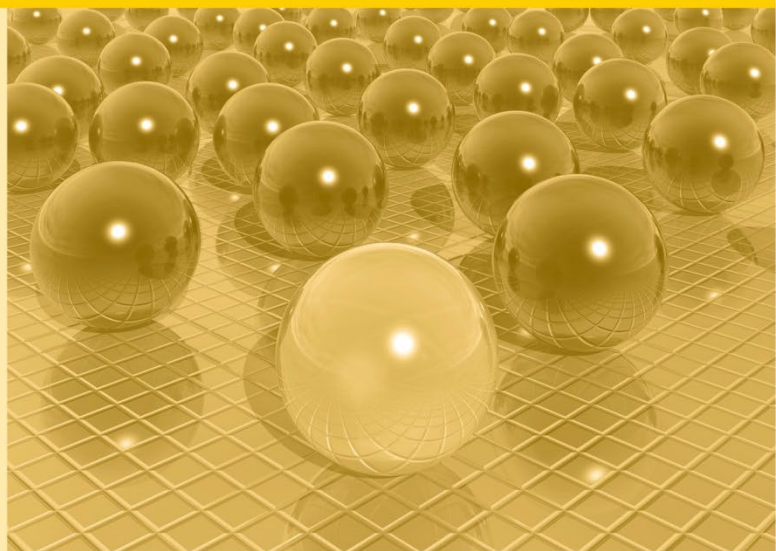


Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung
April 2022 per Scientific Use File (EVAS: 62361)

DOI: 10.21242/62361.2022.04.00.3.1.0

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 November 2023

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

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2023
(im Auftrag der Herausbergemeinschaft)

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

Fotorechte Umschlag: © artSILENCE – Fotolia.com

Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung April 2022 per Scientific Use File (EVAS-Nummer: 62361). Version 1. DOI: 10.21242/62361.2022.04.00.3.1.0. Wiesbaden 2023.

Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung der Verdiensterhebung
April 2022 per Scientific Use File (EVAS: 62361)

DOI: 10.21242/62361.2022.04.00.3.1.0

Version 1

Inhalt

1	Datenaufbereitung in den FDZ	2
1.1	Datenaufbereitung	2
1.2	Anonymisierungsmaßnahmen	3
1.3	Methodik der Verknüpfung	10
2	Produkt	10
2.1	Merkmale und Merkmalsbeschreibung	11
2.1.1	Datensatzbeschreibung /Schlüsselverzeichnis	11
2.1.2	Merkmalsdefinitionen	11
2.2	Vergleichbarkeit der Merkmale über die Zeit	35
2.3	Eckwerte relevanter Merkmale und Merkmalskombinationen	36
2.4	Auswertbare regionale Ebene	36
3	Praktische Hinweise	37
3.1	Hinweise zur Geheimhaltung	37
3.1.1	Gesetzliche Grundlagen der statistischen Geheimhaltung	37
3.1.2	Geheimhaltung von Ergebnissen	38
3.1.3	Praktische Tipps zur Vermeidung von Geheimhaltungsfällen	39
3.2	FAQ	39
3.3	Verfügbare Tools	39
	Anhang	40

1 Datenaufbereitung in den FDZ

1.1 Datenaufbereitung

Scientific Use Files (SUF) sind standardisierte Einzeldatensätze, die von den Statistischen Ämtern für wissenschaftliche Vorhaben bereitgestellt werden. Im Gegensatz zu Gastwissenschaftsarbbeitsplätzen oder der kontrollierten Datenfernverarbeitung können SUF außerhalb der geschützten Räume der amtlichen Statistik genutzt werden.

§ 16 Abs. 6 Bundesstatistikgesetz (BStatG) ermöglicht den Statistischen Ämtern des Bundes und der Länder Einzelangaben an Hochschulen und andere Einrichtungen der unabhängigen wissenschaftlichen Forschung zu übermitteln. Voraussetzung hierfür ist jedoch, dass die Einzelangaben nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können. „Unverhältnismäßig“ bedeutet, dass der Aufwand für eine Reidentifikation höher ist, als der Nutzen, der daraus gezogen wird.

