis 47.077 verknüpften Betrieben.

Trotz der genannten Einschränkung bei der Ziehung der Personendaten umfasst das LIAB LM 7524 die Befragungsdaten zu allen Betrieben, die im IAB-Betriebspanel zwischen 1993 und 2024 befragt wurden und für die eine Zustimmung zur Verknüpfung der Daten vorliegt. Weiterhin können administrative Betriebsinformationen aus dem Betriebs-Historik-Panel (BHP) zugespielt werden (siehe https://fdz.iab.de/unsere datenprodukte/betriebsdaten/bhp/). Die Informationen aus dem BHP sind auch für die Betriebe enthalten, die ausschließlich im IAB-Betriebspanel vorkommen.

In Abgrenzung zum Querschnittmodell der LIAB umfasst das Längsschnittmodell zeitraumbezogene Erwerbsbiografien der verknüpften Personen für eine Untergruppe von Betrieben, die im IAB-Betriebspanel befragt wurden. Für diese Personen lassen sich Erwerbsbiografien somit auch bei Betriebswechseln und Übergängen in Arbeitslosigkeit verfolgen. Im Querschnittmodell hingegen werden die Erwerbsbiografien jeweils immer nur zum Stichtag 30. Juni abgebildet. Allerdings werden zu allen im IAB-Betriebspanel befragten Betrieben die Erwerbsbiografien aller Beschäftigten zugespielt, die im Befragungsjahr zum Stichtag 30. Juni bei einem dieser Betriebe tätig waren. Eine Übersicht zu den verschiedenen LIAB-Datensätzen findet sich auf den Internetseiten des FDZ (https://fdz.iab.de/unsere-datenprodukte/integrierte-betriebs-und-personendaten/liab/).

Der vorliegende Datenreport beschreibt die Merkmale aus den Personendaten des LIAB LM, welche weitgehend die Originaldaten, d.h. nicht vergröberte Daten, beinhalten. Einige Merkmale sind allerdings als besonders sensibel eingestuft und werden nur auf gesonderten Antrag (siehe Abschnitt 1.2.3) weitergegeben.

Dieser Datenreport gliedert sich wie folgt: Neben der Einleitung enthält Kapitel 1 unter anderem Informationen zum Datenzugang sowie eine Kurzbeschreibung der Daten und das Mengengerüst. Eine Beschreibung der einzelnen Datenquellen findet sich in Kapitel 2. Die Datenaufbereitung sowie die Datenqualität werden in den Kapiteln 3 und 4 diskutiert, während die einzelnen Merkmale in Kapitel 5 dargestellt werden.

1.2 Datennutzung

1.2.1 Datenzugang

Die LIAB liegen in schwach anonymisierter Version vor und können ausschließlich im Rahmen von Gastaufenthalten am FDZ und der 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 der Forschenden geschlossen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können der FDZ-Website (https://fdz.iab.de/) entnommen werden.

1.2.2 Datenhaltung

Die LIAB, die sowohl deutsche als auch englische Labels enthalten,[3] sind modular gegliedert und in mehreren Dateien abgelegt. Ein erstes Modul umfasst alle Wellen der Betriebsbefragung des IAB-Betriebspanels, die jeweils als separater Datensatz abgelegt sind. In den Wellen finden sich jahresspezifische Informationen zu Beschäftigungsentwicklung, Geschäftspolitik und Geschäftsentwicklung, betriebliche Investitionen, Innovationen im Betrieb, öffentliche Förderung, Personalstruktur, Berufsausbildung und Ausbildungsstellen, Personaleinstellungen und -abgänge, Personalsuche, Löhne und Gehälter, betriebliche Arbeitszeiten, Fort- und Weiterbildung oder auch allgemeine Angaben zum Betrieb.

Ein zweites Modul, das im weiteren Verlauf Personendatei genannt wird, enthält Identifikatoren (systemfreie Personen- und Betriebsnummern), persönliche Merkmale, Informationen zu Beschäftigung und Merkmale zum Wohnort sowie technische Merkmale der verknüpften Personen.

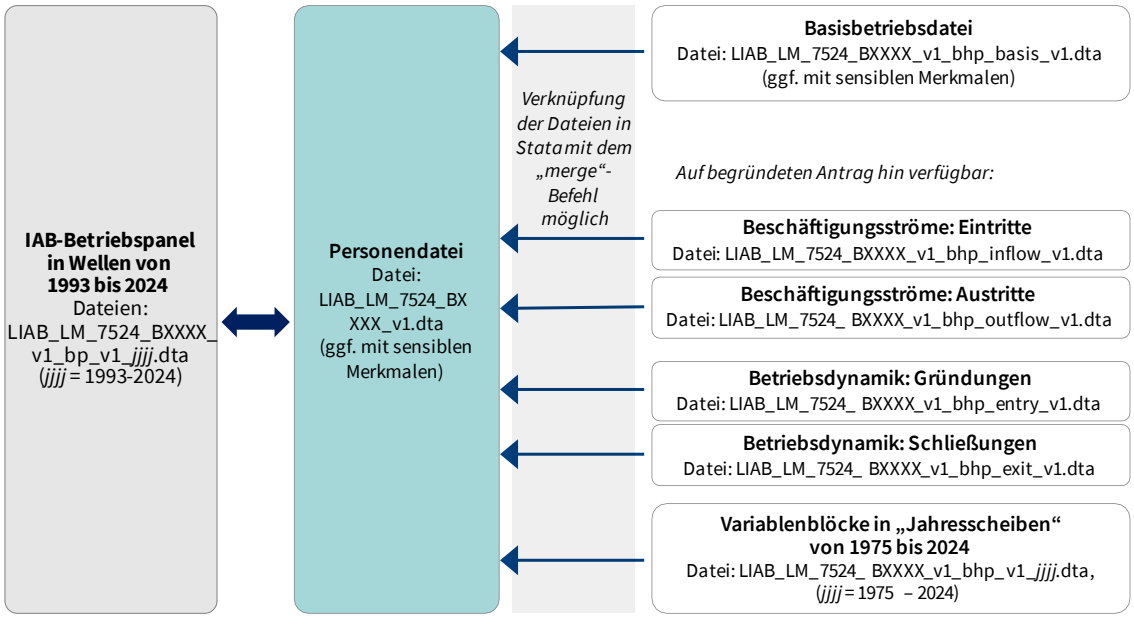
Ein drittes Modul, die Basis-Betriebsdatei, enthält die Betriebs-ID, das Jahr und Merkmale zu Arbeitsort und Wirtschaftszweigen sowie weitere Betriebsinformationen zum Stichtag 30. Juni aus dem Betriebs-Historik-Panel (BHP). Das BHP umfasst alle Betriebe in Deutschland, die zum 30. Juni eines Jahres mindestens eine oder einen sozialversicherungspflichtig oder seit 1999 auch geringfügig Beschäftigten aufweisen. Quelle des BHP ist die Beschäftigtenhistorik (BeH). Im BHP werden die Beschäftigungsdaten der BeH mit Hilfe der Betriebsnummern auf die Betriebs-Jahrebene aggregiert. Für Ostdeutschland liegen diese administrativen Betriebsdaten erst ab 1992 vor. Nähere Informationen zum BHP finden sich auf den Internetseiten des FDZ (https://fdz.iab.de/unsere-datenprodukte/betriebsdaten/bhp/).

Somit sind Betriebs- und Personenmerkmale organisatorisch voneinander getrennt, was die Struktur der Daten deutlich macht und Speicherplatz spart (vgl. Abbildung 1). Die Verknüpfung der Wellen des IAB-Betriebspanels und der Personendatei erfolgt über die in beiden Modulen enthaltene Betriebs-ID. Die Verknüpfung von Basisbetriebs- und Personendatei erfolgt über die Betriebs-ID und das Jahr des Datensatzes. Zu beachten ist, dass das für die Verknüpfung der Module notwendige Merkmal „betnr_liab_lm“ sowohl in den Wellen des IAB-Betriebspanels, in der

[3] Mit den Stata-Befehlen *label language en* oder *label language de* kann auf englisch- bzw. deutschsprachige Labels umgestellt werden.

Personen- als auch in der Basis-Betriebsdatei vorhanden ist. Alle Dateinamen im LIAB LM 7524 sind ebenfalls Abbildung 1 zu entnehmen.

Abbildung 1: Datenhaltung der Linked-Employer-Employee-Daten im LIAB LM 7524



Anmerkung: Das LIAB LM 7524 ist für vier Basisperioden verfügbar: 1993-2000 (B9300), 2000-2007 (B0007), 2007-2017 (B0717), 2014-2024 (B1424). Diese stehen für die Befragungswellen des IAB-Betriebspanels, die der Ziehung der Personendaten zugrunde liegen; siehe Abschnitt 3.3. Der Platzhalter „BXXXX" in den Dateinamen ist entsprechend zu ergänzen. Durch die oder den Forschenden ist jeweils die Basisperiode auszuwählen, die für die Durchführung des Projektes am besten geeignet ist. Es ist nicht möglich, mehrere Basisperioden zu beantragen oder Basisperioden zu kombinieren bzw. zu erweitern.

Die Verknüpfung der Betriebs- und Personendateien erfolgt über die programmspezifischen Kommandos der zur Datenaufbereitung und -analyse genutzten Softwarepakete. In Stata können die Daten beispielsweise unter Verwendung des „`merge`"-Befehls und der Jahresangabe sowie der Betriebs-IDs in Verbindung mit der entsprechenden Pfadangabe verknüpft werden (siehe Kasten 1). Es gilt zu beachten, dass das Beispiel in Kasten 1 lediglich zur Illustration dient. In der Praxis werden die Personen-, administrativen Betriebs- und Befragungsdaten üblicherweise zuerst separat aufbereitet, bevor sie zusammengespielt werden.

Kasten 1: Beispielcode zur Verbindung von Personen- und Betriebsdaten in Stata 19

```
** Verknüpfung der Daten am Beispiel der Basisperiode 2014 bis 2024
(B1424) [alternativ: B9300, B0007 oder B0717]


* Laden des IAB-Betriebspanels (hier: Welle 2024)

use betnr_liab_lm x using $orig/LIAB_LM_7524_B1424_v1_bp_v1_2024.dta,
clear      //  x=für  die  Analyse  benötigte  Variablen  aus  den
Befragungsdaten der Betriebe für das Jahr 2024
```

```
* Anspielen der Personendatei
merge 1:m betnr_liab_lm using $orig/LIAB_LM_7524_B1424_v1.dta, keepusing(persnr_liab_lm begepi y) // y=für die Analyse benötigte Variablen aus der Personendatei
drop _merge

* Anspielen der Basis-Betriebsdatei des BHP
gen int jahr = year(begepi)
sort betnr_liab_lm jahr
merge m:1 betnr_liab_lm jahr using $orig/LIAB_LM_7524_B1424_v1_bhp_basis_v1.dta, keepusing(z) // z=für die Analyse benötigte Variablen aus den administrativen Betriebsdaten
drop if _merge==2 // Administrative Daten von Betrieben entfernen, in denen keine Person der aktuellen Auswahl beschäftigt ist
drop _merge
```

1.2.3 Sensible Merkmale und Zusatzvariablen

Bestimmte, sogenannte sensible Merkmale, die eine Deanonymisierung von Personen oder Betrieben erleichtern könnten, werden nur dann im Original weitergegeben, wenn es für das Analyseziel zwingend notwendig ist und dies im Antrag auf den Datenzugang stichhaltig begründet wird. Es ist zudem zu beachten, dass die sensiblen Merkmale in der Regel in vergröberter Form bereits in den Daten enthalten sind. Die sensiblen Merkmale werden nur dann detailliert weitergegeben, wenn die Information im vergröberten Zustand zur Erreichung des Forschungsziels nicht ausreichend ist. Bei der Bereitstellung eines Datensatzes mit sensiblen Merkmalen werden die entsprechenden nicht-sensiblen Merkmale gelöscht, wenn diese von den Nutzenden selbst wieder generiert werden können, z. B. Bundesland aus Kreisinformationen.

Diese aus datenschutzrechtlicher Sicht besonders sensiblen Merkmale sind:

<u>Personendatei:</u>

- Staatsangehörigkeit (nation)
- Geburtsmonat (gebmon)
- Berufsuntergruppe (beruf2010_4)
- Wohnort - Arbeitsagentur (wo_aa)
- Wohnort - Kreis (wo_kreis)

<u>Basis-Betriebsdatei (einzelne Merkmale):</u>

- Arbeitsort - Kreis (ao_kreis)
- Wirtschaftszweig - Wirtschaftsunterklasse (w93_5)
- Wirtschaftszweig - Wirtschaftsunterklasse (w03_5)
- Wirtschaftszweig - Wirtschaftsunterklasse (w08_5)

Auf begründeten Antrag können aus dem BHP weitere Betriebsmerkmale in Variablenblöcken bereitgestellt werden (vgl. Abbildung 1). Zudem können die Erweiterungsmodule Beschäftigungsströme (Eintritte/Austritte) und Betriebsdynamik (Gründungen/Schließungen) mit gesonderter Begründung beantragt werden.

