Evaluation Criteria for the Selection of a SDC Method

Johannes Rohde, Christiane Seifert (IT. NRW)
Sarah Giessing (Destatis)

Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality
The Hague, October 29-31, 2019
Background

• Situation in German Official Statistics:
  – One Federal Statistical Office, 14 Federal State Statistical Offices
  – Joint decisions are brought out by working groups related to specific statistical domains
    ➢ This includes decisions regarding SDC

• **Goal:** Developing a document which provides support to the working groups during the selection process of a SDC method
Approach

• Criteria focus on practical, organizational, technical and legal issues

• Identification of interdependencies and conflicting aims

• Argumentation
  – from the perspectives of statistical offices as well as user groups
  – only at a general level

• Criteria are derived from the European Code of Practice (CoP)
Structure

• Structure of the document:
  – Part 1: Definition of twenty evaluation criteria
  – Part 2: Evaluation of four SDC methods along the criteria

• Part 1 consists of four main topics of criteria:
  1. Protection against Disclosure Risks
  2. Data Quality
  3. Accessibility and Clarity
  4. Cost Effectiveness

• Evaluated SDC methods in Part 2:
  – Cell suppression, Deterministic Rounding, Stochastic Perturbation, SAFE (micro aggregation method)
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 1: Protection against Disclosure Risks

1.1 Primary Protection of Confidentiality
1.2 Secondary Protection of Confidentiality
1.3 Harmonized Protection of Confidentiality
1.4 Confidentiality in the Case of Third Party Access

(→ Principle 5, CoP)
Main Topics of Criteria

The twenty criteria are grouped into four main topics:

Topic 1: Protection against Disclosure Risks

1.1 Primary Protection of Confidentiality
1.2 Secondary Protection of Confidentiality
1.3 Harmonized Protection of Confidentiality
→ Integrative SDC concept for the entire publication program within a Statistical System!
1.4 Confidentiality in the Case of Third Party Access
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 1: Protection against Disclosure Risks

1.1 Primary Protection of Confidentiality
1.2 Secondary Protection of Confidentiality
1.3 Harmonized Protection of Confidentiality
1.4 Confidentiality in the Case of Third Party Access

→ SDC must be maintained for publications of third parties!
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 2: Data Quality

2.1 Relevance
2.2 Accuracy
2.3 Timeliness
2.4 Comparability
2.5 Additivity
2.6 Consistency

(➔ Principles 11-14, CoP)
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 2: Data Quality

2.1 Relevance
→ The SDC method should allow that the needs of the data users can be satisfied!

2.2 Accuracy

2.3 Timeliness

2.4 Comparability

2.5 Additivity

2.6 Consistency

(→ Principles 11-14, CoP)
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 2: Data Quality

  2.1 Relevance
  2.2 Accuracy  
  → The information loss induced by a SDC method should be kept at a minimum!
  2.3 Timeliness
  2.4 Comparability
  2.5 Additivity
  2.6 Consistency
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 3: Accessibility and Clarity

3.1 Complexity and Transparency
3.2 Replicability
3.3 Scope of Application
3.4 Flexibility of Publications
3.5 Web Access
3.6 Data Use for Academic Purposes via Research Data Centers

(→ Principle 15, CoP)
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 3: Accessibility and Clarity

3.1 Complexity and Transparency
→ The SDC method should allow accurate interpretations of results and meaningful comparisons!

3.2 Replicability

3.3 Scope of Application

3.4 Flexibility of Publications

3.5 Web Access

3.6 Data Use for Academic Purposes via Research Data Centers
Main Topics of Criteria

• The twenty criteria are grouped into four main topics:

Topic 3: Accessibility and Clarity

3.1 Complexity and Transparency
3.2 Replicability
3.3 Scope of Application
3.4 Flexibility of Publications
→ The SDC method should allow the easy preparation of customer-designed analyses!
3.5 Web Access
3.6 Data Use for Academic Purposes via Research Data Centers

(→ Principle 15, CoP)
Main Topics of Criteria

- The twenty criteria are grouped into four main topics:

**Topic 4: Cost Effectiveness**

4.1 Implementation Costs
4.2 Operating Costs
4.3 Potential for Standardized IT Solutions
4.4 Usability of Modern Information and Communication Technologies on the Data Dissemination

(→ Principle 10, CoP)
Main Topics of Criteria

- The twenty criteria are grouped into four main topics:

**Topic 4: Cost Effectiveness**

4.1 Implementation Costs
4.2 Operating Costs
   - The SDC method should allow the fast conduction of the SDC process!
4.3 Potential for Standardized IT Solutions
4.4 Usability of Modern Information and Communication Technologies on the Data Dissemination

(→ Principle 10, CoP)
Summary

• The entire document “Evaluation Criteria for the Selection of a SDC method” has been developed by a SDC working group

• Detailed evaluation of four SDC methods along the defined criteria

• Mandatory framework for the selection of a SDC method in German Official Statistics since 2018

• Available on request only
Contact details

Johannes Rohde
Tel. (+49)211 9449-2882
Johannes.Rohde@it.nrw.de

Christiane Seifert
Tel. (+49)211 9449-2898
Christiane.Seifert@it.nrw.de

Sarah Giessing
Tel. (+49)611 75-2701
Sarah.Giessing@destatis.de