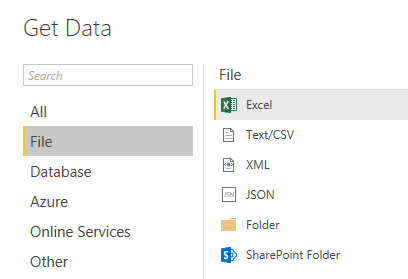
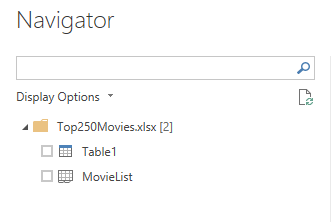
Loading Data with Power BI Desktop

# Exercise 1

We’re going to read data about the top 250 popular movies. The list itself is stored in an Excel file. Start up a new query and select Excel from the list of possible data sources.



Navigate to the directory where the file *Top250Movies.xlsx* is located. When you select this file, it’s possible 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.



* Promote the header.
* Split out the rank and the title of the movie. The easiest options are to split on delimiter or to extract a specific number of characters from the start.
* Clean up the resulting columns (use the TRIM function) and assign appropriate data types.

The result should look like this:

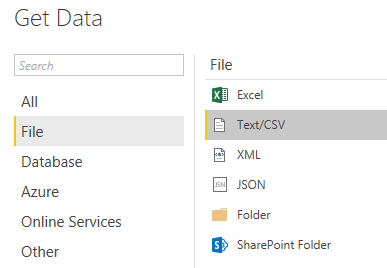


Rename the query to *TopMovies*.

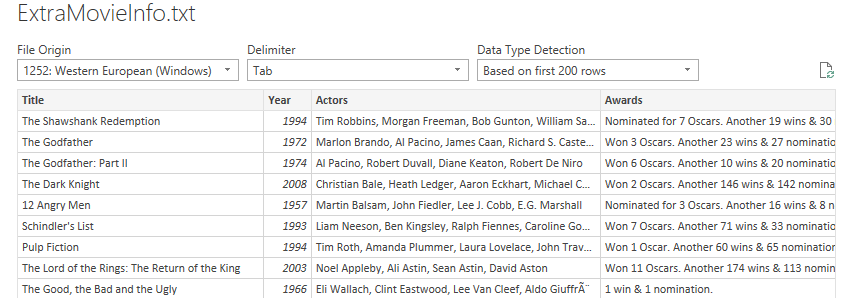
# Exercise 2

We’re going to create a new query which will fetch more information about our movies.

Create a new query, but this time read from a **text file**. Navigate to the *ExtraMovieInfo.txt* file. This is a tab-delimited file.

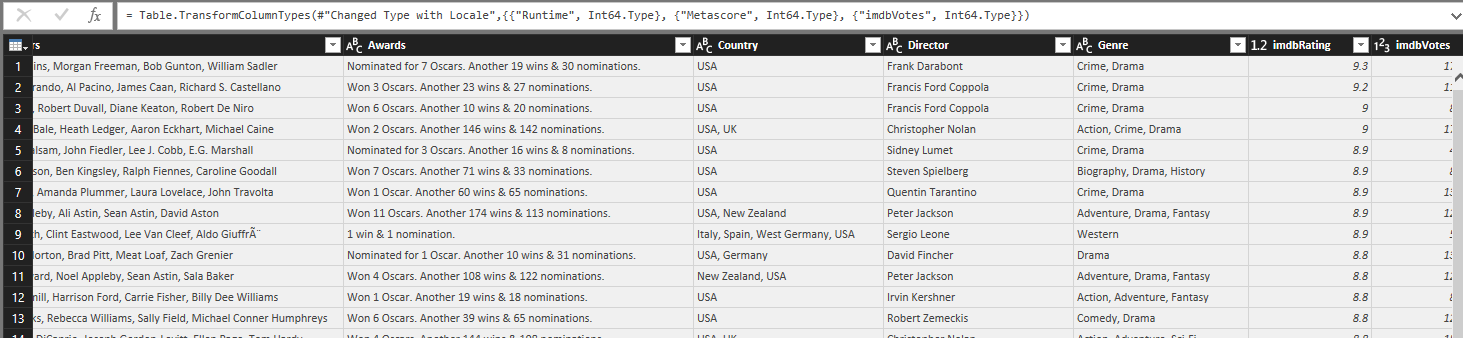


Power BI Desktop will automatically detect the nature of the file: the delimiter, but also the code page used to save the file.

