Intro:	Hello and welcome to the CPA Australia Podcast, your weekly source of business, leadership, and public practise accounting information.
Neale Blackwood:	In this month's podcast, I wanted to talk about Power Query in a little more detail. I have mentioned Power Query in previous podcasts, but in this one I wanted to go into a little more depth and cover a few more topics relating to Power Query. Now, what is Power Query? Well, Power Query is basically Excel's data import package now. It's built into Excel 2016 and later versions. It's an add-in to Excel 2010 and 2013. It did get a name change in 2016 to Get and Transform, but most people honestly call it Power Query.
	Now Power Query is also in Power BI. BI stands for Business Intelligence. Power Query is pretty much the same in Excel as it is in Power BI. The main different in Power BI is you're getting more updates regularly. So Power BI gets updated every month, whereas Excel is every quarter or six months if you're on the subscription version, and if you're buying the full version then it's every two or three years. It's got the same functionality in both, so it can be a good place to start.
	Talking about starting. The most important part of your Excel files is getting your data right. If you get the data right, structured right, then much of the reporting side of things is very straightforward. You can use things like pivot tables, you can use a lot of Excel's functions a lot more easily if you are using tables as your data structure. Now Power Query fits right in with that because what it can do is it can take the raw materials, let's say your CSV files, your text files, your databases, so you might have SQL databases, you might have other system databases, and Power Query can extract from those and bring the data into Excel, and then transform it. For example, you might have some issues with the way the dates come into Excel. So Power Query can actually fix those dates and make them real dates so they work just like dates in Excel, which sometimes they'll come in, dates might come in like text, and so you can't really use them in any sort of date related calculation. So part of Power Query's usefulness is that it will automate the importation process.
	So what it does is it first identifies the data, and so you can specify where it's coming from. It brings it in. Now, it sort of brings it into its own interface, and in that interface there are so many built-in ways to work with ... It can be called data cleansing as well. Basically fixing up the data. So things like missing codes, or you need to copy codes down a column because they're not in every row, or fixing the dates, removing leading and trailing spaces which has always been an issue for some databases. Well, that's automated. The beauty is it works very much like a macro in that you record the steps one after the other, so it's a very sequential operation, and then when you refresh the query it goes through and does those steps in sequence again. Now, it's even better than a macro because you can click on any of the steps, and it'll actually take you back to what the

data looked like at that stage. So it's very interactive, it's very user friendly, it's a lot easier to work with than macros, for example.

Now in the creation of the steps, it uses something called the M language. Now there's a formula bar there that you can learn the M language by just looking at it. If the formula bar is not visible, then you can go into the view ribbon tab in the Power Query window and you can make the formula bar visible, just in case it isn't there. The M language is there, but you don't really need to use it because just following the interface will actually create the language. But the more and more you use Power Query, the more likely it is that you might need to tweak the language. So you can use the formula bar to learn the language, and again there's more and more resources online that can help you understand it.

Ultimately there will be some intellisense coming out. Intellisense is where you start tying and the app will actually help you and show you the words that you could be meaning. There's none of that at the moment with the M language, you've got to know what you want to type, whereas the intellisense is coming, it is a future upgrade. I think it is part of ... It might be a preview in Power BI already in the Power Query side of things. So it will be coming, you just have to be patient there.

So you can use almost anything as your data importation source. So you can use existing Excel data tables, most external data sets or databases can be accessed through Power Query. Obviously you'll need your IT people involved to give you the access and also to give you where they are, and what the passwords are, and things like that. But technically it can import data from most data sets.

One of the beauties of using Power Query, and it's something I'm going to cover in an upcoming article, is mapping tables. So one of the areas that you can really make strides in your reports is bringing data together from separate databases. They tend to be siloed in the past, and so you couldn't report on them together. But mapping tables allow you to basically bring the data together and create reports that can bring out relationships and trends that you might not be able to pick up from each data set in isolation. So there's an article in the April edition of INTHEBLACK which will cover mapping tables and how you can use those to bring your data together.

