

# Statistische Geheimhaltung

Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit



## Impressum

<b>Produktlinie/Reihe:</b>	Grundlagen: Methodenbericht
<b>Titel:</b>	Statistische Geheimhaltung: Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit
<b>Veröffentlichung:</b>	Oktober 2018
<b>Herausgeberin:</b>	Bundesagentur für Arbeit Statistik/Arbeitsmarktberichterstattung
<b>Rückfragen an:</b>	Zentraler Statistik-Service Regensburger Straße 104 90478 Nürnberg
<b>E-Mail:</b>	<a href="mailto:Zentraler-Statistik-Service@arbeitsagentur.de">Zentraler-Statistik-Service@arbeitsagentur.de</a>
<b>Telefon:</b>	0911 179-3632
<b>Fax:</b>	0911 179-1131

### Weiterführende statistische Informationen:

<b>Internet:</b>	<a href="http://statistik.arbeitsagentur.de">http://statistik.arbeitsagentur.de</a>
<b>Zitierhinweis:</b>	Statistik der Bundesagentur für Arbeit, Grundlagen: Definitionen – Statistische Geheimhaltung: Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit, Nürnberg, Oktober 2018

**Nutzungsbedingungen:** © Statistik der Bundesagentur für Arbeit

Sie können Informationen speichern, (auch auszugsweise) mit Quellenangabe weitergeben, vervielfältigen und verbreiten. Die Inhalte dürfen nicht verändert oder verfälscht werden. Eigene Berechnungen sind erlaubt, jedoch als solche kenntlich zu machen.

Im Falle einer Zugänglichmachung im Internet soll dies in Form einer Verlinkung auf die Homepage der Statistik der Bundesagentur für Arbeit erfolgen.

Die Nutzung der Inhalte für gewerbliche Zwecke, ausgenommen Presse, Rundfunk und Fernsehen und wissenschaftliche Publikationen, bedarf der Genehmigung durch die Statistik der Bundesagentur für Arbeit.

## Inhaltsverzeichnis

1	Rechtliche Grundlagen der statistischen Geheimhaltung .....	4
2	Regeln der statistischen Geheimhaltung.....	5
2.1	Primäre Geheimhaltung (Mindestfallzahlregel u. Dominanzregel).....	5
2.2	Sekundäre Geheimhaltung (Ausschluss der Rückrechenbarkeit) .....	7
3	Maßnahmen zur Reduzierung des Informationsverlustes.....	10

## 1 Rechtliche Grundlagen der statistischen Geheimhaltung

Bei der Erstellung einer Statistik geht es um die Verarbeitung personenbezogener Daten, auch wenn das Interesse der Statistik dabei nicht der Person als Individuum, sondern nur als „Merkmalsträger“ innerhalb der statistischen Grundgesamtheit gilt. Unter Verarbeitung versteht man alle Phasen der statistischen Arbeit von der Erhebung bis zur Veröffentlichung. Für die Statistik der BA gilt als oberster Grundsatz die statistische Geheimhaltung gemäß § 16 des Bundesstatistikgesetzes (BStatG). Der statistische Produktionsprozess, also die Verarbeitung von personenbezogenen Daten, unterliegt den Vorschriften der Verordnung 2016/679 des europäischen Parlaments und des Rates vom 27. April 2016 (EU-DSGVO) (EU-DSVGO). Bei Zweifelsfragen wird das Bundesdatenschutzgesetz (BDSG) ergänzend herangezogen. Darüber hinaus orientiert sich die Statistik BA an der Verordnung EG Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken.

Unter Statistikgeheimnis versteht man die Verschwiegenheitspflicht, welcher die Mitarbeiterinnen und Mitarbeiter von statistischen Stellen - und somit auch die der Statistik der BA - unterliegen. Dem Statistikgeheimnis wurde im Volkszählungsurteil von 1983 (BVerfGE 65,1) eine herausragende Bedeutung zugewiesen, es gilt als das zentrale Fundament der Beziehungen von statistischen Behörden zu den Auskunft gebenden Stellen. Die Regelungen sind strenger als die des allgemeinen Datenschutzes und des Sozialgeheimnisses. Mit Artikel 89 EU-DSGVO wurde die statistische Arbeit privilegiert, soweit Garantien für die Rechte und Freiheiten der betroffenen Person eingehalten werden.

