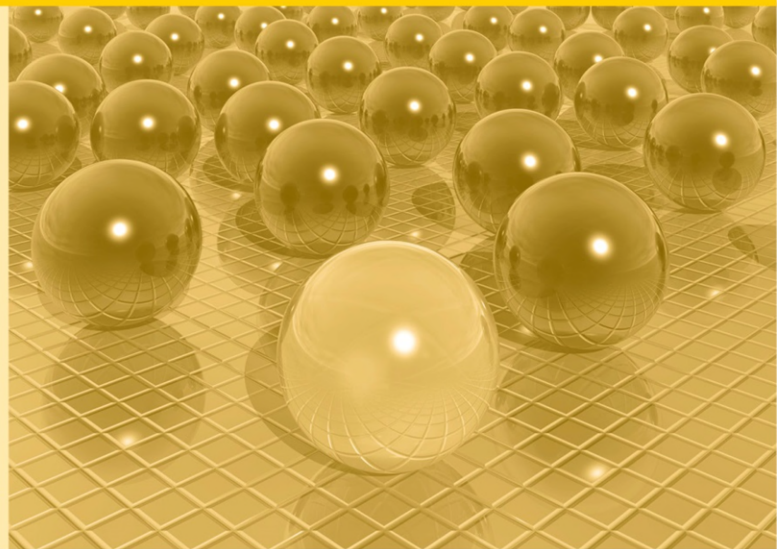


Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung -
Beschäftigtenpanel 2022-2023 per On-Site-Nutzung (EVAS-Nummer: 62361)

DOI: 10.21242/62361.2023.00.01.1.1.0

Version 1

Impressum

Herausgeber: Statistische Ämter des Bundes und der Länder
Herstellung: Information und Technik Nordrhein-Westfalen
Telefon 0211 9449-01 • Telefax 0211 9449-8000
Internet: www.forschungsdatenzentrum.de
E-Mail: forschungsdatenzentrum@it.nrw.de

Fachliche Informationen zu dieser Veröffentlichung:

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Standort Hessen –
Tel.: 0611 3802-822
Fax: 0611 3802-890
forschungsdatenzentrum@statistik.hessen.de

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum

Tel.: 0611 75-2420
Fax: 0611 72-3915
forschungsdatenzentrum@destatis.de

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Geschäftsstelle –
Tel.: 0211 9449-2883
Fax: 0211 9449-8087
forschungsdatenzentrum@it.nrw.de

Erscheinungsfolge: unregelmäßig
Erschienen im Oktober 2025

Diese Publikation wird kostenlos als PDF-Datei zum Download unter www.forschungsdatenzentrum.de angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2025
(im Auftrag der Herausbergemeinschaft)

Vervielfältigung und Verbreitung, nur auszugsweise, mit Quellenangabe gestattet. Alle übrigen Rechte bleiben vorbehalten.

Fotorechte Umschlag: ©artSILENCEcom –Fotolia.com

Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung - Beschäftigtenpanel 2022-2023 per On-Site-Nutzung (EVAS-Nummer: 62361). Version 1. DOI: 10.21242/62361.2023.00.01.1.1.0. Wiesbaden 2025.

Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung -
Beschäftigtenpanel 2022-2023 per On-Site-Nutzung (EVAS-Nummer: 62361)

DOI: 10.21242/62361.2023.00.01.1.1.0

Version 1

Inhalt

1	Datenaufbereitung in den FDZ	2
1.1	Datenaufbereitung	2
1.2	Anonymisierungsmaßnahmen	2
1.3	Speicherstruktur des Beschäftigtenpanels.....	2
1.4	Methodik der Verknüpfung.....	2
2	Produkt	2
2.1	Merkmale und Merkmalsbeschreibung	4
2.1.1	Datensatzbeschreibung/Schlüsselverzeichnis	5
2.1.2	Merkmalsdefinitionen	5
2.2	Vergleichbarkeit der Merkmale über die Zeit	25
2.3	Eckwerte relevanter Merkmale und Merkmalskombinationen	26
2.4	Auswertbare regionale Ebene.....	27
3	Praktische Hinweise	28
3.1	Hinweise zur Geheimhaltung	28
3.1.1	Gesetzliche Grundlagen der statistischen Geheimhaltung.....	28
3.1.2	Geheimhaltung von Ergebnissen.....	29
3.1.3	Praktische Tipps zur Vermeidung von Geheimhaltungsfällen.	30
3.2	FAQ	30
3.3	Verfügbare Tools	30
	Anhang	31

1 Datenaufbereitung in den FDZ

1.1 Datenaufbereitung

Es wurden keine weiteren Schritte zur Aufbereitung der Daten vorgenommen. Aufbereitungsschritte, die durch die Fachseite erfolgten, werden im Metadatenreport Teil I beschrieben.

1.2 Anonymisierungsmaßnahmen

Die Gemeindekennziffer für das Bundesland Bayern steht am Gastwissenschaftsarbeitsplatz lediglich pseudoanonymisiert zur Verfügung.

1.3 Speicherstruktur des Beschäftigtenpanels

Das Beschäftigtenpanel wird im Long-Format gespeichert, dabei bildet jede Zeile eine Kombination aus Berichtsmonat und Beschäftigten ab. Die strukturierte Trennung der Zeitperioden in einzelne Zeilen erleichtert die Verarbeitung gängiger Statistiksoftware.

1.4 Methodik der Verknüpfung

Die Verknüpfung wurde gemäß den Angaben im Metadatenreport Teil I von der Fachseite des Statistischen Bundesamtes vorgenommen.

2 Produkt

Die dem Beschäftigtenpanel zu Grunde liegenden Daten werden von der Verdiensterhebung ab dem Jahr 2022 monatlich erhoben.

Die Jahreswellen werden im FDZ jährlich zur Verfügung gestellt. Es liegen Angaben zu Betrieben sowie zu Arbeitnehmerinnen und Arbeitnehmern vor. Die

Daten eignen sich gut zur Analyse geschlechtsspezifischer Lohnunterschiede sowie zur Untersuchung der Verdienstunterschiede in tarifgebundenen Betrieben im Vergleich zu solchen, die nach freier Vereinbarung vergüten. Außerdem sind sie geeignet, um Verdienstverläufe im Kontext von arbeitsmarktpolitischen Eingriffen wie beispielsweise dem Mindestlohn zu analysieren.

