

Statistische Geheimhaltung

Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit



Impressum

Produktlinie/Reihe:	Grundlagen: Definitionen
Titel:	Statistische Geheimhaltung: Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit
Veröffentlichung:	Oktober 2021
Herausgeberin:	Bundesagentur für Arbeit Statistik/Arbeitsmarktberichterstattung
Rückfragen an:	Zentraler Statistik-Service Regensburger Straße 104 90478 Nürnberg
E-Mail:	Zentraler-Statistik-Service@arbeitsagentur.de
Telefon:	0911 179-3632
Fax:	0911 179-1131

Weiterführende statistische Informationen:

Internet:	http://statistik.arbeitsagentur.de
Zitierhinweis:	Statistik der Bundesagentur für Arbeit, Grundlagen: Definitionen – Statistische Geheimhaltung: Rechtliche Grundlagen und fachliche Regelungen der Statistik der Bundesagentur für Arbeit, Nürnberg, Oktober 2021

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	Statistische Geheimhaltung in der Statistik der Bundesagentur für Arbeit	5
2.1	Mindestfallzahl.....	5
2.2	Dominanzregel	6
3	Verfahren zur Sicherstellung der statistischen Geheimhaltung	7
3.1	Zellsperrungsverfahren	7
3.1.1	Primäre Geheimhaltung	7
3.1.2	Sekundäre Geheimhaltung	8
3.1.3	Maßnahmen zur Reduzierung des Informationsverlustes	10
3.2	Rundungsverfahren	11
3.2.1	Summen.....	13
3.2.2	Arithmetisches Mittel.....	14
3.2.3	Kennzahlen und Quoten	14
3.2.4	Absolute Veränderungen	14
3.2.5	Relative Veränderungen und Anteilswerte	14
3.3	Vergleich der Verfahren	16
3.3.1	Sicherheit	16
3.3.2	Informationsverlust.....	17
3.3.3	Konsistenz	17
3.3.4	(technische) Umsetzung	17

1 Rechtliche Grundlagen der statistischen Geheimhaltung

Bei der Erstellung einer Statistik geht es in der Regel um die Erhebung und Verarbeitung personenbezogener Daten, auch wenn das Interesse der Statistik dabei nicht dem Bürger als Individuum, sondern nur als "Merkmalsträger" innerhalb der statistischen Grundgesamtheit gilt. Für die Statistik der Bundesagentur für Arbeit (BA) gilt der Grundsatz der Statistischen Geheimhaltung gemäß § 281 Abs. 3 SGB III in Verbindung mit § 16 des Bundesstatistikgesetzes (BStatG). Darüber hinaus orientiert sich die Statistik der BA an den Regeln der Verordnung EG Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken und an der DSGVO. Unter statistischer Geheimhaltung in diesem Sinne versteht sich grundlegend die Verschwiegenheitspflicht, welcher die Mitarbeiter von statistischen Stellen – zu welcher die Statistik der BA als einzelstaatliche Stelle für die Entwicklung, Erstellung und Verbreitung europäischer Statistiken i. S. d. Art. 5 Abs. 2 der o. g. Verordnung zählt – 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. Diese Regelungen sind strenger als die des allgemeinen Datenschutzes und des Sozialdatenschutzes. Für die Statistiken der Bundesagentur gelten die Vorschriften der Geheimhaltung nach § 16 des BStatG entsprechend. Das Statistikgeheimnis ist durch technische und organisatorische Maßnahmen der Trennung zwischen statistischen und nichtstatistischen Aufgaben zu gewährleisten.

