



Tip or Technique

IBM Cognos 8.3 BI Excel 2007¹ Notes

Product(s): IBM Cognos 8.3

Area of Interest: Reporting

¹ Excel 2007 refers to Microsoft Office Excel 2007, a trademark of Microsoft Corporation.

Copyright

Copyright © 2008 Cognos ULC (formerly Cognos Incorporated). Cognos ULC is an IBM Company. While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Cognos does not accept responsibility for any kind of loss resulting from the use of information contained in this document. This document shows the publication date. The information contained in this document is subject to change without notice. Any improvements or changes to the information contained in this document will be documented in subsequent editions. This document contains proprietary information of Cognos. All rights are reserved. No part of this document may be copied, photocopied, reproduced, stored in a retrieval system, transmitted in any form or by any means, or translated into another language without the prior written consent of Cognos. Cognos and the Cognos logo are trademarks of Cognos ULC (formerly Cognos Incorporated) in the United States and/or other countries. IBM and the IBM logo are trademarks of International Business Machines Corporation in the United States, or other countries, or both. All other names are trademarks or registered trademarks of their respective companies. Information about Cognos products can be found at **www.cognos.com**

This document is maintained by the Best Practices, Product and Technology team. You can send comments, suggestions, and additions to cscogpp@ca.ibm.com.

Contents

1	INTRODUCTION	4
2	WHY EXCEL 2007?	4
3	COMPATIBILITY WITH OLDER VERSIONS OF EXCEL	4
4	FILE SIZE COMPARISON	5
5	ARE EXCEL 2007 FILES <i>ALWAYS</i> SMALLER?	7
6	WHAT ABOUT CHARTS?	8
7	WHAT ABOUT MAPS?	12
8	DOES EXCEL 2007 HAVE A LARGER CAPACITY THAN EXCEL 2002?	12
9	WHAT ABOUT EXISTING EXCEL FORMATS?.....	12

1 Introduction

Native Excel 2007 spreadsheets can be created in IBM Cognos 8.3 as a report output format. This document discusses some aspects of the new native Excel 2007 format which you might find interesting.

2 Why Excel 2007?

This new XML format has several advantages over the old binary Excel formats:

- Being XML based it is easy to process and create
- It is smaller in size than the old binary Excel formats
- It can be opened by older versions on Excel using Microsoft's compatibility packs
- The native Excel 2007 format is an XML based open standard Office Open format supported by ECMA International (<http://www.ecma-international.org>). The specifications can be found at <http://www.ecma-international.org/publications/standards/Ecma-376.htm>.

3 Compatibility with Older Versions of Excel

How do I open Excel 2007 files in older versions of Excel?

You install the compatibility pack from Microsoft following these steps:

Open your IE browser and go to <http://www.microsoft.com/office>

In Welcome to Office Online search for Compatibility Pack.

Take the link to Microsoft Office Compatibility Pack for Word, Excel, and PowerPoint 2007 File Formats.

Download and install the compatibility pack.

4 File Size Comparison

It is well known that the older Excel 2000 and Excel 2002 format files created by IBM Cognos ReportNet and IBM Cognos 8.1/8.2 were very large when compared to native Excel 2000 and 2002 files due to the use of the HTML/MHT format.

Excel 2007 files are compressed resulting in dramatically smaller files sizes than the HTML/MHT format and in most case even smaller than the old binary Excel format (BIF-8).

Let's try a little test.

Open Query Studio with the GO Sales and Retailers package and create a new list report:

- Product line
- Product type
- Product name
- Order number
- Retailer name
- Order year
- Order method
- Quantity
- Revenue

Save this report as Excel Test 1.

Go to IBM Cognos Connection and run as PDF, Excel 2002, Excel 2007 and CSV. This will take a few minutes as there are 45,000 rows of data to process.

Download all of these files to your PC.

How large are they? Here are the file sizes largest to smallest.

Excel 2002 (.XLS²): 24.6 MB

PDF: 9.7 MB

CSV: 7.8 MB

Excel 2007 (.XLSX³): 1.6 MB

² Regardless of the file extension, Excel 2000 and Excel 2002 files created by IBM Cognos ReportNet and IBM Cognos 8 are HTML/MHT format files and NOT binary files.

³ XLSX is the file extension used for Excel 2007 XML format files.

Clearly Excel 2007 files are small. Even smaller than CSV!! That's Excel 2007 file compression at work.

And Excel 2007 files are fast. The Excel 2007 spreadsheet opens in less than 5 seconds on my PC. In contrast, the Excel 2002 HTML/MHT file takes 3 minutes and 15 seconds ...

Yes, but that's not really the whole story. How does native Excel 2007 compare to binary Excel 2002?

Open the Excel 2002 (.XLS) file and save it as a native Excel file. Note the file size is reduced dramatically from 24.6 MB to 6 MB. It's still larger than the Excel 2007 XLSX format.

