

Statistical disclosure control

The (legal) context

Aleksandra Bujnowska, Wim Kloek
Eurostat, Luxembourg

GRUPO DE TRABAJO
DE FORTALECIMIENTO
INSTITUCIONAL
CEA- CEPAL



Coordinado por:  **DANE** INFORMACIÓN
ESTRATÉGICA

4 July 2018

Luxembourg

Contents

- **Types of confidential data/ information**
- **Statistical confidentiality**

Types of confidential data/information

Personal
data

Personal
data
protection

Statistical
data

Statistical
confidentiality

Secret
information

- internal procedures

Personal data

Examples:

- Recruitment: CVs, motivation letters
- Daily work: meetings, contracts

Legal framework:

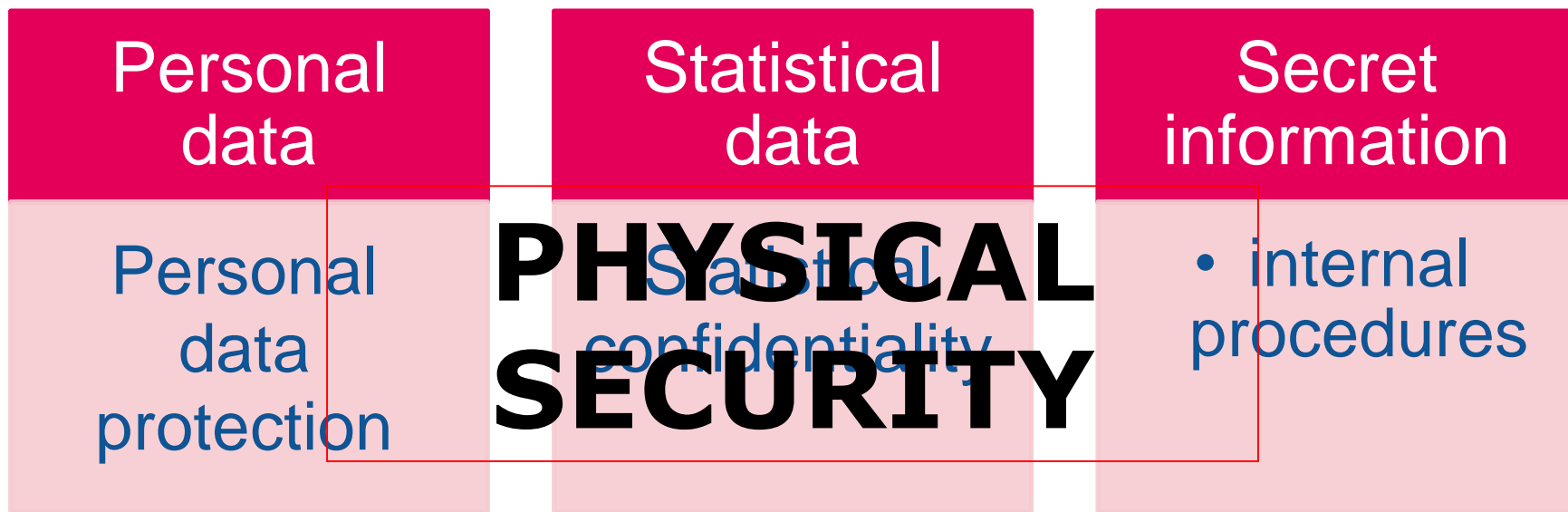
General Data Protection Regulation (GDPR)
valid for all countries in the EU

Secret information

Examples

- Statistical data before their official release
- Commercial and financial information

Types of confidential data/information



Statistical confidentiality in the EU-Treaty

The production of Union statistics shall conform to impartiality, reliability, objectivity, scientific independence, cost-effectiveness **and statistical confidentiality**; it shall not entail excessive burdens on economic operators

Statistical Confidentiality

- European Statistics Code of Practice
- 'Statistical law' – EU Regulation 223/2009

Statistical confidentiality

The **protection of confidential data** related to **single statistical units** which are obtained directly for **statistical purposes** or indirectly from administrative or other sources and implying the prohibition of use for non-statistical purposes of the data obtained and of their unlawful disclosure

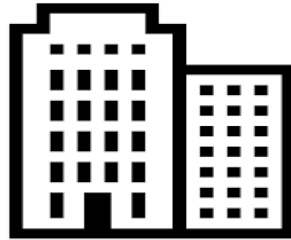


Confidential data: definition

Confidential data means:

data which allow **statistical units** to be **identified**, either **directly or indirectly**, thereby **disclosing** individual information

Statistical unit



Direct and indirect identifiers

Name

Age

Address

Sex

ID number

Nationality



Direct and indirect identifiers

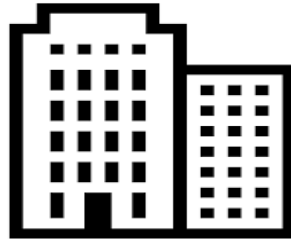
Name

Headquarters

Telephone number Turnover

VAT-number

Branch of activity



Enterprises by economic activity and region

	Region1	Region2	Region3	Total
Ship building	113	304	1	418
Mining	8	2	7	17
Production of wine	1	8	116	125
Total	122	314	124	560

Disclosure

- Reveal individual information
- Identification is not sufficient, as you do not learn anything new about a persons or company
- Not all countries in Europe apply the same rules; for some identification already requires protection

Beware of linked tables

Turnover in ship building by region

	Region1	Region2	Region3	Total
Ship building	877665	223334	77889	1178888
Ship building	877665	x	x	1178888

Statistical disclosure control

- Methods to protect statistical output against disclosure risk
- Aims to maintain the usefulness as much as possible
- Find the right balance between protection and information loss

Statistical output types

- **Tables**
- Microdata
- Graphs
- Models
- Indicators
-

Primary confidentiality

Minimum number of observations per cell

No dominant company:

- (1,90) rule: largest contributes over 90%
- (2,80) rule: largest two contribute over 80%
- p% rule: cell total minus the two largest contributors is less than p% of the largest

Protection methods

- Suppression
- Rounding
- Table redesign (aggregation)
- Perturbative methods (adding noise)

Suppression advantages/disadvantages

Pro: easy to understand, familiar, safe, tools are available

Contra: difficult to analyse the table as a whole, difficult to manage over a set of tables and over time, requires expert judgement

Developments

- Moving away from suppression towards perturbative methods: protection of census 2021, tailor-made tables in social surveys, eventually in all social surveys
- Not for business statistics

Business statistics

Accounting variables and skewed distributions: perturbation would damage patterns in the data

- Agreement to publish (waiver)
- Passive confidentiality (protection only on request of the company)



GRUPO DE TRABAJO
DE FORTALECIMIENTO
INSTITUCIONAL
CEA- CEPAL

Coordinado por:



DANE INFORMACIÓN
ESTRÉGICA