United Nations Statistics Division

DATA PROCESSING IN

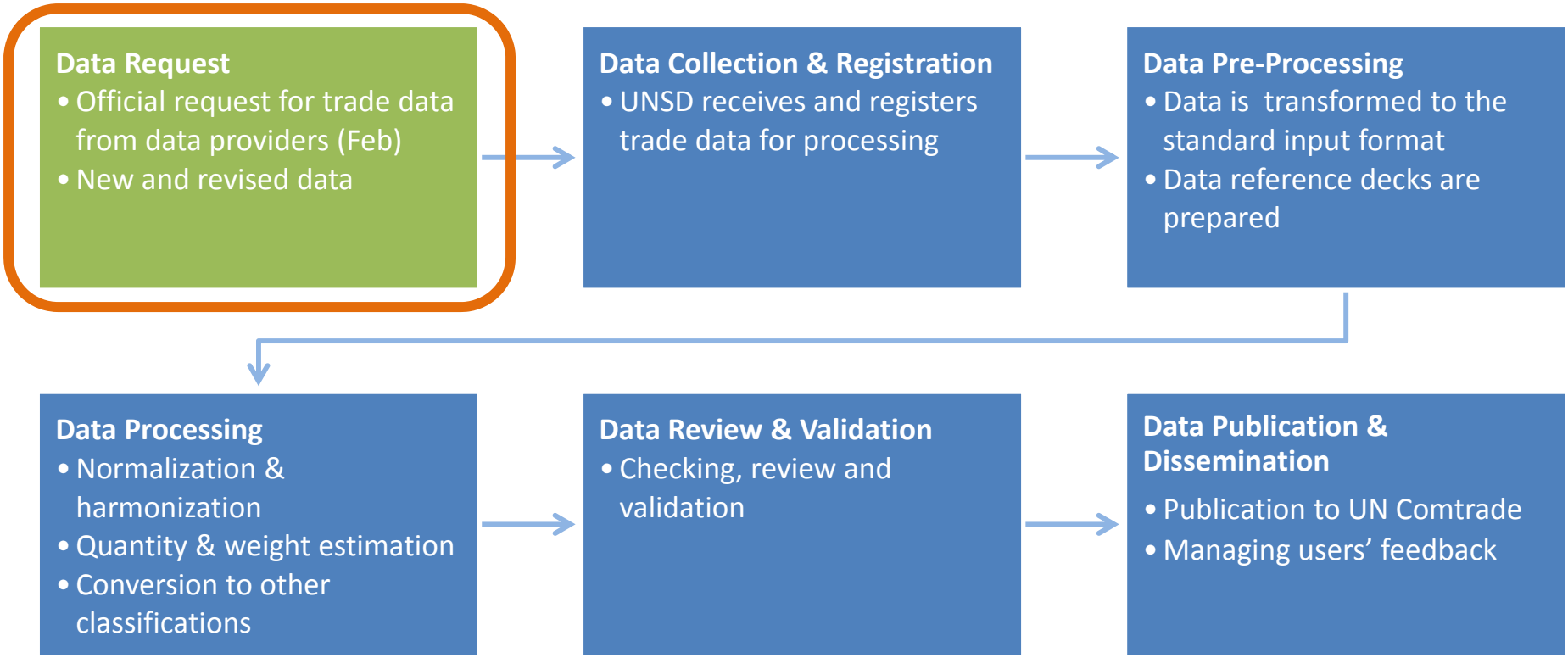
Regional Workshop on IMTS – Malé, Maldives

3 April 2017

UN Comtrade Database

- United Nations International Trade Statistics Database
- Comprehensive repository of official detailed international trade data since 1962
- It contains both merchandise and services trade data
- https://comtrade.un.org/

Production Process: Data Request



Data Items

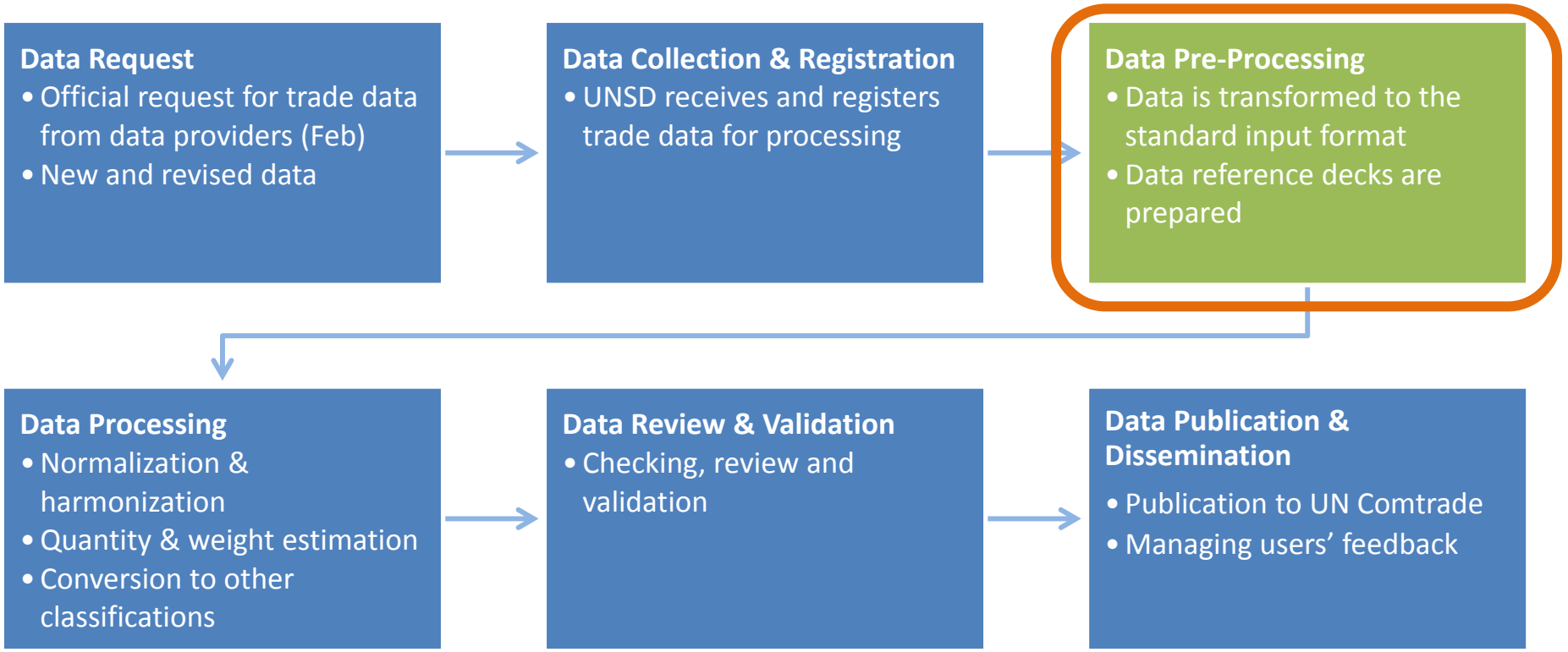
Pre-IMTS 2010

- Commodity
- Partner
- Trade values
  - CIF for Imports
  - FOB for Exports
- Net weight
- Quantity and units
- Official totals

IMTS 2010 (new items)

- Partner
  - Consignment
- Trade values
  - FOB for Imports
- Mode of transport
- Customs procedure codes

Production Process: Pre-processing



Pre-Processing Objectives

- **Transform** national data into the input format required by the internal processing system
- Maintain the original level of detail
- 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
- Produce an output file in .mdb or tab-delimited format

National Data

- Formats:
  - Text files (.txt, .dat, .tsv)
  - Worksheets (.xlsx)
  - Database (.mdb/.accdb, .dbf)
- Methods: Email, FTP, Direct download, CD
- Content:
  - Detailed trade data
  - Reference decks
  - Official Totals
  - Data layout
  - Explanatory notes
  - Other supplementary files

Tools and Software

- MS Access
- Excel
- Notepad ++
- SQL
- SQL Server Integrated Services (SSIS) and fATA
- R
- Etc.

Pre-Processing Procedure

Import the original data file

- Maintain and keep level of information originally submitted (do not aggregate)
- 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
- Ensure that data layout is accurate

Pre-Processing

Inspect data

- Check trade values against metadata. Trade values may sometimes be broken down into value, insurance and freight.

  Example: USA data – Imports at C.I.F., Exports at F.O.B.
  In order to arrive at C.I.F., we need to add [General Imports Customs Value] and [General Imports Charges]

- Check if there are negative values
- Check if data has aggregated values

Pre-Processing

Inspect data (cont.)

- 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

Pre-Processing

Inspect data (cont.)

- Supplementary files containing detailed trade data provided

Example: Russian Federation
Main data file has record for unknown commodity and partner, with trade value.
Supplementary file provides gas and oil data with unknown partner and trade value

→ *We insert a record with the correct commodity code for gas and oil and the corresponding reported value.*
→ *In order for accumulated totals to remain as reported, we adjust the unknown commodity and partner trade value record by removing the corresponding value for gas and oil. This minimizes the value under unknown commodities, and provides the reported value for gas and oil in the dataset.*

Pre-Processing

Inspect data (cont.)

- Oil producing countries – Verify that related oil trade data is reported
  - If not available, request data from country
  - We can still process data with an explanatory note that the dataset excludes oil data

Pre-Processing

Inspect data (cont.)

- Confidential values
  - Suppressed data under pseudo headings
    - '9999AA', where 'AA' reflects chapter (i.e. UK)
  - Data provided in commodity detail and submitted with supplementary file containing confidential commodities requested to be suppressed during processing by mapping to <chapter>+"0000" (i.e. Switzerland)
  - Reported as negative values under unknown commodity (i.e. Australia)

Pre-Processing

Handle reported high-level data

- High-level totals must be greater than the aggregated totals of detailed records
  - Check and compare the high-level vs the detailed totals
  - We use the feature of EXCEL to calculate and compare TOTALS BY PARTNER, keeping in mind that the overall totals must be the same.

Pre-Processing

Create reference decks

- Create reference decks for national data items and map these unique national codes to standard reference decks
  - Partner Codes: National Codes -> M49
  - Quantity Units: National Codes -> WCO units
  - Modes of Transport
  - Customs Procedure Codes: National Codes -> Revised Kyoto Convention Codes

Challenges in Pre-Processing

- Data format is difficult, but not impossible, to convert into standard format
  - Example: .pdf or .doc
  - Data in worksheet format instead of record format
  - Multiple and large files
- No data layout information
- Incomplete metadata/explanatory notes
  - Trade system
  - Valuation
  - Partner country attribution
  - Weight – net or gross?

Challenges in Pre-Processing

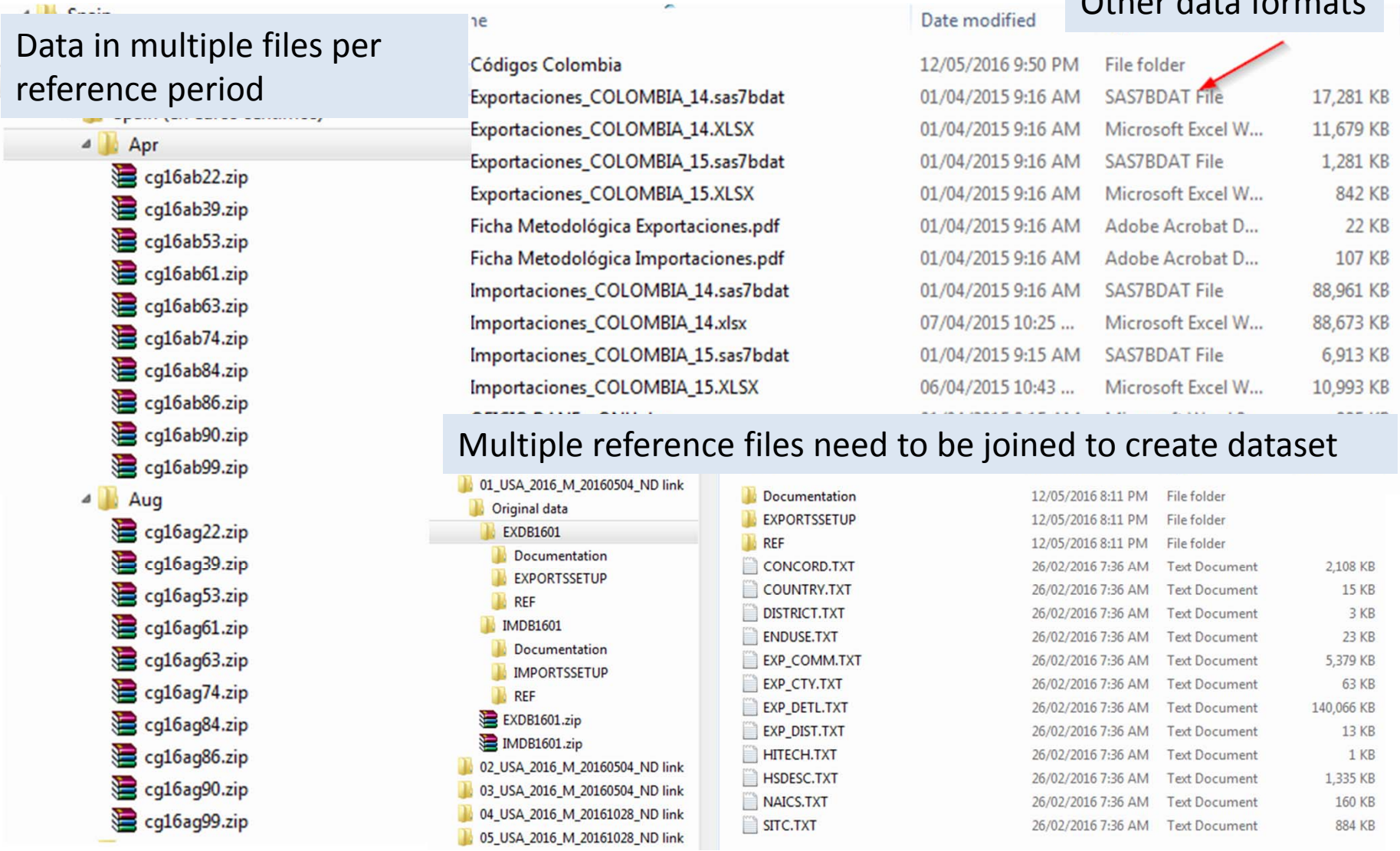
- Reference decks not provided
- Other languages
- Exports data for oil not available for oil producing countries
- High-level data; commodities not in 6-digit level; confidential data
- High percentage of invalid commodity codes resulting big memorandum items

