

## Leitfaden zur absoluten Anonymisierung für die Erstellung eines Campus-File aus den Einzeldaten der Gehalts- und Lohnstrukturerhebung 2001

### 1. Vorbemerkungen

Im Jahr 1987 wurde mit § 16 Abs. 6 des Bundesstatistikgesetzes<sup>1</sup> der Wissenschaft ein privilegierter Zugang zu Mikrodaten der amtlichen Statistik eingeräumt. Hiernach ist die Übermittlung von Einzeldaten an die Wissenschaft erlaubt, sofern diese nur mit unverhältnismäßig großem Aufwand an Zeit, Kosten und Arbeitskraft reidentifiziert werden können (faktische Anonymität). „Unverhältnismäßig“ bedeutet hier, dass der Aufwand einer Reidentifikation deren Nutzen übersteigt. Die Deanonymisierung von Einzelangaben in einem faktisch anonymen Datensatz kann nicht mit absoluter Sicherheit ausgeschlossen werden. Die Schutzmaßnahmen für einen, im Rahmen der Wissenschaft und Lehre frei zugänglichen Datensatz, dem so genannten Campus-File müssen weitaus höher angesetzt werden. In diesem Fall ist es nur zulässig, absolut anonymisierte Daten weiterzugeben, d.h. die eindeutige Identifikation von Fällen ist ausgeschlossen. Der vorliegende Leitfaden behandelt absolut anonymisierte Datensätze für die Wissenschaft und Lehre, generiert aus den Daten der Gehalts- und Lohnstrukturerhebung (GLS) 2001. Der Leitfaden ist in Kooperation zwischen dem Hessischen Statistischen Landesamt und dem Statistischen Bundesamt entstanden.

### 2. Basismaterial

Die Gehalts- und Lohnstrukturerhebung 2001 wurde als Stichprobe bei 27000 Betrieben mit 10 und mehr Beschäftigten durchgeführt. Dabei wurden nach Ländern differenzierte Auswahlsätze vorgegeben. In kleineren Ländern wurden höhere und in großen Ländern niedrigere Auswahlsätze angelegt. So lagen die Auswahlsätze zwischen 5,3% (Nordrhein-Westfalen) und 19,4% (Bremen). Die ausgewählten Betriebe bezogen bundesweit rund 940.000 Beschäftigte ein. Es handelt sich um ein zweistufiges Auswahlverfahren mit Betrieben in der 1. Stufe und Beschäftigten in der 2. Stufe. In den kleineren Betrieben wurden alle Beschäftigten erfasst, während in den größeren Betrieben nach einem Zufallsverfahren nur ein Teil der Beschäftigten ausgewählt wurde.

Bei der GLS werden Angaben über Arbeiter/innen und Angestellte im Individualverfahren erhoben. Damit können individuelle Merkmale der Arbeitnehmer/innen – mit dem Arbeitsplatz verbundene Merkmale einerseits und persönliche Merkmale andererseits – zum Verdienst in Beziehung gesetzt werden. Neben den Bruttoverdiensten werden auch Nettoverdienste und Sozialversicherungsbeiträge der Arbeitnehmer/innen nachgewiesen. Brutto- und Nettoverdienste sowie Sonderzahlungen beziehen sich auf das gesamte Jahr 2001, Brutto- und Nettomonatsverdienste und Arbeitszeiten auf den Berichtsmonat Oktober 2001.

---

<sup>1</sup> Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) vom 22. Januar 1987 (BGBl. I S. 462, 565), zuletzt geändert durch Artikel 2 des Gesetzes vom 9. Juni 2005 (BGBl. I S. 1534).

Eine vollständige Liste der im vorgestellten Campus-File enthaltenen Merkmale findet sich in der Datensatzbeschreibung im Anhang.

