

Data sources for demographic and social statistics

Training workshop
on data science and big data for official statistics and the SDGs
6-8 November 2023



UNITED NATIONS

الاسواق
ESCWA

Shared Prosperity **Dignified Life**



Content

1. Introduction
2. Different data sources for demographic and social statistics: pros and cons
 - a. Traditional
 - b. Non traditional
3. Conclusion remarks

1. Introduction

- The need for greater coverage, timeliness and disaggregation of the indicators to fit SDG data requirements implies reconsidering the ways statistics organizations work and plan for the future.
- The COVID-19 crisis stressed the need for alternative methods of data collection as the pandemic disrupted traditional operations undertaken to collect and compile basic statistics (censuses, surveys, etc...)
- The need for data from open sources: Conflict and humanitarian settings.
- The emerging of new technologies in data collection, analysis and dissemination

2. Different data sources for demographic and social statistics: Pros and cons

A. Traditional data sources include: Censuses, surveys, population registers, vital statistics registries and administrative records.

Pros	Cons
Reliability: Data from official sources like national statistics offices is considered highly reliable	Costly and Time-consuming: Traditional methods such as face-to-face interviews or paper-based surveys require significant resources and time
Standardization: These data usually follow international standards, making it easier to compare across countries and time	Inflexibility: Once a survey or census is designed, it is difficult to adapt or change it quickly to capture new trends
Longitudinal Data: Traditional sources often provide longitudinal data, which is valuable for understanding trends and changes over time	Coverage and disaggregation:

2. Different data sources for demographic and social statistics: Pros and cons(cont'd)

Traditional data sources include: censuses, surveys, registries (population, vital statistics,...) and administrative records.

Pros	Cons
Representativeness: Traditional surveys are designed to be representative of the entire population	Response Burden: Respondents may suffer from survey fatigue, especially when surveys are lengthy or frequent
Legal Backing: The collection of official statistics is often supported by a legal framework ensuring compliance and data privacy	Lag in Availability: There can be a considerable time lag between data collection and availability, which can make the data less relevant

2. Different data sources for demographic and social statistics: Pros and cons (cont'd)

B. Non-traditional data sources include: geospatial, bigdata, mobile phone, web script, social media, etc...

- **Geospatial information:** in all stages of statistical process, census, population estimates in humanitarian setting, 28 SDGs indicators, addresses

- **Big data: Big Data:**

1. Mobile Phone Records: Call Detail Records (CDRs) and mobile phone usage patterns offer insights into human behavior and mobility.

2. Internet Data: Browsing patterns, search queries, and online transactions provide data on consumer behavior and social trends.

3. Social Media: Data from platforms like Twitter, Facebook, and Instagram can reveal trends in public opinion, real-time population sentiment, and social dynamics.

Public Cameras and Surveillance Footage:

•With the appropriate privacy safeguards, video data can be analyzed for crowd counting, traffic flow, and urban planning.

2. Different data sources for demographic and social statistics: Pros and cons (cont'd)

❖ **Opportunities that non-traditional data sources provide :**

- improved timeliness and accuracy, time and cost effectiveness, detailed data, increased disaggregation capabilities, real-time insights

❖ Countries have identified, **the following issues in using such data sources**, in order of importance: (as cited in their answers to a questionnaire sent by ESCWA in 2020)

- (a) inadequate legal framework, and limited access to data;
- (b) human resources not skilled enough, and high costs of accessing the data; and
- (c) lack of technological tools, and statisticians' perception of Big Data

2. Different data sources for demographic and social statistics: Pros and cons (cont'd)

Challenges	Possible solutions
<p>Coverage and Representativeness: Non-traditional sources may not be representative of the broader population, as they often don't cover all the population and over-represent certain demographic groups</p>	<p>Stratification and Weighting: Employ statistical techniques to correct for biases and ensure that the data is representative of the broader population.</p> <p>Combining Data Sources: Integrate non-traditional data with traditional datasets to complement and enhance representativeness.</p>
<p>Data Quality: The quality and accuracy of the data can be variable, and the data is not collected for statistical purposes</p>	<p>Validation and Verification: Use statistical techniques and cross-referencing with traditional data sources to validate the accuracy and reliability of non-traditional data.</p> <p>Standardization of Methods: Work towards standardizing data collection and processing methods to ensure consistency and reliability.</p>

2. Different data sources for demographic and social statistics: Pros and cons (cont'd)

Challenges	Possible solutions
Data Ownership and Access - Get data from sources	Negotiating Access: Work with private companies to negotiate access to data, possibly through partnerships or data-sharing agreements. Open Data Initiatives: Support and participate in open data initiatives that promote the sharing of non-traditional data for public benefit.
Inadequate Legal and Ethical Issues: There might be legal and ethical limitations on accessing and using certain types of non-traditional data	Stay Informed on Regulations: Keep updated on the latest regulations concerning data privacy and adhere to them strictly. Legal Expertise: Engage legal experts to navigate the complex landscape of data rights and regulations.
Human resources/expertise: Staff need to be trained on	Capacity building

3. Conclusion remarks

- Non-traditional data sources are becoming increasingly important for enhancing and complementing traditional statistical methods. When combined, they can provide a more comprehensive and nuanced picture of societal trends.
- Statisticians and researchers must try to overcome the challenges faced and ensure that the information they provide is both accurate and ethical.



Shared Prosperity **Dignified Life**



Thank you