How Big Data Can Ensure Overall Data Quality

Fill data gaps: supplement statistics where traditional sources are not sufficient.

Gain efficiencies: cheaper and faster to produce, allowing NSOs to achieve more with less, while reducing the response burden.

Enhance data accuracy: Produce real-time insights for faster evidence-based decision making.

Remaining relevant: Engage in partnerships with a variety of stakeholders such as the private sector, academia and civil society.

Future Opportunities

National Statistical Offices worldwide have identified the need to address the following challenges in incorporating Big Data into their business processes.

- Skills and training for Big Data
- Quality frameworks
- Access to Big Data
- Identifying funding sources
- Collaboration with academia and the private sector

Big Data are very large data sets that are characterized by three Vs: velocity, variety and volume.

Resources

If you want to learn more, please contact the United Nation Statistics Division at **bigdata@un.org** or visit **http://unstats.un.org/unsd/trade/bigdata**

Download the PDF version of this brochure

Big Data for Official Statistics

BigData
UN Global Working Group

The United Nations Global Working Group on Big Data for Official Statistics (GWG)

The GWG brings together the global community of official statisticians to participate, contribute and share knowledge and experience in using Big Data to produce official statistics. This working group aims to be a resource for producers and users of statistics from the government, private sector, academia, civil society and other organizations.

Our Mission

The GWG **prepares case studies, methods, tools, resources, as well as conducts pilot projects and surveys** about using Big Data for official statistics.

Our work is organized in **8 task teams**:

- Advocacy and communication
- Using Big Data for the Sustainable Development Goals (SDGs)
- Data access and partnerships
- Training, skills and capacity building
- Cross-cutting issues
- Mobile phone data
- Satellite imagery and remote sensing data
- Social media and web-scraping data

Why We Need Big Data

Policy makers, the media and society as a whole expect and demand better, faster and more detailed statistics.

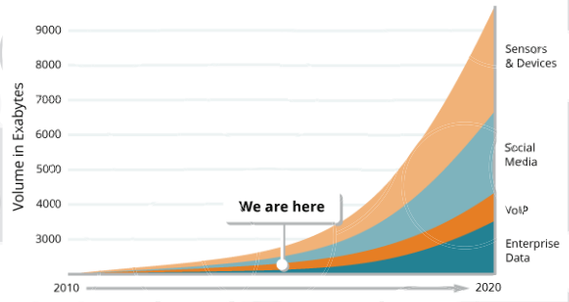
The 2030 Agenda for Sustainable Development insists on accountability and no one left behind.

Emerging issues are difficult to measure with traditional data sources. Given these realities, the statistical community must modernize to stay relevant.

The Sustainable Development Goals have highlighted opportunities to use big data sources and techniques to fill data gaps, offer insights with higher frequency and resolution, and to use technology in new ways to measure the world around us.

Big Data Sources

Big Data comes from a variety of sources: mobile and smart phones, web communications and transactions, automated banking and other automated services, scanner data, satellite imagery and other geodata, social media data, smart meters and the other sources.



Source: IBM

Examples of Big Data Usage for Official Statistics

Mobile phone data

- **Understanding population movement,** migration and tourism using location data
- **Mapping poverty** based on the behavioural patterns of mobile phone users and airtime credit purchases
- **Providing real-time population statistics** and population movements

Satellite imagery and remote sensing data

- **Measuring crop production and harvests,** as well as estimating the rate of land consumption
- **Measuring and mapping land cover,** land use, the condition of ecosystems and natural disaster impacts
- **Assessing road traffic** using road sensors

Social media data

- **Using social media** messages to extract information on populations
- **Capturing real-time food prices** using Twitter messages
- **Surveying labour market dynamics** through Google searches and trends

Other data

- **Estimating the job vacancies** based on information from job portals
- **Generating price statistics** by web scraping of online retailers
- **Using supermarket transaction data** for price indices and retail trade statistics