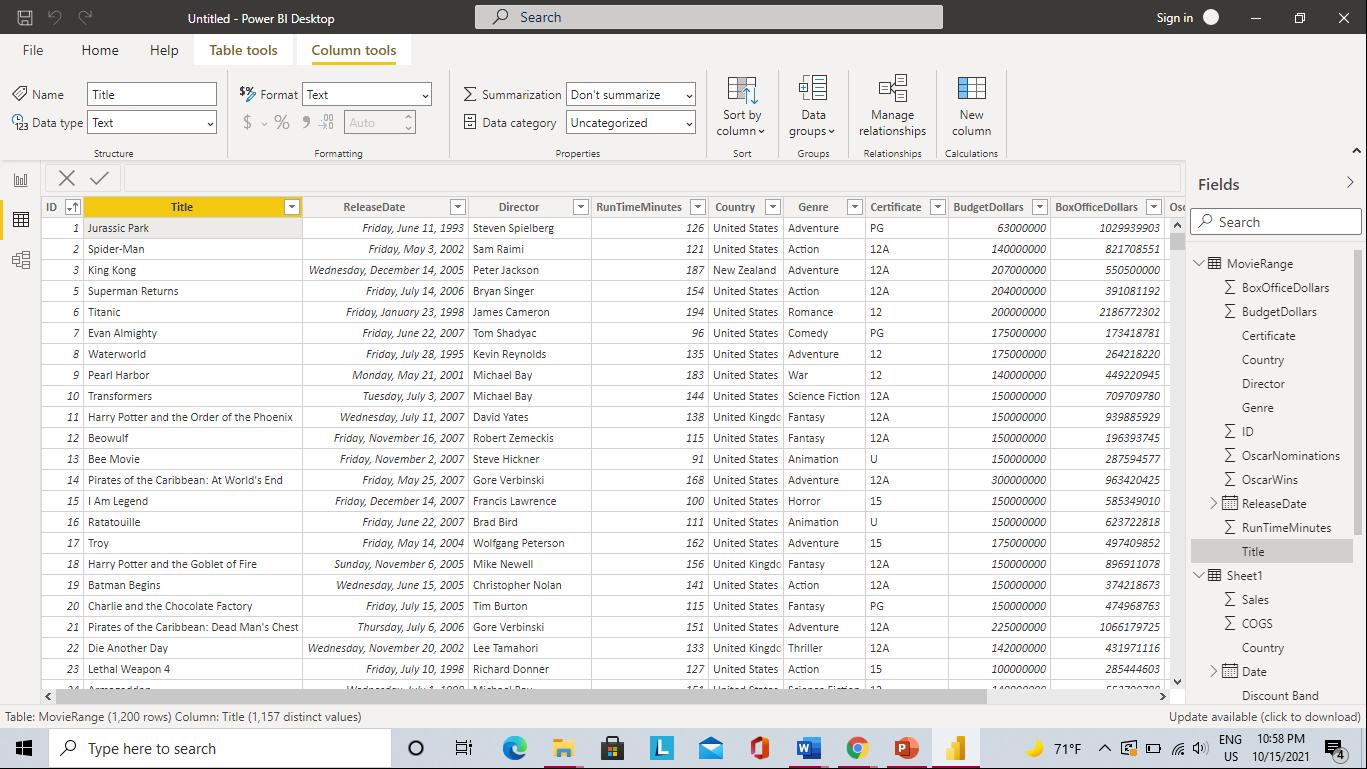
IT 7123 – Power BI Lab 3:

Data Preparation

By Anna Kurenkov

4. C. WiseOwlTutorial

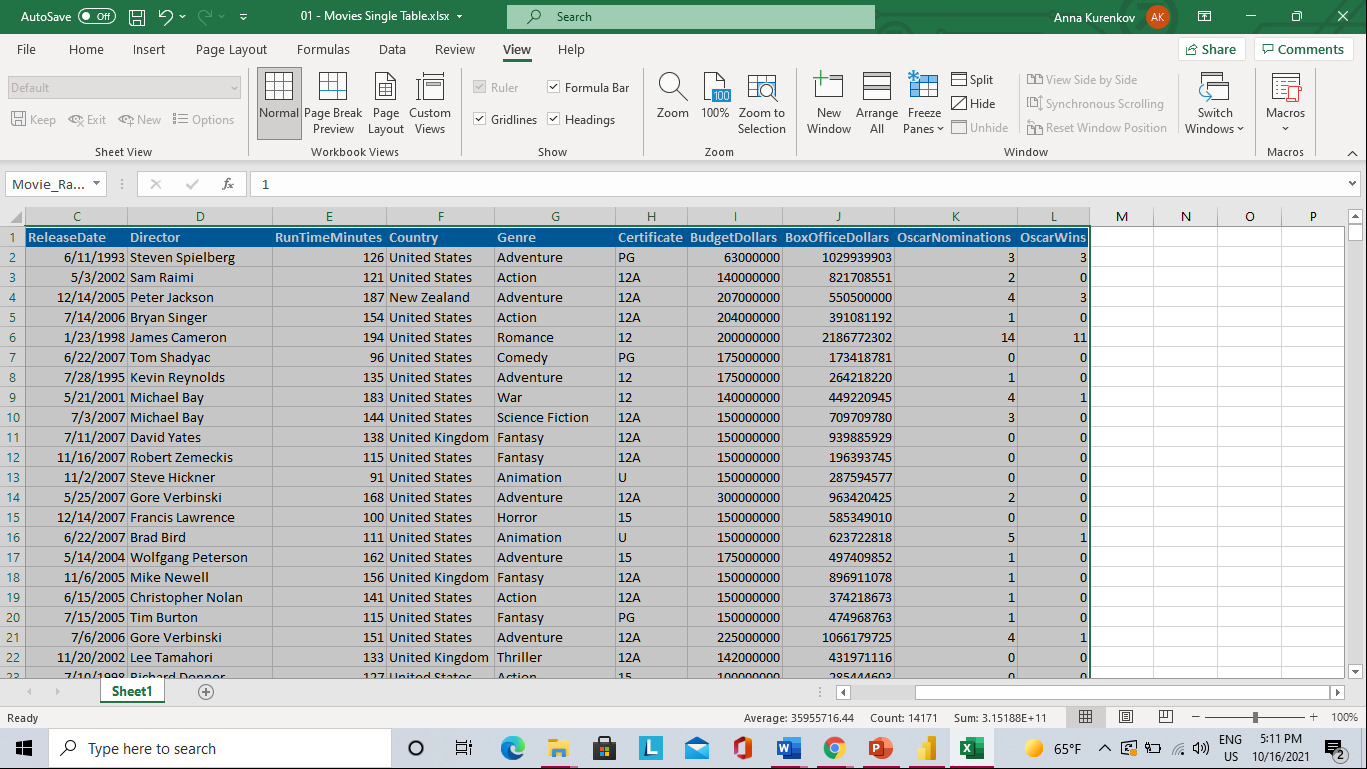
3.1



The sheet from the excel file was imported.

The first modeling activity was changing the date format to a shorter, numeric one as well as changing budget values to currency format.

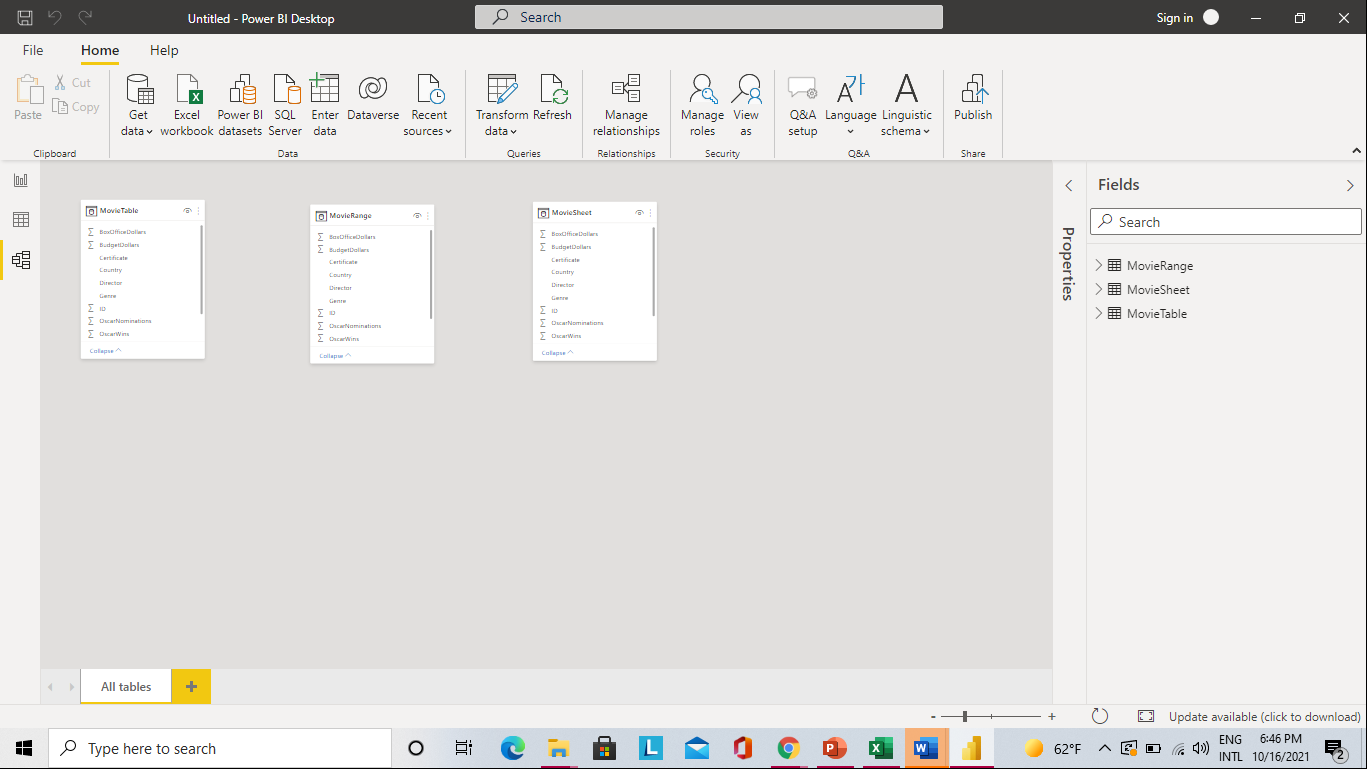
In excel a custom range was created and called “Movie\_Range”