### 3. Anonymisierungsmaßnahmen

Nachfolgende Maßnahmen wurden durchgeführt:

#### **Stichprobenziehung**

Die Stichprobenziehung erfolgt in 2 Stufen:

- (1) Zunächst werden 10% der Betriebe des Originalmaterials gezogen, geschichtet nach den unten angegebenen 9 Wirtschaftskategorien, alten und neuen Bundesländern und 6 Beschäftigtengrößenklassen.
- (2) Aus den im Basismaterial enthaltenen Beschäftigten dieser ausgewählten Betriebe wird dann im zweiten Schritt eine Zufallsstichprobe gezogen, wobei der Auswahlsatz in Abhängigkeit von der Größenklasse des Betriebes zufällig ermittelt wird. Für Größenklassen 1 und 2 liegt der Faktor zwischen 0.6 und 0.9; für Größenklassen 3 und 4 zwischen 0.3 und 0.7 und für Größenklassen 5 und 6 zwischen 0.2 und 0.8. Der genaue Auswahlsatz pro Betrieb wird nicht weitergegeben. Damit soll verhindert werden, dass von der Anzahl der Beschäftigten des Betriebes in der Stichprobe Rückschlüsse auf die Anzahl der Beschäftigten des Betriebes möglich sind, was insbesondere für die größten Betriebe eine Reidentifikation ermöglichen könnte.

#### **Vergroberung der Regionalangabe**

Es werden die beiden Regionen alte Bundesländer (mit Berlin) und neue Bundesländer (ohne Berlin) ausgewiesen.

#### **Vergroberung der Wirtschaftsgruppen**

Es werden 9 Wirtschaftskategorien ausgewiesen. Diese entsprechen denjenigen im Campus-File der zweiten Europäischen Erhebung zur beruflichen Weiterbildung (CVTS2) ausgewiesenen. Damit wird die Verwendung mehrerer wirtschaftsstatistischer Campus-Files und die Kombination der Ergebnisse von Analysen erleichtert.

#### **Anzahl der Beschäftigten eines Unternehmens**

Es werden nur Beschäftigtengrößenklassen ausgewiesen, wobei die letzte Gruppe Unternehmen mit 500 und mehr Beschäftigten umfasst. Da es in den neuen Bundesländern weniger große Unternehmen gibt, werden hier zusätzlich die beiden höchsten Klassen zusammengefasst, so dass hier die letzte Gruppe aus den Unternehmen mit 250 und mehr Beschäftigten besteht.

#### **Anzahl der Beschäftigten eines Betriebes**

Es wird ausgewiesen, welcher originale Anteil der Beschäftigten des Unternehmens im Betrieb arbeitet.

### Arbeiter und Angestellte des Betriebes nach Geschlecht

Die originalen Anteile der jeweiligen Gruppen werden ausgewiesen.

### Einfluss der öffentlichen Hand auf das Unternehmen

Die Tatsache, dass Betriebe, die unter einem Einfluss der öffentlichen Hand stehen, sich auf einige wenige Wirtschaftszweige konzentrieren, erleichtert eine mögliche Reidentifikation. Die entsprechenden Risiken liegen unter dem Kriterium der faktischen Anonymität unterhalb der kritischen Schwelle. Unter der Auflage, dass die Angaben in einem Campus-File aber absolut anonym sein müssen, lässt sich dieses Merkmal nicht beibehalten. Es wird daher aus dem Datensatz entfernt.

### Tarifvertragsschlüssel

Vom Tarifvertragsschlüssel wird nur die zweite Stelle ausgewiesen, die angibt, ob es sich um einen Kollektiv- oder Firmentarifvertrag oder eine Betriebsvereinbarung handelt. Dabei werden zusätzlich die Ausprägungen *Firmentarifvertrag* und *Betriebsvereinbarung* zusammengefasst.

### Ausgeübte Tätigkeit

Die ausgeübten Tätigkeiten werden zu 20 Berufsgruppen zusammengefasst (s. Datensatzbeschreibung im Anhang).

### Brutto-Monats- und Jahresverdienst

Für hohe Verdienste ab 7000 Euro brutto im Monat und ab 84000 Euro brutto im Jahr wird ein Topcoding durchgeführt; d. h. für diese Verdienste werden nur 7000 Euro und mehr bzw. 84000 Euro und mehr ausgewiesen.

### Sonstige Verdienstmerkmale / Sozialabgaben

Für die sonstigen Verdienstangaben wird der Anteil am Bruttomonatsverdienst (bzw. bei Angaben, die sich auf das Jahr beziehen, am Bruttojahresverdienst) in Prozent ausgewiesen. Analog wird bei den Sozialabgaben verfahren.