Da die Stichprobenauswahl auf Bundeslandebene erfolgt, lassen sich für kleinräumigere regionale Gliederungen keine repräsentativen Ergebnisse erzielen.

Das Beschäftigtenpanel enthält Informationen zur Person (Geschlecht, Geburtsdatum (MMJJJJ), Ausbildung), zur Tätigkeit (Tätigkeitsschlüssel, Arbeitszeit, Eintrittsdatum (MMJJJJ)) und zum Verdienst (Brutto, Zulagen für Schicht-/Nachtarbeit, Sonderzahlungen). Auf Betriebsebene gibt es beispielsweise Angaben zum Sitz des Betriebes, dem Wirtschaftszweig und der Anzahl der Beschäftigten im zugehörigen Unternehmen.

Die Einzeldaten des Beschäftigtenpanels umfassen monatlich durchschnittlich 52 000 Betriebs- und ca. 8,8 Mio. Beschäftigtendatensätze ohne Imputation von Antwortausfällen. Mit der Imputation von Antwortausfällen sind es monatlich durchschnittlich 53 000 Betriebs- und ca. 8,9 Mio. Beschäftigtendatensätze.

Die Auswahlgrundlage für die Stichprobenziehung bildet die Bundeskopie des statistischen Unternehmensregisters zum vorhergehenden Berichtsjahr. Jährlich wird ein Sechstel der Betriebe aus der Stichprobe entlassen und durch neue Betriebe ersetzt.

Für jedes Jahr werden eine Welle (die verknüpften 12 Monate eines Berichtsjahres) sowie die aktuellsten Hochrechnungsfaktoren im FDZ bereitgestellt und können in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder On-Site, also am Gastwissenschafts Arbeitsplatz und über die

kontrollierte Datenfernverarbeitung genutzt werden. Die Verknüpfung der Beschäftigten ist mit Hilfe des Merkmals EF3 über die Berichtsmonate hinweg individuell möglich.

2.1 Merkmale und Merkmalsbeschreibung

Im Beschäftigtenpanel lassen sich bei einzelnen Merkmalen, die in der Praxis typischerweise über die Zeit konstant bleiben – wie beispielsweise das Geschlecht oder die Staatsangehörigkeit –, vereinzelt abweichende Angaben innerhalb eines Beschäftigungsverhältnisses beobachten. So liegt der Anteil an Meldungen mit wechselnder Geschlechtsangabe im Jahr 2022 bei 0,02 % und im Jahr 2023 bei 0,01 %. Solche Veränderungen können reale biografische Entwicklungen widerspiegeln, etwa eine Änderung der Staatsangehörigkeit oder des eingetragenen Geschlechts. Angesichts des geringen Anteils und des Musters in den Daten erscheint es jedoch plausibel, dass ein erheblicher Teil dieser Abweichungen auf Eingabefehler oder uneinheitliche Meldungen zurückzuführen ist.

Dies legt auch der Vergleich nach Betriebsgrößenklasse nahe: In Kleinbetrieben treten solche Abweichungen beim Geschlecht deutlich häufiger auf als in großen Betrieben. In allen Berichtsjahren ist der Anteil der Meldungen mit wechselnder Geschlechtsangabe in der kleinsten Größenklasse mindestens zehnmal so hoch wie in der größten.

2.1.1 Datensatzbeschreibung/Schlüsselverzeichnis

Die Angaben beziehen sich auf das Material VEBP050X.

2.1.2 Merkmalsdefinitionen

EF1 – Berichtsmonat

EF2 – Berichtsjahr

EF3 – Beschäftigten-ID

Bei der Beschäftigten-ID handelt es sich um einen anonymisierten, eindeutigen Identifikator des Beschäftigten.

EF4 – BerichtseinheitID (Identnummer des Betriebs (URS))

Bei der Identnummer aus dem Unternehmensregister (URS) handelt es sich um eine fortlaufende Nummer, die von den jeweiligen Landesämtern zur Nummerierung der Betriebe im URS verwendet wird.

EF5 – Geschlecht

Ausprägungen:

1 = männlich

2 = weiblich

Der Fragebogen ermöglicht auch die Angabe „divers“. Aus Geheimhaltungsgründen werden diese Angaben aber nicht ausgewiesen, sondern hälftig den beiden anderen Geschlechtskategorien zugeordnet.

EF6 – Geburtsdatum

Das Geburtsdatum ist nach dem Muster MMJJJJ hinterlegt.

EF7 – Erste Staatsangehörigkeit

[Staatsangehörigkeit nach der Staats- und Gebietssystematik \(Stand: 01. April 2019\).](#)

EF8 – Eintrittsdatum

Monat und Jahr des Eintritts in das Unternehmen im Format MMJJJJ. Bei Unterbrechungen des Beschäftigungsverhältnisses ist das Eintrittsdatum anzugeben, das der Betrieb für seine internen Zwecke verwendet. Es entspricht dem Datum des Beschäftigungsbeginns laut Entgeltbescheinigungsverordnung (EBV) § 1 Absatz 1.

EF9 – Personengruppenschlüssel

Schlüsselzahlen für Personengruppen in den Meldungen nach der Datenerfassungs- und Übermittlungsverordnung (DEÜV). Für Beschäftigte, die nicht der Sozialversicherung gemeldet werden, z. B. Beamtinnen und Beamte, werden die Schlüsselzahlen analog ermittelt.¹

Schlüssel	Personenkreis
101	Sozialversicherungspflichtig Beschäftigte ohne besondere Merkmale
102	Auszubildende ohne besondere Merkmale
103	Beschäftigte in Altersteilzeit
105	Praktikanten
106	Werkstudenten
109	Geringfügig entlohnte Beschäftigte nach § 8 Absatz 1 Nummer 1 des Vierten Buches Sozialgesetzbuch (SGB IV)

¹ Die 800er und 900er-Schlüssel sind keine offiziellen Schlüssel lt. DEÜV, sondern Hilfsschlüssel der amtlichen Statistik.