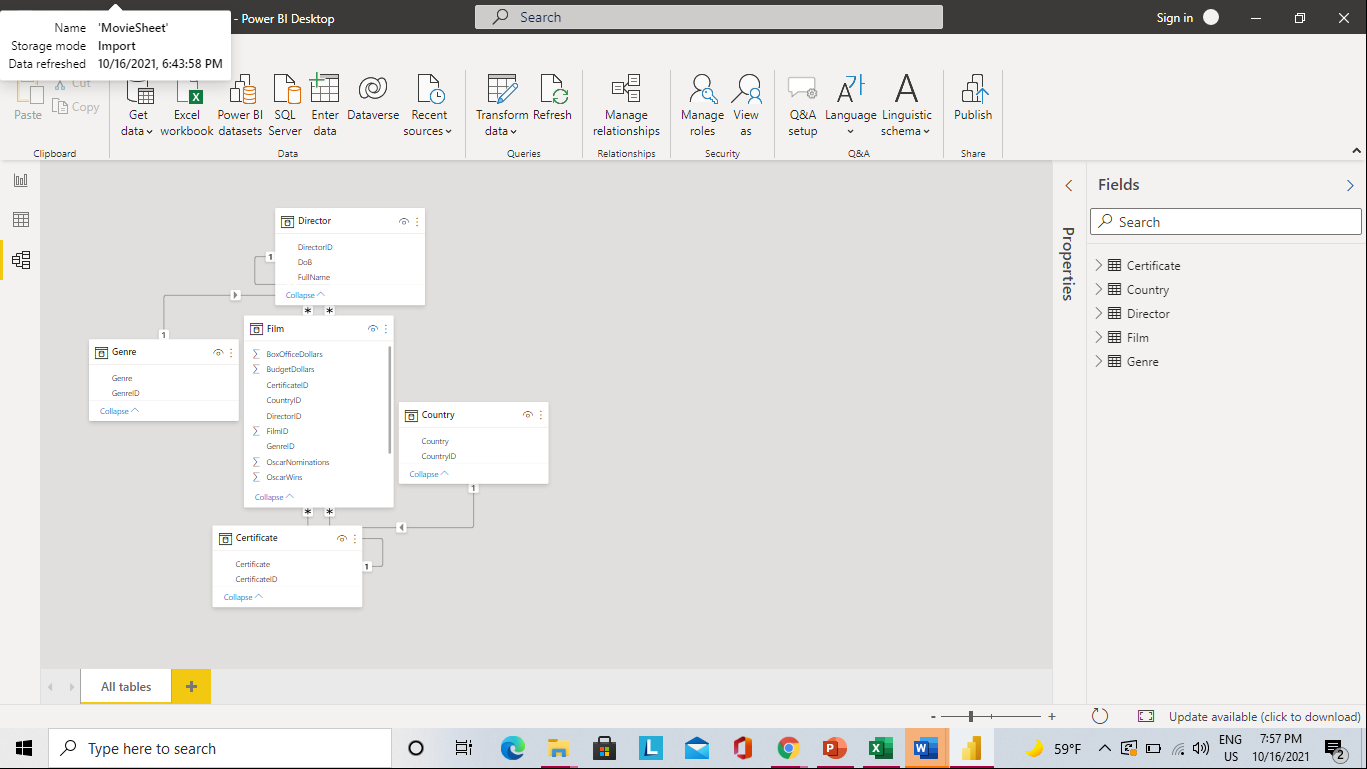
A table was then created from the cells in the range and renamed



All data gets loaded with the MovieRange, MovieSheet and MovieTable displayed in the model view

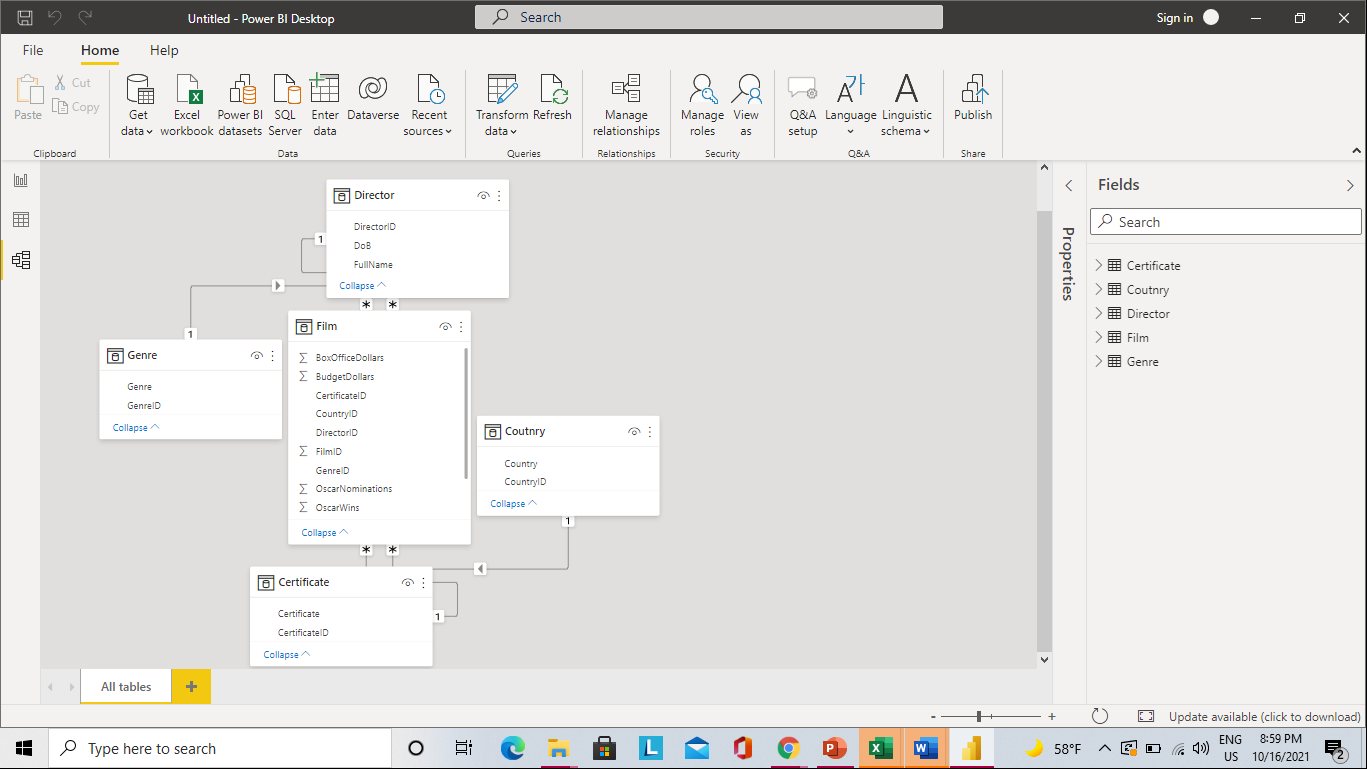


The tables were deleted from Model view and replaced with Movie Related Sheets and their relationships are defined automatically

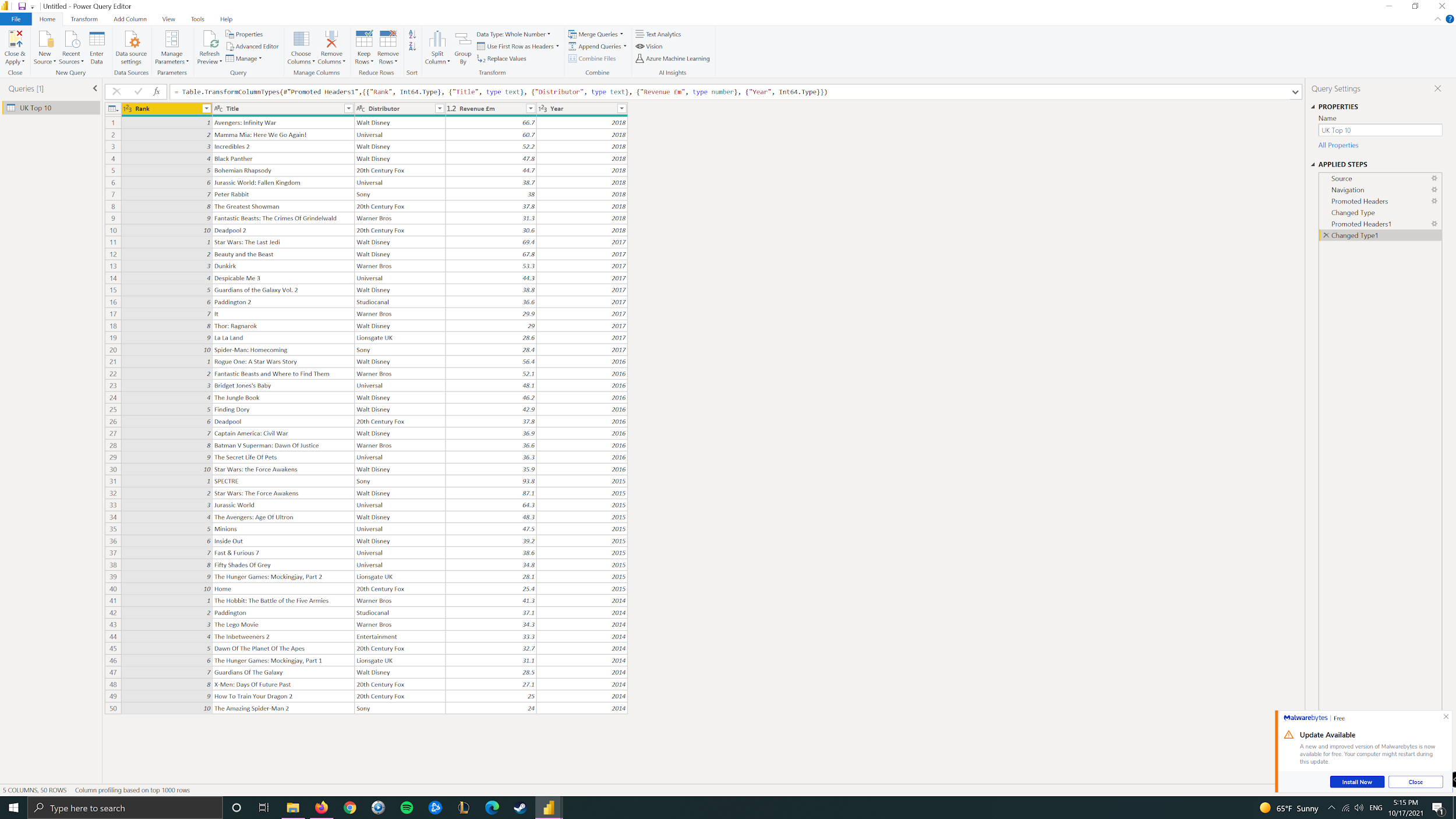


Autodetect relationships then removed and relationships added back manually resulting in the same model view.

