



In association with **Microsoft**[®]

How to use PowerPivot to understand your data

MARK WHITEHORN PROVIDES A STEP-BY-STEP GUIDE TO ONE OF THE EXCEL 2010'S MOST POWERFUL TOOLS

The closest many companies come to analysing their data is to use Excel's PivotTables, but PowerPivot takes things to a new level. Think of it as Excel PivotTable on steroids (*see box, right*).

But it's more than that. Yes, PowerPivot works as an Add-in for Excel 2010 that enables users to analyse very large sets of data, very rapidly, in real-time, but it should be viewed as a vital component in Microsoft's Business Intelligence suite.

We'd highlight two key skills: the ability work with simple and/or complex data (simple being essentially Excel worksheets, complex being relational databases); and the ability to combine data from different sources.

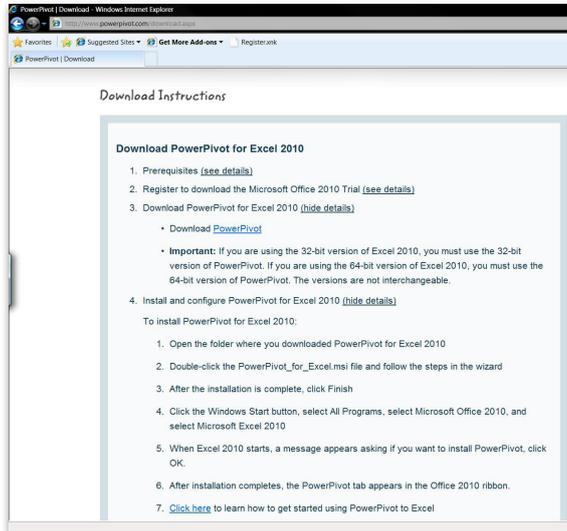
In this article, we're going to provide a step-by-step guide to using it as a tool in your business to make sense of your data.

WHAT ARE PIVOTTABLES?

PivotTables make it easy to analyse complex data. Take a spreadsheet of sales data, as shown on p4 of this whitepaper. We know the quantity of items each customer bought. We know (for most customers) in which county/region they live. If we want to analyse which items sell best in which county/region, we simply select Quantity as the Value, CountyRegion as the 'Column Labels' and 'CD ROM Title' as the 'Row Labels' and the PivotTable will automatically total it all up for us. See p5 for more.

Download

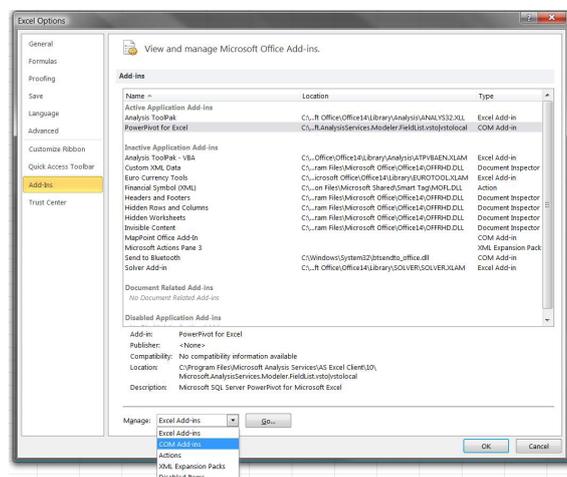
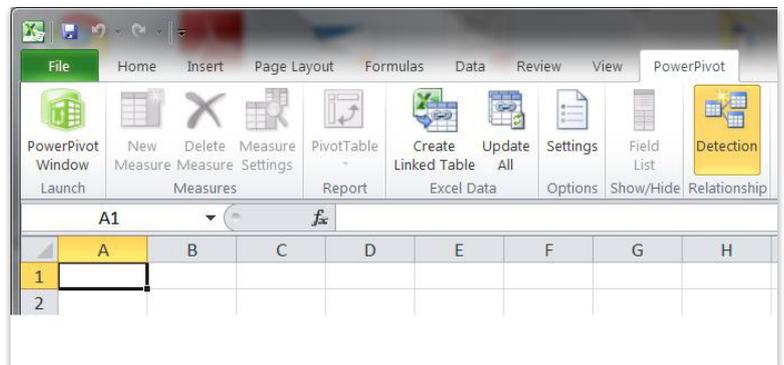
You will need Excel 2010 already installed. Go to www.powerpivot.com, click on “Download PowerPivot now!” and follow the instructions.