SUF haben daher stets zwei sehr gegensätzliche Anforderungen zu erfüllen. Auf der einen Seite sollen sie für einen möglichst großen Teil an Forschungsvorhaben ein Höchstmaß an Analysepotenzial bieten, zum anderen müssen sie den strengen Maßgaben der faktischen Anonymität genügen.

Den Anforderungen des BStatG zur Sicherstellung der faktischen Anonymität wird bei einem SUF zum einen dadurch entsprochen, dass mit dem Datennutzer ein Vertrag geschlossen wird, in dem ein De-Anonymisierungsverbot festgeschrieben ist und Sanktionen bei Nichteinhaltung dieses Verbotes festge-

setzt werden. Zum anderen wird die faktische Anonymität durch datenverändernde Maßnahmen erreicht. Im Folgenden werden die datenverändernden Maßnahmen beschrieben.

Für detaillierte Informationen zur Verdiensterhebung (VE) wird an dieser Stelle auf die Metadaten zur Erhebung (Teil I) verwiesen.

1.2 Anonymisierungsmaßnahmen

Für den SUF zur VE April 2022 werden nur die Merkmale bereitgestellt, die sich für die Durchführung wissenschaftlicher Vorhaben eignen. Hilfsmerkmale oder direkte Identifikatoren wie beispielsweise die Betriebsnummer aus dem Unternehmensregister sind im SUF nicht enthalten. Eine Übersicht über die im SUF enthaltenen Merkmale findet sich weiter unten.

Neben der Löschung von Merkmalen sind weitere Maßnahmen erforderlich, um die faktische Anonymität der Daten sicherzustellen. Diese Maßnahmen stellen im Wesentlichen auf die Merkmale ab, die sowohl in den vertraulichen Daten als auch in dem möglichen Zusatzwissen eines potentiellen Datenangreifers enthalten sein können (so genannte Überschneidungsmerkmale). Bei der Recherche des möglichen Zusatzwissens haben sich insbesondere die Regionalangabe, der Wirtschaftszweig sowie die Anzahl der Beschäftigten als mögliche Überschneidungsmerkmale herausgestellt.

In der VE meldet ein Betrieb die angeforderten Angaben für alle Beschäftigten, die sich zum Zeitpunkt der Erhebung im Betrieb befinden. Durch die Kenntnis aller Beschäftigten in einem Betrieb und deren betrieblicher und persönlicher Eigenschaften kann es jedoch möglich sein, dass einzelne Beschäftigte iden-

tifiziert werden können (z.B. ein Betrieb in einer kleinen Branche, der insgesamt 30 Beschäftigte und davon zwei weibliche Beschäftigte hat). Um die faktische Anonymität sicherzustellen, wird im SUF eine Stichprobe der Beschäftigten pro Betrieb bereitgestellt. Die Ziehung der Stichprobe erfolgt zufällig und umfasst einen zufällig vergebenen Auswahlatz zwischen 60 und 80 Prozent der gemeldeten Beschäftigten. Hochgerechnete absolute Zahlen des SUF weichen daher um ca. 30 Prozent von den Hochrechnungen mit dem On-Site Material ab.

Bei einzelnen Betrieben kann auch nach Durchführung aller unten aufgeführten Maßnahmen kein ausreichender Schutz erreicht werden. Die betroffenen Datensätze der Beschäftigten werden daher vollständig aus dem Datensatz entfernt.

Im Folgenden werden die Maßnahmen beschrieben, die auf die einzelnen Merkmale angewendet wurden:

F3 – Identnummer des Betriebs (URS)

Das Merkmal wird durch eine systemfreie Nummer ersetzt.

