Gender Statistics
Database quality criteria

Irene Riobóo Lestón
Rey Juan Carlos University

Expert meeting 7 November 2017, Vilna
Without data you are just another person with an opinion

W. E. Deming
1. Introduction
2. Data quality: A general framework
3. Improving the general framework by mainstreaming gender
4. Selection of sources for EIGE’s Gender Statistics Database
QUALITY is a complex, multidimensional concept that measures the fitness of data for their purpose.
EIGE’s focus on data quality:

Data quality assessment
Metadata section
Methodological reports
Report on data quality
Introduction

Report “Quality criteria of the EIGE’s Gender Statistics Database”

- Engage producers of statistics in measuring gender equality progress
- Clarify and share common standards on production and use of gender statistics
- Explain the quality criteria for selecting indicators, statistics, data and metadata
- Describe the technical aspects with which data and metadata should comply with
Data quality: A general framework

The general framework EIGE uses to evaluate data quality is borrowed from the quality assessment and assurance frameworks of the European Statistical System (ESS).

With the adoption of the European Statistics Code of Practice Eurostat and the statistical authorities of the EU MSs have committed themselves to an approach towards high quality statistics.

The Code sets the standard for developing, producing and disseminating European statistics.
### Data quality: A general framework

#### Institutional environment
1. Professional independence
2. Mandate for data collection
3. Adequacy of resources
4. Commitment to quality
5. Statistical confidentiality
6. Impartiality and objectivity

#### Statistical processes
7. Sound methodology
8. Appropriate statistical procedures
9. Non-excessive burden on respondents
10. Cost effectiveness

#### Statistical output
11. Relevance
12. Accuracy and reliability
13. Timeliness and punctuality
14. Coherence and comparability
15. Accessibility and clarity
## Data quality: A general framework

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relevance</strong></td>
<td>Extent to which outputs meet current and potential needs of users.</td>
<td>• User needs</td>
<td>WMID data are the primary source for indicators to monitor the implementation of Area G of the Beijing Platform for Action. The data are therefore widely used by EC and EIGE for analysis in this area and for reporting to the Council of the European Union. The data are also widely used by researchers in this area.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• User satisfaction</td>
<td>No user satisfaction surveys are carried out.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Completeness</td>
<td>Data are complete.</td>
</tr>
</tbody>
</table>
# Data quality: A general framework

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy and reliability</strong></td>
<td>Accuracy is the closeness of estimates and computations to the exact or true values. Reliability is the closeness of initial estimated values to subsequent estimated values.</td>
<td>• Overall accuracy</td>
<td>Data can be considered fully accurate to describe the situation for the area of decision-making concerned,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sampling error</td>
<td>Not applicable.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Non-sampling error (coverage error, measurement error, non-response error, processing error)</td>
<td>Not applicable.</td>
</tr>
</tbody>
</table>
### Data quality: A general framework

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timeliness and punctuality</strong></td>
<td>Timeliness reflects the length of time between data availability and the event or phenomenon the data describe.</td>
<td>• Timeliness</td>
<td>Data are released within one month of collection.</td>
</tr>
<tr>
<td></td>
<td>Punctuality refers to the time lag between actual and scheduled data release dates.</td>
<td>• Punctuality</td>
<td>Punctuality is 100%.</td>
</tr>
<tr>
<td>Dimension</td>
<td>Definition</td>
<td>Sub-dimension</td>
<td>Women and men in the European Parliament</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Coherence and comparability</strong></td>
<td>Coherence refers both to the extent to which a dataset is internally consistent and to the degree to which different datasets can be reconciled and combined.</td>
<td>• Cross-domain</td>
<td>The official website of the institution publishes the distribution of women and men in the opening sessions of each parliamentary term. However, data are not presented in absolute numbers (only the share by gender are shown), and do not reflect the variations between elections nor the current composition.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Internal</td>
<td>It is ensured through careful application of the WMID methodology, and routine data validation. Additionally, data collected goes through a verification process done by a senior researcher.</td>
</tr>
</tbody>
</table>
## Data quality: A general framework

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence and comparability</td>
<td>Comparability measures the impact of differences in applied statistical concepts, definitions, and methodologies when statistics are compared between geographical areas, non-geographical dimensions, or over time.</td>
<td>• Geographical</td>
<td>Note that seats are allocated based on the population of each Member State and therefore the number of members by country will vary. Apart from this point, data are comparable between countries.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Over time</td>
<td>In general, WMID data are comparable through time in each area of decision-making. In the case of the European Parliament, there are changes in the number of members of the European Parliament (MEPs) due to EU enlargement.</td>
</tr>
</tbody>
</table>
## Data quality: A general framework

