

Sources for International Trade, Prices, Production, and Consumption

Abstract

Lengthy time-series of international trade, prices, production and consumption of commodities are becoming increasingly easy to find and use via the Internet. Major international organizations see openly available data sources as key elements of their service mission. At the same time, international organizations maintain use policies increasingly out of sync with the features that their technological innovations offer. In contrast, even though the United States government offers a wide range of international data with no use limits, user interfaces are often significantly less well developed. In addition to international and U. S. government sources, one entirely privately produced resource, LexisNexis Statistical Insight, is also included and notable for falling somewhere in between these two poles. This article describes the general data availability and stated use policies for the databases *Comtrade*, OECD iLibrary, UNCTAD Commodity Price Statistics, Minerals Yearbook, FAOSTAT, PSD Online, and Statistical Insight with respect to data on commodity trade, prices, production, and consumption

Keywords: Trade; Commodities; Minerals; Agriculture;

Introduction

Thanks to the Internet, librarians and researchers have the best kind of problem these days: sifting through the vast quantities of increasingly freely available statistical time-series for just the statistics needed for their research. While this paper will not attempt to provide an exhaustive catalog of every database, indicator, country or year of coverage in each of the sources discussed, it will provide some broad outlines for each database's trade, prices, production, and consumption data.

The databases covered are

- United Nations (UN) *Comtrade*
- Organisation for Economic Cooperation and Development's OECD *iLibrary*
- United Nations Conference on Trade and Development (UNCTAD) *Commodity Price*

by Amy West ¹

Statistics

- United States Geological Survey *Minerals Yearbook*
- Food and Agriculture Organization (FAO) *FAOSTAT*
- United States Department of Agriculture (USDA) *Production, Supply and Distribution Online (PSD Online)*
- LexisNexis *Statistical Insight*

Sources covered have at least five decades of data and one or more of the following indicators:

- bilateral commodity trade
- international prices
- international production
- international consumption

These and many more links can all be found in the author's del.icio.us account at <http://www.delicious.com/umdatalib>.

This paper will focus on the parts of these databases that relate to bilateral commodity trade, prices, production and consumption, but each contains time-series for many additional topics. Where database providers make such features available, the paper will also cover copyright statements, citation tools, alert services, and visualization options.

Terminology note: while data and statistics are different things (specifically, data are what one uses to create statistics), in practice, the two terms are routinely used interchangeably. Most of the databases in this paper use "data" to describe their content and I will follow the same convention.

UN *Comtrade*

The most comprehensive source of international bilateral commodity trade data is the United Nations *Comtrade* database at <http://Comtrade.un.org/db/default.aspx>.

Comtrade contains data from 170 reporting countries and covers 1962 to the present as completely as possible. *Comtrade* data processing practices include conversion of commodity values from national currency to US dollars using reporter-supplied exchange rates, conversion of reported quantities to metric units, and conversion from the reporter-supplied classification into all versions of the Harmonized System (HS), all versions of the Standard International Trade Classification (SITC), and the Broad Economic Categories (BEC) classification. The United Nations says of its received data, "The data are permanently stored in the UN Comtrade database server." The *Comtrade* "About" page (<http://unstats.un.org/unsd/tradekb/Knowledgebase/What-is-UN-Comtrade>) contains the statement "Browser Warning: UN Comtrade works best with Internet Explorer Web Browser. In other browsers, some pages might not work properly. " Mac users can not use *Internet Explorer* unless they have a Windows emulator, but they should have no trouble using *Comtrade* with *Firefox*. On the other hand, users of *Google Chrome* and *Safari* may find some limited functionality.

Anyone may view *Comtrade* data and anyone may download up to 50,000 records per user per session. If needed, a person or institution may subscribe for additional services including unlimited downloads.

Visualization options are very limited. *Comtrade* includes some automatically generated pie charts and a map based explorer. However, at the time of writing, the "Explorer with Map" remains broken in all platforms/browsers checked.