F5 – Regionalschlüssel (AGS)

Der in den Originaldaten bis auf die Gemeindeebene ausgewiesene Regionalschlüssel wird zu fünf (bzw. für WZ-Abschnitt B sechs) Regionen vergrößert:

Region 1: Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, Berlin

Region 2: Nordrhein-Westfalen

Region 3: Hessen, Rheinland-Pfalz, Saarland

Region 4: Baden-Württemberg, Bayern

Region 5: Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen

Region 6: Für Betriebe mit weniger als 500 Beschäftigten wurde im Wirtschaftszweig-Abschnitt B „Bergbau und Gewinnung von Steinen und Erden“ ein erhöhtes Reidentifikationsrisiko festgestellt. Dieses Risiko ist darauf zurück zu führen, dass Nordrhein-Westfalen eine eigene Kategorie (Region 2) bildet und hier Betriebe leichter zu identifizieren sein können. Für den WZ-Abschnitt B wurde daher die Regionalangabe weiter vergrößert indem die Regionen 1 bis 4 zu einer Region 6 zusammengefasst wurden.

F10 – Tarifbindung

Sofern es bei einem Wirtschaftszweig (vgl. „F11“) in einer Region („F5“) weniger als drei Betriebe gibt, die einen Firmentarifvertrag oder eine Betriebsvereinbarung anwenden, wird in dieser Region für diesen Wirtschaftszweig nur ausgewiesen, ob ein Tarifvertrag angewendet wird oder nicht. Die Art des Tarifvertrages wird dagegen nicht angegeben.

F11 – Wirtschaftszweig

Durch die Kombination aus regionaler Ebene und Wirtschaftszweig kann ein hohes Risiko für die Identifikation eines Betriebes ausgehen. Aus diesem Grund wird bei der Anonymisierung der Wirtschaftszweigangabe auch die Verteilung der Betriebe auf die 5 gebildeten Regionen berücksichtigt. Wie bei früheren SUFs, die zur Verdienststrukturerhebung (VSE) erstellt wurden, soll

ein Wirtschaftszweig in einer Region mindestens 50 Betriebe umfassen. Ausgangspunkt für die Anonymisierung der Wirtschaftszweige sind die 85 Abteilungen der WZ 2008 (WZ-2-Steller).

Sofern eine Abteilung eines Wirtschaftszweiges in einer Region zu schwach besetzt ist, wird diese mit einer benachbarten bzw. inhaltlich verwandten zusammengefasst. Eine Übersicht zu den Zusammenfassungen bei den Wirtschaftszweigen findet sich bei der Beschreibung des Merkmals „F11“.

Eine Ausnahme ergibt sich für den WZ-Abschnitt B „Bergbau und Gewinnung von Steinen und Erden“. Für Betriebe mit weniger als 500 Beschäftigten wurde in diesem Abschnitt ein erhöhtes Reidentifikationsrisiko festgestellt. Dieses Risiko ergibt sich daraus, dass Nordrhein-Westfalen eine eigene Region (Region 2) umfasst und damit Betriebe aus dieser Region leichter zu identifizieren sind. Eine Zusammenfassung des WZ Abschnitts B mit einem anderen ist aus inhaltlichen Gründen wenig sinnvoll. Es wurde daher für diesen Bereich, wie bereits bei den vorangegangenen SUFs zur VSE, die Regionalangabe stärker vergrößert. Hier liegen in der Folge die beiden Kategorien „neue Bundesländer (ohne Berlin)“ und „alte Bundesländer (einschließlich Berlin)“ vor.

F13 – Geburtsdatum

Es wird nur das Jahr der Geburt angegeben, der Geburtsmonat wird entfernt. Bei Beschäftigten die älter als 65 Jahre sind, wird ein Bottom-Coding, bei Beschäftigten, die jünger als 16 Jahre sind, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass das Geburtsjahr „1957 oder früher“ oder „2006 oder später“ war.

F14 – Erste Staatsangehörigkeit

Die Angaben zur Staatsgenhörigkeit werden stark vergrößert und lediglich der Kontinent des zugehörigen Staates ausgewiesen.