Worksheet, different language

Код ТН ВЭД	Наименование товара	Наименование дополнительной единицы измерения	Экспорт			Импорт		
			количество		тыс. долларов США	количество		тыс. долларов США
			тонн	дополнительная ед. измерения		тонн	дополнительная ед. измерения	
	Всего		**12,992,504.7**		**4,169,110.6**	**2,352,431.3**		**2,751,962.2**
	в том числе:							
010121	**Лошади живые: чистопородные племенные животные**	**Штука**				**20.4**	**42.0**	**80.2**
	Страны СНГ					17.2	34.0	70.2
	Россия					17.2	34.0	70.2
	Остальные страны мира					3.2	8.0	10.1
	Франция					0.4	1.0	1.3
	Соеденненное Королевство					0.4	1.0	1.2
	Ирландия					1.2	3.0	3.8
	Польша					0.8	2.0	2.6
	Соединенные Штаты Америки					0.4	1.0	1.3
010129	**Лошади прочие**	**Штука**	**2.1**	**7.0**	**5.1**	**9.9**	**19.0**	**6.6**
	Страны СНГ		2.1	7.0	5.1	9.9	19.0	6.6
	Киргизстан		2.1	7.0	5.1			
	Украина					9.9	19.0	6.6
010221	**Чистопородный племенной домашний крупный рогатый скот**	**Штука**	**4.8**	**19.0**	**10.2**	**28.4**	**82.0**	**136.3**
	Страны СНГ		4.8	19.0	10.2	28.4	82.0	136.3
	Россия		4.8	19.0	10.2	28.4	82.0	136.3

Data in multiple files per reference period

Other data formats

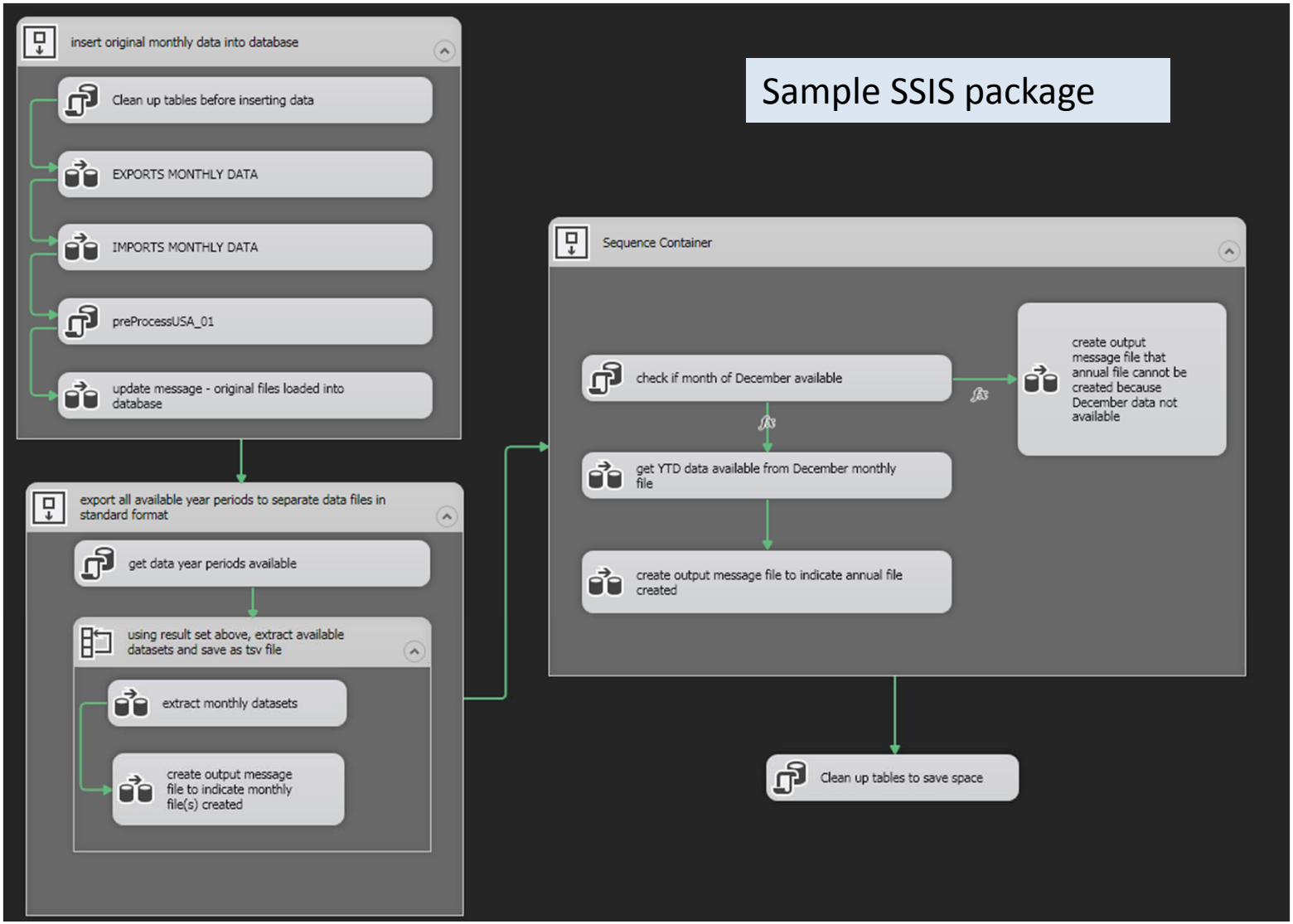


Multiple reference files need to be joined to create dataset

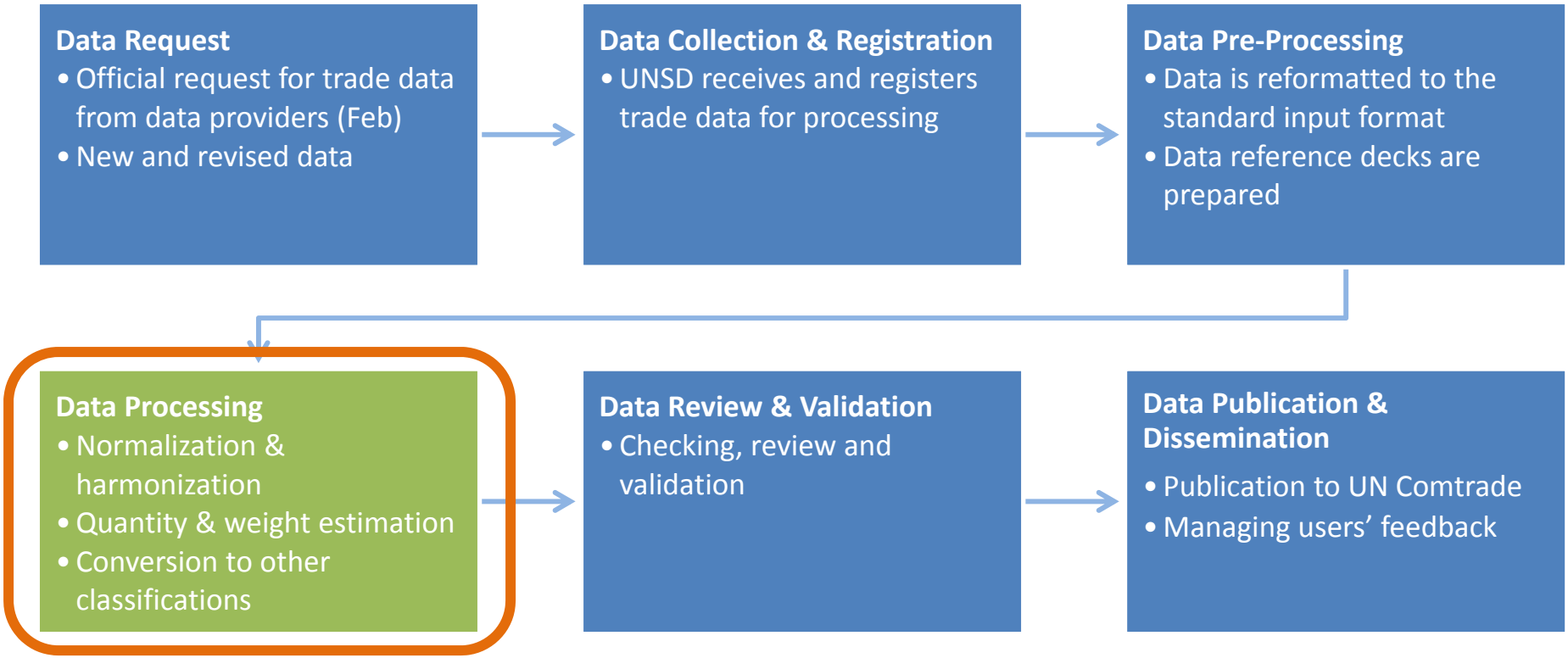
Pre-processing tool:
SQL Server Integrated Services (SSIS)

- Advantages:
  - Can handle multiple data sources and formats in a single package
  - Great for complex transformations, multi-step operations
  - Can store the pre-processing steps in a package ensuring consistency in method
  - Able to join and loop through multiple files
- Disadvantages
  - Learning curve
  - Needs software and technical support
  - Assumes that national data format is stable; Any change in input format will require an update to package.

Sample SSIS package



Production Process: Processing



CoprA: Internal Processing System

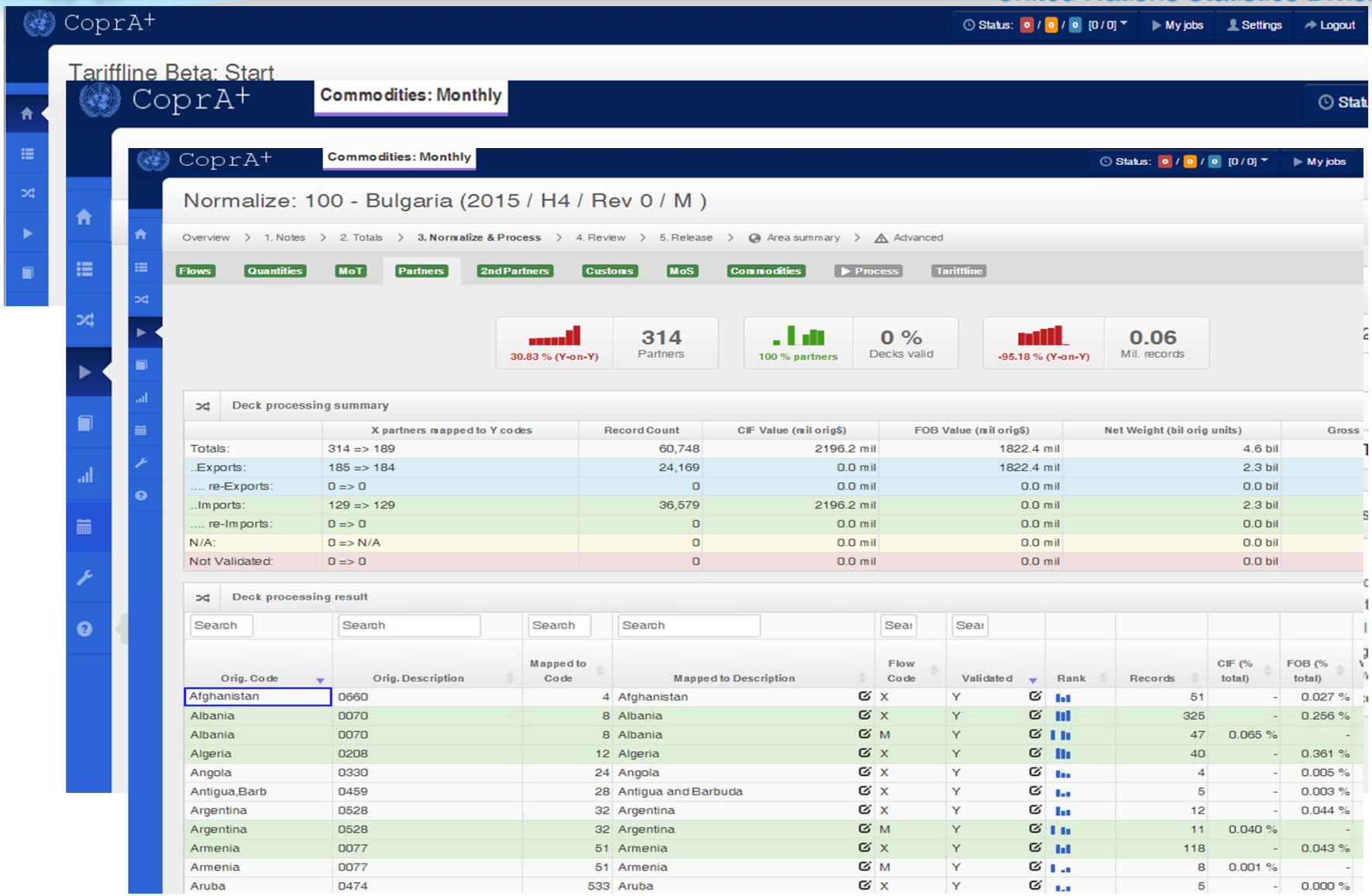
"Combined Processing System"