Diese Zusatzblöcke bzw. -module sind:

BHP-Variablenblöcke in „Jahresscheiben“:

- Struktur der Beschäftigten allgemein (az_f az_reg az_azubi az_atz az_tz az_f_vz az_f_tz az_reg_vz)
- Struktur der Beschäftigten nach Schul- und Berufsbildung (az_gq az_mq az_hq az_gq_vz az_mq_vz az_hq_vz)
- Altersstruktur der Beschäftigten (az_15_19 az_15_19_vz az_20_24 az_20_24_vz az_25_29 az_25_29_vz az_30_34 az_30_34_vz az_35_39 az_35_39_vz az_40_44 az_40_44_vz az_45_49 az_45_49_vz az_50_54 az_50_54_vz az_55_59 az_55_59_vz az_60_64 az_60_64_vz az_ab65 az_ab65_vz alter_mw alter_mw_vz)
- Aktivitäten Forschung und Entwicklung (az_ingnat)
- Anzahl atypisch Beschäftigter (az_leih az_bfr)
- Struktur der Beschäftigten nach Staatsangehörigkeit (az_d az_d_vz az_eu)
- Struktur der Beschäftigten nach Blossfeld Berufsqualifikation (az_bf_agr az_bf_emb az_bf_edi az_bf_evb az_bf_qmb az_bf_qdi az_bf_qvb az_bf_tec az_bf_semi az_bf_ing az_bf_prof az_bf_man)
- Struktur der Beschäftigten nach Tätigkeitsniveau (az_niv1 az_niv2 az_niv3 az_niv4)
- Gehaltsstruktur der Vollzeitbeschäftigten (az_zens te_imp_med te_imp_p25 te_imp_p75 te_imp_mw_f te_imp_med_f te_imp_med_m te_imp_mw_gq te_imp_med_gq te_imp_mw_mq te_imp_med_mq te_imp_mw_hq te_imp_med_hq te_imp_med_uq te_imp_mw_d te_imp_med_d te_imp_med_a te_imp_sum)

BHP-Erweiterungsmodule:

- Beschäftigtenströme (ein_ges ein_f ein_reg ein_gf ein_azubi ein_vz ein_tz ein_f_vz ein_f_tz ein_reg_vz ein_wdr ein_wdr_f ein_bw ein_bw_f ein_bf_agr ein_bf_emb ein_bf_edi ein_bf_evb ein_bf_qmb ein_bf_qdi ein_bf_qvb ein_bf_tec ein_bf_semi ein_bf_ing ein_bf_prof ein_bf_man ein_15_19 ein_20_24 ein_25_29 ein_30_34 ein_35_39 ein_40_44 ein_45_49 ein_50_54 ein_55_59 ein_60_64 ein_ab65 aus_ges aus_f aus_reg aus_gf aus_azubi aus_vz aus_tz aus_f_vz aus_f_tz aus_reg_vz aus_temp aus_temp_f aus_bw aus_bw_f aus_bf_agr aus_bf_emb aus_bf_edi aus_bf_evb aus_bf_qmb aus_bf_qdi aus_bf_qvb aus_bf_tec aus_bf_semi aus_bf_ing aus_bf_prof aus_bf_man aus_15_19 aus_20_24 aus_25_29 aus_30_34 aus_35_39 aus_40_44 aus_45_49 aus_50_54 aus_55_59 aus_60_64 aus_ab65 aus_senio_1 aus_senio_2 aus_senio_3)
- Betriebsdynamik (eintritt besch besch_vor status_vor inflow austritt besch besch_nach status_nach outflow)

Nähere Informationen zum BHP finden sich unter https://fdz.iab.de/betriebsdaten/betriebs-historik-panel-bhp-version-7524-v1/ und im FDZ-Datenreport 13/2025 (Ganzer et al., 2025).

Zudem stehen für das LIAB LM 7524 personen- und betriebsspezifische Effekte (AKM-Effekte) für die Jahre 1985 bis 2023 zur Verfügung (Dateinamen: LIAB_LM_7524_BXXXX_v1_akm_estab_8523_v1.dta; LIAB_LM_7524_BXXXX_v1_akm_pers_8523_v1.dta). Weitere Informationen finden Sie im FDZ-Methodenreport 03/2025 (Lochner/Wolter, 2025). Bei Neuanträgen können die AKM-Effekte über das Feld für die sensiblen Merkmale und Ergänzungsmodule beantragt werden. Bei bestehenden Projekten genügt eine formlose E-Mail an iab.fdz@iab.de.

1.3 Veränderungen zum LIAB LM 7521

1.3.1 Stichprobe und Verknüpfung

Das LIAB LM wurde von Grund auf neu konzipiert. In den früheren Versionen wurde eine Subgruppe von wiederholt befragten „Panelbetrieben“ aus einer konkreten Befragungswelle ausgewählt. Für diese Betriebe wurden dann die Daten aller Beschäftigten eines bestimmten „Panelzeitraums“ verknüpft. Diese Art der Ziehung schränkt sowohl das Analysepotential als auch die Repräsentativität des Datensatzes ein.

Im Gegensatz hierzu werden im LIAB LM 7524 immer alle Betriebe einer Befragungswelle für die Verknüpfung zugrunde gelegt. Um die Daten handhabbar zu halten aber weiterhin Panelanalysen zu ermöglichen, werden mehrere Wellen einer sogenannten Basisperiode kombiniert und dann die darin befragten Betriebe verknüpft, sofern eine entsprechende Erlaubnis vorliegt. Aufgrund von Größenbeschränkungen können alle Beschäftigten dieser Betriebe allerdings nur noch dann bereitgestellt werden, wenn ein Betrieb eine bestimmte Größe nicht überschreitet. Für die übrigen Betriebe wird stattdessen eine Zufallsstichprobe der Beschäftigten pro Betrieb bereitgestellt. Details finden sich in Abschnitt 3.3.

1.3.2 Datenquellen und Merkmalskanon

Aufgrund von Größenbeschränkungen mussten im LIAB LM 7524 die administrativen Personendaten auf die Quelle BeH eingeschränkt werden. Variablen, die keine Informationen aus der BeH enthalten, wurden entfernt.

Die Variable „Status der Betriebsnummer (betr_st)“ wurde neu definiert, um der veränderten Ziehungslogik zu entsprechen. Sie gibt nun an, ob ein Betrieb für die Ziehung der Personendaten ausgewählt wurde, oder nicht. Wurde er ausgewählt, dann zeigt sie außerdem an, ob alle Beschäftigten der jeweiligen Basisperiode (und ein Jahr davor) dieses Betriebes gezogen wurden, oder ob eine Zufallsstichprobe gezogen wurde.

In den finalen Daten sind die Erwerbsbiografien der verknüpften Personen auch dann enthalten, wenn sie in Betriebe wechseln, aus denen sie nicht bei der Verknüpfung der Daten gezogen wurden. Um dies sichtbar zu machen, wurde eine weitere Variable „Beobachtung wurde gezogen (smpl_st)“ aufgenommen, die anzeigt, wenn eine Beobachtung einer Person zu einem Betrieb gehört, für den diese Person ursprünglich verknüpft wurde. In Verbindung mit der Variable „betr_st“ kann dann erkannt werden, ob eine Person zu einem Betrieb gehört, für den die gesamte Belegschaft der Basisperiode (und ein Jahr davor) gezogen wurde, ob sie Teil der 20 Prozent Stichprobe eines Betriebs ist oder ob sie zu einem Zeitpunkt vor oder nach der eigentlichen Ziehung in einem anderen Betrieb tätig war, der entweder nicht für die Verknüpfung ausgewählt

wurde oder zu dessen Stichprobe die Person nicht gehört hat. Da „smpl_st“ nur für die Jahre der Basisperiode (und ein Jahr davor) definiert ist, kann diese Variable auch verwendet werden, um die Personendaten auf den Ziehungszeitraum einzugrenzen.

Weiterhin wurden im LIAB LM 7524 gegenüber der Version 7521 neue Merkmale zum Tagesentgelt aufgenommen, die bei der Anrechnung von Einmalzahlungen unterstützen und eine Imputation der zensierten Werte rechts der Beitragsbemessungsgrenze[4] anbieten. Die zugrundeliegenden Datenarbeiten werden in Drechsler et al. (2023) beschrieben. Weitere Informationen zu den neuen Merkmalen finden sich unter „Tagesentgelt (inkl. Einmalzahlungen) (tentgelt_bonus)“ und „Tagesentgelt (imputiert) (tentgelt_imp)“.

Die beiden Merkmale „Erstes Auftreten Betriebsnummer (grd_dat)“ und „Letztes Auftreten Betriebsnummer (lzt_dat)“ werden nicht länger angeboten, jahresgenaue Angaben zum ersten und letzten Auftreten eines Betriebes sind aber weiter vorhanden.

Tabelle 1 gibt einen Überblick über die Änderungen im Merkmalskanon.

Tabelle 1: Neue, entfallene und veränderte Merkmale im LIAB LM 7524

Merkmal	**Erläuterung**
Tagesentgelt (inkl. Einmalzahlungen) (tentgelt_bonus)	Neues Merkmal (siehe Abschnitt 5.5.2)
Tagesentgelt (imputiert) (tentgelt_imp)	Neues Merkmal (siehe Abschnitt 5.5.3)
Quelle des Satzes (quelle)	Wird nicht länger angeboten, da nur noch Quelle BeH enthalten
Familienstand (famst)	Wird nicht länger angeboten, da nur noch Quelle BeH enthalten
Kinderzahl (kind)	Wird nicht länger angeboten, da nur noch Quelle BeH enthalten
Restanspruch/geplante Dauer (restanspruch)	Wird nicht länger angeboten, da nur noch Quelle BeH enthalten
Trägerart (traeger)	Wird nicht länger angeboten, da nur noch Quelle BeH enthalten
Beginndatum der Arbeitslosigkeit (alo_beg)	Wird nicht länger angeboten; „Dauer der Arbeitslosigkeit (alo_dau)“ weiter vorhanden
Erstes Auftreten Betriebsnummer (grd_dat)	Wird nicht länger angeboten; „Jahr erstes Auftreten Betriebsnummer (grd_jahr)“ weiter vorhanden
Letztes Auftreten Betriebsnummer (lzt_dat)	Wird nicht länger angeboten; „Jahr letztes Auftreten Betriebsnummer (lzt_jahr)“ weiter vorhanden
Status der Betriebsnummer (betr_st)	Neue Definition der Variable zur Kennzeichnung der Ziehungsbetriebe je Basisperiode (siehe Abschnitt 5.2.1)
Beobachtung wurde gezogen (smpl_st)	Neue Variable zur Kennzeichnung, welche Beobachtung einer Person ursprünglich verknüpft wurde (siehe Abschnitt 5.2.2)

1.3.3 Ziehung der Daten aus dem BHP

In früheren Versionen des LIAB LM waren Informationen zu Betrieben aus dem BHP nur für die Jahre verfügbar, in denen eine Person aus der LIAB-Stichprobe dort beschäftigt war. Im LIAB LM

[4] Die von Stüber et al. (2023) bereitgestellten Skripte können zur Aufbereitung der Personendaten des LIAB LM genutzt werden. Das darin enthaltene Verfahren zur Imputation der Löhne kann nun durch die neu angebotenen Variablen ersetzt werden.

7524 hingegen sind sämtliche Informationen aus dem BHP enthalten, wenn ein Betrieb zumindest einmal an der Befragung des IAB-Betriebspanels teilgenommen hat. Für alle anderen Betriebe gilt die Einschränkung weiterhin.

1.3.4 Datendokumentation

Im vorliegenden Datenreport wurden die meisten Abschnitte aufgrund der Neukonzeption des Datenprodukts stark überarbeitet.

Zudem wurden die Beschreibungen für folgende Merkmale überarbeitet bzw. ergänzt:

- Status der Betriebsnummer (betr_st)
- Beobachtung wurde gezogen (smpl_st)
- Jahr (jahr)
- Tagesentgelt (inkl. Einmalzahlungen) (tentgelt_bonus)
- Tagesentgelt (imputiert) (tentgelt_imp)
- Gleitzone (gleitz)
- Wohnort - Kreis (wo_kreis)
- Wohnort - Arbeitsagentur (wo_aa)
- Jahr letztes Auftreten Betriebsnummer (lzt_jahr)
- Arbeitsort - Kreis (ao_kreis)

Der ehemalige Abschnitt „1.5 Merkmalsliste“ wurde gelöscht. Als Ersatz dienen die detaillierten Übersichten zu den einzelnen Merkmalen (Label, Auszählungen, fehlende Werte) im Online-Anhang zum Datenreport (siehe Abschnitt 7.1).

1.4 Steckbrief

Tabelle 2: Steckbrief

Kategorie	Beschreibung
Themen/ Merkmals-gruppen	**Beschäftigtenhistorik (BeH)**: Jahres- und Abmeldungen zur Sozialversicherung für sozialversicherungspflichtig und geringfügig Beschäftigte **IAB-Betriebspanel**: Informationen aus jährlicher Betriebsbefragung mit wechselnden Schwerpunktthemen
Untersuchungs-einheit	Sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Beschäftigungsbetriebe
Fallzahlen	Die Fallzahlen sind abhängig vom Ziehungszeitraum. Verknüpfte Betriebe: 13.264 bis 40.024 pro Jahr, 18.505 bis 47.077 insgesamt Personen: 151.733 bis 3.014.593 pro Jahr, 2.338.291 bis 3.772.719 insgesamt Überschneidungsfreie Sätze (nach Episodensplitting): 87.124.938 bis 132.486.139

Kategorie	Beschreibung
Zeitraum	Der Zeitraum hängt von der Datenquelle ab. BeH: 01.01.1975 - 31.12.2024 (bis 2021: 36-Monats-Datei 2022 und 2023: 18-Monats-Datei, 2024: 6-Monats-Datei) IAB-Betriebspanel: Wellen 1993 - 2024
Zeitbezug	**Betriebe**: Stichtag 30. Juni (BHP) bzw. variierend (IAB-Betriebspanel) **Personen**: tagesgenau
Regionale Gliederung	Bundesländer, Kreise
Gebietsstand	Gebietsstandsbereinigt zum 31.12.2024
Erhebungsdesign	Linked-Employer-Employee-Daten; Auswahl bestimmter Betriebe aus Betriebsbefragung und Zuspielen der Erwerbsbiografien ihrer Beschäftigten aus Prozessdaten.
Beteiligte Institutionen	Sozialversicherungsträger, Bundesagentur für Arbeit, kommunale Träger Auftraggeber IAB-Betriebspanel: Forschungsbereich "Betriebe und Beschäftigung" des IAB, Ministerien/Senatsverwaltungen der Bundesländer und auswertende Institute. Durchführung IAB-Betriebspanel: Verian (ehemals Kantar Public) in Zusammenarbeit mit dem Institut SÖSTRA Sozialökonomische Strukturanalysen GmbH und dem Institut GEFRA – Gesellschaft für Finanz- und Regionalanalysen GbR. Partnerinstitute: https://iab.de/das-iab/befragungen/iab-betriebspanel/beteiligte-partnerinstitute-in-den-bundeslaendern/
Frequenz der Datensammlung	**Administrative Daten**: laufend **Daten des IAB-Betriebspanels**: einmal jährlich
Dateiformat/ -größe	Die Dateigrößen sind abhängig vom Ziehungszeitraum. Im Stata-Format: Betriebsdaten (ohne sensible Merkmale): IAB-Betriebspanel: ca. 1,6 bis 9,9 MB pro Jahr Basis-Betriebsdatei: ca. 833 bis 1.129 MB Personendatei (ohne sensible Merkmale): ca. 5,9 bis 9,0 GB
Datei-organisation	Die Personendaten werden in einer Datei bereitgestellt. Hinzu kommt eine Datei mit den betriebsbezogenen Informationen aus dem BHP für alle Jahre (die Basis-Betriebsdatei). Auf begründeten Antrag werden weitere Dateien mit zusätzlichen Betriebsinformationen aus dem BHP bereitgestellt. Die Dateien des IAB-Betriebspanels beschränken sich auf Betriebe, für die eine Zustimmung zur Datenverknüpfung vorliegt, und werden als eine Datei pro Jahr bereitgestellt.
Datenzugang	Remote Desktop, Gastaufenthalt oder Datenfernverarbeitung
Anonymisierungsgrad	Schwach anonymisiert

Kategorie	Beschreibung
Sensible Merkmale	**Personendatei**: Geburtsmonat (gebmon), Staatsangehörigkeit (nation), Berufsuntergruppe (beruf2010_4), Wohnort – Kreis (wo_kreis), Wohnort – Arbeitsagentur (wo_aa) **Basis-Betriebsdatei**: Arbeitsort - Kreis (ao_kreis), Wirtschaftsunterklasse (w93_5), Wirtschaftsunterklasse (w03_5), Wirtschaftsunterklasse (w08_5)
Zitierung der Daten und Datendokumentation	**Daten**: Filser, Andreas; Graf, Tobias; Grießemer, Stephan; Köhler, Markus; Moczall, Andreas; Oertel, Martina; Schmucker, Alexandra; Schneider, Andreas; Umkehrer, Matthias; Viehring, Joshua; Walz, Hannes (2026): "Linked-Employer-Employee-Daten des IAB, Längsschnittmodell (LIAB LM) – Version 7524 v1". Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). DOI: 10.5164/IAB.LIABLM7524.de.en.v1 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. **Datendokumentation**: Filser, Andreas; Schmucker, Alexandra; Umkehrer, Matthias (2026): Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell (LIAB LM) 1975-2024. FDZ-Datenreport, 09/2026 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2609.de.v1
Datensatzversion	Linked-Employer-Employee-Daten des IAB, Längsschnittmodell 1975-2024 (LIAB LM 7524) – Version 7524 v1; DOI: 10.5164/IAB.LIABLM7524.de.en.v1

1.5 Mengengerüst

Die folgende Aufstellung zeigt ausgewählte Fallzahlen, separat für jede der vier Basisperioden.