Next, the third table contained undivided tables in Excel. I went in excel and created separate tables before loading them into model view

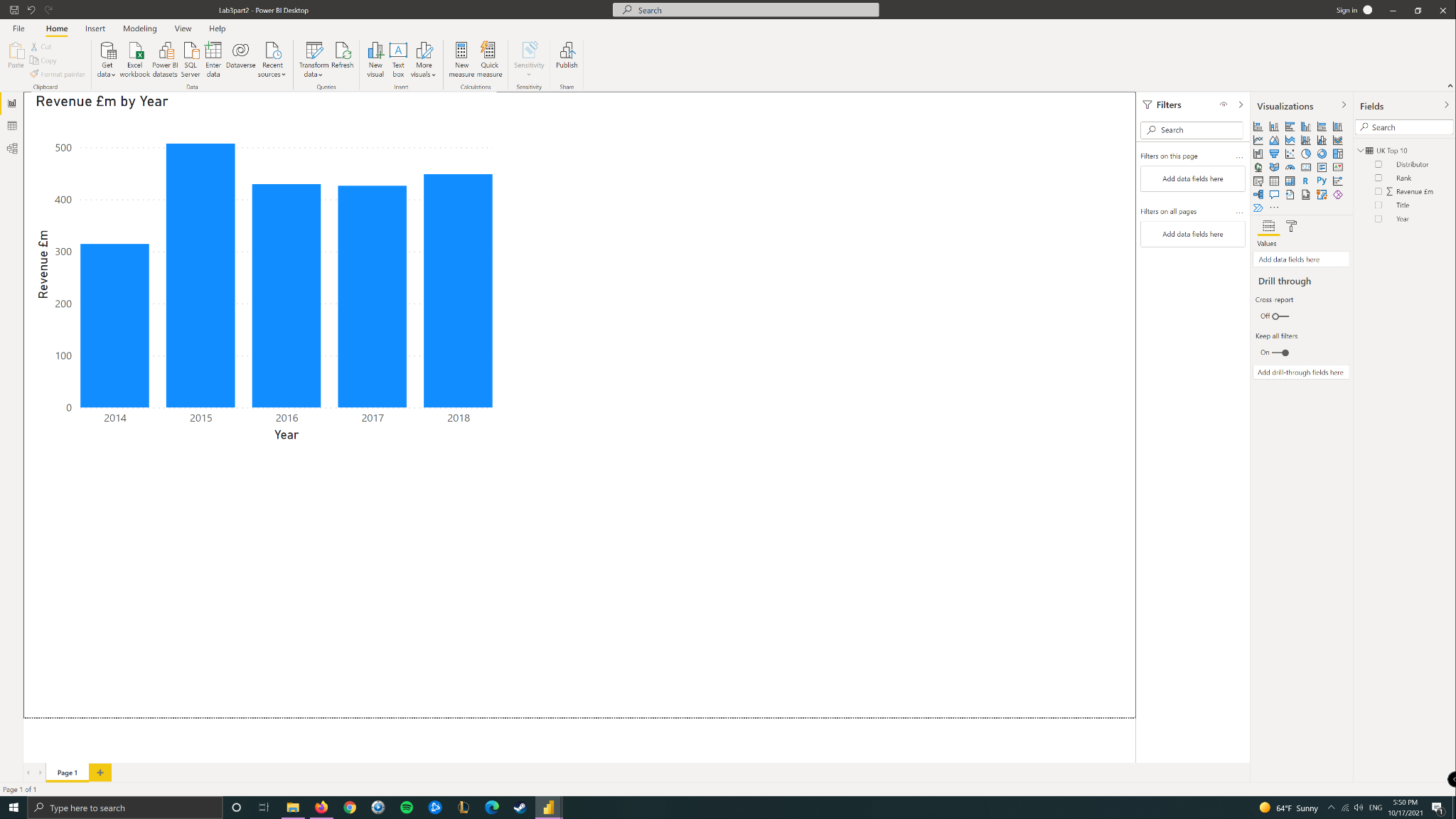


Data transformation in table 4 done with the “use first row as header” key as well as “delete first row”





Summarization turned off so “year” is displayed correctly

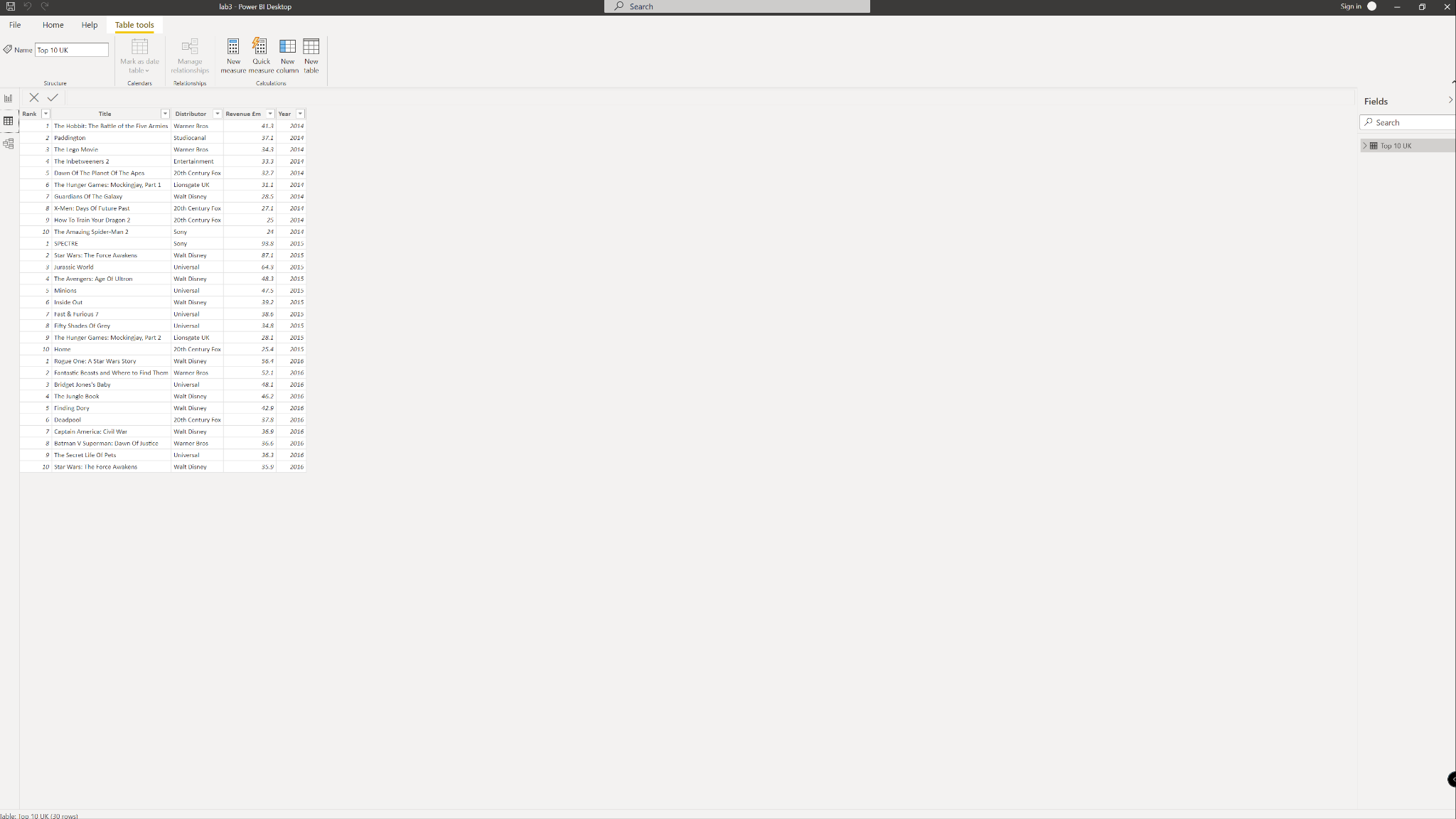


(Further with no screenshots until 3.2)

* Removed “alternate rows” so that rows in regular intervals are removed.
* Merging tables with “append tables as new” & cleaning table. Other tables are removed from view but are not deleted due to dependency reasons.
* Merging queries- all foreign keys matching fields in other tables merged with master table

3.2

Three text files were loaded with no transformation needed. The tables were merged into 1 and the preexisting tables were removed from model



The files were then loaded from the folder. The year was extracted from the first column using delimiters.