Schlüssel	Personenkreis
110	Kurzfristig Beschäftigte nach § 8 Absatz 1 Nummer 2 SGB IV
113	Nebenerwerbslandwirte
114	Nebenerwerbslandwirte - saisonal beschäftigt
116	Ausgleichsgeldempfänger nach dem Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit (FELEG)
117	Nicht berufsmäßig unständig Beschäftigte
118	Berufsmäßig unständig Beschäftigte
119	Versicherungsfreie Altersvollrentner und Versorgungsbezieher wegen Alters
120	Versicherungspflichtige Altersvollrentner
121	Auszubildende, deren Arbeitsentgelt die Geringverdienergrenze nach § 20 Absatz 3 Satz 1 Nummer 1 SGB IV nicht übersteigt
122	Auszubildende in einer außerbetrieblichen Einrichtung
124	Heimarbeiter ohne Anspruch auf Entgeltfortzahlung im Krankheitsfall
140	Seeleute
141	Auszubildende in der Seefahrt ohne besondere Merkmale
142	Seeleute in Altersteilzeit
143	Seelotsen
144	Auszubildende in der Seefahrt, deren Arbeitsentgelt die Geringverdienergrenze nach § 20 Absatz 3 Satz 1 Nummer 1 SGB IV nicht übersteigt
149	In der Seefahrt beschäftigte versicherungsfreie Altersvollrentner und Versorgungsbezieher wegen Alters
150	In der Seefahrt beschäftigte versicherungspflichtige Altersvollrentner
190	Beschäftigte, die ausschließlich gesetzlich unfallversichert sind
801	Beamtinnen und Beamte ohne besondere Merkmale
802	Beamtinnen und Beamte -Auszubildende
803	Beamtinnen und Beamte -Altersteilzeit
910	Leitende Angestellte (auch Geschäftsführerinnen und Geschäftsführer einer GmbH und Vorstände einer AG) mit einem Arbeitsvertrag, die zumindest teilweise feste, d.h. gewinnunabhängige Verdienstbestandteile für die geleistete Arbeit erhalten
920	Saison- und Gelegenheitsarbeiterinnen und -arbeiter auch wenn sie nicht in der deutschen Sozialversicherung gemeldet sind

EF10 – Tätigkeitsschlüssel 2010

Es handelt sich um einen neunstelligen Schlüssel, der durch den Arbeitgeber oder die Arbeitgeberin zur Sozialversicherung an die Krankenkassen gemeldet wird. Er ist nicht als neunstelliger Schlüssel im Datensatz enthalten, sondern setzt sich aus den folgenden Merkmalen zusammen:

EF10U1 - Ziffer 1-5: Berufsschlüssel aus der Klassifikation der Berufe (KldB) 2010; überarbeitete Fassung 2020

Für Beschäftigte, die nicht der Sozialversicherung gemeldet werden, wie beispielsweise Beamtinnen und Beamte, werden die Schlüsselzahlen analog ermittelt.

Eine ausführliche Erläuterung des 5-stelligen Schlüssels der KldB 2010 findet sich in: Statistik der Bundesagentur für Arbeit, Grundlagen: Methodenbericht – Einführung der „Klassifikation der Berufe 2010 – überarbeitete Fassung 2020“ in die Arbeitsmarktstatistiken, Nürnberg, Januar 2021.

EF10U2 - Ziffer 6: Höchster allgemeinbildender Schulabschluss

Ausprägungen:

1 = Ohne Schulabschluss

2 = Haupt-/Volksschulabschluss

3 = Mittlere Reife oder gleichwertiger Abschluss

4 = Abitur

9 = Abschluss unbekannt

Die Ausprägung "Abschluss unbekannt" wurde nach dem Nearest-Neighbour-Prinzip mit der kanadischen Software CANCEIS durch imputierte Werte ersetzt. In 2022 wurden manuell plausibilisierte Beschäftigtendatensätze jedoch nicht durch CANCEIS imputiert, sodass die Angabe 9 enthalten ist.

EF10U3 - Stelle 7: Höchster beruflicher Ausbildungsabschluss

Ausprägungen:

- 1 = Ohne beruflichen Ausbildungsabschluss
- 2 = Abschluss einer anerkannten Berufsausbildung
- 3 = Meister-/Techniker- oder gleichwertiger Fachschulabschluss
- 4 = Bachelor
- 5 = Diplom/Magister/Master/Staatsexamen
- 6 = Promotion
- 9 = Abschluss unbekannt

Die Ausprägung "Abschluss unbekannt" wurde nach dem Nearest-Neighbour-Prinzip mit der kanadischen Software CANCEIS durch imputierte Werte ersetzt. In 2022 wurden manuell plausibilisierte Beschäftigtendatensätze jedoch nicht durch CANCEIS imputiert, sodass die Angabe 9 enthalten ist.

EF10U4 - Stelle 8: Arbeitnehmerüberlassung

Das Merkmal zeigt an, ob eine Arbeitnehmerin/ein Arbeitnehmer als Zeitarbeiterin/Zeitarbeiter eingesetzt wird.

Ausprägungen:

- 1 = Nein
- 2 = Ja

EF10U5 - Stelle 9: Vertragsform

Ausprägungen:

1 = Vollzeit, unbefristet

2 = Teilzeit, unbefristet

3 = Vollzeit, befristet

4 = Teilzeit, befristet

EF11 – Bezahlte Arbeitsstunden ohne Überstunden im Monat

Das Merkmal gibt für jeden Beschäftigten die bezahlten Arbeitsstunden im Monat an, unabhängig davon, ob die Entlohnung anhand der Arbeitsstunden errechnet wird oder ob Monatsgehälter gezahlt werden. Für Beschäftigte, die nicht nach Stunden bezahlt werden, wird hier die vertraglich vereinbarte Arbeitszeit im Monat angegeben. Bei geringfügig Beschäftigten, für die in den Betrieben keine Stundenangaben vorliegen, sollten diese von den Betrieben geschätzt werden.

EF12 – Bezahlte Überstunden im Monat

Als Überstunden gelten in der Berichtsperiode bezahlte Arbeitsstunden, die über die vereinbarte Arbeitszeit hinaus geleistet und nicht durch Gewährung von Freizeit an anderen Tagen ausgeglichen wurden, unabhängig davon, ob für diese Stunden ein Zuschlag bezahlt wurde oder nicht. Hierzu zählen auch über die Normalarbeitszeit hinaus geleistete, zuschlagfreie Stunden im Rahmen von Arbeitszeitkorridoren. Wenn eine tarifliche Wochenarbeitszeitverkürzung durch freie Tage realisiert wird, sind die vorher eingearbeiteten Stunden dementsprechend nicht als Überstunden anzugeben.