### Geburtsjahr

Für Beschäftigte, die 17 Jahre und jünger sind, wird *1984 oder später* ausgewiesen; für Beschäftigte, die 60 Jahre und älter sind, wird *1941 oder früher* ausgewiesen.

### Kinderfreibeträge

Fälle mit 10 bzw. 20 werden in 1 bzw. 2 umcodiert, da dies offensichtliche Eingabefehler sind. Die übrigen Einzelfälle mit 7 und mehr Kinderfreibeträgen werden aus dem Datensatz entfernt. Bei den verbleibenden Fällen wird ab 4 Freibeträgen *4 und mehr* ausgewiesen.

## Hochrechnungsfaktoren

Jeder Betrieb und jeder Beschäftigte einer Größenklasse einer Wirtschaftskategorie erhält pro Region einen einheitlichen Hochrechnungsfaktor. Der Hochrechnungsfaktor für die Betriebe ergibt sich als Quotient aus der Anzahl der Betriebe der Größenklasse der Wirtschaftskategorie in der Region in der Grundgesamtheit und der entsprechenden Anzahl in der Stichprobe. Die Berechnung des Hochrechnungsfaktors für die Beschäftigten erfolgt analog unter Verwendung der entsprechenden Beschäftigtenanzahlen.

4. Mitzeichnung Statistisches Bundesamt  
Bereits erfolgt
5. Mitzeichnung FDZ der Länder
6. z.d.A. IB-FDZ

## Anhang : Merkmalsliste des Campus-Files

<i>Merkmal</i>	<i>Inhalt</i>	<i>Bemerkung</i>
REGION	Region	1 = alte Bundesländer (mit Berlin) 2 = neue Bundesländer (ohne Berlin)
BETR_ID	Nummer für Betrieb	systemfrei
BES_ID	Nummer für Beschäftigten	Fortlaufend je Betrieb
EF2	Beschäftigtengruppe	1 = Arbeiter/innen 2 = Angestellte
WZGRUPPE	Wirtschaftsgruppe, in der Haupttätigkeitsbereich des Betriebes liegt	1 = Ernährungsgewerbe und Tabakverarbeitung; Textil- und Bekleidungs-gewerbe; Ledergewerbe 2 = Verarbeitendes Gewerbe ohne Ernährungsgewerbe, Tabakverarbeitung, Textil- und Bekleidungs-gewerbe, Ledergewerbe 3 = Energie- und Wasserversorgung

		<p>4 = Baugewerbe</p> <p>5 = Handelsvermittlung und Großhandel; Einzelhandel; Reparatur von Gebrauchsgütern; Kraftfahrzeughandel; Instandhaltung und Reparatur von Kraftfahrzeugen; Tankstellen</p> <p>6 = Gastgewerbe</p> <p>7 = Verkehr; Nachrichtenübermittlung</p> <p>8 = Kreditgewerbe; Versicherungsgewerbe; Mit dem Kredit- und Versicherungsgewerbe verbundene Tätigkeiten</p> <p>9 = Grundstücks- und Wohnungswesen; Vermietung beweglicher Sachen ohne Bedienungspersonal; Erbringung von sonstigen Dienstleistungen überwiegend für Unternehmen</p>
EF5	Monats- oder Stundenlohn	<p>Nur für Arbeiter/innen</p> <p>1 = Monatslohn</p> <p>2 = Stundenlohn</p>
EF11a	Leistungsgruppe	<p>0 = qualifizierte/r Facharbeiter/in</p> <p>1 = Facharbeiter/in</p> <p>2 = angelernte/r Arbeiter/in</p> <p>3 = ungelernete/r Arbeiter/in</p> <p>4 = Leitende/r Angestellte/r</p> <p>5 = Angestellte/r mit besonderen Erfahrungen und besonderen Leistungen</p> <p>6 = Angestellte/r mit besonderen Fachkenntnissen</p> <p>7 = Angestellte/r mit nach allgemeiner Anweisung selbständiger Tätigkeit</p> <p>8 = Angestellte/r ohne Entscheidungsbezugnis</p> <p>9 = Angestellte/r mit einfachen Tätigkeiten</p>
EF12	Geschlecht	<p>1 = männlich</p> <p>2 = weiblich</p>
EF13U2	Geburtsjahr	<p>1984 = 1984 und später</p> <p>1941 = 1941 und früher</p>
EF14U2	Jahr des Eintritts in das Unternehmen	
EF16U1	Lohnsteuerklasse	Nur Abschnitte C-F, G und J der WZ93
EF16U2	Anzahl der Kinderfreibeträge	<p>10 -&gt; 1, 20 -&gt; 2;</p> <p>weitere Fälle mit 7 und mehr entfernt</p> <p>4 = 4 und mehr</p> <p>Nur Abschnitte C-F, G und J der WZ93</p>
BERUF	Berufsgruppe	<p>1 = Pflanzenbauer, Tierzüchter, Fischereiberufe (KdB 01 – 06)</p> <p>2 = Bergleute, Mineralgewinner; Herstellung</p>