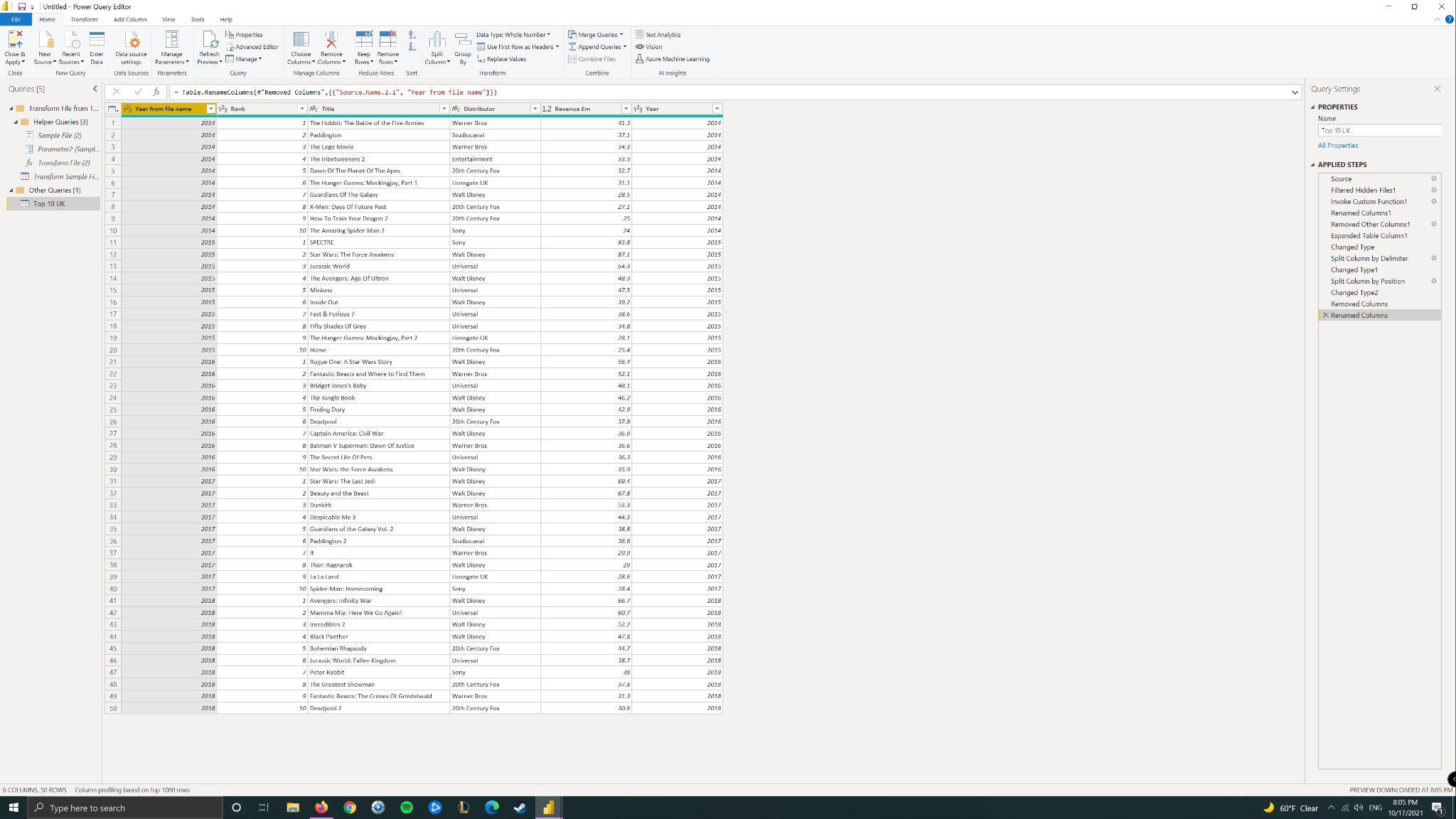
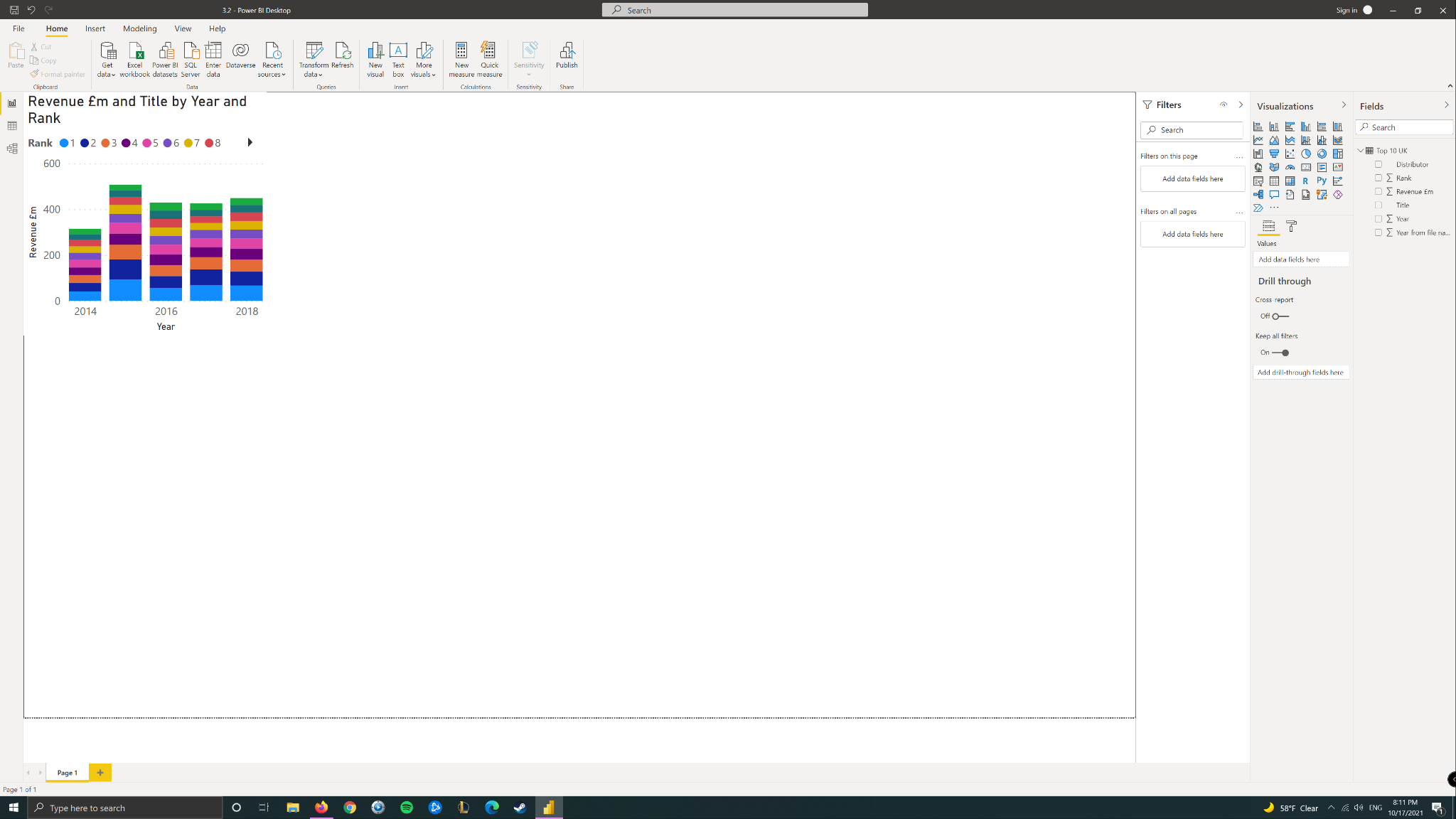
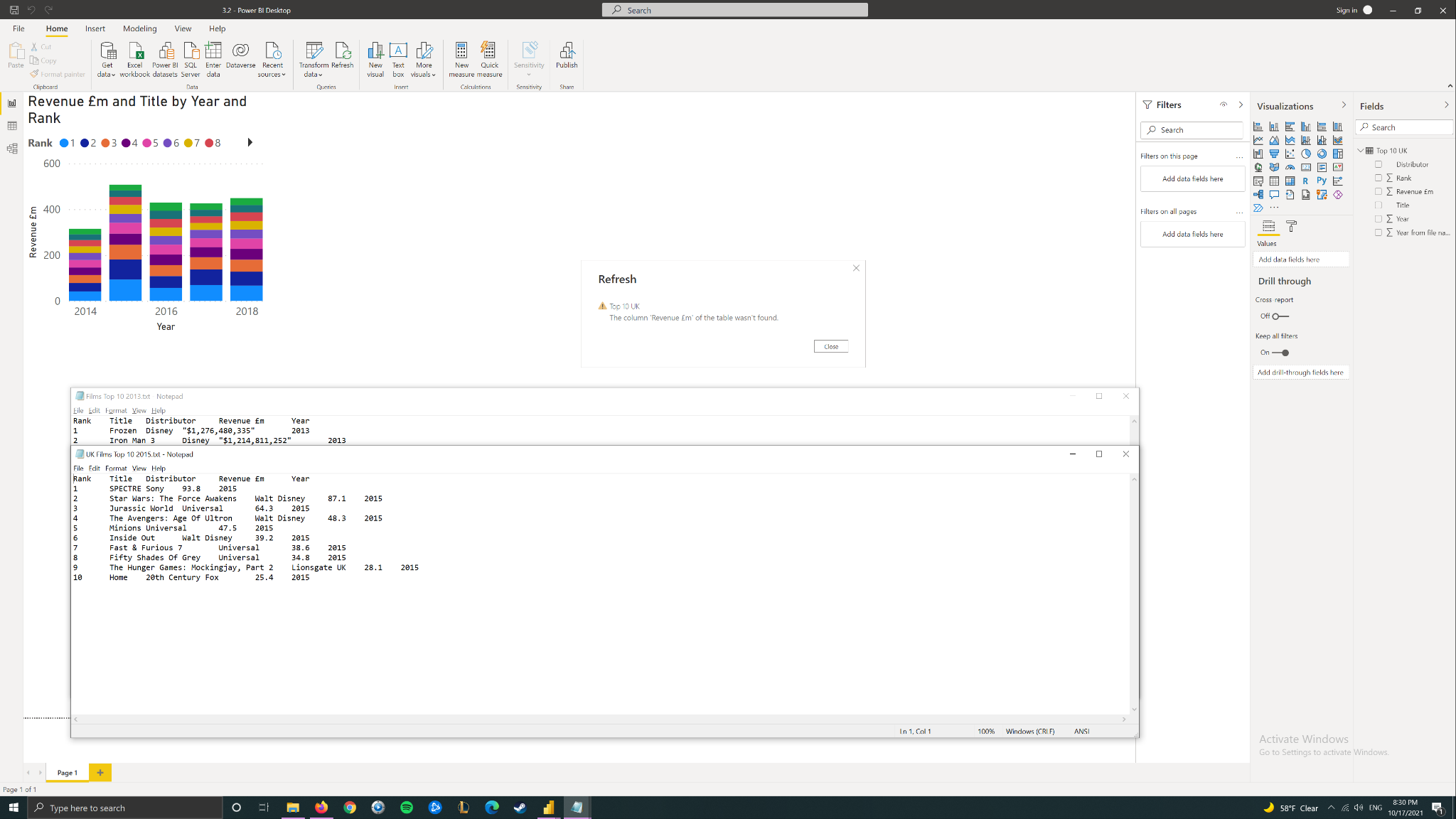


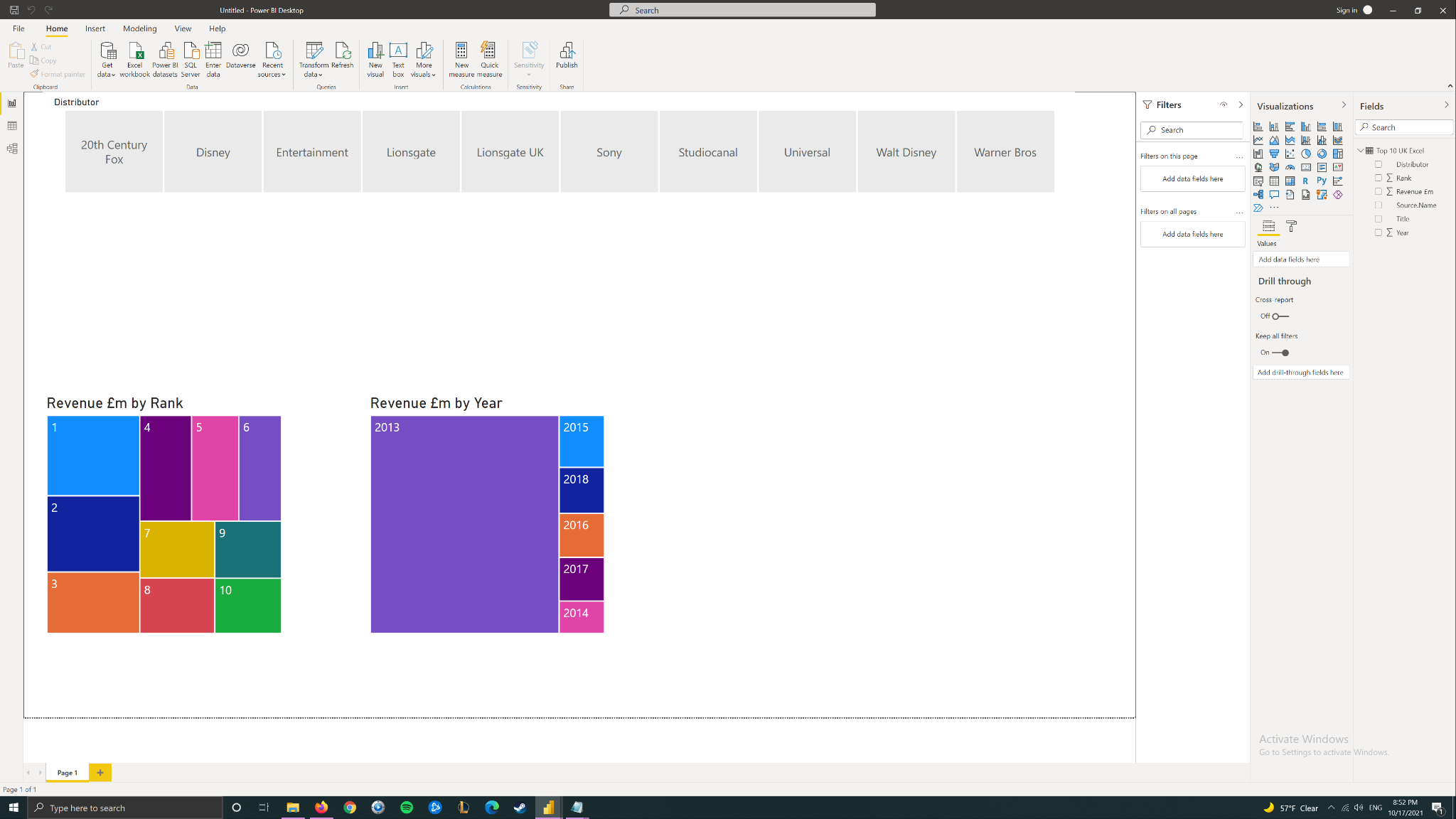
Chart was created from the multiple fields

