ICT394 Business Intelligence Application Development

# Lab 02: Analyzing Data in Power BI

This Lab will continue our exploration of Power BI Desktop. We will be loading and relating data from several different sources and performing some more complex analysis than we did in Lab 01. This will also give you the opportunity to become more familiar with the Power BI Desktop (PBID) interface which is similar to several other products in this market.

**Please note that you will be required to upload a screen shot to Moodle in the quiz, and may need to answer some questions based on the outcomes of the Lab.**

You will also be asked in the quiz for one or more values from the charts you create in this Lab. As such, you should complete the quiz on a machine that has Power BI Desktop installed and running with your .pbix.

## Aims:

At the completion of this Lab you should be able to:

* Connect to several external data sources in Power BI Desktop
* Use Relationships to combine data from multiple data sources
* Demonstrate a basic understanding of the “shaping and combining” processes in Power BI Desktop
* Build a visualization that will allow you to interact with the data in order to filter the data in order to see more details

### For this Lab, you will need:

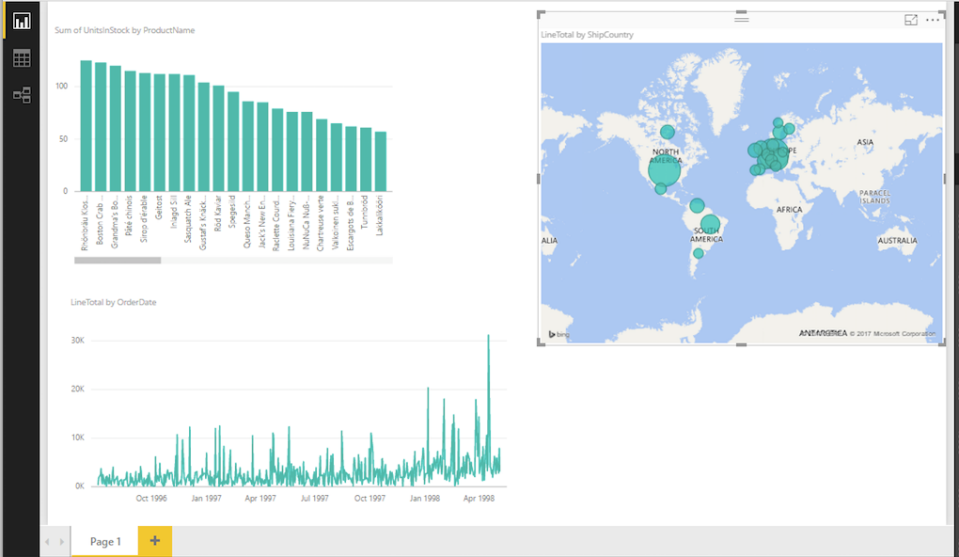
* Power BI Desktop installed
  + This is installed in the lab for internal students
  + It is a free download, so my suggestion is that all students should try to download and install it on their machines themselves.
  + NB: there is no MAC or Unix version
  + There are a number of installation guides for Power BI Desktop, see: <https://www.microsoft.com/en-us/download/details.aspx?id=45331>

This Lab is based largely on “Tutorial: Analyzing sales data from Excel and an OData feed”

* <https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-tutorial-analyzing-sales-data-from-excel-and-an-odata-feed/>
* Please note that this document includes background to assist you with the lab exercise. You should follow the steps in the tutorial above. Where there are differences between the tutorial instructions and the latest version of PBID, I will highlight them here.

