



---

**United Nations Economic Commission for Europe  
Statistical Division**

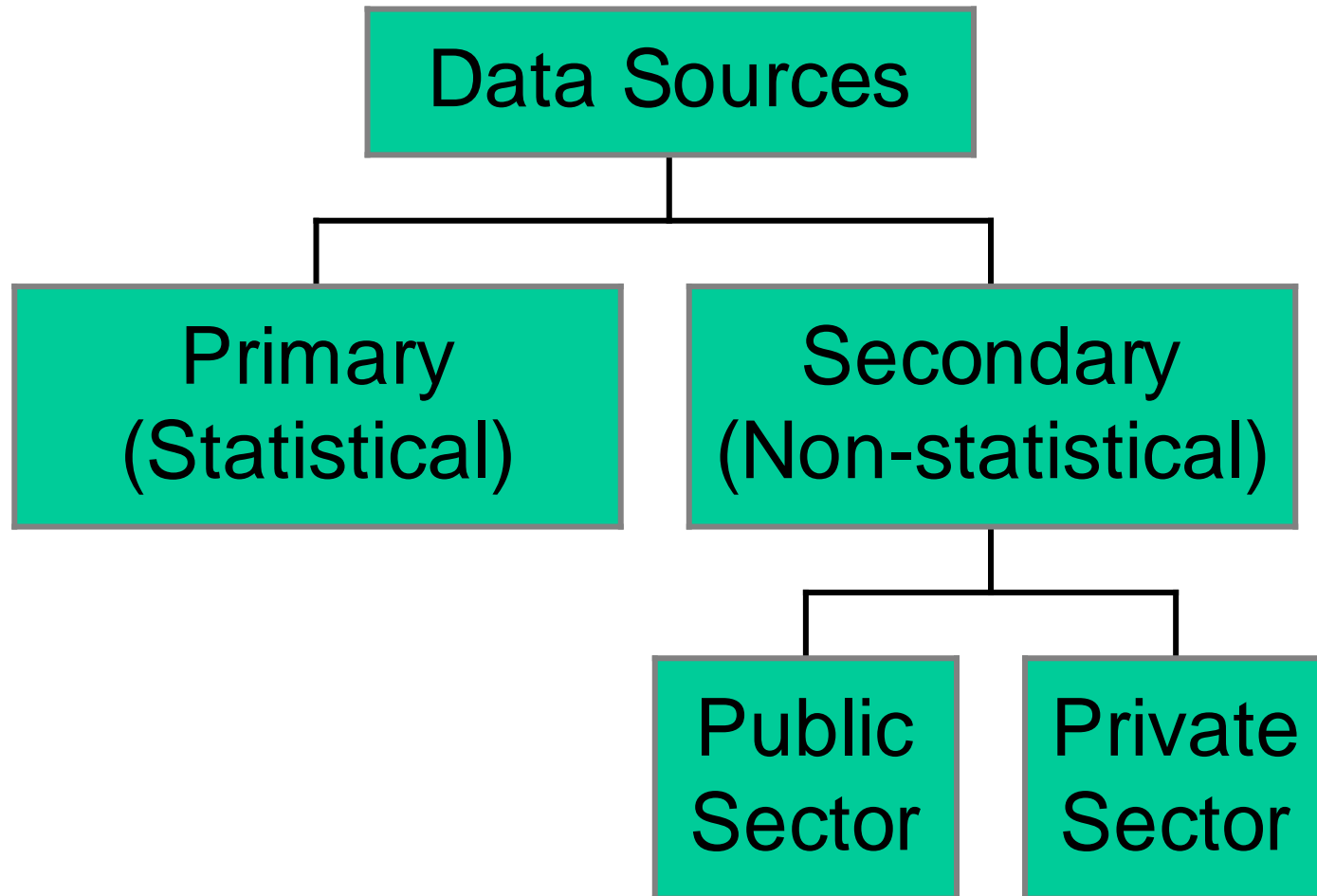
# **Introduction to Administrative Data**

