A Diagnostic Tool for Administrative Registers

11th Meeting of the Steering Committee of the Arab Statistics Initiative (Arabstat)

Abu Dhabi – November 13-14, 2024

THE WORLD BANK

IBRD • IDA | WORLD BANK GROUP

Motivation

- Administrative registers are primarily used in the production of goods and services in public or private institutions or companies or that the information is a result of such production.

- Statistical registers are created by processing data from administrative registers and are a rich source of data that can be exploited for statistical analysis as a complement or in substitution of data from censuses and surveys.

- However, the quality of administrative data must be verified and possibly enhanced before using them for statistical analysis.

- For example, administrative data on income at individual and household level can be used to produce a measure of disposable income (Canberra Handbook) as a potential input to the compilation of statistics on living standards, for M&E of social programs, for national accounts.

- The World Bank developed a diagnostic tool to diagnose the quality, identify weaknesses and strengths, and design an action plan to strengthen the quality of administrative records.

Administrative Data

❑ Perhaps the most comprehensive of the traditional definitions were set out by Gordon Brackstone of Statistics Canada's (1987) paper "Statistical Issues of Administrative Data: Issues and Challenges."

❑ Four key distinguishing features of administrative data:

The agent that supplies the data to the statistical agency and the unit to which the data relate are different (in contrast to most statistical surveys);
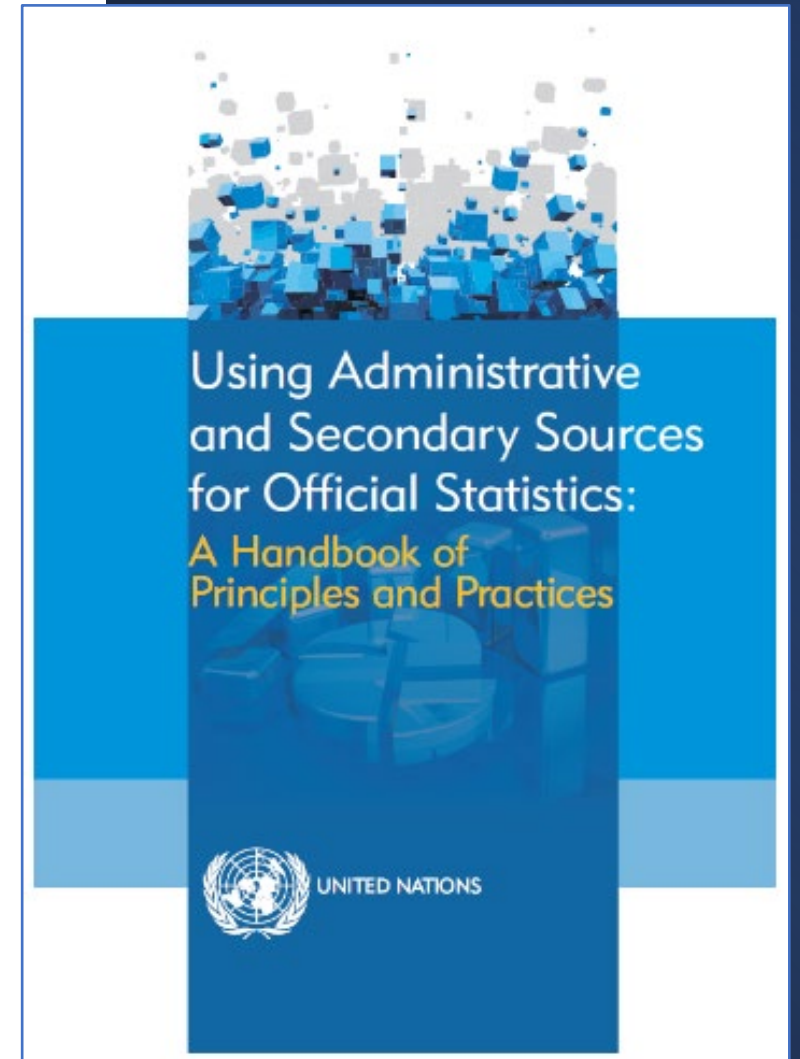
The data were originally collected for a definite non-statistical purpose that might affect the treatment of the source unit;

Complete coverage of the target population is the aim;

Control of the methods by which the administrative data are collected and processed rests with the administrative agency.

Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices

UNITED NATIONS

Administrative and Statistical Registers

- ❑ **Administrative registers** are registers primarily used in administrative information systems. That means that the registers are used in the production of goods and services in public or private institutions or companies or that the information is a result of such production.

- ❑ As part of a register-based statistical system, administrative registers are sometimes called **primary registers**.

- ❑ **Administrative base registers** are kept as a basic resource for public administration. The function of these registers is typically to keep stock of the population, i.e., to record the birth and death of units and to keep track of what units are present at any given time in the population. In addition, base registers have to maintain identification information to be used by other administrative registers. (you can mentioned here the relevance of the unit of analysis as the key to organize base registries such as individuals, firms, holdings services, and dwellings)

- ❑ **Statistical registers** are created by processing data from administrative registers. Statistical registers could be based on a single administrative register, but they are more frequently based on combined data from several administrative sources.