## Introduction

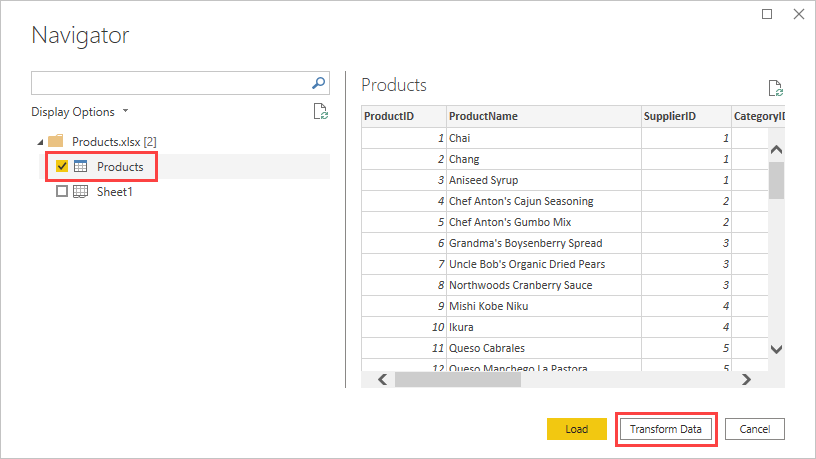
The report you are being asked to create in this lab will look something like this.



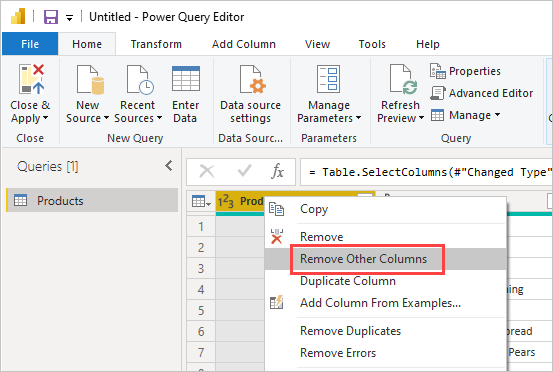
### Task 1: Get data from an Excel workbook

You will need to download the file “*Products.xlsx*” from <https://download.microsoft.com/download/1/4/E/14EDED28-6C58-4055-A65C-23B4DA81C4DE/Products.xlsx> and save it to your local machine.

🡪 Once you have downloaded and saved the file, Follow the instructions in **Task 1: Step 1: Connect to an Excel Workbook**. Make sure that you select **TRANSFORM DATA** rather than LOAD (NB: the online instructions suggest you should select Edit … this is what TRANSFORM DATA was called in an earlier version of PBID).



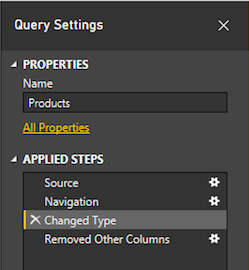
In Step 2, **you are removing all of the columns** from Products.xlsx **apart from** ProductID, ProductName, UnitsInStock and QuantityPerUnit. To do this, you select these columns and then Remove Other Columns from the Manage Columns tool.



In Step 3, **if PBID has not already determined that the UnitsInStock column is a whole number** type, you will need to change the datatype of the UnitsInStock column to Whole Number.

At this point, you should take the time to look at the steps that have been created and applied. Note that in the background, Power BI is creating code that will be executed against the data in the order you have specified. The code is in a language called M. You do not need to bother too much with that now but may do later in the unit.

**You will notice** the steps created might be different to those in the screenshot in the tutorial. It is not a problem if this is the case. Mine looked like this