**Steven Vale  
UNECE**

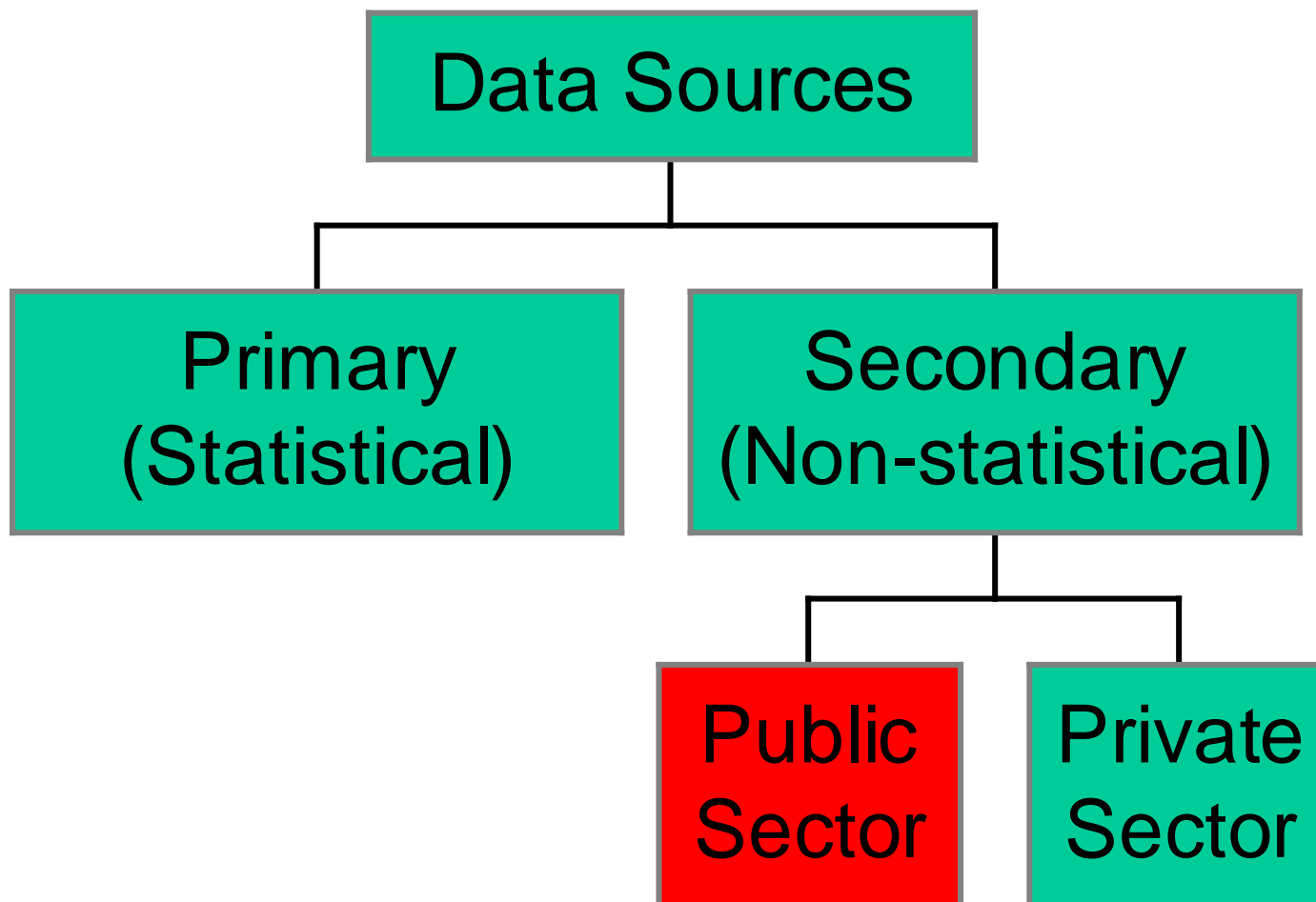
[steven.vale@unece.org](mailto:steven.vale@unece.org)