5 Are Excel 2007 Files *Always* Smaller?

No. If you use images in your report they may be larger because Excel 2007 format stores the images in the .XLSX file and Excel 2000/2002 only store references to the images on the web server.

Open Query Studio with the GO Sales and Retailers package and create a new list report:

- Product name
- Product image

The Product image should be visible within the report.

Save the report as Excel Test 2

Go to IBM Cognos Connection and run as PDF, Excel 2002, Excel 2007 and CSV.

Download all of these files to your PC.

How large are they now? Here are the file sizes largest to smallest.

PDF:	841 KB
Excel 2007:	782 KB
Excel 2002:	68 KB
CSV:	12 KB

Why is Excel 2007 now larger than Excel 2002? The Excel 2002 file does not contain the Product images. Being an HTML document it refers to the images on the web server. Of course, this means that anyone reading the spreadsheet must have access to the IBM Cognos 8 server or the images will not appear. This makes it difficult to mail the document to customers or users outside the firewall.

On the other hand, the Excel 2007 document contains the images. As the images I used are jpeg files, they are already compressed so the file compression which is a natural part of the Excel 2007 file structure has no effect on the size of the images.

Note that PDF is also relatively large because of the images. The images are stored in PDF native (vector image) form so the compression that is a natural part of the PDF file structure has some effect on the size of the files but they still take space.

If you'd like to see how large the original Product images are, they are stored in x:\Program Files\Cognos\c8\webcontent\samples\images\p*.jpg and you'll see the total size of the files is about 758 KB.

6 What about Charts?

Do charts increase the Excel 2007 file size the way that images do?

No. All versions of Excel that IBM Cognos ReportNet or IBM Cognos 8.x produces use native Excel charts. This means we provide the data in the spreadsheet and Excel draws the chart when you open the file.

There are differences however. Let's take a look.

Open Query Studio with the GO Sales and Retailers package and create a new list report:

- Product line
- Order method
- Years
- Quantity

Select Product line and Create Sections .

Add a Chart selecting all default options.

You should now see one chart and crosstab per Product line.

Save this report as Excel Test 3.

Go to IBM Cognos Connection and run as PDF, Excel 2002, Excel 2007 and CSV.

Download all of these files to your PC.

How large are they now? Here are the file sizes largest to smallest.

Excel 2002:	71 KB
PDF:	55 KB
Excel 2007:	18 KB
CSV:	1 KB

The Excel 2002 is large because it's HTML. Open this file in Excel and you'll see that the charts are native Excel charts. This means the data and not the chart image is stored in the .XLS file.

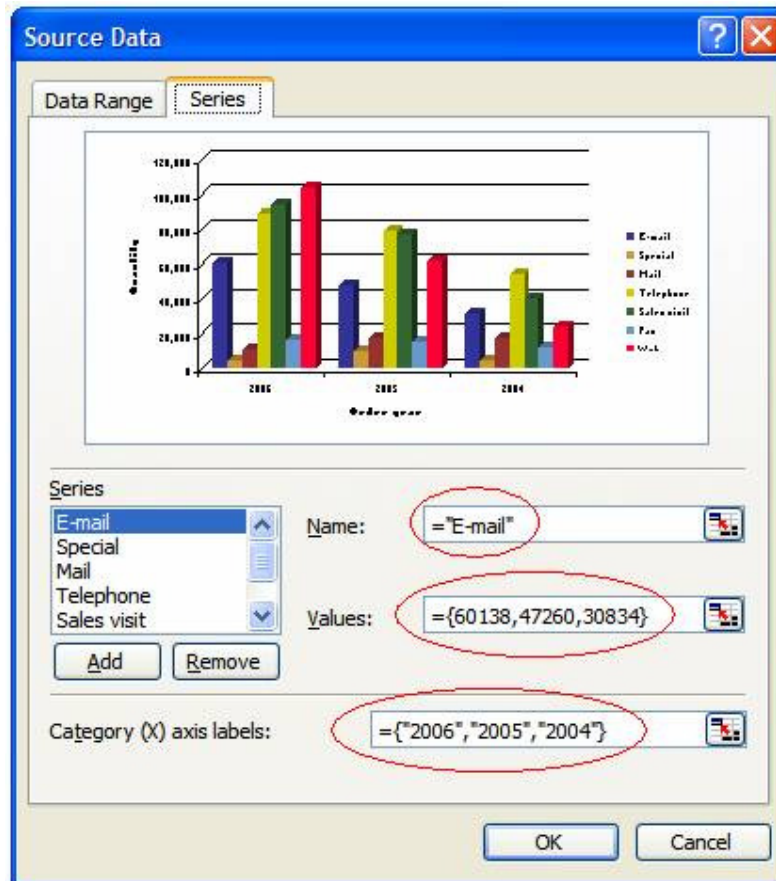
The Excel 2007 also uses native Excel charting so the data, not the chart image, is stored in the .XLSX file.