CoprA: Pre-IMTS 2010

- Three separate processing systems for annual merchandise, monthly merchandise, annual trade in services data
- MS Access is used as an intermediate format between national data format and SQL server

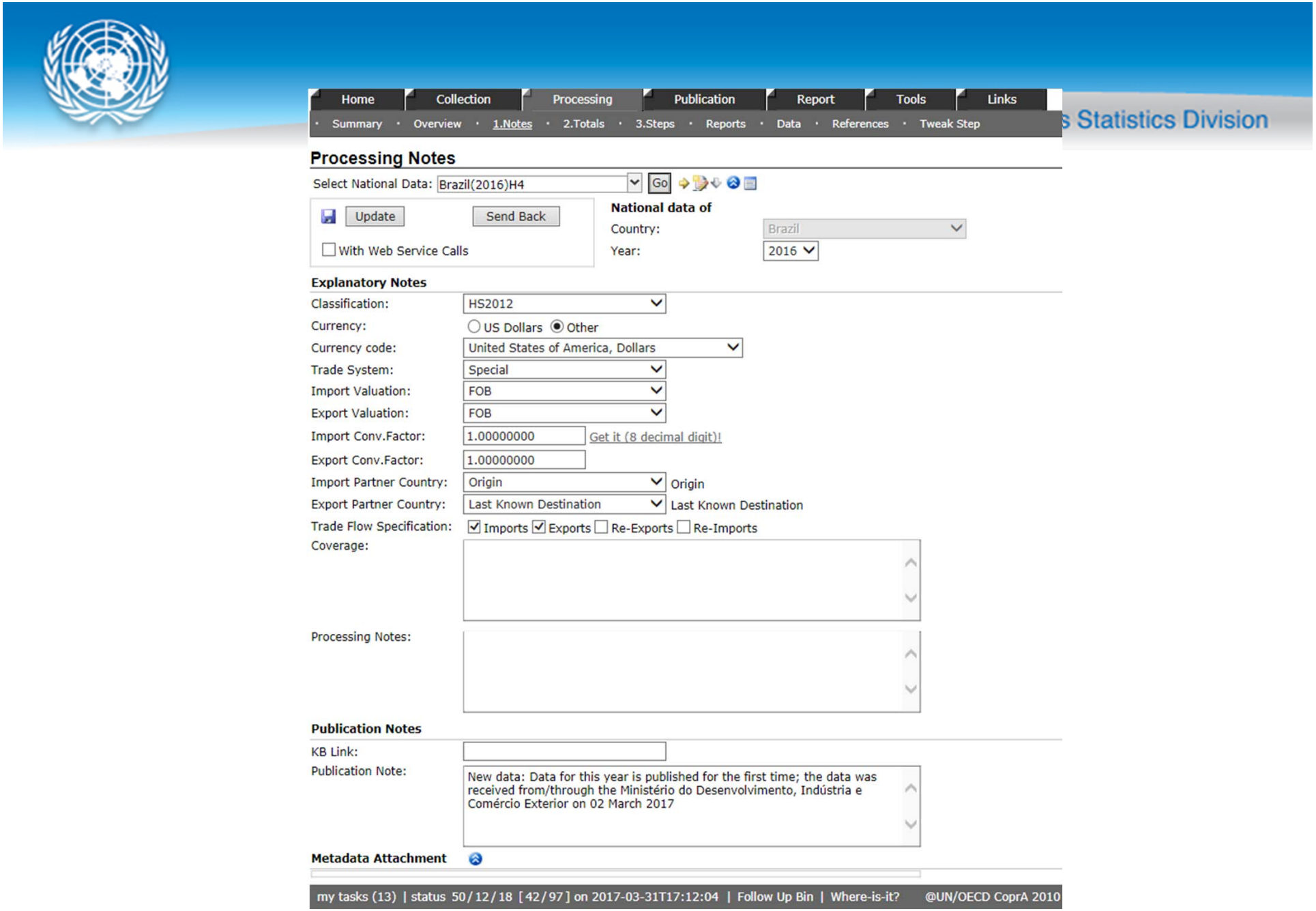
CopraPlus: IMTS 2010

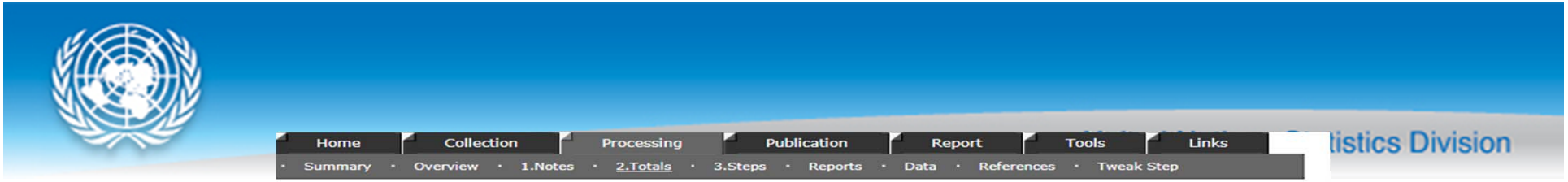
- Single processing system for annual and monthly merchandise, as well as annual trade in services data
- Includes IMTS 2010 data items
- Text file is used as an intermediate format



	X partners mapped to Y codes	Record Count	CIF Value (mil orig$)	FOB Value (mil orig$)	Net Weight (bil orig units)	Gross
Totals:	314 => 189	60,748	2196.2 mil	1822.4 mil	4.6 bil	
..Exports:	185 => 184	24,169	0.0 mil	1822.4 mil	2.3 bil	
.... re-Exports:	0 => 0	0	0.0 mil	0.0 mil	0.0 bil	
..Imports:	129 => 129	36,579	2196.2 mil	0.0 mil	2.3 bil	
.... re-Imports:	0 => 0	0	0.0 mil	0.0 mil	0.0 bil	
N/A:	0 => N/A	0	0.0 mil	0.0 mil	0.0 bil	
Not Validated:	0 => 0	0	0.0 mil	0.0 mil	0.0 bil	

Orig. Code	Orig. Description	Mapped to Code	Mapped to Description	Flow Code	Validated	Rank	Records	CIF (% total)	FOB (% total)
Afghanistan	0660	4	Afghanistan	X	Y		51	-	0.027 %
Albania	0070	8	Albania	X	Y		325	-	0.256 %
Albania	0070	8	Albania	M	Y		47	0.065 %	-
Algeria	0208	12	Algeria	X	Y		40	-	0.361 %
Angola	0330	24	Angola	X	Y		4	-	0.005 %
Antigua,Barb	0459	28	Antigua and Barbuda	X	Y		5	-	0.003 %
Argentina	0528	32	Argentina	X	Y		12	-	0.044 %
Argentina	0528	32	Argentina	M	Y		11	0.040 %	-
Armenia	0077	51	Armenia	X	Y		118	-	0.043 %
Armenia	0077	51	Armenia	M	Y		8	0.001 %	-
Aruba	0474	533	Aruba	X	Y		5	-	0.000 %


Home
Collection
Processing
Publication
Report
Tools
Links
Summary
Overview
1.Notes
2.Totals
3.Steps
Reports
Data
References
Tweak Step
s Statistics Division
Processing Notes
Select National Data: Brazil(2016)H4
Go
Update
Send Back
With Web Service Calls
National data of
Country: Brazil
Year: 2016
Explanatory Notes
Classification: HS2012
Currency: US Dollars Other
Currency code: United States of America, Dollars
Trade System: Special
Import Valuation: FOB
Export Valuation: FOB
Import Conv.Factor: 1.00000000 Get it (8 decimal digit)!
Export Conv.Factor: 1.00000000
Import Partner Country: Origin Origin
Export Partner Country: Last Known Destination Last Known Destination
Trade Flow Specification: Imports Exports Re-Exports Re-Imports
Coverage:
Processing Notes:
Publication Notes
KB Link:
Publication Note: New data: Data for this year is published for the first time; the data was received from/through the Ministério do Desenvolvimento, Indústria e Comércio Exterior on 02 March 2017
Metadata Attachment
my tasks (13) | status 50/12/18 [42/97] on 2017-03-31T17:12:04 | Follow Up Bin | Where-is-it? @UN/OECD CoprA 2010



Annual Totals

Select National Data: Brazil(2016)H4 Go

Update

National data of

Country: Brazil

Year: 2016

Conversion Factor

	M	X	Last Update
	1.00000000	1.00000000	2017.03.31

TTrade/Comtrade/Country Totals

	M	X	RM	RX
TTrade (2016 /X):[1]	$142,267,225,799.00	$185,280,000,000.00	$0.00	$0.00
Comtrade (2015):[2]	$171,446,211,662.00	$191,126,885,834.00	$0.00	$0.00
[0]-Gold(USD):[5]	$0.00	$0.00	$0.00	$0.00
CountryTotals(NC):[0]	0.00	0.00	0.00	0.00
--data source				

Accumulated Totals

	M	X	RM	RX
Total (NC):	137,552,002,856.00	185,235,400,805.00	0.00	0.00
Gold (NC):	390.00	1,704.00	0.00	0.00
Total-Gold(NC):	137,552,002,466.00	185,235,399,101.00	0.00	0.00
Total(USD):	$137,552,002,856.00	$185,235,400,805.00	$0.00	$0.00
Gold(USD):	$390.00	$1,704.00	$0.00	$0.00
Total-Gold(USD):[3]	$137,552,002,466.00	$185,235,399,101.00	$0.00	$0.00

Reported Totals (or containing Totals calculated by UNSD)

	M	X	RM	RX
Reported (NC):	0.00	0.00	0.00	0.00
Reported (USD):	$0.00	$0.00	$0.00	$0.00
Total-Gold(USD):[4]	$137,552,002,466.00	$185,235,399,101.00	$0.00	$0.00

Note: if it is necessary, [4] is used to generate 99AT00

Differences to adopted total

	M	X	RM	RX
W/TTrade:[4]-[1]	($4,715,223,333.00) -3.31 %	($44,600,899.00) -0.02 %	$0.00 NaN	$0.00 NaN
W/Comtrade:[4]-[2]	($33,894,209,196.00) -19.77 %	($5,891,486,733.00) -3.08 %	$0.00 NaN	$0.00 NaN
W/CtyTotals:[4]-[5]	$0.00 Infinity	$0.00 Infinity	$0.00 NaN	$0.00 NaN
Expected 99AT00:[4]-[3]	$0.00 0.00 %	$0.00 0.00 %	$0.00 NaN	$0.00 NaN
Real 99AT00:	$0.00	$0.00	$0.00	$0.00