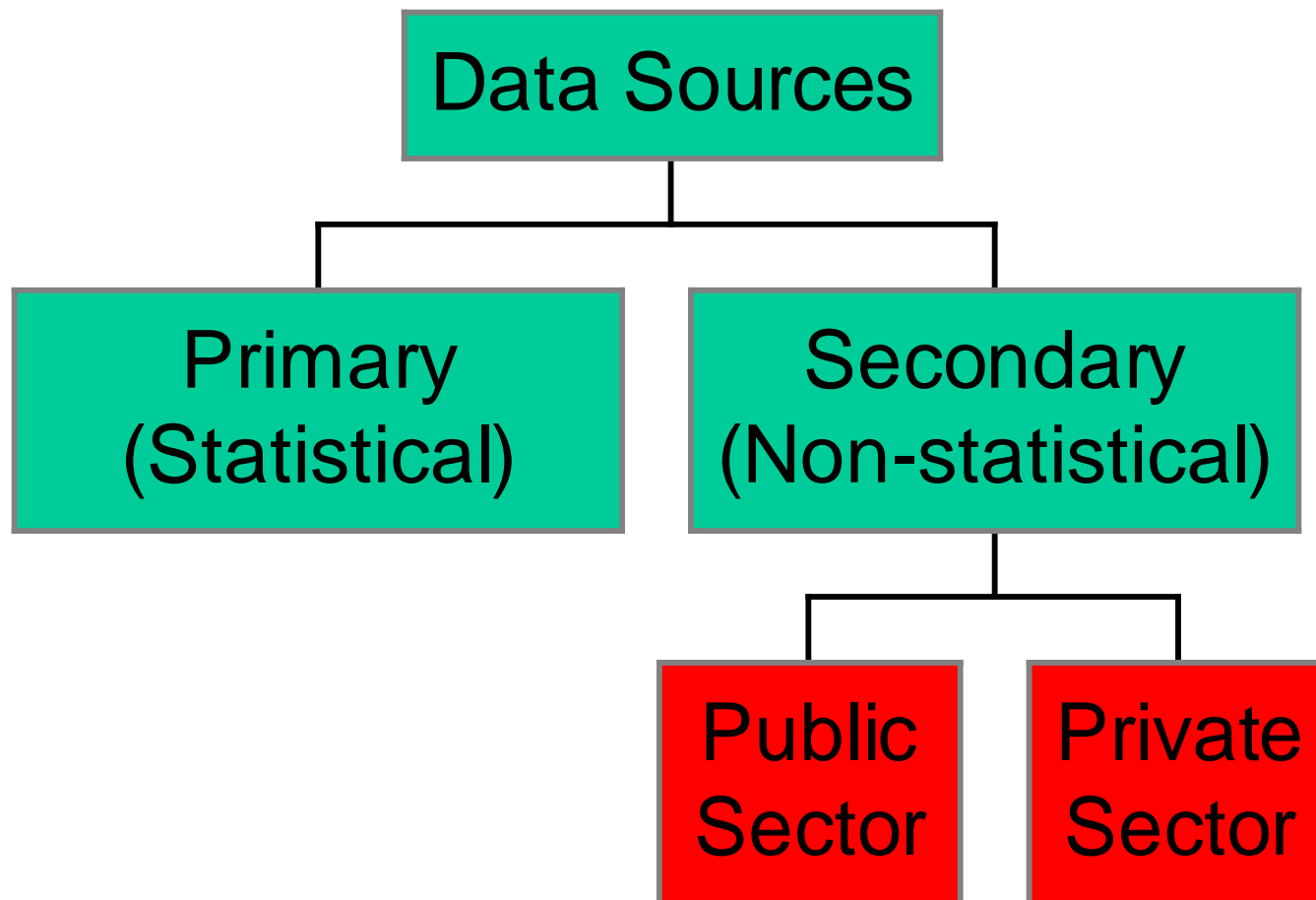
# Types of Data Source



# Narrow Definition



# Wider Definition





**Administrative sources  
contain information which  
is not primarily collected  
for statistical purposes.**

---

# Reasons for this definition

- ❖ Increasing privatisation of government functions
- ❖ Growth of private sector data and “value-added re-sellers”
- ❖ User interest in new types of data



# Examples of Administrative Sources

- Tax data
  - Personal income tax
  - Value Added Tax (VAT)
  - Business / profits tax
- Social security data
- Health / education records
- Registration systems for persons / businesses / property / vehicles





- Identity cards / passports / driving licenses
- Electoral register
- Register of farms
- Local council registers
- Building permits
- Licensing systems e.g. television, sale of restricted goods, import / export

- Published business accounts
- Internal accounting data
- Data held by private businesses:
  - credit agencies
  - business analysts
  - utility companies
  - telephone directories
  - retailers with store cards etc.



# **Advantages of Using Administrative Sources**

---

# Cost



- ❖ Surveys are expensive, a census is worse, data from administrative sources are often “free”
- ❖ Less staff are needed to process administrative data - no need for response chasing.