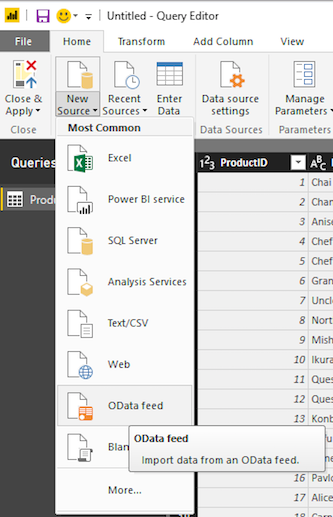
****

### Task 2: Import order data from an OData feed

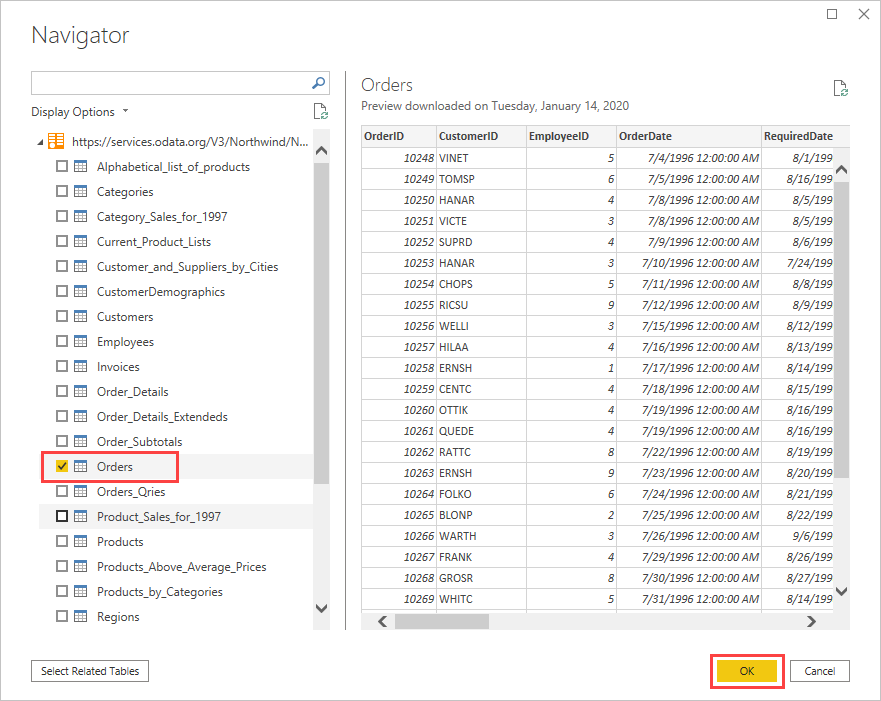
In this task, you will be importing data from an Open Data Protocol (also known as OData) feed. To learn more about OData, see: <https://en.wikipedia.org/wiki/Open_Data_Protocol> or if you have time, see <https://youtu.be/EZU-7k_Wv_s>

In Step 1, you will be connecting to the Orders table from the Northwind Odata feed at <http://services.odata.org/V3/Northwind/Northwind.svc/> (Northwind is a sample dataset that Microsoft has used for many years for its data-focused products such as MSSQL, Analysis Services etc).

In Power Query Editor, select **New Source** and then, from the **Most Common** menu, select **OData feed**.

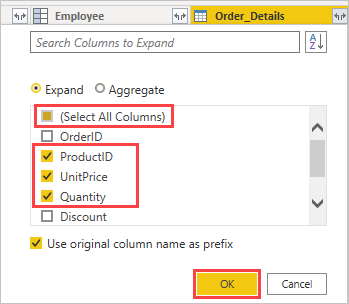
****

In Step 2, you will be “expanding” the **Orders table** to include data from the Order\_Details table. If you think of Orders and Order\_Details as having a 1:M relationship, you will be expanding the Order details with the data from the many side of the relationship – Order\_Details. Note, **there are several other Foreign Key values in the Orders table**.



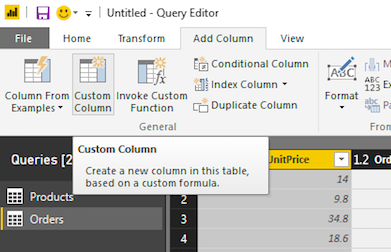
**Write these down** because you might need to know which they are **for the quiz**.

You can use the expand operation to add the **ProductID**, **UnitPrice**, and **Quantity** columns from the related **Order\_Details** table into the subject (**Orders**) table. Select **ProductID**, **UnitPrice**, and **Quantity**, and then select **OK**.