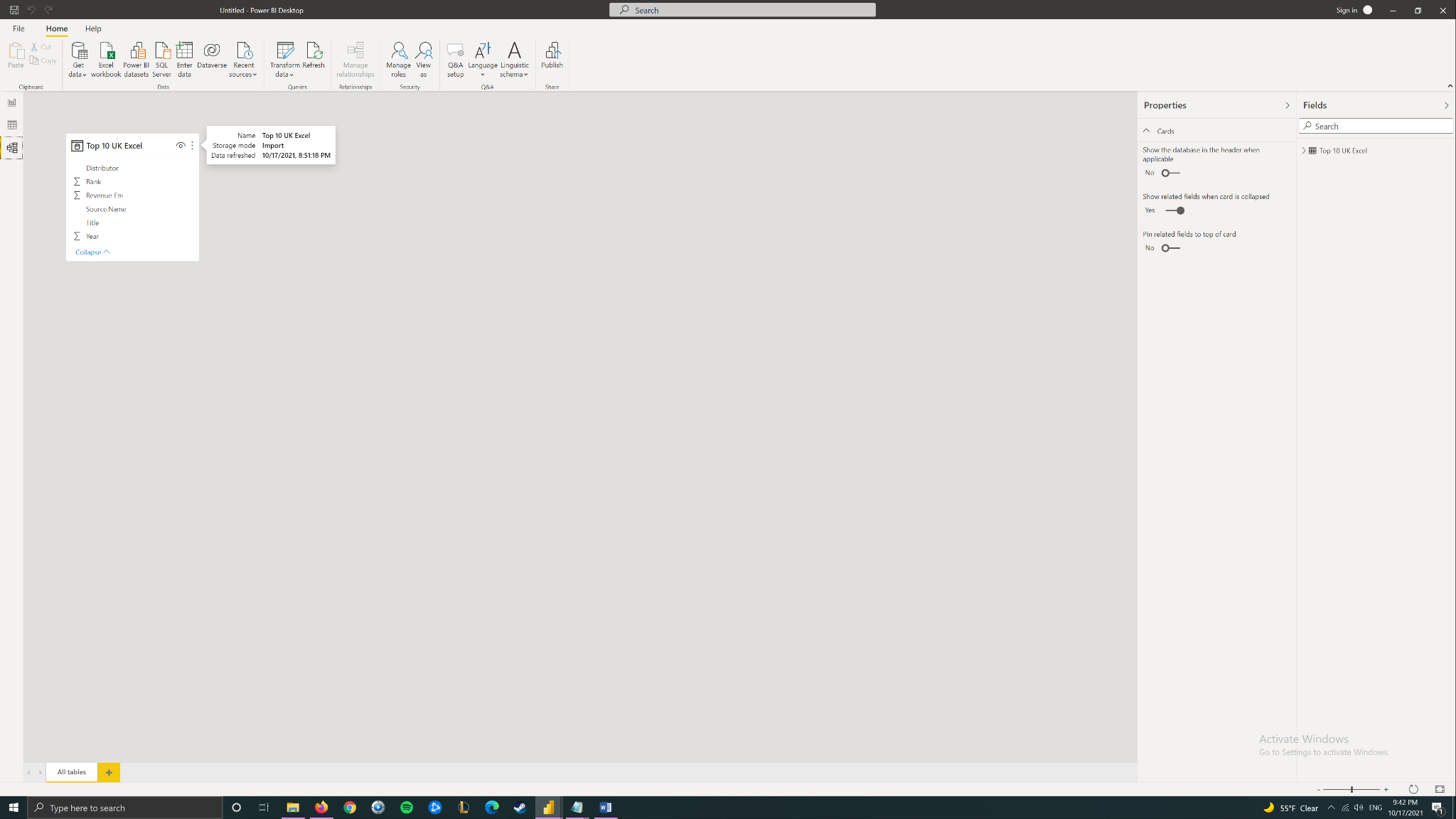


When the 2013 text file was imported, I changed the headings but still got the error

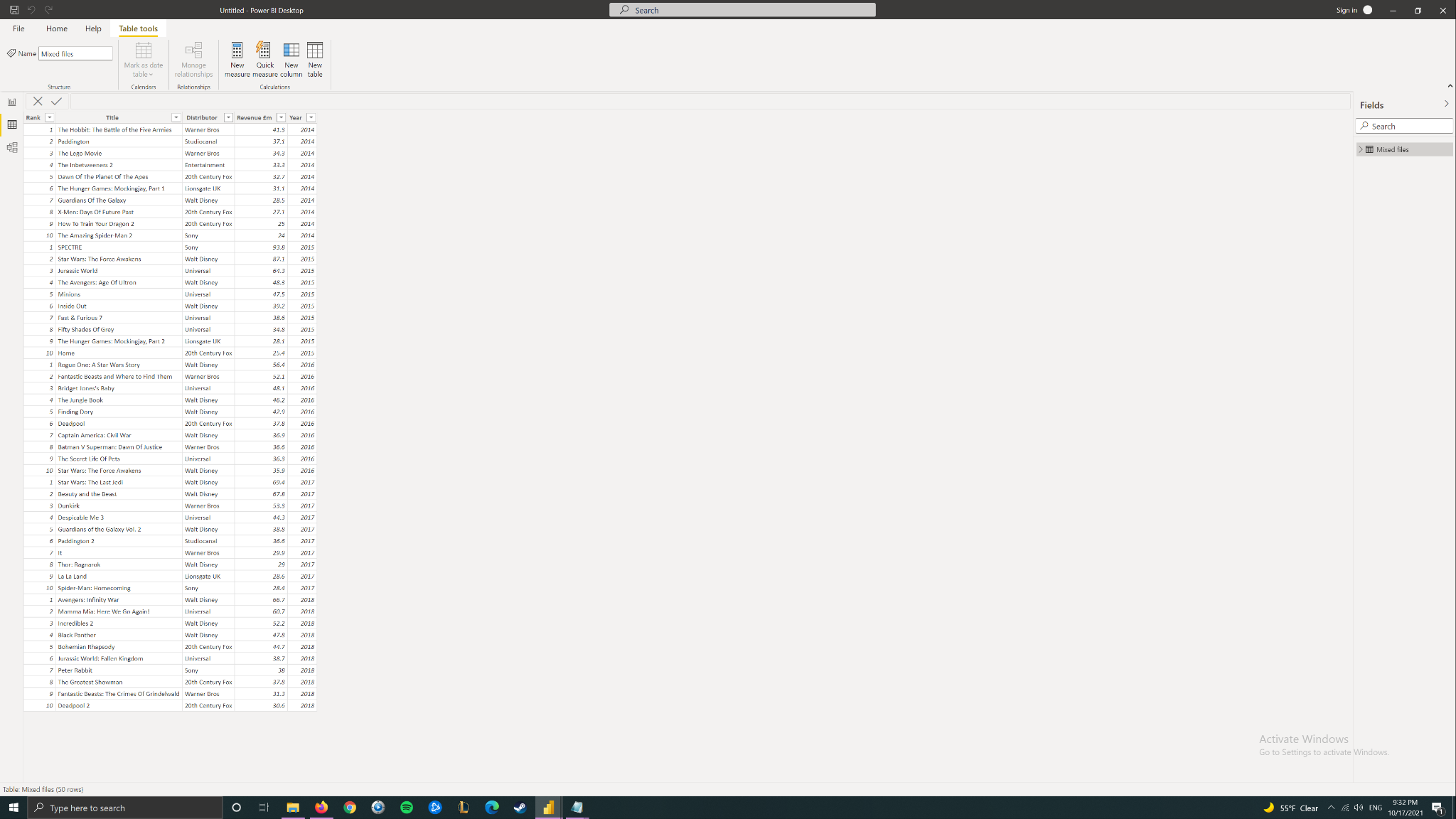


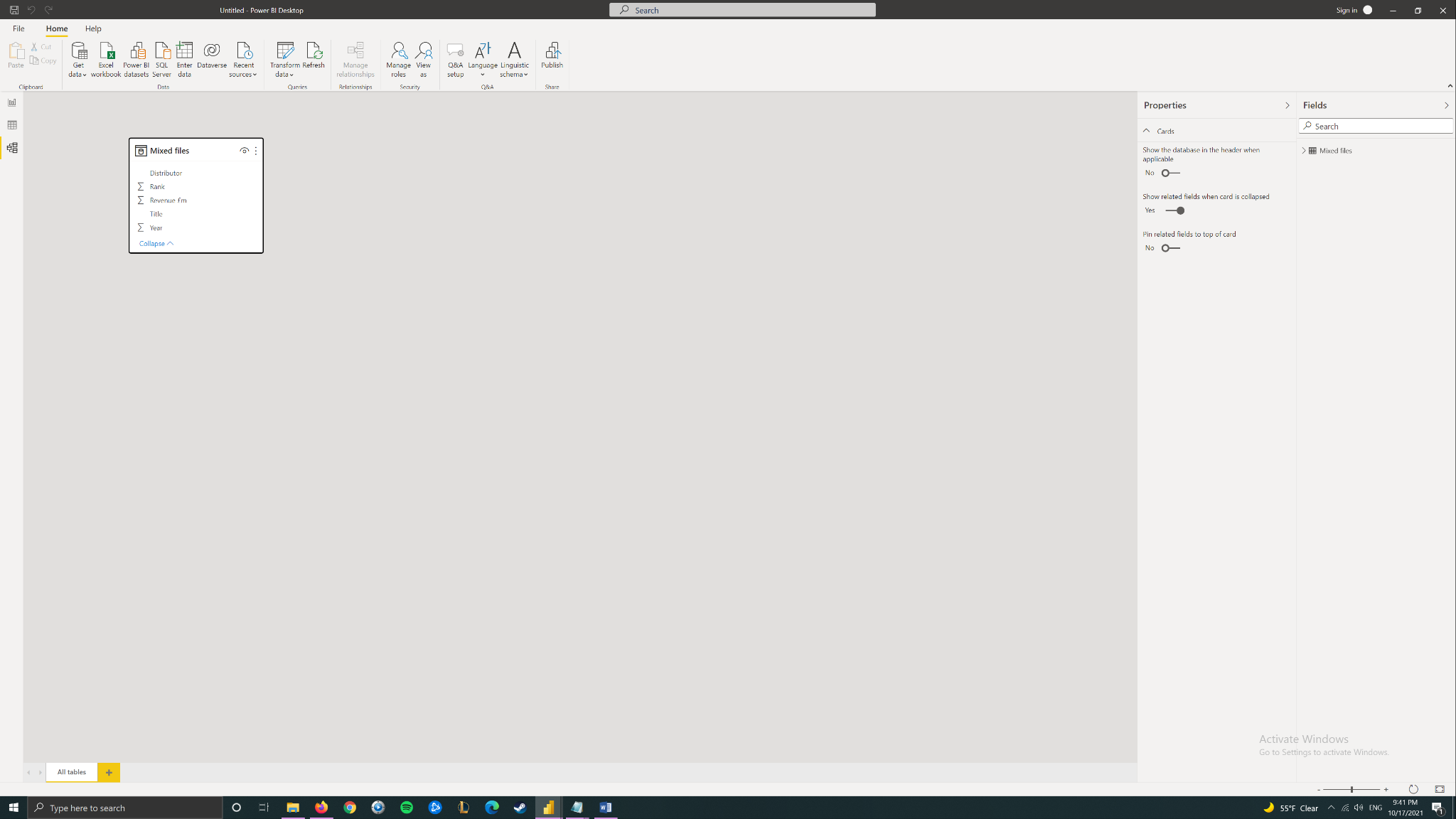
Visualization created from imported Excel file folder with two treemaps and a slicer. 2013 successfully added after changing worksheet name