---

# 2001 Population census costs



- ❖ UK, €367m, €6.20 per person
- ❖ Finland, €0.8m, €0.20 per person

Source: Eurostat – Documentation of the 2000 round of population and Housing censuses in the EU, EFTA and Candidate Countries; Table 22

---

# Response Burden

- ❖ Using administrative sources:
  - Reduces the burden on data suppliers
  - Allows statistics to be compiled more frequently with no extra burden
- ❖ Data suppliers complain if they are asked to provide the same information many times by different government departments

---

# Coverage

- ❖ Administrative sources usually offer better coverage of target populations, and can make statistics more accurate:
  - No survey errors
  - No (or low) non-response
- ❖ Better coverage gives:
  - Better small-area data
  - More detailed information

---

# Timeliness

- ❖ Producing statistics from administrative sources can sometimes be quicker than using surveys
- ❖ No need for:
  - forms design;
  - pilot surveys;
  - sample design etc.





---

# Public Image

- ❖ Making more use of existing data can enhance the prestige of a statistical organisation by making it seem more efficient
- ❖ The concept of “Joined-up government” is politically appealing



# **Frameworks for the Access and Use of Administrative Data**



# Legal Frameworks



---

# National Legal Frameworks

- Most statistical institutions have legislation defining their roles and responsibilities, e.g. a ‘Statistical Law’
- Some statistical laws have been revised to include provisions for the use of administrative data

---

# Generic Law on Official Statistics



- ❖ **Article 17 - Access to administrative data**
- ❖ 17.1 All national and local authorities are obliged to provide the Producers of Official Statistics, free of charge, with data in their possession at the level of detail necessary for the production of official statistics and with the metadata that enable assessing data quality. Special confidentiality or secrecy provisions in other legislation cannot be invoked unless the legislation explicitly excludes the use of data for statistical purposes.

---

# Generic Law on Official Statistics



- ❖ 17.2 If the providers of administrative data plan to develop a new data collection or carry out a major revision in their data collection or processing in a way that may affect data provided for official statistics, they shall consult the National Statistical Office and, where appropriate, the Other Producers of Official Statistics, in advance of the decision.



# Policy Frameworks

---

# International Codes of Practice



## United Nations Fundamental Principles of Official Statistics

Principle 5: Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents



---

# International Codes of Practice



## European Statistical System Code of Practice

Principle 2: Statistical authorities must have a clear legal mandate to collect information for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities



---

# National Codes of Practice

- ❖ Can reassure the public that data will only be used for specific and reasonable purposes
- ❖ Should be made available to the public, e.g. via the internet site of the national statistical institute.



# Organisational Frameworks



---

# Access to Administrative Data in Practice

The use of administrative data is enabled by a legal framework, in the context of a policy framework

But - these are not usually detailed enough to cover all the administrative arrangements for access and use

---

# What Else is Needed?

- Some sort of agreement:
  - Service Level Agreement (SLA)
  - Administrative Protocol
  - Contract
  - Informal or verbal agreement
  - Other type of agreement according to national customs and practices

---

# Formal Agreements

- May not be legally enforceable, but are seen as a formal commitment
- Set out terms for the provision and use of data
- Signed by all parties
- Can cover one or more data flows (e.g. Government contracts)



---

# Agreement Contents

- ❖ Legal basis
- ❖ Names of persons transferring and receiving data
- ❖ Detailed description of data covered
- ❖ Frequency of data supply
- ❖ Quality standards
- ❖ Confidentiality rules

---

# Agreement Contents

- ❖ Technical standards
- ❖ Provision of metadata
- ❖ Provisions for payment for supply data
- ❖ Period of agreement
- ❖ Contingencies for changes in circumstances
- ❖ Procedure for resolving disputes





# Technical Frameworks



---

# Technical Frameworks

- ❖ Mechanisms for data transfer
  - Paper
  - Computer files (off-line)
  - Secure on-line connection
- ❖ File formats
- ❖ Data / metadata standards



# **Integrating Data from Different Sources**

---

# Why Integrate?

- Combining data sets can give more information than is available from individual data sets
- Reduce response burden
- Build efficient sampling frames
- Impute missing data

