

# Platforms for aggregated output databases

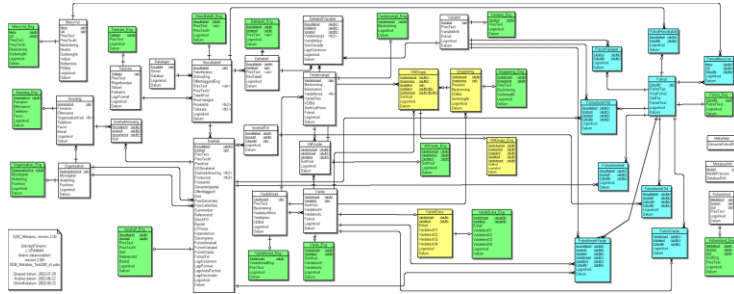
Presentation originally developed by Lars Knudsen  
(lak@dst.dk)

## What is an output database?

- A contextual definition:  
*a collection of statistical information on the Internet, that is organized so that it can easily be accessed, managed, and updated.*
- An example: [EuroStat](#)
- Micro or **Macro** data?
- Datawarehouse – Terminology?
- Most are aggregated!



## A relational data model (example)

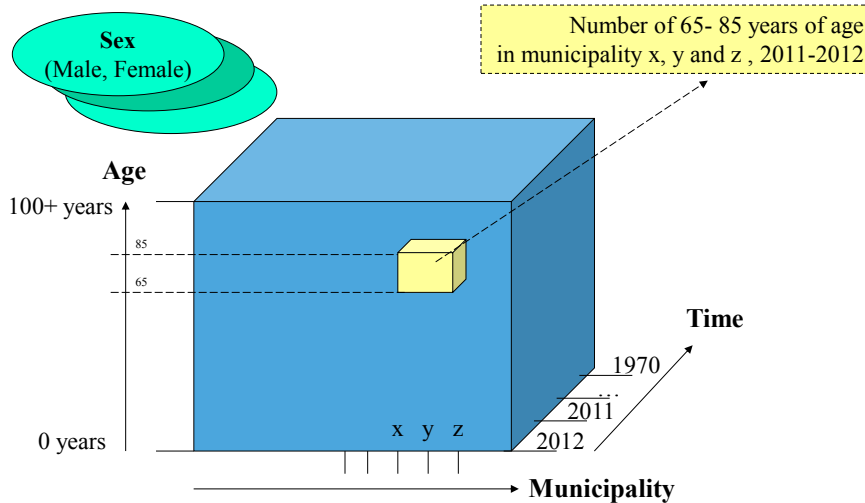


## Related Dataset (example)

Kode "M"  
for "Male"

| MUNICIPALITY | SEX | AGE   | CIVILSTATUS | TIME   | CONTENT |
|--------------|-----|-------|-------------|--------|---------|
| 000          | K   | 16-24 | OFF         | 2000K1 | 33      |
| 000          | M   | 16-24 | OFF         | 2000K1 | 57      |
| 000          | TOT | 16-24 | OFF         | 2000K1 | 90      |
| 000          | K   | 25-29 | OFF         | 2000K1 | 204     |
| 000          | M   | 25-29 | OFF         | 2000K1 | 110     |
| 000          | TOT | 25-29 | OFF         | 2000K1 | 314     |
| 000          | K   | 30-34 | OFF         | 2000K1 | 416     |
| 000          | M   | 30-34 | OFF         | 2000K1 | 183     |
| 000          | TOT | 30-34 | OFF         | 2000K1 | 599     |
| 000          | K   | 35-39 | OFF         | 2000K1 | 752     |
| 000          | M   | 35-39 | OFF         | 2000K1 | 378     |
| 000          | TOT | 35-39 | OFF         | 2000K1 | 1130    |
| 000          | K   | 40-44 | OFF         | 2000K1 | 1104    |
| 000          | M   | 40-44 | OFF         | 2000K1 | 477     |

## Slice - and - dice



## Why use an output database?

- Web is recognized as the main dissemination channel
- Output databases offer tailor-made outputs to versatile users
- No paper limits (size of tables)
- Can offer long time series
- Can offer detailed groupings of geography, industries, commodities etc.
- Feed your publications
- Feed your web site
- Offer API's – programmatic access

## Web pages vs. Database

Shop



Storage



## Consistency with publications, web etc.

- One source serving multiple needs
- Prepare publications using (unpublished) data from dissemination system
- Easy updating (demo)?
- Error checking
- Consistent metadata?



## Demo: From database to web pages

### Example:

- [Statbank](#) ->
- [Web](#) / News release / Publications etc...



