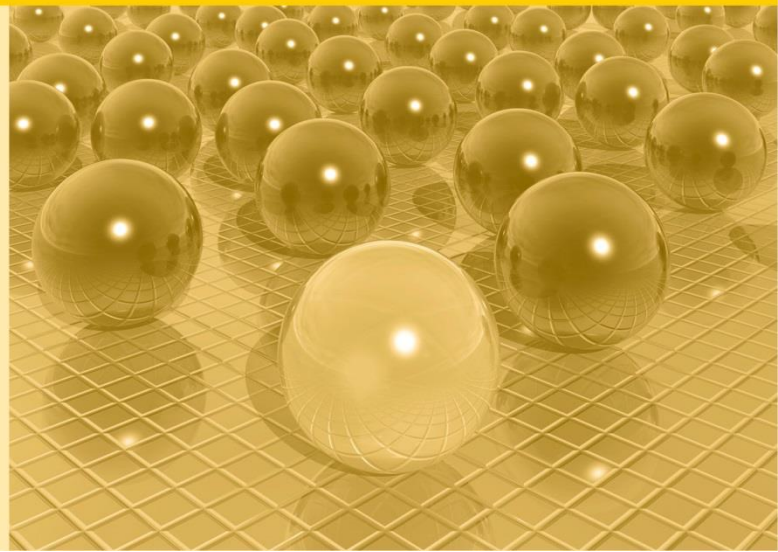


# Metadatenreport



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

DOI: 10.21242/62361.2023.04.00.1.1.0

Version 1

# Metadatenreport

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

DOI: 10.21242/62361.2023.04.00.1.1.0

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](http://www.forschungsdatenzentrum.de)  
E-Mail: [forschungsdatenzentrum@it.nrw.de](mailto: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](mailto:forschungsdatenzentrum@statistik.hessen.de)

### Informationen zum Datenangebot:

Statistisches Bundesamt  
Forschungsdatenzentrum  
Tel.: 0611 75-2420  
Fax: 0611 72-3915  
[forschungsdatenzentrum@destatis.de](mailto:forschungsdatenzentrum@destatis.de)

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

Erscheinungsfolge: unregelmäßig  
Erschienen im April 2024

Diese Publikation wird kostenlos als PDF-Datei zum Download unter [www.forschungsdatenzentrum.de](http://www.forschungsdatenzentrum.de) angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2024  
(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 April 2023 per On-Site-Nutzung (EVAS-Nummer: 62361). Version 1. DOI: 10.21242/62361.2023.04.00.1.1.0. Wiesbaden 2024.

## Inhalt

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

# **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.

Aus Sicht des Statistischen Landesamts Sachsen reicht der Anonymisierungsschutz nicht aus, sodass einzelne Datensätze Sachsens aus den On-Site-Datensätzen entfernt werden.

## **1.3 Methodik der Verknüpfung**

Da zur Erstellung dieses Produkts keine Daten verknüpft wurden, entfällt dieser Punkt.

# **2 Produkt**

Die Verdiensterhebung (VE) wird monatlich erhoben. Für die Jahresbetrachtung der absoluten Werte wird jedoch der Monat April herangezogen, da nur hier eine Vollimputation der Betriebe sowie die gebundene Hochrechnung durchgeführt wird. Es liegen Angaben zu Betrieben und Arbeitnehmerinnen und Arbeitnehmern vor. Die Daten eignen sich gut zur Analyse der Verdienstverteilung, von geschlechtsspezifischen Lohnunterschieden sowie zur Untersuchung der Verdienstunterschiede in tarifgebundenen Betrieben im Vergleich

zu solchen, die nach freier Vereinbarung vergüten. Da die Stichprobenauswahl auf Bundeslandebene erfolgt, lassen sich für kleinräumigere regionale Gliederungen keine repräsentativen Ergebnisse erzielen.

Die Statistik enthält Informationen zur Person (Geschlecht, Geburtsdatum (MMJJJJ), Staatsangehörigkeit, Schul- und Ausbildungsabschluss), zur Tätigkeit (Tätigkeitsschlüssel, Personengruppenschlüssel, Eintrittsdatum (MMJJJJ)), zur Arbeitszeit (bezahlte Stunden und Überstunden) und zum Verdienst (Gesamtbruttoentgelt, Zulagen für Schicht-/Nachtarbeit, Überstundenverdienste, Betriebliche Altersvorsorge (Entgeltumwandlung), Sonderzahlungen). Auf Betriebsebene gibt es beispielsweise Angaben zum Sitz des Betriebes, dem Wirtschaftszweig, der Tarifbindung und der Anzahl der Beschäftigten im zugehörigen Unternehmen.

