

Geheimhaltung der Zensusdaten

Sophia Degen, 08.12.2022
Jahrestagung der AG Nord-West im VDSt 2022

Grundlagen der Geheimhaltung

Warum verwenden wir ein Geheimhaltungsverfahren? Welche Geheimhaltungsverfahren werden im Zensus verwendet?



Gesetzliche Grundlage

- **Volkszählungsurteil** des Bundesverfassungsgerichts 1983: Recht auf informationelle Selbstbestimmung
- **Bundesstatistikgesetz, §16:** „**Einzelangaben** über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik gemacht werden, sind von den **Amtsträgern und Amtsträgerinnen** und für den öffentlichen Dienst besonders Verpflichteten, die mit der Durchführung von Bundesstatistiken betraut sind, **geheim zu halten** [...]“



Geheimhaltung im Zensus 2022

- **Keine Geheimhaltung** bei amtlichen Bevölkerungszahlen für administrative Gebietseinheiten (Kreise, Gemeinden, etc.)
- Automatische Geheimhaltung durch **Hochrechnung und Rundung auf das Vielfache von 10** bei Daten, die ausschließlich aus Personenerhebung kommen (z.B. zu Bildung, Erwerbstätigkeit)
- Geheimhaltung durch **Cell-Key-Methode (CKM)** bei allen anderen Daten zu Personen, Gebäuden, Wohnungen, Haushalten, Familien

Die **Cell-Key-Methode** (CKM) im Zensus 2022

Wie funktioniert die Cell-Key-Methode? Was sind die Besonderheiten der Cell-Key-Methode?



Funktionsweise der Cell-Key-Methode (CKM)

- Idee: **Überlagerung** jeder Zelle innerhalb der Ergebnistabelle (Tabelleninnenwerte und Randwerte) mit einer **Zahl auf zufälliger Basis**

Altersgruppen	0-18	18-65	65+	Gesamt
Originalwerte	2	16	11	29
Zufallszahl	+2	-1	0	-1
Ergebniswerte	4	15	11	28



Herausforderungen der Cell-Key-Methode (CKM)

- **Tabellen nicht additiv:** Randwerte ergeben sich nicht aus Innenwerten

		Altersgruppe				
Geschlecht		0-18	18-65	65+	Total	%
	M	2	7	5	12	43
	W	0	8	7	16	57
	Total	4	15	11	28	100
	%	(14)	54	39	100	

$5+7=12$



Herausforderungen der Cell-Key-Methode (CKM)

- **Ungenauigkeiten bei Verhältniszahlen** möglich
 - Einklammern von zu großen Abweichungen vom Originalwert
 - Anwendung einer Formel, durch die keine Werte >100% möglich sind

		Altersgruppe				
Geschlecht		0-18	18-65	65+	Total	%
	M	2	7	5	12	43
	W	0	8	7	16	57
	Total	4	15	11	28	100
	%	(14)	54	39	100	

↑
Einklammern ungenauer Verhältniszahlen



Stärken der Cell-Key-Methode (CKM)

- **Hohe Sicherheit:** durch zufällige Überlagerungswerte ist das Aufdeckungsrisiko sehr gering
- **Hohe Genauigkeit:** kleine Abweichungen vom Originalwert, auch bei Randwerten
- **Tabellenübergreifende Konsistenz:** gleiche Merkmalskombinationen ergeben immer gleiche Häufigkeiten
- **Flexible Tabellenerstellung** durch Nutzende möglich

Anwendung der Cell-Key-Methode (CKM) in den kommunalen Statistikstellen

Wie kann die CKM in der eigenen Statistikstelle auf die Zensusdaten angewendet werden?



Anwendung der Cell-Key-Methode (CKM) in den kommunalen Statistikstellen

- Wie genau die CKM in der Praxis auf die Zensusdaten angewendet werden soll, muss noch festgelegt werden, aber:
 - Detaillierte **Anweisungen** zur Anwendung der Geheimhaltung werden **mit den Zensusdaten mitgeschickt**
 - Geheimhaltung wird „**per Knopfdruck**“ möglich sein



Quellenangaben

- Bundesstatistikgesetz, §16 (1987). https://www.gesetze-im-internet.de/bstatg_1987/BJNR004620987.html
- Bundesverfassungsgericht. (1983). *Urteil vom 15. Dezember 1983, 1BvR 209/83.*
https://www.bundesverfassungsgericht.de/SharedDocs/Downloads/DE/1983/12/rs19831215_1bvr020983.html
- Enderle, T., & Vollmar, M. (2019). Geheimhaltung in der Hochschulstatistik. *Wista*, 6, 87-98.
- Giessing, S., Schmidt, K., Höninger, J., Höhne, J., & Rohde, J. (2021). *Statistische Geheimhaltung in Tabellen*. Statistisches Bundesamt.