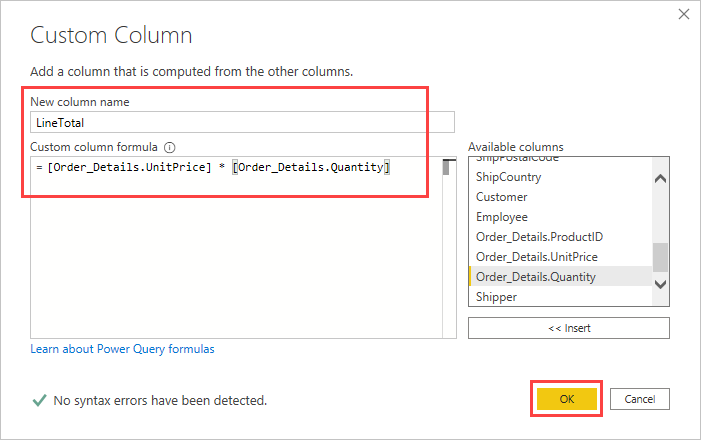


In Step 3, you will be removing several columns so you will only have the columns of interest.

In Step 4, you will be adding a calculated column so you can calculate the line total for each of the Order\_Details row by multiplying the **Quantity** value by the **UnitPrice**.

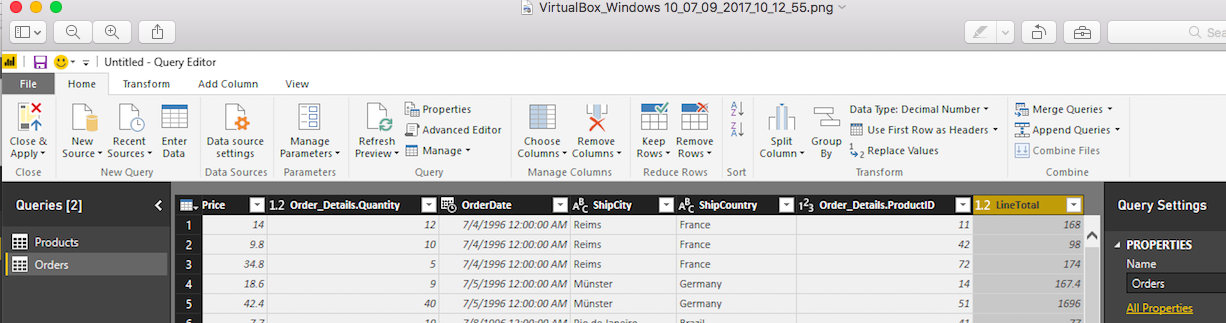


In the **Custom column formula** field after the **=**, enter **[Order\_Details.UnitPrice]** \* **[Order\_Details.Quantity]**. You can also select the field names from the **Available columns** scroll box and select **<< Insert**, instead of typing them.



Step 5 has you changing the value of the new column to **Fixed decimal number**, and Step 6 renames the columns.

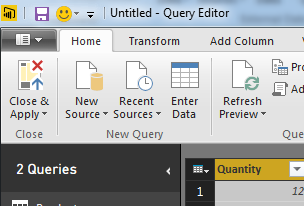
You should now have something that looks like this:



Step 6 has you renaming and re-ordering the columns in the query.

### Task 3: Combine the Products and Total Sales queries

This task has you creating a relationship between the Products and Orders queries you have created so far. To do this, you will load the two queries in Power BI Desktop (all you have done so far is to specify how the data sources will be queried when they are loaded; this step executes the queries and loads the data for analysis), and follow the instructions to create the relationship. (In my version of PBID, Close & Load as shown in the tutorial, is called Close & Apply).



### Task 4: Build visuals using your data and interact with the report to further analyse the data

This task requires you to create several visualisations using the data and queries you have created. You will then use cross-highlighting and filters to examine the trends that are present in the data.

## For this topic’s quiz!

You will be asked in the quiz for one or more values from the charts you created in this Lab. As such, you should complete the quiz on a machine that has Power BI Desktop installed and running with your .pbix so you can check your values.

## Challenge Task

For this challenge, see if you can create a chart that shows the line totals for the quarters of 1997. Paste your chart to the discussion forum in Moodle so we can discuss how you did it!