Ausprägungen:

0 = Deutschland

1 = Europa, ohne Deutschland

2 = Afrika

3 = Amerika

4 = Asien

5 = Australien, Ozeanien, Antarktis

9 = unbekannt/staatenlos

F15 – Eintrittsdatum

Es wird nur das Jahr des Eintrittsdatums angegeben, der Eintrittsmonat wird entfernt. Bei Beschäftigten mit einer Unternehmenszugehörigkeit von 45 und mehr Jahren wird ein Bottom-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass das Eintrittsjahr „1977 oder früher“ war.

F17 – Tätigkeitsschlüssel

Klassifikation der Berufe (KldB) 2010 (überarbeitete Fassung 2020) (Ziffer 1-5):

Aufgrund der ausgeübten Tätigkeiten können sich Rückschlüsse auf den Wirtschaftszweig des Betriebes ziehen lassen. Aus diesem Grund werden im SUF

nur die ersten beiden Stellen ausgewiesen. Die letzten drei Stellen werden durch „@@@“ ersetzt.

Bei einzelnen Schlüsseln können sich aufgrund hoher Häufigkeiten bei einzelnen Wirtschaftszweigen, trotz eines Ausweisens der ersten beiden Stellen, Rückschlüsse auf den Wirtschaftszweig ziehen lassen. Aus diesem Grund erfolgt hier eine Umkodierung zum Schlüssel 99 „ohne nähere Tätigkeitsangabe / ohne Angabe“.

Höchster beruflicher Ausbildungsabschluss (Ziffer 7)

Aufgrund geringer Fallzahlen und der erhöhten Gefahr eines Rückschlusses auf einzelne Beschäftigte eines Betriebes, die einen Dokortitel als Namenszusatz tragen, wird die Ausprägung 6 „Promotion“ mit den anderen universitären Abschlüssen bei Ausprägung 5 „Diplom/Magister/Master/Staatsexamen“ zusammengefasst.

F21 – Bruttomonatsverdienst - Gesamtbruttoentgelt

Bei einem Bruttomonatsverdienst von 62.500 Euro¹ und mehr wird ein Top-Coding durchgeführt. In der Folge wird bei einem Bruttomonatsverdienst ab 62.500 Euro nicht der genaue Verdienst ausgewiesen, sondern nur, dass der Beschäftigte 62.500 Euro oder mehr verdient.

¹ Dies entspricht einem Jahresverdienst von 750.000 Euro, der in den vorigen SUF zur VSE als Obergrenze verwendet wurde.

F22 – Sonstige Bezüge (EBV § 1 Absatz 2 Nr. 2a)

F23 – Bruttomonatsverdienst für bezahlte Überstunden

F24 – Bruttomonatsverdienst für Zuschläge

F25 – Entgeltumwandlung

Für die Darunter-Positionen „F22“ bis „F25“ eines Gesamtbruttoentgelts aus „F21“ von 62.500 Euro und mehr wird nicht der absolute Wert, sondern der prozentuale Anteil am Bruttomonatsverdienst („F21“) ausgewiesen.

F33 – Anzahl Beschäftigte im Unternehmen (URS)

Die Anzahl der Beschäftigten in dem zugehörigen Unternehmen wird zu drei Klassen („weniger als 50“, „50 bis unter 250“ sowie „250 und mehr“) zusammengefasst.

Zusätzliche Maßnahmen bei bestimmten Beschäftigten

Bei einem geringen Anteil von Beschäftigten (ca. 0,01% der Originalstichprobe) müssen zusätzliche Maßnahmen durchgeführt werden, um die faktische Anonymität der Daten zu erreichen. Bei diesen Fällen erfolgt das Top-Coding bereits ab einem Bruttomonatsverdienst („F21“) von 10.000 Euro (statt 62.500 Euro). Die mit dem Top-Coding zusammenhängende Anpassung der Merkmale „F22“ – „EF25“ erfolgt, wie oben beschrieben. Darüber hinaus werden für die Merkmale „F13“ (Geburtsdatum) und „F15“ (Eintrittsdatum) keine Werte ausgewiesen. Die Ziffern 1 bis 5 des Tätigkeitsschlüssels („F17“; Beruf nach KldB) werden auf 99 gesetzt.