Die Einzeldaten der VE April 2023 umfassen 9,4 Mio. Beschäftigtendatensätze, die sich auf 68 861 Betriebe aus der Stichprobenziehung, der Personalstandstatistik und dem Verwaltungsdatenspeicher aufteilen. Die Stichprobenziehung basiert auf den Daten der Bundeskopie des statistischen Unternehmensregisters aus dem Berichtsjahr 2021 mit Stand August 2022 und umfasst maximal 58 000 Betriebe. Bei 10 000 Betrieben mit ausschließlich geringfügig entlohnten Beschäftigten wurden die Angaben nicht erhoben oder aus der Personalstandstatistik bezogen, sondern aus dem Verwaltungsdatenspeicher imputiert. Die im Folgenden beschriebenen Daten können in den Forschungszentren der Statistischen Ämter des Bundes und der Länder On-Site, also am Gastwissenschaftsarbeitsplatz und über die kontrollierte Datenfernverarbeitung genutzt werden.

## **2.1 Merkmale und Merkmalsbeschreibung**

### **2.1.1 Datensatzbeschreibung /Schlüsselverzeichnis**

Die Angaben für April 2023 beziehen sich auf das endgültige Material VE050X.

### **2.1.2 Merkmalsdefinitionen**

#### **F1 – Berichtsmonat**

#### **F2 – Berichtsjahr**

Es handelt sich um die erhobenen Einzeldaten für den April 2023.

#### **F3 – 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 und als Identifikator der Betriebe im URS verwendet wird.

Betriebe ohne SV-pflichtige Beschäftigte, deren Angaben geschätzt werden, haben an der ersten Stelle der Identnummer eine „8“, Betriebe aus der Personalstandstatistik beginnen mit einer „9“.



## **F4 – Bundesland**

Amtlicher Schlüssel des Bundeslandes, in dem sich der Betrieb zum Zeitpunkt der Stichprobenziehung (Herbst 2022) 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 „F5“ nicht übereinstimmen (sog. Landeswechsler).

## **F5 – 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.

Stelle 4+5: Ergibt zusammen mit den Stellen 1-3 die Kennziffer des Kreises.

Stelle 6-8: Ergibt zusammen mit den Stellen 1-5 die Kennziffer der Gemeindekennzahl.

Für Betriebe aus der Personalstandstatistik liegt der AGS nicht vor, sodass nur der zweistellige Länderschlüssel vorhanden ist.

Aufgrund der unterschiedlichen Zeitpunkte des Abzugs der Informationen (Stichprobenziehung und Berichtsjahr), kann bei Umzügen der Betriebe die Angabe mit denen aus „F4“ nicht übereinstimmen (sog. Landeswechsler).

## **F6 – Design-Hochrechnungsfaktor**

Design-Hochrechnungsfaktor bei freier Hochrechnung. Er ergibt sich aus der Anzahl aller Betriebe in der Auswahlgrundlage einer Schicht geteilt durch die Anzahl der Betriebe der Schicht, die in der Stichprobe enthalten sind.

## **F7 – Ergänzungsfaktor (ERGF)**

Der Ergänzungsfaktor dient zur Berücksichtigung der „echten Antwortausfälle“ (Antwortverweigerungen) bei der freien Hochrechnung. Er errechnet sich durch die Division der Anzahl der angeschriebenen Betriebe durch die Anzahl der Betriebe mit brauchbaren Antworten einschließlich der „unechten Ausfälle“ (z. B. wegen Konkurs oder weil der Betrieb nicht mehr zur Auswahlgesamtheit gehört), die dabei als Antworten gezählt werden. Bei der Verwendung von „F9“ für Hochrechnungen spielt dieser Faktor keine Rolle, da bei „F9“ Korrekturen bereits berücksichtigt sind.

## **F8 – Hochrechnungsfaktor (Produkt aus Design-HRF und ERGF)**

Freier Hochrechnungsfaktor, der sich aus dem Produkt des Design-HRF und des Ergänzungsfaktors ergibt.