Also, relationships are in the data menu from Excel 2013 onward. Relationships allow you to create relationships between tables using the mapping tables, and like a relational database in terms of being able to report based on multiple tables. There are a few earlier articles that I've written on Power Query as well, and it's all on the INTHEBLACK website, just do a search for that and there will be some links in the show notes as well.

So the types of things you can do to cleanse the data is removing the top or bottom rows. So you can specify how many. You can remove columns. If you don't need the column, it can be a waste of a resource to bring it in, so you can

remove the columns to only bring in the columns that you require. You can specify the column type, so you can say it's a date, you can go text, numbers, all that sort of thing. You can transform. So the transformations that you can apply does change based on the type of data. So the types of things you can do with dates is you can convert it into a month or a year. The text, you can as I mentioned, fix the leading and trailing spaces using trimming. You can also truncate, so you can actually only bring in certain length of text if you need to. Things like that. So you can maybe just bring in the first 20 characters if you've got some sort of description. Again, depending on what you're trying to achieve. You can replace things, so maybe you've got some characters in there that you don't want, so you can replace them with nothing or you can replace them with something else. That's all automated. All of those steps that you do are captured by Power Query and just replayed when you basically refresh it.

So once you've gone through and made all the changes, you can basically load it to your spreadsheet. There's a couple of options there. You can load it as a table within the file, so you'll seal the data. You can also load it into what's called a connection only. Now connection only means the data's captured, but you don't see it. So it's in the memory, if you like. So it's in there, it's part of what's called the data model, and you can use it to report on it, you can run a pivot table off it, but it's not taking up space in terms of putting data within in a sheet. Again, the connection option is often useful when you start using Power Query more and more, you'll actually find that you'll have steps in your sequence and you might create some data sets that you don't really need as an end, they're just a means to get to the end. So they can just go as a connection, and you can use them later.

You can split columns. So if you've got columns that there may be some codes in there that you want to split up, you can do that. You can fill up, you can fill down. So where there's some reports that come in, they have a code on one row and that code is supposed to be on the rows below it down to the next time there's another code, and you can fill up and fill down to fix those type of things.

A useful ability is the ability to create what's called a calendar table in Power Query. Calendar tables allow you to take basically any transactional data and convert them to things like months, or years, financial years, that can be the other thing, quarters, financial year quarters. So a calendar table allows you to achieve those types of things.

You can do groupings. So you might have a very large data set that's transactional, and you might want to reduce the size of that data set. Well, what you can do is you can group. For example, you could group by date. So you could ... It's like a summary if you like, so it summarises by date, and then you specify the columns that you want it to group by. So it might be, say product, and value, and things like that. So what it does, is it takes a huge data set and it groups it by basically summarising it into a date-based grouping based on whatever columns you've told it to do, to group by. So it's a very powerful feature, especially when you're working on large transactional data sets and you

still want to bring it in, but you don't want to take up the resources of hundreds of thousands rows.

Another feature that is part of the cleansing is something called unpivot. It's an unusual name, it's obviously related to pivot tables. Here's how it works. So if you take the standard budget layout. So let's say we've got accounts going down the left-hand side, we've got the months of the year going across the page, and then you've got values underneath those. Now that layout is a typical budget layout, but it's actually more a reporting layout. That's not how you present data. Data tends to be column-based and going down the page. A budget layout is like a pivot table, and so what you can do is take that type of layout where you've got ... Let's take months as the example. So you've got months going across the page. The most efficient way to create that in a database way is to have a column for months, and then a column for value. So imagine it as a very thin database, but it's very long. So there's lots of transactions. So for one particular account for example, you might have 12 months worth of data in the budget, but in the data set that would actually be 12 rows. So there'd be a row for July, August, September, etc. Okay? That's what unpivot does.

So what you can do is you can specify the columns that you want to unpivot, and then what it does is it converts it from, if you like a wide report, into thin data. It does it at the click of a button. In the old days, you used to have to use a macro to do something like that and it took a while to develop, where this just happens at the click of a button. Depending the way you want to work it, you can actually specify the columns you want to keep and then unpivot the other columns as well. So it's a couple of ways to do that unpivoting. But that's what unpivot means. Basically converting what would be called a report layout into a data layout, which is much easier to work with, especially with pivot tables and things.

