

Geheimhaltung im Zensus 2022

Informationsveranstaltung für Kommunen mit abgeschotteter
Statistikstelle

Digital, 27.06.2024

Statistisches Bundesamt, Statistisches Landesamt Bremen,
Stadt Dortmund, KOSIS-Gemeinschaft DUVA

Herzlich Willkommen Informationsveranstaltung „Geheimhaltung im Zensus 2022“



Schalten Sie Ihr [Mikrofon auf stumm](#), wenn Sie nicht sprechen



Nutzen Sie die [Chatfunktion](#), um eine Frage zu stellen



Begrüßung und Vorstellung der Agenda



Agenda

1. Begrüßung
2. Geheimhaltungsmethode
3. Vorstellung KORHB App
4. Vorstellung DUVA
5. Abschluss



Geheimhaltungsmethode - Hintergrund

- §32 Abs. 2 ZensG2022 - Zensus-Einzeldatenmaterial für KaSS erhältlich
- Lieferumfang:
 - Kommunenhandbuch,
 - Einzeldaten,
 - Merkmalskatalog,
 - Datensatzbeschreibung
- Bereitstellung über Statistische Landesämter
- Auswertungen innerhalb der abgeschotteten Statistikstelle auf Originaldaten möglich



Geheimhaltungsmethode - Hintergrund



- Vertrauen der Auskunftgebenden und der Schutz ihrer Daten wichtig
 - §16 BStatG => Veröffentlichungen nur mit Geheimhaltung
- > Veröffentlichung meint alle Auswertungen, die den Bereich der abgeschotteten Statistikstelle verlassen



Geheimhaltungsmethode beim Zensus 2022



- Traditionelle Verfahren nicht für den Zensus geeignet
- Untersuchungen zum Einsatz datenverändernder Verfahren
- ➔ Entscheidung von Bund und Ländern zu post-tabularem Verfahren:

Cell-Key-Methode



Geheimhaltungsmethode - Cell-Key-Methode

- Überlagerung durch Addition eines „zufälligen“ Störterms
- Anwendung auf Originalwerten
- Störterm kann auch 0 sein

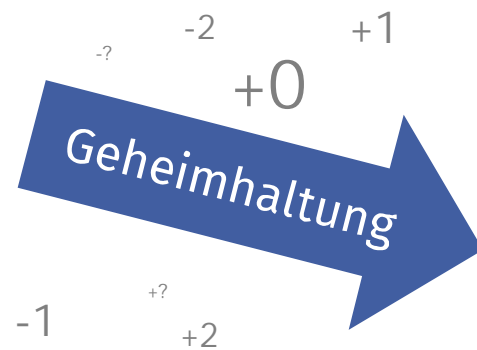
- Unterschiedliche Verfahren abhängig vom Merkmalstyp
 - Bsp. Fallzahlen, Anteile, Durchschnitte, ...
- Tabellenfelder unabhängig voneinander überlagert



Geheimhaltungsmethode - Cell-Key-Methode

Originale Tabelle

Gebäudetyp	Art des Gebäudes		Insgesamt
	bewohnt	unbewohnt	
Frei	4	1	5
Doppel	3	5	8
Reihe	0	2	2
Insgesamt	7	8	15



Veröffentlichungstabelle

Gebäudetyp	Art des Gebäudes		Insgesamt
	bewohnt	unbewohnt	
Frei	6	3	5
Doppel	4	5	7
Reihe	0	4	4
Insgesamt	8	6	16



Geheimhaltungsmethode - Cell-Key-Methode

Vorteile

- Hohe, vorgegebene **Genauigkeit**
- **Konsistenz** über die Tabellen hinweg

Herausforderungen

- **Fehlende Additivität** wegen separater Überlagerung von Innenfelder und Randsummen
- **Kennzahlen** (z.B. Anteilswerte)
- Sehr geringes Restaufdeckungsrisiko



