Agenda item 13:
Data compilation and data quality assurance at Customs and NSO

Presentation by Markie Muryawan
Statistician, International Trade Statistics Branch
Trade Statistics Branch, United Nations Statistics Division
E-mail: markie@un.org
Improvement at main data source

- All participants responded that Customs Declarations are the main source of data
- Data compilation and quality assurance at Customs will greatly improve quality of trade statistics
- Using computerized system to capture customs declarations will ensure basic data validation at data entry
What can be done at Customs?

- Full coverage of customs declarations
  - By consolidating data from automated and non-automated sites

- Partner country validation
  - Double check with shipping manifest, invoice, certificate of origin
  - Cross-checking on commodity level (e.g., not possible to import banana from Sweden)

- Ensure proper classification of goods
- Use price validation to identify possible incorrect values and/or quantity information
- Education programmes for customs officers and traders
Data Quality Assurance at NSO

• Ensure full coverage
  – Use non customs data source
  – Estimate missing values and/or quantities information

• Proper coding for commodity, partner country
  – Using standard code list

• Establish data cross-checking with common sense
  – Quantity unit for “electrical energy must be kwh
  – Oil imports probably via seaports / pipes not airports
  – Etc …
Data Quality Assurance at NSO (cont.)

- **Unit Value Validation**
  - Against standard / historical unit

- **Macro level checks**
  - Growth rates and composition of aggregates (structure of trade)
  - Against non-customs statistics, e.g., domestic production

- **Data Reconciliation and Exchange**
  - Experience indicates that reconciliation of data and subsequent data exchange improve the quality

- **Feedback from International Organizations**
  - E.g., data processed by UNSD (UN Comtrade)
Human vs. Machine

• It is true that experienced customs officers or statisticians can be very effective in detecting errors, however:
  – They have limited time (only work days)
  – They can be overwhelmed by high work load, which lead to decrease in their effectiveness
  – They may not be available or replaced by less experienced person
Therefore, the needs of validation tools

• Validation tools are not intended to replace resource person
  but
✓ just like other tools, it served to increase effectiveness and productivity of resource person
✓ Major errors, to be verified further (usually by experienced customs officers or statisticians), can be identified faster and more accurately
✓ In addition, automatic data correction (for basic error) can help reducing workload of resource person
Use of IT

• IT improvement in past years has enabled the use of technology for validation of “very large” data can be done in relatively inexpensive system such as PC
• It is worth to mention that the main problem in the implementation is not availability of tools but:
  ✓ How to reform/change established working culture?
  ✓ What is the right software for data validation?
  ✓ Has the software implemented best practices and national validation procedures?
  ✓ Do staffs have enough training?
What to do if errors are detected?

1. Categorize them into minor and major errors based on importance of trading partner/commodity and trade value

2. Ask clarification for major errors to data providing agencies, by explaining the reason (for example, unit value is outside acceptable range)

3. If errors seem systematic (for example, missing quantity from specific trader), coordinate with data providing agencies for more permanent solution
Thank you for your attention!