Memorandum Items

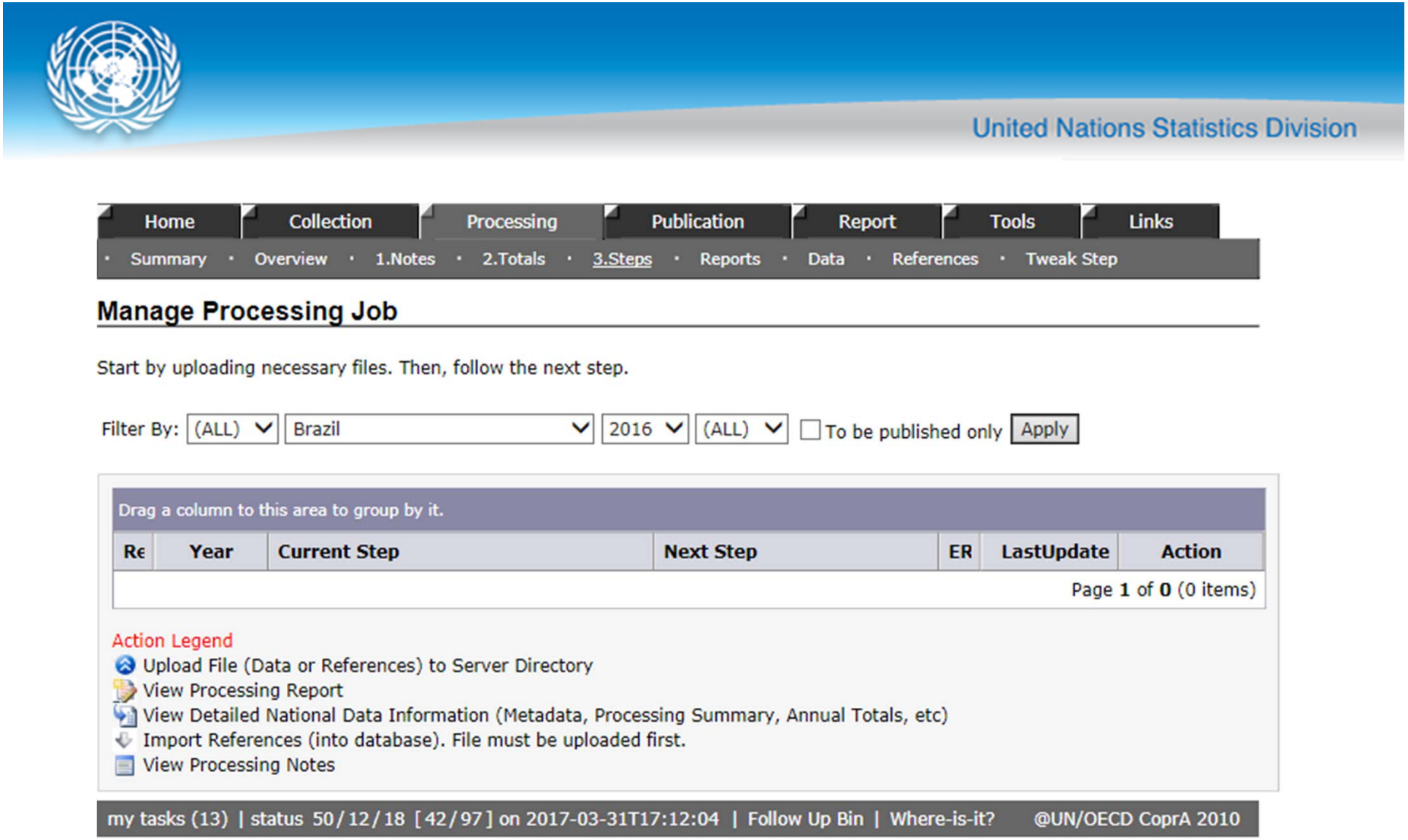
	M	X	RM	RX
Total	$1,017,194.00 0.00 %	$2,128,273,299.00 1.15 %	$0.00 NaN	$0.00 NaN
Assigned to C.01-97	$1,017,194.00	$3,304,486.00	$0.00	$0.00
Assigned to C.99	$0.00	$2,124,968,813.00	$0.00	$0.00

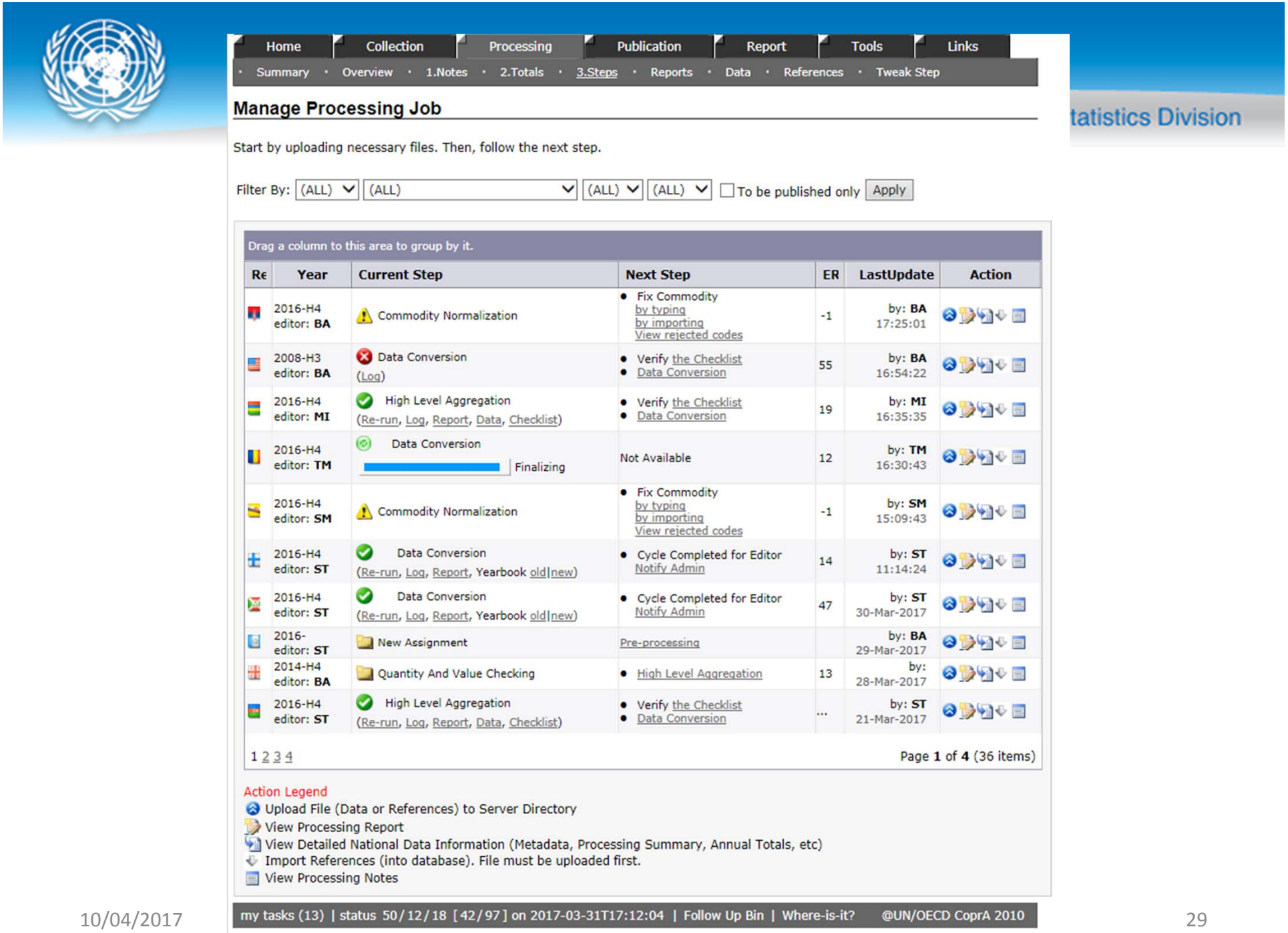
Special Partners (Tariff Line/Final)

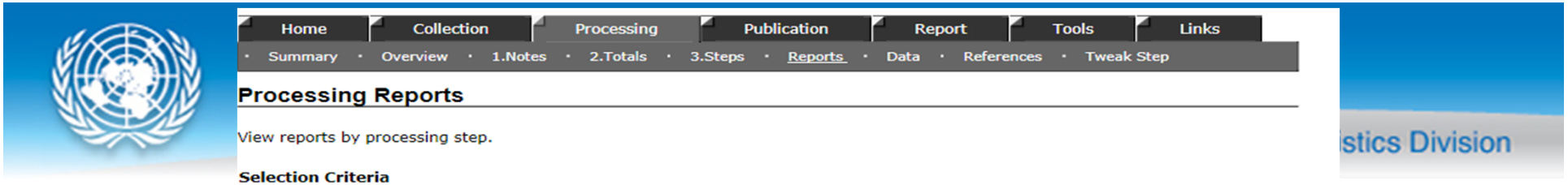
	M	X	RM	RX
839 and 899	0.00 %/0.00 %	0.00 %/0.00 %	NaN/NaN	NaN/NaN
Bunkers:837	0.00 %/0.00 %	0.93 %/0.93 %	NaN/NaN	NaN/NaN
Free Zones: 838	0.22 %/0.22 %	0.00 %/0.00 %	NaN/NaN	NaN/NaN
It Self	0.00 %/0.00 %	0.00 %/0.00 %	NaN/NaN	NaN/NaN

Totals Note

Explanation:


United Nations Statistics Division
Home
Collection
Processing
Publication
Report
Tools
Links
Summary
Overview
1.Notes
2.Totals
3.Steps
Reports
Data
References
Tweak Step
Manage Processing Job
Start by uploading necessary files. Then, follow the next step.
Filter By: (ALL) Brazil 2016 (ALL) To be published only Apply
Drag a column to this area to group by it.
Re
Year
Current Step
Next Step
ER
LastUpdate
Action
Page 1 of 0 (0 items)
Action Legend
Upload File (Data or References) to Server Directory
View Processing Report
View Detailed National Data Information (Metadata, Processing Summary, Annual Totals, etc)
Import References (into database). File must be uploaded first.
View Processing Notes
my tasks (13) | status 50/12/18 [42/97] on 2017-03-31T17:12:04 | Follow Up Bin | Where-is-it? @UN/OECD CoprA 2010


Home
Collection
Processing
Publication
Report
Tools
Links
Summary · Overview · 1.Notes · 2.Totals · 3.Steps · Reports · Data · References · Tweak Step
tatistics Division
Manage Processing Job
Start by uploading necessary files. Then, follow the next step.
Filter By: (ALL) (ALL) (ALL) (ALL) To be published only Apply
Drag a column to this area to group by it.
Re
Year
Current Step
Next Step
ER
LastUpdate
Action
2016-H4 editor: BA
Commodity Normalization
Fix Commodity by typing by importing View rejected codes
-1
by: BA 17:25:01
2008-H3 editor: BA
Data Conversion (Log)
Verify the Checklist
Data Conversion
55
by: BA 16:54:22
2016-H4 editor: MI
High Level Aggregation (Re-run, Log, Report, Data, Checklist)
Verify the Checklist
Data Conversion
19
by: MI 16:35:35
2016-H4 editor: TM
Data Conversion
Finalizing
Not Available
12
by: TM 16:30:43
2016-H4 editor: SM
Commodity Normalization
Fix Commodity by typing by importing View rejected codes
-1
by: SM 15:09:43
2016-H4 editor: ST
Data Conversion (Re-run, Log, Report, Yearbook old|new)
Cycle Completed for Editor
Notify Admin
14
by: ST 11:14:24
2016-H4 editor: ST
Data Conversion (Re-run, Log, Report, Yearbook old|new)
Cycle Completed for Editor
Notify Admin
47
by: ST 30-Mar-2017
2016- editor: ST
New Assignment
Pre-processing
by: BA 29-Mar-2017
2014-H4 editor: BA
Quantity And Value Checking
High Level Aggregation
13
by: 28-Mar-2017
2016-H4 editor: ST
High Level Aggregation (Re-run, Log, Report, Data, Checklist)
Verify the Checklist
Data Conversion
...
by: ST 21-Mar-2017
1 2 3 4
Page 1 of 4 (36 items)
Action Legend
Upload File (Data or References) to Server Directory
View Processing Report
View Detailed National Data Information (Metadata, Processing Summary, Annual Totals, etc)
Import References (into database). File must be uploaded first.
View Processing Notes
my tasks (13) | status 50/12/18 [42/97] on 2017-03-31T17:12:04 | Follow Up Bin | Where-is-it? @UN/OECD CoprA 2010



Processing Reports

View reports by processing step.