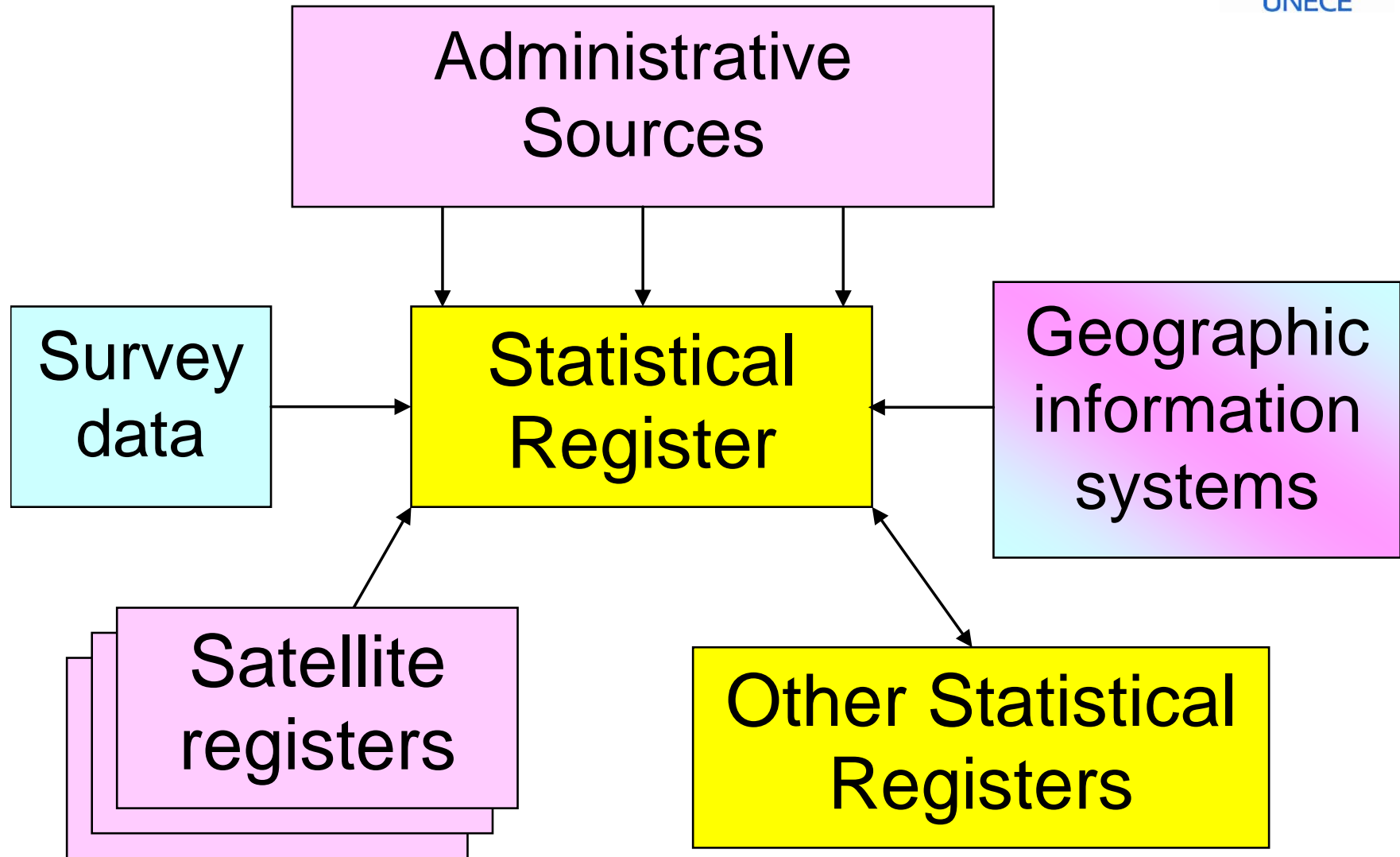


---

# Models for Data Integration

- ❖ Statistical registers
- ❖ Statistics from mixed source models
  - Split population model
  - Split data approach
  - Pre-filled questionnaires
  - Using administrative data for non-responders
  - Using administrative data for estimation
- ❖ Register-based statistical systems

# Statistical Registers





---

# Mixed Source Models

- ❖ Traditionally one statistical output was based on one statistical survey
- ❖ Very little integration or coherence
- ❖ Now there is a move towards more integrated statistical systems
- ❖ Outputs are based on several sources