Quellenangaben

- Kleber, B., Enderle, T., Giessing, S., Tent, R., & Rohde, J. (2019). *Fachkonzept Geheimhaltung V1.0*. Statistisches Bundesamt.
- Radek, S., & Paulus, O. (2021). *Fachkonzept zu Internes Auswertungssystem (IAWS) Version 1.2*. Statistisches Bundesamt.
- Rohde, J. (2018). *Informationsveranstaltung zur Statistischen Geheimhaltung: Projekt Zensus 2021*. IT.NRW.
- Statistische Ämter des Bundes und der Länder. Geheimhaltung mit Cell-Key-Methode: Schutz vor Einzelangaben in der amtlichen Statistik.
<https://www.statistikportal.de/de/cell-key-methode>



Quellenangaben

- Statistisches Bundesamt. (2017). *Sitzungsunterlage für die Referentenbesprechung „Zensus“ am 11. und 12. September 2017: TOP 6.4 Grundsatzentscheidung Geheimhaltung*. Referentenbesprechung „Zensus“ am 11. und 12. September, Wiesbaden, Deutschland.
- Statistisches Bundesamt. Die Geheimhaltung mit der Cell-Key-Methode. <https://www.zensus2022.de/DE/Zensusdatenbank/Geheimhaltung.html>

Danke für Ihre Aufmerksamkeit!

Sophia Degen, 08.12.2022
Jahrestagung der AG Nord-West im VDSt 2022

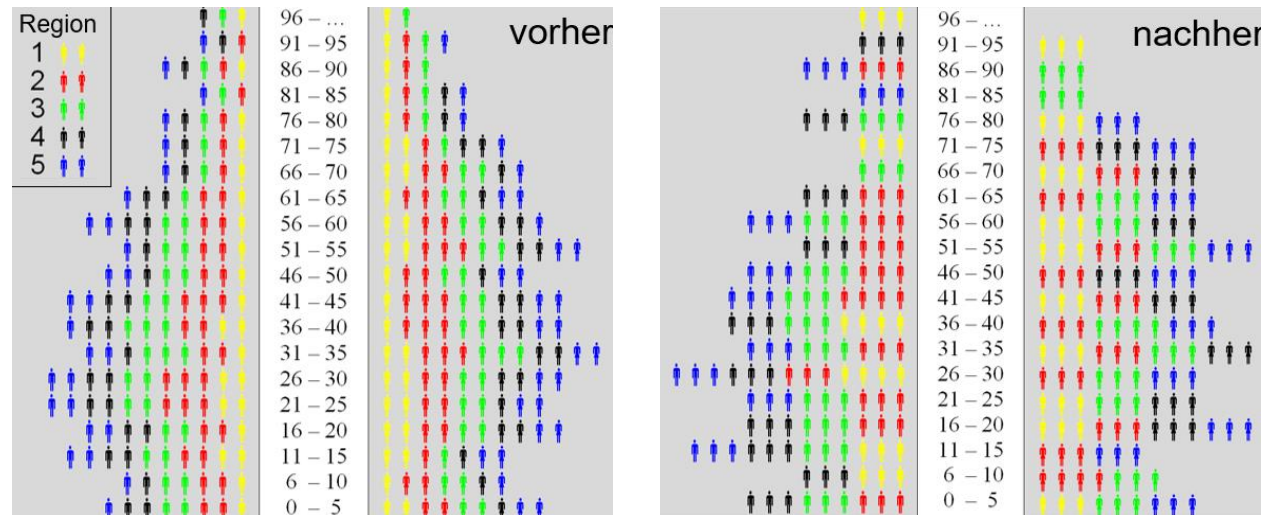
Exkurs: **Vergleich** der CKM (Zensus 2022) mit dem SAFE- Verfahren (Zensus 2011)

Wie funktioniert das SAFE-Verfahren? Was sind die Unterschiede zur Cell-Key-Methode (CKM)?



Funktionsweise des SAFE-Verfahrens

- Idee: Veränderung der Ausgangsdaten, sodass **immer drei statistische Einheiten** (Personen, Haushalte, Gebäude, etc.) die **exakt gleichen Merkmalsausprägungen** haben
- Geheimhaltungsverfahren im Zensus 2011, aber CKM wird seit kurzem rückwirkend auf die Daten des Zensus 2011 angewendet





Vergleich zwischen Cell-Key-Methode (CKM) und SAFE-Verfahren

	CKM	SAFE
Sicherheit	+	+
Tabellenübergreifende Konsistenz	+	+
Genauigkeit der Ergebnisse	+	-
Additivität	-	+
Flexible Tabellenerstellung	+	-

Exkurs: Funktionsweise der Cell-Key-Methode (CKM) im Detail

Erklärung der Cell-Key-Methode (CKM) in 4 Schritten



Funktionsweise der Cell-Key-Methode (CKM)

1. Zuordnung einer **Zufallszahl zwischen 0 und 1 (record key)** zu jeder statistischen Einheit (Person, Haushalt, Gebäude, etc.)