EF13 – Bruttomonatsverdienst – Gesamtbruttoentgelt

Als Bruttomonatsverdienst für den Berichtsmonat ist das Gesamtbruttoentgelt gemäß Entgeltbescheinigungsverordnung (EBV) § 1 Absatz 2 Nummer 2 c anzugeben.

Das Gesamtbruttoentgelt ist gesetzlich auf jeder Lohnabrechnung auszuweisen und kann dadurch nicht nur leicht und eindeutig von den Betrieben übermittelt werden, sondern ist auch für die Datennutzenden leicht und zuverlässig mit der eigenen Lohnabrechnung vergleichbar.

Sehr kleine Werte (z. B. unter einem Euro) wurden ebenfalls durch die Betriebe gemeldet und hängen ggf. mit Kurzarbeit zusammen. Sie wurden durch die Statistischen Landesämter im Plausibilisierungsverfahren geprüft und bestätigt.

EF14 – Sonstige Bezüge (EBV § 1 Absatz 2 Nummer 2a)

Es handelt sich um eine Darunter-Position des Gesamtbruttoentgelts. Die „Sonstigen Bezüge“ laut § 1 Absatz 2 Nummer 2a Entgeltbescheinigungsverordnung (EBV) plus Zuschüsse des Arbeitgebers/der Arbeitgeberin zum Kurzarbeitergeld.

EF15 – Bruttomonatsverdienst für bezahlte Überstunden

Es handelt sich um eine Darunter-Position des Gesamtbruttoentgelts. Hier sind nicht nur die Zuschläge für Überstunden erfasst, sondern die Gesamtvergütung für Überstunden.

EF16 – Bruttomonatsverdienst für Zuschläge für Schicht-, Samstags-, Sonntags-, Feiertags- oder Nachtarbeit

Es handelt sich um eine Darunter-Position des Gesamtbruttoentgelts. Das Merkmal erfasst nur die Zuschläge für Schicht-, Samstags-, Sonntags-, Feiertags- oder Nachtarbeit und nicht den Gesamtverdienst der mit Zulagen vergüteten Stunden. Bereits im Überstundenverdienst gemeldete Zulagen werden ebenfalls nicht nochmals erfasst.

EF17 – Bruttomonatsverdienst für Zwecke der Entgeltumwandlung

Es handelt sich um eine Darunter-Position des Gesamtbruttoentgelts. Bei der Entgeltumwandlung (Gehaltsverzicht) wird zwischen Arbeitgeberin/Arbeitgeber und Beschäftigten vereinbart, Teile des Bruttoverdienstes zu Gunsten einer betrieblichen Altersversorgung einzusetzen. Dieser Bestandteil wird in einen Vertrag eingezahlt, aufgrund dessen im Rentenalter eine einmalige Leistung oder eine laufende Rente geleistet wird.

Finanziert werden können die Beiträge aus dem laufenden Arbeitsentgelt, vermögenswirksamen Leistungen oder Einmal- und Sonderzahlungen wie Urlaubs- oder Weihnachtsgeld oder einem 13. Monatsgehalt.

Der angegebene Wert ist der Betrag an Entgeltumwandlung des Berichtsmonats seitens der oder des Beschäftigten. Einbezogen sind alle Durchführungswege (Pensionskasse, Pensionsfonds, Direktversicherung, auch Direktzusage und Unterstützungskasse) und Besteuerungsformen (steuerfrei, pauschal, individuell versteuert).

EF18 – Bundesland

Amtlicher Schlüssel des Bundeslandes, in dem sich der Betrieb zum Zeitpunkt der Stichprobenziehung befindet:

Ausprägungen:

01 = Schleswig-Holstein	09 = Bayern
02 = Hamburg	10 = Saarland
03 = Niedersachsen	11 = Berlin
04 = Bremen	12 = Brandenburg
05 = Nordrhein-Westfalen	13 = Mecklenburg-Vorpommern
06 = Hessen	14 = Sachsen
07 = Rheinland-Pfalz	15 = Sachsen-Anhalt
08 = Baden-Württemberg	16 = Thüringen

Aufgrund der unterschiedlichen Zeitpunkte des Abzugs der Informationen (Stichprobenziehung und Berichtsjahr), kann bei Umzügen der Betriebe die Angabe mit denen aus „EF19“ nicht übereinstimmen (sog. Landeswechsler).

EF19 – Regionalschlüssel (AGS)

Amtlicher Gemeindeschlüssel der Gemeinde, in welcher der Betrieb seinen Sitz zum Zeitpunkt des Berichtsmonats der VE hat.

Der Gemeindeschlüssel (Regionalschlüssel) setzt sich zusammen aus:

Stelle 1 + 2: Länderschlüssel (2-Steller)

Stelle 3: Ergibt zusammen mit dem Länderschlüssel die Kennziffer des Regierungsbezirkes.

EF22 – Tarifbindung

Angabe des Betriebs, welche Tarifbindung mehrheitlich im Betrieb Anwendung findet.

1 = Branchentarifvertrag

2 = Firmentarifvertrag

3 = kein Tarifvertrag, Orientierung der Verdienste an einem Branchentarifvertrag

4 = kein Tarifvertrag, freie Verdienstvereinbarung

Das Merkmal zur Tarifbindung lässt sich zwischen der Verdienststrukturhebung (VSE) und der Verdiensterhebung (VE) im Längsschnitt nur eingeschränkt vergleichen. Bei der VSE wurde dieses Merkmal auf Beschäftigtenebene erhoben, während die Tarifbindung bei der VE auf Betriebsebene erhoben wird. Bei der VE geben Betriebe an, welcher Tarifvertrag in ihrem Betrieb überwiegend Anwendung findet. Hierbei kann von den Betrieben nur eine Antwortoption ausgewählt werden, auch wenn de facto unterschiedliche Tarifbindungen innerhalb eines Betriebs vorliegen. Bei der VSE hingegen wurde diese Information für jeden einzelnen Beschäftigten geliefert, sodass es zu heterogenen Angaben hinsichtlich der Tarifbindung innerhalb eines Betriebs kommen konnte.

