Data processing in UN Comtrade

United Nations Statistics Division
Regional Workshop on International Merchandise Trade Statistics
Suzhou, China – September 2017
UN Comtrade Overview

- Official trade statistics of almost 200 countries/areas
- Most comprehensive trade database with more than 1 billion records
- Annual trade in goods since 1962, monthly data since 2010
- Annual trade in services since 2000

https://comtrade.un.org/

Free access to detailed global trade data. UN Comtrade is a repository of official international trade statistics and relevant analytical tables. All data is accessible through API.
### UN Comtrade Current Data Elements

- Trade Flow
- Partner
- Commodity
- Trade values
  - CIF for Imports
  - FOB for Exports
- Net weight
- Supplementary Quantity and units
- Estimation Flag

<table>
<thead>
<tr>
<th>Period</th>
<th>Trade Flow</th>
<th>Reporter</th>
<th>Partner</th>
<th>Commodity Code</th>
<th>Trade Value (US$)</th>
<th>Netweight (kg)</th>
<th>Qty Unit</th>
<th>Qty</th>
<th>Flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>Import</td>
<td>China</td>
<td>World</td>
<td>7201</td>
<td>$50,122,175</td>
<td>190,291,180</td>
<td>Weight in kilograms</td>
<td>190,291,180</td>
<td>6</td>
</tr>
<tr>
<td>2016</td>
<td>Export</td>
<td>China</td>
<td>World</td>
<td>7201</td>
<td>$27,967,757</td>
<td>133,061,303</td>
<td>Weight in kilograms</td>
<td>133,061,303</td>
<td>0</td>
</tr>
</tbody>
</table>
• Second Partner
  • Country of Consignment

• Trade values
  • FOB for Imports (as well as CIF)

• Mode of transport

• Customs procedure codes
Data Request
• Official request for trade data from data providers (Feb)
• New and revised data

Data Collection & Registration
• UNSD receives and registers trade data for processing

Data Pre-Processing
• Data is transformed to the standard input format
• Data reference decks are prepared

Data Processing
• Normalization & harmonization
• Quantity & weight estimation
• Conversion to other classifications

Data Review & Validation
• Checking, review and validation

Data Publication & Dissemination
• Publication to UN Comtrade
• Managing users’ feedback
Pre-Processing Objectives

- **Transform** national data into the input format required by the internal processing system
- **Ensure** that national code references are available so that we are able to map national codes to standard codes correctly during processing stage
- **Handle** unique and special cases
- **Maintain** and level of information originally submitted
- **Ensure that the necessary parameters/fields are available**
  - Trade flow, commodity at detailed 6-digit level, partner code, trade value, weight, quantity and quantity unit code
- **Produce an output** file in .mdb or tab-delimited format
Pre-processing: Data validation

- **Check trade values against metadata.** Trade values may sometimes be broken down into value, insurance and freight.
- **Check if there are negative values**
- **Accumulated/Aggregated Totals**
  - Aggregate detailed data to get totals by trade flow, and compare with official or reported totals.
  - New vs Revised datasets: We compare aggregated totals in detail to determine if we should re-process revised data
- **Verify that related oil trade data is reported**
- **Confidential values**
  - Suppressed data under pseudo headings
    - ‘9999AA’, where ‘AA’ reflects chapter
  - Reported negative values under unknown commodity
- **High-level totals** must be greater than the aggregated totals of detailed records
Data Request
• Official request for trade data from data providers (Feb)
• New and revised data

Data Collection & Registration
• UNSD receives and registers trade data for processing

Data Pre-Processing
• Data is reformatted to the standard input format
• Data reference decks are prepared

Data Processing
• Normalization & harmonization
• Quantity & weight estimation
• Conversion to other classifications

Data Review & Validation
• Checking, review and validation

Data Publication & Dissemination
• Publication to UN Comtrade
• Managing users’ feedback
## Processing & Harmonization

### Current Deck Processing Summary

<table>
<thead>
<tr>
<th>Category</th>
<th>Partners</th>
<th>Deck Valid</th>
<th>Percent Change (Y-o-Y)</th>
<th>Total Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totals</td>
<td>314</td>
<td>100%</td>
<td>0%</td>
<td>60746</td>
</tr>
<tr>
<td>Exports</td>
<td>185</td>
<td>100%</td>
<td>-4.6%</td>
<td>2195.2 mil</td>
</tr>
<tr>
<td>Imports</td>
<td>129</td>
<td>100%</td>
<td>0%</td>
<td>2196.2 mil</td>
</tr>
<tr>
<td>N/A</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0.0 mil</td>
</tr>
<tr>
<td>Not Validated</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0.0 mil</td>
</tr>
</tbody>
</table>