ID	Geschlecht	Altersgruppe	Record Key
1	W	B	0,76
2	M	C	0,09
3	M	C	0,81
4	W	A	0,32



Funktionsweise der Cell-Key-Methode (CKM)

2. Zusammenzählen aller Zufallszahlen, die zum Tabellenwert beitragen und Wegfall der Stellen vor dem Komma

ID	Geschlecht	Altersgruppe	Record Key
1	W	B	0,76
2	M	C	0,09
3	M	C	0,81
4	W	A	0,32

		Altersgruppe			
Geschlecht		A	B	C	Total
	M			?	
	W				?
	Total				



Funktionsweise der Cell-Key-Methode (CKM)

2. Zusammenzählen aller Zufallszahlen, die zum Tabellenwert beitragen und Wegfall der Stellen vor dem Komma

ID	Geschlecht	Altersgruppe	Record Key
1	W	B	0,76
2	M	C	0,09
3	M	C	0,81
4	W	A	0,32

		Altersgruppe			
Geschlecht		A	B	C	Total
	M			?	
	W				?
	Total				



Funktionsweise der Cell-Key-Methode (CKM)

2. Zusammenzählen aller Zufallszahlen, die zum Tabellenwert beitragen und Wegfall der Stellen vor dem Komma

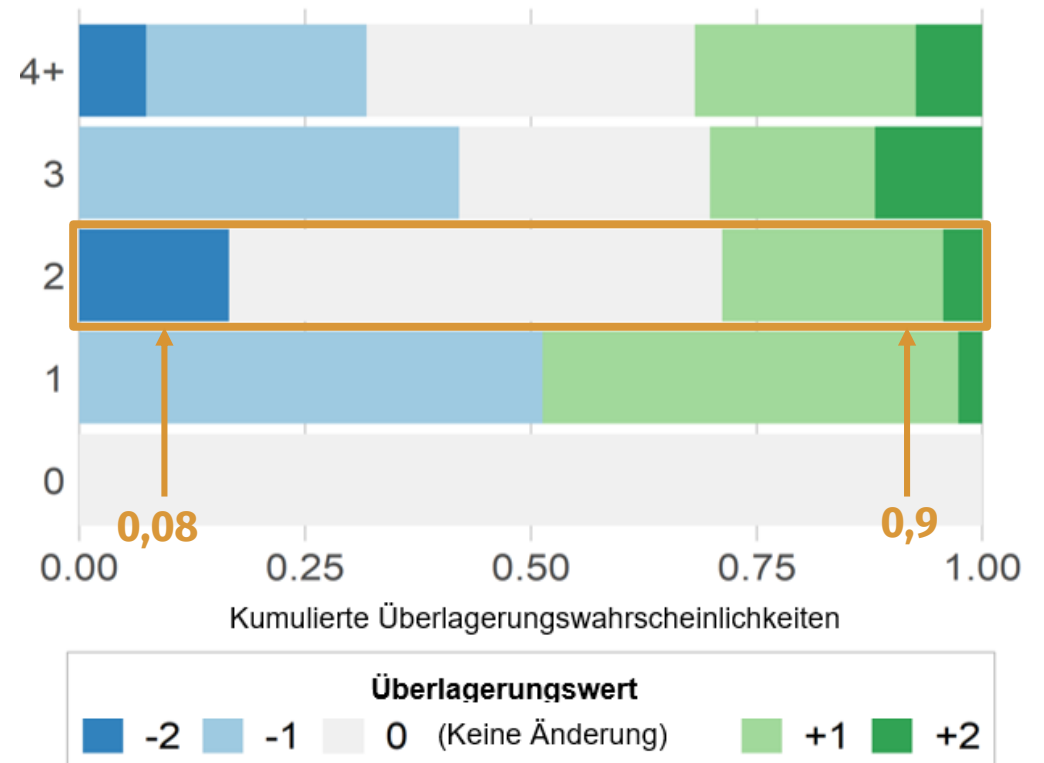
ID	Geschlecht	Altersgruppe	Record Key
1	W	B	0,76
2	M	C	0,09
3	M	C	0,81
4	W	A	0,32

Altersgruppe				
Geschlecht	A	B	C	Total
M			0,09+0,81=0,9	
W				0,76+0,32=1,08
Total				



Funktionsweise der Cell-Key-Methode (CKM)

3. Ablesen des **Überlagerungswerts** von einem festgelegten **Überlagerungstableau** anhand des **Originalwerts** und der **Summe der Zufallszahlen**





Funktionsweise der Cell-Key-Methode (CKM)

4. Addieren des **Überlagerungswerts** zum **Originalwert**

		Altersgruppe			
Geschlecht		A	B	C	Total
	M			$2+1=3$	
	W				$2-2=0$
	Total				