

01

Protection of the European Census 2021

An overview of used methods



Co-funded by
the European Union

**Marieke de Vries¹, Reinhard Tent²,
Manca Golmajer³, Sarah Giessing²,
Peter-Paul de Wolf¹**

¹Statistics Netherlands, ²Destatis,

³Statistics Slovenia

**UNECE Expert Meeting for Statistical
Data Confidentiality, Wiesbaden,
September 27**

02

Agenda

- STACE project
- Survey Results
- Some Perspectives
- Issues of Connected Tables
- Summary and Outlook

03

STACE project

WP2 in project “Statistical methods and tools for time series, seasonal adjustment and statistical disclosure control”

- User support for and maintenance of SDC tools
- Support to producers on implementation and use of the tools
- Provide guidance (guidelines) on SDC methodology



Co-funded by
the European Union



DISTATIS
Statistisches Bundesamt



 Główny
Urząd Statystyczny


Insee
Mesurer pour comprendre


STATISTIK AUSTRIA
Die Informationsmanager

 **Hagstofa Íslands**


НАЦИОНАЛЕН СТАТИСТИЧЕСКИ ИНСТИТУТ
NATIONAL STATISTICAL INSTITUTE

04

STACE project – current work

Some deliverables

- Guidelines on SDC for geo-referenced data
- Update general ESSnet handbook SDC
- Guidelines on SDC for population statistics, after Census 2021

Workshop organized on Census Data protection, May 23rd 2023

- Countries were asked to present their approach for protecting 2021 census
- Presentations by France, Ireland, Portugal, Slovenia, Romania
- Survey on protection of the census sent out to all countries before meeting

05

Online Survey

- Sent out before workshop.
- Completed by 30 countries.
- Questions on:
 - methods,
 - communication,
 - differences in national publication, and
 - challenges during implementation.