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. Im vorliegenden SUF sind zudem nur 5 Regionen aus zusammengefassten Bundesländern als regionale Einheit verfügbar (siehe „F4“, „F5“).

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), Sonderzahlun-

gen). Auf Betriebsebene gibt es beispielsweise Angaben zum Sitz des Betriebes, dem Wirtschaftszweig, der Tarifbindung und der Anzahl der Beschäftigten im zugehörigen Unternehmen.

Die Einzeldaten des SUF der VE April 2022 umfassen 6,1 Mio. Beschäftigten Datensätze, die sich auf 64.803 Betriebe aus der Stichprobenziehung, der Personalstandstatistik und dem Verwaltungsdatenspeicher aufteilen. Die Stichprobenziehung basiert auf den Daten der Bundeskopie des statistischen Unternehmensregisters aus dem Berichtsjahr 2020 mit Stand August 2021 und umfasst maximal 58.000 Betriebe. Bei 10.000 Betrieben mit ausschließlich geringfügig entlohnnten Beschäftigten wurden die Angaben nicht erhoben oder aus der Personalstandstatistik bezogen, sondern aus dem Verwaltungsdatenspeicher imputiert. Die Erstellung des SUF ist im Hessischen Statistischen Landesamt (HSL) erfolgt.

2.1 Merkmale und Merkmalsbeschreibung

2.1.1 Datensatzbeschreibung /Schlüsselverzeichnis

Die Angaben für April 2022 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 2022.

F3 – BERICHTSEINHEITID (Identnummer des Betriebs (URS))

Systemfreie Betriebsnummer für die Zuordnung der Beschäftigten zu Betrieben.

F4 – Bundesland

Amtlicher Schlüssel des Bundeslandes, in dem sich der Betrieb zum Zeitpunkt der Stichprobenziehung (Herbst 2021) befindet. Entsprechend der Anonymisierungsmaßnahmen, werden hier nur die rekodierten Regionen angegeben.

Ausprägungen:

- 1 = Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, Berlin
- 2 = Nordrhein-Westfalen
- 3 = Hessen, Rheinland-Pfalz, Saarland
- 4 = Baden-Württemberg, Bayern
- 5 = Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen
- 6 = Wenn EF6 = 5 (Bergbau und Gewinnung von Steinen und Erden) dann sind Regionen 1 bis 4 zusammengefasst.

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. Entsprechend der Anonymisierungsmaßnahmen, werden hier nur die rekodierten Regionen angegeben.

Ausprägungen:

- 1 = Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, Berlin
- 2 = Nordrhein-Westfalen
- 3 = Hessen, Rheinland-Pfalz, Saarland
- 4 = Baden-Württemberg, Bayern
- 5 = Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen
- 6 = Wenn EF6 = 5 (Bergbau und Gewinnung von Steinen und Erden) dann sind Regionen 1 bis 4 zusammengefasst.

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 April 2022 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 entlohnenden Beschäftigten des Betriebs sowie die Zahl der Betriebe mit abhängig Beschäftigten laut Verwaltungsdatenspeicher der Statistischen Ämter.

Es ist zu beachten, dass **hochgerechnete Ergebnisse mit dem SUF durch die Anonymisierungsmaßnahmen nicht den hochgerechneten Ergebnissen entsprechen, die mit dem On-Site Material berechnet werden können.** Dies betrifft vor allem absolute Häufigkeiten, im geringeren Umfang aber auch andere statistische Werte wie Standardabweichungen, Mittelwerte oder Regressionskoeffizienten.