Basisperiode: 1993 bis 2000 (B9300)

- Betriebe mit Verknüpfung zu ihren Beschäftigten innerhalb der Basisperiode (und ein Jahr davor): 13.264 bis 17.338 pro Jahr, 18.505 insgesamt
    - mit Verknüpfung aller Beschäftigten: 10.992 insgesamt
    - mit Verknüpfung einer 20 Prozent Stichprobe: 7.513 insgesamt
- Personen: 489.684 bis 1.827.108 pro Jahr, 2.338.291 insgesamt
- Überschneidungsfreie Sätze (nach Episodensplitting): 87.124.938 insgesamt

Basisperiode: 2000 bis 2007 (B0007)

- Betriebe mit Verknüpfung zu ihren Beschäftigten innerhalb der Basisperiode (und ein Jahr davor): 28.068 bis 30.102 pro Jahr, 34.574 insgesamt
    - mit Verknüpfung aller Beschäftigten: 23.279 insgesamt
    - mit Verknüpfung einer 20 Prozent Stichprobe: 11.295 insgesamt
- Personen: 450.699 bis 2.557.938 pro Jahr, 3.213.863 insgesamt

- Überschneidungsfreie Sätze (nach Episodensplitting): 126.804.808 insgesamt

Basisperiode: 2007 bis 2017 (B0717)

- Betriebe mit Verknüpfung zu ihren Beschäftigten innerhalb der Basisperiode (und ein Jahr davor): 28.428 bis 32.635 pro Jahr, 39.636 insgesamt
    - mit Verknüpfung aller Beschäftigten: 29.001 insgesamt
    - mit Verknüpfung einer 20 Prozent Stichprobe: 10.635 insgesamt
- Personen: 274.406 bis 2.830.039 pro Jahr, 3.531.495 insgesamt
- Überschneidungsfreie Sätze (nach Episodensplitting): 132.486.139 insgesamt

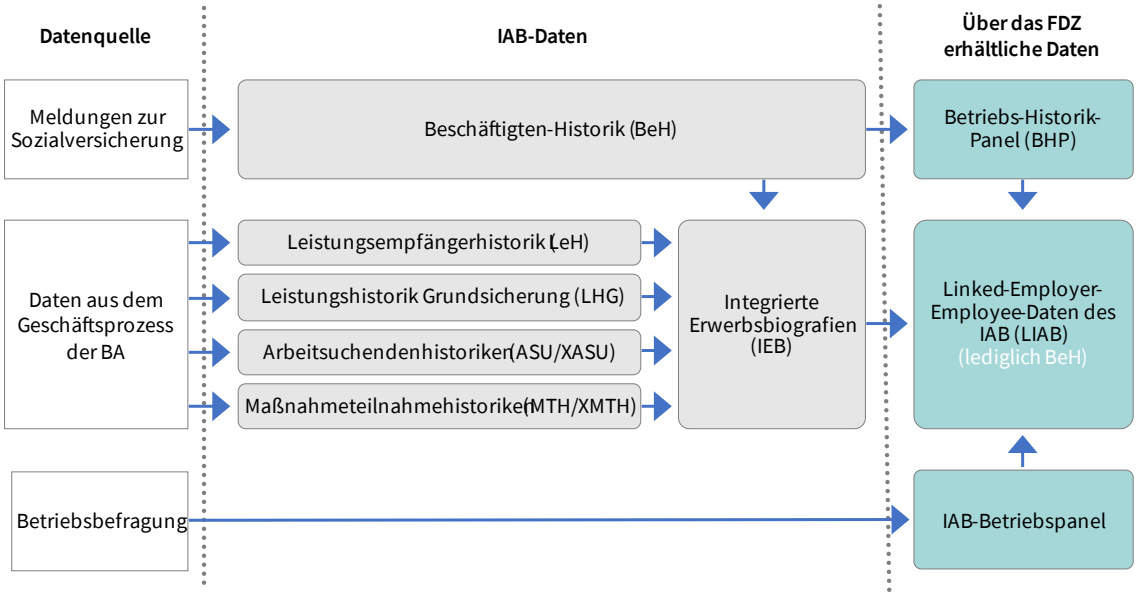
Basisperiode: 2014 bis 2024 (B1424)

- Betriebe mit Verknüpfung zu ihren Beschäftigten innerhalb der Basisperiode (und ein Jahr davor): 34.966 bis 40.024 pro Jahr, 47.077 insgesamt
    - mit Verknüpfung aller Beschäftigten: 35.409 insgesamt
    - mit Verknüpfung einer 20 Prozent Stichprobe: 11.668 insgesamt
- Personen: 151.733 bis 3.014.593 pro Jahr, 3.772.719 insgesamt
- Überschneidungsfreie Sätze (nach Episodensplitting): 122.902.641 insgesamt

2 Datenquellen

Die administrativen Personendaten werden aus den Integrierten Erwerbsbiografien (IEB) des IAB gezogen. 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 zum LIAB LM führen und stellt deren Beziehung zu anderen Datenprodukten des FDZ dar.

Abbildung 2: Quellen des LIAB LM



Die Quelle für Informationen über Beschäftigung ist die Beschäftigtenhistorik (BeH) des IAB. Die Grundlage der Daten ist das mit Wirkung vom 01.01.1973 eingeführte (und seit dem 01.01.1991 auf Ostdeutschland ausgedehnte) integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung (DEÜV; früher DEVO/DÜVO; 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 an die zuständigen Sozialversicherungsträger abzugeben. Die BeH erfasst alle Arbeiter*innen und Angestellten sowie alle Auszubildenden, soweit sie nicht von der Sozialversicherungspflicht befreit sind. Dies bedeutet, dass verbeamtete und selbstständige Tätigkeiten sowie Studierende[5] (vgl. Cramer, 1985) grundsätzlich nicht in der BeH erfasst werden. Seit der Änderung des Meldeverfahrens zum 01.01.1999 werden auch geringfügig Beschäftigte und mithelfende Familienangehörige erfasst (enthalten erst ab 01.04.1999). Die Daten werden von den Krankenkassen aufgenommen, von der Bundesagentur für Arbeit gesammelt und aufbereitet und anschließend vom IAB in dessen Historikdatei integriert.

Die Beschäftigungsepisoden in den administrativen Personendaten werden durch Betriebsdaten (Basisbetriebsdaten und BHP-Erweiterungsmodule) ergänzt. Diese werden dem Betriebs-Historik-Panel (BHP) entnommen, das ebenso auf der BeH basiert. Für alle Betriebe, die mindestens einmal Teil des IAB-Betriebspanels sind, werden sämtliche verfügbare Informationen aus den Basisbetriebsdaten und BHP-Erweiterungsmodulen bereitgestellt. Für alle anderen Betriebe umfassen die Daten aus dem BHP nur Angaben zu Betrieben, für die mindestens eine beschäftigte Person in der LIAB-LM-Stichprobe enthalten ist und für die Jahre, in denen die beschäftigten Personen in der LIAB-LM-Stichprobe vorzufinden sind.

Bei der Verknüpfung von Personen- mit Betriebsdaten muss berücksichtigt werden, dass die Merkmale in der Basis-Betriebsdatei sowie den BHP-Erweiterungsmodulen zum 30.06. eines

[5] Studierende können dennoch in der BeH erscheinen, wenn sie parallel zum Studium z.B. geringfügig beschäftigt waren.

Jahres aggregiert werden. Die Betriebsmerkmale sind deshalb, im Unterschied zu den Personenangaben, keine zeitraumbezogenen Daten, sondern nur am 30.06. jedes Jahres exakt gültig.[6]

Ein Beispielcode zur Verknüpfung von Personen- und Betriebsdaten findet sich in Abschnitt 1.2.2.

3 Datenaufbereitung und Stichprobenziehung

3.1 Bereinigungen und Aufbereitungen

Bevor die Daten aus den in Kapitel 2 genannten Quellen zur IEB zusammengefügt werden, durchlaufen sie quellenspezifische Bereinigungsverfahren (siehe 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.

Die Daten der BeH werden wie folgt aufbereitet und bereinigt:

- 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 ab 2011 neu eingeführt.
- Aus datenschutzrechtlichen Gründen werden die Personengruppen 107, 111, 113, 114, 127 und 204 zur Personengruppe "Sonstige Erwerbstätige“ (599) zusammengefasst.
- Ab dem Meldejahr 2012 wurden Auszubildende mit den Personengruppen 121 und 122 neu aufgenommen.

[6] Extrembeispiel: Es liegen Beschäftigungsmeldungen vom 1.1. bis zum 30.05.2006 vor, der Betrieb meldet jedoch Konkurs an, weshalb alle Beschäftigten entlassen werden. Dann gibt es im BHP zu diesem Betrieb für 2006 keine Informationen.

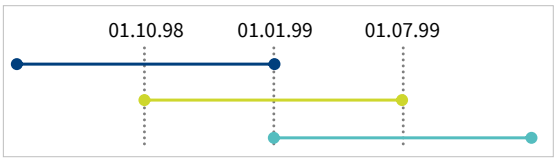
- Sätze mit fehlendem Entgelt sowie der Ausprägung 101 beim Merkmal Erwerbsstatus und dem Wert 50 beim Abgabegrund (Jahresmeldung) werden nicht in die IEB übernommen.
- Geschlecht und Geburtsdatum werden aus dem Data Warehouse (DWH) der Statistik der BA übernommen. Dort liegen diese Angaben quellenübergreifend harmonisiert vor.
- Der Gebietsstand für Arbeitsort und Wohnort entspricht dem 31.12.2024.

3.2 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).

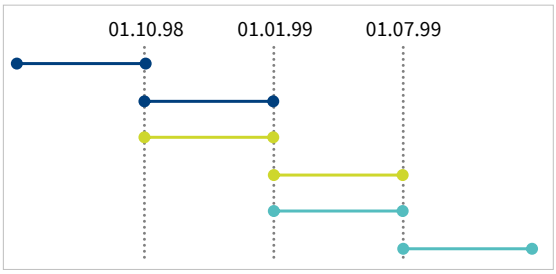
Die ursprünglichen Datumsvariablen zum Beginn und Ende des Originalsatzes (begorig und endorig) bleiben erhalten, die Variablen Beginn und Ende der Episode (begepi und endepi) kennzeichnen Anfang und Ende der gesplitteten Episoden. Ob Sätze gesplittet wurden, kann an einem Vergleich des ursprünglichen Zeitraums (begorig und endorig) mit dem Episodenzeitraum (begepi und endepi) festgestellt werden. Um die Originaldaten ohne Episodensplitting wiederherzustellen bzw. die durch das Episodensplitting künstlich erzeugten Datensätze zu löschen, selektiert man alle Sätze, bei denen der Beginn des Originalsatzes gleich dem Beginn der gesplitteten Episode ist (`begepi == begorig`).

Abbildung 3: Episodensplitting



persnr	*begorig*	*endorig*	*level*
4711	01.01.98	31.12.98	-
4711	01.10.98	30.06.99	-
4711	01.01.99	31.12.99	-

Episodensplitting
Aufteilung der Zeitintervalle in sich nicht überschneidende Intervalle



persnr	*begepi*	*endepi*	*level*
4711	01.01.98	30.09.98	0
4711	01.10.98	31.12.98	0
4711	01.10.98	31.12.98	1
4711	01.01.99	30.06.99	0
4711	01.01.99	30.06.99	1
4711	01.07.99	31.12.99	0

Es ist ratsam, die durch das Splitting erzeugten, vollständig parallelen Sätze einheitlich zu sortieren. Die früher in anderen Datenprodukten, wie z. B. der SIAB, enthaltene Variable Satzzähler pro Episode (level2) kann bei Bedarf mit folgendem Stata-Befehl selbst erzeugt werden:

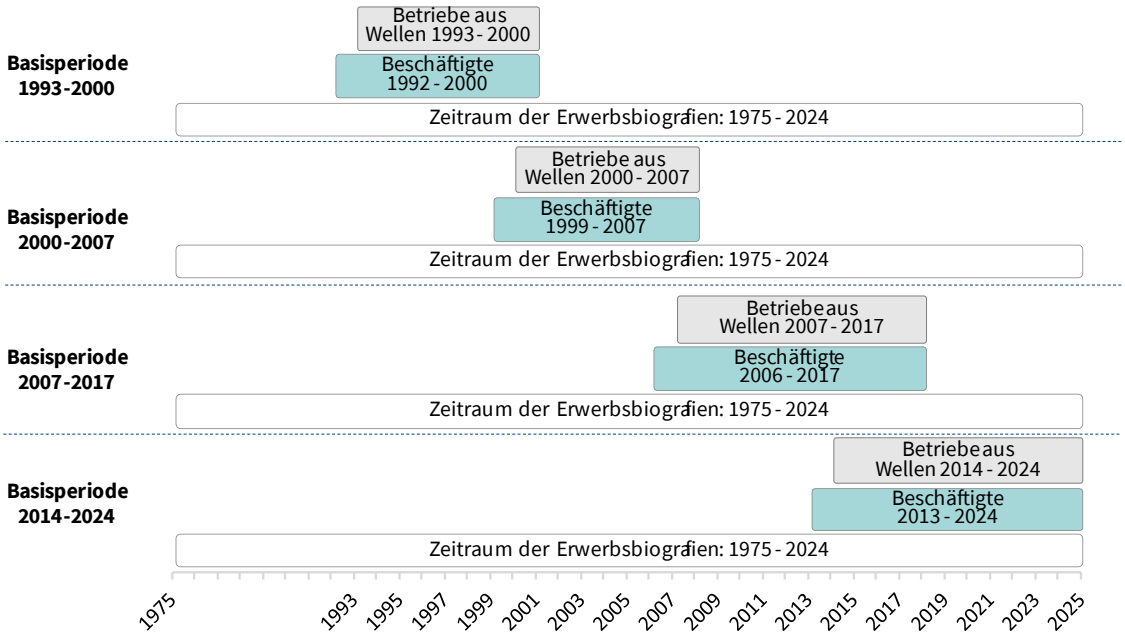
Kasten 2: Beispielcode zum Erstellen weiterer Satzzähler in Stata 19

```
bysort persnr_liab_lm begepi (spell): gen byte level2 = _n-1
```

3.3 Stichprobenziehung

Die Befragung des IAB-Betriebspanels (Version 9324 v1) bildet den Ausgangspunkt für die Ziehung der Personendaten des LIAB LM 7524. Hieraus werden alle Betriebe ausgewählt, die in der jeweiligen Basisperiode Teil des IAB-Betriebspanels sind und für die eine Zustimmung zur Verknüpfung der Daten vorliegt. Für diese Betriebe werden aus den IEB alle Beschäftigten ermittelt, die in der Basisperiode oder ein Jahr davor mindestens einen Tag in einem der Betriebe beschäftigt waren. Die verschiedenen Basisperioden und die zugehörigen Ziehungszeiträume sind in Abbildung 4 dargestellt.