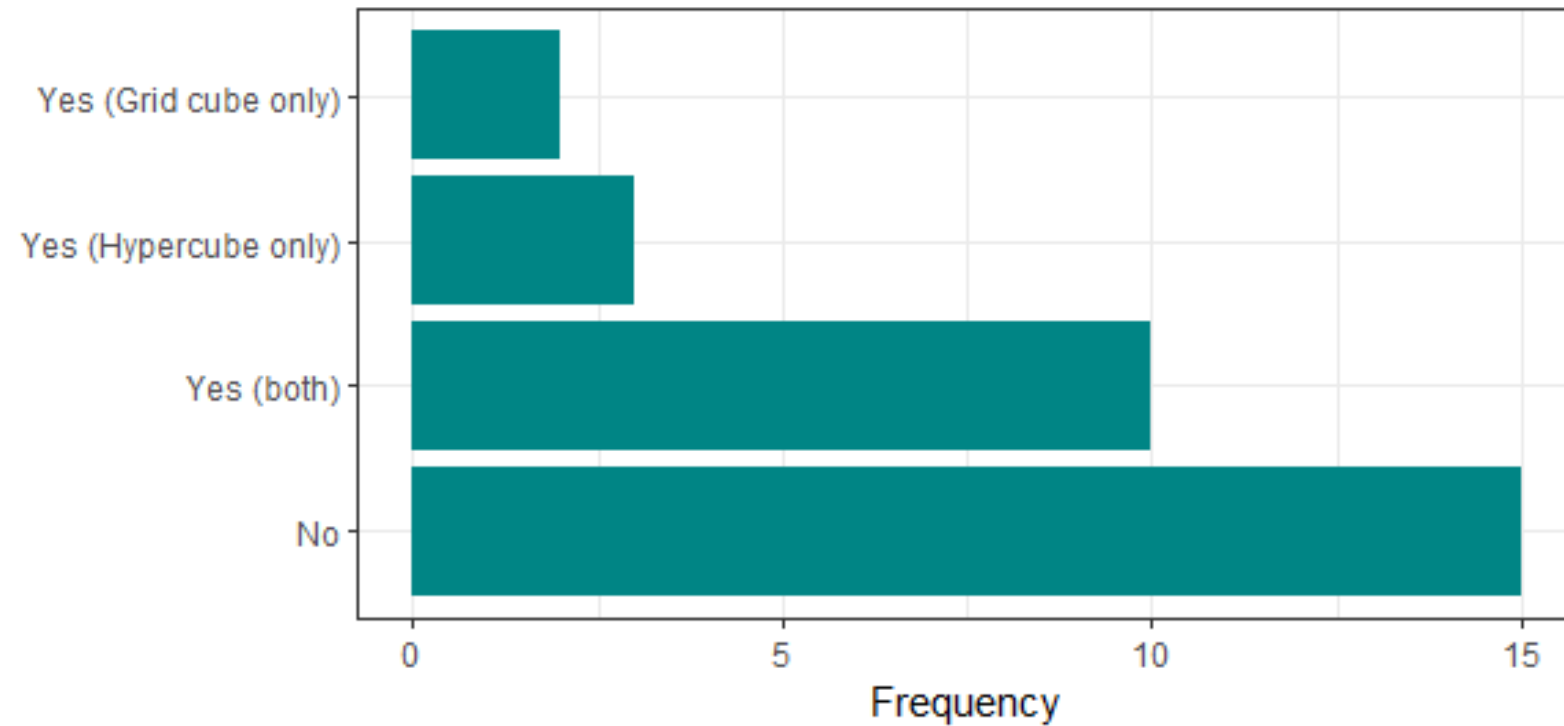
06

Results

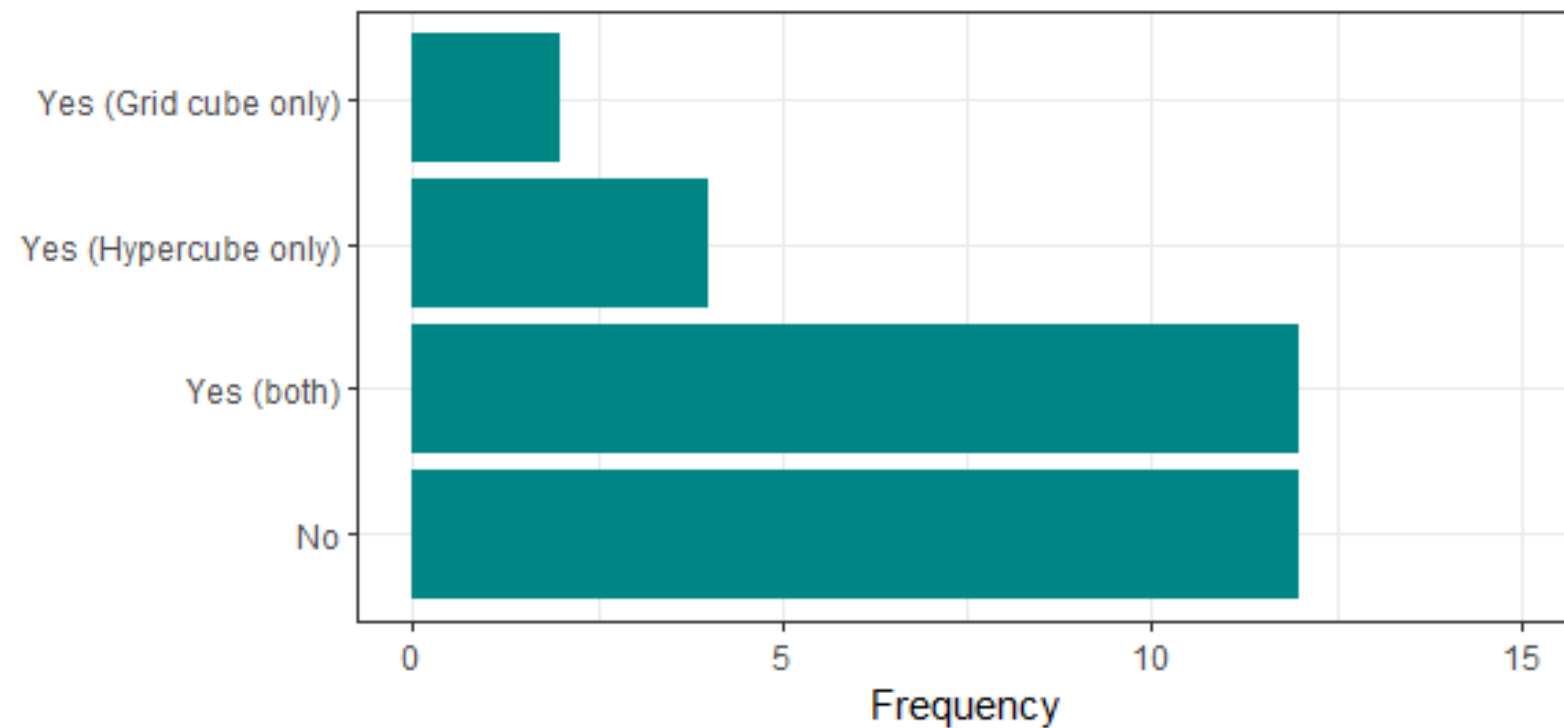
Use of CKM more popular than TRS