Geheimhaltungsmethode – Beispiel Fallzahlen

Einzeldaten

ID	Art des Gebäudes	Gebäudetyp	Baujahr	n	Record-Key
1	bewohnt	Frei	Vor 1950	1	0,35
2	bewohnt	Frei	Vor 1950	1	0,91
3	bewohnt	Frei	1990 und später	1	0,70
4	unbewohnt	Frei	1950-1969	1	0,67
5	bewohnt	Frei	1970-1989	1	0,01

Häufigkeitstabelle:

Art des Gebäudes

Art des Gebäudes	i_original	Cell-Key
bewohnt	4	
unbewohnt	1	
Insgesamt	5	



Geheimhaltungsmethode – Beispiel Fallzahlen

Einzeldaten

ID	Art des Gebäudes	Gebäudetyp	Baujahr	n	Record-Key
1	bewohnt	Frei	Vor 1950	1	0,35
2	bewohnt	Frei	Vor 1950	1	0,91
3	bewohnt	Frei	1990 und später	1	0,70
4	unbewohnt	Frei	1950-1969	1	0,67
5	bewohnt	Frei	1970-1989	1	0,01

Häufigkeitstabelle:

Art des Gebäudes

Art des Gebäudes	i_original	Cell-Key
bewohnt	4	0,97
unbewohnt	1	0,67
Insgesamt	5	0,64

Cell-Key Gebäude, bewohnt:

$$0,35 + 0,91 + 0,70 + 0,01 = 1,97$$

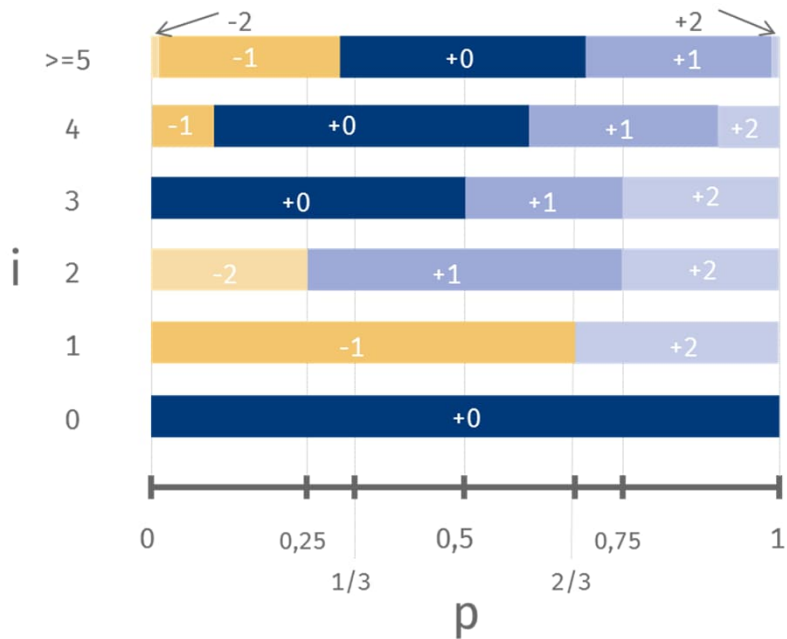
Auf Nachkommastellen reduziert:

$$\rightarrow 0,97$$



Cell-Key-Methode – Beispiel Fallzahlen

Überlagerungstableau (Beispiel)



Häufigkeitstabelle:

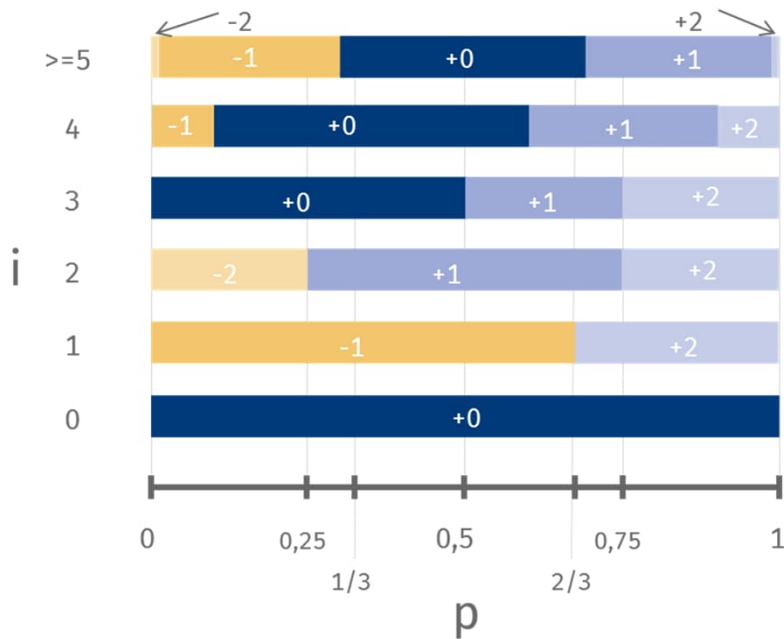
Art des Gebäudes

Art des Gebäudes	i_original	Cell-Key	Differenz	j_überlagert
bewohnt	4	0,97	+2	6
unbewohnt	1	0,67	+2	3
Insgesamt	5	0,64	0	5



Cell-Key-Methode – Erweitertes Beispiel Fallzahlen

Überlagerungstableau (Beispiel)



Häufigkeitstabelle:

Art des Gebäudes x Gebäudetyp

Art des Gebäudes	Gebäudetyp	i_original	CK	Diff.	j_überlagert
bewohnt	Frei	4	0,97	+2	6
bewohnt	Doppel	3	0,74	+1	4
bewohnt	Reihe	0	0	0	0
unbewohnt	Frei	1	0,67	+2	3
unbewohnt	Doppel	5	0,47	0	5
unbewohnt	Reihe	2	0,99	+2	4
bewohnt	Insgesamt	7	0,71	+1	8
unbewohnt	Insgesamt	8	0,13	-1	7
Insgesamt	Frei	5	0,64	0	5
Insgesamt	Doppel	8	0,21	-1	7
Insgesamt	Reihe	2	0,99	+2	4
Insgesamt	Insgesamt	15	0,84	+1	16



Cell-Key-Methode – Vorteil: Konsistenz

Häufigkeitstabelle:

Art des Gebäudes x Gebäudetyt

Art des Gebäudes	Gebäudetyt	i_original	CK	Diff.	j_überlagert
bewohnt	Frei	4	0,97	+2	6
bewohnt	Doppel	3	0,74	+1	4
bewohnt	Reihe	0	0	0	0
unbewohnt	Frei	1	0,67	+2	3
unbewohnt	Doppel	5	0,47	0	5
unbewohnt	Reihe	2	0,99	+2	4
bewohnt	Insgesamt	7	0,71	+1	8
unbewohnt	Insgesamt	8	0,13	-1	7
Insgesamt	Frei	5	0,64	0	5
Insgesamt	Doppel	8	0,21	-1	7
Insgesamt	Reihe	2	0,99	+2	4
Insgesamt	Insgesamt	15	0,84	+1	16

Häufigkeitstabelle:

Baujahr (univariat/eindimensional)

Baujahr	i_original	CK	Diff.	j_überlagert
Vor 1950	5	0,74		
1950-1969				
1970-1989		0,94		
1990-2009				
2010-2019	3	0,86		
Insgesamt	15	0,84	+1	16

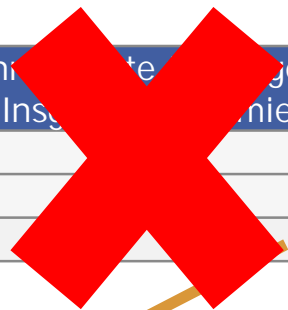
Konsistenz über alle Tabellen



Cell-Key-Methode – Herausforderung: Additivität

Wir überlagern separat:

Art des Gebäudes	i_original	Insgesamt überlagern; Insgesamt zählen	Alle Tabellenfelder separat überlagert
bewohnt	4	6	6
unbewohnt	1	3	3
Insgesamt	5	9	5



Differenz von 4!