Abbildung 4: Zeitdimension im LIAB LM 7524



Aufgrund von Größenbeschränkungen können jedoch nicht die kompletten Erwerbsbiografien von all diesen Personen angeboten werden. Einerseits ist es nötig, die Personendaten auf Beschäftigungsmeldungen aus der Quelle BeH einzuschränken. Andererseits können nur für „kleinere“ Betriebe die administrativen Personendaten zu allen Beschäftigten innerhalb der Basisperiode (und ein Jahr davor) bereitgestellt werden.

Die Datenverknüpfung erfolgt in vier Schritten: Zuerst werden die Betriebe einer Basisperiode anhand der Größenklassen-Angabe aus der Schichtungsmatrix des IAB-Betriebspanels in „kleine“ und „große“ Betriebe eingeteilt, wobei Betriebe mit weniger als 50 sozialversicherungspflichtig Beschäftigten als Kleinbetriebe und Betriebe mit mindestens 50 sozialversicherungspflichtig Beschäftigten als Großbetriebe definiert werden. Wird ein Betrieb einmal als Kleinbetrieb klassifiziert, dann behält er diesen Status auch dann bei, wenn er in einer anderen Befragungswelle

als Großbetrieb gilt. Hierdurch wird sichergestellt, dass sämtliche Analysen im Bereich der Schichtungsmatrix mit Betrieben unterhalb von 50 sozialversicherungspflichtig Beschäftigten mit den Personendaten der gesamten Belegschaft durchgeführt werden können.

Im zweiten Schritt werden alle Personen, die mindesten einen Tag in der Basisperiode (und ein Jahr davor) in einem der Betriebe sozialversicherungspflichtig oder geringfügig beschäftigt waren, aus den IEB gezogen. Für die Kleinbetriebe gehen alle so ermittelten Personen in das LIAB LM ein, für die Großbetriebe wird eine 20 Prozent Zufallsstichprobe der Personen pro Betrieb gezogen. Es gibt jedoch auch Personen, die im Laufe einer Basisperiode sowohl in einem der enthaltenen Klein- wie auch in einem Großbetrieb beschäftigt waren. Diese Personen werden von der Stichprobenziehung der Großbetriebe ausgeschlossen, da sie über die Zugehörigkeit zu einem Kleinbetrieb ohnehin in den Personendaten des LIAB LM enthalten sind. Außerdem gibt es Personen, die für die Ziehung der Stichprobe aus einem Betrieb ausgewählt wurden, im weiteren Verlauf dann aber für einen anderen Betrieben tätig werden, für den ebenfalls eine Stichprobe gezogen wurde, die Person aber nicht Teil dieser zweiten Stichprobe war. Um diese verschiedenen Fälle in den finalen Daten sichtbar zu machen, wurden die beiden Variablen „Status der Betriebsnummer (betr_st)“ und „Beobachtung wurde gezogen (smpl_st)“ aufgenommen. In Kombination haben diese technischen Merkmale folgende Bedeutung:

- betr_st==1 & smpl_st==1 -> alle Beschäftigten des Betriebs wurden verknüpft; die Person gehört dazu
- betr_st==1 & smpl_st==0 -> kommt nicht vor
- betr_st==2 & smpl_st==1 -> eine Stichprobe der Beschäftigten des Betriebs wurde verknüpft; die Person gehört dazu
- betr_st==2 & smpl_st==0 -> eine Stichprobe der Beschäftigten des Betriebs wurde verknüpft, die Person gehört nicht dazu
- betr_st==3 & smpl_st==1 -> kommt nicht vor
- betr_st==3 & smpl_st==0 -> der Betrieb ist nur im Datensatz, weil eine verknüpfte Person einmal dort beschäftigt war, der Betrieb selbst wurde aber nie direkt verknüpft

Im dritten Schritt werden die tagesgenauen Erwerbsbiografien dieser Personen für die Quelle BeH aus den IEB gezogen. Die Erwerbsbiografien umfassen dabei den Zeitraum 1975 bis 2024 und enthalten somit sowohl Beschäftigungszeiten bei den Betrieben der Basisperiode als auch bei anderen Betrieben.

Im vierten Schritt werden für alle Beschäftigungsbetriebe die Angaben aus dem BHP gezogen, wobei für die Betriebe aus dem IAB-Betriebspanel alle vorliegenden Informationen seit 1975 zur Verfügung gestellt werden. Für andere Betriebe werden lediglich für die Jahre, in denen mindestens eine dort beschäftigte Person in den Daten enthalten ist, BHP-Informationen vorgehalten.

3.4 Qualität der Verknüpfung

Es ist zu beachten, dass nicht für alle Betriebe, die für die jeweilige Basisperiode potentiell verknüpfbar sind, auch Personendaten vorliegen. Somit unterscheiden sich die Fallzahlen der verknüpfbaren Betriebe aus dem IAB-Betriebspanel von den Fallzahlen der tatsächlich

verknüpften Betriebe in den Personendaten (betr_st==1 oder betr_st==2). Je nach Basisperiode können zwischen 6 und 15 Prozent der Betriebe nicht verknüpft werden. Ursache hierfür ist zum einen, dass Betriebe, für die im IAB-Betriebspanel keine Zuordnung zu den Größenklassen vorliegt, von der Ziehung der Personen ausgeschlossen werden. Zum anderen fehlen potentiell verknüpfbare Betriebe in den Personendaten, wenn keine Beschäftigten im relevanten Zeitraum in den administrativen Daten gefunden werden. Dies kann entweder bedeuten, dass es tatsächlich keine Beschäftigten gab oder dass für die vorliegenden Betriebsnummern keine Informationen vorliegen. Dieses Problem rührt insbesondere daher, dass sich Betriebsnummern, die ursprünglich einmal für die Befragung gezogen wurden, über die Zeit ändern können, sodass sie in den administrativen Daten nicht mehr gefunden werden können, u. a., weil es bei Wiederholungsbefragungen keinen Betriebsnummernabgleich gibt.

Diese Problematik zeigt sich auch bei der Verknüpfung der Befragungsdaten mit dem BHP. Auch hier gilt es zu beachten, dass nicht für jeden Betrieb des IAB-Betriebspanels eine Information aus dem BHP verfügbar ist. Somit beträgt im LIAB LM 7524 die Quote der erfolgreichen Matches zwischen IAB-Betriebspanel und der Basis-Betriebsdatei des BHP im Jahr 1993 noch 98 Prozent, fällt dann aber auf 79 Prozent im Jahr 2024. Unabhängig hiervon sollte immer berücksichtigt werden, dass die Angaben aus den administrativen Daten grundsätzlich von den Daten aus der Befragung zum IAB-Betriebspanel abweichen können.

3.5 Gewichtung

Schließlich sei darauf hingewiesen, dass die Stichprobe des IAB-Betriebspanels stratifiziert gezogen wird; nach Betriebsgrößenklasse, Branche und Bundesland. Somit sind Ergebnisse auf Grundlage dieser Befragungsdaten nur mit entsprechender Gewichtung repräsentativ. Da das IAB-Betriebspanel die Grundlage des LIAB LM bildet, sind die verknüpften Personen- und Betriebsdaten ebenfalls nicht als repräsentativ anzusehen. Um die Stratifizierung auszugleichen, werden im FDZ-Datenreport 01/2026 (Bellmann et al., 2026) zum IAB-Betriebspanel unter https://fdz.iab.de/betriebsdaten/iab-betriebspanel-iab-bp-version-9324-v1/ mehrere Möglichkeiten diskutiert. Hier sei insbesondere auf Kapitel 5 verwiesen.

3.6 Fehlende Werte

In den Daten gilt die folgende Codierung von fehlenden Werten:

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.

4 Datenqualität und Probleme

4.1 Gesamte IEB

Die IEB beinhalten Erwerbsbiografien. Allerdings ist in den administrativen Daten nicht jede Art von Erwerbstätigkeit erfasst. Bei entsprechenden Lebensverläufen sind bestimmte Personen überhaupt nicht in den IEB abgebildet.

Für Auswertungszwecke ist es oft relevant, Lücken in den abgebildeten Biografien zu kennen (z. B. für Kontrollgruppenbildung, Lebensverlaufsanalysen etc.). Die im Folgenden aufgeführten Lücken sind definiert als Zeiträume nach Ende der Schulausbildung, für die kein Satz in den IEB enthalten ist. Diese Lücken können unterschieden werden in

- Lücken, für die keinerlei Information vorliegt und
- Lücken, für die anhand des Merkmals Abgabegrund/Beendigungsgrund/SGB-II-Einstellungsgrund/Abmeldegrund des direkten Vorgängersatzes eine Information vorliegen kann (sofern ein entsprechender Satz vorhanden ist).

Diese Lücken wurden identifiziert anhand der Ausprägungen der Merkmale Abmeldegrund/Abgabegrund/Beendigungsgrund und Erwerbsstatus in den verschiedenen Quellen. Die Auflistung in Tabelle 3 erhebt keinen Anspruch auf Vollständigkeit.

Tabelle 3: Biografische Lücken und Identifikationsmöglichkeiten

Biografische Lücke	**Information über Lücke, ggf. identifizierbar anhand der Ausprägung des Merkmals „grund“ im Vorgängersatz der Quelle**[7]
Beamt*innen, Berufssoldat*innen, Richter*innen, Angestellte von Körperschaften oder Stiftungen des öffentlichen Rechts	XASU
Ungeförderte Selbstständige	LeH, ASU
Studierende, Personen in schulischer Weiterbildung	LeH, LHG, ASU, XASU
Personen, die länger als 6 Wochen krank / arbeitsunfähig sind (Krankheit während Arbeitslosigkeit ist aber unter bestimmten Bedingungen in der Quelle ASU abgebildet)	BeH, LeH, ASU
Personen in Altersrente ohne Beschäftigung, wenn nicht Mitglied in einer BG	LeH, LHG, ASU
Personen in Mutterschutz / Elternzeit	XASU
Beziehende von Vorruhestandsgeld	LeH, ASU
Hausgewerbetreibende	
Kurzfristig Beschäftigte	ASU
Personen in Einrichtungen der Jugendhilfe, Berufsbildungswerken, anerkannten Werkstätten oder	ASU

[7] Bitte beachten, dass die Quellen XASU, ASU, LeH und LHG im LIAB-LM-Datensatz nicht gefüllt sind.

Biografische Lücke	Information über Lücke, ggf. identifizierbar anhand der Ausprägung des Merkmals „grund“ im Vorgängersatz der Quelle[7]
ähnlichen Einrichtungen für Menschen mit Behinderung	
Teilnehmende an Leistungen zur Teilhabe am Arbeitsleben (Rehabilitanden)	ASU
(Nebenerwerbs-)Landwirt*innen	
Pflegepersonen i.S. von § 19 SBG XI	
Grundwehrdienstleistende	BeH, LeH, LHG, ASU, XASU
Wehrübungsleistende	BeH, LeH, LHG, ASU, XASU
Zivildienstleistende	BeH, LeH, LHG, ASU, XASU
Ableistende eines freiwilligen sozialen bzw. ökologischen Jahres anstelle des Zivildienstes	
Sonstige Personen, die nicht bei der Rentenversicherung oder der BA registriert sind (z.B. Sabbatjahr, Finanzierung aus privatem Vermögen oder Rentenbezügen, Emigration, Beschäftigung im Ausland, ehrenamtliche Tätigkeit, etc.)	BeH, LeH, ASU
Streikende bei Streikdauer über einem Monat	LeH
Sozialhilfe-Beziehende (vor Einführung des SGB II 2005), Sozialgeld-Beziehende (im Sinne des SGB II)	
SGB-II-Leistungsberechtigte, bei deren Trägern Lieferausfälle vorkamen	
Ausgleichsempfangende nach dem FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit)	

4.2 Beschäftigtenhistorik

- Die Beschäftigungshistorik enthält in wenigen Fällen (kleiner 1 Prozent) Duplikate in den Merkmalen Personennummer, Betriebsnummer, Erwerbsstatus, Beginndatum Originalzeitraum und Enddatum Originalzeitraum. Das Meldeverfahren sieht solche Duplikate eigentlich nicht vor. In den meisten Fällen unterscheiden sich die einzelnen Sätze der Duplikategruppe im Tagesentgelt, in selteneren Fällen auch in anderen Merkmalen. Der Grund für solche Duplikate sind vermutlich nicht korrekt verarbeitete Stornomeldungen.
- Die Übermittlung von Informationen zur Ausbildung, ausgeübten Tätigkeit und Stellung im Beruf erfolgt durch Meldungen der Arbeitgeber nach DEÜV (siehe Abschnitt 2) über einen sogenannten Tätigkeitsschlüssel. Für Zeiträume mit einem Enddatum nach dem 30.11.2011 erfolgte 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.[8]

- Das Merkmal „Stellung im Beruf und Arbeitszeit (stib)“ ist nur für Meldungen gefüllt, die vor Einführung des neuen Tätigkeitsschlüssels abgegeben wurden. Das FDZ nimmt keine Extrapolation oder Imputation für spätere Meldungen vor. Die Kategorien der Variable „stib“, die über den gesamten Beobachtungszeitraum konsistent beobachtet werden können (Erwerbsstatus als Auszubildender, Unterscheidung von Teil- und Vollzeit), sind in den Variablen „Erwerbsstatus (erwstat)“ und „Teilzeit (teilzeit)“ nachzuvollziehen.
- 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,[9] 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 im LIAB LM 7524 verwendet. Für die Lücken in den anderen Merkmalen werden keine Imputationen vorgenommen.
- Aufgrund der Einführung des Meldeverfahrens in den neuen Bundesländern kann erst ab 1993 von einer hinreichenden Vollständigkeit der Meldungen ostdeutscher Betriebe ausgegangen werden. Aus demselben Grund weisen sehr viele Sätze des Jahrs 1991 für einige Merkmale (u.a. Ausbildung, Erwerbsstatus und Tagesentgelt) fehlende Werte auf.
- Der Anstieg der Zahl der BeH-Sätze ab 1999 folgt aus der Einführung der Meldepflicht für geringfügig Beschäftigte ab dem 01.04.1999.
- Es lässt sich insbesondere 1999 ein deutlicher Anstieg bei den Meldungen zu Teilzeitbeschäftigung feststellen. Verursacht wird dieser Befund sowohl von einem tatsächlich zu beobachtenden Anstieg der Teilzeitbeschäftigung als auch von der Tatsache, dass seit 1999 die Beschäftigtenmeldungen generell besser ausgefüllt werden.
- 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

[8] Siehe Paulus/Matthes, 2013, für Details zur Klassifikation der Berufe 2010.