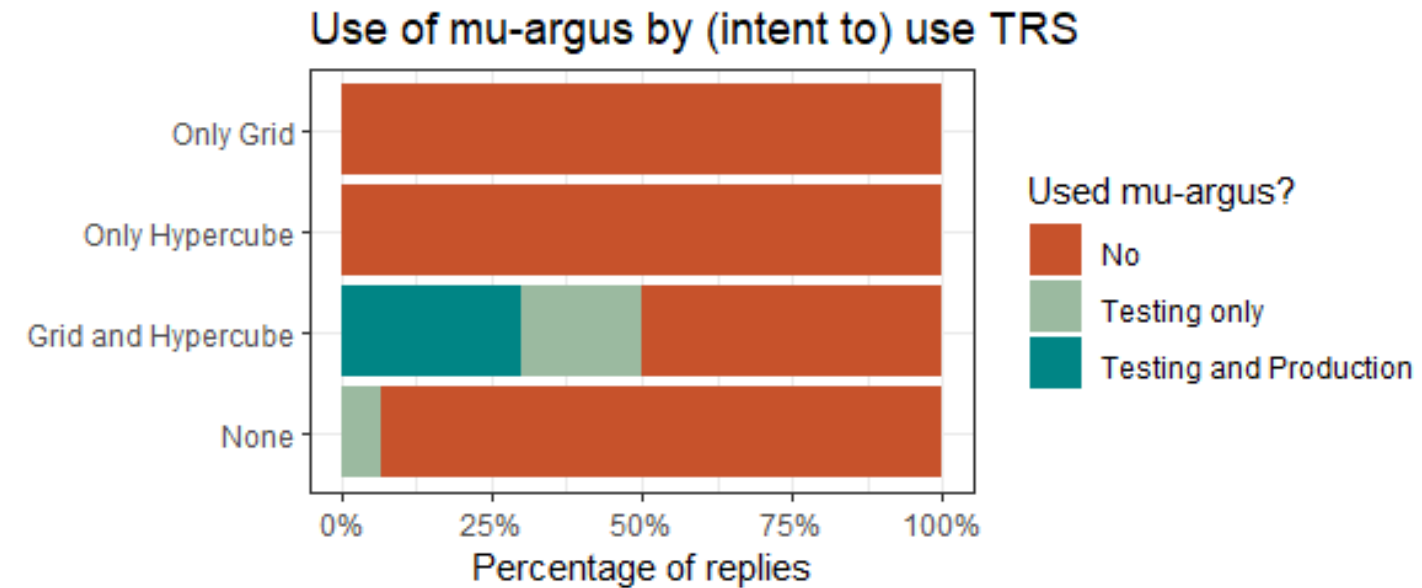
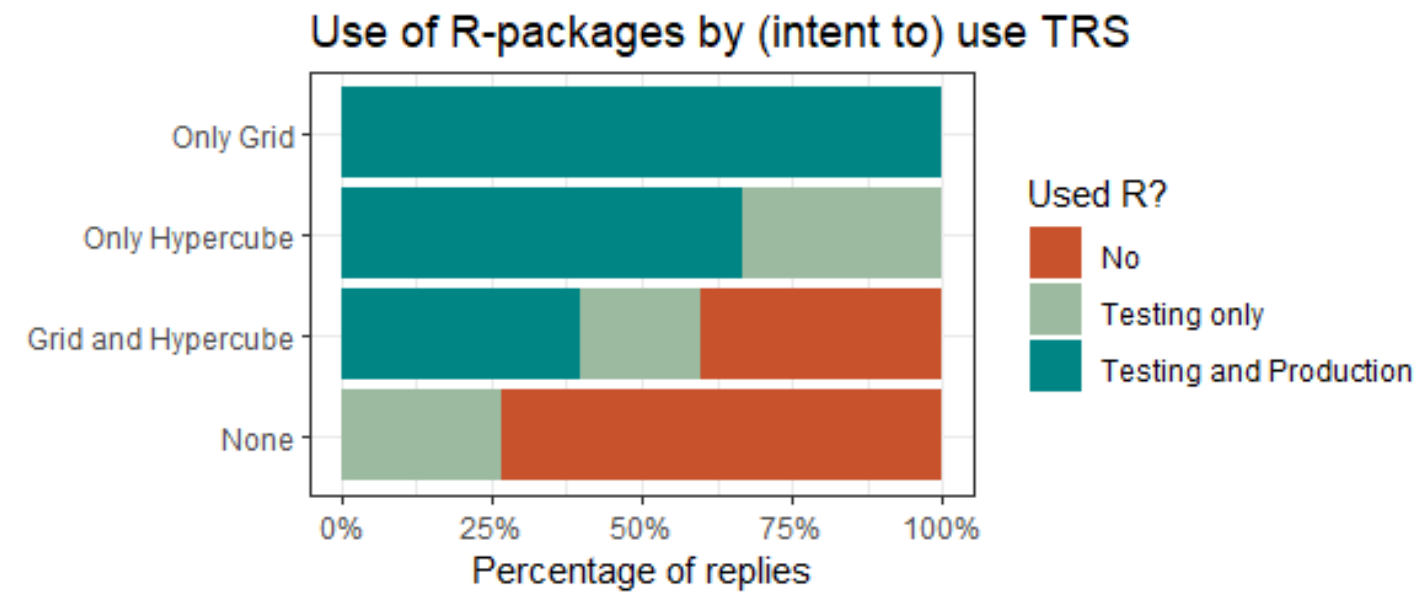
- TRS applied by 50%, CKM by 60%.
Combination of TRS and CKM only 33%.
- Other methods include:
 - Specific methodology designed by NSI
 - SDC methods imposed by law