There is no charge for PowerPivot; you don't even have to register to download it. Make sure you download the correct version (32-bit and 64-bit versions are not interchangeable). Once installed, fire up Excel and a message should appear asking if you want to install PowerPivot into Excel; click OK. After the installation finishes, you should find a new PowerPivot tab appears in the ribbon.

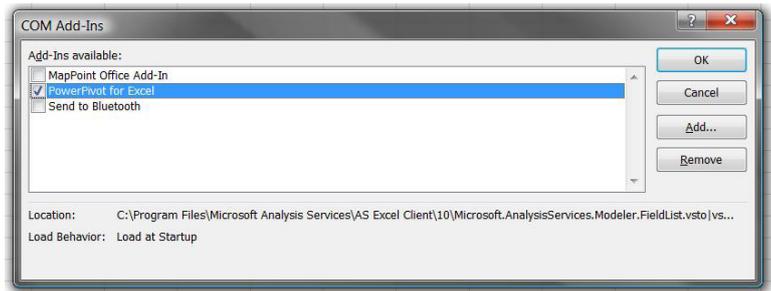
Disable and re-enable

The PowerPivot Add-in certainly slows down Excel's load time (it's three times slower). Your mileage may vary but it's worth knowing that you can easily disable the Add-in when it's not required. Click the File tab, Options, Add-Ins. You should see PowerPivot for Excel listed.



Note (in the 64-bit version at least) that it is a COM Add-in. To disable it, select COM Add-ins in the Manage: dialog box bottom left and press Go. This dialog appears...

Deselect PowerPivot for Excel and click OK. Re-enabling is the same except, of course, you select PowerPivot for Excel.

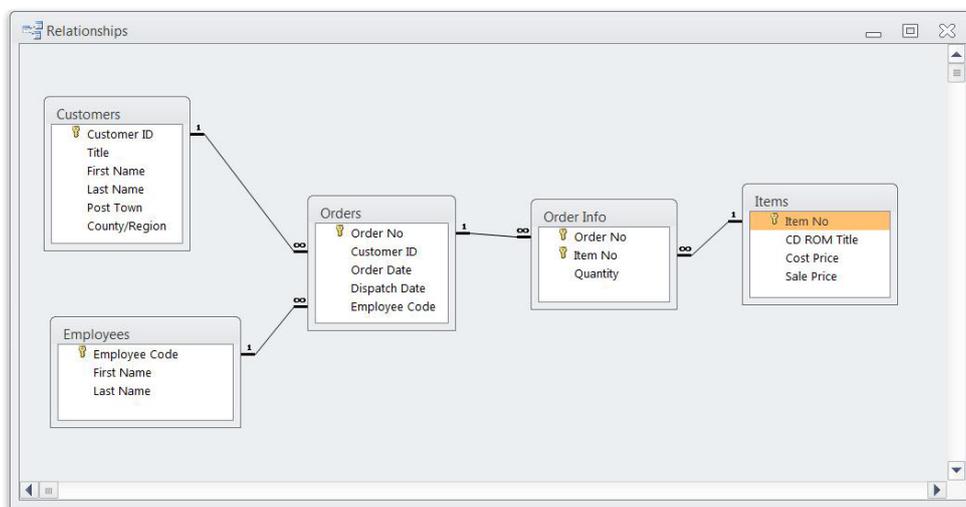


Sample data

It makes sense to practice with some of your own data (always a copy, never the original) because:

- You will have a better understanding of that data
- You may well actually uncover something interesting by analysing it

However we've also provided the sample Access database that we used. This data is entirely fabricated (no real people appear here). There are customers, products, sales etc, so it's standard sales data, available as a set of relational, mostly normalised tables.

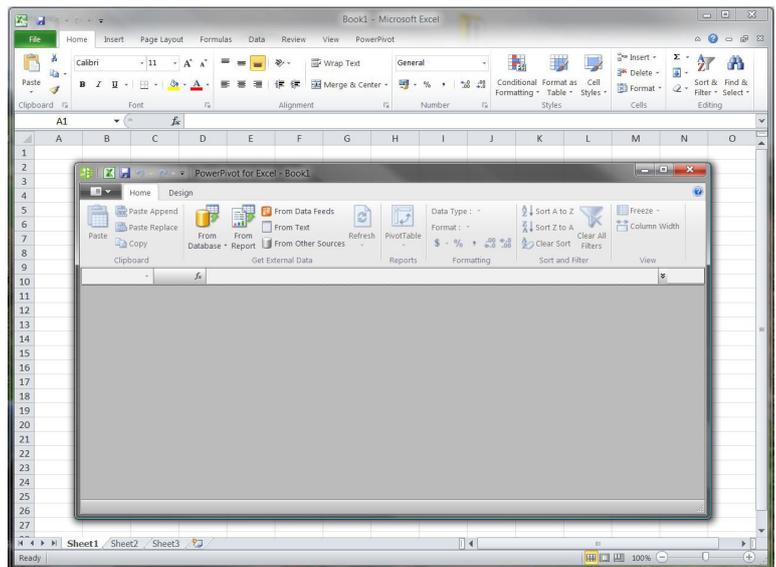


It's also available as a single query (called FlatFile) that pulls some of the data together into a simple flat file. Download the sample data from the PC Pro website.