EF23 – Regionsgrundtyp

Regionen des Betriebssitzes nach der [Abgrenzung des Bundesamts für Bauwesen und Raumordnung \(BBR\)](#).

Ausprägungen:

1 = Städtische Regionen

Regionen, in denen mindestens 50% der Bevölkerung in Groß- und Mittelstädten leben und in der sich eine Großstadt mit rund 500 000 Einwohnern und mehr befindet sowie Regionen mit einer Einwohnerdichte ohne Berücksichtigung der Großstädte von mindestens 300 Einwohner pro km².

2 = Regionen mit Verdichtungsansätzen

Regionen, in denen mindestens 33% der Bevölkerung in Groß- und Mittelstädten leben, mit einer Einwohnerdichte zwischen 150 und 300 Einwohner pro km² sowie Regionen, in denen sich mindestens eine Großstadt befindet und die eine Einwohnerdichte ohne Berücksichtigung der Großstädte von mindestens 100 Einwohner pro km² aufweisen.

3 = Ländliche Regionen

Regionen, in denen weniger als 33% der Bevölkerung in Groß- und Mittelstädten leben, mit einer Einwohnerdichte unter 150 Einwohner pro km² sowie Regionen, in denen sich zwar eine Großstadt befindet, aber die eine Einwohnerdichte ohne Berücksichtigung der Großstädte unter 100 Einwohner pro km² hat.

EF24 – Differenzierter Regionstyp

Regionen des Betriebssitzes nach der [Abgrenzung des Bundesamts für Bauwesen und Raumordnung \(BBR\)](#).

Ausprägungen:

01 = „Städtischer Raum“ (Kreisfreie Großstädte und städtische Kreise)

02 = „Ländlicher Raum“ (Ländliche Kreise)

EF25 – Kreistyp

Typen des Betriebssitzes nach der [Abgrenzung des Bundesamts für Bauwesen und Raumordnung \(BBR\)](#). Die siedlungsstrukturellen Kreistypen dienen dem intraregionalen Vergleich. Es wird nach "Kernstädten" und sonstigen Kreisen bzw. Kreisregionen unterschieden. Als Kernstädte werden kreisfreie Städte mit mehr als 100 000 Einwohnern ausgewiesen. Kreisfreie Städte unterhalb dieser Größe werden mit ihrem Umland zu Kreisregionen zusammengefasst. Die Typisierung der Kreise und Kreisregionen erfolgt außerhalb der Kernstädte nach der Bevölkerungsdichte. Um den großräumigen Kontext zu berücksichtigen, wird dann weiter nach der Lage im siedlungsstrukturellen Regionstyp differenziert. (Quelle: [GV-ISys - Verzeichnis der Gebietseinheiten - Definitionen und Beschreibungen 2020](#))

Ausprägungen:

1 = Kreisfreie Großstädte

Kreisfreie Städte mit mind. 100 000 Einwohnern.

2 = Städtische Kreise

Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten von mind. 50% und einer Einwohnerdichte von mind. 150 E./km²; sowie

Kreise mit einer Einwohnerdichte ohne Groß- und Mittelstädte von mind. 150 E./km².

3 = Ländliche Kreise

Kreis mit Verdichtungsansätzen: Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten von mind. 50%, aber einer Einwohnerdichte unter 150 E./km², sowie Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten von unter 50% und mit einer Einwohnerdichte ohne Groß- und Mittelstädte von mind. 100 E./km².

4 = Dünn besiedelte ländliche Kreise

Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten unter 50% und einer Einwohnerdichte ohne Groß- und Mittelstädte unter 100 E./km².

EF26 – Gemeindetyp

Typen des Betriebssitzes nach der [Abgrenzung des Bundesamts für Bauwesen und Raumordnung \(BBR\)](#).

Ausprägungen:

11 = Große Großstädte

22 = Kleinere Mittelstädte

12 = Kleinere Großstädte

40 = Kleine Kleinstädte

30 = Größere Kleinstädte

50 = Landgemeinden

21 = Größere Mittelstädte

EF27 – Nachzügler

Information, ob die Meldung erst nach dem vorläufigen Monatsabschluss eingegangen ist.

Ausprägungen:

0 = Nein

1 = Ja

EF28 – Anzahl Beschäftigte im Unternehmen (URS)

Anzahl der Beschäftigten im gesamten Unternehmen zum 31.12. des Vorjahres des Berichtjahres, beispielsweise der 31.12.2021 für das Kalenderjahr 2022.

EF29– STIA - Schichtnummer

Die Schichtnummer ist eine fortlaufende Nummer zur Kennzeichnung der Stichprobenschichten. Die Einteilung der Schichten erfolgt mittels Wirtschaftsabteilungen (2-Steller der WZ 2008 bis zum Berichtsjahr 2026, ab dem Berichtsjahr 2027 der WZ 2025) und Betriebsbeschäftigtengrößenklassen. Die Zuteilung der Betriebe zu den Beschäftigtengrößenklassen richtet sich nach den Angaben im Unternehmensregister, die nicht immer aktuell sind. Die aktuelle Beschäftigtenzahl entspricht daher in manchen Fällen nicht dieser Beschäftigtengrößenklasse.

EF30 – Gruppennummer

Die Gruppennummer ist eine fortlaufende Nummer zur Kennzeichnung der Stichprobenschichten. Die Einteilung der Gruppen erfolgt mittels Wirtschaftsgruppen (zusammengefasste 2-Steller der WZ 2008 bis zum Berichtsjahr 2026, ab dem Berichtsjahr 2027 der WZ 2025). Im Gegensatz zur Schichtnummer differenziert die Gruppennummer die Betriebe nicht nach Betriebsgrößenklassen.

EF31 – Größenklasse

Bei diesem Merkmal handelt es sich um die Größenklassengliederung im Rahmen der Stichprobenziehung.