Selection Criteria

National Data of:

Step:

Report Type:

(ALL)
Data Import
Code Normalization
Completeness Check
Commodity Normalization
Basic Level Aggregation
Quantity & Value Checking
High Level Aggregation
Data Conversion
Data Publication

Report Tables

Brazil-2016 (Classification: H4 --

Printable Format

CONTENT REPORTS

DataImport

Table.1 Data Import Report: of records

Type	# of records	%	date
CMDLENGTH:8	2,792,077	100.00 %	2017-03-10T18:08:52
NON-BLANK:weight	2,524,101	90.40 %	2017-03-10T18:14:04
NON-BLANK:qty	1,144,568	40.99 %	2017-03-10T18:14:36
NON-BLANK:value	2,787,655	99.84 %	2017-03-10T18:15:00
FLOW:E	1,092,225	39.12 %	2017-03-10T18:09:28
FLOW:imp	1,699,852	60.88 %	2017-03-10T18:09:28
COMP:weight=qty(8)	0	0.00 %	2017-03-10T18:15:23
COMP:weight=qty(5)	0	0.00 %	2017-03-10T18:15:49
QU:11	1,059,373	37.94 %	2017-03-10T18:10:19
QU:12	434	0.02 %	2017-03-10T18:10:19
QU:13	28,023	1.00 %	2017-03-10T18:10:19
QU:14	43	0.00 %	2017-03-10T18:10:19
QU:15	18,990	0.68 %	2017-03-10T18:10:19
QU:16	14,872	0.53 %	2017-03-10T18:10:19
QU:17	7,488	0.27 %	2017-03-10T18:10:19
QU:18	15	0.00 %	2017-03-10T18:10:19
QU:19	1,741	0.06 %	2017-03-10T18:10:19
QU:20	494	0.02 %	2017-03-10T18:10:19
QU:21	12,861	0.46 %	2017-03-10T18:10:19
QU:22	234	0.01 %	2017-03-10T18:10:19

Table.2 Data Import Report: of uniqueness

Type	# of unique values	date	
UNIQUE:qtyunit	13	2017-03-10T18:11:32	View
UNIQUE:partner	246	2017-03-10T18:12:09	View
UNIQUE:flow	2	2017-03-10T18:12:52	View
UNIQUE:commodity	8,987	2017-03-10T18:13:29	View

Table.3 Data Source Info

Type	Value	date
FileSource	\\dfs-desast-54\Trade\ProcessingDataSources\Brazil (76)\2016\Brazil2016 (2017_03_10_TM_OR_ND).mdb	2017-03-10T18:17:13
TableName	Import_export_2016	2017-03-

Standardization and Harmonization

Code and Commodity Normalization

- 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 and Weight in UN Comtrade

Objectives

- Standardize quantity and weight information according to WCO recommended units in order to provide users with comparable and comprehensive information as possible
- Estimation will be done at the most detailed reported level (tariff line data)

Estimation Methodology

- Mathematical conversion (pounds to kg)
- Estimation using FAO conversion factors for selected commodities (i.e., 1 l of heavy cream = 1.02 kg)
- Estimation using partially reported quantity and/or net weight
- Estimation using standard unit values
- Net weight can also be estimated using the same approach

Use of country provided information

- Quantity conversion using mathematical conversion
- Net weight is copied into the supplementary quantity, and vice-versa, in cases where only one of them is available
- FAO conversion factors for selected commodities applied during code normalization

Weighted Unit Value

Use partially reported quantity in the same commodity-flow

New Quantity = Value / Partially Reported Unit Value

Original Data							
HS Code	Description		Value	Quantity	Unit	WCO RU	
080810	Fresh Apple		100	50	5 (no of items)	8 (kilogram)	
080810	Fresh Apple		400	10	8 (kilogram)	8 (kilogram)	
080810	Fresh Apple		100	1	11 (dozen)	8 (kilogram)	

Original Data with share							
HS Code	Description	Value Share	Value	Quantity	Unit	WCO RU	UV
080810	Fresh Apple	**33.33%**	200	**62**	**5 (no of items)**	8 (kilogram)	**3.2**
080810	Fresh Apple	**66.66%**	400	10	8 (kilogram)	8 (kilogram)	**40**

Final data (estimated record is highlighted)							
HS Code	Description		Value	Quantity	Unit	WCO RU	UV
080810	Fresh Apple		200	***5***	***8 (kilogram)***	8 (kilogram)	***40***
080810	Fresh Apple		400	10	8 (kilogram)	8 (kilogram)	40

Standard Unit Value

Use standard unit value in the same commodity-flow

New Quantity = Value / Standard Unit Value

Original Data

HS Code	Description		Value	Quantity	Unit	WCO RU	
080810	Fresh Apple		**1000**	50	5 (no of items)	8 (kilogram)	
080810	Fresh Apple		400	10	8 (kilogram)	8 (kilogram)	
080810	Fresh Apple		100	1	11 (dozen)	8 (kilogram)	

HS Code	Description	Value Share	Value	Quantity	Unit	WCO RU	UV
080810	Fresh Apple	**73.33%**	1100	**62**	**5 (no of items)**	8 (kilogram)	**17.74**
080810	Fresh Apple	**26.66%**	400	10	8 (kilogram)	8 (kilogram)	40

Standard Unit Value for HS Code 080810

HS Code	Description	Flow	Unit	SUV	Low UV	High UV
080810	Fresh Apple	Imports	8 (kilogram)	0.898	0.181	3.689
080810	Fresh Apple	Exports	8 (kilogram)	0.655	0.249	1.863

Final data (estimated record is highlighted)

HS Code	Description		Value	Quantity	Unit	WCO RU	UV
080810	Fresh Apple		1100	***1225***	***8 (kilogram)***	8 (kilogram)	***0.898***
080810	Fresh Apple		400	10	8 (kilogram)	8 (kilogram)	40

Extreme quantity/weight and value data

Objective: to detect and remove extreme outliers in quantity

Three checks:

- Historical Unit Value Check
- Standard Unit Value Check
- Historical Value Check

Historical Unit Value (UV) Check

Quantity and weight unit values are considered extreme if

- UV > maximum historical unit value * 10, or
- UV < minimum historical unit value / 10
  - *Threshold: commodity flows above 0.005% of total export or import are considered*
  - *Historical unit values are the unit values of the last five years of the reporting country*

Example

CmdFlow	Unit	SUV	5UV	5V	Ratio	SUV
850423-2	5	33437.86<1.53<5195565.79			21,815.40	444,076.77
841861-1	80	3.99<0.00<62.16			4,517.88	15.71
151219-2	8	0.61<2168.77<2.05			1,055.98	1.09
710811-2	8	3441.01<10.42<139938.28			330.34	24,655.70
846019-2	80	1.68<21300.19<81.20			262.32	9.50
732111-1	5	9.64<0.04<1692.58	2.00<0.04<9.06		233.99	128.86

Outlier detected for imports commodity 732111, quantity provided in number of units (5)			
Calculated UV	0.04		
Minimum UV	9.64	Maximum UV	1692.58
Minimum UV / 10	0.9642	Maximum UV * 10	16,925.8
Is SUV < (Min UV / 10)	**Yes**	Is SUV > (Max UV * 10)	No

Standard Unit Value (SUV) Check

Quantity and weight unit values are considered extreme if

- UV > upper limit of standard unit value * 2, or
- UV < lower limit of standard unit value / 2
  - *Threshold: commodity flows greater than $50,000*

Example

CmdFlow	Unit	SUV	5UV	5V	Ratio	SUV
850423-2	5	33437.86<1.53<5195565.79			21,815.40	444,076.77
841861-1	80	3.99<0.00<62.16			4,517.88	15.71
151219-2	8	0.61<2168.77<2.05			1,055.98	1.09
710811-2	8	3441.01<10.42<139938.28			330.34	24,655.70
846019-2	80	1.68<21300.19<81.20			262.32	9.50
732111-1	5	9.64<0.04<1692.58	2.00<0.04<9.06		233.99	128.86

Outlier detected for exports commodity 850423, quantity provided in number of units (5)			
Calculated UV	1.53		
Minimum UV	33,437.86	Maximum UV	5,195,565.79
Minimum UV / 2	16,718.93	Maximum UV * 2	10,391,131.58
Is SUV < (Min UV / 2)	Yes	Is SUV > (Max UV * 2)	No

Historical Value Check

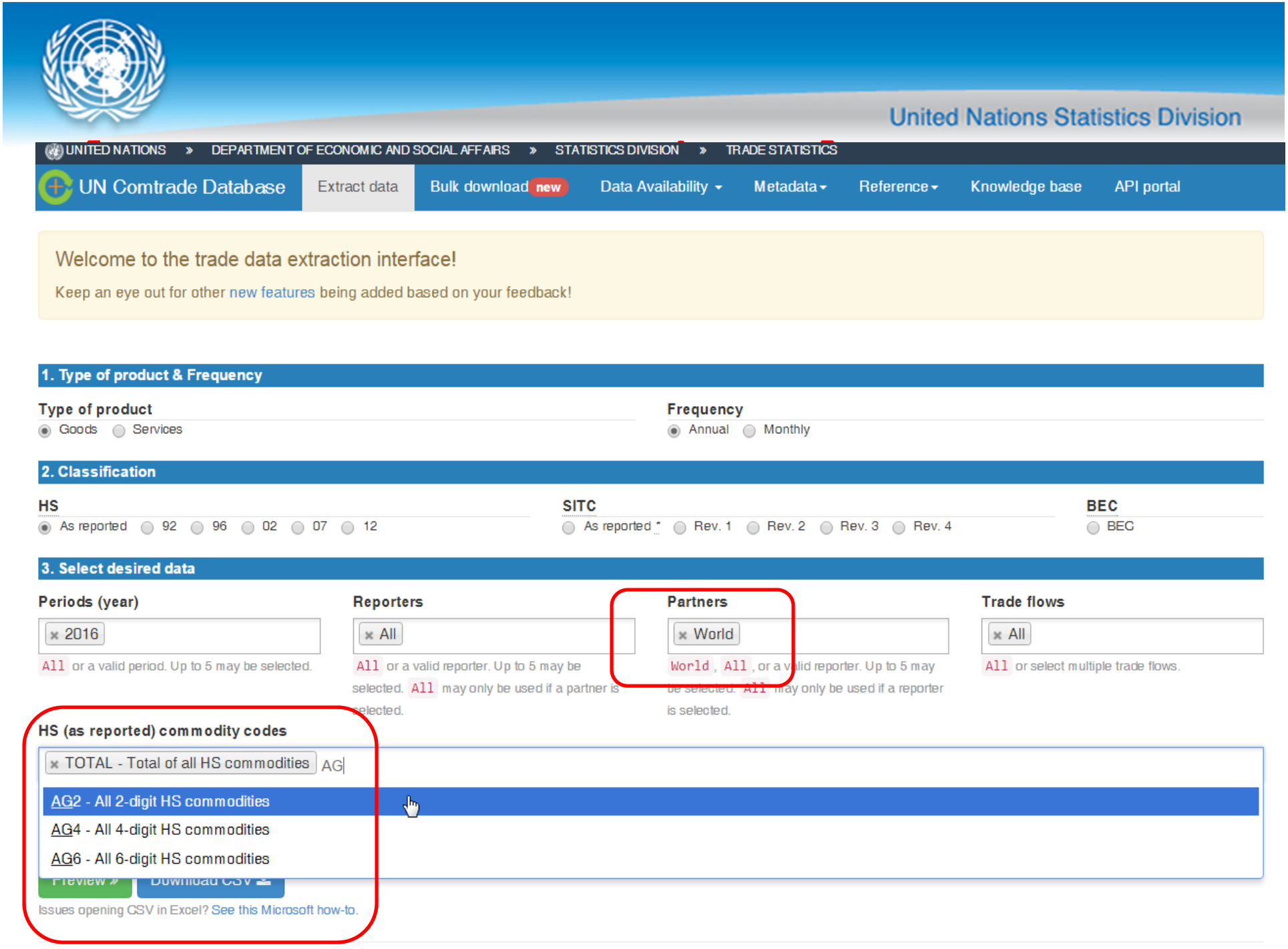
Values are considered extreme if

- Value > maximum historical value * 100, or
- Value < minimum historical value / 100
  - *Threshold: commodity flows greater than $1,000,000*
  - *Historical values are the values of the last five years of a particular country.*

Data Aggregation

Data Aggregation in UN Comtrade

- We provide pre-aggregated data for our users
  - HS: 2, 4, 6 digit
  - SITC: 1, 3, 5 digit
  - BEC: 1, 2, 3 digit
- Partner “World”
- Commodity “Total”
- Future: add pre-aggregations on IMTS 2010 (extended data sets)


United Nations Statistics Division
UNITED NATIONS » DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS » STATISTICS DIVISION » TRADE STATISTICS
UN Comtrade Database
Extract data
Bulk download new
Data Availability
Metadata
Reference
Knowledge base
API portal
Welcome to the trade data extraction interface!
Keep an eye out for other new features being added based on your feedback!
1. Type of product & Frequency
Type of product
Goods Services
Frequency
Annual Monthly
2. Classification
HS
As reported 92 96 02 07 12
SITC
As reported * Rev. 1 Rev. 2 Rev. 3 Rev. 4
BEC
BEC
3. Select desired data
Periods (year)
× 2016
All or a valid period. Up to 5 may be selected.
Reporters
× All
All or a valid reporter. Up to 5 may be selected. All may only be used if a partner is selected.
Partners
× World
World, All, or a valid reporter. Up to 5 may be selected. All may only be used if a reporter is selected.
Trade flows
× All
All or select multiple trade flows.
HS (as reported) commodity codes
× TOTAL - Total of all HS commodities AG
AG2 - All 2-digit HS commodities
AG4 - All 4-digit HS commodities
AG6 - All 6-digit HS commodities
Preview »
Download CSV
Issues opening CSV in Excel? See this Microsoft how-to.

Example

1. Type of product & Frequency

Type of product
(•) Goods ( ) Services

Frequency
(•) Annual ( ) Monthly

2. Classification

HS
(•) As reported ( ) 92 ( ) 96 ( ) 02 ( ) 07 ( ) 12

SITC
( ) As reported * ( ) Rev. 1 ( ) Rev. 2 ( ) Rev. 3 ( ) Rev. 4

BEC
( ) BEC

3. Select desired data

Periods (year)
× 2015
`All` or a valid period. Up to 5 may be selected.