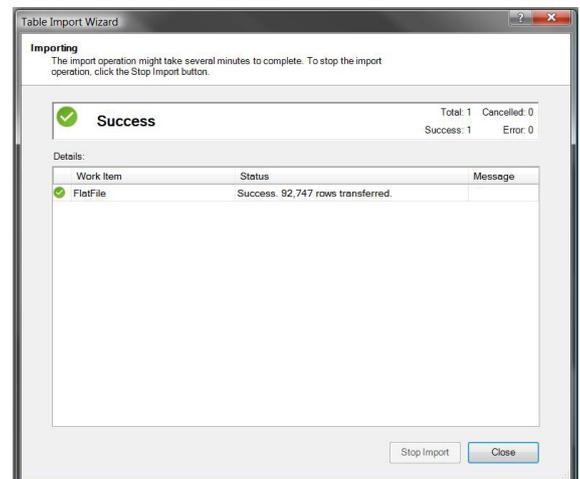
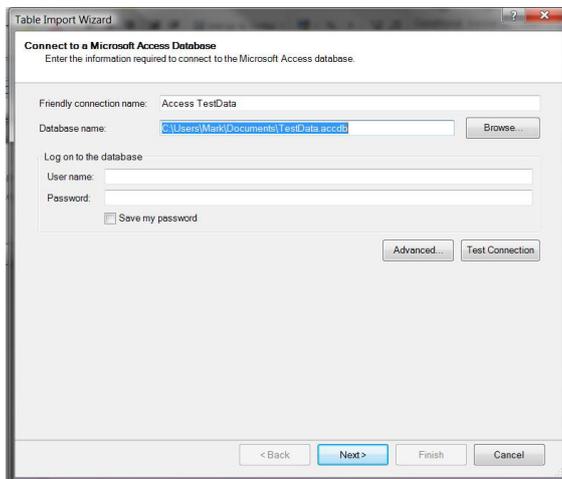
Connect PowerPivot to the data

First we'll connect to the simple flatfile of data so fire up Excel, click on the PowerPivot tab, then on the PowerPivot Window button (see the screen shot on page 2). This will open a new window, separate from the Excel window.

If you're running Excel full screen, it is not at all clear that this is a separate window. This can become a problem later on as we flip between PowerPivot and Excel. So I would recommend that initially you run both Excel and PowerPivot in non-maximised windows which makes it much easier to keep track of what's going on.



The next step is to establish a connection with a data source, in this case an Access database file. Click on the From Database button and choose From Access. A dialog box opens up: you don't have to provide a user name or password, just use the Browse button to locate the Access file (TestData.accdb), select it, click Open.



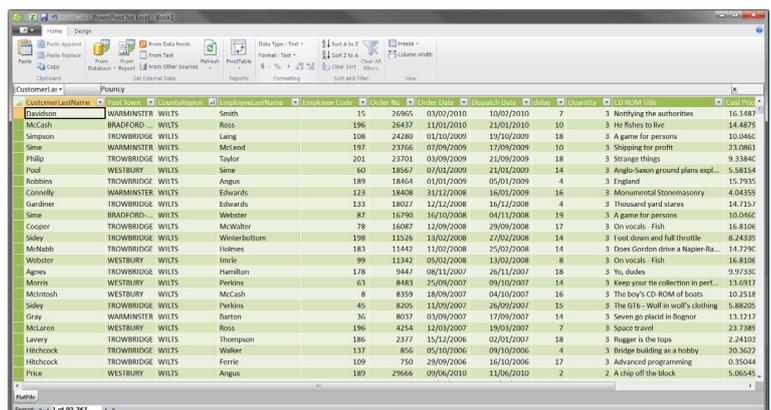
Then click Next. The default in the next section is fine so press Next, select FlatFile and Finish.