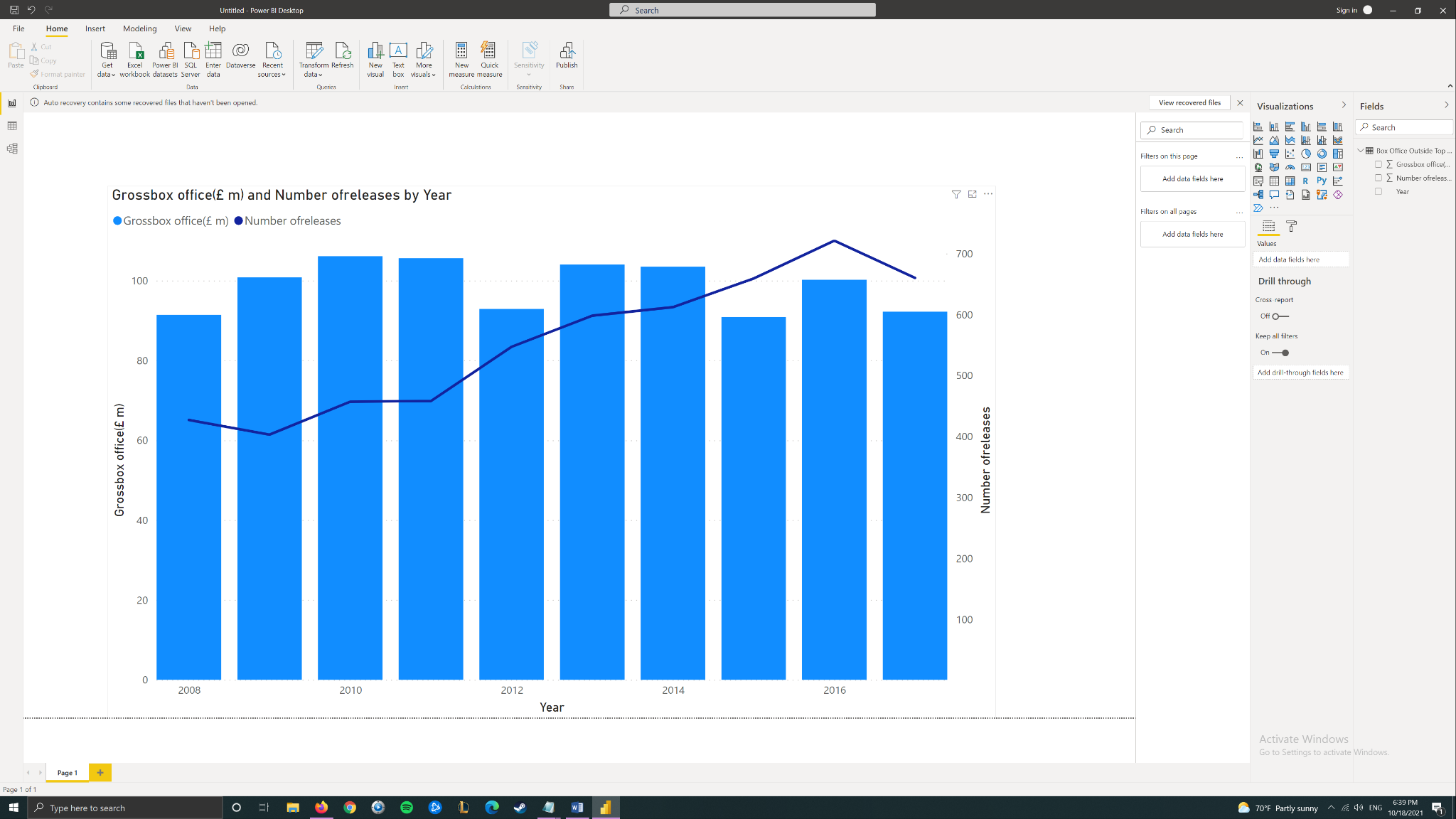
After importing a folder with mixed format files, query editor was used to filter out the excel files





3.3

A table was imported from a PDF file and data was cleaned in PowerQuery. A clustered bar and column chart was created.



(Further without screenshots)

- Another table loaded and data cleaned (removing rows, using top row as header, replacing values)

- Data from the second table was added to the report view, with one-to-one relationship detected between both “year” fields.

-Again we practiced disabling “automatic relationship detection”

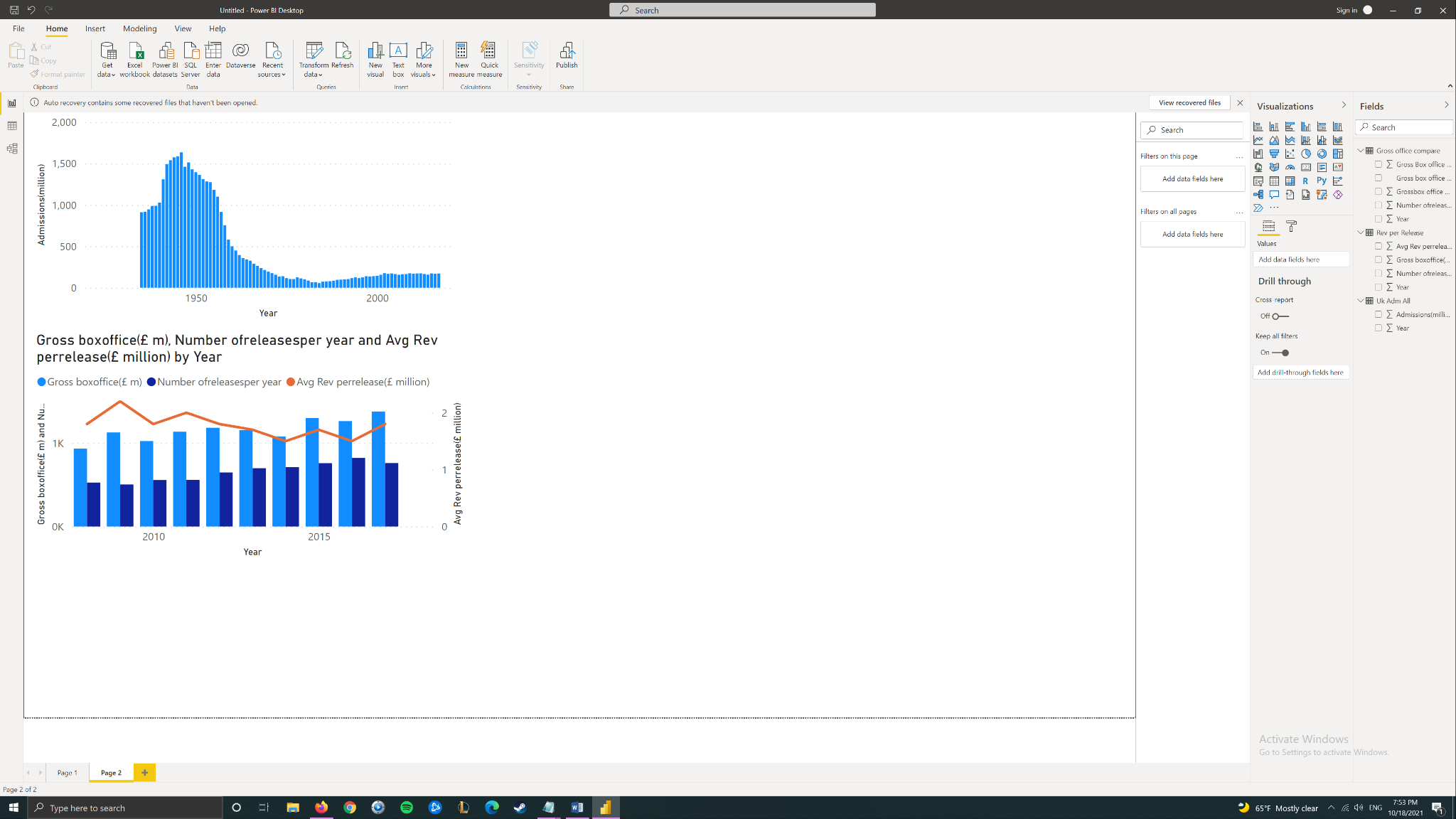
- A merged table was created based on “year”, the previous removed from model.

- Calculated column using the two other columns was created.

- New table created with new column.

- New 3 tables loaded & combined into one. New chart created from that table.