Reporters
× Maldives
`All` or a valid reporter. Up to 5 may be selected. `All` may only be used if a partner is selected.

Partners
× World
`World`, `All`, or a valid reporter. Up to 5 may be selected. `All` may only be used if a reporter is selected.

Trade flows
× All
`All` or select multiple trade flows.

HS (as reported) commodity codes
× AG2 - All 2-digit HS commodities
`All`, `Total`, `AG[X]` or a valid code. Up to 20 may be selected. If you know the code number, e.g. `01 - Live animals`, type `01`. To search by description type a word, e.g. `rice`.

4. See the results

Preview » Download CSV

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

5. Preview (109 records)

Show 25 entries

Period	Trade Flow	Reporter	Partner	Commodity Code	Trade Value (US$)	Netweight (kg)	Qty Unit	Qty	Flag
2015	Import	Maldives	World	01	$207,064	0	No Quantity	0	0
2015	Import	Maldives	World	02	$48,341,449	0	No Quantity	0	0
2015	Import	Maldives	World	03	$22,079,876	0	No Quantity	0	0
2015	Export	Maldives	World	03	$126,081,807	0	No Quantity	0	0
2015	Import	Maldives	World	04	$51,333,963	0	No Quantity	0	0
2015	Import	Maldives	World	05	$171,117	0	No Quantity	0	0
2015	Import	Maldives	World	06	$1,289,943	0	No Quantity	0	0
2015	Import	Maldives	World	07	$43,895,803	0	No Quantity	0	0
2015	Import	Maldives	World	08	$42,853,856	0	No Quantity	0	0
2015	Export	Maldives	World	08	$16,590	0	No Quantity	0	0
2015	Import	Maldives	World	09	$11,751,130	0	No Quantity	0	0
2015	Import	Maldives	World	10	$20,532,939	0	No Quantity	0	0
2015	Import	Maldives	World	11	$12,284,127	0	No Quantity	0	0
2015	Import	Maldives	World	12	$1,347,467	0	No Quantity	0	0
2015	Import	Maldives	World	13	$163,179	0	No Quantity	0	0
2015	Import	Maldives	World	14	$335,987	0	No Quantity	0	0
2015	Import	Maldives	World	15	$10,317,232	0	No Quantity	0	0
2015	Import	Maldives	World	16	$10,358,806	0	No Quantity	0	0
2015	Export	Maldives	World	16	$13,915,443	0	No Quantity	0	0
2015	Import	Maldives	World	17	$10,988,772	0	No Quantity	0	0
2015	Import	Maldives	World	18	$10,123,875	0	No Quantity	0	0
2015	Import	Maldives	World	19	$26,405,312	0	No Quantity	0	0
2015	Import	Maldives	World	20	$21,141,540	0	No Quantity	0	0
2015	Import	Maldives	World	21	$24,024,615	0	No Quantity	0	0
2015	Export	Maldives	World	21	$5	0	No Quantity	0	0

Showing 1 to 25 of 109 entries

First Previous 1 2 3 4 5 Next Last

Conversion of Commodity Classification

Data Conversion Objectives

- To provide comparable time series data in UN Comtrade for all countries in specific classification (i.e., HS 1988)
- 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
- https://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm

Illustration: relationships

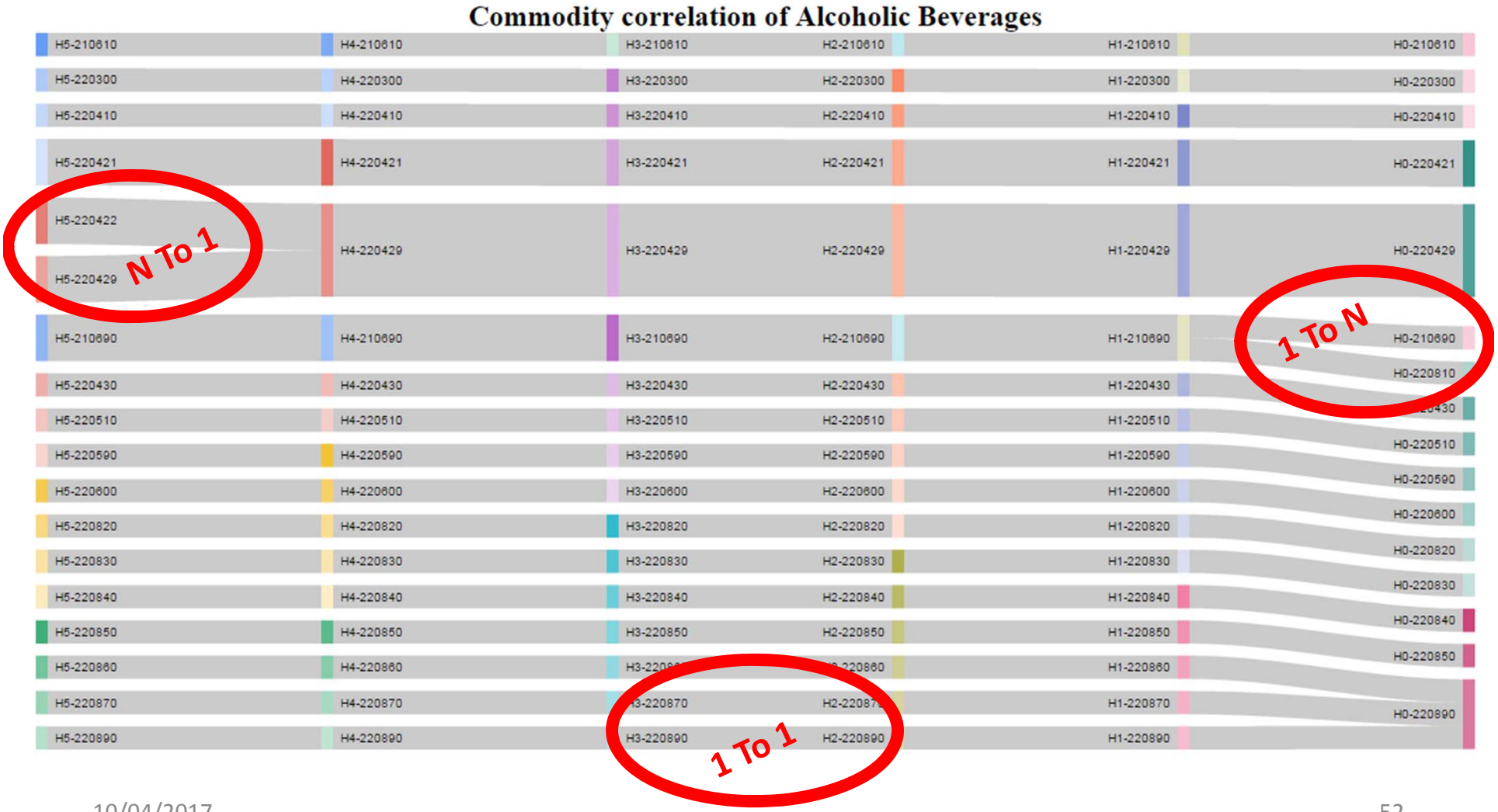
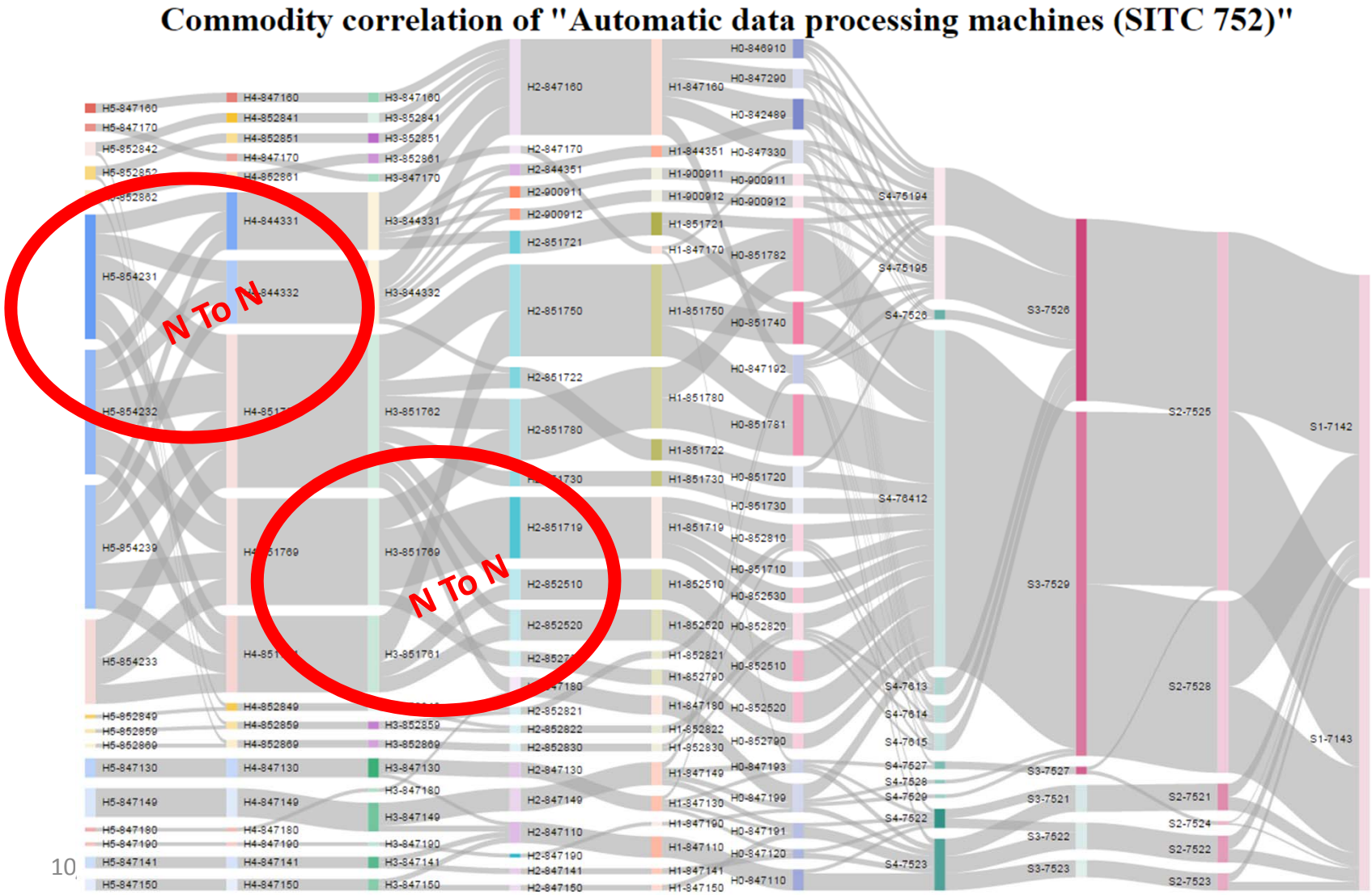


Illustration: relationships



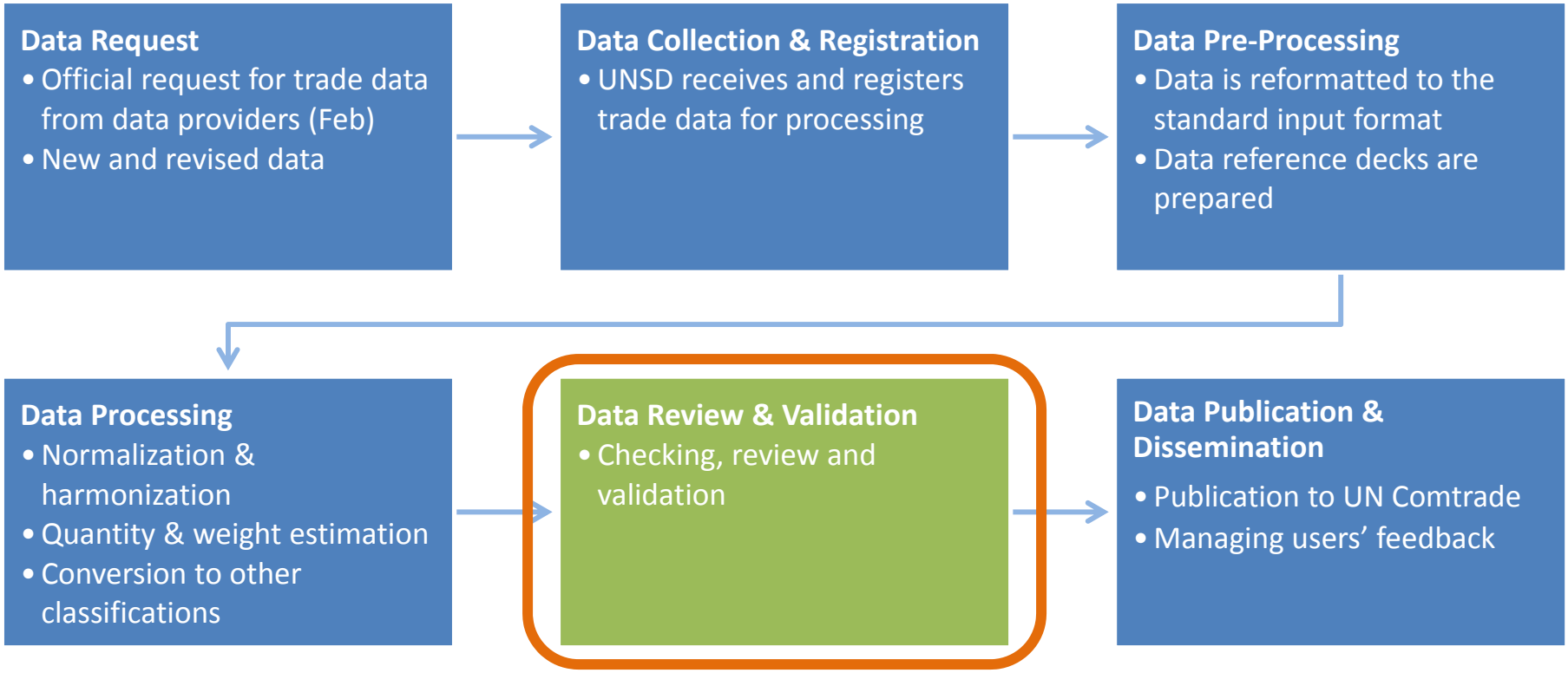
Method of conversion

- Our practice is to convert most detailed level in original classification into one single code in the target classification, without splitting.
- The target code could be at detailed level or at higher level in aggregation structure
- We flatten 1:n and n:n to 1:1 and/or n:1 relationships