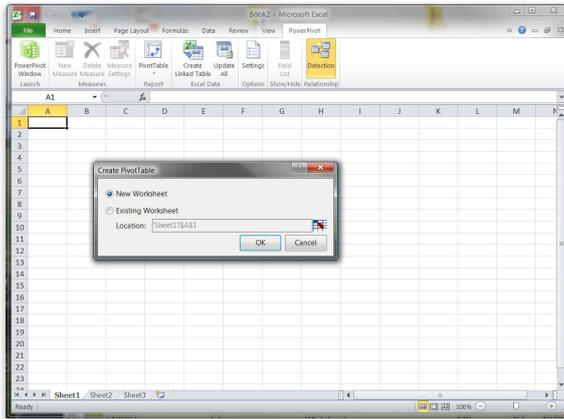
92,747 rows of data should be imported and you can click Close to get rid of the dialog box.

You now have a PowerPivot window full of data.

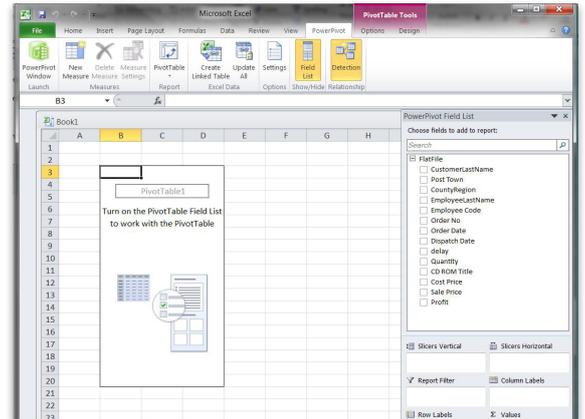
Using PowerPivot to analyse the data

To start analysing it, click on the PivotTable button. As discussed earlier, it may not be apparent but you will just have jumped back to the Excel window.

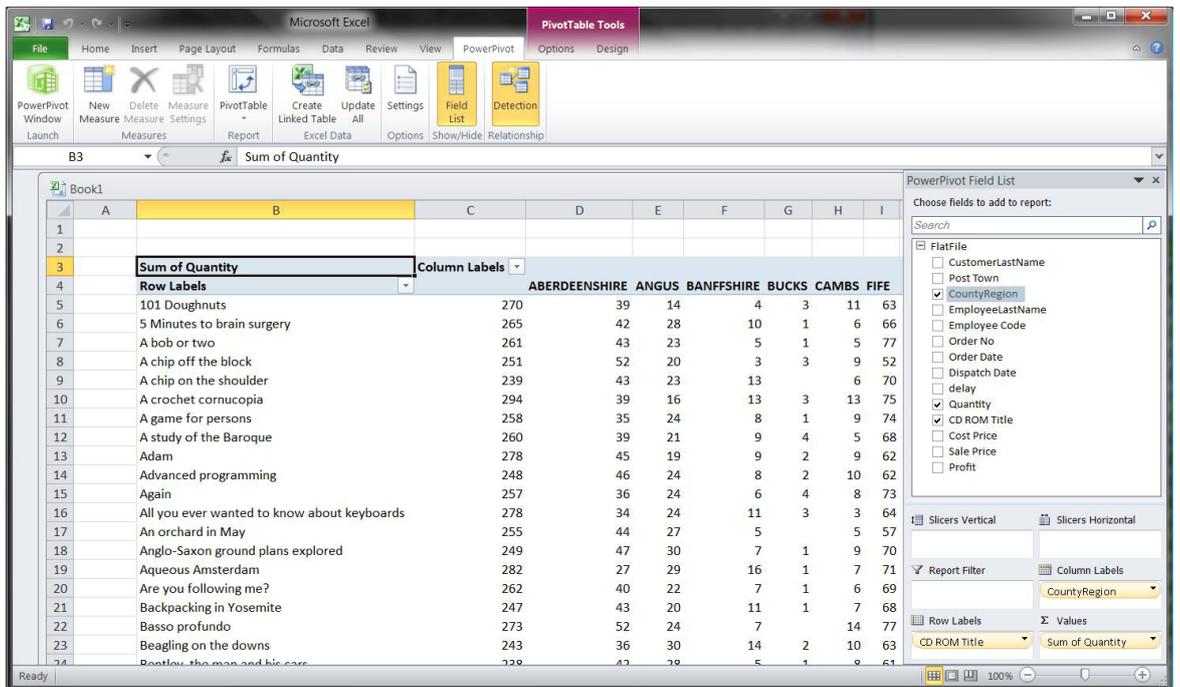




Click OK and you will find an environment that will look familiar to PivotTable users.

