

Distr.  
GENERAL

WP.16  
14 April 2011

ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION  
FOR EUROPE (UNECE)  
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION  
STATISTICAL OFFICE OF THE EUROPEAN  
UNION (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION  
AND DEVELOPMENT (OECD)  
STATISTICS DIRECTORATE**

**Meeting on the Management of Statistical Information Systems (MSIS 2011)**  
(Luxembourg, 23-25 May 2011)

Topic (ii): From local to corporate perspective (industrialization and standardization)

## **Economic Data Management at the IMF**

### **Supporting Paper**

Prepared by Shri Narayanan, International Monetary Fund (IMF)

## **I. Introduction**

1. The International Monetary Fund collects Economic Data from member country authorities during their surveillance missions and other times. Data collected are per country and the data are stored mostly in Excel spreadsheets. Most of the analysis and calculations are done using Excel. Data stored did not have much of a structure, with little or no metadata. Data management and data sharing were extremely cumbersome.

## **II. Background**

2. In order to address the above issue, over the last several years, IMF has developed many versions of the Data Management Guidelines (DMGs) for their economists and research assistants. One of the important guidelines is that the data collected is stored in a structured database for ease of use, maintenance, and sharing. However, these were guidelines and were not adopted as a standard. The absence of a sophisticated tool for data management was also a factor in the non-adoption of these standards.

3. As Excel is neither a database nor a time series data management tool, maintaining and manipulating data in Excel can become quite difficult and unwieldy over time, and open the door for introducing data errors. Data stored in unstructured form in Excel is not retrievable, difficult to find, and prone to errors. Without proper metadata it is difficult to ensure correct understanding and seamless sharing of data within country teams. Therefore, tools are needed to help staff move towards using structured databases that better facilitate storage, manipulation, documentation, and sharing of time series data.

4. Since early 2008, migration of data out of Excel workbooks and into structured databases has been a key element of IMF-wide efforts to improve data management. Economic data and supporting documentation

are essential inputs for the organization's work. As recommended by the DMGs, economic data need to be maintained in structured databases to preserve the institutional memory of the IMF, and to facilitate improvements in data quality and data sharing among desk economists in support of their operational work across the organization and to the outside world more generally.

### III. The Evolution of Data Management for Excel (DMX)

5. To support the need of a structured data base at the IMF, the Information Technology group of the IMF developed a tool called Data Management for Excel (DMX) as an initial prototype for a few country desk data management. Over the years, this product has become quite sophisticated and robust.

6. DMX is developed using Microsoft Visual Studio (.Net) tool set and uses MS Access and SQL Server 2008 as database store. Excel is the front-end GUI for this application, which is the overwhelming choice of the IMF's economists and research officers and data managers.

7. Initially, the development of DMX was sponsored by the African Department of the IMF to support their data management needs. This department had stored their country data in many spreadsheets which had become quite cumbersome over the years. To assist this department, the IT department developed and deployed the DMX software and also provided support for cleansing and migrating the data from the spreadsheets to the DMX database as an initial deployment. The economists and data managers were then able to use the system on their own.

8. In early 2011, in the context of the Economic Data Management Initiative (EDMI) at the IMF, the overseeing task force recommended and endorsed the use of DMX as the standard tool for maintenance and management of country desk data at the Fund. With the selection of DMX as the data management tool, the use of Excel would be limited to data manipulation, analysis, and presentation, while the structure of Excel files should be designed to facilitate the transfer of data from Excel to DMX and vice versa.

### IV. What is DMX?

9. Database Management for Excel (DMX) is an Excel add-in that provides time series data management capabilities to end-users thus facilitating the adoption of the data management guidelines. DMX combines the **best of both worlds** – the **familiar Excel interface** with **time series data management capabilities** that overcome the limitations of Excel as a time series management system. The result is a *portable* system that facilitates management and sharing both at the IMF headquarters and on while on mission to member countries.