F10 – Tarifbindung

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

1 = Branchentarifvertrag, gültig in allen, mehreren oder einem Bundesland

2 = Firmentarifvertrag

3 = kein Tarifvertrag, Orientierung der Verdienste an einem Branchentarifvertrag

4 = kein Tarifvertrag, freie Verdienstvereinbarung

Bei Wirtschaftszweigen in Regionen mit weniger als 3 Betrieben:

5 = kein Tarifvertrag im Betrieb

6 = Tarifvertrag wird im Betrieb angewandt

Das Merkmal zur Tarifbindung lässt sich zwischen VSE und 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, welche Art der Tarifregelung 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 der genau für jeden einzelnen Beschäftigten gültige Tarifvertrag angegeben, sodass es zu heterogenen Angaben hinsichtlich der Tarifbindung innerhalb eines Betriebs kommen konnte.

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.

Ausprägung im Datensatz	Klassifikation WZ 2008	
	WZ-2-Steller	Bezeichnung
1	01	Landwirtschaft, Jagd und damit verbundene Tätigkeiten
2	02 – 03	Forstwirtschaft und Holzeinschlag sowie Fischerei und Aquakultur
5	05 – 09	Abschnitt B - Bergbau und Gewinnung von Steinen und Erden
10	10 – 12	Herstellung von Nahrungs- und Futtermitteln, Getränkeherstellung, Tabakverarbeitung

Ausprägung im Datensatz	Klassifikation WZ 2008	
	WZ-2-Steller	Bezeichnung
13	13 – 15	Herstellung von Textilien, Herstellung von Bekleidung, Herstellung von Leder, Lederwaren und Schuhen
16	16 – 17	Herstellung von Holz-, Flecht-, Korb- und Korkwaren (ohne Möbel), Herstellung von Papier, Pappe und Waren daraus
18	18	Herstellung von Druckerzeugnissen; Vervielfältigung von bespielten Ton-, Bild- und Datenträgern
19	19 – 21	Kokerei und Mineralölverarbeitung, Herstellung von chemischen Erzeugnissen, Herstellung von pharmazeutischen Erzeugnissen
22	22	Herstellung von Gummi- und Kunststoffwaren
23	23	Herstellung von Glas und Glaswaren, Keramik, Verarbeitung von Steinen und Erden
24	24 – 25	Metallerzeugung und -bearbeitung, Herstellung von Metallerzeugnissen
26	26	Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen
27	27	Herstellung von elektrischen Ausrüstungen
28	28	Maschinenbau
29	29 – 30	Herstellung von Kraftwagen und Kraftwagenteilen, Sonstiger Fahrzeugbau
31	31 – 32	Herstellung von Möbeln, Herstellung von sonstigen Waren
33	33	Reparatur und Installation von Maschinen und Ausrüstungen
35	D / 35 und E / 36 – 39	Energieversorgung; Wasserversorgung; Abwasser- und Abfallentsorgung
41	41	Hochbau
42	42	Tiefbau
43	43	Vorbereitende Baustellenarbeiten, Bauinstallation und sonstiges Ausbaugewerbe
45	45	Handel mit Kraftfahrzeugen; Instandhaltung und Reparatur von Kraftfahrzeugen
46	46	Großhandel (ohne Handel mit Kraftfahrzeugen)
47	47	Einzelhandel (ohne Handel mit Kraftfahrzeugen)