[9] Die im Rahmen des Meldeverfahrens eingesetzten Prüfprogramme ließen bis einschließlich Mai 2012 leere Angaben beim Tätigkeitsschlüssel 2010 zu.

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 Definition einen Füllgrad von 100 Prozent aufweist.

- In der zugrundeliegenden Version der IEB ist 2021 das letzte Jahr mit einem Füllgrad der BeH-Meldungen von 100 Prozent. Für die Jahre 2022 und 2023 wurden 18-Monats-Dateien verwendet, für 2024 eine 6-Monats-Datei. Es ist daher davon auszugehen, dass Beschäftigungsmeldungen für die Jahre 2022 bis 2023 leicht und für 2024 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 die Betriebsangaben, z.B. die Betriebsgröße, grob falsch sind und sich in Folgeversionen stark verändern werden.
- Im Jahr 1984 erfolgte eine Änderung des Meldeverfahrens. Einmalig gezahltes Bruttoarbeitsentgelt wurde ab diesem Zeitpunkt als Teil des beitragspflichtigen Jahresentgelts gemeldet, was zu einer Erhöhung des durchschnittlichen Tagesentgelts führt. Insbesondere steigt ab diesem Jahr der Anteil der Entgelte oberhalb der Beitragsbemessungsgrenze stark an (vgl. Bender et al., 1996).
- Für die Jahre 1992 bis 2000 wurden in den IEB auffällige Einbrüche und Anstiege in der Anzahl der Meldungen festgestellt. Es lassen sich v.a. Einbrüche bei folgenden zehn Kreisen beobachten: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) und Kempten (Allgäu) (09763). Grund hierfür sind Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen.
- Nennenswerte Einbrüche verzeichnen auch die Kreise Salzgitter (03102) und Hoyerswerda (14264).
- Bei den Meldungen zu Vollzeitbeschäftigung in den IEB sind besonders die Kreise Main-Taunus-Kreis (06436) und Alzey-Worms (07331) erwähnenswert. Hier sind überdurchschnittliche Anstiege zu beobachten. Als Grund sind ebenfalls Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen anzuführen.
- In den Jahren 1996 bis 1998 kommen die Werte 841-844 (Ärzt*innen und Apotheker*innen) in der Variable „Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf)“ im Vergleich zu benachbarten Jahren sehr selten vor. Die Gründe dafür sind nicht bekannt.
- In den Jahren 1975 bis 1977 gab es bisher deutlich zu viele Beschäftigte mit einem Arbeitsort in der niedersächsischen Gemeinde Lahn (Kreis Emsland). Statt im zu erwartenden zweistelligen Bereich gab es pro Jahr bis zu 90.000 Beschäftigtenmeldungen mit diesem Arbeitsort in der Grundgesamtheit der BeH. Ursächlich ist eine historische Fehlzuordnung der Beschäftigtenmeldungen aus der Stadt Lahn. Hierbei handelte es sich um einen nach kurzer Zeit wieder aufgelösten Zusammenschluss der hessischen Gemeinden Wetzlar, Gießen, Heuchelheim, Wettenberg und Lahnau (Kreise Gießen und Lahn-Dill-Kreis). Auf Basis dieser Erkenntnisse und unter der Annahme, dass es keine realen Betriebsumzüge zwischen diesen Regionen gab, wurde auf Ebene der Betriebsnummer folgende Bereinigungsregel umgesetzt:

Sobald der Betriebsort in den Jahren 1975 bis 1978 von der Gemeinde Lahn auf eine der aufgeführten hessischen Gemeinden wechselt, wird die frühere Angabe Gemeinde Lahn mit der späteren korrekten Angabe überschrieben. Über diese Regel wird der Überhang deutlich reduziert und die niedersächsische Gemeinde Lahn weist danach in den kritischen Jahren zumindest nur noch unter 2.000 Beschäftigtenmeldungen in der Grundgesamtheit der BeH auf.

5 Variablenbeschreibung

Dieses Kapitel beschreibt die im Datenprodukt enthaltenen Variablen im Detail. Weitere detaillierte Übersichten zu den einzelnen Merkmalen (Label, Auszählungen, fehlende Werte) sind dem Online-Anhang zum Datenreport zu entnehmen (siehe Abschnitt 7.1).

5.1 Identifikatoren

5.1.1 Systemfreie Personennummer (persnr_liab_lm)

Kategorie	**Beschreibung**
Variablenlabel	Systemfreie Personennummer
Variablenname	persnr_liab_lm
Kategorie	Identifikatoren
Datentyp	Numerisch
Detailbeschreibung	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. 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. Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.

5.1.2 Systemfreie Betriebsnummer (betnr_liab_lm)

Kategorie	**Beschreibung**
Variablenlabel	Systemfreie Betriebsnummer
Variablenname	betnr_liab_lm
Kategorie	Identifikatoren
Datentyp	Numerisch
Detailbeschreibung	Die systemfreie Betriebsnummer zeigt an, welche Sätze zum gleichen Betrieb gehören. Sie basiert auf der von der BA vergebenen Betriebsnummer, die durch eine systemfreie Nummer ersetzt wurde (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 https://www.arbeitsagentur.de/betriebsnummern-service/alles-wichtige). Die Betriebsnummer kann zusammen mit der Jahresangabe benutzt werden, um Personen- und Betriebsinformationen zusammenzuspielen. Für die Betriebsnummer sollte Folgendes grundsätzlich beachtet werden: 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. Hat das Unternehmen **mehrere** Niederlassungen in **einer** 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 **nicht** derselben Wirtschaftsklasse angehören.

Kategorie	Beschreibung
	Hat das Unternehmen **mehrere** Niederlassungen in **mehreren** Gemeinden, so ist jede dieser Niederlassungen ein Betrieb und erhält je eine Betriebsnummer. In diesem Zusammenhang sind auch folgende Definitionen im Sinne der Betriebsnummernvergabe im Rahmen des Meldeverfahrens zur Sozialversicherung zu beachten: Ein **Betrieb** ist eine regional und wirtschaftsfachlich abgegrenzte Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer vergeben wird. Eine **Arbeitsstätte** ist eine Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen keine Betriebsnummer zu vergeben ist. Ein **Unternehmen** als Begriff fasst Betriebs- und Arbeitsstätten desselben Arbeitgebers zusammen. Ein **Arbeitgeber** ist jede natürliche oder juristische Person, für die mindestens ein sozialversicherungspflichtig bzw. geringfügig Beschäftigter tätig ist. Betrieb und Betriebsstätten sind synonyme Begriffe; Niederlassung ist dann ein synonymer Begriff für Filialen, Zweigstellen, Auslagerungen, Dependancen, Arbeitsstätten u.ä., wenn sie kein Betrieb ist.
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, Betriebsgröße usw.) über die Betriebsnummer zugespielt werden, fehlen sie in diesen Meldungen.

5.2 Generierte technische Merkmale

5.2.1 Status der Betriebsnummer (betr_st)

Kategorie	Beschreibung
Variablenlabel	Status der Betriebsnummer
Variablenname	betr_st
Kategorie	Generierte technische Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal zeigt für die einzelnen Beschäftigungsepisoden in den Personendaten an, ob der zugrunde liegende Betrieb in der jeweiligen Basisperiode aus dem IAB-Betriebspanel stammt und ob die gesamte Belegschaft oder nur eine Stichprobe gezogen wurde. Die folgenden Ausprägungen existieren: 1. Alle Beschaeftigte BBBB bis EEEE verknuepft 2. 20 Prozent der Besch. BBBB bis EEEE verknuepft 3. Betrieb nicht Teil der Basisperiode SSSS bis EEEE BBBB steht dabei für den Beginn des Verknüpfungszeitraums, also von ein Jahr vor Beginn der jeweiligen Basisperiode bis zum Ende der Basisperiode (EEEE). SSSS steht für den Beginn der jeweiligen Basisperiode. Beispiel: Für die Basisperiode 2014 bis 2024 wurden die Betriebe des IAB-Betriebspanels der Wellen 2014 bis 2024 für die Verknüpfung zugrunde gelegt. Für diese Betriebe wurden die Personen identifiziert, die zwischen 2013 und 2024 zumindest einen Tag dort beschäftigt waren.

Kategorie	Beschreibung
	Für Betriebe mit der Ausprägung „betr_st==1“ sind die Erwerbsbiografien all dieser Personen in der Personendatei enthalten. Für Betriebe mit der Ausprägung „betr_st==2“ sind die Erwerbsbiografien einer 20 Prozent Zufallsstichprobe dieser Personen pro Betrieb über den gesamten Verknüpfungszeitraum enthalten. Betriebe mit der Ausprägung „betr_st==3“ wurden nicht verknüpft, sondern gelangen nur über die Erwerbsbiografien der verknüpften Personen in die Personendatei. Somit lassen sich für Betriebe mit „betr_st==1 | betr_st==2“ Variablen auf Betriebsebene aus den Personendaten generieren, allerdings nur innerhalb des jeweiligen Verknüpfungszeitraum.

5.2.2 Beobachtung wurde gezogen (smpl_st)

Kategorie	Beschreibung
Variablenlabel	Beobachtung wurde gezogen
Variablenname	smpl_st
Kategorie	Generierte technische Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal zeigt an, ob die Person aus dem Betrieb gezogen wurde, der der aktuellen Beschäftigungsepisode zugrunde liegt. Dies können sowohl Betriebe sein, für die alle Beschäftigten innerhalb des Verknüpfungszeitraum (d. h. der Basisperiode und ein Jahr davor) gezogen wurden, als auch Betriebe, für die eine 20 Prozent Zufallsstichprobe der Beschäftigten innerhalb des Verknüpfungszeitraums gezogen wurde. Dabei gilt einerseits zu beachten, dass Personen von der Stichprobenziehung ausgenommen wurden, wenn sie bereits im Rahmen der Ziehung der gesamten Belegschaft eines Betriebes ausgewählt wurden. Andererseits können Personen, die im Rahmen der Stichprobeziehung aus einem Betrieb ausgewählt wurden, auch in anderen Betrieben, für die ebenfalls eine Stichprobe gezogen wurde, auftreten, ohne Teil dieser zweiten Stichprobe zu sein (für diesen Fall nimmt „smpl_st“ den Wert „0“ an). Um nur die Personen einer 20 Prozent Zufallsstichprobe zu betrachten, müssen diese Fälle entsprechend ausgeschlossen werden. In Verbindung mit der Variable „Status der Betriebsnummer (betr_st)“ lassen sich diese Konstellationen identifizieren, um die Beobachtungen zu selektieren, für die Variablen auf Betriebsebene aus den Personendaten generiert werden können. Da dies nur innerhalb des Verknüpfungszeitraums möglich ist, nimmt „smpl_st“ für Episoden, die außerhalb des Verknüpfungszeitraums liegen, den Wert „.n“ an.

5.2.3 Satzzähler pro Konto (spell)

Kategorie	Beschreibung
Variablenlabel	Satzzähler pro Konto
Variablenname	spell
Kategorie	Generierte technische Merkmale
Datentyp	Numerisch
Detailbeschreibung	Der Satzzähler pro Konto nummeriert aufsteigend die Sätze einer Person. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze. Mit dem Merkmal „Satzzähler pro Konto“ kann die ursprüngliche Sortierung

Kategorie	Beschreibung
	auf einfache Art und Weise wiederhergestellt werden. Die Sortierung erfolgt grundsätzlich zunächst nach dem Beginndatum der Episode und dann nach der Quelle. Innerhalb der Beschäftigungsmeldungen werden sozialversicherungspflichtige vor geringfügige Beschäftigungsmeldungen und höhere vor niedrigere Tagesentgelte sortiert. Eine Ausnahme bilden Einmalzahlungen, die nach hinten sortiert werden.

5.2.4 Jahr (jahr)

Kategorie	Beschreibung
Variablenlabel	Jahr
Variablenname	jahr
Kategorie	Generierte technische Merkmale
Datentyp	Numerisch

Detailbeschreibung	Dieses Merkmal ist nur in der Betriebsdatei enthalten. Es gibt das Gültigkeitsjahr der Betriebsinformationen zum Stichtag 30.06 an.
	Mit Hilfe dieses Merkmals und der Betriebsnummer können Personen- und Betriebsdatei verknüpft werden.
	Beispielcode mit dem „merge“-Befehl in Stata 19 für die Basisperiode 2014 bis 2024:
	`use $orig/LIAB_LM_7524_B1424_v1.dta`
	`gen int jahr = year(begepi)`
	`sort betnr_liab_lm jahr`
	`merge m:1 betnr_liab_lm jahr using`
	`$orig/LIAB_LM_7524_B1424_v1_bhp_basis_v1.dta`

5.3 Gültigkeitszeitraum

5.3.1 Beginndatum Originalbeobachtung (begorig)

Kategorie	Beschreibung
Variablenlabel	Beginndatum Originalbeobachtung
Variablenname	begorig
Kategorie	Gültigkeitszeitraum
Datentyp	Datum
Detailbeschreibung	Das Beginndatum des Originalsatzes entspricht dem ursprünglichen Beginndatum der Meldung. Dies kann vom Beginndatum der betreffenden Datenzeile, dem sogenannten Episoden-Beginndatum, abweichen (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2). In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.