You could add extra columns to your data. There's a brand new option called based on examples. So you might not know how to get the data in the structure that you want, in terms of a column. So what you can do is you can give it some examples, and depending on what you're trying to do, it might actually be able to write the M language code for you and create the extra code. So that's a brand new feature that you can have a play around with.

Merging. That's something else. You can merge two tables together to bring the data into a single table. That's also what I'll be covering in the April article in the INTHEBLACK Magazine. When you are merging, by the way, good idea to keep the data types the same of the columns you are using to merge. By the way, you can even merge on multiple columns. Usually you have one column, and that's the column that it looked up, if you like. But there is the ability to hold the control key down and you can merge based on two columns, or more columns, you just need to do them in sequence.

Something else you can do with the data is append it. Now that's very common. What you might do is you might get, let's say a CSV file, and you might get one

CSV file per month. So that single CSV file contains all the data for that month, which is okay because you might be doing a monthly report. But you might want to do a year-to-date report as well. So what you can do with the append operation is basically take, let's say you've got all 12 CSVs for the year, you can just put them one under the other. An easy way to do that is ... There's a couple of ways to do that. You can create multiple queries and then append each one, or the easy way is all you do is drop all of the CSV files in the one folder, and then you can get Power Query to basically append each file, one under the other, and it pretty much does it for you. I think I even covered that in an article a couple of years ago, on actually using multiple CSV files. So you can check that out on the INTHEBLACK website as well.

As I mentioned before, you can get data from Excel files. So one of the things that Power Query can be used for is budgeting. So for example, you might create a ... Let's say we create a single sheet budget that you send out to your managers. So you can send that out, they fill it in, and they send them all back. Now typically that's a cut and paste job that you've got to do to actually bring the data together. But again, you can put all of those returned files in a folder, and then run a Power Query on it to be able to bring all of that data back together again. Now again, the steps might be that the first thing is to bring the data together, and again, typically the layout that you would use for inputting a budget would have let's say accounts and names going down the left-hand side, and again the months of the year going across the page. Now that looks good and it's understandable, but it's not a good data structure. So again that unpivot can work.

So you can initially bring all of the reports together from all of the departments, or states, however you break up your data. Then once you've brought it together, then again, you just run the unpivot to create a listing that has the month as a column, and then you'll have a value column as well. Then again, that data tends to be long and thin. That can also be needed to upload into your accounting systems. Again, different accounting systems take different things. Sometimes accounting systems will actually take a layout that looks like the budget going across the page, other systems need it in more of a data layout, which again is the long thin type of layout where you have a column for month and you have a column for value.

So I've covered a whole lot of stuff. Power Query is one of the most exciting things that's happened to Excel in the last 10 years. There's some exciting things coming as well, so there's some brand new functions coming that are really, really powerful. So keep an eye out for that. In terms of getting the data into Excel, then Power Query is definitely worth learning. I highly recommend any accountant should learn Power Query because it can basically simplify getting the data in, and it also gets the data into Excel in a table layout, which so many of Excel's functions and features work well with tables. So Power Query be part of getting the data right. Then you've got things like pivot tables which can allow you to do the reports, etc.

So I hope I've piqued your interest in Power Query. There's lots of resources online where you can learn Power Query. It's also reasonably intuitive, so you can just jump in and have a play and see. By the way, you're not affecting the data in any way. So for example, a CSV file that's there, you are not changing the CSV file in any way. So you can have a play, if it doesn't work, just start again. So there's not real dilemma there.

Oh by the way, there is one thing about Power Query that may be a little unusual. There's no undo. Okay? So you can't undo. There's no undo button, put it that way. You can undo things, because you can delete things, but you can't hit a button and have it undo what you last did. You can delete the step that you've just done, but just be a head's up, there's no undo, which is a little unusual these days.

So thanks for listening, and hopefully you'll have some fun with Power Query.

Outro:	Thank you for listening to the CPA Australia Podcast. To download the transcript and to access the show notes for this episode, please visit www.cpaaustralia.com.au/podcast/98.