10. With DMX, you still use Excel but in addition, you can do the following:

- Store data as series in a database, thereby eliminating the need to maintain large datasets in Excel workbooks
- **Share data through the database**, while eliminating the need to establish links between workbooks and worksheets
- **Convert high frequency data to lower frequencies**, thereby eliminating the need to maintain and link separate worksheets for each frequency
- **Perform manipulations in the database**, thereby eliminating the need to create and maintain links between worksheets and workbooks; manipulated data are automatically updated when the component values are updated

- Import data from CEIC and Haver directly into DMX
- Access data in different data sources, such as DMX and the IMF Data Warehouse
- **Track changes** made to the data over time
- Use **Excel-like functions for statistical calculations**, such as splicing, rebasing, percentage change, compound growth, HPFilter, and seasonal adjustment
- Maintain **documentation of data through metadata** such as source, descriptor, scale, and units, thereby making it easier to share data over time and between different members of a country team
- **Organize data** into smaller chunks called Tables for easy access and viewing
- **Clone time series** within or across databases
- **Create links** between databases

## V. Adoption

11. Two of the IMF's five area departments have already made substantial progress in migrating country databases to structured databases, using DMX. Currently, over 60 countries have migrated their data from Excel to DMX. Twenty more countries will be completed in the next year. Other departments have engaged in the migration with the intent to migrate their respective countries within the next three years.

12. At this stage, migration of data to DMX is intended to be largely limited to country desks in area departments. This reflects the need to preserve the operational flexibility of Excel commonly used by desk economists, which is fully compatible with DMX. Functional departments, and at some level area departments, use alternative database management tools (e.g., EcOS – Economic Outlook Suite – A software developed by Prognoz Inc, <http://www.prognoz.com>) which are better suited for manipulating large volumes of cross-country data (including enhanced statistical functions) and powerful graphical capabilities frequently used for professional standard publications.

## VI. User Feedback

13. A user survey was conducted to evaluate the benefits of the migration to structured databases using DMX as the platform in early 2011. Some of the key findings are:

- Over 90 percent of respondents said that the DMX migration greatly or somewhat improved data management in their Department. Only, 9 percent said it somewhat or greatly worsened data management practices
- 45 percent of the respondents rated their DMX expertise as excellent or above average.
- Reduced workload or data requests
- Improved Data Quality
- Improved Data Management for country desk
- Consistency across data in country data bases

## VII. Dmx Outside of IMF

14. An earlier version of DMX has been distributed to Central Banks in member countries for their data management work. Countries include Brunei, The Islamic Republic of Iran, Libya, Papua New Guinea, and a few others.

15. DMX is a Microsoft Visual Studio (.Net) based tool with a small footprint on the desktop or a laptop. The latest version of DMX needs Excel 2007 for proper functioning. The software is distributed on a CD-Rom.

## VIII. Sample Screen Shots of DMX Application

Simple interface within Excel to search and browse series and their metadata

The screenshot shows the DMX application running within Microsoft Excel. The interface is divided into several panes:

- Data Navigator:** Shows a list of series codes and their descriptors. The selected series is 'A111PCPI' (Consumer Price Index (2000=1)).
- Metadata:** Shows details for the selected series, including its alias, country (United States), and various attributes like 'Prices' and 'Department'.
- Chart:** Displays a line graph of the Consumer Price Index (CPI) for the United States from 1990 to 2008, showing a steady upward trend.