There is a considerable difference in how the data is stored for the charts in Excel 2002 and 2007.

For Excel 2000 and Excel 2002, using the HTML/MHT format, we stored the data as part of the chart definition.

Open the Excel 2002 file. Note that there is one worksheet.

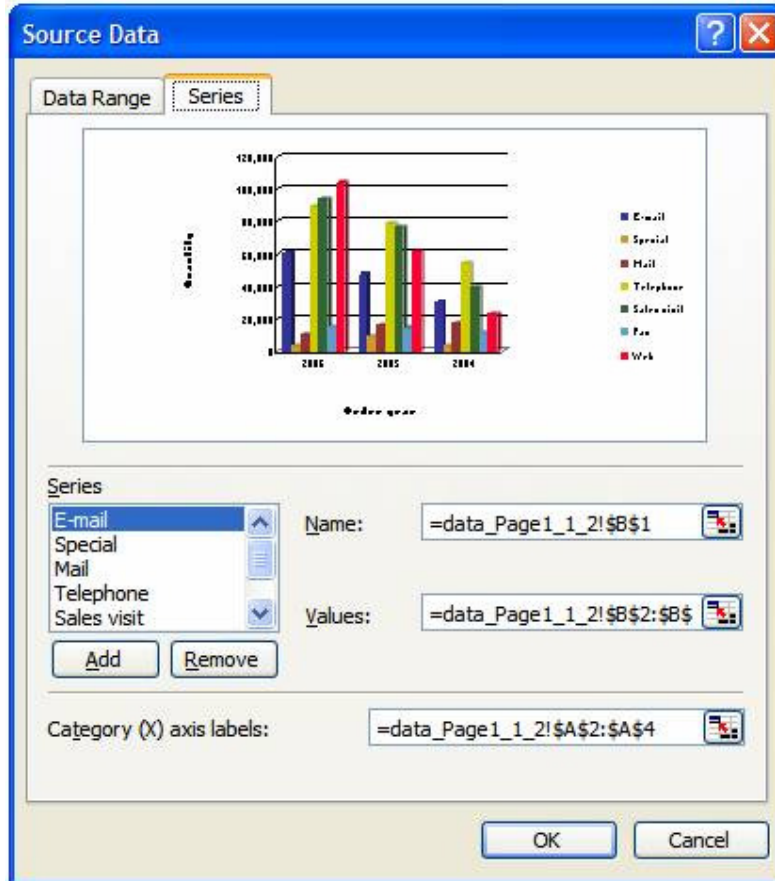
Right click on a chart and select Source Data. You will see that data for that chart right in the chart dialog:



Open the Excel 2007 file.

Note that there are 6 worksheets. Why??

Right click on a chart and select Source Data. You will see Excel references to the data:



Note that each reference starts with a page. Excel 2007 format stores the data for each chart on a separate page.

Why did we take a different approach with Excel 2007? Storing the data in that chart imposes data size limits which resulted in bug logged against our Excel 2002 format. We are unable to fix these bugs as the data size limits are imposed by Excel not IBM Cognos. There is no data size limit (that we are aware of) when the data is stored in a worksheet. Note that a user who does not want to see these worksheets can hide them. Select the sheet and the /Format/Sheet/Hide menu option.

The PDF is large because it has the chart images. The PDF compression of the vector images helps reduce the size.

7 What about Maps?

Do maps increase the Excel 2007 file size the way that images do?

Yes. Because Excel does not provide native map support, all maps must be rendered as images.

8 Does Excel 2007 Have a Larger Capacity than Excel 2002?

No, not in the way IBM Cognos generates the files.

While Microsoft states that Excel 2007 can manage 1 millions rows by 16,000 columns (<http://download.microsoft.com/download/3/3/c/33c4374f-372a-4a3a-bf27-f6d24ade4e32/Excel2007ProductGuide.doc>)...

... the Excel 2007 files generated by IBM Cognos 8.3 do not take advantage of this capability at this time instead using the Excel 2002 limits of 64K rows and 256 columns.

This is done as at this time very few companies have deployed Microsoft Office 2007 so the bulk of usage of this new format is expected to be older versions of Excel using the compatibility pack.

While the compatibility pack allows the older versions of Excel to read the new Excel 2007 format, it does not change the capacity limits. Were you to create a 1 million row Excel 2007 file and try to open it in an older version of Excel it would fail to open as it exceeds the 64K row limit.

9 What about Existing Excel Formats?

The existing Excel 2000 and Excel 2002 formats have not changed in the IBM Cognos 8.3 and continue to use HTML/MHT format output.