Expansive user access to *Comtrade* stands in contrast to the stated policies on re-use and dissemination of *Comtrade* data. There are two slightly different copyright statements plus a separate policy on use and re-dissemination. On the "About" page (<http://unstats.un.org/unsd/tradekb/Knowledgebase/What-is-UN-Comtrade>) is this statement: "The UNSD holds the copyright to COMTRADE data. All rights reserved. Users cannot disseminate *Comtrade* raw data without the express permission of the UNSD." However, *Comtrade* provides no information on what "COMTRADE raw data" would be on this or any other pages. The footer of every *Comtrade* page contains this copyright statement: "No part of this material in either its printed or electronic format may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner." *Comtrade* has a third description of rights and permissions at (http://Comtrade.un.org/db/help/PolicyOnUseAndRedissemination_11Aug2010.pdf) which states that re-dissemination of *Comtrade* data within an institution does not require permission. So, a student including data or tables in a paper can do so without requesting permission provided the paper only goes to the professor. The *Comtrade* policy

on use also states that use of a "few tables or graphs in newspaper articles, journals, other magazines or books" doesn't require permission. In an email from staff in the UN Publications Board and Exhibits Committee, the author was told "Therefore, whenever someone, after downloading some available data, wishes to reproduce it (say, in an article or as photocopies to be used in a class) or forward it (say, to other users, subscribers, include it in a database, etc), that someone needs a permission."²

To recap, the UN says that using data in papers within a single institution or using a few tables or graphs is fine to do without seeking permission. Actually reproducing the data does require permission and depending on the quantity reproduced, a fee as well.

What is dissonant about this use policy is that one of the best things about *Comtrade* is the ease with which a user may access the data for precisely these kinds of uses. Because librarians ought to direct their users in the proper use of the resources to which they send their users, the author anticipates that in the future she will recommend that, regardless of whether users download data or not, when they cite their sources they should point to the table on the *Comtrade* website that they generated when downloading data for analysis. In pointing to *Comtrade* itself, the user is not re-disseminating data and therefore does not need to seek permission.

Because there are so many caveats that come with using data from so many different sources, *Comtrade* takes the unusual step of inserting a "ReadMe" screen before the display of a table. (<http://Comtrade.un.org/db/help/uReadMeFirst.aspx>) Much like click-wrap licenses for software purchased online, to get past the screen and to the data, users must check a box indicating that they have read the information provided describing data coverage and limitations before they can see the table that they have just constructed. The ReadMe page also appears when a user saves the URL to a particular table. Thus, if a table is cited in a paper and another user (a reviewer or professor for example) goes to see the table, then that user will also have to indicate that she has read the form.

Finally, for users who would like to develop their own interface to the *Comtrade* data, the United Nations provides an API. Some aspects of the *Comtrade* database are freely available via the API, but others are limited to institutions with a license to the database. Details are available at <http://Comtrade.un.org/wsl>.

OECD iLibrary

The OECD has two versions of its bilateral commodity trade database. *SourceOECD* is the older version and is in the process of being phased out. Its replacement is the *OECD iLibrary*, officially launched on 20/9/10.³

The OECD is a 33-nation member organization established in 1961. As the OECD's publications serve the needs of member organizations, the *OECD iLibrary* has a different scope from *Comtrade*. The reporting countries are the member nations, although for the newest members little data may be found so far. *OECD iLibrary* also reports data from a few non-member entities including the EU-15 Extra EU⁴ and Taiwan (labeled as Chinese Taipei and for which data are hard to find). However, each member reports all of the countries with which it trades, so while the reporters are limited to the OECD membership, the database is still quite large. The overall time period covered is 1960-present. *OECD iLibrary* classifies its data according to the HS and SITC systems. *OECD iLibrary* may also include volume and value of trade for price calculation if provided by the reporting countries.

While some *OECD iLibrary* content is freely available, full access is fee-based. The cost can be quite high. However, the *OECD iLibrary* also includes all of the OECD's textual materials as well as their statistical databases, so subscribers get a lot of content for the cost.

Several very useful features in the *OECD iLibrary* make it even more worthwhile for institutions that can afford to subscribe. While providing citations and links to citation management software is pretty standard now in journal databases, similar features are much less common in statistical databases. *OECD iLibrary* provides citations to whole datasets, links to the major citation management software and will provide citations down to the table level for tables in textual publications. *OECD iLibrary* uses Digital Object Identifiers (DOI) as part of its classification structure. In doing so, they are adhering to emerging international standards for data citation. *OECD iLibrary* also provides the markup needed to use browser-based citation tools like Zotero. RSS feeds that alert subscribers to new data as added are another feature of use to researchers studying long-term trends .