Ausprägung im Datensatz	Klassifikation WZ 2008	
	WZ-2-Steller	Bezeichnung
49	49 – 51	Landverkehr und Transport in Rohrfernleitungen, Schifffahrt, Luftfahrt
52	52	Lagerei sowie Erbringung von sonstigen Dienstleistungen für den Verkehr
53	53	Post-, Kurier- und Expressdienste
55	55	Beherbergung
56	56	Gastronomie
58	58	Verlagswesen
59	59-60	Herstellung, Verleih und Vertrieb von Filmen und Fernsehprogrammen; Kinos
61	61 – 62	Telekommunikation, Erbringung von Dienstleistungen der Informationstechnologie, Informationsdienstleistungen
63	63	Informationsdienstleistungen
64	64	Erbringung von Finanzdienstleistungen
65	65 – 66	Versicherungen, Rückversicherungen und Pensionskassen (ohne Sozialversicherung), Mit Finanz- und Versicherungsdienstleistungen verbundene Tätigkeiten
68	L / 68	Grundstücks- und Wohnungswesen
69	69	Rechts- und Steuerberatung, Wirtschaftsprüfung
70	70	Verwaltung und Führung von Unternehmen und Betrieben; Unternehmensberatung
71	71	Architektur- und Ingenieurbüros; technische, physikalische und chemische Untersuchung
72	72	Forschung und Entwicklung
73	73	Werbung und Marktforschung
74	74	Sonstige freiberufliche, wissenschaftliche und technische Tätigkeiten
77	77	Vermietung von beweglichen Sachen
78	78	Vermittlung und Überlassung von Arbeitskräften
79	79 – 80	Reisebüros, Reiseveranstalter und Erbringung sonstiger Reservierungsdienstleistungen, Wach- und Sicherheitsdienste sowie Detekteien
81	81	Gebäudebetreuung; Garten- und Landschaftsbau

Ausprägung im Datensatz	Klassifikation WZ 2008	
	WZ-2-Steller	Bezeichnung
82	82	Erbringung von wirtschaftlichen Dienstleistungen für Unternehmen und Privatpersonen a. n. g.
84	O / 84	Öffentliche Verwaltung, Verteidigung; Sozialversicherung
85	P / 85	Erziehung und Unterricht
86	86, 75	Gesundheitswesen, Veterinärwesen
87	87	Heime (ohne Erholungs- und Ferienheime)
88	88	Sozialwesen (ohne Heime)
90	R / 90 – 93	Kunst, Unterhaltung und Erholung
94	94	Interessenvertretungen sowie kirchliche und sonstige religiöse Vereinigungen (ohne Sozialwesen und Sport)
95	95 – 96	Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern, Erbringung von sonstigen überwiegend persönlichen Dienstleistungen

F12 – Personalnummer (gehasht)

Laufende systemfreie Nummer der Beschäftigten.

F13 – Geburtsdatum

Das Geburtsdatum ist nach dem Muster JJJJ hinterlegt. Es wurde ein Bottom- und Top-Coding durchgeführt.

1957 = 1957 oder früher

2006 = 2006 oder später

F14 – Erste Staatsangehörigkeit

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

Für Beschäftigte aus der Personalstandstatistik liegen keine Angaben vor.

Ausprägungen:

0 = Deutschland

1 = Europa, ohne Deutschland

2 = Afrika

3 = Amerika

4 = Asien

5 = Australien, Ozeanien, Antarktis

9 = unbekannt/staatenlos

F15 – Eintrittsdatum

Jahr des Eintritts in das Unternehmen im Format JJJJ. 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. Es wird ein Bottom-Coding für Beschäftigte durchgeführt, die länger als 45 Jahre im Betrieb sind.

1977 = 1977 oder früher

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.²

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

² 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 die Arbeitgeberin oder 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.

Aufgrund der ausgeübten Tätigkeiten können sich Rückschlüsse auf den Wirtschaftszweig des Betriebes ziehen lassen. Aus diesem Grund werden im SUF nur die ersten beiden Stellen ausgewiesen. Die letzten drei Stellen werden durch „@@@“ ersetzt.

Bei einzelnen Schlüsseln können sich aufgrund bestimmter Häufungen bei einzelnen Wirtschaftszweigen, trotz eines Ausweisens der ersten beiden Stellen, Rückschlüsse auf den Wirtschaftszweig ziehen lassen. Aus diesem Grund erfolgt hier eine Umkodierung zum Schlüssel 99 „ohne nähere Tätigkeitsangabe / ohne Angabe“.

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. In 2022 wurden manuell plausibilisierte Beschäftigtendatensätze jedoch nicht durch CANCEIS imputiert, sodass die Angabe 9 enthalten ist.