Ausprägungen:

Betriebe mit ... bis ... Beschäftigten

1 = 1 – 9

2 = 10 – 49

3 = 50 – 99

4 = 100 – 249

5 = 250 – 499

6 = 500 – 999

7 = 1000 und mehr

EF32 – Fusion/Aufspaltung

Ausprägungen:

0 = Nein

1 = Ja

EF33 – Konjunkturelle Kurzarbeit

Ausprägungen:

0 = Nein

1 = Ja

EF34 – Saisonale Kurzarbeit

Ausprägungen:

0 = Nein

1 = Ja

EF35 – Streik

Ausprägungen:

0 = Nein

1 = Ja

EF36 – Sonstige Gründe

Ausprägungen:

0 = Nein

1 = Ja

EF37 – Zustandsanzeiger

Ausprägungen:

0 = nicht geschätzt

1 = geschätzt

EF38 – Längsschnitthochrechnungsfaktor Berichtsjahr (t) mit imputierten Beschäftigungsverhältnissen

Faktor zur Verwendung bei Analysen über zwei oder mehr Monate innerhalb des Jahres t (z. B. zum Vergleich der Verdienste in einem bestimmten Sektor im März und Juli des Jahres t oder zur Betrachtung der Entwicklung der Verdienste in einem Bundesland von Januar bis Dezember des Jahres t), einschließlich Betriebe und ihre Beschäftigten, die in bestimmten Monaten nicht gemeldet haben und aus dem Vormonat imputiert wurden. Verbleibender Antwortausfall, der nicht imputiert wurde, wird mit diesem Faktor kompensiert. Dieser Faktor, welcher für jeden Betrieb für jeden Monat enthalten ist, ist bei der Analyse dieser Daten stets zu verwenden. Die Ergebnisse unter Verwendung dieses Faktors sind repräsentativ für die Grundgesamtheit der im Januar des Jahres t aktiven Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten in den Wirtschaftsabschnitten A–S (ohne O und Teile von P).² Die Ergebnisse bei Verwendung dieses Faktors können von den Ergebnissen abweichen, die bei Verwendung des Faktors EF40 erzielt werden. Bei Nutzung

² Obwohl die Januar-Population eines Jahres t repräsentiert wird, ist dieser Faktor auch zu verwenden, um Vergleiche von Monaten innerhalb des Jahres t (z.B. März bis Juli) durchzuführen.

des vollständigen Datensatzes, einschließlich imputierter Betriebe, ist stets dieser Faktor EF38 zu verwenden.

EF39 – Längsschnitthochrechnungsfaktor für zwei Berichtsjahre ($t - 1$ bis t) mit imputierten Beschäftigungsverhältnissen

Faktor zur Verwendung bei Analysen, die Monate aus dem Jahr $t - 1$ und dem Jahr t abdecken (diese Monate müssen nicht aufeinanderfolgen; z.B. der Vergleich der Verdienste im Januar des Jahres $t - 1$ mit dem Januar des Jahres t oder die Entwicklung der Verdienste in einem Bundesland von Juli des Jahres $t - 1$ bis Dezember des Jahres t), einschließlich Betriebe und ihre Beschäftigten, die in bestimmten Monaten nicht gemeldet haben und aus dem Vormonat imputiert wurden. Verbleibender Antwortausfall, der nicht imputiert wurde, wird mit diesem Faktor kompensiert. Dieser Faktor, welcher für jeden Betrieb für jeden Monat enthalten ist, ist bei Auswertungen über Monate aus zwei verschiedenen Jahren stets zu verwenden. Die Ergebnisse bei Verwendung dieses Faktors sind repräsentativ für die Grundgesamtheit der im Januar des Jahres $t - 1$ aktiven Betriebe, die auch im Dezember des Jahres t aktiv waren, mindestens einen sozialversicherungspflichtig Beschäftigten haben und den Wirtschaftsabschnitten A–S (ohne O und Teile von P) angehören.³ Die Ergebnisse bei Verwendung dieses Faktors können von den Ergebnissen abweichen, die bei Verwendung des Faktors EF41 erzielt werden. Bei Nutzung des

³ Obwohl die Januar-Population eines Jahres $t-1$ repräsentiert wird, ist dieser Faktor auch zu verwenden, um Vergleiche von Monaten innerhalb der Jahre $t-1$ und t (z.B. März im Jahr $t-1$ bis Juli im Jahr t) durchzuführen.

vollständigen Datensatzes, einschließlich imputierter Betriebe, ist stets dieser Faktor EF39 zu verwenden.

EF40 – Längsschnitthochrechnungsfaktor Berichtsjahr (t) ohne imputierte Beschäftigungsverhältnisse

Faktor zur Verwendung bei Analysen über zwei oder mehr Monate innerhalb des Jahres t (z. B. zum Vergleich der Verdienste in einem bestimmten Sektor im März und Juli des Jahres t oder zur Betrachtung der Entwicklung der Verdienste in einem Bundesland von Januar bis Dezember des Jahres t). Antwortausfall wird mit diesem Faktor, welcher für jeden Betrieb für jeden Monat enthalten ist, kompensiert und ist bei der Analyse dieser Daten stets zu verwenden. Die Ergebnisse unter Verwendung dieses Faktors sind repräsentativ für die Grundgesamtheit der im Januar des Jahres t aktiven Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten in den Wirtschaftsschnitten A–S (ohne O und Teile von P).⁴ Die Ergebnisse bei Verwendung dieses Faktors können von den Ergebnissen abweichen, die bei Verwendung des

⁴ Obwohl die Januar-Population eines Jahres t repräsentiert wird, ist dieser Faktor auch zu verwenden, um Vergleiche von Monaten innerhalb des Jahres t (z.B. März bis Juli) durchzuführen.

Faktors EF38 erzielt werden. Bei Nutzung des ausschließlich erhobenen Datensatzes, ohne imputierter Betriebe, ist stets dieser Faktor EF40 zu verwenden.