## Some products on the market

- PX-Web
- Beyond 2020
- SuperWeb
- OECD.Stat
- StatLine (Netherlands)
- TGM (EuroStat)
  
- Other tailor-made solutions



## About the web user interface

- Should not require software installation by users
- Should be user-friendly without additional learning (cognitive perception)
- Support versatile ways of data extraction
- Provide different download formats



## Common flow ...

- Define or redefine selection
- Rearrange stub and heading
- Choose multiple output formats
- Save tables / extractions for later use
- Notification of changes / updates
- Additional documentation- included / external links

## Some considerations ...

- Should it be free or not?
- Should it contain all NSI data?
- Cater for more than one system?
- Relationship between micro and macro data
- In more than one language?
- Data Cubes versus Time series
- etc...

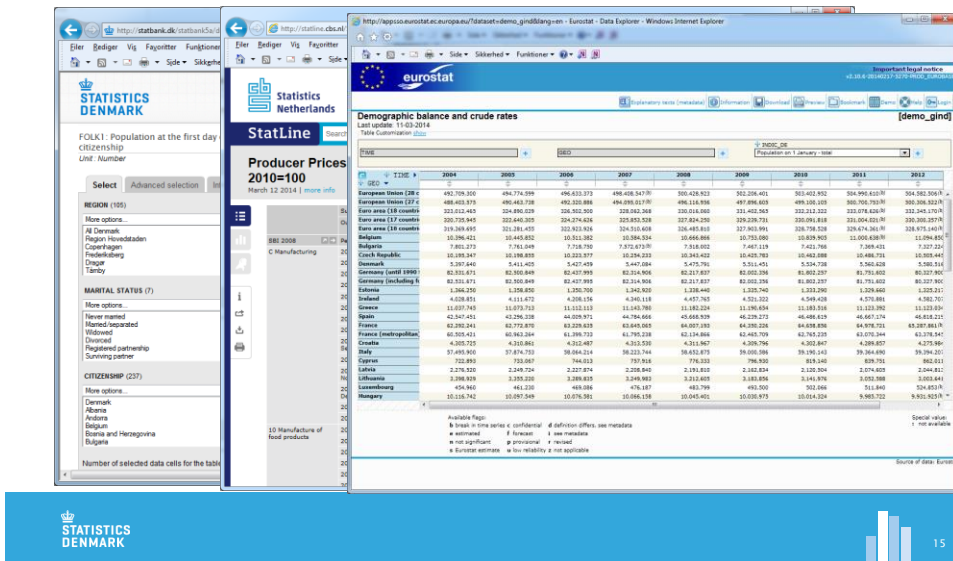


## Demo / examples – Different but same

- <http://epp.eurostat.ec.europa.eu> ...
- <http://stats.oecd.org>
- <http://www.ine.es>
- <http://www.statbank.dk>
- <http://statline.cbs.nl/statweb/?LA=en>
- <http://www.stat.si/pxweb/>



# Predefined table before selection?



## Data and MetaData

- What is a Dataset?
  - Only codes and figures
- What is MetaData?
  - Everything else related to data
  - Title, texts, footnotes etc.

*Example*

## Public and private metadata

- Devide access to metadata depending on the nature
- Public metadata:  
Centralize adm.
- Private metadata:  
Decentralize adm.

Population 1. January by age, sex, region and time

|               | 2004   | 2003   | 2002   | 2001   |
|---------------|--------|--------|--------|--------|
| 33 years      |        |        |        |        |
| Women         |        |        |        |        |
| All Denmark   | 37 267 | 36 865 | 38 065 | 41 068 |
| Copenhagen    | 4 818  | 4 547  | 4 602  | 4 972  |
| Frederiksberg | 878    | 804    | 919    | 939    |

Before 1 January 2003 Bornholm includes Christianse.

Unit : Number

Contact : Jeanne Rasmussen [jso@dst.dk](mailto:jso@dst.dk) ☎ +45 39173308

## Need for coordinated metadata

- How to organise database administration
- Common- and table specific metadata (Public and private)
- Coherence between tables and valuesets
- Policies are important for success

## Organisational structures

- How to organise database administration
- Common- and table specific metadata (Public and private)
- Coherence between tables and valuesets
- Policies are important for success