Alle mit der Durchführung von amtlichen Statistiken betrauten Personen sind verpflichtet, statistische Einzelangaben geheim zu halten und nur für statistische Zwecke zu verwenden. Statistische Geheimhaltung bedeutet, dass für statistische Zwecke übermittelte oder aus eigenen administrativen oder sonstigen Quellen erhaltene Angaben über einzelne statistische Einheiten, d.h. Personen, geschützt werden müssen. Die Verwendung der vorhandenen Einzelangaben für nichtstatistische Zwecke oder ihre unrechtmäßige Offenlegung sind untersagt. Als vertraulich zu schützen sind alle Daten, die eine direkte oder indirekte Identifizierung statistischer Einheiten möglich machen und dadurch Einzelinformationen offenlegen. Eine Person ist nicht identifizierbar, wenn ein Dritter nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft die statistische Einheit deanonymisieren kann.

Daher muss durch geeignete Maßnahmen sichergestellt werden, dass eine Verwendung der zusammengeführten Daten für andere als statistische Zwecke ausgeschlossen ist. Insbesondere sind Einzeldaten zu Personen genauso wie zu Betrieben geheim zu halten. Diese Maßnahmen gelten für die von der Bundesagentur für Arbeit zu verantwortende Arbeitsmarkt- und Grundsicherungsstatistik in gleichem Maße wie für andere statistische Stellen.

Aus den gesetzlichen Vorgaben und den darin beschriebenen Anforderungen leitet die Statistik der BA folgende Regelungen ab:

## 2 Regeln der statistischen Geheimhaltung

### 2.1 Primäre Geheimhaltung (Mindestfallzahlregel u. Dominanzregel)

Die wichtigste Vorgabe der primären Geheimhaltung lautet, dass zu jedem Tabellenwert eine Mindestzahl an Fällen (z.B. Personen, Maßnahmen, Betriebe) vorhanden sein muss. Die Zahl der Mindestfälle richtet sich nach dem Risiko bzw. der Wahrscheinlichkeit des Offenlegens von Einzelangaben. Als Standard für die Mindestfallzahlregel gilt eine Mindestfallzahl von drei. Bei erhöhtem Risiko aufgrund eines hohen Detaillierungsgrades der Daten kann die Mindestfallzahl auch mit größer als 3 festgesetzt werden.

#### Beispiel 1: Mindestfallzahlregel

	Merkmal A	Merkmal B	Merkmal C
Region X	2	123	111
Region Y	88	1	60

Die beiden grau markierten Felder sind mit den Fallzahlen 1 und 2 belegt und sind dementsprechend zu anonymisieren.

Ergebnis:

	Merkmal A	Merkmal B	Merkmal C
Region X	*	123	111
Region Y	88	*	60

Neben der Mindestfallzahlregel von drei für jede Zellbesetzung muss stets geprüft werden, ob sich die dargestellten Informationen, die in einer hierarchischen Beziehung zueinanderstehen, auf nur sehr wenige Merkmalsträger beziehen, z.B. Betriebe in der in der Beschäftigungsstatistik. Am Beispiel der Beschäftigungsnachweise, die sich auf mindestens drei Betriebe beziehen müssen, wird dies im Folgenden erläutert.

*Beispiel 2: Mindestfallzahlregel (Schutz von Betrieben)*

	SvB	Betriebe
Gemeinde A	2	1
Gemeinde B	15	6
Gemeinde C	25	2
Gemeinde D	42	4

Die Zahl der Sozialversicherungspflichtig Beschäftigten (SvB) in der Gemeinde A muss anonymisiert werden, da die Mindestfallzahlregel dies unmittelbar erfordert. Die Zahl der Sozialversicherungspflichtig Beschäftigten in der Gemeinde C ist ebenfalls zu anonymisieren, da sie sich auf nur 2 Betriebe bezieht und somit auch hier die Mindestfallzahlregel anzuwenden ist.