5.3.2 Enddatum Originalbeobachtung (endorig)

Kategorie	Beschreibung
Variablenlabel	Enddatum Originalbeobachtung

Kategorie	Beschreibung
Variablenname	endorig
Kategorie	Gültigkeitszeitraum
Datentyp	Datum
Detailbeschreibung	Das Enddatum des Originalsatzes entspricht dem ursprünglichen Enddatum der Meldung. Dies kann vom Enddatum der betreffenden Datenzeile, dem sogenannten Episoden-Enddatum, abweichen (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2). In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.

5.3.3 Beginndatum Episode (begepi)

Kategorie	Beschreibung
Variablenlabel	Beginndatum Episode
Variablenname	begepi
Kategorie	Generierter Gültigkeitszeitraum
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 3.2).

5.3.4 Enddatum Episode (endepi)

Kategorie	Beschreibung
Variablenlabel	Enddatum Episode
Variablenname	endepi
Kategorie	Generierter Gültigkeitszeitraum
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 3.2).

5.4 Informationen zur Person

5.4.1 Geschlecht (frau)

Kategorie	Beschreibung
Variablenlabel	Geschlecht
Variablenname	frau
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Geschlechter-Dummy (0 - Mann, 1 - Frau). Die Geschlechtsangabe ist innerhalb eines Personenkontos konstant.

5.4.2 Geburtsjahr (gebjahr)

Kategorie	Beschreibung
Variablenlabel	Geburtsjahr
Variablenname	gebjahr
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Geburtsjahr ist innerhalb eines Personenkontos konstant.

5.4.3 Geburtsmonat (gebmon)

Kategorie	Beschreibung
Variablenlabel	Geburtsmonat
Variablenname	gebmon
Kategorie	Persönliche Merkmale
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, kann folgender Befehl verwendet werden:
	`gen int gebdat = ym(gebjahr, gebmon)`
	`format gebdat %tm`

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)

Kategorie	Beschreibung
Variablenlabel	Staatsangehörigkeit
Variablenname	nation
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die Staatenschlüssel des Statistischen Bundesamtes (Statistisches Bundesamt, 2026).
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird als Staatsangehörigkeitsmerkmal lediglich die Staatsangehörigkeit vergröbert (nation_gr) bereitgestellt.

5.4.5 Staatsangehörigkeit vergröbert (nation_gr)

Kategorie	Beschreibung
Variablenlabel	Staatsangehörigkeit vergröbert
Variablenname	nation_gr
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die zu Gruppen zusammengefassten Staatenschlüssel des Statistischen Bundesamtes (Statistisches Bundesamt, 2026). Eine Übersicht über diese Gruppen findet sich in der „Übersicht der

Kategorie	Beschreibung
	Klassifikationen“ unter den zentralen Arbeitshilfen (https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/).

5.4.6 Ausbildung (ausbildung)

Kategorie	Beschreibung
Variablenlabel	Ausbildung
Variablenname	ausbildung
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	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. Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Berufsausbildung. Folgende Ausprägungen kommen vor: 1 ohne abgeschlossene Berufsausbildung 2 betriebliche/außerbetriebliche Ausbildung 11 FH ohne nähere Spezifikation 12 Hochschulabschluss ohne nähere Spezifikation Bei Meldungen nach dem neuen Tätigkeitsschlüssel (siehe Abschnitt 4.2) 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.
Qualitätshinweis	„Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammenstellt werden. Wenn ein Beschäftigter z.B. durch berufsbegleitende 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 Prozent. Seit 2014 pendelt sich der Anteil der fehlenden Werte in der BeH bei rund 40 Prozent 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)“ 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“.

5.4.7 Ausbildung (imputiert) (ausbildung_imp)

Kategorie	Beschreibung
Variablenlabel	Ausbildung (imputiert)
Variablenname	ausbildung_imp
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal „Ausbildung (imputiert)“ ist ein Zusatzmerkmal zum Merkmal „Ausbildung (ausbildung)“ und enthält für BeH-Meldungen ergänzte und harmonisierte Angaben zur Berufsausbildung von Beschäftigten. Das Merkmal bietet damit einen Lösungsansatz für die in Abschnitt 5.4.6 beschriebenen Probleme der Variable „Ausbildung (ausbildung)“ bei BeH-Meldungen. Das Imputationsverfahren wird in Thomsen et al. (2018) beschrieben, welches auf der Arbeit von Fitzenberger et al. (2006) basiert. Da das Merkmal nur Bildungsangaben der BeH-Meldungen nutzt, und weil die Bildungskategorien des alten und des neuen Tätigkeitsschlüssels für das Merkmal „Ausbildung (ausbildung)“ harmonisiert werden mussten, bietet das Merkmal „Ausbildung (imputiert) (ausbildung_imp)“ andere Kategorien als das Merkmal „Ausbildung (ausbildung)“. Eine Einordnung der Abschlüsse in die Internationale Standardklassifikation des Bildungswesens (ISCED 2011) ist den zentralen Arbeitshilfen des FDZ (https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/) zu entnehmen.
Qualitätshinweis	Trotz des Imputationsverfahrens nimmt der Anteil der fehlenden Angaben im Zeitverlauf zu und erreichen 2019 ca. 9 Prozent. 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.

5.4.8 Schulausbildung (schule)

Kategorie	Beschreibung
Variablenlabel	Schulausbildung
Variablenname	schule
Kategorie	Persönliche Merkmale
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält den Schulabschluss. Eine Einordnung der Abschlüsse in die Internationale Standardklassifikation des Bildungswesens (ISCED 2011) ist den zentralen Arbeitshilfen des FDZ (https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/) zu entnehmen. Je nach Quelle sind unterschiedliche Ausprägungen möglich. Mit der Umstellung des Tätigkeitsschlüssels in 2011 ändern sich die möglichen Ausprägungen (siehe Abschnitt 4.2). Davor gelten: 5 Volks-/Hauptschule o./m. Abschluss, Mittlere Reife oder gleichwertige Schulbildung 8 Fachhochschulreife oder Abitur/Hochschulreife 9 Abitur/Hochschulreife Danach gelten: 1 Kein Hauptschulabschluss 4 Hauptschulabschluss/Volksschulabschluss 6 Mittlere Reife

Kategorie	Beschreibung
	8 Fachhochschulreife oder Abitur/Hochschulreife 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).
Qualitätshinweis	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.

5.5 Informationen zu Beschäftigung

5.5.1 Tagesentgelt/ täglicher Leistungssatz (tentgelt)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt/täglicher Leistungssatz
Variablenname	tentgelt
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	In BeH-Sätzen weist das Merkmal jeweils das Bruttotagesentgelt eines Beschäftigten aus. Es wird berechnet aus dem vom Betrieb gemeldeten Zeitraumentgelt sowie der Dauer des Zeitraums der (ungesplitteten) Originalmeldung in Kalendertagen. Das Tagesentgelt ist in Euro angegeben. Bis 1998 meldeten Betriebe 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 zwei Beitragsbemessungsgrenzen. Die Grenze der Knappschaftlichen Rentenversicherung liegt meist höher als die Grenze der Arbeiter- und Angestelltenrentenversicherung. Jedoch ist eine Unterscheidung der beiden Versicherungsträger in den Daten nicht möglich. Seit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren zum 01.04.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze erfasst. Die Beitragsbemessungsgrenze als obere Kappungsgrenze 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 Entgelt-Informationen hingegen können auf Grund ihrer Versicherungsrelevanz als weniger fehleranfällig gesehen werden.) 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 den zentralen Arbeitshilfen (https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/). Eine Angabe von 0 Euro ist auf "Unterbrechungsmeldungen" zurückzuführen. In diesen Zeiten besteht das Beschäftigungsverhältnis zwar rechtlich fort, es wird

Kategorie	Beschreibung
	aber kein Entgelt gezahlt. Dies ist der Fall bei Krankheit nach Ende der Lohnfortzahlung, Mutterschaftszeiten und Sabbaticals. Ab 2013 nimmt die Anzahl der Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) sprunghaft zu (siehe Abschnitt 5.5.14). 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) (tentgelt_bonus)“). 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.

5.5.2 Tagesentgelt (inkl. Einmalzahlungen) (tentgelt_bonus)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt (inkl. Einmalzahlungen)
Variablenname	tentgelt_bonus
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal Tagesentgelt (inkl. Einmalzahlungen) ist ein Zusatzmerkmal zum Merkmal Tagesentgelt/ täglicher Leistungssatz (tentgelt) 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. Das Verfahren wird in Abschnitt 4.1. von Drechsler et al. (2023) beschrieben.
Qualitätshinweis	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.

5.5.3 Tagesentgelt (imputiert) (tentgelt_imp)

Kategorie	Beschreibung
Variablenlabel	Tagesentgelt (imputiert)
Variablenname	tentgelt_imp
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal „Tagesentgelt (imputiert)“ ist ein Zusatzmerkmal zum Merkmal „Tagesentgelt/ täglicher Leistungssatz (tentgelt)“ 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) (tentgelt_bonus)“ auf. Tagesentgelte dieser Subpopulation, die die Beitragsbemessungsgrenze nicht überschreiten, werden aus „tentgelt_bonus“ übernommen.

Kategorie	Beschreibung
	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.

5.5.4 Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf)

Kategorie	Beschreibung
Variablenlabel	Beruf – ausgeübte/letzte Tätigkeit (KldB 1988)
Variablenname	beruf
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
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 4.2). 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. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 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.

5.5.5 Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3)

Kategorie	Beschreibung
Variablenlabel	Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010), 3-Steller
Variablenname	beruf2010_3
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch

Kategorie	Beschreibung
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 „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2021). 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 – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)“ zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). 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. Die KldB2010 wurde 2020 aktualisiert. Dabei wurden zwei neue Berufsuntergruppen (4-Steller) und 14 neue Berufsgattungen (5-Steller) geschaffen. Zudem wurden eine Berufsuntergruppe und eine Berufsgattung innerhalb der Systematik umgezogen. In der Quelle BeH sind die Änderungen ab 2020 enthalten.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 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.

5.5.6 Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4)

Kategorie	Beschreibung
Variablenlabel	Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010), 4-Steller
Variablenname	beruf2010_4
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
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 „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2021). 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

Kategorie	Beschreibung
	„Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)“ zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). 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. Die KldB2010 wurde 2020 aktualisiert. Dabei wurden zwei neue Berufsuntergruppen (4-Steller) und 14 neue Berufsgattungen (5-Steller) geschaffen. Zudem wurden eine Berufsuntergruppe und eine Berufsgattung innerhalb der Systematik umgezogen. In der Quelle BeH sind die Änderungen ab 2020 enthalten. Die Berufsgruppe lässt sich über die Berufshauptgruppe (erste zwei Stellen des Schlüssels) den Berufssektoren und Berufssegmenten nach Klassifikation der Bundesagentur für Arbeit und des IAB zuordnen. Eine Einordnung ist den zentralen Arbeitshilfen des FDZ (https://fdz.iab.de/anleitungen-und-hilfen/zentrale-arbeitshilfen/) zu entnehmen. Die als Meister tätigen Beschäftigten können, abgesehen von wenigen Ausnahmen in den Verkehrs-, Handels- und personenbezogenen Dienstleistungsberufen (beruf2010_3: 514, 621, 632, 633, 713, 714, 732, 813, 831, 832, 921, 944 und 946), durch eine Kombination des Werts 9 an der vierten Stelle (Berufsuntergruppe) mit dem Wert 3 im „Anforderungsniveau (niveau)“ identifiziert werden.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 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.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die „Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3)“ bereitgestellt.

5.5.7 Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)

Kategorie	Beschreibung
Variablenlabel	Anforderungsniveau - ausgeübte/letzte Tätigkeit (KldB 2010)
Variablenname	niveau
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
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 „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2021). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Das an der fünften Stelle erfasste Anforderungsniveau einer

Kategorie	Beschreibung
	Tätigkeit wird gesondert in der Variable „Anforderungsniveau (niveau)“ zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). 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. Die KldB2010 wurde 2020 aktualisiert. Dabei wurden zwei neue Berufsuntergruppen (4-Steller) und 14 neue Berufsgattungen (5-Steller) geschaffen. Zudem wurden eine Berufsuntergruppe und eine Berufsgattung innerhalb der Systematik umgezogen. In der Quelle BeH sind die Änderungen ab 2020 enthalten.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 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.

5.5.8 Teilzeit (teilzeit)

Kategorie	Beschreibung
Variablenlabel	Teilzeit
Variablenname	teilzeit
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal „teilzeit“ unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit.
Qualitätshinweis	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. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. Hier wird keine Imputation vorgenommen. 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. Eine weiterführende Diskussion des Problems und ein möglicher Korrekturansatz findet sich in Fitzenberger/Seidlitz (2020).

5.5.9 Stellung im Beruf und Arbeitszeit (stib)

Kategorie	Beschreibung
Variablenlabel	Stellung im Beruf und Arbeitszeit
Variablenname	stib
Kategorie	Informationen zu Beschäftigung
Datentyp	Numerisch

Kategorie	Beschreibung
Detailbeschreibung	Die berufliche Stellung eines Beschäftigten im Meldezeitraum wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt. Das Merkmal „Stellung im Beruf und Arbeitszeit (stib)“ 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 lag bis 1978 bei 20 Wochenstunden, zwischen 1979 und 1987 bei 15 Wochenstunden und seit 1988 bei 18 Wochenstunden. Nur für Vollzeitbeschäftigte liefert die Variable tatsächlich Angaben zur beruflichen Stellung, wobei unter anderem zwischen vollzeitbeschäftigten Arbeitern und Angestellten sowie Auszubildenden unterschieden wird. Die Unterscheidung zwischen vollzeitbeschäftigten Arbeitern und Angestellten richtet sich ausschließlich nach der Art der Rentenversicherungspflicht (BfA für Angestellte und LVA für Arbeiter). Meister und Poliere werden nur dann in einer eigenen Kategorie erfasst, wenn sie in der Arbeiterrentenversicherung pflichtversichert sind. Die Zuordnung eines Meisters oder Poliers zu den Arbeitern oder Angestellten kann nur über den jeweiligen Rententräger erfolgen. Neben den Auszubildenden, Praktikanten und Volontären zählen auch Anlernlinge, Schüler an Schulen des Gesundheitswesens und Teilnehmer an geförderten Maßnahmen zur beruflichen Fortbildung, Umschulung und betrieblichen Einarbeitung zur Kategorie Beschäftigte in beruflicher Ausbildung. Falls mehrere Schlüsselzahlen in Betracht kommen, ist der Arbeitgeber dazu angehalten, die Einstufung nach der überwiegend ausgeübten Tätigkeit vorzunehmen. Lässt sich diese nicht eindeutig feststellen, ist die Schlüsselzahl der höherwertigen Stellung im Beruf einzutragen (vgl. Bundesagentur für Arbeit 2005, S. VI). Mit der Einführung des neuen Tätigkeitsschlüssels (siehe Abschnitt 4.2) fiel diese Unterscheidung jedoch weg. Die Variable „stib“ ist daher nur für Meldungen gefüllt, die noch vor Einführung des neuen Tätigkeitsschlüssels abgegeben wurden. Es besteht jedoch die Möglichkeit, Tätigkeiten, die eine Meisterausbildung erfordern, anhand der KldB 2010 zu identifiziert. Weitere Informationen finden Sie unter der Variablen „Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4)“.
Qualitätshinweis	In 1991 kommt es aufgrund der Wiedervereinigung zu einer deutlichen Häufung der fehlenden Angaben.

