

Microdata availability

Cees Steijn



Statistics
Netherlands

Microdata availability

- Introduction;
- Overview available datasets;
- Work process to get data to Remote Access/
On Site environment;
- Key variables
- Datasets research project

History of datasets

From 30 datasets available in 2003
to 3,416 datasets in February 2017.

441 unique filenames

273 active filenames

Yearly between 400 to 600 files are placed in
the secured environment.

Datasets per theme

| Theme | Unique fn | Number datasets |
|-------------------------------------|-----------|-----------------|
| • Labour and social security | 154 | 1338 |
| • Enterprises | 45 | 506 |
| • Population | 44 | 168 |
| • Construction and dwellings | 34 | 182 |
| • Health and welfare | 38 | 268 |
| • Manufacturing industry and energy | 2 | 18 |
| • Income and expenditure | 34 | 270 |
| • International trade | 1 | 6 |
| • Agriculture | 3 | 12 |
| • Nature and environment | 1 | 8 |
| • Education | 42 | 349 |
| • Government and politics | 5 | 27 |
| • Prices | 4 | 31 |
| • Security and justice | 21 | 181 |
| • Traffic and transport | 5 | 25 |
| • Leisure and culture | 8 | 27 |
| Total | 441 | 3416 |

Use by projects

441 unique filenames of which 364 has been used by one or more projects

77 filenames are never used by RA-projects

| | |
|---|----|
| Filenames Used by 1 Project | 49 |
| Filenames Used by 2 -5 Projects | 95 |
| Filenames Used by 6 -10 Projects | 58 |
| Filenames Used by 11 - 20 Projects | 71 |
| Filenames Used by 21 - 30 Projects | 34 |
| Filenames Used by 31 - 50 Projects | 24 |
| Filenames Used by more than 50 Projects | 33 |

Work process

- Data available intern Statistics NL
 - Data Service Center (DSC);
 - Administrated in Commence;
 - Transpose micro- and meta-data to standard format
 - Micro- and Metadata available for projects

From DSC to RA-environment 1

- Get notification by e-mail dataset is available
- Administrate in Commence
- Starting the documentation process
- Using tools to transpose the microdata and metadata to standard format for RA/OS environment.
 - Microdata to SPSS file by batch
 - Metadata to PDF file

From DSC to RA-environment 2

- Place files in the intern environment and the secured environment .
- Notify the active projects who are granted on this file
- If its an new version, we give the reason of the version usually its from provisional changed into definite.
- If its an new period than it speak for itself.



Key variables 1

- Persons identified by:
 - RINPERSOONS and RINPERSOON
 - RINPERSOONS together with the RINPERSOON, identifies a natural person. The code indicates the source (register) from which the personal ID is derived.
- Households identified in some datasets by:
 - RINPERSOONSKERN and RINPERSOONKERN, which is the key/identifier of the household.

Key variables 2

- Addresses identified by :
 - SOORTOBJECTNUMMER and RINOBJECTNUMMER
- Companies identified by:
 - Business Identification Number (BE_ID)
 - or Chamber of Commerce file number (KvK)
 - or Tax Identification Number (Finr)

datasets research project

- Importing own datasets
- Instructions for file format and upload process
- anonymise datasets and create the link with the relevant CBS microdata
- In 2016 it concerns about 120 data files

Examples of English Documentation

GBAPERSOONTAB

Personal characteristics of all persons entered in the Municipal Personal Records Database (GBA)

HOOGSTEOPLTAB

Highest level of education attained and followed and stream of education of the population in the Netherlands

INTEGRAAL HUISHOUDENS INKOMEN

Integrated Household Income

BAANKENMERKENBUS

Characteristics of employee jobs

ABR

General Business Register

Favourite datasets

| Theme | datasetname | number of projects |
|------------------------------|-------------------------------|--------------------|
| • Population | GBAPERSOONTAB | 339 |
| • Population | GBAHUISHOUDENSBUS | 187 |
| • Education | HOOGSTEOPLTAB | 159 |
| • Population | GBAADRESBUS | 152 |
| • Income and expenditure | INTEGRAAL HUISHOUDENS INKOMEN | 151 |
| • Labour and social security | BAANKENMERKENBUS | 150 |
| • Labour and social security | EBB | 144 |
| • Population | GBAADRESOBJECTBUS | 129 |
| • Income and expenditure | SECMBUS | 118 |
| • Labour and social security | BAANSOMMENTAB | 115 |
| • Enterprises | ABR | 112 |
| • Health and welfare | DO | 99 |

Internet address

- <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research>