Series Code	Descriptor
108	A111NCPCL_D_SA_3M
109	A111NCPCL_R_SA
110	A111NCPCL_R_SA_3M
111	A111NFB_R_SA
112	A111NFB_R_SA
113	A111NFIPR_R_SA
114	A111NGDP
115	A111NGDP_R_SA
116	A111NGDP_SA
117	A111NGDP_SA1
118	A111NGDP_SAD
119	A111NM_R_SA
120	A111NX_R_SA
121	A111PCPI
122	A111PCPLCORE
123	A111PCPLCORE_PCYA
124	A111PCPLCORE_SA
125	A111PCPLCORE_SA_3M
126	A111PCPLCORE_SA_3M
127	A111PCPL_SA
128	A111PCPL_SA_3M
129	A111PCPIE
130	A111PCPIE_PCYA

Series Code	Descriptor	Country Code	Country Name
A111PCPI	Consumer Price Index (2000=1)	111	United States

Series Code	Value
1990A1	75.88
1991A1	79.09
1992A1	81.48
1993A1	83.89
1994A1	86.08
1995A1	88.49
1996A1	91.09
1997A1	93.22
1998A1	94.66
1999A1	96.73
2000A1	100.00
2001A1	102.83
2002A1	104.46
2003A1	106.83
2004A1	109.69

## Interface within Excel to view the data stored in the database.

Database	Series_Code	Descriptor	Formula	Scale	Units	Y:2002	Y:2003	Y:2004	Y:2005	Y:2006	Y:2007
1	Wimfdata\ecor	AE11PCPI_CC	Consumer F	X12 (AE11PC	Unit	103.967	105.774	108.454	110.875	113.324	116.248
2	Wimfdata\ecor	AE11PCPI_CC	Consum	( ( AN04PCPI	Unit	103.978	105.774	108.455	110.872	113.318	116.237
3	Wimfdata\ecor	AA05PPPWT	PPP Va	A536PPPWT	Unit	3730291859.38	1536714935.01	1976085620.47	1147630503.44	1340924903.44	38822808256.4
4	Wimfdata\econ\dr	A536PPPWT	PPP \		Billion	563.712	603.226	650.198	705.162	767.988	838.479
5	Wimfdata\econ\dr	A548PPPWT	PPP \		Billion	234.995	253.886	277.648	301.306	328.973	359.271
6	Wimfdata\econ\dr	A566PPPWT	PPP \		Billion	197.245	211.372	230.936	250.24	272.25	299.673
7	Wimfdata\econ\dr	A578PPPWT	PPP \		Billion	346.317	378.938	413.277	445.367	483.194	519.786
8	Wimfdata\econ\dr	A582PPPWT	PPP \		Billion	131.461	144.114	159.916	178.073	198.935	221.614
9	Wimfdata\ecor	AN04PPPWT	PPP Va	A532PPPWT	Unit	2123421644.38	1908395437.44	4300345673.13	7538103192.49	1024867068.94	9616527095.99
10	Wimfdata\econ\dr	A542PPPWT	PPP \		Billion	845.778	890.525	957.115	1027.374	1114.914	1201.866
11	Wimfdata\econ\dr	A532PPPWT	PPP \		Billion	187.348	197.085	219.555	243.081	268.523	293.311
12	Wimfdata\econ\dr	A576PPPWT	PPP \		Billion	140.121	148.116	166.642	184.85	206.393	228.303
13	Wimfdata\econ\dr	A528PPPWT	PPP \		Billion	478.876	506.182	550.988	592.234	641.194	696.137
14	Wimfdata\econ\dr	A924PPPWT	PPP Val		Billion	3701.095	4157.822	4697.901	5314.372	6121.904	7034.838
15	Wimfdata\econ\dr	A534PPPWT	PPP Val		Billion	1719.582	1876.582	2096.198	2354.448	2668.832	2996.588
16	Wimfdata\ecor	AE11PPPWT	PPP Va	AN04PPPWT	Unit	6531281082.88	7848982355.61	20374865313.1	196505820454.6	73101344940.1	89865402786.5
17	Wimfdata\ecor	AA05PPPWT	PPP \	A536PPPWT	Unit	3730291859.38	1536714935.01	1976085620.47	1147630503.44	1340924903.44	38822808256.4
18	Wimfdata\econ\dr	A536PPPWT	PP		Billion	563.712	603.226	650.198	705.162	767.988	838.479
19	Wimfdata\econ\dr	A548PPPWT	PP		Billion	234.995	253.886	277.648	301.306	328.973	359.271
20	Wimfdata\econ\dr	A566PPPWT	PP		Billion	197.245	211.372	230.936	250.24	272.25	299.673
21	Wimfdata\econ\dr	A578PPPWT	PP		Billion	346.317	378.938	413.277	445.367	483.194	519.786
22	Wimfdata\econ\dr	A582PPPWT	PP		Billion	131.461	144.114	159.916	178.073	198.935	221.614
23	Wimfdata\ecor	AN04PPPWT	PPP \	A532PPPWT	Unit	2123421644.38	1908395437.44	4300345673.13	7538103192.49	1024867068.94	9616527095.99
24	Wimfdata\econ\dr	A542PPPWT	PPP \		Billion	845.778	890.525	957.115	1027.374	1114.914	1201.866
25	Wimfdata\econ\dr	A532PPPWT	PP		Billion	187.348	197.085	219.555	243.081	268.523	293.311
26	Wimfdata\econ\dr	A576PPPWT	PP		Billion	140.121	148.116	166.642	184.85	206.393	228.303
27	Wimfdata\econ\dr	A528PPPWT	PP		Billion	478.876	506.182	550.988	592.234	641.194	696.137
28	Wimfdata\econ\dr	A924PPPWT	PPP \		Billion	3701.095	4157.822	4697.901	5314.372	6121.904	7034.838
29	Wimfdata\econ\dr	A534PPPWT	PPP \		Billion	1719.582	1876.582	2096.198	2354.448	2668.832	2996.588
30	Wimfdata\econ\dr	A924PCPI_CDR	Consum		Unit	100.096	99.732	100.111	100.829	101.435	102.41
31	Wimfdata\econ\dr	A534PCPI_CDR	Wholes		Unit	107.713	113.459	120.989	125.353	129.833	137.121
32	Wimfdata\ecor	AA05PCPI_CC	Consur	( ( A536PCPI	Unit	110.305	114.09	118.56	124.762	132.066	137.419