## **F9 – Hochrechnungsfaktor (gebunden)**

Hochrechnungsfaktor der gebundenen Hochrechnung, mit dem die Angaben für die Betriebe und die Beschäftigten gewichtet werden müssen (vgl. Metadaten zur VE Teil I; Abschnitt 2.5). Die freie Hochrechnung („F8“) unterschätzt die tatsächlichen absoluten Anzahlen und Summen der Grundgesamtheit. Das liegt daran, dass die Auswahlgrundlage der Stichprobe nicht aus dem Berichtsjahr stammt, sondern älter ist (siehe [Qualitätsbericht der VE](#)). Das führt sowohl zu einer Überabdeckung der Stichprobe (bei Betriebsschließungen) als auch zu einer Unterabdeckung (bei Betriebsgründungen). Die Unterabdeckung verursacht die Unterschätzung der absoluten Statistiken. Relative Statistiken, wie Anteile oder Mittelwerte, sind davon kaum betroffen. *Dieser Hochrechnungsfaktor wird daher zur Verwendung empfohlen.*

Die gebundene Hochrechnung der VE 2023 erfolgte auf Betriebsebene nach der Methode Generalised regression estimator (siehe [Qualitätsbericht der VE](#)). Die für dieses Verfahren typischen Hilfsvariablen waren die Anzahl der sozialversicherungspflichtig Beschäftigten und der geringfügig entlohnten Beschäftigten des Betriebs sowie die Zahl der Betriebe mit abhängig Beschäftigten laut Verwaltungsdatenspeicher der Statistischen Ämter. Hochgerechnete Anzahlen von Betrieben und Beschäftigungsverhältnissen der VE 2023 sind dadurch kohärent zu anderen Statistiken.

## **F10 – Tarifbindung**

- 1 = Branchentarifvertrag, gültig in allen oder mehreren Bundesländern
- 2 = Firmentarifvertrag
- 3 = kein Tarifvertrag, Orientierung der Verdienste an einem Branchentarifvertrag
- 4 = kein Tarifvertrag, freie Verdienstvereinbarung

## **F11 – Wirtschaftszweig der Klassifikation WZ 2008**

Als Wirtschaftszweig oder Branche bezeichnet man eine Gruppe von Firmen, die ähnliche Produkte herstellen oder ähnliche Dienstleistungen erbringen und somit den wirtschaftlichen Schwerpunkt ihrer Firma setzen.

Die Codes und die zugehörigen Wirtschaftszweige finden sich in der Klassifikation der Wirtschaftszweige WZ 2008.

## **F12 – Personalnummer (gehasht)**

Bei der Personalnummer handelt es sich um einen (betrieblichen) Identifikator des Arbeitnehmers/der Arbeitnehmerin. Sollte keine Personalnummer vorliegen, wird an dieser Stelle vom Betrieb eine eindeutige, im Zeitverlauf gleichbleibende Ordnungsnummer für die Beschäftigte/den Beschäftigten angegeben.

Die Personalnummern der Beschäftigten aus den erhobenen Betrieben sowie der Personalstandstatistik eignen sich für eine Längsschnittverknüpfung der Beschäftigten zwischen verschiedenen Zeitpunkten der VE. Dies ist mit den

Beschäftigten aus den imputierten Betrieben (siehe F34) hingegen nicht möglich, da gleichen Nummern unterschiedliche Beschäftigte zu Grunde liegen können.

Für eine eindeutige Identifikation der Beschäftigten ist die Kombination aus F3, F12 und F13 zu verwenden.

Für Analysen im FDZ wird eine systemfreie laufende Nummer zur Verfügung gestellt.

### **F13 – Geburtsdatum**

Das Geburtsdatum ist nach dem Muster MMJJJJ hinterlegt.

### **F14 – Erste Staatsangehörigkeit**

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

Für Beschäftigte aus der Personalstandstatistik liegen keine Angaben vor.

### **F15 – 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 Beschäftigungsbegins laut Entgeltbescheinigungsverordnung (EBV) § 1 Absatz 1.

