

Behind the Scenes: Crafting Hungary's New Census Database

Melinda Oparin-Salamon



EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL





EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Development of the database

- 2022 idea emerged
- From February 2023
 - 2011 census data from dissemination database
 - Further development on opinion research
- June 2023
 - First publication of 2022 census data on settlements
- September 2023
 - Publication of population data
 - Smaller improvements
- January 2024: over 200000 tables downloaded from 45 datasets



eurostat 

The conference is partly
financed by the European Union



EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

In-house development

- Tailored solutions
- Control and ownership
- Flexibility and adaptability
- Cost-effectiveness
- Expertise utilization
- Data security and confidentiality



EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

SDMX in the Census Database

- SDMX is not just a file format
- Standard tools and definitions
- Recommendations for technical implementation
 - URL patterns
 - API
 - Solutions for technological limitations



eurostat 

The conference is partly
financed by the European Union



Data pipeline

- Data in SAS
- Data and meta data in Excel files from Population Census and Demographic Statistics Department
- From meta to sdmx-json (Node.js)
- For backend jq conversion (dataset definition, codelists)
- Definition of attributes
- Data conversion (ssconvert), data transformation (Python)
- Aggregations with scripts (Bash)
- Templates for operations, dimensions, dataflows
- Indexing
- Deployment
 - Scripts
 - Docker



Census database

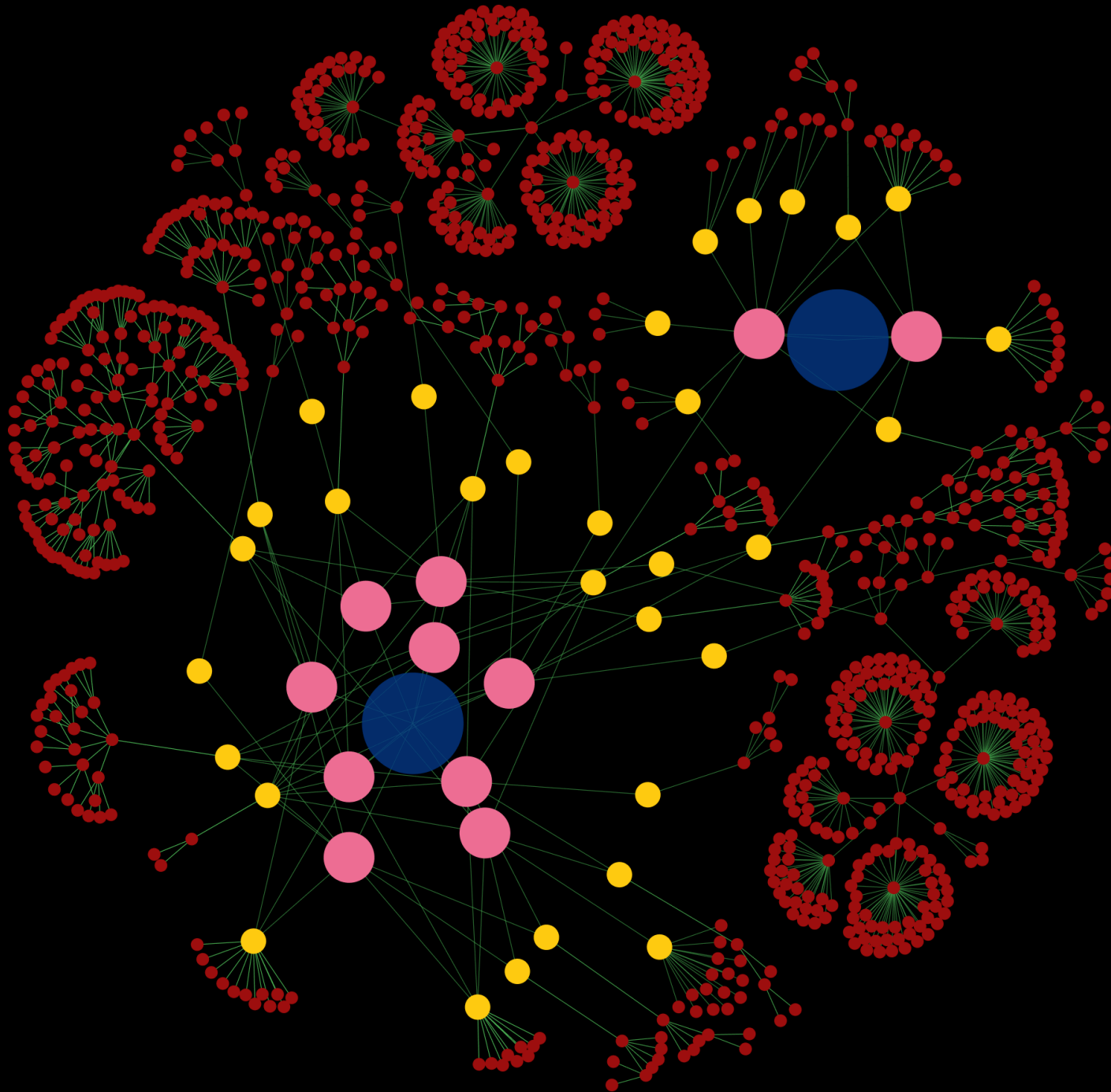
Census database



Population by economic activity, county and type of settlement

¹

	Sex							Male	
	Total	Lower than general (primary) school (8th grade)	General (primary) school the 8th grade completed	Secondary level without final examination, with final vocational exam	Secondary level with final examination	Tertiary education	Persons under 15 years of age	Total	Lower than general (primary) school (8th grade)
Highest level of education completed	Total								
Types of settlement									
Census year: 2022									
Total	9 603 634	187 268	1 739 464	1 733 767	2 716 187	1 833 716	1 393 232	4 620 846	69 61
· Towns	6 737 468	106 170	1 059 253	1 081 922	2 034 058	1 528 263	927 802	3 204 480	38 90
· · Capital (Budapest)	1 685 342	17 418	195 230	160 344	528 740	573 877	209 733	791 042	6 93
· · Towns of county rank	1 958 227	24 776	279 482	310 591	626 638	462 467	254 273	923 432	9 08
· · Other towns	3 093 899	63 976	584 541	610 987	878 680	491 919	463 796	1 490 006	22 89
· Village	2 866 166	81 098	680 211	651 845	682 129	305 453	465 430	1 416 366	30 70



EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Metadata Graph of the Census Database



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat 

The conference is partly
financed by the European Union



EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Future plans

- Further developments
- New dissemination database on the basis of the census database
 - Huge challenge
 - About 1000 datasets

Thank you for your attention!

Melinda Oparin-Salamon
melinda.salamon@ksh.hu



EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS 2024 ESTORIL - PORTUGAL