The OECD also claims copyright to its materials and has a separate terms and conditions document at <http://www.oecd-ilibrary.org/about/terms> for subscribers to the *OECD iLibrary*. However, in each case acceptable uses are clearly defined and for most academic purposes, the OECD maintains a liberal use policy. They summarize acceptable use this way: "OECD encourages the use of its content (textual, statistical, and multimedia). You can copy, download or print content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. When available, you should use the "Cite As" tool within the *OECD iLibrary*. When the "Cite as" tool is not available, you should cite the Title of the material, © OECD, publication year (if available) and page

number or URL (uniform resource locator) as applicable."⁵

Whereas *Comtrade* has the greatest volume of trade data, *OECD iLibrary* provides the most researcher-friendly interface by integrating citation into the database and providing alert services via RSS feeds on a database-by-database or topic-by-topic basis. The *OECD iLibrary* also has a more clearly worded statement on acceptable use of their content than does *Comtrade*.

UNCTAD Commodity Price Statistics

The *UNCTAD Commodity Price Statistics* at <http://www.unctad.org/templates/Page.asp?intItemID=1889&lang=1> contains commodity prices and price indices for the period 1960-present. The focus here is not on comprehensive lists of commodities, but instead on raw materials, especially mineral commodities. Prices are in US dollars per ton unless indicated otherwise. The *Commodity Price Statistics* works in multiple browsers and operating systems. The web version of this database is freely available. The *UNCTAD Commodity Price Statistics Terms and Conditions* document generally "grants permission to Users to visit the Site and to download and copy the information, documents and materials (collectively, "Materials") from the Site for the User's personal, non-commercial use".⁶

For researchers looking for prices on minerals and raw materials, this is an excellent choice.

USGS Minerals Yearbook

The USGS *Minerals Yearbook* at <http://minerals.usgs.gov/minerals/pubs/myb.html> provides price and production data either by country with reports from the mid-90s-present or by mineral with statistics from 1900-present. It is freely available online and, as a United States government publication, is not subject to copyright. The web version of the *Minerals Yearbook* remains essentially like its print predecessor with a few format modifications. The *Minerals Yearbook* retains the organization of the print with volume one covering minerals, volume two covering domestic prices and production and volume three covering international prices and production. Most of the content is provided as PDF files or Microsoft Excel files. There is no single database to search across nor are there any built-in citation tools, or APIs. The USGS does provide an RSS feed of new publications on minerals that supports researchers with ongoing needs for data. Individual tables may include recommended citations. While the *Minerals Yearbook* is not very fancy, it is free, easy to find and use, regularly updated, not subject to usage restrictions of any kind, and backed up with print editions in case the site were to become unavailable.

FAOSTAT

The FAO database *FAOSTAT* contains a wealth of agricultural and resource data. It is the primary source for comprehensive (or as near to comprehensive as can be)

international

- production data (1961-present)
- trade (1961-present) - quantity, unit value, value
- prices (1966-present) - in local currency, standard local currency, US dollars
- food supply (1961-present) - quantity produced, producer price, area harvested, yield per hectare

FAOSTAT units are metric and values are expressed in US dollars unless noted otherwise. Production, price, and supply data can be converted to HS classification from FAO codes, but so far, trade data are only available by FAO code. *FAOSTAT* is free to use in its entirety as of July 2010, but for researchers in need of very large batch downloads, a free registration is required.

FAOSTAT says of its copyright “The data of the *FAOSTAT* database shown on this internet site are copyrighted by the Food and Agriculture Organization of the United Nations and are provided for your internal use only. They may not be re-disseminated in any form without written permission of the FAO Statistics Division.”⁷ There are no separate Terms of Use. A slightly more positive expression of rights from the FAO is “All rights reserved. FAO encourages the reproduction and dissemination of material published on this Web site. Non-commercial uses will be authorized free of charge, upon request.”⁸

FAOSTAT does not have an RSS feed for the whole database or parts thereof. No citation tools are built in either. Indeed, all FAO says about citation is “All references to *FAOSTAT* data will have to be mentioned with the proper URL and the access date.”⁹ Unfortunately, *FAOSTAT* does not support persistent URLs for individual tables, so presumably the proper URL is the URL for the overall dataset such as TradeSTAT rather than US imports of shelled almonds for 2006 by importing country.

