



REPUBLIKA SLOVENIJA



STATISTIČNI URAD REPUBLIKE SLOVENIJE  
STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA



[www.stat.si](http://www.stat.si)

THE SLOVENIAN  
NATIONAL  
STATISTICS

TRUSTWORTHY  
AND  
USER-ORIENTED

# Dissemination Strategy and Statistical Disclosure Control of the 2011 Register-based Census in Slovenia

Danilo Dolenc

Statistical Office of the Republic of Slovenia

ESSnet Workshop on Statistical Disclosure Control of Census Data,  
Luxembourg, 19-20 April 2012

# Background

---

- Census 1991
  - Only data on ethnicity at the lowest territorial level suppressed
    - Threshold 30 inhabitants
- Census 2002
  - Cell suppression method
    - Threshold 3 used in general
  - Additional very strict rules for data on ethnicity
  - Negative feedback from users
  - Huge workload

# 2011 Register-based Census

---

- Reference date 1 January
- Processing already concluded
  - Except some housing data
- Four stages of the statistical process followed by dissemination of final results
  - Basic demographic data (30 April 2011)
  - Households and dwellings (30 June 2011)
  - Other population topics (30 December 2011)
    - Including occupied dwellings – preliminary data
  - Housing (by the end of June 2012)

# New dissemination strategy

- Starting points for consideration for 2011:
  - Users needs
    - Including demand for small area and micro data
  - Information privacy should be respected
  - Assure maximum information value
  - Are public data on individual basis available
  - Sensitive variables (e. g. citizenship)
  - Human resources and SDC tools
  - Smallness of population
    - Administrative division includes 211 municipalities and 6,000 settlements (one third with less than 30 persons)

# Dissemination policy

---

- Interdependence of territory and breakdowns of variables
  - Higher level of territory (NUTS 0 to NUTS 3)
    - More detailed breakdowns
    - More variables in aggregation
      - Limited to 5 including geographical area
  - Lowest level of territory (NUTS 5 and lower)
    - Threshold is set up (30 inhabitant) for NUTS 7
    - Only basic very aggregated data available
    - Maximum 2 variables
      - Exception is age-sex structure

# Example

- ❑ No SDC applied to the levels of municipality
  - ❑ Every six settlement data suppressed in whole
  - ❑ Very positive users reactions

Variable	Geographical area				
	NUTS 0	NUTS 2	NUTS 3	NUTS 5 (LAU 2)	NUTS 7
	Slovenia	Cohesion region	Statistical region	Municipality	Settlement
AGE	1 - year age groups, 15-85+	5 - year age groups, 15-65+	5 - year age groups, 15-65+	no	no
SEX	yes	yes	yes	yes	no
EDUCATION	11 categories	9 categories	7 categories	7 categories	3 categories

22 categories of educational attainment in code-book

# Hypercubes and PMDs

## □ Main findings

- Average number of variables in hypercubes
  - 7 at NUTS 2 level
  - 5 at NUTS 3 level
  - 3 at LAU 2 level
    - Harmonized with our dissemination policy
- Average number of variables in PMDs
  - more than 5 at NUTS 2 level
  - 4 at NUTS 3 level
  - no PMD at LAU 2 level

## SDC in case of hypercube 2

---

- 8 variables in hypercube
  - Not possible to execute SDC
  - 3,276 mio cells including totals
  - 127,008 cells without totals
- 6 PMDs
  - Threshold 3 applied
- ARGUS version 3.5.0 used for SDS
  - Hypercube method selected
  - No cross-protection



# Hypercube 2 and PMDs

Hyper-cube	No of categories *	PMD 1	PMD 2	PMD 3	PMD 4	PMD 5	PMD 6
GEO.L.	<b>3</b>	GEO.L.	GEO.L.	GEO.L.	GEO.L.	GEO.L.	GEO.L.
SEX.	<b>3</b>	SEX.	SEX.	SEX.	SEX.	SEX.	SEX.
HST.H.	<b>13</b>	HST.H.	HST.H.	HST.H.	HST.H.	HST.H.	HST.H.
EDU.	<b>8</b>	EDU.	EDU.	EDU.	CAS.L.	POB.L.	COC.L.
CAS.L.	<b>5</b>	AGE.M.	CAS.L.	CAS.L.	AGE.M.	AGE.M.	AGE.M.
POB.L.	<b>5</b>		POB.L.	COC.L.			
COC.L.	<b>5</b>						
AGE.M.	<b>28</b>						
* categories not relevant for Slovenia not included							

# PMDs- indicators

	PMD_1	PMD_2	PMD_3	PMD_4	PMD_5	PMD_6
Nr. of cells	26208	23400	23400	16380	16380	16380
Empty cells	8103	3978	5197	4893	1929	2727
Numerically safe cells	10310	3086	2621	8504	8379	5905
Numerically unsafe cells	7795	16336	15582	2983	6072	7748
Primary confidential cells	945	1237	1476	206	540	886
Secondary suppressions	6850	15099	14106	2777	5532	6862
Information loss (%)	29,7	69,8	66,6	18,2	37,1	47,3
Maximum value suppressed	226148	2050189	2050189	293905	355435	575287

# Conclusion

---

- ❑ Does EU really need so detailed data
  - Hypercubes are easy to produce but difficult to protect
- ❑ Is there an appropriate tool for efficient SDC
  - Our aim is to transmit valuable data and not just suppressions
- ❑ Do we have alternatives for SDC
  - Additional aggregation of categories that are most sensitive
- ❑ 2011 Census data will be already outdated in 2014

Thank you for your attention!

---