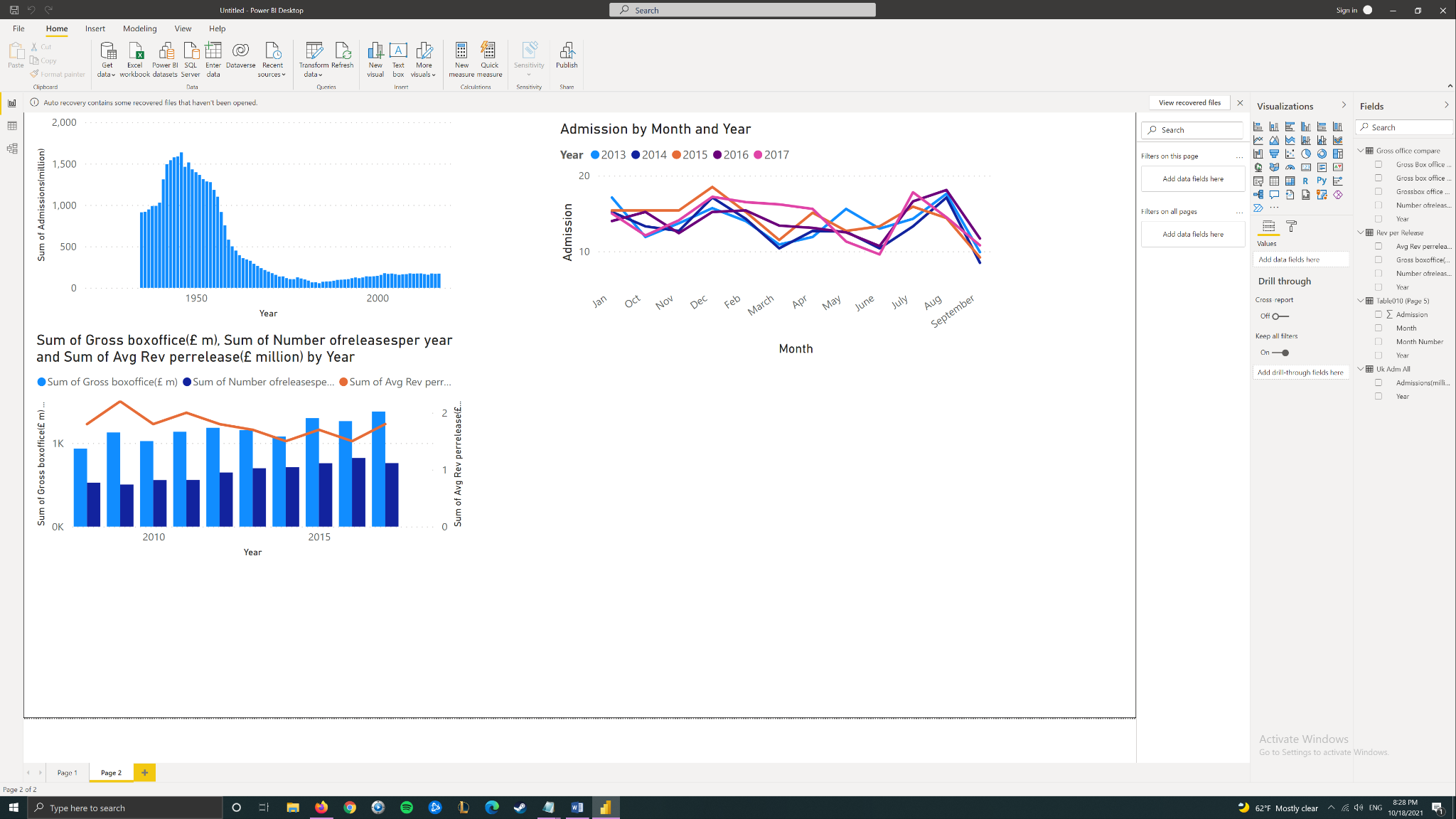
- Transposed table created and added to model view



- Imported new table and transformed “use headers as first row”

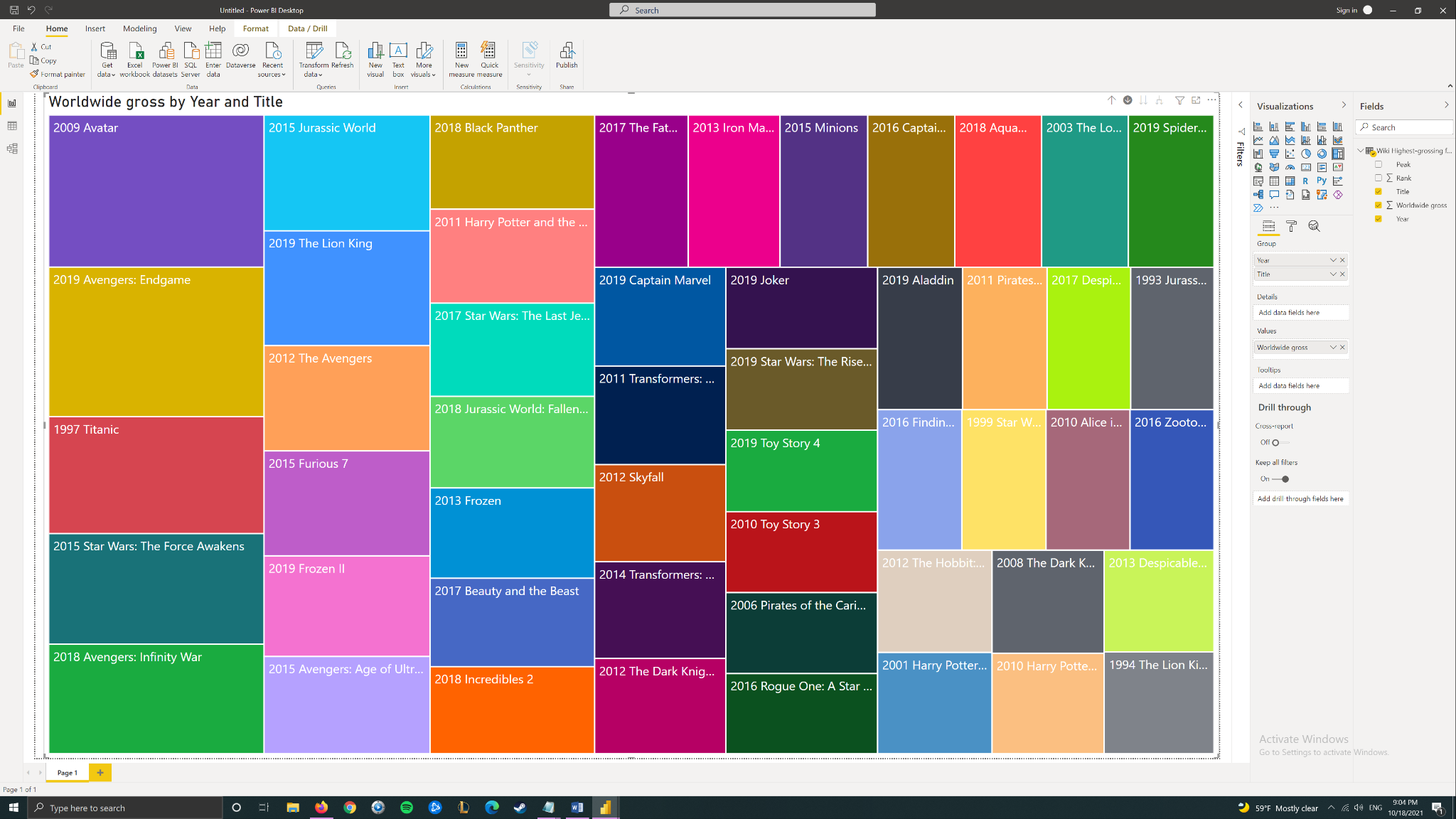
- Unpivoting the table, and followed the steps to create a column in DAX that guesses month numbers.

- Sorting by Month Number added to month column but for whatever reason it did not change the viewport order as it did in the tutorial

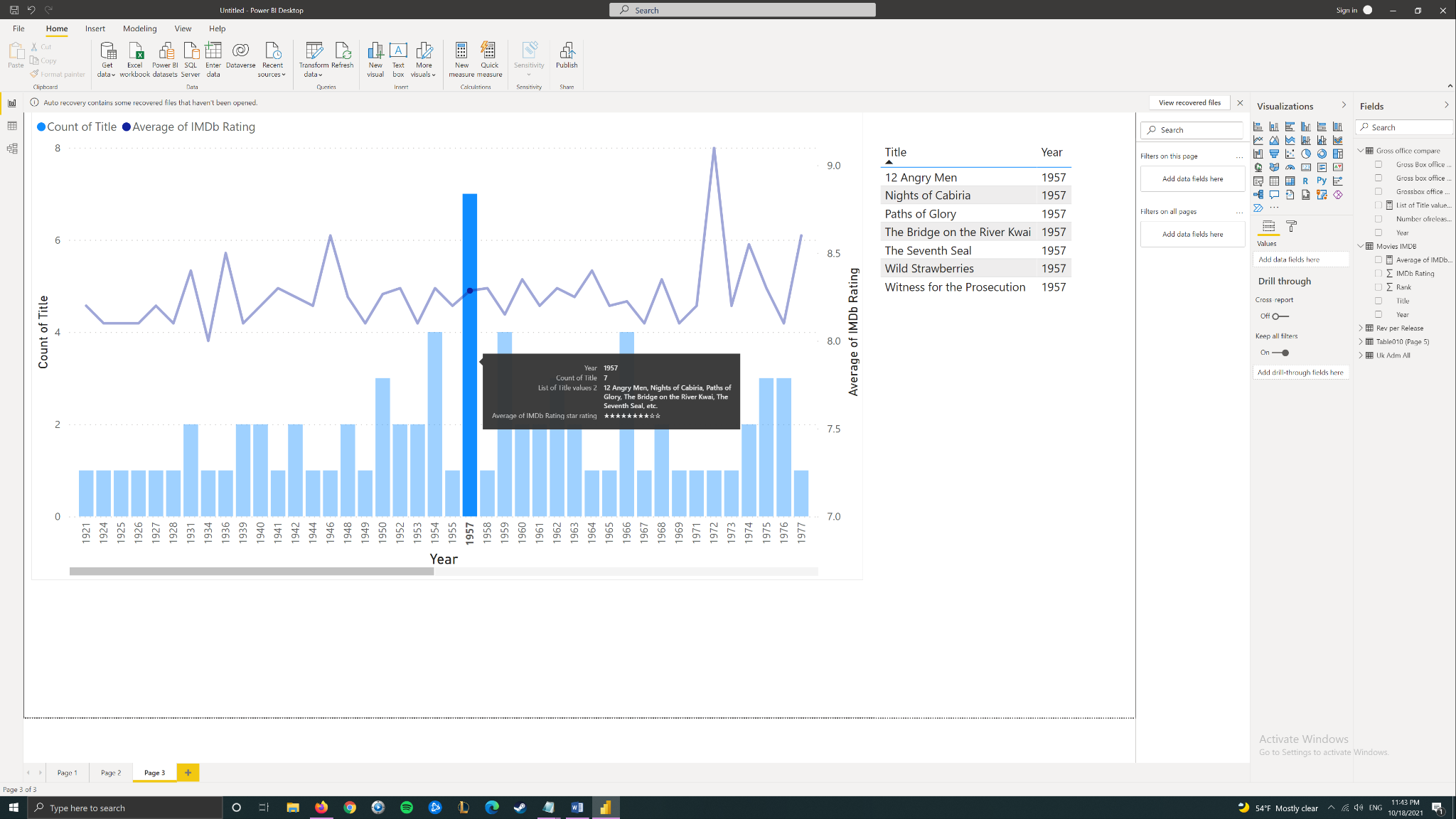


3.4

- Data received from Wikipedia and transformed. Summarization removed. Treemap was created with different Drill Down levels



A table was then clustered column line chart was created with two calculated measures to show rating and titles