Alle mit der Durchführung von amtlichen Statistiken betrauten Personen sind verpflichtet, statistische Einzelangaben geheim zu halten und grundsätzlich nur für statistische Zwecke zu verwenden. Dies beinhaltet, dass direkt für statistische Zwecke oder indirekt aus administrativen oder sonstigen Quellen eingeholte vertrauliche Angaben über einzelne statistische Einheiten geschützt werden müssen, wobei die Verwendung der eingeholten Angaben für nichtstatistische Zwecke und ihre unrechtmäßige Offenlegung untersagt sind. Als vertraulich zu schützen sind alle Daten, die eine direkte oder indirekte Identifizierung statistischer Einheiten möglich machen und dadurch Einzelinformationen offenlegen. Bei der Entscheidung, ob eine statistische Einheit identifizierbar ist, sind alle Mittel zu berücksichtigen, die nach vernünftigem Ermessen von einem Dritten angewendet werden können, um die statistische Einheit zu identifizieren.

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 und Betrieben geheim zu halten. Diese Maßnahmen gelten für die von der BA zu verantwortende Arbeitsmarkt- und Grundsicherungsstatistik in gleichem Maße wie für andere statistische Stellen.

Die Erstellung von Arbeitsmarktstatistiken und die Übermittlung statistischer Daten durch die BA ist im Wesentlichen im SGB III im Kapitel sieben, erster Abschnitt durch die §§ 280 ff. im Rahmen der Aufgabenstellung „Statistiken, Arbeitsmarkt- und Berufsforschung“ und im Kapitel sieben des SGB II in den §§ 53 ff. „Statistik und Forschung“ geregelt. In den Bestimmungen zur Übermittlung von Daten wird dabei im § 281 Abs. 3 SGB III und § 53 Abs. 7 SGB II explizit auf die Geheimhaltungsnormen des BStatG verwiesen. Als Spezialfall des Datenschutzes werden die Vorschriften der Statistischen Geheimhaltung auch in den Datenschutzbestimmungen der BA behandelt.

2 Statistische Geheimhaltung in der Statistik der Bundesagentur für Arbeit

2.1 Mindestfallzahl

Aus den rechtlichen Vorgaben zum Datenschutz bzw. zur statistischen Geheimhaltung ergibt sich, dass im Zuge der statistischen Berichterstattung kein Rückschluss auf Einzelangaben möglich sein darf. Das bedeutet, dass solche Werte nicht zu veröffentlichen sind, die sich aus wenigen Einzelangaben zusammensetzen bzw. die nur wenige Einzelangaben repräsentieren, wenn damit eine exakte bzw. näherungsweise Offenlegung enthaltener Einzelangaben erreicht werden kann.

Um die Offenlegung zu verhindern, wird bestimmt, 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 3. Bei erhöhtem Risiko aufgrund eines hohen Detaillierungsgrades der Daten kann die Mindestfallzahl auch mit größer als 3 festgesetzt werden.

Beispiel 1: Mindestfallzahl

	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.

Geheimhaltungsbedürftig im Sinne der Mindestfallzahlregel ist nicht nur, wenn der Tabellenwert selbst kleiner als 3 ist, sondern auch, wenn hinter diesem Wert (der durchaus deutlich größer sein kann) weniger als 3 Fälle stehen. Dieser Sachverhalt tritt immer dann auf, wenn die Daten in einer hierarchischen Beziehung zueinander stehen.

In der Statistik der BA sind insbesondere die folgenden Beziehungen:

Betriebe	<->	Beschäftigte
Betriebe	<->	Personen in Kurzarbeit / Anzeigen Kurzarbeit
Bedarfsgemeinschaften	<->	Personen in Bedarfsgemeinschaften und alle weiteren Merkmale
Anträge zur Vorfinanzierung von Insolvenzgeld (Zustimmung, Auszahlung)	<->	In Anträgen genannte Personen und Auszahlungshöhen

Beispiel 2: Schutz von Daten in hierarchischen Beziehungen

	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.

2.2 Dominanzregel

Darüber hinaus besteht ein erhöhtes Risiko der Offenlegung von Einzelangaben bei Daten, die in einer hierarchischen Beziehung zueinander stehen, 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“¹. Ein Tabellenwert ist geheim zu halten, wenn der Anteil des größten Einzelwerts mehr als k % beträgt.

Diese Regel wird im Folgenden am Beispiel 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.

¹ Die Statistik der BA wendet die (1,k)-Dominanzregel an.