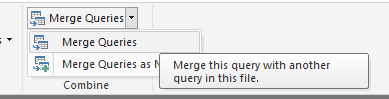


* Remove the step *Change Type*. This step was automatically added and it converts all columns to its detected data type. However, most columns are text and don’t need conversion. Also, depending on your locale – Power BI can make a mistake in the *imdbRating* column: a value of 9 will be converted to 9 but a value of 9.1 to 91. When you select the file, there is a dropdown for *Data Type Detection* (see previous screenshot), but turning it off still creates the *Change Type* step.
* If the headers were not promoted, make sure they do.
* Convert the column *imdbRating* to the decimal number data type and the column *Released* to the date data type, taking into account the locale used for both columns.
* Convert the columns *Runtime*, *Metascore* and *imdbVotes* to integers.

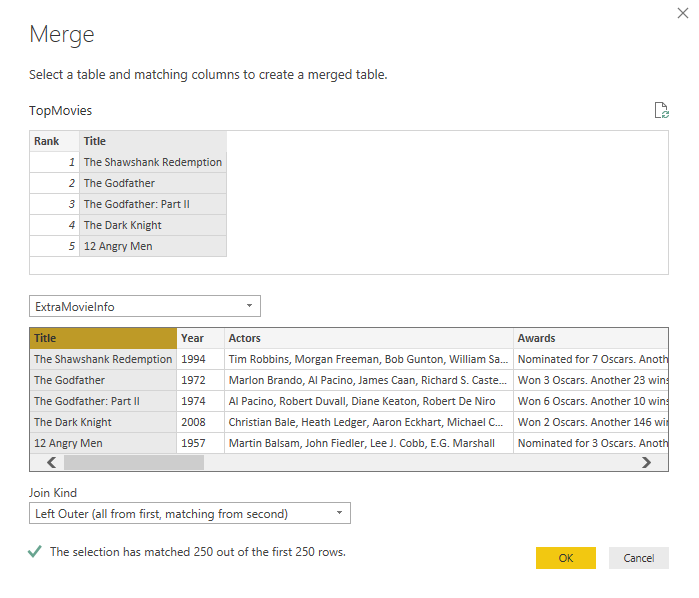
The query now has the following format:



Rename the query *ExtraMovieInfo*. Let’s merge the query from exercise 1 with our newly created query. While the query TopMovies is selected, click on *Merge Queries* in the ribbon.

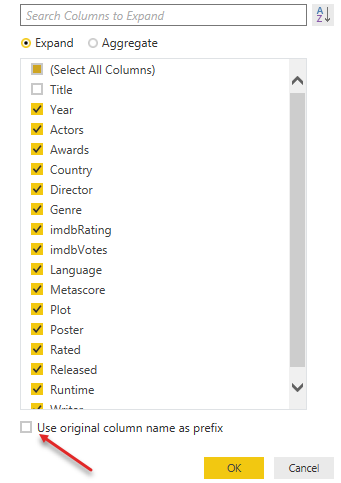


In the Merge Query editor, select the ExtraMovieInfo query. In both queries, select the *Title* column. These columns will be used to join the two queries together.

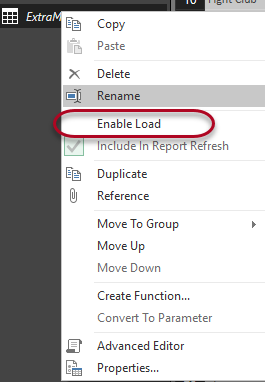


The default *Join Kind* can be kept to **Left Outer**, but **Inner** will work as well.

By clicking on the double arrow icon in our new column, we can extract the columns from the ExtraMovieInfo query. Almost all columns can be used, except Title since we already have that column. Also drop the prefix of the original column (see checkbox at the bottom).



You can disable the load for the query ExtraMovieInfo since all of its data is now included in the TopMovies query.



Finally, the data can be loaded to the model.