## F16 – 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.<sup>1</sup>

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)
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

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

## **F17 – Tätigkeitsschlüssel**

Es handelt sich um einen neunstelligen Schlüssel der Bundesagentur für Arbeit, der durch den Arbeitgeber zur Sozialversicherung an die Krankenkassen gemeldet wird. Er setzt sich wie folgt zusammen:

Stelle 1-5: Ausgeübte Tätigkeit (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: Wiemer, S.; Reimer, K. und Lewerenz, J.: Einführung der Klassifikation der Berufe 2010 in die Arbeitsmarktstatistik, Nürnberg 2011.

Stelle 6: Höchster allgemeinbildender Schulabschluss

Ausprägungen:

1 = Ohne Schulabschluss

2 = Haupt-/Volksschulabschluss

3 = Mittlere Reife oder gleichwertiger Abschluss

4 = Abitur/Fachabitur

9 = Abschluss unbekannt

Die Ausprägung "Abschluss unbekannt" wurde nach dem Nearest-Neighbour-Prinzip mit der kanadischen Software CANCEIS durch imputierte Werte ersetzt.

### 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.

### Stelle 8: Arbeitnehmerüberlassung

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

Ausprägungen:

1 = Nein

2 = Ja



### Stelle 9: Vertragsform

Ausprägungen:

1 = Vollzeit, unbefristet

2 = Teilzeit, unbefristet

3 = Vollzeit, befristet

4 = Teilzeit, befristet

### **F18 – 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.

### **F19 – Bezahlte Arbeitsstunden ohne Überstunden im Monat**

Das Merkmal gibt für jeden Beschäftigten die bezahlten Arbeitsstunden im Monat an, unabhängig davon, ob deren 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.

## **F20 – 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.

## **F21 – 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 Datennutzer 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.

### **F22 – Sonstige Bezüge (EBV § 1 Absatz 2 Nummer 2a)**

Es handelt sich um die Darunter-Position des Gesamtbruttoentgelts „Sonstige Bezüge“ laut § 1 Absatz 2 Nummer 2a Entgeltbescheinigungsverordnung (EBV) plus Zuschüsse des Arbeitgebers zum Kurzarbeitergeld.

### **F23 – 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.

### **F24 – 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.

### **F25 – 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 des Beschäftigten. Einbezogen sind alle Durchführungswege (Pensionskasse, Pensionsfonds, Direktversicherung, auch Direktzusage und Unterstützungskasse) und Besteuerungsformen (steuerfrei, pauschal, individuell versteuert).

## **F26 – Regionsgrundtyp**

Regionen 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 lebt 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 km<sup>2</sup>.)

2 = Regionen mit Verdichtungsansätzen

(Regionen, in denen mindestens 33% der Bevölkerung in Groß- und Mittelstädten lebt, mit einer Einwohnerdichte zwischen 150 und 300 Einwohner pro km<sup>2</sup> sowie Regionen, in denen sich mindestens eine Großstadt befindet und die eine Einwohnerdichte ohne Berücksichtigung der Großstädte von mindestens 100 Einwohner pro km<sup>2</sup> aufweisen.)

3 = Ländliche Regionen

(Regionen, in denen weniger als 33% der Bevölkerung in Groß- und Mittelstädten lebt, mit einer Einwohnerdichte unter 150 Einwohner pro km<sup>2</sup> sowie Regionen, in denen sich zwar eine Großstadt befindet, aber die eine Einwohnerdichte ohne Berücksichtigung der Großstädte unter 100 Einwohner pro km<sup>2</sup> hat.)

Missing = Fehlende Angabe

(Beschäftigte, die aus der Personalstandstatistik stammen.)

## **F27 – Differenzierter Regionstyp**

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

Ausprägungen:

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

02 = „Ländlicher Raum“ (Ländliche Kreise)

Missing = Fehlende Angabe

(Beschäftigte, die aus der Personalstandstatistik stammen.)

## **F28 – Kreistyp**

Typen 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 2015](#))

#### 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./km<sup>2</sup>, sowie Kreise mit einer Einwohnerdichte ohne Groß- und Mittelstädte von mind. 150 E./km<sup>2</sup>.)

3 = Ländliche Kreise

(Kreis mit Verdichtungsansätzen: Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten von mind. 50%, aber einer Einwohnerdichte unter 150 E./km<sup>2</sup>, sowie Kreise mit einem Bevölkerungsanteil in Groß- und Mittelstädten von unter 50% und mit einer Einwohnerdichte ohne Groß- und Mittelstädte von mind. 100 E./km<sup>2</sup>.)