### Deck Processing Result

<table>
<thead>
<tr>
<th>Country</th>
<th>Code</th>
<th>Original Description</th>
<th>Mapped to Code</th>
<th>Mapped to Description</th>
<th>Flow Code</th>
<th>Validated</th>
<th>Rank</th>
<th>Records</th>
<th>CIF Value (mld orig)</th>
<th>FOB Value (mld orig)</th>
<th>Net Weight (mld org units)</th>
<th>Gross Value (mld org units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>0560</td>
<td>Afghanistan</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>51</td>
<td>-</td>
<td>0.037%</td>
<td>0.255%</td>
<td>0.301%</td>
<td>0.255%</td>
</tr>
<tr>
<td>Albania</td>
<td>0070</td>
<td>Albania</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>325</td>
<td>-</td>
<td>0.005%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Angola</td>
<td>0300</td>
<td>Angola</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>4</td>
<td>-</td>
<td>0.005%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Antigua &amp; Barbuda</td>
<td>0459</td>
<td>Antigua and Barbuda</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>5</td>
<td>-</td>
<td>0.003%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Argentina</td>
<td>0528</td>
<td>Argentina</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>12</td>
<td>-</td>
<td>0.004%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Armenia</td>
<td>0077</td>
<td>Armenia</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>11</td>
<td>-</td>
<td>0.040%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Aruba</td>
<td>0474</td>
<td>Aruba</td>
<td>X Y</td>
<td>X Y</td>
<td></td>
<td></td>
<td>8</td>
<td>-</td>
<td>0.001%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
• **Code Normalization**
  – National data items are mapped to standard codes (M49, WCO, Revised Kyoto Convention Codes)
  – Currency conversion to USD
  – Output: Tariffline data

• **Commodity Normalization**
  – Removes monetary gold [HS: 711890, 710820] from dataset
  – Non-standard codes are addressed
  – Special codes are mapped (i.e. pseudo-codes)
Quantity Information

- UN Comtrade includes **net weight** and, when available, **supplementary quantity**
- Quantity units are standardized according to World Customs Organization recommended units to provide comparable data
- **Non-standard quantities** are converted using specific conversion factors or FAO conversions
- UN Comtrade **estimates missing quantity** at the most detailed reported level when possible using the reporter’s reported weighted unit value within the same 6-digit commodity flow or median standard unit value of all reporters for the same 6-digit commodity flow the previous year, after elimination of outliers
- Detect and remove extreme outliers in quantity based on:
  - Historical Unit Value Check
  - Standard Unit Value Check
  - Historical Value Check
Conversion to other classifications

- Data reported in latest HS classification is converted to earlier versions of HS
- Use of WCO provided correlation tables between latest HS and previous HS version
- Our practice is to convert most detailed level in original classification into one single code in the target classification, without splitting.
Correlation and conversion

The complex relationships (left) are simplified (right)
## Conversion to other classifications

### Conversion and Correlation Tables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HS 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS 2012</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS 2007</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS 2002</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS 1996</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS 1992</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SITC 4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SITC 3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SITC 2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SITC 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Download Complete correlations among HS, SITC and BEC

Download Conversion Methodology General Note (2017)

Data Review

- Consolidated report contains result of each processing phase
- Reviews both by editor and supervisor
- Check overall totals against official sources, time series, trends
- Check level of memorandum items (overall value of non-standard commodity codes)
- Verify the mapping of partners, commodities and quantity units are correct
- Check integrity reports (possible wrong classification, conversion factors changed more than 1% from previous processing, export to itself, general trade system but data has free zones, partner world vs sum of individual partners)
### Table 6: Partner Value Comparison

<table>
<thead>
<tr>
<th>Flow</th>
<th>Partner</th>
<th>Value.A</th>
<th>Value.B</th>
<th>%diff</th>
<th>diff</th>
<th>%share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AAA</td>
<td>$611,364,435,458.00</td>
<td>$611,364,435,458.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>100.00%</td>
</tr>
<tr>
<td>1</td>
<td>AFG</td>
<td>$2,506,231.00</td>
<td>$2,506,231.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>ALB</td>
<td>$23,238,224.00</td>
<td>$23,238,224.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>ANT</td>
<td>$16,470.00</td>
<td>$16,470.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>DZA</td>
<td>$4,627,196,986.00</td>
<td>$4,627,196,986.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.76%</td>
</tr>
<tr>
<td>1</td>
<td>ASM</td>
<td>$32,941.00</td>
<td>$32,941.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>AND</td>
<td>$20,263,962.00</td>
<td>$20,263,962.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>AGO</td>
<td>$2,370,503,388.00</td>
<td>$2,370,503,388.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.39%</td>
</tr>
<tr>
<td>1</td>
<td>ATG</td>
<td>$2,952,650.00</td>
<td>$2,952,650.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>1</td>
<td>AZE</td>
<td>$2,260,091,970.00</td>
<td>$2,260,091,970.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.37%</td>
</tr>
<tr>
<td>1</td>
<td>ARG</td>
<td>$807,654,225.00</td>
<td>$807,654,225.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.13%</td>
</tr>
<tr>
<td>1</td>
<td>AUS</td>
<td>$1,610,338,536.00</td>
<td>$1,610,338,536.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.26%</td>
</tr>
<tr>
<td>1</td>
<td>AUT</td>
<td>$5,339,687,175.00</td>
<td>$5,339,687,175.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.87%</td>
</tr>
</tbody>
</table>