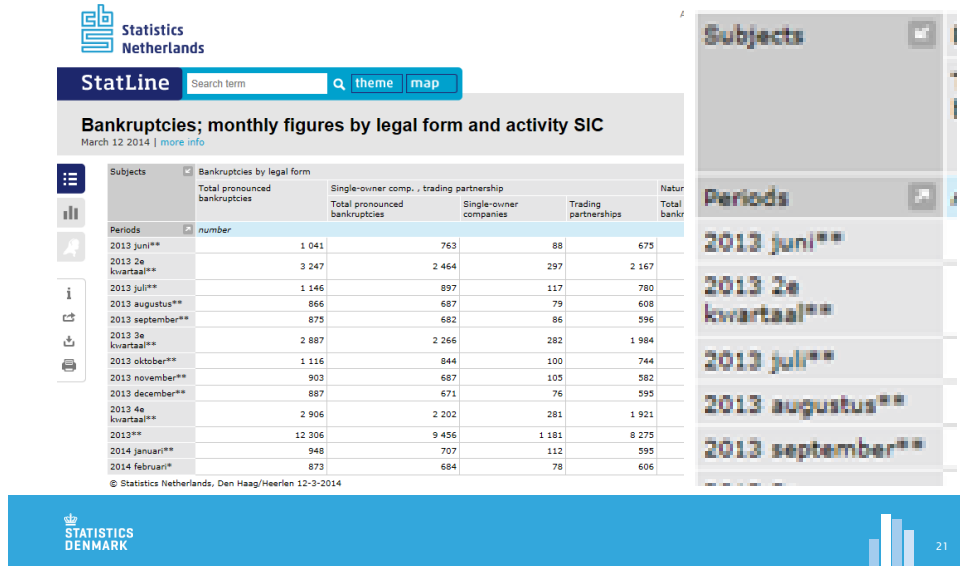
## Key word:

|            |            |            |            |               |
|------------|------------|------------|------------|---------------|
| 1 000 DKK  | 1 000 DKK. | DKK 1 000  | Dkk. 1 000 | Thousand DKK. |
| 1.000 DKK  | 1.000 DKK. | DKK 1.000  | Dkk. 1.000 | Thousand DKK  |
| 1000 DKK   | 1000 DKK.  | DKK 1000   | Dkk. 1000  | Thousand Dkk  |
| 1 000 Dkk  | 1 000 dkk  | DKK. 1 000 | dkk 1 000  | Thousand Dkk. |
| 1.000 Dkk  | 1.000 dkk  | DKK. 1.000 | dkk 1.000  | thousand DKK  |
| 1000 Dkk   | 1000 dkk   | DKK. 1000  | dkk 1000   | thousand DKK. |
| 1 000 Dkk. | 1 000 dkk. | Dkk 1 000  | dkk 1 000  | thousand Dkk  |
| 1.000 Dkk. | 1.000 dkk. | Dkk 1.000  | dkk. 1.000 | thousand Dkk. |
| 1000 Dkk.  | 1000 dkk.  | Dkk 1000   | dkk. 1000  | thousand dkk  |
|            |            |            |            | thousand dkk. |

Standardization



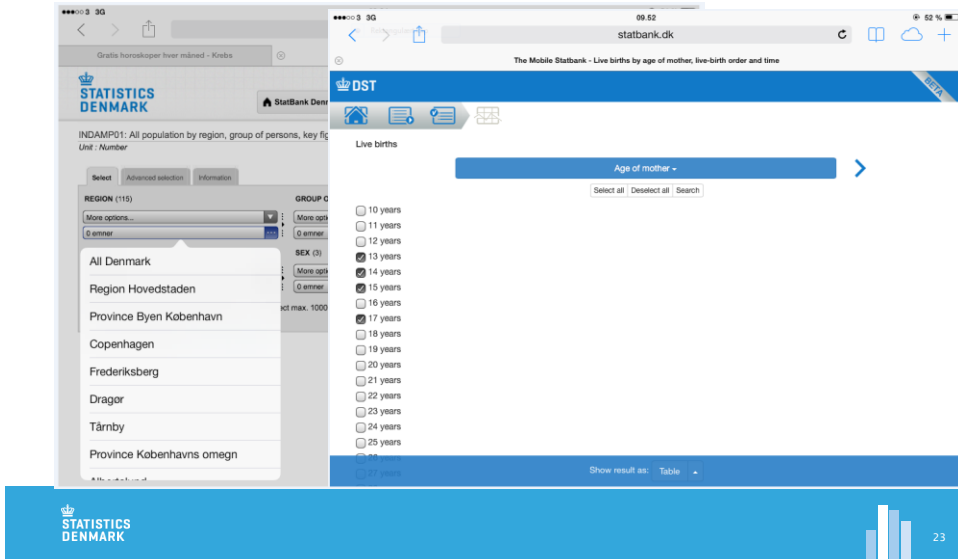
# Centralization



## Interfaces to output databases

- Graphical User Interfaces (GUI's)
  - web and other software user interfaces
  - Platform challenges
- Application Programming Interfaces (API's)
- Web Services (SOAP, Rest)
- E-mail services
- SMS services
- Etc ...

## Devices- I



## Devices II