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	1	100
Abschnitt B	9	15	9	15	-	-	-	-
Abschnitt C	11	100	10	12	1	88	-	-

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.

3 Verfahren zur Sicherstellung der statistischen Geheimhaltung

Die Statistik setzt zwei Verfahren zur Sicherstellung der statistischen Geheimhaltung ein, das Zellsperungsverfahren und ein Rundungsverfahren.

3.1 Zellsperungsverfahren

Die Statistik der BA wendet in ihren Veröffentlichungen mit statistischen Tabellen das Zellsperverfahren zur Sicherstellung der statistischen Geheimhaltung an. Das bedeutet, ein zu schützender Wert wird in der Tabelle entfernt und der Wert wird durch ein „*“ ersetzt.

3.1.1 Primäre Geheimhaltung

Zunächst werden alle Werte, die unterhalb der Mindestfallzahl liegen und/oder der Dominanzregel unterworfen werden müssen, in einer Tabelle identifiziert, und die Zelle wird gesperrt.

Beispiel 4: Zellsperungsverfahren – Primäre Geheimhaltung

	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	-

Der dunkelgrau hinterlegte Wert unterliegt der primären Geheimhaltung und wird daher durch ein „*“ ersetzt.

Zwischenergebnis:

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

3.1.2 Sekundäre Geheimhaltung

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 5: Zellsperungsverfahren – Sekundäre Geheimhaltung***Schritt 1:**

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

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

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	*	*	-

3.1.3 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 stehen die [Statistik-Services](#) bei Anfragen im Zusammenhang mit der Berücksichtigung datenschutzrechtlicher Belange gerne beratend zur Verfügung.

3.2 Rundungsverfahren

In ihren webbasierten Anwendungen verwendet die Statistik der BA ein Rundungsverfahren zur Sicherstellung der statistischen Geheimhaltung im Hinblick auf die Mindestfallzahl. Der Schutz von Daten in hierarchischer Beziehung und der Schutz von Dominanzfällen ist mit dem Rundungsverfahren nicht möglich.

Hinweis:

Um in den webbasierten Anwendungen der Statistik der BA dennoch Daten in hierarchischer Beziehung und geschützte Dominanzfällen zu veröffentlichen, wird in diesen Fällen eine Kombination aus Rundungsverfahren und Zellsperungsverfahren angewendet.

Bei der deterministischen Rundung wird jeder Wert einer Tabelle auf das jeweils nächstliegende Vielfache einer festgelegten Basis gerundet.

Rundungsbasis für die Ergebnisdarstellungen ist 10. Die Abweichung von den Originalwerten ist bei dieser Basis höher als bei kleineren Basen. Dieses Argument tritt jedoch hinter den folgenden beiden zurück.

Sicherheit vor Aufdeckung: Für kleinere Rundungsbasen – hier am Beispiel von 3 und 5 – lassen sich Konstellationen finden, die eine Aufdeckung der wahren Werte mithilfe der Randsummen zulassen.

Runden auf 3: $0 + 0 = 3$ kann nur aus den Originaldaten $1 + 1 = 2$ hervorgegangen sein.

Runden auf 5: Hinter $0 + 0 + 0 + 0 = 10$ kann nur die Konstellation $2 + 2 + 2 + 2 = 8$ stehen.

Runden auf 10: Dieses Problem kann nicht auftreten, falls zwischen einem auf null abgerundeten Wert und einem nicht vorhandenen Wert nicht unterschieden wird.

Transparenz: Anhand der dargestellten Werte ist vermutlich für viele Nutzer offensichtlich, dass es sich um gerundete Werte handelt. Selbst wenn Nutzer die entsprechenden methodischen Hinweise nicht beachten oder in der Weiterverwendung der Daten nicht an Auswertungen anbringen, bleibt das Rundungsverfahren leicht erkennbar.

Jeder Wert wird also auf das jeweils nächstliegende Vielfache von 10 gerundet. Bis zur Endziffer 4 wird abgerundet, ab 5 aufgerundet (kaufmännisches Runden). Die größte mögliche Abweichung eines dargestellten Werts vom Originalwert beträgt somit 5.