- ❑ Statistical registers are also referred to as **secondary registers**.

- ❑ **Statistical base registers** are based on the corresponding administrative registers. These are registers of great importance for the whole register system. Their principal tasks are to define important populations and contain links to other base registers.

Characteristics

- Censuses and surveys are produced for statistical purposes and on formally established conceptual and methodological foundations, and they follow international standards.

- ARs are not necessarily designed or produced for statistical purposes. Some are consolidated by national statistical offices as vital statistics (ARs of births, deaths, marriages).

- Although there are international standards on ARs, such data frequently lack definitions, classifications, and identifiers, have missing data, and are fragmented. This limits their statistical use.

ARs: A Cost-Effective Strategy for the production of statistics

- Provide relevant information related to various topics: demographic, social, economic, environmental

- Low production cost, if already produced

- No sampling errors

- Comprehensive coverage

- Useful geographical breakdown for monitoring and evaluation programs and policies

- Allow integration with other sources of information

- Support the continuity of statistics

- Constitute a sampling frame for surveys (for example, directories of schools, health clinics)

AN OVERVIEW OF THE WORLD BANK DIAGNOSTIC TOOL

Objectives

- To enable staff responsible for collecting and managing ARs and staff who statistically exploit ARs to diagnose data quality.

- The application of the tool is designed to rely on international best practices in evaluating administrative records and in a comprehensive analysis of data management, data quality, and data use.

- The tool is applied based on an assisted self-evaluation method through a series of questions reflecting a high degree of objectivity.

- The full version includes over 130 questions; a mini version is available with 28 questions aimed at tailoring the assessment based on attributes of agency (data producer, data user, or both).

- The assessment takes place in workshops in which public officials in entities responsible for administrative registers act as facilitators in discussion groups.