5.5.10 Erwerbsstatus (erwstat)

Kategorie	Beschreibung
Variablenlabel	Erwerbsstatus
Variablenname	erwstat
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Für BeH-Sätze entspricht das Merkmal Erwerbsstatus der im neuen Meldeverfahren (DEÜV) ab dem 01.01.1999 erfassten Personengruppe. Sie weist beitrags- oder leistungsrechtliche Besonderheiten des Beschäftigungsverhältnisses aus. Treffen für eine Beschäftigungsmeldung mehrere Schlüssel zu, ist vom meldenden Arbeitgeber der kleinste anzugeben. Mehrheitlich handelt es sich dabei um sozialversicherungspflichtige Beschäftigungsverhältnisse ohne Besonderheiten, die unter der Schlüsselnummer 101 geführt werden. Entsprechend ist es möglich, dass diese Beschäftigungen leicht überschätzt werden.

Kategorie	Beschreibung
	Im Meldeverfahren ist geregelt, dass Wechsel im Erwerbsstatus - z. B. die Übernahme eines Auszubildenden nach dem Ende der Berufsausbildung im Ausbildungsbetrieb - durch eine erneute Meldung erfolgen müssen. 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. Diese übrigen Meldungen wurden der Personengruppe "Sozialversicherungspflichtig Beschäftigte ohne besondere Merkmale" (101) zugeordnet. Seit dem 01.04.1999 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.

5.5.11 Gleitzone (gleitz)

Kategorie	Beschreibung
Variablenlabel	Gleitzone
Variablenname	gleitz
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal ist erst ab 2003 und nur für BeH-Datensätze gefüllt. Es 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 entsprechen Entgelt als Gleitzone gekennzeichnet. Die Zuordnung erfolgt über das Bruttoarbeitsentgelt und ändern sich im Zeitverlauf: 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 01.07.2019 bis 30.09.2022: 450,01 bis 1.300,00 Euro 01.10.2022 bis 31.12.2022: 520,01 bis 1.600 Euro 01.01.2023 bis 31.12.2023: 520,01 bis 2.000 Euro 01.01.2024 bis 31.12.2024: 538,01 bis 2.000 Euro Abweichende untere Grenzen existieren für Auszubildende und Beschäftigte in Freiwilligendiensten. Bei mehreren Beschäftigungsverhältnissen ist das insgesamt erzielte Arbeitsentgelt maßgebend. Die entsprechende gesetzliche Regelung besteht seit dem 01.04.2003.

5.5.12 Leiharbeit (leih)

Kategorie	Beschreibung
Variablenlabel	Leiharbeit

Kategorie	Beschreibung
Variablenname	leih
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	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.
Qualitätshinweis	In 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Im Jahr 2012 beträgt die Missingquote nur noch rund 3 Prozent und sinkt bis 2019 auf unter 1 Prozent.

5.5.13 Befristung (befrist)

Kategorie	Beschreibung
Variablenlabel	Befristung
Variablenname	befrist
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
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 noch rund 3 Prozent und sinkt bis 2019 auf unter 1 Prozent.

5.5.14 Abmeldegrund / Abgabegrund / Beendigungsgrund (grund)

Kategorie	Beschreibung
Variablenlabel	Abmeldegrund / Abgabegrund / Beendigungsgrund
Variablenname	grund
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
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 möglichen 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 Sonderzahlungen, die bis März des Folgejahres ausgezahlt

Kategorie	Beschreibung
	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.

5.5.15 Dauer der Arbeitslosigkeit (alo_dau)

Kategorie	Beschreibung
Variablenlabel	Dauer der Arbeitslosigkeit
Variablenname	alo_dau
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Dauer (in Tagen) einer ununterbrochenen Folge von Arbeitslosigkeitsperioden an und ist zu Beginn des Satzes (siehe „Beginndatum Episode (begepi)“) gültig. Es wird auf Basis weiterer Quellen zu Leistungsbezug, Arbeitssuche und Maßnahmenteilnahme in der IEB berechnet (siehe dazu Schmucker und vom Berge, 2025). Folgende Lücken führen nicht zu einer Unterbrechung der Arbeitslosigkeit: • jede Lücke kleiner oder gleich 7 Tage • Krankheitslücken bis zu 42 Tagen Bei der Berechnung der Dauer werden diese Lücken jedoch nicht addiert. Daher kann sich die Dauer in diesem Merkmal von dem Ergebnis der Dauerberechnung mit dem "Beginndatum der Arbeitslosigkeit" unterscheiden. Vor 1997 bedeutet der Wert „0“ nicht, dass keine Arbeitslosigkeit vorliegt, da hier die Quellen ASU/XASU nicht gefüllt sind.

5.6 Ortsangaben

5.6.1 Wohnort - Kreis (wo_kreis)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Kreis
Variablenname	wo_kreis
Kategorie	Ortsangaben
Datentyp	Numerisch
Hierarchie	Bundesland | Kreis
Detailbeschreibung	In BeH- Sätzen liegt der Kreis des Wohnortes erst für die Jahre ab 1999 vor. Darin ist jeweils der Kreis (kreisfreie Stadt oder Landkreis) angegeben, 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. In der BeH 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.

Kategorie	Beschreibung
	Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben für alle Quellen auf den Gebietsstand vom 31.12.2024 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2024 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/die betreffende Versicherte umgezogen ist.
Qualitätshinweis	Hinsichtlich der Wohnortangaben 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. Im Jahr 2015 kommt es in den Daten zu einer Reduktion der Wohnortwechsel von ca. 10-15 Prozent gegenüber dem üblichen Niveau der umliegenden Jahre, wobei regionale Unterschiede auftreten. Der genaue Grund für diese Abweichung konnte bisher nicht ermittelt werden.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das „Wohnort - Bundesland (wo_bula)“ bereitgestellt.

5.6.2 Wohnort - Bundesland (wo_bula)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Bundesland
Variablenname	wo_bula
Kategorie	Ortsangaben
Datentyp	Numerisch
Hierarchie	Bundesland | Kreis
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal Wohnort - Kreis auf die 16 deutschen Bundesländer. In BeH Sätzen liegt das Bundesland des Wohnortes erst für die Jahre ab 1999 vor. Weitere Informationen zum Kreis des Wohnortes finden sich unter „Wohnort - Kreis (wo_kreis)“.

5.6.3 Wohnort - Arbeitsagentur (wo_aa)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Arbeitsagentur
Variablenname	wo_aa
Kategorie	Ortsangaben
Datentyp	Numerisch
Hierarchie	Regionaldirektion | Arbeitsagentur
Detailbeschreibung	Das Merkmal enthält ab 1999 den Agenturbezirk der am Wohnort des Beschäftigten/Kunden zuständigen Arbeitsagentur. Sie wird aus der Wohnadresse ermittelt. In der BeH enthält das Merkmal den Wohnort zum Beginn der Meldung. Mit der Dauer des Spells steigt also das Risiko, dass der Wohnort veraltet ist und

Kategorie	Beschreibung
	die Angaben zu späteren Stichtagen falsch sind. Es ist für die BeH erst ab 1999 vorhanden. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben für alle Quellen auf den Gebietsstand vom 31.12.2024 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Agenturbezirk nach den Grenzen, welche die Bezirke zum 31.12.2024 hatten. Da sich die Grenzen im Zeitverlauf geändert haben, würden ohne diese Gebietsstandsaktualisierung Fälle auftreten, bei denen der Bezirk einer Person wechselt, ohne dass diese Person umgezogen ist. Einen Problemfall stellt hinsichtlich der Gebietsstandsaktualisierung allerdings Berlin dar: Die Berliner Arbeitsagenturbezirke wurden im Laufe der Jahre immer wieder neu zugeschnitten, was auch durch die Gebietsstandsumschlüsselung nicht vollständig bereinigt werden konnte. Berlin sollte möglichst als Einheit und nicht getrennt nach Ost und West analysiert werden, da seit der Gebietsreform der Arbeitsämter in Berlin zum 01.07.1997 deren eindeutige Zuordnung zu Ost bzw. West nicht mehr möglich ist. Von der Statistik der BA wird Berlin einheitlich dem Osten zugeordnet.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird lediglich das Gebiet der Regionaldirektion, in dem der Wohnort des Versicherten liegt, ausgewiesen („Wohnort - Regionaldirektion (wo_rd)“).

5.6.4 Wohnort - Regionaldirektion (wo_rd)

Kategorie	Beschreibung
Variablenlabel	Wohnort - Regionaldirektion
Variablenname	wo_rd
Kategorie	Ortsangaben
Datentyp	Numerisch
Hierarchie	Regionaldirektion | Arbeitsagentur
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal „Wohnort - Arbeitsagentur (wo_aa)“ auf die Ebene der Regionaldirektionen. Weitere Informationen finden sich dort.

5.7 Betriebsmerkmale

5.7.1 WZ73 3-Steller (w73_3)

Kategorie	Beschreibung
Variablenlabel	WZ73 3-Steller
Variablenname	w73_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabteilung (1-Steller) | Wirtschaftsgruppe (2-Steller) | Wirtschaftsklasse (3-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WS73 als 3-Steller aus und ist für die Jahre 1975 bis einschließlich 2002 gefüllt.

Kategorie	Beschreibung
	WS73 steht für die "Klassifikation der Wirtschaftszweige 1973“ (Bundesagentur für Arbeit, 2009). Mittels eines dreistelligen Zahlencodes werden 269 Wirtschaftsklassen unterschieden, wobei die erste Stelle dieses Codes die Wirtschaftsabteilung, insgesamt 10, definiert und die beiden ersten Stellen zusammen die jeweilige Gruppe, insgesamt 95, festlegen. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Die Zuordnung zu der entsprechenden Wirtschaftsklasse erfolgt unter Beachtung der institutionellen Ausrichtung des Betriebes.

5.7.2 WZ93 5-Steller (w93_5)

Kategorie	Beschreibung
Variablenlabel	WZ93 5-Steller
Variablenname	w93_5
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 5-Steller aus und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt, 2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature génerale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w93_3) bereitgestellt.

5.7.3 WZ93 3-Steller (w93_3)

Kategorie	Beschreibung
Variablenlabel	WZ93 3-Steller
Variablenname	w93_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 3-Steller aus und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt,

Kategorie	Beschreibung
	2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature générale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.

5.7.4 WZ03 5-Steller (w03_5)

Kategorie	Beschreibung
Variablenlabel	WZ03 5-Steller
Variablenname	w03_5
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ03 als 5-Steller aus und ist für die Jahre 2003 bis 2008 gefüllt. WZ03 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2003" (Statistisches Bundesamt, 2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 auf (siehe Variablenbeschreibung w93_3, w93_5). Die Klassifikationen der Wirtschaftszweige wurden aktualisiert, wobei die Struktur der WZ93 weitgehend beibehalten wurde. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w03_3) bereitgestellt.

5.7.5 WZ03 3-Steller (w03_3)

Kategorie	Beschreibung
Variablenlabel	WZ03 3-Steller
Variablenname	w03_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ03 als 3-Steller aus und ist für die Jahre 2003 bis 2008 gefüllt. WZ03 steht für die "Klassifikation der Wirtschaftszweige Ausgabe 2003" (Statistisches Bundesamt,

Kategorie	Beschreibung
	2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 auf (siehe Variablenbeschreibung w93_3, w93_5). Die Klassifikationen der Wirtschaftszweige wurden aktualisiert, wobei die Struktur der WZ93 weitgehend beibehalten wurde. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.

5.7.6 WZ08 5-Steller (w08_5)

Kategorie	Beschreibung
Variablenlabel	WZ08 5-Steller
Variablenname	w08_5
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ08 als 5-Steller aus und ist für die Jahre ab 2008 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. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird der Wirtschaftszweig lediglich als 3-Steller (w08_3) bereitgestellt.

5.7.7 WZ08 3-Steller (w08_3)

Kategorie	Beschreibung
Variablenlabel	WZ08 3-Steller
Variablenname	w08_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ08 als 3-Steller aus und ist für die Jahre ab 2008 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.

Kategorie	Beschreibung
	Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widergespiegelt werden.

5.7.8 w73_3 vervollständigt durch Extrapolation/Imputation (w73_3_gen)

Kategorie	Beschreibung
Variablenlabel	w73_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w73_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabteilung (1-Steller) | Wirtschaftsgruppe (2-Steller) | Wirtschaftsklasse (3-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WS73 als 3-Steller aus. Von 1975 bis einschließlich 2002 enthält das Merkmal die Originalwerte aus „w73_3“. Ab 2003 wird die Information entweder fortgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt. Das Merkmal bietet somit eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WS73. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al., 2011. Weitere Informationen zur Klassifikation WS73 finden sich bei der Variablenbeschreibung zum Merkmal „w73_3“.

5.7.9 Art der Vervollständigung w73_3 (group_w73_3)

Kategorie	Beschreibung
Variablenlabel	Art der Vervollständigung w73_3
Variablenname	group_w73_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal weist die Art der Vervollständigung für das Merkmal „w73_3_gen“ aus. Es ist angegeben, ob der jeweilige Wert in „w73_3_gen“ dem Originalwert von „w73_3“ entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde. Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al., 2011.

5.7.10 w93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen)

Kategorie	Beschreibung
Variablenlabel	w93_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w93_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch

Kategorie	Beschreibung
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ93 als 3-Steller aus. Von 1998 bis einschließlich 2003 enthält das Merkmal die Originalwerte aus „w93_3“. Vor 1998 und nach 2003 wird die Information entweder rück- oder fortgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt, sodass das Merkmal eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WZ93 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al., 2011. Weitere Informationen zur Klassifikation WZ93 finden sich bei der Variablenbeschreibung zum Merkmal „w93_3“.