Beispiel 6: Runden auf ein Vielfaches von 10

Originalwert	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gerundeter Wert	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	20	20

<i>Abweichung (absolut)</i>	0	-1	-2	-3	-4	5	4	3	2	1	0	-1	-2	-3	-4	5	4
-----------------------------	---	----	----	----	----	---	---	---	---	---	---	----	----	----	----	---	---

Mit dem dargestellten Rundungsverfahren wird die Geheimhaltung sichergestellt, da nicht auf Einzelangaben geschlossen werden kann. Eine sekundäre Geheimhaltung ist beim Rundungsverfahren nicht erforderlich.

3.2.1 Summen

Summen werden auf Basis der Originalwerte berechnet und anschließend gerundet. Das führt möglicherweise zu einer Abweichung zwischen den dargestellten Summanden und der zugehörigen Summe. Dafür liegt die angezeigte Summe grundsätzlich näher am Originalwert. Zusätzlich wird jeder Wert immer gleich angezeigt, egal ob er in der aktuellen Darstellung als Einzelwert oder als Summe erscheint.

Beispiel 7: Rundungsverfahren - Summen

Die Summen sind fett dargestellt.

	Anzahl Original	Summe gerundet	Anzahl gerundet
A	74		70
B	13		10
C	11		10
D	3		0
Summe (A-D)	101	100	90

	Anzahl Original	Summe gerundet	Anzahl gerundet
A-D	101		100
E-H	134		130
I-M	113		110
N-Z	232		230
Summe (A-Z)	580	580	570

3.2.2 Arithmetisches Mittel

Bei der Durchschnittsbildung wird die gerundete Summe der Originalwerte verwendet, um größere Abweichungen des angezeigten Durchschnitts von dem auf Basis der Summe der Originalwerte errechneten Durchschnitt zu vermeiden (vgl. Tabelle 3).

Beispiel 8: Rundungsverfahren – Arithmetisches Mittel

	Anzahl Original	Summe gerundet	Anzahl gerundet
A	74		70
B	13		10
C	11		10
D	3		0
Summe	101	100	90
Durchschnitt	25,3	25,0	22,5

3.2.3 Kennzahlen und Quoten

Kennzahlen und Quoten (z. B. Arbeitslosen- /Unterbeschäftigungsquoten) werden grundsätzlich auf Basis der Originalwerte berechnet.

3.2.4 Absolute Veränderungen

Absolute Veränderungen werden auf Basis der gerundeten Werte errechnet.

Beispiel 9: Rundungsverfahren – Absolute Veränderungen

	Aktueller Monat	Vormonat	Veränderung absolut
Original	254	250	4
Gerundet	250	250	0

3.2.5 Relative Veränderungen und Anteilswerte

Relative Veränderungen und Anteilswerte werden auf Basis der Originalwerte berechnet und anschließend auf ganze Prozentzahlen gerundet (vgl. Beispiele 10 und 11). Dieses Verfahren wird der Berechnung von Anteilen und Veränderungen auf Basis gerundeter Werte vorgezogen.

Die Vorteile einer Berechnung auf Basis gerundeter Werte bestünden vor allem darin, dass sie vom Nutzer anhand der dargestellten Absolutwerte nachvollzogen werden könnten. Zudem lägen die dargestellten Werte bei großen Fallzahlen oft näher am Originalwert als bei einer Berechnung auf Basis der Originalwerte und anschließender Rundung. Das Verfahren hat jedoch auch Nachteile; so wichen die dargestellten Werte bei kleineren Fallzahlen (unterhalb 1.000) oft stärker ab als bei einer Berechnung mit Originalwerten. Die ausgewiesene Kommastelle suggerierte eine Genauigkeit, die de facto (zumindest bei kleineren Fallzahlen) nicht gegeben ist. Schließlich ließen sich die dargestellten Anteile meist nicht zu 100 % aufsummieren. Anteile von über 105 % oder unter 95 % in der Summe könnten Nutzer irritieren.