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, 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.

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 jede Beschäftigte und 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.

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.

Bei einem Bruttomonatsverdienst von 62.500 Euro und mehr wird ein Top-Coding durchgeführt. In der Folge wird bei einem Bruttomonatsverdienst ab 62.500 Euro nicht der genaue Verdienst ausgewiesen, sondern nur, dass der Beschäftigte 62.500 Euro oder mehr verdient.

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.

Bei einem Bruttomonatsverdienst („F21“) von 62.500 Euro und mehr wird nicht der absolute Wert, sondern der prozentuale Anteil am Bruttomonatsverdienst („F21“) ausgewiesen.

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.

Bei einem Bruttomonatsverdienst („F21“) von 62.500 Euro und mehr wird nicht der absolute Wert, sondern der prozentuale Anteil am Bruttomonatsverdienst („F21“) ausgewiesen.

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.

Bei einem Bruttomonatsverdienst („F21“) von 62.500 Euro und mehr wird nicht der absolute Wert, sondern der prozentuale Anteil am Bruttomonatsverdienst („F21“) ausgewiesen.

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).

Bei einem Bruttomonatsverdienst („F21“) von 62.500 Euro und mehr wird nicht der absolute Wert, sondern der prozentuale Anteil am Bruttomonatsverdienst („F21“) ausgewiesen.

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 km².)

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 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 lebt, 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 haben.)

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./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.)

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

Missing = Fehlende Angabe

(Beschäftigte, die aus der Personalstandstatistik stammen.)

F32 – Zustandsanzeiger

Ausprägungen:

0 = nicht geschätzt

1 = geschätzt

F33 – Anzahl Beschäftigte im Unternehmen (URS)

Klassierte Anzahl der Beschäftigten im gesamten Unternehmen zum 31.12.2021.

1 = weniger als 50

2 = 50 bis unter 250

3 = 250 und mehr

9 = k. A. (Personalstandstatistik)

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.

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

2.2 Vergleichbarkeit der Merkmale über die Zeit

Berichtsjahre der VE können miteinander verglichen werden. Da der SUF zur Verdienststrukturerhebung (VSE) (EVAS 62111) ab dem Berichtsjahr 2014 überwiegend nach demselben methodischen Vorgehen durchgeführt wurde, kann auch die VSE für Zeitvergleiche herangezogen werden. Lediglich die Aggregationen der Wirtschaftszweige unterscheiden sich geringfügig und sollten entsprechend beachtet werden.

2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

Tabelle 1: Häufigkeiten der ungewichteten Beschäftigten und Betriebe je Region (Merkmal F5) im SUF der VE April 2022

Bundesland	Anzahl Beschäftigte	Anzahl Betriebe
Region 1 Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, Berlin	1 318 792	16 706
Region 2 Nordrhein-Westfalen	1 019 103	7 541
Region 3 Hessen, Rheinland-Pfalz, Saarland	986 550	11 510
Region 4 Baden-Württemberg, Bayern	1 948 280	13 606
Region 5 Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen	844 260	15 196
Region 6 Region 1-4 im WZ-Abschnitt B	8 984	244
Insgesamt	6 125 969	64 803

2.4 Auswertbare regionale Ebene

Entsprechend der in Abschnitt 1.2 geschilderten Anonymisierungsmaßnahmen stehen als regionale Einheit fünf Regionen zur Verfügung, die die Bundesländer zusammenfassen.

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 und keiner 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 dargelegt. 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 Gastwissenschaftsarbeitsplatz 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 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

Merkmal	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
F32	Zustandsanzeiger
F33	Anzahl Beschäftigte im Unternehmen (URS)
F34	STIA
F35	Gruppennummer
F38	Fusion/Aufspaltung
F39	Konjunkturelle Kurzarbeit

Merkmal	Bezeichnung
F40	Saisonale Kurzarbeit
F41	Streik
F42	Sonstige Gründe
F43	Handwerkszugehörigkeit