Ergebnis:

	SvB
Gemeinde A	*
Gemeinde B	15
Gemeinde C	*
Gemeinde D	42

Darüber hinaus besteht ein erhöhtes Risiko der Offenlegung von Einzelangaben auch dann, wenn ein Tabellenwert durch einen einzelnen dahinterstehenden Fall dominiert wird. Vor diesem Hintergrund gilt zur Gewährleistung der statistischen Geheimhaltung zusätzlich die sogenannte „Dominanzregel“. Diese Regel wird im Folgenden am Beispiel der der Beschäftigungsstatistik erläutert:

Selbst wenn hinter der Beschäftigtenzahl eines Wirtschaftszweiges drei oder mehr Betriebe stehen, kann einer dieser Betriebe einen so hohen Beschäftigtenanteil auf sich vereinen, dass die Beschäftigtenzahl praktisch eine Einzelangabe über den Branchenführer darstellt. Würde dann der Wert der Beschäftigten für diese Branche in dieser Region ausgewiesen, so könnte leicht auf die Beschäftigtenzahl des Dominanzbetriebs rückgeschlossen werden.

Es sind folgende Regelungen zu beachten:

- Bei 3 bis 9 Betrieben, die hinter einer Beschäftigtenzahl eines Wirtschaftszweiges stehen, darf keiner der Betriebe 50 oder mehr Prozent der Beschäftigten auf sich vereinen.
- Bei 10 oder mehr Betrieben dürfen auf keine Betriebseinheit 85 oder mehr Prozent der Beschäftigten des Wirtschaftszweiges entfallen.

Die Identifikation von Dominanzbetrieben ist über eine Differenzierung von Betrieben und Beschäftigten nach Betriebsgrößenklassen möglich.

Beispiel 3: Dominanzregel

WZ	Betriebe	SvB	dav.					
			Betriebsgrößen- klasse 1 - 9		Betriebsgrößen- klasse 10 - 99		Betriebsgrößen- klasse 100 - 249	
			Betriebe	SvB	Betriebe	SvB	Betriebe	SvB
Abschnitt A	5	165	2	15	2	50	<b>1</b>	<b>100</b>
Abschnitt B	9	15	9	15	-	-	-	-
Abschnitt C	11	100	10	12	<b>1</b>	<b>88</b>	-	-

Der Abschnitt A umfasst 5 Betriebe mit 165 SvB. In der Betriebsgrößenklasse 100-249 werden in einem Betrieb 100 sozialversicherungspflichtig Beschäftigte (SvB) ausgewiesen. Dieser Betrieb stellt damit mehr als 50 Prozent der SvB dieses Abschnitts. Es ist deshalb eine Anonymisierung vorzunehmen. Im Abschnitt C sind es 11 Betriebe mit 100 SvB. Eine Betriebsstätte beschäftigt 88 SvB, dies sind 88 Prozent der Beschäftigten dieses Abschnitts. Es ist eine Anonymisierung vorzunehmen.

Ergebnis:

WZ	SvB
Abschnitt A	*
Abschnitt B	15
Abschnitt C	*

2.2 Sekundäre Geheimhaltung (Ausschluss der Rückrechenbarkeit)

Neben der primären Geheimhaltung müssen statistische Darstellungen auch einer Prüfung der sekundären Geheimhaltung unterworfen werden. Das ist immer dann notwendig, wenn Einzelfelder durch Einbeziehung von Spalten- oder Zeilensummen oder andere Vergleichswerte miteinander in einem rechnerischen Zusammenhang stehen, die es ermöglichen, den Inhalt bereits gesperrter Tabellenfelder über einfache Rechenoperationen zu ermitteln.

Für jedes Feld, dessen Inhalt aufgrund primärer Geheimhaltung gesperrt wurde, muss geprüft werden, ob sich

- aus einer übergeordneten Einheit,
- in Verbindung mit einer oder mehreren gleichrangigen Einheiten, die die übergeordnete Einheit gemeinsam mit dem gesperrten Wert vollständig abbilden,

der gesperrte Wert wieder errechnen lässt. Gleiches gilt für durch obige Herangehensweise gesperrte Tabellenwerte, d.h., es kann notwendig sein, mehrere aufeinander aufbauende Prüfgänge zu durchlaufen und weitere Werte zu sperren.

