# Data Cleaning:

1. No duplicate records found. Checked with “remove duplicates.
2. Removed white spaces in Title column using “=Trim()” function.
3. Found 30 rows what doesn’t had English letter using filter tool. Removing them changed the Number of rows from 15480 to 15450.
4. Found 1728 rows has blanks value in ‘Genre’. Removing them changed the Number of rows from 15450 to 13752
5. Found 36 rows has blanks value in ‘Tags’. Removing them changed the Number of reows form 13752 to 13716
6. Separated the Genre to individual columns using Text to Columns by setting the delimiter to ‘,’.
7. Deleting the Previous Genre column and having the top two genre as Genre1, Genre2.
8. Separated the Tags column using Text to Columns by setting the delimiter to ‘,’.
9. Separated the Production House to individual columns using Text to Columns by setting the delimiter to ‘,’.
10. Deleting the Previous Production column and having the top two genre as Production house1, production house 2.
11. Separated the Country Availability column using Text to Columns by setting the delimiter to ‘,’.
12. Deleting the previous Country availability and having the first county mentioned.
13. Found 281 blank rows in Language Column, Filling the blanks column with respective of TV Tag, and their country available. For example a Movie in Japan has the language as Japanese.
14. If language not found in Tag, setting their language according to their county
15. Column Name ‘Box office’ has mismatched budget with different data types. To deal with this, coping the column and converting it to number without the ‘$’ using =right(y2,len(y2)-1) and removing the comma(,) using data formatting
16. Converting from text to number using =INT() and converting it to Currency using data formatting.