### Table 8: Commodity Value Comparison

<table>
<thead>
<tr>
<th>Flow</th>
<th>Code</th>
<th>Value.A</th>
<th>Value.B</th>
<th>%diff</th>
<th>diff</th>
<th>%share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TOTAL All Commodities</td>
<td>$611,364,435,458.00</td>
<td>$611,364,435,458.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>100.00%</td>
</tr>
<tr>
<td>1</td>
<td>01 Live animals; animal prod</td>
<td>$352,104,872.00</td>
<td>$352,104,872.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.06%</td>
</tr>
<tr>
<td>1</td>
<td>02 Meat and edible meat offa</td>
<td>$4,818,724,643.00</td>
<td>$4,818,724,643.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.79%</td>
</tr>
<tr>
<td>1</td>
<td>03 Fish and crustaceans, mol</td>
<td>$4,164,666,970.00</td>
<td>$4,164,666,970.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.68%</td>
</tr>
<tr>
<td>1</td>
<td>04 Dairy produce; birds' egg</td>
<td>$3,483,250,314.00</td>
<td>$3,483,250,314.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.57%</td>
</tr>
<tr>
<td>1</td>
<td>05 Products of animal origin</td>
<td>$316,305,297.00</td>
<td>$316,305,297.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.05%</td>
</tr>
<tr>
<td>1</td>
<td>06 Live trees and other plan</td>
<td>$1,429,273,637.00</td>
<td>$1,429,273,637.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.23%</td>
</tr>
<tr>
<td>1</td>
<td>07 Edible vegetables and cer</td>
<td>$3,023,608,529.00</td>
<td>$3,023,608,529.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.49%</td>
</tr>
<tr>
<td>1</td>
<td>08 Edible fruit and nuts; pe</td>
<td>$4,163,397,384.00</td>
<td>$4,163,397,384.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.68%</td>
</tr>
<tr>
<td>1</td>
<td>09 Coffee, tea, maté and spi</td>
<td>$1,388,285,917.00</td>
<td>$1,388,285,917.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.23%</td>
</tr>
<tr>
<td>1</td>
<td>10 Cereals</td>
<td>$910,063,821.00</td>
<td>$910,063,821.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.15%</td>
</tr>
<tr>
<td>1</td>
<td>11 Products of the milling i</td>
<td>$380,426,930.00</td>
<td>$380,426,930.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.06%</td>
</tr>
<tr>
<td>1</td>
<td>12 Oil seeds and oleaginous</td>
<td>$862,062,497.00</td>
<td>$862,062,497.00</td>
<td>0.00%</td>
<td>$0.00</td>
<td>0.14%</td>
</tr>
</tbody>
</table>

- **Table 6: Partner Value Comparison**
  - Check major differences by partners
- **Table 8: Commodity Value Comparison**
  - Check major differences by commodities
## New UN Comtrade interface

### 1. Type of product & Frequency

<table>
<thead>
<tr>
<th>Type of product</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goods</td>
<td>Annual</td>
</tr>
<tr>
<td>Services</td>
<td>Monthly</td>
</tr>
</tbody>
</table>

### 2. Classification

- **HS**
  - As reported
  - 02
  - 96
  - 02
  - 07
  - 12

- **SITC**
  - As reported
  - Rev. 1
  - Rev. 2
  - Rev. 3
  - Rev. 4

- **BEC**
  - BEC

### 3. Select desired data

- **Periods (year)**: 2016
- **Reporters**: All
- **Partners**: World
- **Trade flows**: All

**HS (as reported) commodity codes**

- TOTAL - Total of all HS commodities

### 4. See the results

- [Preview](#)
- [Download CSV](#)

Issues opening CSV in Excel? See this Microsoft how-to.
Metadata

Step 1. Select Selection Filter

Reporter: Albania
Year: 2016
Loaded: (All)

Step 2. View Search Results

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Reported Classification</th>
<th>Reported Currency</th>
<th>Trade Flow</th>
<th>Currency Conversion Factor</th>
<th>Trade System</th>
<th>Valuation</th>
<th>Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albania(2016)</td>
<td>HS2012</td>
<td>ALL</td>
<td>Import</td>
<td>0.0008061</td>
<td>Special</td>
<td>CIF</td>
<td>Origin</td>
</tr>
<tr>
<td>Albania(2016)</td>
<td>HS2012</td>
<td>ALL</td>
<td>Export</td>
<td>0.008058</td>
<td>Special</td>
<td>FOB</td>
<td>Last Known Destination</td>
</tr>
</tbody>
</table>

**Reporter/Year** Publication Note

Finland (2016) New Data: Data for this year is published for the first time; the data was compiled from the monthly data set from National Board of Customs (Tulli-Tullen - Finnish Customs), with the last submission on 28 February 2017.

First/Last Publication Date

2017.09.08 View
2017.09.08 History
Data Release

- Ongoing and continuous release and publication of datasets

**API:**

User support mailbox
[comtrade@un.org](mailto:comtrade@un.org)
Thank You!