Homicide Data Quality Assessment Framework

Joint Second Meeting of
UN-CTS Focal Points and ICCS Technical Advisory Group
Lima, 8 June 2018

Enrico Bisogno
Enrico.Bisogno@un.org
Chief, Data Development and Dissemination Unit
Research and Trends Analysis Branch
UNODC Homicide Data

- Collection grounded on national data
- Transparency
- Stimuli to improve
- Monitoring (SDG 16.1.1) Focus on Trends
UNODC Approach in Developing Homicide Data

Collection
- Data collection ← National Focal Points
  - Enriching data through open sources (national data)

Processing
- Data validation ↔ Adjustments
  - Estimating regional trends
    - Calculating uncertainty ranges
  - Estimating global trend

Management
- Monitoring (SDG 16.1.1)
  - UNODC Data Portal
  - Global Study on Homicide
UNODC Approach in Developing Homicide Data

**Collection**
- Data collection ← National Focal Points
- Enriching data through open sources (national data)

**Processing**
- Data validation ↔ Adjustments
- Estimating regional trends
  - Calculating uncertainty ranges
- Estimating global trend

**Dissemination**
- Monitoring (SDG 16.1.1)
- UNODC Data Portal
- Global Study on Homicide

**Management**
- Quality
What is the need for a framework to assess the quality of UNODC homicide data?

Challenge: The quality of homicide data varies between data sources, countries, and years ...

→ Setting minimum quality requirements while making as much use as possible of data submitted by MSs

→ Inform data users and general public about the quality of data published by UNODC

→ Bring issue of data quality to the attention of MSs and incentivise quality improvements

→ Guidance to improve data quality as an alternative to the blurring of quality problems (e.g. by including model-based estimates)
Homicide Data Quality Assessment Framework

- Country-based
- Conducted for published data that conforms to minimum quality standards
- Combines 3 quality dimensions
- Produces overall quality scores ranging from 0 to 100
- May be transposed to broad quality categories for easy dissemination (e.g. traffic light rating)

1. Validity
   - 2.1 Time series
   - 2.2 Timeliness
   - 2.3 Disaggregate variables (homicide)
   - 2.4 CTS variables other than homicide

2. Completeness
   - 3.1 Internal consistency
   - 3.2 External consistency
The Data Quality Dimensions in Detail (1/3)

1. Validity

- Compliance with ICCS definitions
- Full geographic and institutional coverage

→ Scoring according to CTS metadata categories
2. Completeness

Component 2.1  
Time series  
• Counting values of variable “total homicide” (1990-2016)

Component 2.2  
Timeliness  
• Weighted counting of values of variable “total homicide” (2012-2016)

Component 2.3  
Disaggregate homicide variables  
• Counting values for disaggregate homicide variables (2012-2016)  
→ Homicide by sex, age, mechanism, situational context

Component 2.4  
Variables other than homicide  
Counting values of variables other than homicide (2012-2016)  
→ Data on perpetrators and criminal justice (CTS)
The Data Quality Dimensions in Detail (3/3)

3. Consistency

Component 3.1  
**Internal consistency**
- Calculating average percentage differences for:
  Total homicide ↔ Homicide by sex, age, mechanism, situational context

Component 3.2  
**External consistency**
- Calculating average percentage differences for:
  Criminal justice homicide data ↔ Public health homicide data (e.g. WHO)
### Example Output (1/4): Scoring ↔ Rating

<table>
<thead>
<tr>
<th>Country</th>
<th>Region</th>
<th>Count.Values</th>
<th>Score.Relative</th>
<th>Rating.All</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXXXXXX</td>
<td>Asia</td>
<td>4</td>
<td>0.740740741</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Europe</td>
<td>313</td>
<td>57.962962962</td>
<td>Yellow</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Africa</td>
<td>85</td>
<td>15.74074074</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Oceania</td>
<td>16</td>
<td>2.962962962</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Europe</td>
<td>144</td>
<td>26.666666667</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Africa</td>
<td>2</td>
<td>0.37037037</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Americas</td>
<td>95</td>
<td>17.59259259</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Americas</td>
<td>42</td>
<td>7.777777777</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Americas</td>
<td>38</td>
<td>7.037037037</td>
<td>Red</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Asia</td>
<td>383</td>
<td>70.92592593</td>
<td>Green</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Americas</td>
<td>245</td>
<td>45.37037037</td>
<td>Yellow</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Oceania</td>
<td>378</td>
<td>70</td>
<td>Green</td>
</tr>
<tr>
<td>XXXXXXX</td>
<td>Europe</td>
<td>296</td>
<td>54.81481481</td>
<td>Yellow</td>
</tr>
</tbody>
</table>
Example Output (2/4)

Traffic Light Rating - 'Dimension 2. Completeness/Component 2.1 Time Series'

- 35% Red
- 28% Green
- 37% Yellow

Percentage of countries
Example Output (3/4)

Traffic Light Rating (Regional) - 'Dimension 2. Completeness/Component 2.1 Time Series'

<table>
<thead>
<tr>
<th>Region</th>
<th>Countries/Entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>30</td>
</tr>
<tr>
<td>Americas</td>
<td>15</td>
</tr>
<tr>
<td>Asia</td>
<td>20</td>
</tr>
<tr>
<td>Europe</td>
<td>10</td>
</tr>
<tr>
<td>Oceania</td>
<td>5</td>
</tr>
</tbody>
</table>
Example Output (4/4)

Aggregate Homicide Rate Estimate + Uncertainty Range
Next steps

• Sharing data quality scores with MSs as part of the prepublication process