The challenge of embracing the potential of multiple data sources for 2030 agenda

UN Friday seminar on emerging issues
4 Mar 2016, New York

Mariana Kotzeva
Deputy Director General, Eurostat
Drivers for change

Data revolution
- New data sources
- New statistical and analytical methods and tools
- Competition among data providers
- New users' needs

Official Statistics
- Data collection: New & multi-mode access to multiple new sources
- Data Integration: Combining surveys & model based & algorithm based statistics
- Data Analytics: New inference methods & visualisation tools
- Dissemination: From data as a product to data as a service
Using Multiple Data Sources for monitoring of 2030 agenda

- Geodata
- Administrative Data
- Frames / Registers
- Surveys
- Big Data
Challenges in using multiple and new data sources in 2030 Agenda

1. Specific nature and challenges of:
   • Big data;
   • Administrative sources
   • Business registers

2. Integration of multiple data sources for building registers and producing statistics

3. Concrete steps towards using multiple data sources
Meeting challenges: examples from European Statistical System (ESS)

ESS Vision 2020

Key areas

Focus on users

Strive for quality

Harness new data sources

Promote efficiency in production processes

Improve dissemination and communication
Specific nature and challenges of big data

- One source could be used in various domains;
- Complement existing statistics and improve their quality rather than replace them;
- Conditions for using big data equally important to conducting pilots
ESS Big Data Action plan

Policy
Legislation
Ethics / Communication
Quality
Skills
Methodology
IT Infrastructure
Experience sharing

Looking for a job?
Lesson learnt from ESS experience with administrative data

- Legal basis providing access to the administrative sources is a necessary but not sufficient condition;
- NSI involvement in establishment and changes of administrative sources;
- Strong political support and good permanent collaboration with data owners.
- Availability of unique identifier allowing for data linking
Administrative data: Emerging topics

- Coverage control of administrative data before use;
- Improving sample survey frames and design;
- Privacy and data protection issues associated with the use of administrative sources;
- Challenges to the comparability across countries;
- Measuring quality of statistical outputs based on multiple sources
ESS project on using Administrative Data

Challenges:
- Access to data
- Influencing quality and design of admin data
- Quality of frames
- Quality of integrated outputs
- Integration methods
Sources to establish and update business registers in the ESS

- Surveys
- Administrative Registers
- Big Data and other Sources

Profiling

Business Registers
European Network of Business Registers

- EGR = Euro group register of multinational enterprises;
- National BR = backbone of business statistics
Successful ESS projects linking data from business registers with other micro data

- International trade of goods by enterprise characteristics since 2009;
- In 2016, statistics on export performance of young (and/or small) enterprises;
- International trade of services by enterprise characteristics

A unique identifier a precondition for a successful linking
Integration of various sources

Domain STATISTICS (Tourism)

- Mobile phone data
- Payment cards data
- Web (prices)
- Bookings (nowcast forecast)
- Households & Business Surveys
- Administrative data

Eurostat
Integrated Business Statistics (FRIBS)
Improve our capabilities for statistical matching and data fusion
**Smart Statistics: the emerging new data sources**

- *Massive increase of data production and exchange*
  - Not manageable by humans
  - Exchange of data without human intervention
  - **Autonomous Agents**
    - Filtering, aggregation
    - Decisions
  - **Cognitive systems**

- *Can we use this new internet for producing statistics in an intelligent way?*
Conclusions – How to make it happen

• **Vision**  
e.g. ESS Vision: high level collective and individual commitment

• **Action**  
e.g. in the ESS: enabling projects and frameworks

• **Prepare for the next wave**  
e.g. Smart Statistics

• **Cannot be done by statistical community alone**  
  Partnerships, strong political support and investment
Thank you for your attention