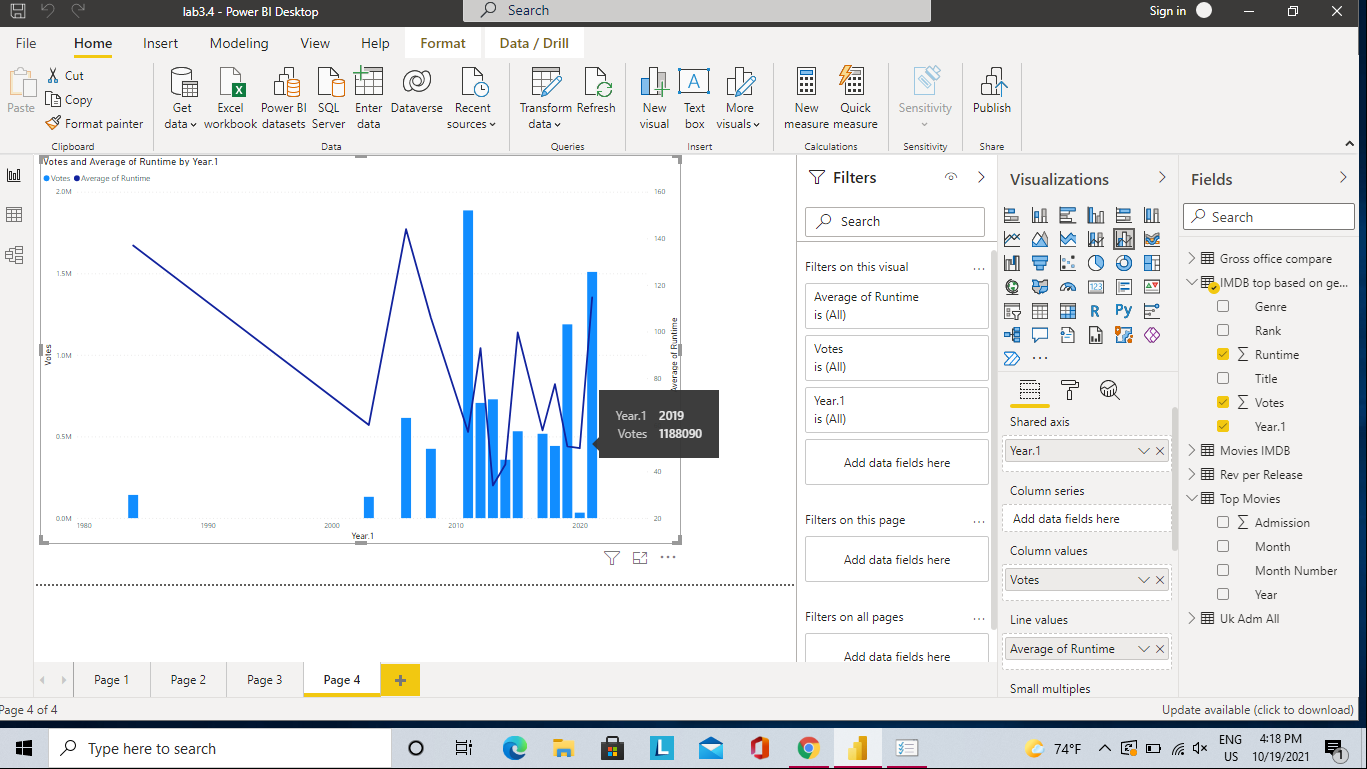
More data was loaded from the web, but this time no tables were detected and I had to “load from example.” All data needed was found on the page and put in a table. After then some transformation including splitting columns, removing values using delimiters was done. Finally, the table was loaded into a stacked column chart.

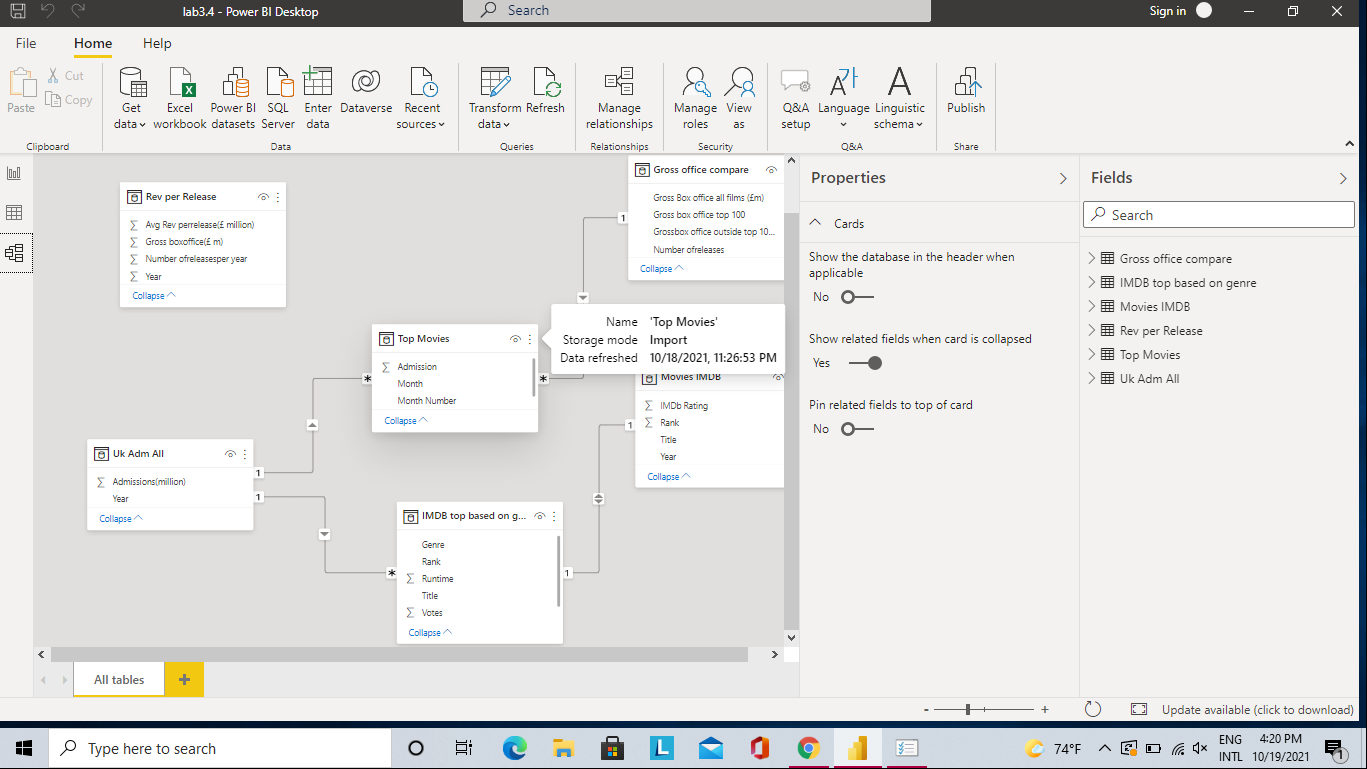
A parameter was created to let us switch between genres.

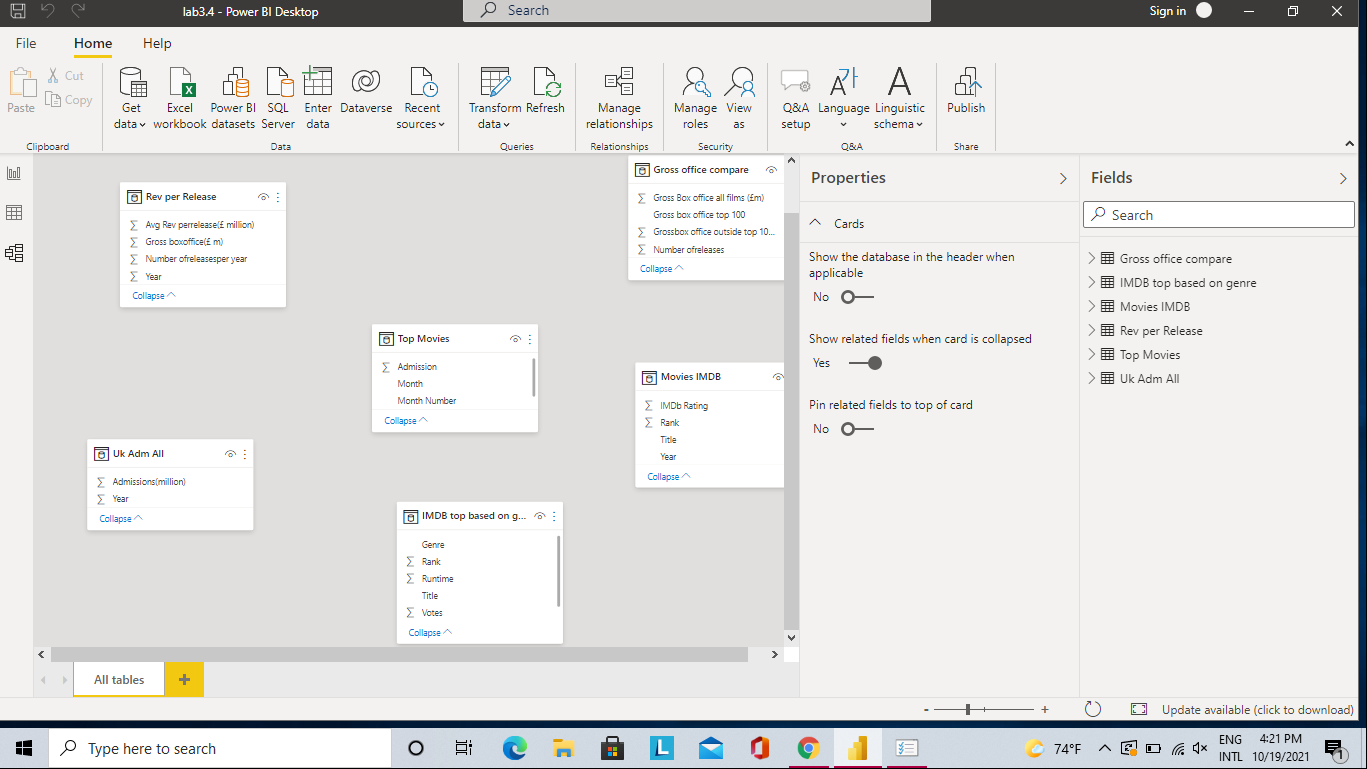
Based on parameter I can view Crime



Or action etc.



The model view here autodetects relationships. Since most /all tables come from different sources relationships can be probably removed as they’re not related. 



6.

Fact Table created with the following steps:

-Removed unnecessary columns

-Renamed tables and renamed matching column names the same across tables

* CourseCode table extracted using delimeters in FactRegistration

-Created index in FactRegistration

-Merge inner-join between Dim tables and FactRegistration across InstructorID, SectionCode, SemesterCode and CourseCode

- Split Original, Actual, and Cap registration in FactRegistration using delimeters based on “/”

- Reordered columns (may not be reflected in Data view - does not match PowerQuery order, it’s a feature)