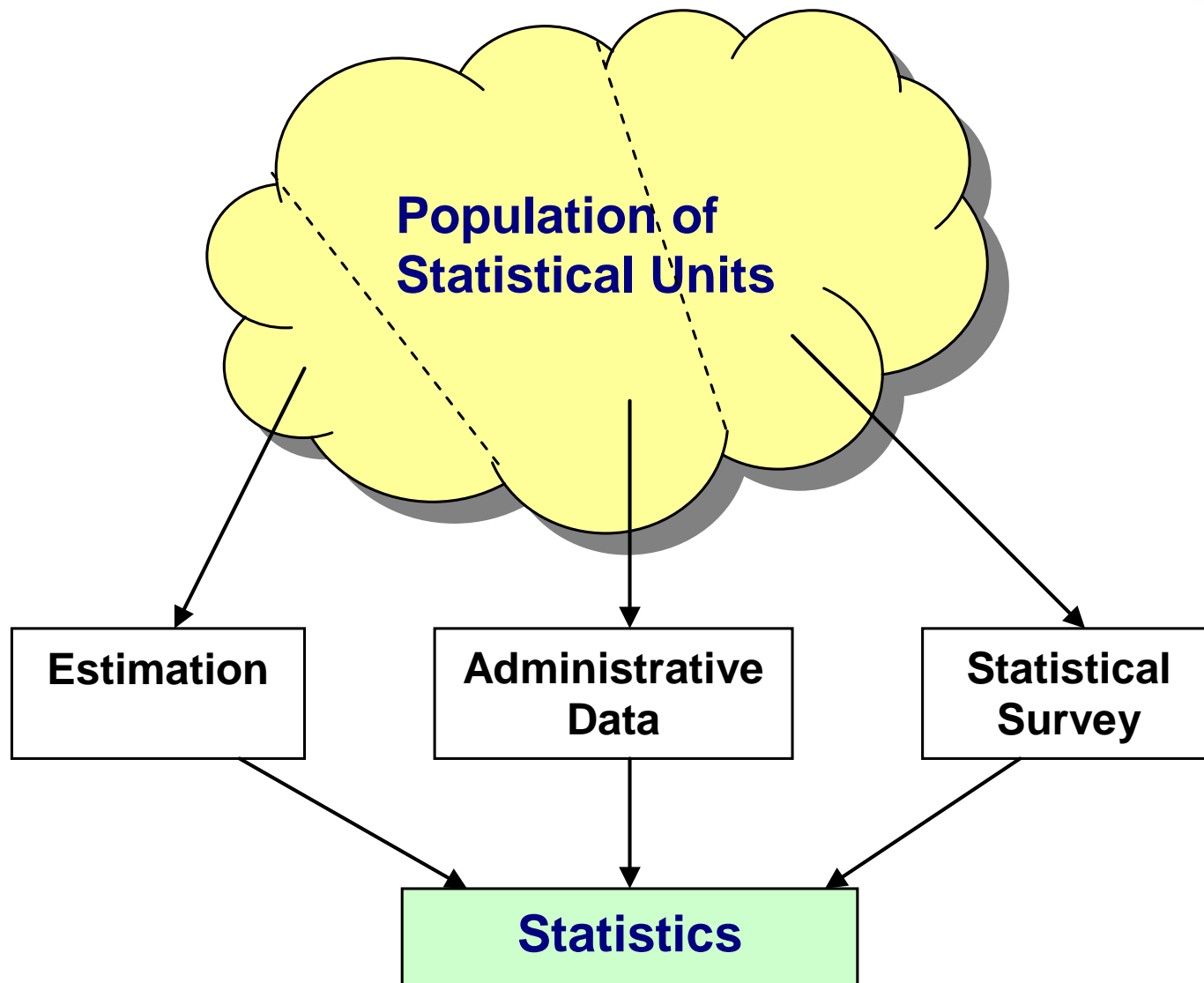
---

# Split Population Model

- ❖ One source of data for each unit
- ❖ Different sources for different parts of the population