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./km<sup>2</sup>.)

Missing = Fehlende Angabe

(Beschäftigte, die aus der Personalstandstatistik stammen.)

## **F29 – Gemeindetyp**

Typen 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

. = Fehlende Angabe

(Beschäftigte, die aus der Personalstandstatistik stammen.)

## **F30 – Nachzügler**

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

Ausprägungen:

0 = Nein

1 = Ja

### **F31 – Referenzdatensatz aus dem Vormonat**

Information, ob es sich um plausibilisierte Vormonatsdaten handelt, mit denen die komplett identisch zur Vorperiode gemeldeten Datensätze ersetzt wurden.

#### Ausprägungen:

0 = Nein

1 = Ja

### **F32 – Zustandsanzeiger**

#### Ausprägungen:

0 = nicht geschätzt

1 = geschätzt

### **F33 – Anzahl Beschäftigte im Unternehmen (URS)**

Anzahl der Beschäftigten im gesamten Unternehmen zum 31.12.2022. Bei den Datensätzen, die aus der Personalstandstatistik generiert wurden, ist das Merkmal durchgängig mit 99999 kodiert.

### **F34 – STIA - Schichtnummer**

Die Schichtnummer ist eine fortlaufende Nummer zur Kennzeichnung der Stichprobenschichten. Die Einteilung der Schichten erfolgt mittels Wirtschaftsabteilungen (2-Steller der WZ 2008) 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.

Die Herkunft der Betriebe für den Datensatz lässt sich wie folgt ermitteln:

< 600 = Betriebe aus der Erhebung

600 = Betriebe aus der Personalstandstatistik

>= 700 = Imputierte Betriebe aus dem Verwaltungsdatenspeicher

### **F35 – Gruppennummer**

Die Gruppennummer ist eine fortlaufende Nummer zur Kennzeichnung der Stichprobenschichten. Die Einteilung der Gruppen erfolgt mittels Wirtschaftsgruppen (zusammengefasste 2-Steller der WZ 2008). Im Gegensatz zur Schichtnummer differenziert die Gruppennummer die Betriebe nicht nach Betriebsgrößenklassen.

### **F36 – Größenklasse**

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

#### Ausprägungen:

Betriebe mit ... bis ... Arbeitnehmern

1 = 1 – 9

2 = 10 – 49

3 = 50 – 99

4 = 100 – 249

5 = 250 – 499

6 = 500 – 999

7 = 1000 und mehr

### **F37 – Bundesland lt. Stichprobe**

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

#### Ausprägungen:

01 = Schleswig-Holstein

07 = Rheinland-Pfalz

02 = Hamburg

08 = Baden-Württemberg

03 = Niedersachsen

09 = Bayern

04 = Bremen

10 = Saarland

05 = Nordrhein-Westfalen

11 = Berlin

06 = Hessen

12 = Brandenburg

13 = Mecklenburg-Vorpommern

14 = Sachsen

15 = Sachsen-Anhalt

16 = Thüringen

### **F38 – Fusion/Aufspaltung**

Ausprägungen:

0 = Nein

1 = Ja

### **F39 – Konjunkturelle Kurzarbeit**

Ausprägungen:

0 = Nein

1 = Ja

### **F40 – Saisonale Kurzarbeit**

Ausprägungen:

0 = Nein

1 = Ja

### **F41 – Streik**

Ausprägungen:

0 = Nein

1 = Ja

## **F42 – Sonstige Gründe**

### Ausprägungen:

0 = Nein

1 = Ja

## **F43 – Handwerkszugehörigkeit**

### Ausprägungen:

0 = nicht im Handwerk

1 = im Handwerk

## **F44 bis F53 (AN bis NRESP)**

Bei den Merkmalen AN bis NRESP handelt es sich um Felder, die für die gebundene Hochrechnung benötigt wurden bzw. für die Abschätzung des relativen Standardfehlers mit der Software %CLAN unter SAS benötigt werden. Für wissenschaftliche Forschungsvorhaben sind diese Merkmale ungeeignet und stehen daher auch nicht zur Verfügung, sie werden hier nur der Vollständigkeit halber aufgeführt.