Correlation and Conversion of "Automatic data processing machines (SITC 752)"

The complex relationships (left) are simplified (right)

Production Process: Data Review



Data Review and Validation

- Consolidated report (PDF) contains result of each processing phase
- Checklist
- Reviews both by editor and supervisor
- Reporting services

Sample Data Processing Reports

DataImport

Table.1 Data Import Report: of records

Type	# of records	%
CMDLENGTH:8	300,775	100.00 %
NON-BLANK:weight	300,775	100.00 %
NON-BLANK:qty	121,683	40.46 %
NON-BLANK:value	300,775	100.00 %
FLOW:1	239,980	79.79 %
FLOW:2	60,795	20.21 %
COMP:weight=qty(8)	0	0.00 %
COMP:weight=qty(5)	0	0.00 %
QU:B	107	0.04 %
QU:C	65	0.02 %
QU:F	19	0.01 %
QU:G	768	0.26 %
QU:K	144	0.05 %
QU:L	1,217	0.40 %
QU:M	223	0.07 %
QU:P	5,977	1.99 %
QU:S	109,475	36.40 %

- **Data Import: Table1**
  - In CMDLENGTH, shows first indication of the existence of Reported High Level Data (any length less than 6).
  - Length is usually even-numbered, but in some cases, odd-numbered. Possible cause is missing first digit '0' in the comm. code. Re-check if odd-numbered codes reported.
  - Check missing flow
  - Check unusual record counts for Quantity unit

Table.4 HS Commodity Rejection

Version	# of rejections
H0	741
H1	504
H2	238
H3	2

- **Data Import: Table4**
  - Shows the no. of rejected codes for each classification
  - If no. of rejection between classifications do not differ that much, further investigation is required

Sample Data Processing Reports

United Nations Statistics Division

Table.7 the Checklist

Name	M	X	RX	RM	
Currency*	NOK	NOK	NOK	NOK	
Conversion Factor	0.19061767	0.19112836	0.19112836	0.19061767	
Length of Original Commodity	8	8			
#of records in original table	239980	60795			
#of records in tariff line table	239980	60795			
#of records in final table w/o areas	129412	83693			
#of records in final table areas only	46156	25447			
Accumulated Total in LC	495,686,042,620.00	929,053,001,431.00			
Accumulated Total in USD 1	94,486,518,495.75	177,568,376,516.58			
Accumulated Total in USD 2	94,486,518,495.74	177,568,376,516.58			
Official Total	94,486,518,495.75	177,568,376,516.58	0.00	0.00	
Total Trade	79,777,826,078.40	137,975,150,546.85			
Comtrade Total	80,297,682,385.00	136,357,020,656.00			
% Diff with Total Trade	18.44%	28.70%			
% Diff with Comtrade Total	17.67%	30.22%			
99AT00 Value					
9999 Value	461,749,402.84	6,795,638,420.65			
999999 Value	461,749,402.84	6,795,638,420.65			
# of Memorandum Items	1	1			
# of Reported Data					
ER Index*	5	5	5	5	
%SUV Applied for Qty/Weight in QV*	0.50%	0.50%	0.50%	0.50%	
# of records rejected by SUV	9	36			
# of records rejected by 5UV	2	1			
# of records rejected by 5V	1	2			
Accumulated Gold in Original Data					
Accumulated Gold in Tariff Line Data	0.00	0.00	0.00		
# of commodity wrongly assigned as MI					
# of non standard					

- Table7: Checklist
  - **Summary of previous reports**

Sample Data Processing Reports

DataConversion

Table.A World vs Sum of Partners

Flow	World	Sum Of Partners	Diff %	Diff Value	date
1	$611,364,435,458.00	$611,364,435,464.00	0.00 %	$6.00	2009-02-04T12:10:37
2	$539,730,711,536.00	$539,730,711,545.00	0.00 %	$9.00	2009-02-04T12:10:37
4	$6,681,094,493.00	$6,681,094,493.00	0.00 %	$0.00	2009-02-04T12:10:37

Table.5. Quantity Unit Comparison

Code		# of Recs.A	# of Recs.B	%share.A	%share.B	date
0	Total	547,861	547,166	100.00 %	100.00 %	2009-02-04T12:10:37
80	Netweight	489,081	497,864	89.27 %	90.99 %	2009-02-04T12:10:37
1	No Quantity	91,138	91,349	16.64 %	16.69 %	2009-02-04T12:10:37
2	Area in square metres	4,211	4,207	0.77 %	0.77 %	2009-02-04T12:10:37
3	Electrical energy in thousands of kilowatt-hours	30	30	0.01 %	0.01 %	2009-02-04T12:10:37
4	Length in metres	508	508	0.09 %	0.09 %	2009-02-04T12:10:37
5	Number of items	90,998	90,881	16.61 %	16.61 %	2009-02-04T12:10:37
6	Number of pairs	3,887	3,879	0.71 %	0.71 %	2009-02-04T12:10:37
7	Volume in liters	4,665	4,398	0.85 %	0.80 %	2009-02-04T12:10:37
8	Weight in kilograms	350,270	349,915	63.93 %	63.95 %	2009-02-04T12:10:37
9	Thousands of items	37	37	0.01 %	0.01 %	2009-02-04T12:10:37
10	Number of packages	79	79	0.01 %	0.01 %	2009-02-04T12:10:37
12	Volume in cubic meters	1,693	1,692	0.31 %	0.31 %	2009-02-04T12:10:37
13	Weight in carats	345	191	0.06 %	0.03 %	2009-02-04T12:10:37

- **TableA: World vs. Sum of Partners**
  - **Make sure that differences in value between World and Sum of Partners is zero (or very minimal)**
- **Table5: Quantity Unit Comparison**
  - **Make sure that comparison between quantity units used do not differ that much in terms of %share**

Sample Data Processing Reports

Table.6. Partner Value Comparison

Flow	Partner		Value.A	Value.B	%diff	diff	%share
1	0	AAA	$611,364,435,458.00	$611,364,435,458.00	0.00 %	$0.00	100.00 %
1	4	AFG	$2,506,231.00	$2,506,231.00	0.00 %	$0.00	0.00 %
1	8	ALB	$23,238,224.00	$23,238,224.00	0.00 %	$0.00	0.00 %
1	10	ANT	$16,470.00	$16,470.00	0.00 %	$0.00	0.00 %
1	12	DZA	$4,627,196,986.00	$4,627,196,986.00	0.00 %	$0.00	0.76 %
1	16	ASM	$32,941.00	$32,941.00	0.00 %	$0.00	0.00 %
1	20	AND	$20,263,962.00	$20,263,962.00	0.00 %	$0.00	0.00 %
1	24	AGO	$2,370,503,388.00	$2,370,503,388.00	0.00 %	$0.00	0.39 %
1	28	ATG	$2,972,890.00	$2,972,890.00	0.00 %	$0.00	0.00 %
1	31	AZE	$2,260,091,970.00	$2,260,091,970.00	0.00 %	$0.00	0.37 %
1	32	ARG	$807,654,225.00	$807,654,225.00	0.00 %	$0.00	0.13 %
1	36	AUS	$1,610,338,536.00	$1,610,338,536.00	0.00 %	$0.00	0.26 %
1	40	AUT	$5,339,687,175.00	$5,339,687,175.00	0.00 %	$0.00	0.87 %

Table.8. Commodity Value Comparison

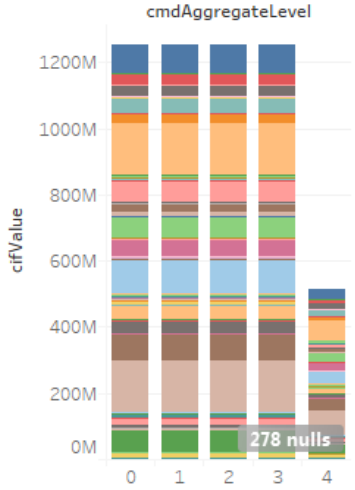
Flow	Code		Value.A	Value.B	%diff	diff	%share
1	TOTAL	All Commodities	$611,364,435,458.00	$611,364,435,458.00	0.00 %	$0.00	100.00 %
1	01	Live animals; animal prod	$352,104,872.00	$352,104,872.00	0.00 %	$0.00	0.06 %
1	02	Meat and edible meat offa	$4,818,724,643.00	$4,818,724,643.00	0.00 %	$0.00	0.79 %
1	03	Fish and crustaceans, mol	$4,164,666,970.00	$4,164,666,970.00	0.00 %	$0.00	0.68 %
1	04	Dairy produce; birds' egg	$3,463,250,314.00	$3,463,250,314.00	0.00 %	$0.00	0.57 %
1	05	Products of animal origin	$316,305,297.00	$316,305,297.00	0.00 %	$0.00	0.05 %
1	06	Live trees and other plan	$1,429,273,637.00	$1,429,273,637.00	0.00 %	$0.00	0.23 %
1	07	Edible vegetables and cer	$3,023,608,529.00	$3,023,608,529.00	0.00 %	$0.00	0.49 %
1	08	Edible fruit and nuts; pe	$4,163,397,384.00	$4,163,397,384.00	0.00 %	$0.00	0.68 %
1	09	Coffee, tea, maté and spi	$1,388,285,917.00	$1,388,285,917.00	0.00 %	$0.00	0.23 %
1	10	Cereals	$910,063,821.00	$910,063,821.00	0.00 %	$0.00	0.15 %
1	11	Products of the milling i	$380,426,930.00	$380,426,930.00	0.00 %	$0.00	0.06 %
1	12	Oil seeds and oleaginous	$862,062,497.00	$862,062,497.00	0.00 %	$0.00	0.14 %

- **Table6: Partner Value Comparison**
  - **Check major differences by partners**
- **Table8: Commodity Value Comparison**
  - **Check major differences by commodities**

Data Visualization

United Nations Statistics Division

CIF by partner



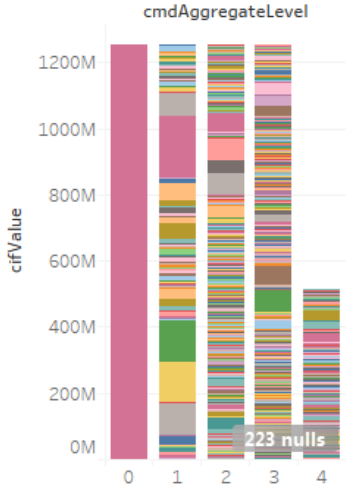
Customs

flowDescription		
customCod..	Export	Import
All customs..	9,198,271,154	11,033,72

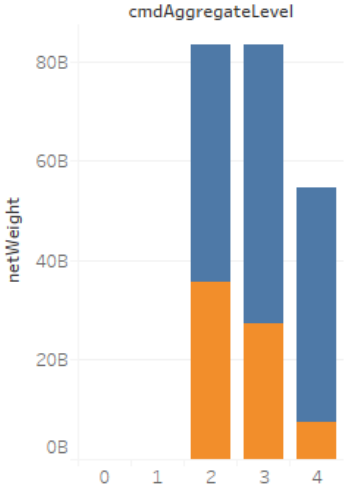
Mode

flowDescription		
mosDescrip..	Export	Import
All Modes o..	9,198,271,154	11,033,72