Die Vorteile einer Berechnung auf Basis der Originalwerte liegen vor allem darin, dass die dargestellten Werte bei kleineren Fallzahlen oft deutlich näher am Originalwert liegen als bei der Berechnung mit gerundeten Werten. Anteilswerte lassen sich zudem meist zu 100 % (Abweichungen von höchstens einem Prozentpunkt) aufsummieren. Ein Nachteil liegt darin, dass die Berechnung vom Nutzer nicht nachvollzogen werden kann, weil er die Originalwerte nicht kennt und das Ergebnis als „falsch“ fehlinterpretiert. Diese Gefahr ist jedoch auch bei der Berechnung von Summen auf Basis der Originalwerte gegeben, und das einheitliche Vorgehen bei der Summenbildung und bei der Berechnung von Anteilen und Veränderungen ist aus Nutzersicht ein Vorteil. Nachteilig ist zudem, dass absolute und relative Veränderungen auf Basis unterschiedlicher Ausgangswerte berechnet werden sowie dass die dargestellten Werte bei großen Fallzahlen (über 1.000) mitunter stärker von den Originalwerten abweichen als bei einer Berechnung mit gerundeten Werten (ohne anschließende Rundung des Ergebnisses). Wir nehmen jedoch an, dass für Nutzer durchgängig geringe Abweichungen in ähnlicher Größenordnung akzeptabler sind als hohe Abweichungen bei kleinen Fallzahlen und zunehmende Präzision bei hohen Fallzahlen. Dies kommt Nutzern von Daten auf regionaler Ebene, die tendenziell kleinere Werte betrachten, entgegen und mindert den Nutzen für Betrachter eher großer Zahlen nicht deutlich.

Beispiel 10: Rundungsverfahren – Berechnung von relativen Veränderungen und Anteilswerten

Relative Veränderung:

	Aktueller Monat	Vormonat	Veränderung	
			absolut	relativ in %
Original	254	250	4	1,6
Gerundet	250	250	0	2,0

Anteil:

	Gesamt	Unter 25	25 bis unter 50	50 und älter	Anteile in %			Summe Anteile
					Unter 25	25 bis unter 50	50 und älter	
Original	255	8	246	1	3,1	96,5	0,4	100
Gerundet	260	10	250	0	3,0	96,0	0,0	99

Um die Aufdeckung geschützter Werte zu verhindern, ist bei der Berechnung von relativen Veränderungen und Anteilen auf Basis der Originalwerte und anschließender Rundung des Ergebnisses auf ganze Prozentzahlen eine Mindestfallzahl für den Nenner festzulegen. Für Veröffentlichungen im Rahmen der neuen webbasierten Anwendungen wird diese auf 250 festgelegt und bietet damit Sicherheit vor Aufdeckung auch in sehr seltenen Wertekonstellationen.

Beispiel 11: Rundungsverfahren - Darstellung von Veränderungen und Anteilen bei Unterschreitung der Mindestfallzahl 250

Relative Veränderung

	Aktueller Monat	Vormonat	Veränderung	
			absolut	relativ in %
Original	3	2	1	50,0
Gerundet	0	0	0	(50,0) x

Anteil:

	Gesamt	U25	25 bis U50	50 und älter	Anteile in %		
					U25	25 bis U50	50 und älter
Original	94	7	86	1	7,4	91,5	1,1
Gerundet	90	10	90	0	(7,0) x	(91,0) x	(1,0) x

3.3 Vergleich der Verfahren

3.3.1 Sicherheit

Beide Geheimhaltungsverfahren, welche in der Statistik der BA angewendet werden, stellen zuverlässig die statistische Geheimhaltung sicher. Allerdings ist das Rundungsverfahren für Daten, die in einer hierarchischen Beziehung zueinanderstehen, ungeeignet, denn ein Rückschluss auf z.B. die Zahl der Beschäftigten eines Betriebes ist weiterhin möglich.