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accessibility and clarity</strong></td>
<td>Accessibility if the ease and conditions under which statistical information can be obtained. Clarity assesses whether data are accompanied by appropriate metadata to the extent necessary for full understanding of the data.</td>
<td>• News release</td>
<td>No regular news release.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Publications</td>
<td>From 2017, EIGE will publish general documents on gender statistics. The EC’s Annual Report on Equality usually includes a section on this topic.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Online Database</td>
<td>EIGE’s Gender Statistics Database</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Microdata</td>
<td>Not made available</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Other</td>
<td>WMID data are the primary source of information for indicators to monitor the implementation of Area G of the Beijing Platform for Action.</td>
</tr>
</tbody>
</table>
### Dimension: Accessibility and Clarity

<table>
<thead>
<tr>
<th>Definition</th>
<th>Sub-dimension</th>
<th>Women and men in the European Parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Documentation on methodology</strong></td>
<td><strong>WMID Methodology published on EIGE's website.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Quality documentation</strong></td>
<td><strong>Not applicable.</strong></td>
<td></td>
</tr>
</tbody>
</table>
Improving the general framework by mainstreaming gender

Gender statistics:

- Based on concepts and definitions that reflect the diversity of women and men and capture all aspects of their lives, reflecting their conditions, roles and contributions in society, their needs, their specific problems and inequalities.

- Collected taking into account stereotypes and social and cultural factors that may induce gender biases.

- Collected and presented disaggregated by sex as a primary and overall classification.
Potential uses for Gender Statistics:

- Promote understanding of the actual situation of women and men in society, and their differences
- Advance gender analysis and research
- Monitor progress towards gender equality
- Develop and monitor policies and programmes
- Support gender mainstreaming in development and poverty reduction policies
- Facilitate decision-making
Improving the general framework by mainstreaming gender

Gender sensitive indicators:

- Represent statistical data for a specified time, place, and other characteristics.

- Corrected for at least one dimension (usually size) in order to be expressed in units that allow for meaningful comparisons between populations across space and time.

- Have a normative nature, in the sense that a change from an identified reference point (a norm or a benchmark) in a particular direction can be interpreted as ‘good’ or ‘bad’.
Improving the general framework by mainstreaming gender

Gender sensitive data production:

The overarching general principle in the production of high quality gender-sensitive data is that of gender mainstreaming.

Gender issues and gender-based biases are systematically taken into account in the production of all official statistics and at all stages of data production.
Improving the general framework by mainstreaming gender equality

Selection of research topics

- Decision-making power
- Economic opportunities
- Work-life balance
- Gender-based violence
- Gender stereotypes
- Reproductive health
- Intra-household resources
Improving the general framework by mainstreaming gender

Definitions, concepts, and research methods

- Avoid hidden inequalities
- Avoid gender biases
- Ensure coherence and comparability

Ensure coherence and comparability

Avoid hidden inequalities

Avoid gender biases
Administrative register design

- Revise concepts and definitions used.
- Application of unified identification systems across different sources based on the use of common identification codes.
- High level of cooperation among the administrative authorities is needed.
Improving the general framework by mainstreaming gender

Survey design

- Appropriate choice of sampling and analysis units
- Selecting sampling frames in ways that does not cause gender-biased
- Selecting sufficient sample sizes to enable inference about subgroups
- Selecting the proper interview mode
- Questionnaire simple and unambiguous
- Use of neutral language
- Avoid leading questions
- Control by international differences in norms and expectations
- Training interviewers
- Appropriate data processing
Presentation of data and metadata

- Data should be disaggregated by sex and often also by additional characteristics
- Microdata should be anonymised and properly labelled
- All data should be accompanied by appropriate metadata
Improving the general framework by mainstreaming gender

Good data tells a story...
Selection of sources for EIGE’s Gender Statistics Database

- Disaggregation by sex and other relevant characteristics.
- Metadata analysis.
- Analysis of the response rates.
History is herstory too.

Author unknown
Thank you

www.eige.europa.eu

EuroGender Network

https://twitter.com/eurogender

facebook.com/eige.europa.eu

youtube.com/user/eurogender

eige.europa.eu/newsletter