→ Genauigkeitsverlust

NICHT immer additiv

d.h. Insgesamt != Summe der Teile

$$6 + 3 \neq 5$$

Jedes Tabellenfeld +/- 2
(in diesem Beispiel!)

Vorteil:
Höhere Genauigkeit



Kennzeichnungen

- Herausforderungen bei Kennzahlen (Anteile/Quoten, Mittelwerte, etc.)
 - Korrektur: Werte $>100\%$
 - Klammerung: Werte mit hohen Abweichungen zum Original
 - Sperrungen
- ➔ Kennzeichnungen werden in den KOSIS Werkzeugen KORHB App und DUVA enthalten sein.



Kennzeichnungen: Nullen & Klammern I

Original

Art des Gebäudes	Anzahl	%
bewohnt	4	80,0
unbewohnt	1	20,0
Insgesamt	5	100,0

Überlagert

Art des Gebäudes	Anzahl	%
bewohnt	6	100,0
unbewohnt	3	(60,0)
Insgesamt	5	100,0

Zeichenerklärung

- Genau Null oder auf Null geändert
- () Aussagewert eingeschränkt, da der Zahlenwert durch das Geheimhaltungsverfahren relativ stark verändert wurde
- Zahlenwert unbekannt oder geheim

Überlagert (Alternative)

Art des Gebäudes	Anzahl	%
bewohnt	6	100,0
unbewohnt	-	-
Insgesamt	5	100,0



Kennzeichnungen: Nullen & Klammern II

Original (neues Beispiel!)

Art des Gebäudes	Anzahl	%
Frei	4	80,0
Doppel	0	0,0
Reihe	1	20,0
Insgesamt	5	100,0

Zeichenerklärung

- Genau Null oder auf Null geändert
- () Aussagewert eingeschränkt, da der Zahlenwert durch das Geheimhaltungsverfahren relativ stark verändert wurde
- Zahlenwert unbekannt oder geheim

Überlagert

Art des Gebäudes	Anzahl	%
Frei	6	100,0
Doppel	-	-
Reihe	3	(60,0)
Insgesamt	5	100,0

Überlagert (Alternative)

Art des Gebäudes	Anzahl	%
Frei	6	100,0
Doppel	-	-
Reihe	-	-
Insgesamt	5	100,0



Geheimhaltungsmethode – Kennzeichnungen II

Original

Art des Gebäudes	Anzahl	%
bewohnt	4	80,0
unbewohnt	1	20,0
Insgesamt	5	100,0

Überlagert (Alternative)

Art des Gebäudes	Anzahl	%
bewohnt	.	.
unbewohnt	.	.
Insgesamt	5	100,0

Zeichenerklärung

- Genau Null oder auf Null geändert
 - () Aussagewert eingeschränkt, da der Zahlenwert durch das Geheimhaltungsverfahren relativ stark verändert wurde
 - .
- Zahlenwert unbekannt oder geheim

Überlagert (Alternative)

Art des Gebäudes	Anzahl	%
bewohnt	.	.
unbewohnt	.	.
Insgesamt	.	.



Zusammenfassung

1. Alle Auswertungen, die den Bereich der abgeschotteten Statistikstelle verlassen sind mit der Cell-Key-Methode zu überlagern.
2. Die Überlagerung muss auf Basis der von den Statistischen Ämtern beschlossenen und angewandten Einstellungen erfolgen.
3. Hierzu stehen zwei KOSIS Auswertungswerkzeuge zur Verfügung
 - KORHB App
 - DUVA
4. Beide Werkzeuge verwenden die CKM-Anwendung (JAVA) des Statistischen Bundesamtes (inkl. Nutzungsbedingungen)



Vielen Dank.