Subcategories do not add up to higher level codes

Subcategories do not add up to higher level codes (Part I)

Q: In some cases, why the sum of detailed HS 6 digit level by partners are the same with TOTAL by partners? See example of Russian Federation 2009

<table>
<thead>
<tr>
<th>YEAR</th>
<th>FLOW</th>
<th>COUNTRY</th>
<th>PARTNER</th>
<th>CODE</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-490110</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-490199</td>
<td></td>
<td>1,168</td>
</tr>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-490290</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-491000</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-970110</td>
<td></td>
<td>669</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sum</td>
<td>1,892</td>
</tr>
<tr>
<td>2009</td>
<td>Import</td>
<td>Russian Federation - nes</td>
<td>H2-TOTAL</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

A: This is due to the existence of high-level reported data in certain data sets (in addition to the regular detailed trade data). The combination of detailed reported data and high-level reported data are used by the data provider to report trade data as much as possible while still complying with the confidentiality rules. Certain records at a detailed level may be suppressed due to confidentiality. However, at the aggregated level, these records are no longer confidential. The example above shows that partner distribution for certain HS 6 digit codes are considered confidential, however, not at the aggregated level.

Q: I have the necessary software and hardware to deal with the data. I would like to know, though, if it is possible to find cases where subcategories of SITC do not add up the next higher category level.

A: Yes, it is possible. The table below illustrates Brazil's exports in 2006 in SITC code 676 (Iron and steel bars, rods, angles, shapes and sections) with all available detail at a 4-digit level. The reasons why this can happen are:

1. that the data are collected and reported by Brazil in the HS-2002 classification and
2. that the correlation between the HS-2002 codes and the SITC, Rev.3 codes is such that only for a sub-set of codes a good correlation can be found at the detailed level of SITC; all other HS-2002 codes in this area can only be allocated to the aggregated 3-digit code 676.

The result is – as is shown in the table – that the full amount of trade value for code 676 is correctly reflected and that only for some 4-digit 676 x codes trade can be reproduced.
Subcategories do not add up to higher level codes (Part II)

Values on different levels of HS Classification

Q: We have noticed that, in the Comtrade HS database, total values for a particular chapter do not always correspond to sum of values at HS level 6 or 4 within that chapter. Differences are sometimes considerable. You will find below 2 examples.

We have also noticed that, at least in the examples we have looked at, the following computations give the same results:
- sum of values of all products at HS level 6
- sum of values of all products at HS level 4
- sum of values of all products at HS level 2
- value mentioned for the "TOTAL" product

This would let us think that all products are reported at all levels. The problem could thus possibly be that some products are reallocated to other chapters in HS level2 data.

Could you explain the reason of the observed differences in HS level 2 data? How are aggregated data at level 4, 2 and TOTAL constructed? Does this situation also occurs in Comtrade SITC databases?

First example: The extraction of

Year=2006
Flow=2(Export)
Reporter=376(Israel)
Partner=400(Jordan)
Product=all products at all levels within chapter 71

gives the following data:

Product=71 => value=29639000
Product=7102 => value=436000
Product=710239 => value=436000
Product=7108 => value=10932000
Product=710812 => value=10932000
Product=7113 => value=1200000
Product=711311 => value=22000
Product=711319 => value=1178000

If you sum values at level 6 of HS you get 436000 + 10932000 + 22000 + 1178000 = 12568000.
If you sum values at level 4 of HS you get the same amount: 436000 + 10932000 + 1200000 = 12568000.
This is not the value which is reported for chapter 71: 29639000.

Second example: The extraction of

Year=2006
Flow=2(Export)
Reporter=376(Israel)
Partner=400(Jordan)
Product=all products at all levels within chapter 99:

gives the following data:

Product=99 => value=105000
Product=9999 => value=23291000
Product=999999 => value=23291000

If you sum values at level 6 of HS you get 23291000.
If you sum values at level 4 of HS you get the same amount: 23291000.
This is not the value which is reported for chapter 99: 105000.

A: Please note that countries often provide non-standard codes, which we need to assign to so-called memorandum items on the 2-digit level (to the respective HS chapter if possible or to chapter 99). This essentially means that this information is deleted from the detailed 6-digit data (from which the 4-digit data is aggregated) but maintained on the 2-digit level. In order to allow the detailed data to sum up to the overall total, we create item 999999=Total-(Sum of 6 digit data).

As a result, you find that the detailed data for chapters 01-97 is not necessarily adding up to the total of individual chapters (whenever there are non-standard codes) while chapter 99=999999. For SITC data, this is essentially happening, at least for the most recent data: Any recently published SITC data is based on a data conversion of the detailed HS data provided by the country. In this conversion on the detailed level, the issue of memorandum items plays no role. Yet, for the SITC data, there is the issue that not all HS sub-headings are converted to the most detailed level of SITC because of the lack of correspondence. Hence, some 3-digit groups might not have any or only partial 4-digit data. The 4 digit level adds the difference to code 9310.
Q: Why are quantity units not given at higher levels?
A: Weight or other quantity figures are not given at the higher levels of the classification tree because it would contain too many estimates. The quantity information is given at the 3-, 4- and 5-digit level of the SITC classification. To obtain quantity information at the 2-digit level of the SITC classification, you would need yourself to extract the 3-digit information and aggregate it, adding estimates where necessary.

Q: Why there is discrepancy between partner World (code 0) and sum of all individual partners?
A: It's due to rounding. In the dissemination system (UN Comtrade), trade values are rounded but not in the processing system. The partner World is calculated during data processing, taking into account decimal numbers in each partner country, then the result is rounded. Therefore, there is a possibility of a slight discrepancy between the world partner and the sum of all individual partners.