FAOSTAT targets data users with needs for large amounts of data, but it also provides simple graphical representations of trend data for users with casual data needs.

USDA Production, Supply and Distribution Online (PSD Online)

The USDA *PSD Online* database at <http://www.fas.usda.gov/psdonline/psdhome.aspx> contains trade, prices, supply, demand, and consumption data for agricultural commodities from 1960-present. The United States is the only reporter, but bilateral data between the US and other countries are available where they exist. Like nearly all US governmental databases, *PSD Online* is free. There are also no usage restrictions since this database is excluded from copyright like the *Minerals Yearbook* by virtue of being a US government work. It appears to work in all browsers

and platforms. In Firefox, there are some display quirks, but they do not appear to affect the database’s function or its return of data. Users may create free accounts in the database and thereby save their queries for re-use. Users may also download datasets for each commodity for analysis in their preferred software. *PSD Online* also offers a data availability lookup at <http://www.fas.usda.gov/psdonline/psdAvailability.aspx> that, if the user finds it first, can save them the effort of constructing a query only to get no results.

LexisNexis Statistical Insight

LexisNexis Statistical Insight is very different from the other resources mentioned here. It is probably closest in many ways to a traditional journal database. It does contain some tables in Excel or *.gif, but its value lies in abstracts of print publications from an exceptionally wide range of sources. Even when the data are not directly available in *Statistical Insight*, it still helps users to figure out what kinds of data might exist and the most likely sources for that data. LexisNexis claims that *Statistical Insight* includes data from 1801-present. *Statistical Insight* includes all kinds of statistical publications including those with prices, production, and consumption. At the University of Minnesota, *Statistical Insight* helps to fill a gap in the Libraries collection of market research. Market research is very, very expensive and the University of Minnesota is unfortunately not able to purchase market research reports directly. However, *Statistical Insight* can identify sources of data from trade associations that can cover similar territory to market research. Thus, it helps us serve researchers that otherwise would have to look elsewhere for such data.

Statistical Insight is less expensive than market research reports, but it is still very expensive, especially if an institution licenses access to additional modules that incorporate more data into the database. No part of it is freely available. While there are limits on claims to copyright of factual material in US law, license agreements can overcome those limits. LexisNexis licenses are typically fairly restrictive. On the other hand, to the extent that *Statistical Insight* serves to identify data sources rather than as a source of data itself, rights issues seem a little less important in this case.

Conclusion

Not only are more data sources appearing online, but they are doing so with the explicit goal of making more data available to the greatest number of users. Data sources that may have launched with short time series find that users always want more and the data source producers strive to meet those desires as openly as possible. Many sources provide researcher-friendly features such as update alerts via RSS feed and built-in citation tools. At the same time, data source rights statements do seem to be out of sync with the spirit of expanded open access. In an environment in which data source producers enable features like mass

downloads, the expectation that permission must be acquired on a case-by-case basis in order to use those downloads seems odd. However, these are minor quibbles. The increased quantity of data, increased ease of access and improved interfaces far outweigh any drawbacks identified in this paper.

Notes:

1 Amy West, Data Services Librarian. 10 Wilson Library, University of Minnesota, 309 - 19th Avenue South, Minneapolis, MN 55455-0414. (612) 625-6368. westx045@umn.edu.

2 Morteo, R., 2010. Question about the copyright statement for UN *Comtrade*.

3 Schoenfeld, A., 2010. Official Launch of OECD *iLibrary*.

4 Glossary:EU enlargements - Statistics explained. European Commission - Eurostat. Available at: http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:EU-15 [Accessed October 1, 2010].

5 OECD *iLibrary*: Copyright and Permissions. Available at: <http://www.oecd-iLibrary.org/about/copyright> [Accessed September 30, 2010].

6 UNCTAD.ORG >> Terms & Conditions. Available at: <http://www.unctad.org/Templates/Page.asp?intItemID=2146&lang=1> [Accessed September 30, 2010].

7 FAOSTAT. Available at: <http://faostat.fao.org/site/567/default.aspx#ancor> [Accessed September 30, 2010].

8 FAO: Copyright. Available at: <http://www.fao.org/corp/copyright/en/> [Accessed September 30, 2010].

9 FAOSTAT. Available at: <http://faostat.fao.org/site/567/default.aspx#ancor> [Accessed September 30, 2010].