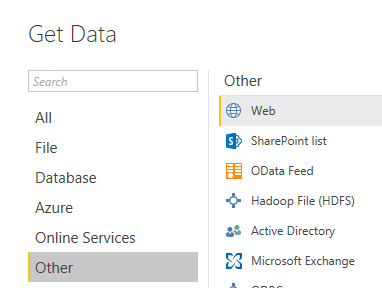
Loading Data with Power BI Desktop

# Exercise 1

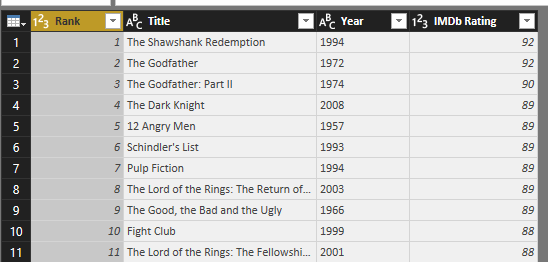
We’re going to read data in from IMDB: the top 250 movies. You can find the list on <http://www.imdb.com/chart/top>. Use the Get Data > Web feature to scrape the data from the website.



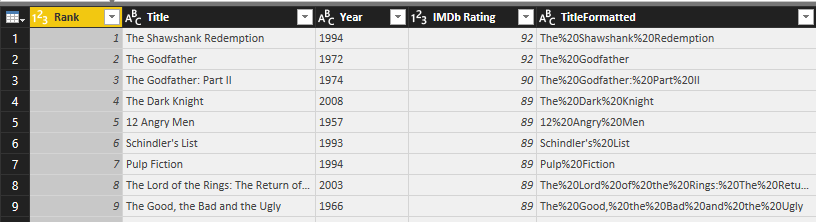
Multiple tables will be returned. Use the preview pane to determine which table you want to load. Then click on **Edit** to go to the query.

* Only keep the columns “Rank & Title” and “IMDB Rating”.
* Change the data type of each column to a more appropriate data type, if necessary.
* Extract the year and the rank out of the “Rank & Title”. The easiest options are to split on delimiter or to extract a specific number of characters from the end or the start.
* Clean up the resulting columns (use the TRIM function) and assign appropriate data types.

The result should look like this:



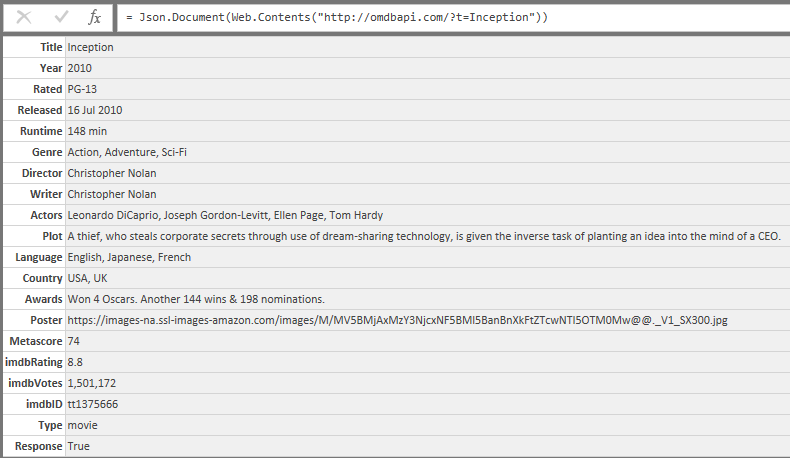
To prepare for the next exercises, we’re going to duplicate the Title column and replace the spaces with “%20”. We will pass this column along to a web service to find more information about the movie title. In order to do this, the title needs to be formatted as an URL. The result should look like this:



# Exercise 2

We’re going to create a function that will call a web service and retrieve more information about a movie title. The parameter of this function will be the movie title: the column we prepared in the previous exercise.

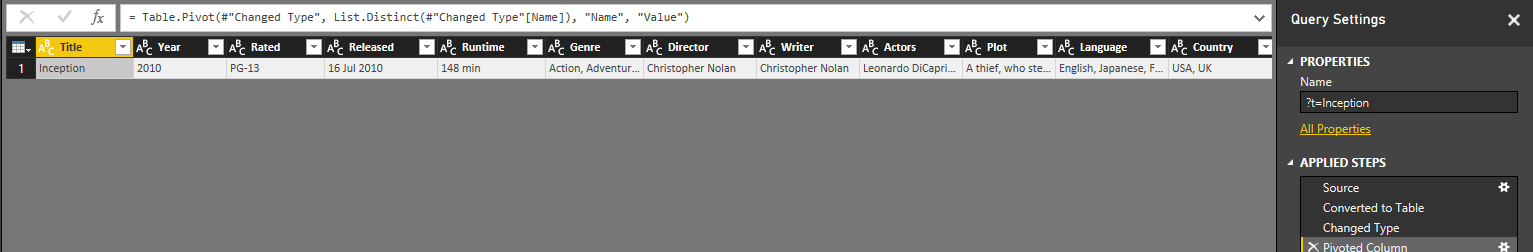
Create a new query, again using the **From Web** source. Enter the URL <http://omdbapi.com/?t=Inception>. We’re using a link with a hardcoded movie title for now, but we will change the movie title to a parameter later on. The results from the web query look like this:



We’re going to turn this into a useable result.