F44 – AN - Anzahl Betriebe

F45 – SV - Anzahl SV-Beschäftigte

F46 – GB - Anzahl geringfügig entlohnte Beschäftigte

F47 – QK - Korrekturfaktor für Homoskedastizität

F48 – BLOCK - Bundesland

F49 – GKL5 - Größenklasse des Betriebs

F50 – WZ18 - Wirtschaftsabschnitt des Betriebs

F51 – STRATID - Schichtidentifikator

F52 – NPOP – Anzahl Grundgesamtheit (geschätzt)

F53 – NRESP - Anzahl Respondenten

## **2.2 Vergleichbarkeit der Merkmale über die Zeit**

Die Angaben innerhalb der Statistik der Verdiensterhebung sind über die Zeit vergleichbar. Aufgrund des ähnlichen Erhebungskonzepts sind die Angaben weitestgehend auch mit den Ergebnissen der Verdienststrukturerhebung 2018 und 2014 vergleichbar.

## **2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen**

-

## **2.4 Auswertbare regionale Ebene**

Die Erhebung erfüllt die Anforderungen von Eurostat. Entsprechend sind die Ergebnisse EU-weit vergleichbar, in tiefster regionaler Ebene nach NUTS1 "Nomenclature des unités territoriales statistiques" (Systematik der Gebiets-einheiten für die Statistik), was in Deutschland den Bundesländern entspricht. 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 Sekundärspernung schnell viele „Löcher“ in den Auswertungen. Da eine einmal zur Sekundärspernung 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 zusammen zu fassen 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**

Merkmalskennung	Bezeichnung
F1	Berichtsmonat
F2	Berichtsjahr
F3	BERICHTSEINHEITID (Identnummer des Betriebs (URS))
F4	Bundesland
F5	Regionalschlüssel (AGS)
F6	Design-Hochrechnungsfaktor
F7	Ergänzungsfaktor (ERGF)
F8	Hochrechnungsfaktor (Produkt aus Design-HRF und ERGF)
F9	Hochrechnungsfaktor (gebunden)
F10	Tarifbindung
F11	Wirtschaftszweig der Klassifikation WZ2008
F12	Personalnummer (gehasht)
F13	Geburtsdatum
F14	Erste Staatsangehörigkeit
F15	Eintrittsdatum
F16	Personengruppenschlüssel
F17	Tätigkeitsschlüssel
F18	Geschlecht
F19	Bezahlte Arbeitsstunden ohne Überstunden im Monat
F20	Bezahlte Überstunden im Monat
F21	Bruttomonatsverdienst – Gesamtbruttoentgelt
F22	Sonstige Bezüge (EBV §1 Absatz 2 Nummer 2a)
F23	Bruttomonatsverdienst für bezahlte Überstunden
F24	Bruttomonatsverdienst für Zuschläge für Schicht-, Samstags-, Sonntags-, Feiertags- oder Nachtarbeit
F25	Bruttomonatsverdienst für Zwecke der Entgeltumwandlung
F26	Regionsgrundtyp
F27	Differenzierter Regionstyp
F28	Kreistyp
F29	Gemeindetyp
F30	Nachzügler
F31	Referenzdatensatz aus dem Vormonat
F32	Zustandsanzeiger
F33	Anzahl Beschäftigte im Unternehmen (URS)

<b>Merkmal</b>	<b>Bezeichnung</b>
F34	STIA
F35	Gruppennummer
F36	Größenklasse
F37	Bundesland lt. Stichprobe
F38	Fusion/Aufspaltung
F39	Konjunkturelle Kurzarbeit
F40	Saisonale Kurzarbeit
F41	Streik
F42	Sonstige Gründe
F43	Handwerkszugehörigkeit
F44	AN - Anzahl Betriebe
F45	SV - Anzahl SV-Beschäftigte
F46	GB - Anzahl geringfügig entlohnte Beschäftigte
F47	QK - Korrekturfaktor für Homoskedastizität
F48	BLOCK - Bundesland
F49	GKL5 - Größenklasse des Betriebs
F50	WZ18 - Wirtschaftsabschnitt des Betriebs
F51	STRATID - Schichtidentifikator
F52	NPOP – Anzahl Grundgesamtheit (geschätzt)
F53	NRESP - Anzahl Respondenten

Statistische Ämter des Bundes und der Länder,  
Metadatenreport – Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung April 2023 per On-  
Site-Nutzung (EVAS-Nummer: 62361)

Fotorechte Umschlag: ©artSILENCEcom – Fotolia.com