

Das Rundungsverfahren zur Sicherstellung der Geheimhaltung in der Berufsbildungsstatistik

Nach § 16 BStatG sind Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik gemacht werden, geheim zu halten. Zur Sicherstellung der Geheimhaltung in der Berufsbildungsstatistik wird ein konventionelles Rundungsverfahren angewendet, sodass die veröffentlichten Ergebnisse keine Rückschlüsse auf eine einzelne Person zulassen.

Alle Ergebnisse werden zunächst ohne Rundung aufbereitet (Echtwerte). Anschließend wird jede Zahl für sich auf ein Vielfaches von 3 auf- bzw. abgerundet. Dabei wird nicht unterschieden, ob für das jeweilige Ergebnis Geheimhaltungsbedarf besteht oder nicht. Durch die Rundung werden die Daten nur geringfügig verzerrt. Je ausgewiesener Datenzelle beträgt die Abweichung vom Echtwert maximal 1 (*Tabelle 1*). Da sich hinter allen gerundeten Werten (mit Ausnahme der 0) jeweils drei verschiedene Echtwerte verbergen können und zugleich eine ausgewiesene 0 immer eine echte 0 oder eine echte 1 sein kann, ist der Rückschluss auf einen Einzelfall nicht möglich.

Tabelle 1

Echtwert	0	1	2	3	4	5	6	7	8	9	10	...	98	99	100	...
Gerundeter Wert	0	3			6			9			...	99			...	

Bei der Darstellung differenzierender Merkmale, z.B. Neuabschlüsse nach allgemeinbildendem Schulabschluss, kann das Verfahren dazu führen, dass die Summe der gerundeten Werte nicht der gerundeten Zahl der Neuabschlüsse insgesamt entspricht. Die Abweichung der Summe der Werte in differenzierter Darstellung vom Echtwert entspricht maximal der Anzahl der Merkmalsausprägungen.

Die Beispieltabelle zeigt die Unterschiede zwischen einer Ergebnistabelle mit Echtwerten und der Darstellung der gleichen Tabellen mit gerundeten Werten (*Tabelle 2*).

Tabelle 2

Jahr	Auszubildende am 31.12.							
	Insgesamt	Industrie und Handel	Handwerk	Landwirtschaft	Öffentlicher Dienst	Freie Berufe	Hauswirtschaft	Seeschifffahrt
Echtwerte								
Männer								
2004	937 064	504 721	377 943	30 652	15 755	6 570	957	466
2005	936 435	513 163	367 769	31 738	15 883	6 349	919	614
2006	950 217	527 748	366 565	32 613	15 692	5 963	887	749
Frauen								
2004	627 000	333 193	111 228	9 746	28 265	132 141	12 406	21
2005	617 002	335 054	109 414	9 574	27 483	124 071	11 380	26
2006	620 398	345 056	110 050	9 412	27 280	117 679	10 891	30
Insgesamt								
2004	1 564 064	837 914	489 171	40 398	44 020	138 711	13 363	487
2005	1 553 437	848 217	477 183	41 312	43 366	130 420	12 299	640
2006	1 570 615	872 804	476 615	42 025	42 972	123 642	11 778	779
Gerundete Werte								
Männer								
2004	937 065	504 720	377 943	30 651	15 756	6 570	957	465
2005	936 435	513 162	367 770	31 737	15 882	6 348	918	615
2006	950 217	527 748	366 564	32 613	15 693	5 964	888	750
Frauen								
2004	627 000	333 192	111 228	9 747	28 266	132 141	12 405	21
2005	617 001	335 055	109 413	9 573	27 483	124 071	11 379	27
2006	620 397	345 057	110 049	9 411	27 279	117 678	10 890	30
Insgesamt								
2004	1 564 065	837 915	489 171	40 398	44 019	138 711	13 362	486
2005	1 553 436	848 217	477 183	41 313	43 365	130 419	12 300	639
2006	1 570 614	872 805	476 616	42 024	42 972	123 642	11 778	780

Die Einzelwerte unterscheiden sich nach Rundung vom jeweiligen Echtwert lediglich um maximal 1. Betrachtet man jedoch die Summe der Einzelwerte in 2006, ergibt dies 1 570 617. Die ausgewiesene (gerundete) Gesamtzahl der Auszubildenden beträgt jedoch 1 570 614. Die berechnete Summe der gerundeten Werte unterscheidet sich also von der ausgewiesenen Gesamtzahl in diesem Fall um 3. Die Abweichung vom Echtwert beträgt 2. Eine Abweichung von maximal 7 (bei 7 Ausbildungsberreichen) vom Echtwert ist in diesem Beispiel möglich.

Bei sehr kleinen Fallzahlen kann die Gesamtsumme relativ betrachtet stark von der Summe der Einzelwerte abweichen. So können bei 5 Merkmalsausprägungen die Echtwerte jeweils 1 betragen. Nach der Rundung werden die Einzelwerte mit 0 ausgewiesen, die gerundete Summe mit 6.

Ebenfalls können bei sehr kleinen Fallzahlen durch das Rundungsverfahren, relativ betrachtet, größere Verzerrungen auftreten. Diese Fälle werden im Folgenden aufgeführt:

1) Beim Ausweisen von sehr kleinen Fallzahlen (z.B. bei der Aufgliederung nach Berufen), kann die prozentuale Abweichung nach Rundung relativ groß sein. Beträgt der Echtwert beispielsweise 2, so wird nach der Rundung die Zahl 3 ausgewiesen. Die ausgewiesene Zahl weicht hier vom Echtwert um 50 % ab.

2) Bei der Anteilsberechnung setzen sich diese Abweichungen fort. Nehmen beispielsweise 5 Auszubildende eines Berufs an der Abschlussprüfung teil, von denen 4 Auszubildende die Prüfung bestehen, entspricht dies einem Anteil von 80 %. Ausgewiesen werden hingegen gerundet 6 Prüfungsteilnehmer, von denen 3 Teilnehmer bestehen. Der Anteil bestandener Prüfungen auf Basis der ausgewiesenen gerundeten Werte wäre nur 50 %.

3) Auch Veränderungsdaten können von diesen Abweichungen betroffen sein. So können in einem Jahr 5 Auszubildende in einem Beruf gemeldet sein, im Vorjahr waren es 4. Dies entspricht einer Veränderungsrate von + 25 %. Nach Rundung werden im aktuellen Jahr 6 Auszubildende ausgewiesen und im Vorjahr 3. Die Veränderungsrate der gerundeten Werte beträgt dann + 100 %.

Bei der Beurteilung dieser Verzerrungen ist allerdings zu bedenken, dass die Interpretation der Ergebnisse bei sehr kleinen Fallzahlen ohnehin problematisch ist (z.B. sind Erfolgsquoten von 100 % in einem Beruf mit kleinen Fallzahlen nicht aussagekräftig und unterliegen zudem starken jährlichen Schwankungen). Die Berechnung von Veränderungsdaten, Anteilen oder Quoten erfolgt daher grundsätzlich mit Echtwerten.

Das Rundungsverfahren bietet einige Vorteile gegenüber anderen Verfahren zur Geheimhaltung. Das Verfahren ist einfach, leicht nachzuvollziehen und verzerrt die Daten nur geringfügig. Im Vergleich zur Geheimhaltung durch Zellsperren ist zudem eine sehr detaillierte Darstellung auch bei Sonderaufbereitungen möglich. Durch das Rundungsverfahren werden Einschränkungen bei der Auswertungs- bzw. Gliederungstiefe zur Sicherstellung der Geheimhaltung vermieden.