DEPARTMENT OF STATISTICS SINGAPORE

Session 1: Introduction to Data Integration

Country presentation: Singapore

Regional Workshop on Data Integration in Asia and the Pacific



About Singapore

700 sq km

5.6 million



Singapore's name is derived from 'Singa Pura' which means "Lion City" due to a sighting of an animal which looked like a lion.



Data integration in Singapore

- Singapore has been practicing data integration since the 1990s.
- In the domain of demographic statistics, two core databases have been set up in the 1990s which draws data from multiple administrative sources.

    1) A population statistical database with basic demographic information

    2) A dwelling statistical database on residential dwellings

- Singapore Citizens, Permanent Residents and foreigners on long term pass are issued with unique identification number (UIN), which facilitates data integration

Example of national data integration in Singapore

- **Product:** Cohort-based marriage and divorce indicators
- **Objective:** To track the outcome of marriages longitudinally and study the stability of marriages across cohorts and duration
- **Type of data source:** Marriage and marital dissolution records
- **Method used:** Deterministic multi-step matching approach

Challenges and solutions

• High data volume and frequency	• Tapping on technology and systems to automate validations, processing, fusion and provision of data
• Unmatched records arising from incomplete data, differences in format or unavailability of matching keys across datasets	• Perform validation checks and cleaning on matching keys to ensure that the data is complete • Establish national statistical standards for classifying information

Final word



- Data comes in heterogenous forms and data cleaning and preparation is an essential step before data integration.
- By combining up-to-date data from multiple public agencies, data integration enables us to support use cases across the public sector



- To keep up to date with countries' best practices on data integration
- To be updated with newest techniques to prepare data for integration

DEPARTMENT OF STATISTICS SINGAPORE

Thank you!