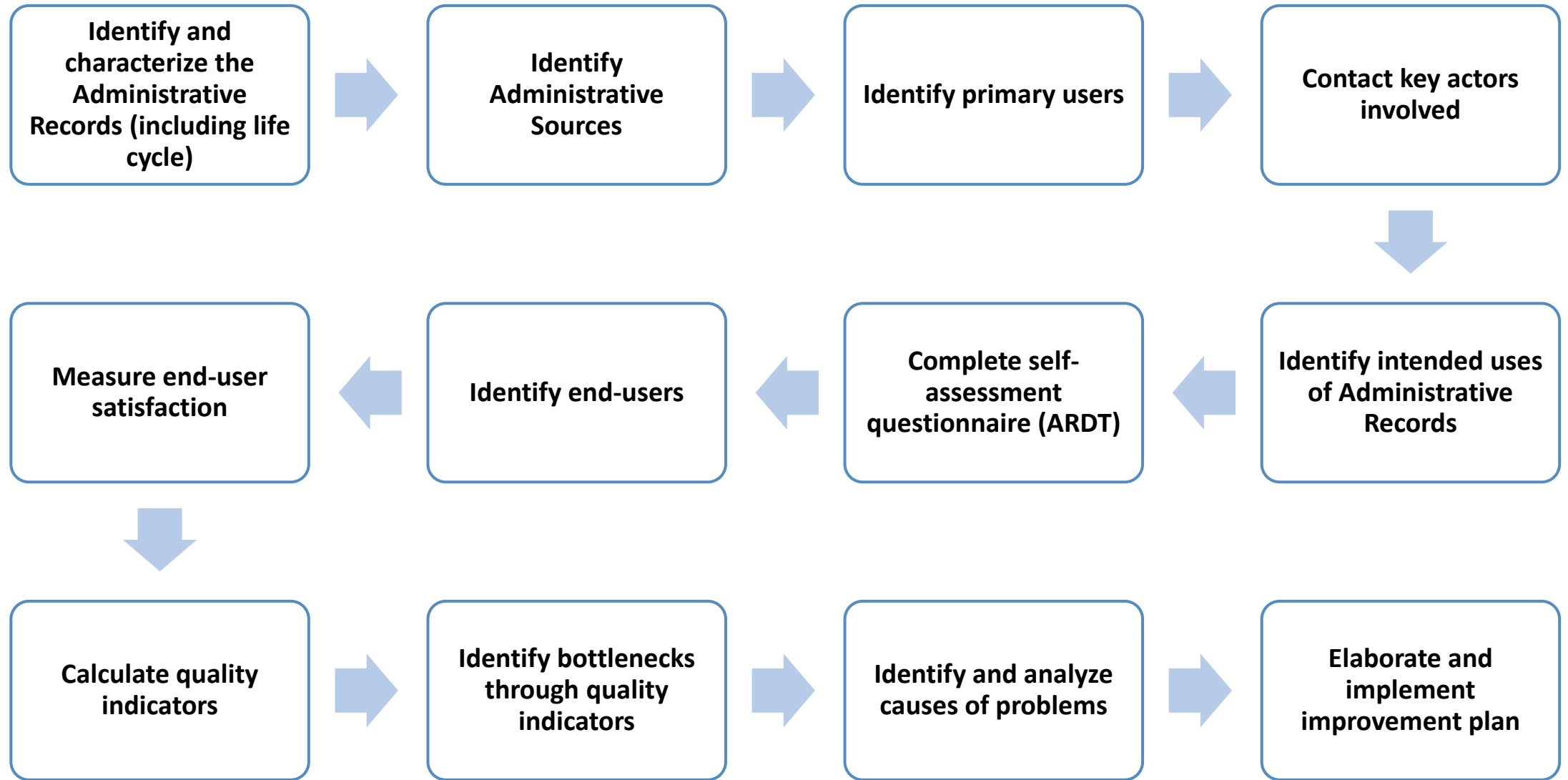# Diagnostic Domains(1/2)
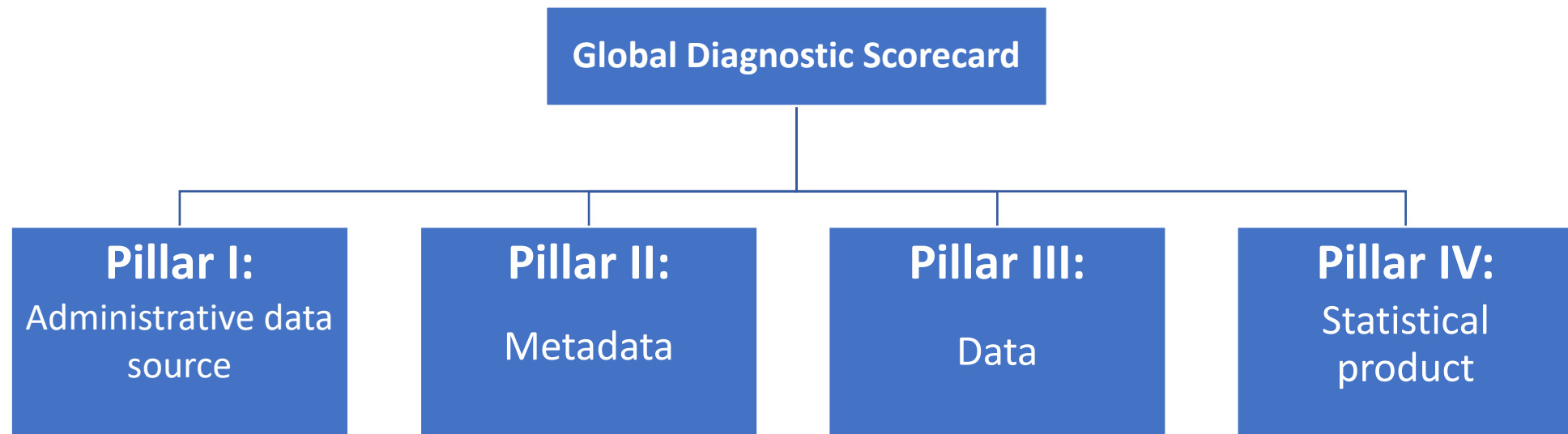
Diagnostic Domains(2/2)

PILLAR I

a diagnosis is carried out of the attributes associated with the source of the administrative data, including the institutional environment, the legal framework to ensure information security, and the procedures for data delivery and data processing.

❑ 50 closed-end questions.

Scoring System

➤ Each pillar comprises several attributes, and each attribute is broken down into specific indicators.

➤ The value of each indicator is obtained based on responses to a set of closed questions.

➤ The responses and indicators are captured in an Excel file that also contains graphs to facilitate the visualization of critical quality points and identify opportunities for improvement.

➤ The overall score summarizes the value of each module. A weighting scheme assigns weights to specific attributes within each module.

➤ The global score ranges from 0 to 5, whereby 5 is the highest value.

Implementation Process

Identify and characterize the Administrative Records (including life cycle) → Identify Administrative Sources → Identify primary users → Contact key actors involved

Measure end-user satisfaction ← Identify end-users ← Complete self-assessment questionnaire (ARDT) ← Identify intended uses of Administrative Records

Calculate quality indicators → Identify bottlenecks through quality indicators → Identify and analyze causes of problems → Elaborate and implement improvement plan

Expected Results

Identify main bottlenecks that affect the quality of the ARs and its statistical products (indicators).

Examples of Recommendations

1. Reorganize original datasets (ARs fragmented in pieces, and at best arrangements of columns and rows)
2. Redesign data collections formats (adding or adjusting variables as needed)
3. Reorganize data collection and processing flows, including legal time, responsible
4. Include standard national or international classifiers (e.g. geographical codes, education levels classifiers, etc.)
5. Include unique identifiers that allow merging one AR with similar datasets
6. Improve data management (security, preservation, integration)
7. Move from paper to digital data collection
8. Improve documentation (metadata)
9. Etc.

Thank you