EF41 – Längsschnitthochrechnungsfaktor für zwei Berichtsjahre ($t - 1$ bis t) ohne imputierte Beschäftigungsverhältnisse

Faktor zur Verwendung bei Analysen, die Monate aus dem Jahr $t - 1$ und dem Jahr t abdecken (diese Monate müssen nicht aufeinanderfolgen; z.B. der Vergleich der Verdienste im Januar des Jahres $t - 1$ mit dem Januar des Jahres t oder die Entwicklung der Verdienste in einem Bundesland von Juli des Jahres $t - 1$ bis Dezember des Jahres t). Antwortausfall wird mit diesem Faktor, welcher für jeden Betrieb für jeden Monat enthalten ist, kompensiert und ist bei Auswertungen über Monate aus zwei verschiedenen Jahren stets zu verwenden. Die Ergebnisse bei Verwendung dieses Faktors sind repräsentativ für die Grundgesamtheit der im Januar des Jahres $t - 1$ aktiven Betriebe, die auch im Dezember des Jahres t aktiv waren, mindestens einen sozialversicherungspflichtig Beschäftigten haben und den Wirtschaftsabschnitten A–S (ohne O und Teile von P) angehören.⁵ Die Ergebnisse bei Verwendung dieses Faktors können von den Ergebnissen abweichen, die bei Verwendung des Faktors EF39 erzielt werden. Bei Nutzung des ausschließlich erhobenen Datensatzes, ohne imputierter Betriebe, ist stets dieser Faktor EF41 zu verwenden.

2.2 Vergleichbarkeit der Merkmale über die Zeit

Die Werte innerhalb der Verdiensterhebung sind über die Zeit hinweg vergleichbar.

⁵ Obwohl die Januar-Population eines Jahres $t-1$ repräsentiert wird, ist dieser Faktor auch zu verwenden, um Vergleiche von Monaten innerhalb der Jahre $t-1$ und t (z.B. März im Jahr $t-1$ bis Juli im Jahr t) durchzuführen.

2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

Tabelle 1: Häufigkeiten der ungewichteten Betriebe und Beschäftigten ohne und mit Imputation für Deutschland

Berichtsmonat	Anzahl Betriebe ohne Imputation	Anzahl Beschäftigte ohne Imputation	Anzahl Betriebe mit Imputation	Anzahl Beschäftigte mit Imputation
Jan 2022	50 583	8 441 078	50 583	8 441 078
Feb 2022	48 983	8 183 955	51 116	8 636 349
März 2022	47 236	8 055 025	50 126	8 472 318
Apr 2022	50 086	8 562 493	51 037	8 667 615
Mai 2022	49 668	8 572 618	50 702	8 704 995
Jun 2022	48 642	8 459 243	50 224	8 668 333
Jul 2022	50 955	8 663 871	51 604	8 743 503
Aug 2022	50 335	8 653 235	51 366	8 759 860
Sep 2022	49 144	8 597 258	50 791	8 776 703
Okt 2022	51 216	8 789 902	51 706	8 863 326
Nov 2022	50 413	8 715 403	51 523	8 859 065
Dez 2022	49 780	8 555 331	51 046	8 728 551
Jan 2023	53 664	9 021 178	54 211	9 110 826
Feb 2023	53 417	9 048 949	54 107	9 133 680
März 2023	52 469	8 931 181	53 763	9 093 967
Apr 2023	54 137	9 169 130	54 508	9 213 932
Mai 2023	53 970	9 125 510	54 526	9 221 990
Jun 2023	52 981	9 051 608	54 249	9 187 504
Jul 2023	54 145	9 126 740	54 538	9 172 093
Aug 2023	53 770	9 126 152	54 371	9 187 995
Sep 2023	52 359	9 075 842	54 016	9 229 802
Okt 2023	53 717	9 210 269	54 184	9 285 528
Nov 2023	53 304	9 199 018	54 065	9 287 513
Dez 2023	52 342	9 105 305	53 697	9 245 449

Tabelle 2: Häufigkeit der Teilnahmen im Panel

Anzahl Teilnahmen im Panel	Ohne Imputation	Mit Imputation
1	1 059 940	1 059 940
2	733 122	765 471
3	712 216	728 515
4	578 589	588 562
5	494 377	503 354
6	505 612	513 405
7	383 133	390 229
8	392 020	398 933
9	458 012	464 614
10	380 371	385 306
11	464 771	474 105
12	1 760 016	1 775 361
13	193 901	196 191
14	211 223	214 579
15	265 931	269 007
16	251 379	254 231
17	230 299	233 369
18	255 113	258 531
19	231 478	234 891
20	298 902	303 863
21	411 227	416 700
22	408 767	415 007
23	646 734	658 617
24	3 933 261	3 980 648
Insgesamt	15 260 394	15 483 429

2.4 Auswertbare regionale Ebene

Die tiefste vergleichbare regionale Ebene sind die Bundesländer. Ergebnisse auf Ebene der Kreise und Gemeinden sind nicht Teil der Zielsetzung der Statistik.

3 Praktische Hinweise

3.1 Hinweise zur Geheimhaltung

3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung

Unter Geheimhaltung versteht man das Herstellen der absoluten Anonymität der Ergebnisse statistischer Analysen. Konkret bedeutet das, dass im Rahmen der Geheimhaltung sichergestellt wird, dass mit den veröffentlichten Ergebnissen keine Rückschlüsse auf einen Einzelfall (z. B. Person, Betrieb, Einrichtung) gezogen werden können. Statistische Geheimhaltung wird überall dort angewendet, wo statistische Ergebnisse oder Einzeldaten die geschützten Räume der amtlichen Statistik verlassen.

Die Geheimhaltung in der amtlichen Statistik ist in § 16 Bundesstatistikgesetz (BStatG) geregelt und beinhaltet, dass Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik angegeben werden, von den jeweils durchführenden statistischen Stellen geheim zu halten sind, soweit es keine anderslautenden Bestimmungen gibt. Dies wird auch als Statistikgeheimnis bezeichnet. Das Statistikgeheimnis verpflichtet die amtliche Statistik, die erhaltenen Informationen zu schützen, d. h. sie in einer Form zu anonymisieren, die keine Rückschlüsse mehr auf die betreffende Person und den dargelegten Sachverhalt enthält. Die Geheimhaltung ist auch im Hinblick auf die informationelle Selbstbestimmung von besonderem Interesse: Viele Erhebungen der amtlichen Statistik unterliegen der Auskunftspflicht, somit steht es den Befragten nicht frei, selbst zu entscheiden, ob sie eine Information weitergeben möchten. Die amtliche Statistik muss deshalb sicherstellen, dass die erhobenen Daten keinem Befragten zugeordnet werden können.