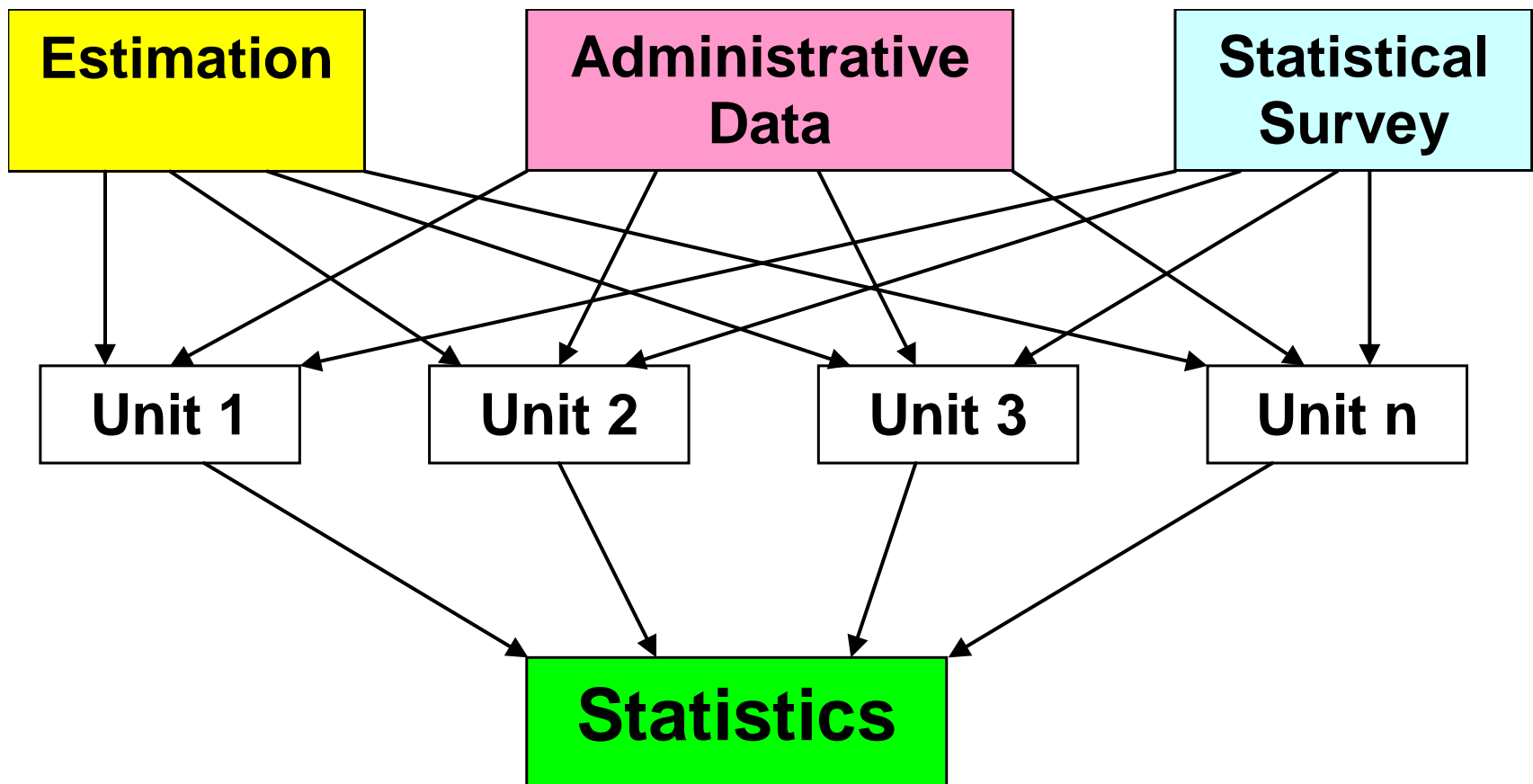


# Split Population Model



# Split Data Approach

- ❖ Several sources of data for each unit



---

# Pre-filled Questionnaires

- ❖ Survey questionnaires are pre-filled with data from other sources where possible
- ❖ Respondents check that the information is correct, rather than completing a blank questionnaire
- ❖ This reduces response burden  
..... but may introduce a bias!



---

# Using Administrative Data for Non-responders

- ❖ Administrative data are used directly to supply variables for units that do not respond to a statistical survey
- ❖ Often used for less important units, so that response-chasing resources can be focused on key units

---

# Using Administrative Data for Estimation



- ❖ Administrative data are used as auxiliary variables to improve the accuracy of statistical estimation
- ❖ Often used to estimate for small sub-populations or small geographic areas



UNECE

# Register-based Statistical Systems