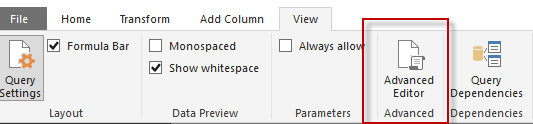
* In the ribbon, the *record tools - convert* tab will be selected. There’s only one button: **Into Table**. This will turn the JSON result into a normal table.
* Convert the Value column to text.
* Pivot the columns to get all the data into one single line. Tip: check the advanced properties of the transformation.

The query now has the following format:



Let’s turn it into a function. Since the latest release, there are ways to do this using the user interface only. However, editing the M code directly is faster and creates less clutter.

In the View tab, click on **Advanced Editor**.



This will show the entire M code of the current query.

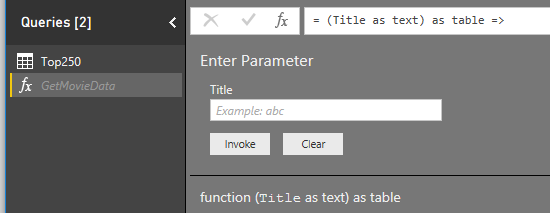


Turning a query into a function is easy. You need to embed the current query code into a function declaration:



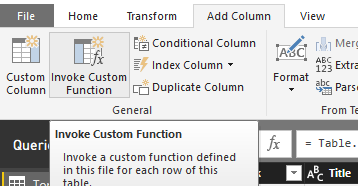
The function in this example is called GetMovieJSON. *Title* is a text parameter and the result is a table. In the first line of our original query, change Inception to parameter Title.

When you click on Done, the query will automatically change to a function:

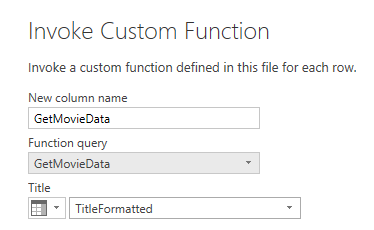


You can test the function by entering a movie title and select *Invoke*. This will create a separate query with the result.

Now we can use this function to get more information about all the movie titles. Let’s go back to our original query. In the Add Column tab there is an **Invoke Custom Function** button.

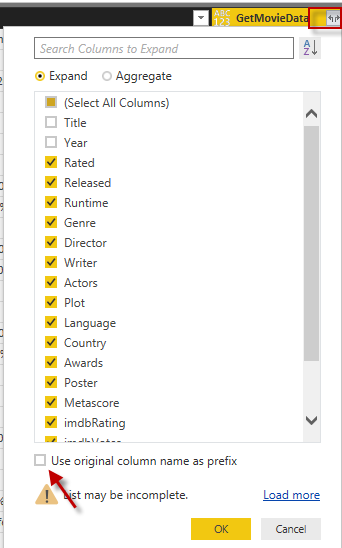


With this transformation, we can call our function for every single row. The result will be an embedded table inserted in each row.

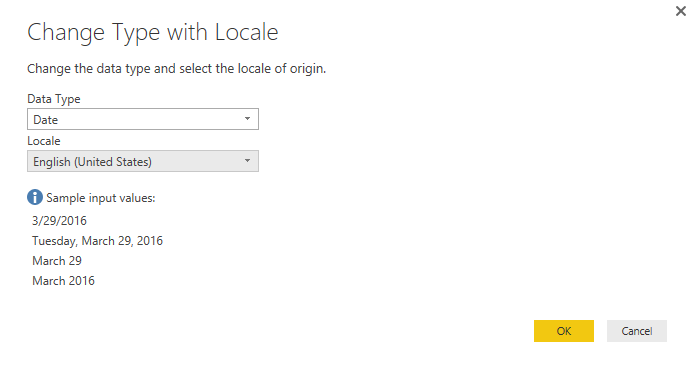


Keep in mind the web service must be called 250 times and the query behind the function will be executed 250 times. In other words, grab a cup of coffee.

By clicking on the double arrow icon in our new column, we can extract the columns generated by the function. Almost all column can be used, except Title, Year (we already have those), imdbID, Type and Response. Also drop the prefix of the original column (see checkbox at the bottom).



Convert all new columns to their appropriate data type. It’s possible that for some columns, you should use a locale. This depends on your computer settings and the settings in Power BI Desktop.



For the Released and the Metascore column, you first need to replace N/A with null, before you can assign a proper data type. The Runtime column should be converted to an integer column, so it can be used in calculations later on.

Finally, the data can be loaded to the model.