*Beispiel 4: Ausschluss der Rückrechenbarkeit*

Schritt 1:

		Insgesamt	dav.		
			Merkmal A	Merkmal B	Merkmal C
Insgesamt		22	6	12	4
dav.	Region X	14	2	8	4
	Region Y	8	4	4	-

Dunkelgrau: primäre Geheimhaltung

Hellgrau: sekundäre Geheimhaltung (aufgrund Summe der Regionen pro Merkmal)

Zwischenergebnis:

		Insgesamt	dav.		
			Merkmal A	Merkmal B	Merkmal C
Insgesamt		22	6	12	4
dav.	Region X	14	*	8	4
	Region Y	8	*	4	-



**Schritt 2:**

		Insgesamt	dav.		
			Merkmal A	Merkmal B	Merkmal C
Insgesamt		22	6	12	4
dav.	Region X	14	*	8	4
	Region Y	8	*	4	-

Der Wert für Merkmal B, Region Y muss gesperrt werden, da er in Verbindung mit der Zeilensumme ermöglicht, den Wert für Merkmal A, Region Y zu errechnen und so auf Merkmal A, Region X geschlossen werden kann.

**Schritt 3:**

		Insgesamt	dav.		
			Merkmal A	Merkmal B	Merkmal C
Insgesamt		22	6	12	4
dav.	Region X	14	*	8	4
	Region Y	8	*	*	-

Der Wert für Merkmal B, Region X muss gesperrt werden, da er in Verbindung mit der Spaltensumme ermöglicht, den Wert für Merkmal B, Region Y zu errechnen.

**Ergebnis:**

		Insgesamt	dav.		
			Merkmal A	Merkmal B	Merkmal C
Insgesamt		22	6	12	4
dav.	Region X	14	*	*	4
	Region Y	8	*	*	-

### **3 Maßnahmen zur Reduzierung des Informationsverlustes**

Um möglichst viele relevante Informationen zu erhalten, sollte man im Vorfeld von statistischen Anfragen überlegen, wie detailliert eine bestimmte Auswertung sein muss. Eine Möglichkeit besteht darin, den Fokus auf bestimmte wesentliche Kategorien anstelle der Abbildung aller einzelnen Ausprägungen eines Merkmals zu legen (Beispiel: Muss es bei berufsfachlichen Darstellungen immer der komplette Zwei-, und Dreisteller nach tief gegliederten regionalen Strukturen sein oder gibt es aggregierte Darstellungen, die vergleichbar informativ sind?). Ein Verzicht auf nicht interessierende Kategorien oder Summen kann ebenfalls die vollständige Darstellung der benötigten Daten ermöglichen. Insbesondere ist es häufig hilfreich, nicht alle Merkmale miteinander zu kreuzen, sondern Kombinationen eher sparsam auszuwählen oder sich auf die "Ränder" von Tabellen zu konzentrieren.

Natürlich steht der Statistik-Service bei Anfragen im Zusammenhang mit der Berücksichtigung datenschutzrechtlicher Belange gerne beratend zur Verfügung.

## Statistik-Infoseite

Im Internet stehen statistische Informationen unterteilt nachfolgenden Themenbereichen zur Verfügung:

[Arbeitsmarkt im Überblick](#)  
[Arbeitslose, Unterbeschäftigung und Stellen](#)  
[Ausbildungsstellenmarkt](#)  
[Beschäftigung](#)  
[Förderungen](#)  
[Grundsicherung für Arbeitsuchende \(SGB II\)](#)  
[Leistungen SGB III](#)  
[Migration](#)  
[Langzeitarbeitslosigkeit](#)  
[Frauen und Männer](#)  
[Statistik nach Berufen](#)  
[Statistik nach Wirtschaftszweigen](#)  
[Zeitreihen](#)  
[Daten zu den Eingliederungsbilanzen](#)  
[Amtliche Nachrichten der BA](#)  
[Kreisdaten](#)

Die [Methodischen Hinweise der Statistik](#) bieten ergänzende Informationen.

Das [Glossar](#) enthält Erläuterungen zu allen statistisch relevanten Begriffen, die in den verschiedenen Produkten der Statistik der BA Verwendung finden.

Abkürzungen und Zeichen, die in den Produkten der Statistik der Bundesagentur für Arbeit vorkommen, werden im [Abkürzungsverzeichnis](#) bzw. der [Zeichenerklärung](#) der Statistik der BA erläutert.