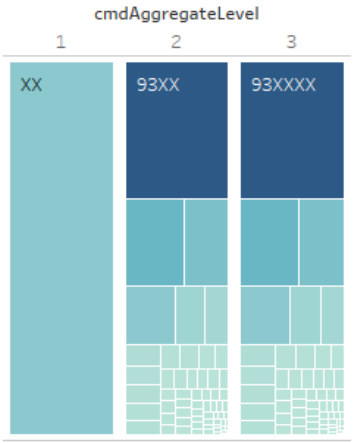
CIF by Commodity



Netweight



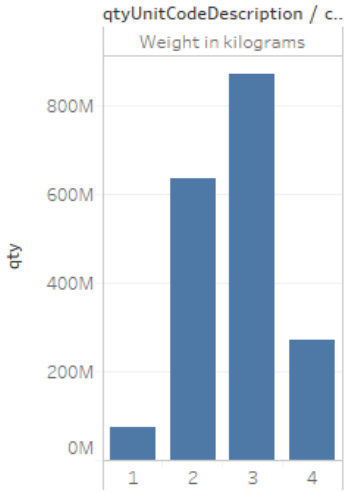
Residuals 6Digit



Qty

	cmdAggregateLevel				
qtyUnitCod..	0	1	2	3	4
All		254,370	551,552,794	967,036,962	310,245,370
Area in squ..	1,199	715,989	4,656,464	4,679,545	2,521,450
Electrical e..			1,073,535	1,073,535	
No quantity					
Number of i..	180	4,754,887	39,852,736	59,225,611	29,230,302
Number of ..		314,717	2,799,242	2,803,931	2,357,353
Thousand o..			1,938	1,939	
Volume in c..	63	51,621	4,141,130	4,162,208	348,380
Volume in li..	14,100	13,468,590	28,473,466	28,522,155	3,650,989
Weight in ki..		75,589,599	634,953,469	868,866,731	272,136,896

QtyWeight



Data Review

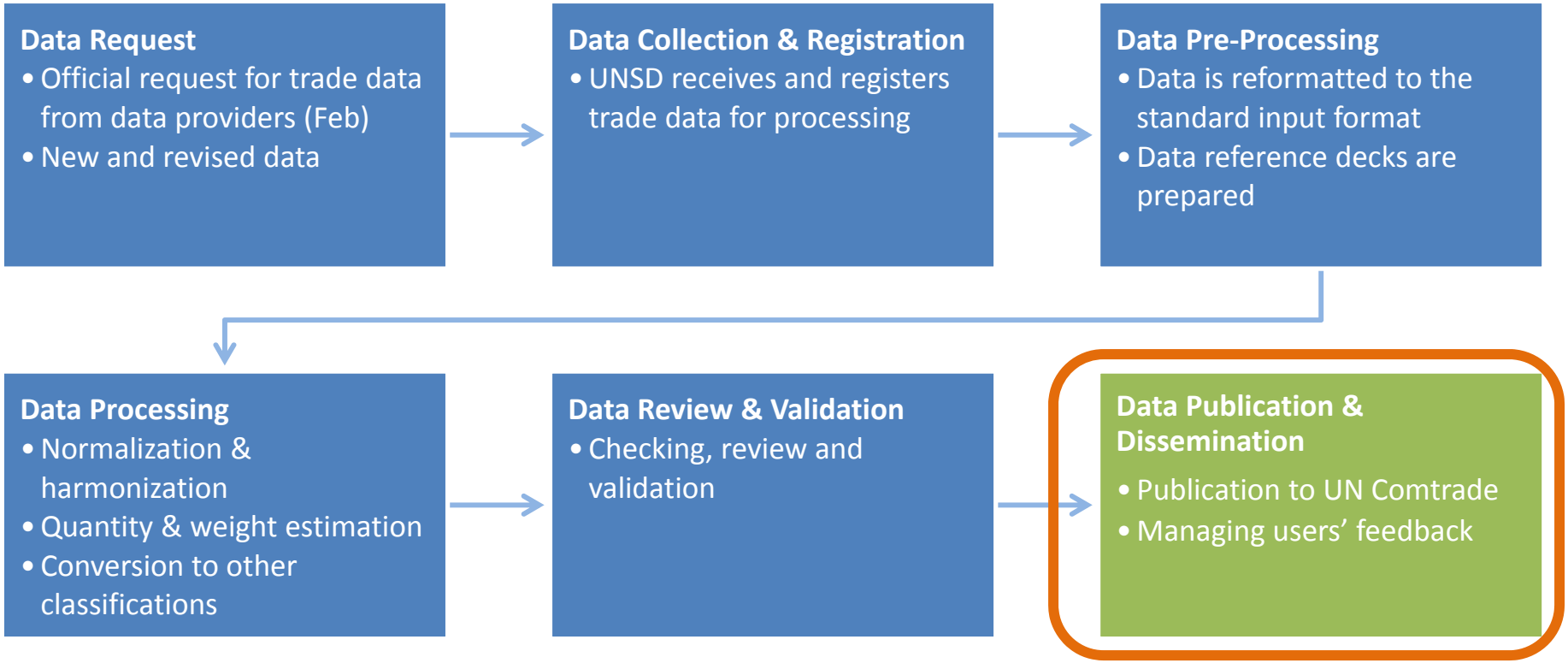
- Check overall totals against official sources, time series, trends
- Check level of memorandum items (overall value of non-standard commodity codes)
- Verify the mapping of partners, commodities and quantity units are correct
- Check integrity reports (possible wrong classification, conversion factors 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)

Management Reports

- Supplementary reports available through reporting services

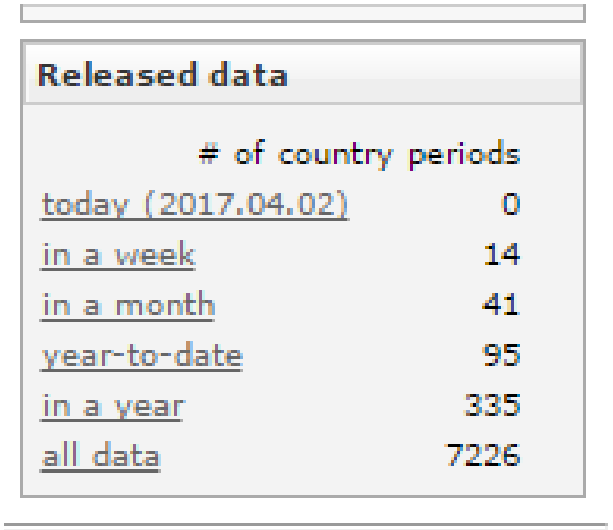
Copra - Management - Combined Processing Status
Copra - Management - Quantity Information
Copra - Management - User Performance
Copra - References - Partner Comparison
Copra - References - Partner Lookup
Copra - References - QtyUnit Comparison
Copra - References - QtyUnit Lookup
Copra_-_Management_-_Data_Release_Failure
CoprA - Management - Estimation Info
CoprA - Management - Historical Processing Overview
Copra - Management - In Processing
Copra - Management - Integrity Checks
Copra - Management - Missing Archived Reports
Copra - Management - Processing Stage
Copra - Management - Processing Status
Copra - Management - Qty Outliers
Copra - Release Checks - Commodity
Copra - Release Checks - Partner
Copra - Release Checks - Top Commodity Partner Matrix
Copra - Release Checks - Total Comparison

Production Process: Data Release



Data Release and Publication

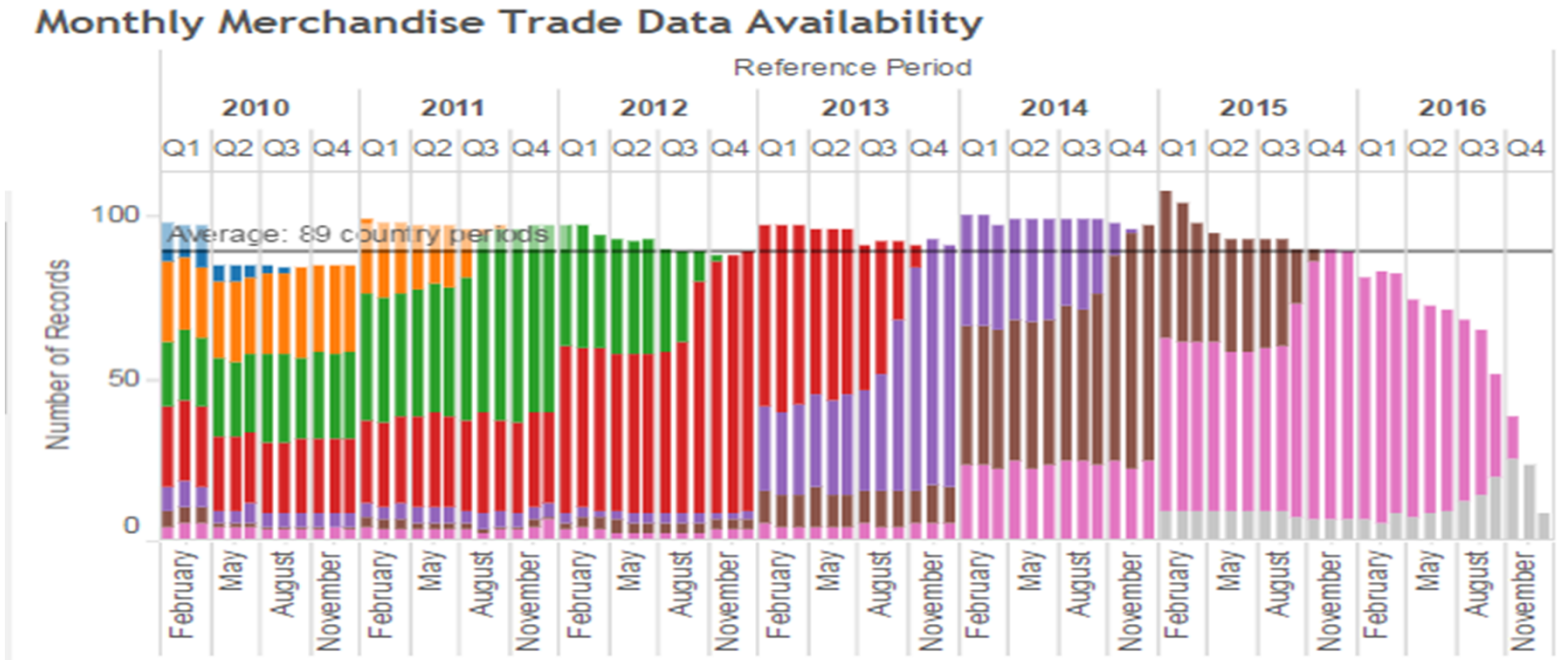
- Ongoing and continuous release and publication of datasets
- Data availability
  - www.comtrade.un.org/data/da
- User support mailbox
  - comtrade@un.org

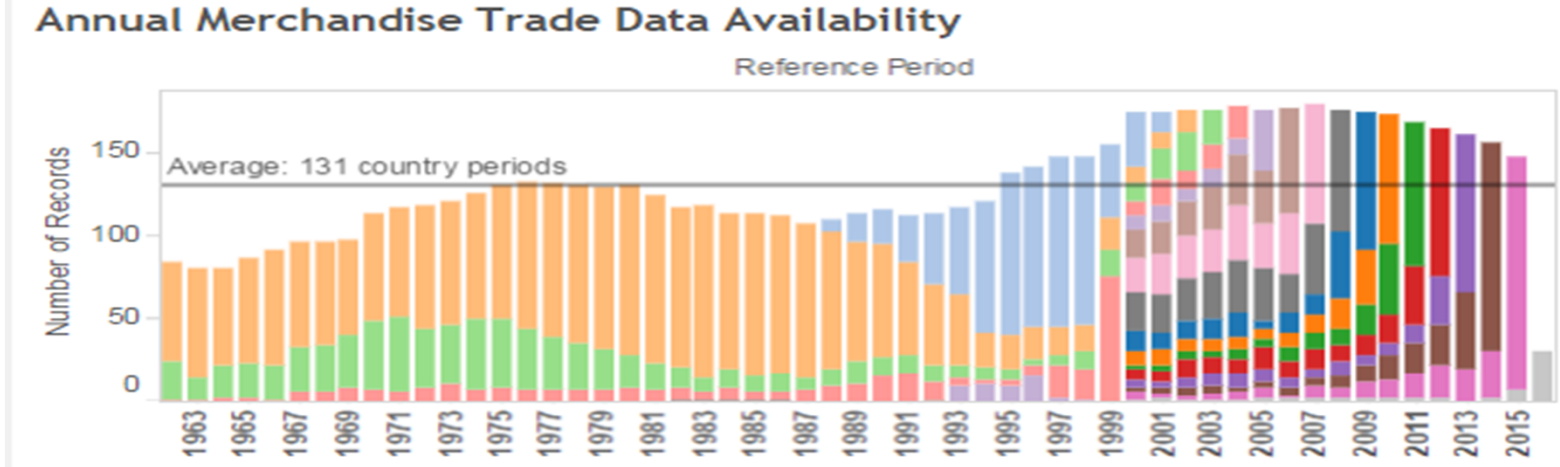


Released data

	# of country periods
today (2017.04.02)	0
in a week	14
in a month	41
year-to-date	95
in a year	335
all data	7226

Data availability in UN Comtrade





Thank you