5.7.11 Art der Vervollständigung w93_3 (group_w93_3)

Kategorie	Beschreibung
Variablenlabel	Art der Vervollständigung w93_3
Variablenname	group_w93_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal weist die Art der Vervollständigung für das Merkmal „w93_3_gen“ aus. Es ist angegeben, ob der jeweilige Wert in „w93_3_gen“ dem Originalwert von „w93_3“ entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde. Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al., 2011.

5.7.12 w08_3 vervollständigt durch Extrapolation/Imputation (w08_3_gen)

Kategorie	Beschreibung
Variablenlabel	w08_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w08_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Wirtschaftsabschnitt (1-Steller) | Wirtschaftsabteilung (2-Steller) | Wirtschaftsgruppe (3-Steller) | Wirtschaftsklasse (4-Steller) | Wirtschaftsunterklasse (5-Steller)
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ08 als 3-Steller aus. Von 2008 bis 2017 enthält das Merkmal die Originalwerte aus „w08_3“. Vor 2008 wird die Information rückgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt, sodass das Merkmal eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WZ08 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al., 2011. Weitere Informationen zur Klassifikation WZ08 finden sich bei der Variablenbeschreibung zum Merkmal „w08_3“.

5.7.13 Art der Vervollständigung w08_3 (group_w08_3)

Kategorie	Beschreibung
Variablenlabel	Art der Vervollständigung w08_3
Variablenname	group_w08_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal weist die Art der Vervollständigung für das Merkmal „w08_3_gen“ aus. Es ist angegeben, ob der jeweilige Wert in „w08_3_gen“ dem Originalwert von „w08_3“ entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde. Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al., 2011.

5.7.14 Jahr erstes Auftreten Betriebsnummer (grd_jahr)

Kategorie	Beschreibung
Variablenlabel	Jahr erstes Auftreten Betriebsnummer
Variablenname	grd_jahr
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Bei dieser Variablen handelt es sich um das erste Auftreten der Nummer des Betriebes im Datensatz. Wird eine Betriebsnummer in Westdeutschland erst nach 1975 bzw. in Ostdeutschland nach 1992 erstmalig nachgewiesen, könnte es sich um das Gründungsdatum des betreffenden Betriebes handeln. Allerdings könnte es sich auch um einen bereits länger existierenden Betrieb handeln, dem nach einem Wechsel des Inhabers oder einer Änderung der Rechtsform eine neue Betriebsnummer zugeordnet wurde (zur Betriebsnummernvergabe siehe Bundesagentur für Arbeit, 2007, S. 9-11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte. Ein Betrieb muss nicht zwingend in dem Jahr des ersten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Gründungsjahr an diesem Stichtag keine Beschäftigte, erscheint er auch nicht im BHP.

5.7.15 Jahr letztes Auftreten Betriebsnummer (lzt_jahr)

Kategorie	Beschreibung
Variablenlabel	Jahr letztes Auftreten Betriebsnummer
Variablenname	lzt_jahr
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Bei dieser Variablen handelt es sich um das letztmalige Auftreten der Nummer des Betriebes im Datensatz (vgl. Bender et al., 1996). Endet die Existenz einer Betriebsnummer im BHP bereits vor dem Jahr 2024, könnte es sich um eine Betriebsschließung handeln. Allerdings kommen als Ursache auch ein „willkürlicher Wechsel der Betriebsnummer bei Wechsel des Inhabers oder

Kategorie	Beschreibung
	Änderung der Rechtsform“, eine „Ausgliederung von Unternehmensteilen unter einer neuen Nummer“ oder andere verwaltungstechnische Änderungen in Betracht (vgl. Bender et al., 1996 oder Bundesagentur für Arbeit, 2007, S. 9-11). Ein Betrieb muss nicht zwingend in dem Jahr des letzten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Schließungsjahr an diesem Stichtag keine Beschäftigte, erscheint er auch nicht im BHP.

5.7.16 Anzahl Beschäftigte gesamt (az_ges)

Kategorie	Beschreibung
Variablenlabel	Anzahl Beschäftigte gesamt
Variablenname	az_ges
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Die Variable enthält die Anzahl aller zum 30.06. eines Jahres zur Sozialversicherung gemeldeten Mitarbeiter eines Betriebes. Seit Einführung der neuen Meldevorschriften im Jahr 1999 sind auch die geringfügig Beschäftigten erfasst. Ruhende Beschäftigungsverhältnisse (Tagesentgelt von Null) sind nicht enthalten.

5.7.17 Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz)

Kategorie	Beschreibung
Variablenlabel	Anzahl Vollzeit (Normalbeschaeftigte + sonstige)
Variablenname	az_vz
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Die Variable enthält die Anzahl der Personen im Betrieb, die die am 30.06. eines Jahres mit den Personengruppenschlüsseln 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 und YYY und als Vollzeitbeschäftigte gemeldet sind. Azubis, geringfügige Beschäftigte und Personen in Altersteilzeit werden nicht berücksichtigt.

5.7.18 Anzahl geringfügig Beschäftigte (az_gf)

Kategorie	Beschreibung
Variablenlabel	Anzahl geringfügig Beschäftigte
Variablenname	az_gf
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Die Anzahl der geringfügig Beschäftigten wird anhand des Personengruppenschlüssels – Ausprägungen 109 und 209 – gebildet. Dieses Merkmal ist erst seit 1999 im Datensatz enthalten, da die geringfügig Beschäftigten erst seit diesem Jahr in das Meldeverfahren zur Sozialversicherung integriert wurden.

5.7.19 Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw)

Kategorie	Beschreibung
Variablenlabel	Mittelwert imp. Bruttotagesentgelt Vollzeitbeschaeftigte
Variablenname	te_imp_mw
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Diese Variable enthält den Mittelwert des imputierten Bruttotagesentgeltes der Vollzeitbeschäftigten eines Betriebes. Nicht enthalten sind die Löhne der geringfügig Beschäftigten, der Azubis sowie der Personen in Altersteilzeit. Die Werte sind für alle Jahre in Euro angegeben. Gemäß den Meldevorschriften zur Sozialversicherung müssen Arbeitgeber angeben, wie hoch das beitragspflichtige Bruttoarbeitsentgelt des Beschäftigten für einen bestimmten Zeitraum war (Zeitraumentgelt). Bis Ende 1998 hatten Arbeitgeber nur den sozialversicherungspflichtigen Bruttolohn zu melden. Damit wurden nur Arbeitsentgelte erfasst, die oberhalb der Geringfügigkeits- und unterhalb der Beitragsbemessungsgrenze lagen. Seit 1999 werden im Rahmen des neuen Meldeverfahrens auch Entgelte gemeldet, die unterhalb der Geringfügigkeitsgrenze liegen. Bruttolöhne, die oberhalb der Beitragsbemessungsgrenze liegen, werden jedoch weiterhin gekappt. Zur Ermittlung des Bruttotageslohnes wird das Zeitraumentgelt durch die Anzahl der Kalendertage des Zeitraums geteilt. Zur Berechnung der Mittelwerte wurden diese zensierten Löhne imputiert (vgl. Abschnitt 3.1.3.3 in Ganzer et al. 2025). Diese Daten wurden dann auf Betriebsebene aggregiert. Die Werte sind auf zwei Dezimalstellen gerundet. Aufgrund des "storage types" in Stata werden jedoch weitere Dezimalstellen angezeigt, die nicht korrekt sind.

5.7.20 Arbeitsort - Kreis (ao_kreis)

Kategorie	Beschreibung
Variablenlabel	Arbeitsort Kreis
Variablenname	ao_kreis
Kategorie	Ortsangaben
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Bundesland | Kreis
Detailbeschreibung	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. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben auf den Gebietsstand vom 31.12.2024 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2024 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.

Kategorie	Beschreibung
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergröbert bereitgestellt. Ansonsten wird als Regionalangabe lediglich das „Arbeitsort - Bundesland (ao_bula)“ bereitgestellt.

5.7.21 Arbeitsort - Bundesland (ao_bula)

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

6 Literatur

Bellmann, Lisa/ Gensicke, Miriam/ Grau, Katrin/ Kohaut, Susanne/ Möller, Iris/ Oberfichtner, Michael/ Schwengler, Barbara/ Tschersich, Nikolai/ Umkehrer, Matthias (2026): IAB-Betriebspanel (IAB-BP) 1993-2024. FDZ-Datenreport 01/2026 (de), Nürnberg, DOI:10.5164/IAB.FDZD.2601.de.v1

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, Dr. Henning/ Zemann, Dr. Christian (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Methodik-Qualitaet/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/MB-Arbeitszeit-ausgeuebte-Taetigkeit-Schul-und-Berufsabschluss.pdf?__blob=publicationFile&v=4 (Zugriff: 01.09.2026).

Bundesagentur für Arbeit (Hrsg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.

Bundesagentur für Arbeit (Hrsg.) (2007): Handbuch für die Betriebsnummernvergabe und –pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Ausgabe Dezember 2007, Nürnberg.

Bundesagentur für Arbeit (2009): Klassifikation der Wirtschaftszweige 1973, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Wirtschaftszweige/Klassifikation-der-Wirtschaftszweige-1973-2003/Generische-Publikationen/WS-1973-Excel.xlsx?__blob=publicationFile&v=6 (Zugriff: 01.09.2026).

Bundesagentur für Arbeit (2021): Klassifikation der Berufe 2010 – überarbeitete Fassung 2020 Band 1: Systematischer und alphabetischer Teil mit Erläuterungen. Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Printausgabe-KldB-2010-Fassung2020/Generische-Publikationen/KldB2010-PDF-Version-Band1-Fassung2020.pdf?__blob=publicationFile&v=21 (Zugriff: 01.09.2026).

Bundesanstalt für Arbeit (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB1975-1992/Generische-Publikationen/KldB1988-Systematisches-und-alphabetisches-Verzeichnis-der-Berufsbenennungen-PDF.pdf?__blob=publicationFile&v=12 (Zugriff: 01.09.2026).

Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: S. 56-68.

Drechsler, Jörg/ Ludsteck, Johannes (2025): Imputation Strategies for Rightcensored Wages in Longitudinal Datasets, arXiv:2502.12967v1.

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/ 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, Jg. 126, H. 3, S. 405-436.

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.

Ganzer, Andreas/ Schmucker, Alexandra/ Wolter, Stefanie (2025): Betriebs-Historik-Panel 1975–2024. FDZ-Datenreport 13/2025 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2513.de.v1

Lochner, Ben; Wolter, Stefanie (2025): AKM effects for German labour market data 1985-2023. (FDZ-Methodenreport, 03/2025 (en)), Nürnberg, 13 S.

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.

Schmucker, Alexandra/ vom Berge, Philipp (2025): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975 - 2023. FDZ-Datenreport, 02/2025 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2502.de.v1

Statistisches Bundesamt (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-1993.html (Zugriff: 01.09.2026).

Statistisches Bundesamt (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2003.html (Zugriff: 01.09.2026).

Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html (Zugriff: 01.09.2026).

Statistisches Bundesamt (2026): Staats- und Gebietssystematik, Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?__blob=publicationFile&v=5 (Zugriff: 01.09.2026).

Statistik der Bundesagentur für Arbeit (2021): Grundlagen: Methodenbericht – Einführung der „Klassifikation der Berufe 2010 – überarbeitete Fassung 2020“ in die Arbeitsmarktstatistiken,

Nürnberg. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Methodik-Qualitaet/Methodenberichte/Uebergreifend/Generische-Publikationen/Methodenbericht-Klassifikation-Berufe-ueberarbeitete-Fassung.pdf?__blob=publicationFile&v=4 (Zugriff: 01.09.2026).

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. Journal for Labour Market Research, Volume 57, Article number: 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.

7 Anhang

7.1 Labels, Auszählungen und fehlende Werte

Der Online-Anhang zum Datenreport stellt detaillierte Übersichten zu den einzelnen Merkmalen im Datenprodukt bereit. Er enthält Labellisten der einzelnen Ausprägungen, Auszählungen sowie Übersichten über den Anteil fehlender Werte (missings) getrennt nach Datenquellen. Sie finden den Anhang in gesonderten Dateien in deutscher und englischer Sprache unter: https://fdz.iab.de/

- https://doku.iab.de/fdz/reporte/2026/DR_09-26_frequencies_labels_de.zip
- https://doku.iab.de/fdz/reporte/2026/DR_09-26_frequencies_labels_en.zip

7.2 Abkürzungsverzeichnis

AA	Agentur für Arbeit / Arbeitsamt
ALG	Arbeitslosengeld
ARGE	Arbeitsgemeinschaft
ASU	Arbeitsuchendenhistorik
A2LL	Arbeitslosengeld II – Leistungen zum Lebensunterhalt
BA	Bundesagentur für Arbeit
BeH	Beschäftigtenhistorik
BfA	Bundesversicherungsanstalt für Angestellte
BG	Bedarfsgemeinschaft
BHP	Betriebs-Historik-Panel
BMAS	Bundesministerium für Arbeit und Soziales
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und -übermittlungsverordnung –
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –
DIM	Daten- und IT-Management
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung
EDV	Elektronische Datenverarbeitung
FDZ	Forschungsdatenzentrum der Bundesagentur für Arbeit am IAB
FELEG	Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit
gAw	Träger mit getrennter Aufgabenwahrnehmung
gE	Gemeinsame Einrichtungen

gT	Getrennte Trägerschaft
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IEB	Integrierte Erwerbsbiografien
ISIC	International Standard Industrial Classification of All Economic Activities
KldB	Klassifikation der Berufe
LeH	Leistungsempfängerhistorik
LHG	Leistungshistorik Grundsicherung
LIAB	Linked-Employer-Employee-Daten des IAB
LM	Längsschnittmodell
LVA	Landesversicherungsanstalt
MTH	Maßnahmeteilnahmehistorik
NACE	Nomenclature génerale des activités économiques dans les communautés européennes
NUTS	Nomenclature des unités territoriales statistiques
QM	Querschnittmodell
SGB	Sozialgesetzbuch
SIAB	Stichprobe der Integrierten Arbeitsmarktbiografien
VerBIS	Vermittlungs- und Beratungsinformationssystem
XASU	Arbeitsuchendenhistorik aus XSozial-BA-SGB II
XMTH	Maßnahmeteilnahmehistorik aus XSozial-BA-SGB II
zkT	Zugelassene kommunale Träger

Impressum

FDZ-Datenreport 09|2026

Veröffentlichungsdatum
04.11.2026

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/2026/DR_09-26.pdf

Dokumentation Version
LIABLM7524_DE_v1_dok1, DOI: 10.5164/IAB.FDZD.2609.de.v1

Datensatz Version
LIAB LM 7524 v1, DOI: 10.5164/IAB.LIABLM7524.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
Matthias Umkehrer
Telefon: 0911 179-6211
E-Mail: Matthias.Umkehrer@iab.de