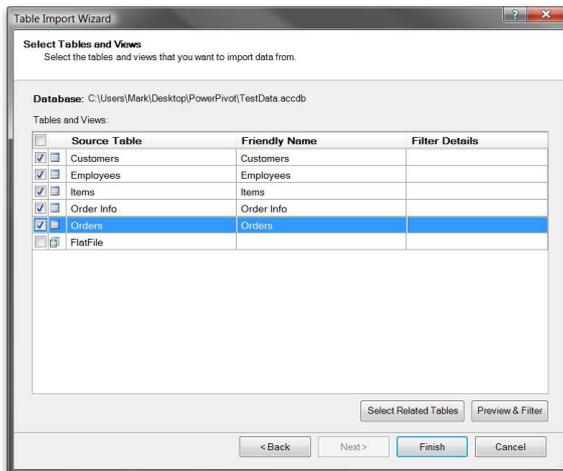


By making selections in the Field List and putting them in the correct places, you can analyse, as here, the sales of different products in different regions.

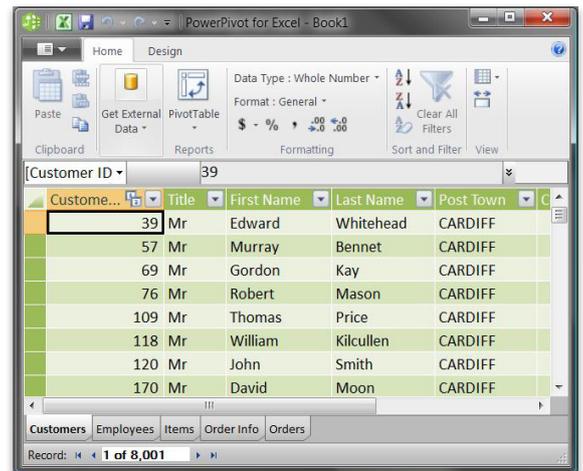


Seasoned PivotTable users can be forgiven for thinking that this is nothing extraordinary because they have been able to do manipulate data like this for years. The crucial difference (and it is crucial) is that PivotTable worked poorly (and painfully slowly) with large sets of data. PowerPivot is holding the data in an in-memory analytical engine (called Vertipac). On a 64-bit machine with enough RAM, this can manipulate millions of rows very, very fast. So this may look and feel like a PivotTable, but it's backed by a much more powerful engine.

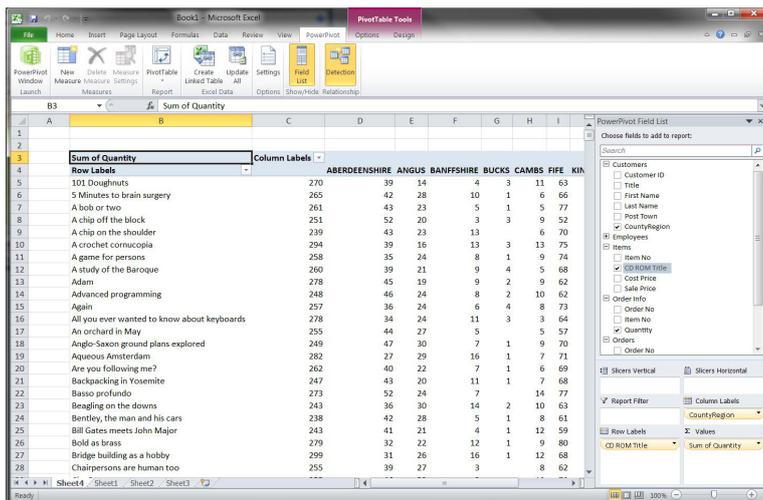
And PowerPivot has a host of other interesting tricks up its sleeve. We'll show you just one of them, but it is impressive. Start again, repeat the steps outlined above, except select the five tables as shown.



This will give you five tabs in PowerPivot.



If you then create a PivotTable as before, you can do the analysis as before. But note how the PowerPivot Field List has changed.



What makes this remarkable is that PowerPivot is reading the relationships between the tables and intelligently interpreting them for you. No-one had to create the query for you to bring the data from the relational tables together.

So an important implication of this is that PowerPivot enables you to reach data rapidly, without waiting for an IT person to make the data available as a query. And it lets you combine data from completely disparate sources.

So you can pull data from, say, your company's sales database and match it against, for example, weather data that you found on the web, to see if sales increase in hot weather.

Summary

PowerPivot is a non-trivial tool and almost painfully easy to use. If you want to know more, check out "PowerPivot for Excel 2010" by M Russo and A Ferrari, ISBN 978-7356-4058-0.

Further reading

PowerPivot Download the Add-in and discover more uses for its power.

PivotTables Overview of PivotTables and PivotChart reports.

Introduction to Excel 2010 Microsoft's guide to its hugely versatile software.