Das BStatG sieht jedoch auch Fälle vor, in denen das Statistikgeheimnis nicht gilt. In § 16 BStatG sind die Ausnahmen von der Geheimhaltungspflicht dargestellt. Unter anderem wird dort festgelegt, unter welchen Umständen die Daten

der amtlichen Statistik für die Wissenschaft zugänglich gemacht werden dürfen und welche Regeln dabei einzuhalten sind.

3.1.2 Geheimhaltung von Ergebnissen

Um die gesetzlich vorgeschriebene Geheimhaltung von Einzelfällen in den Daten sicherzustellen, müssen alle Ergebnisse, die am Gastwissenschaftlerarbeitsplatz oder per Kontrollierter Datenfernverarbeitung erzeugt werden, vor ihrer Freigabe an den Nutzer von den FDZ einer Geheimhaltungsprüfung unterzogen werden. Dabei stellen die FDZ sicher, dass die Ergebnisse absolut anonym sind und eine Reidentifikation einzelner Befragter nach menschlichem Ermessen ausgeschlossen werden kann. Entsprechend handeln auch die Fachabteilungen der Statistischen Ämter vor der Veröffentlichung von Ergebnissen.

Zur Sicherstellung der Geheimhaltung wenden die FDZ verschiedene Geheimhaltungsregeln an, die jeweils individuell auf die jeweilige Statistik zugeschnitten sind. In der Broschüre „Regelungen zur Auswertung von Mikrodaten in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder“ werden die gebräuchlichsten Regeln zur primären Geheimhaltung dargestellt. Diese Regeln werden in den FDZ im Grunde auf alle Statistiken angewendet. Die Anlage dieser Broschüre enthält Informationen darüber, welche Geheimhaltungsregeln auf welche Statistiken anzuwenden sind.

Die Broschüre finden Sie hier: <http://www.forschungsdatenzentrum.de/geheimhaltung.asp>.

3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen

Treten in den erstellten Analysen Geheimhaltungsfälle auf, werden diese Werte von den FDZ zur Sicherstellung der Geheimhaltung durch ein Sperrmuster ersetzt. Gerade in Kreuztabellen entstehen so durch die notwendige Sekundärsperrung schnell viele „Löcher“ in den Auswertungen. Da eine einmal zur Sekundärsperrung herangezogene Tabellenzelle auch in allen folgenden Analysen gesperrt werden muss (tabellenübergreifende Geheimhaltung) – auch, wenn es in der neu erstellten Tabelle nicht nötig wäre – ist es sinnvoll, bei jeder Ergebniserstellung darauf zu achten, dass möglichst keine Geheimhaltungsfälle erzeugt werden. Treten in einem Output Geheimhaltungsfälle auf, steht es dem betreuenden FDZ frei, die Prüfung und Freigabe des Outputs abzulehnen.

Um Geheimhaltungsfälle in den Analysen zu vermeiden, sollte immer darauf geachtet werden, dass die erstellten Analysen auf ausreichend großen Fallzahlen beruhen. Bei geringen Fallzahlen empfiehlt es sich, Variablenausprägungen zusammenzufassen und damit größere Fallzahlen zu erzielen.

3.2 FAQ

Bitte wenden Sie sich bei auftretenden Fragen an den im Impressum für fachliche Informationen genannten FDZ-Standort.

3.3 Verfügbare Tools

Für dieses Produkt werden seitens der Forschungsdatenzentren keine weiterführenden Tools angeboten.

Anhang

Tabelle 1: Merkmalsübersicht

Merkmals	Bezeichnung
EF1	Berichtsmonat
EF2	Berichtsjahr
EF3	Beschäftigten-ID, anonymisierter Identifikator
EF4	BerichtseinheitID (Identnummer des Betriebs (URS))
EF5	Geschlecht
EF6	Geburtsdatum
EF7	Erste Staatsangehörigkeit
EF8	Eintrittsdatum
EF9	Personengruppe
EF10U1	Tätigkeitsschlüssel Stellen 1-5
EF10U2	Tätigkeitsschlüssel Stelle 6
EF10U3	Tätigkeitsschlüssel Stelle 7
EF10U4	Tätigkeitsschlüssel Stelle 8
EF10U5	Tätigkeitsschlüssel Stelle 9
EF11	Bezahlte Arbeitsstunden ohne Überstunden im Monat
EF12	Bezahlte Überstunden im Monat
EF13	Bruttomonatsverdienst – Gesamtbruttoentgelt
EF14	Sonstige Bezüge (EBV §1 Absatz 2 Nummer 2a)
EF15	Bruttomonatsverdienst für bezahlte Überstunden
EF16	Bruttomonatsverdienst für Zuschläge für Schicht-, Samstags-, Sonntags-, Feiertags- oder Nachtarbeit
EF17	Bruttomonatsverdienst für Zwecke der Entgeltumwandlung
EF18	Bundesland
EF19	Regionalschlüssel (AGS)
EF20	Wirtschaftszweig der aktuellen Klassifikation
EF21	Wirtschaftszweig alt
EF22	Tarifbindung des Betriebs
EF23	Regionsgrundtyp
EF24	Differenzierter Regionstyp
EF25	Kreistyp
EF26	Gemeindetyp
EF27	Nachzügler
EF28	Anzahl Beschäftigte im Unternehmen (URS)
EF29	STIA-Schichtnummer
EF30	Gruppennummer
EF31	Größenklasse
EF32	Fusion/Aufspaltung
EF33	Konjunkturelle Kurzarbeit
EF34	Saisonale Kurzarbeit
EF35	Streik

Merkmal	Bezeichnung
EF36	Sonstige Gründe
EF37	Zustandsanzeiger
EF38	Längsschnitthochrechnungsfaktor Berichtsjahr (t) mit imputierten Beschäftigungsverhältnissen
EF39	Längsschnitthochrechnungsfaktor für zwei Berichtsjahre (t-1 bis t) mit imputierten Beschäftigungsverhältnissen
EF40	Längsschnitthochrechnungsfaktor Berichtsjahr (t) ohne imputierte Beschäftigungsverhältnisse
EF41	Längsschnitthochrechnungsfaktor für zwei Berichtsjahre (t-1 bis t) ohne imputierte Beschäftigungsverhältnisse