		<p>und Verarbeitung von Stein, Baustoff, Keramik, Glas, Kunststoff, Papier, Holz und Metall (KdB 07 – 24)</p> <p>3 = Schlosser, Mechaniker, Montierer, Elektriker, Maschinisten (KdB 25 – 32 + 54)</p> <p>4 = Textil- und Ernährungsberufe (KdB 33 – 43)</p> <p>5 = Bau- und Raumausstattungsberufe (KdB 44 – 50)</p> <p>6 = Warenprüfer, Versandfertigmacher; Lager- und Transportarbeiter (KdB 52 + 74)</p> <p>7 = Ingenieure / Naturwissenschaftler (KdB 60 + 61)</p> <p>8 = Techniker / Technische Sonderfachkräfte (KdB 62 + 63)</p> <p>9 = Warenkaufleute (KdB 68)</p> <p>10 = Dienstleistungskaufleute (KdB 69 + 70)</p> <p>11 = Berufe des Verkehrs und Nachrichtenwesens (KdB 71 + 73)</p> <p>12 = Unternehmer, Abgeordnete u. ä. (KdB 75 + 76)</p> <p>13 = Rechnungskaufleute; Datenverarbeitungsfachleute (KdB 77)</p> <p>14 = Bürokräfte (KdB 78)</p> <p>15 = Ordnungs-/Sicherheitsberufe (KdB 79-81)</p> <p>16 = Publizisten, Künstler u. ä. (KdB 82 + 83)</p> <p>17 = Gesundheitsdienstberufe; Sozial- und Erziehungsberufe (KdB 84 – 87)</p> <p>18 = Sonstige geistes- und naturwissenschaftliche Berufe (KdB 88)</p> <p>19 = Allgemeine Dienstleistungsberufe (KdB 90 – 93)</p> <p>20 = Sonstige Berufe (KdB 53, 89, 97 – 99)</p>
EF17U2	Stellung im Beruf	<p>0 = Auszubildende</p> <p>1 = Arbeiter/in, nicht als Facharbeiter/in tätig</p> <p>2 = Facharbeiter/in</p> <p>3 = Meister/in, Polier/in</p> <p>4 = Angestellte/r</p> <p>8 = Teilzeit mit weniger als 18 Wochenstunden</p> <p>9 = Teilzeit mit mindestens 18 Wochenstunden</p>
EF17U3	Ausbildung	1 = Volks-/Hauptschule, mittlere Reife ohne abgeschlossene Berufsausbildung