## Interface within Excel to transform data as needed – retrieve just the data, retrieve rescaled data or transformed data.

The screenshot shows the Microsoft Excel interface with the Data Navigator pane on the left. The pane displays a list of series codes and their descriptors. The series code 'AA05PCPI' is selected, and a context menu is open over it, showing various options for data retrieval and transformation. The main Excel window shows a blank spreadsheet with columns A through P and rows 1 through 32.

Series_Code	Descriptor
3457	AA04PCPIF_SA Consumer Price Index - Food & ( At
3511	AA05PCPIF Consumer Price Index - Food & ( At
3455	AA04PCPIF Consumer Price Index - Food & ( At
3508	AA05PCPL_SA Consumer Price Index - ASEAN - ( At
3453	AA04PCPL_SA Consumer Price Index - ASEAN - ( At
3508	AA05PCPI_PCYA Consumer Price Inflation: ASEA ( At
3452	AA04PCPI_PCYA Consumer Price Inflation: ASEA ( At
3505	AA05PCPI_CORE_PCYA Consumer Price Inflation- Core: ( At
3451	AA04PCPI_CORE_PCYA Consumer Price Inflation- Core: ( At
3504	AA05PCPI Core: ( At
3503	AA05PCPI Core: ( At
3450	AA04PCPI Core: ( At
3441	AA04LWR_S Core: ( At
3440	AA04LWR_F Core: ( At
3498	AA05FSMIZ Core: ( At
3437	AA04FSMIZ Core: ( At
3432	AA04FSDCF Core: ( At
3497	AA05FSDCF Core: ( At
3496	AA05FSDCF Core: ( At
3433	AA04FSDCF Core: ( At
3495	AA05FRM_E Core: ( At
3431	AA04FRM_E Core: ( At
3876	Core: ( At

## Easy access to commonly used statistical functions