3.3.2 Informationsverlust

Sowohl beim Zellsperungsverfahren als auch beim Rundungsverfahren ist von höheren Informationsverlusten bei sehr kleinen Fallzahlen auszugehen. Allerdings bleiben auch bei überwiegend kleinen Fallzahlen die Größenverhältnisse innerhalb der Tabelle beim Rundungsverfahren erkennbar, weil kein Wert mit eventuell höherer Fallzahl zum Ausschluss von Rückrechnungen gesperrt werden muss. Beim Zellsperungsverfahren hingegen werden durch die sekundäre Geheimhaltung auch größere Werte entfernt, dafür werden ansonsten die nicht gerundeten Werte veröffentlicht.

Beispiel 12: Vergleich der Verfahren – Informationsverlust

	Originaldaten	Zellsperungsverfahren		Rundungsverfahren
		Variante 1	Variante 2	
Insgesamt	100	100	*	100
Frauen	2	*	*	0
Männer	98	*	98	100

3.3.3 Konsistenz

Die Ergebnisse beim Rundungsverfahren sind auswertungsübergreifend konsistent, weil jeder Originalwert immer auf den gleichen Wert gerundet wird. Gleichzeitig ist die auswertungsübergreifende Geheimhaltung sichergestellt, weil auch eine Auswertung in anderer Struktur – wie bei der Zellsperung häufig möglich – keine Deanonymisierung erlaubt. Beim Zellsperungsverfahren hingegen sind unterschiedliche Ergebnisse möglich, da entweder ein anderer Summand oder die Summe entfernt werden kann.

3.3.4 (technische) Umsetzung

Für das Zellsperverfahren zur Sicherstellung der statistischen Geheimhaltung verwendet die Statistik der BA eine Arbeitshilfe, welche die primäre und sekundäre Geheimhaltung in statischen Tabellen umsetzt. Für die meisten Ergebnisdarstellungen im Rahmen der webbasierten Anwendungen ist das Zellsperungsverfahren jedoch nicht geeignet. Aufgrund der Komplexität insbesondere des Ausschlusses von Möglichkeiten zur Rückrechnung stehen keine Algorithmen für eine automatisierte Umsetzung des Zellsperverfahrens zur Verfügung. Selbst wenn diese verfügbar wären, würden sie die Performanz des Systems enorm einschränken, weil für jede vom Nutzer individuell erstellte Tabelle umfangreiche Berechnungen erforderlich wären. Das würde zu nicht akzeptablen Reaktionszeiten für die Nutzer führen. Zudem könnten Nutzer des geplanten Self-Service jene Werte, die in einzelnen Ergebnistabellen gesperrt sind, einfach mittels weiterer Auswertungen offenlegen. Das Rundungsverfahren kann bei den webbasierten Anwendungen ohne Einschränkung der Performanz des Systems umgesetzt werden.

Statistik-Infoseite

Im Internet stehen statistische Informationen unterteilt nach folgenden Themenbereichen zur Verfügung:

Fachstatistiken:

[Arbeitsuche, Arbeitslosigkeit und Unterbeschäftigung](#)
[Ausbildungsmarkt](#)
[Beschäftigung](#)
[Einnahmen/Ausgaben](#)
[Förderung und berufliche Rehabilitation](#)
[Gemeldete Arbeitsstellen](#)
[Grundsicherung für Arbeitsuchende \(SGB II\)](#)
[Leistungen SGB III](#)

Themen im Fokus:

[Berufe](#)
[Bildung](#)
[Corona](#)
[Demografie](#)
[Eingliederungsbilanzen](#)
[Entgelt](#)
[Fachkräftebedarf](#)
[Familien und Kinder](#)
[Frauen und Männer](#)
[Langzeitarbeitslosigkeit](#)
[Menschen mit Behinderungen](#)
[Migration](#)
[Regionale Mobilität](#)
[Wirtschaftszweige](#)
[Zeitarbeit](#)

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

Die [Qualitätsberichte](#) der Statistik erläutern die Entstehung und Aussagekraft der jeweiligen Fachstatistik.

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.