(Intent to) use of TRS for protection



(Intent to) use of CKM for protection





07

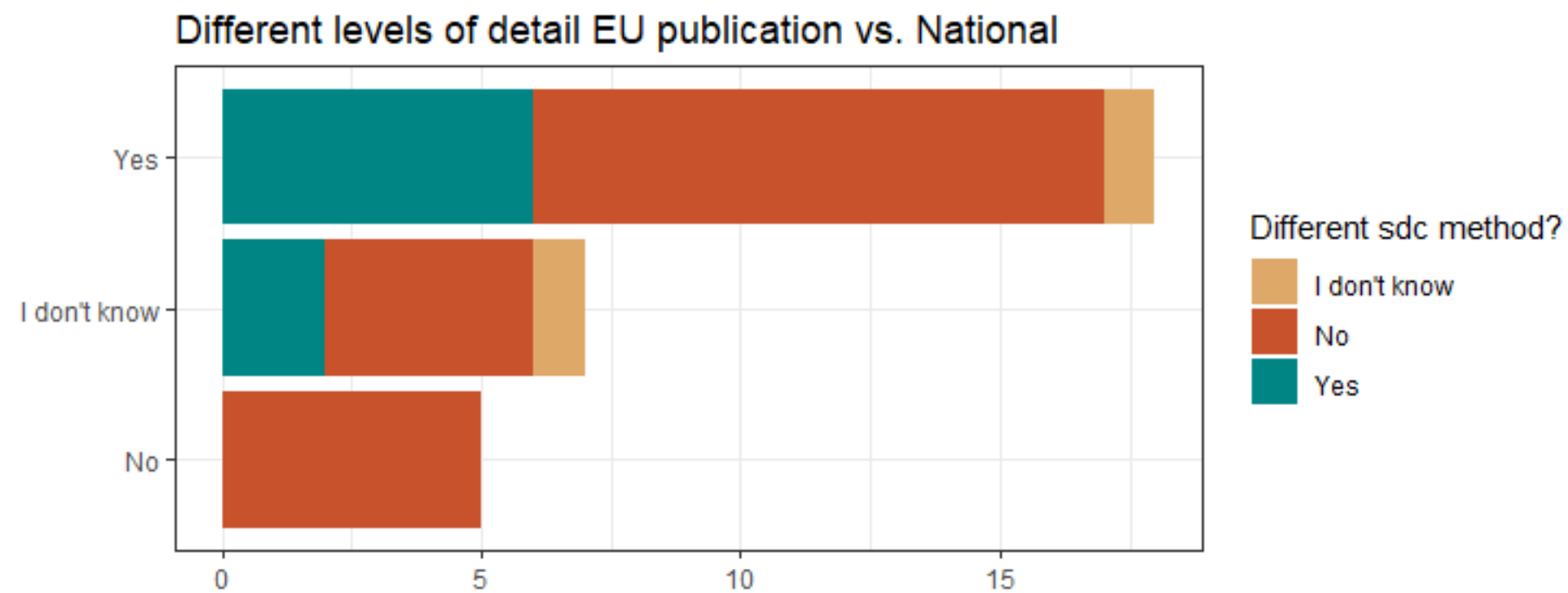
Results

R-packages more popular than mu-argus or tau-argus

- Other methods include:
 - Implementation of TRS or a comparable method using SAS,
 - Building own R scripts and packages for CKM.

08

Results



Majority of countries have different level of detail between publications

- Different sdc methods for EU vs national:
 - Cell suppression,
 - Aggregation levels suppressed
- 11 countries experience consistency issues between publications, including demographic data. Only six solved.

Some perspectives

- Perspectives of France, Ireland, Portugal, Slovenia, Germany,
 - CKM: Slovenia, Germany
 - TRS: France, Ireland, Portugal
- Partially CKM and TRS are combined with additional SDC methods.
 - Coarsening
 - Cell Suppression
- Even if two countries use the same method, implementation may be different.
 - Swapping individuals (France) vs. Swapping households (Ireland).
 - Applying CKM to all data (Germany) vs. applying CKM to grid data and Census hypercubes only.

Pros and cons for both methods that influenced their choices

	CKM	TRS
Maintain additivity	-	+
Evaluation easy/not costly	-	+
Deviation controllable	+	-
Protection on highest aggregation level	+	-

Issue: Connection between Tables

- Overlapping of national annual demographic tables and census tables.
- SDC methods for annual demographic tables might differ from methods for census data.
 - **Issues concerning disclosure risk possible.**
- Theoretical Example:
 - Demographic tables: **Coarsening**
 - Census data: **perturbative disclosure control method**
- If same combinations occur in both tables → Attacker could disclose other values

Two real cases

Slovenia (SURS)

- Hypercubes for Eurostat not yet protected.
 - Very detailed.
 - Overlapping with other publications that use Cell Suppression or no protection at all (due to low level of detail).
- Tests showed: CKM only not sufficient.

Germany (Destatis)

- SDC method for National demographic tables not yet decided.
 - Overlapping with Census publications, which use CKM.

Recommendation of Eurostat

- Usage of both TRS and CKM.
- For consistency: impute the overlapping cell values from the annual tables into the census hypercubes.

Summary and Outlook

- Advised SDC methods for Census 2021 : TRS and CKM
- Survey showed:
 - TRS: Used on at least one of the Census data products by 50% of 30 respondents
 - CKM: Used on at least one of the Census data products by 60% of 30 respondents
- First time that so many European countries apply these techniques.
- Method and code development the most challenging part.
- Intended guidelines on ‘SDC for population statistics’ should discuss the lessons learned.
- Discuss whether the new techniques and experiences can also be used for other statistics
- Continue the current work in the future ESOP* (combines census and demography)
- Consistency issues regarding demography data released for the exact same reference date.
- Guidelines how to apply the new protection methods might be helpful.

15

Thank you!
