Balancing Openness Minimizing Risks

Sri Lanka’s Experience

Anoja Seneviratne
Director
Department of Census and Statistics
SRI LANKA
Outline of the Presentation

- Legal base for data collection
- National Statistics System in Sri Lanka
- Data Dissemination Policy in Sri Lanka and terms and conditions of the policy
- Challenges of Open Data for NSO
- suggestions
Legal Base for the Data Collection

- Data collection is done under the **Census Ordinance and Statistical Ordinance**

- According to the Ordinances, any publication cannot be prepared disclosing any particulars relating to the identification of individuals or businesses

- Data collected under these Ordinances can be used for statistical purposes only, not allowed to use for legal purposes
In Sri Lanka, What we have is a National Statistical System (NSS)

The NSS consists of all statistics-producing entities functioning under the direct responsibility of the government of a country
Data Dissemination Policy in Sri Lanka

Three types of anonymizations have been defined

- **Less geographic and less characteristic details**
  - Geographic details and data dictionary are hidden
  - Geography has been collapsed and variable detail has been reduced

- **Higher level of geographic details**
  - Data dictionary is provided
  - Less geographic details and more variable details

- **Full or more geographic details data dictionary and variable details**
  - Generally only available through data enclaves
Terms and Conditions of the Policy

• Data user have to signed an agreement with department
• Data should be used only for the specific study/analysis mentioned in the agreement
• A nominal cost is associated with each micro dataset except for Government agencies, students engaged in higher education and some international organizations
Challenges of Open Data for NSO

- There is no guarantee that data are used for statistical purposes only
- Some data cannot be opened due to security reasons
- Privacy of individuals, households and businesses should be safeguarded
Suggestions

- Data should be opened with anonymization techniques
- Re-use of statistical data must be addressed very carefully
- NSO should defined the disaggregation levels for release open data