		<p>2 = Volks-/Hauptschule, mittlere Reife mit abgeschlossener Berufsausbildung</p> <p>3 = Abitur ohne abgeschlossene Berufsausbildung</p> <p>4 = Abitur mit abgeschlossener Berufsausbildung</p> <p>5 = Abschluss einer Fachhochschule</p> <p>6 = Hochschulabschluss</p> <p>7 = unbekannt, Angabe nicht möglich</p>
EF18	Art des Arbeitsvertrages	<p>1 = unbefristet</p> <p>2 = befristet ohne Auszubildende und Praktikant/innen</p> <p>3 = Altersteilzeit</p> <p>4 = Auszubildende und Praktikant/innen</p>
EF19A	Lohnform	<p>Nur Arbeiter/innen</p> <p>1 = Zeitlohn</p> <p>2 = Prämienlohn</p> <p>3 = Akkordlohn</p> <p>4 = Prämien- und Akkordlohn</p> <p>5 = Mischlohn</p>
EF19B	Technische oder kaufmännische Tätigkeit	<p>Nur Angestellte</p> <p>1 = Kaufmännische Tätigkeit</p> <p>2 = Technische Tätigkeit</p> <p>3 = Meister/in</p>
EF20	Arbeitsvertraglich vereinbarte wöchentliche Arbeitszeit	
EF21	Bezahlte Normalarbeitsstunden	
EF22	Bezahlte Mehrarbeitsstunden	
EF25	Bruttoverdienst insgesamt im Berichtsmonat in Euro	7000 = 7000 Euro und mehr
EF26	Verdienst aus Mehrarbeitszeit in Prozent von EF25	
EF27	Verdienstminderung im Berichtsmonat	<p>0 = nein</p> <p>1 = ja</p>
EF28	In Monatsverdienst enthaltene Zulagen für Schichtarbeit in Prozent von EF25	
EF29	In Monatsverdienst enthaltene Zulagen für Samstags-/Sonntags-/Feiertagsarbeit in Prozent von EF25	
EF30	In Monatsverdienst enthaltene Zulagen für Nachtarbeit in Prozent von EF25	
EF31	Lohnsteuer incl. Solidaritätszuschlag ohne Kirchensteuer in Prozent von EF25	Nur Abschnitte C-F, G und J der WZ93
EF32	Beiträge der Arbeitnehmer/innen zur Renten- und Arbeitslosenversicherung in Prozent von EF25	Nur Abschnitte C-F, G und J der WZ93
EF33	Beiträge der Arbeitnehmer/innen zur Kranken-	Nur Abschnitte C-F, G und J der WZ93

	und Pflegeversicherung in Prozent von EF25	
EF34	Bruttojahresverdienst Insgesamt in Euro	84000 = 84000 Euro und mehr
EF35	Sonderzahlungen für das ganze Jahr in Prozent von EF34	
EF36	Nettojahresverdienst in Prozent von EF34	Nur Abschnitte C-F, G und J der WZ93
EF37	Verdienstminderung im Jahr 2001	0 = nein 1 = ja
EF38	Urlaubsanspruch auf Basis einer 5-Tage-Woche	
EF39	Arbeitszeit in Stunden bei festem Monatslohn	Nur Arbeiter/innen
EF40	Stundenzahl nach der Arbeitszeitregelung vom 1. bis 28. Oktober 2001	Nur Arbeiter/innen
EF41	Stundenzahl nach der Arbeitszeitregelung vom 29. bis 31. Oktober 2001	Nur Arbeiter/innen
EF42	Umrechnungsfaktor für Wochenstunden	Nur Arbeiter/innen
EF49	Austrittsmonat bei Ausscheiden in 2001	
EF51	Nettomonatsverdienst in Prozent von EF25	Nur Abschnitte C-F, G und J der WZ93
TARIFART	Art des Tarifvertrages	0 = kein Tarifvertrag 1 = Kollektivtarifvertrag 2 = Firmentarifvertrag oder Betriebsvereinbarung
B_EF13	Beschäftigtengrößenklasse des Unternehmens	1 = unter 20 2 = 20 – unter 50 3 = 50 – unter 100 4 = 100 – unter 250 5 = 250 – unter 500 (alte Bundesländer) 6 = 500 und mehr (alte Bundesländer) 56 = 250 und mehr (neue Bundesländer)
B_EF36	Anteil Beschäftigte des Betriebes an Beschäftigten des Unternehmens in Prozent	Hierzu wird die originale Beschäftigtenangabe des Unternehmens verwendet
B_EF14	Anteil der männlichen Arbeiter an den Beschäftigten des Betriebes in Prozent	
B_EF15	Anteil der weiblichen Arbeiter an den Beschäftigten des Betriebes in Prozent	
B_EF16	Anteil der männlichen Angestellten an den Beschäftigten des Betriebes in Prozent	
B_EF17	Anteil der weiblichen Angestellten an den Beschäftigten des Betriebes in Prozent	
B_EF30	Hochrechnungsfaktor 1. Stufe	Neuberechnung s. o.
B_EF31	Hochrechnungsfaktor 2. Stufe	Neuberechnung s. o.