

Task 1

Aim :

To make eda and visualization on the given dataset

Things Done :

1) Libraries Imported

- Numpy for handling arrays if needed
- Pandas for handling the data frames and matrices
- Matplotlib for visualisation
- Seaborn for clear in statistical visualisation

2) Loaded the data and accessed it to gain information

- **info()** - gaining up the data information of what are the features and what are up over there
- **describe()** - gives up the summary statistics, it describes the summary stat for the numerical column

3) Found missing values

Found missing values in several fields and imputed it using

- **Arbitrary imputing** - imputes the value with given input for change

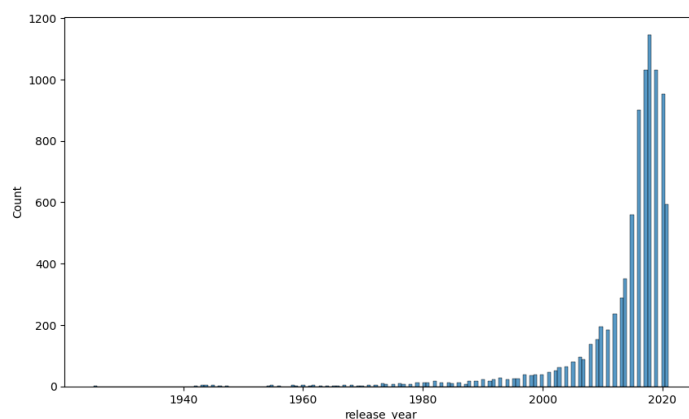
Made the format change for datetime and imputed its missing value

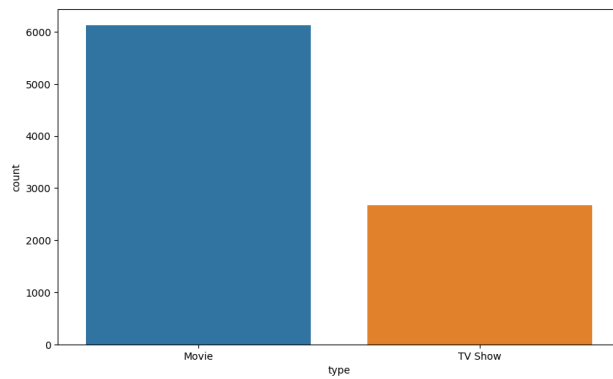
- **ffill** - imputes by filling the previous nonnull value

4) Visualization

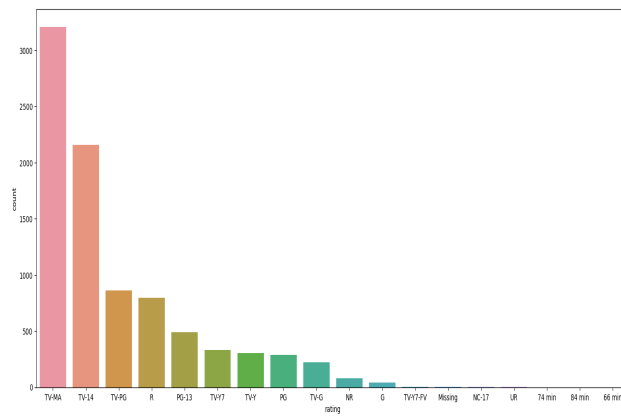
Done some visual implementation

1. Release Year visualization towards year and count

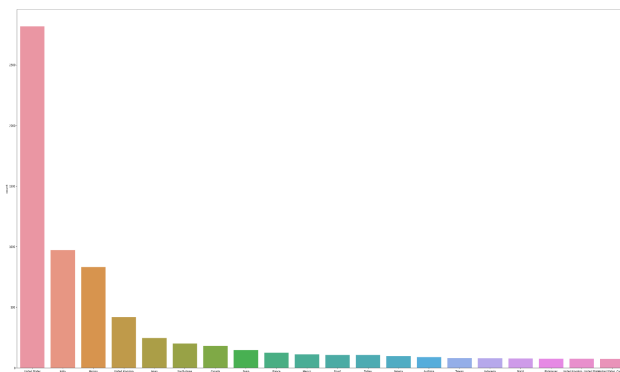




2. Comparing Content as Movies and Tv



3. Comparing Ratings



4. comparing countries and acquiring the top 20 countries

Result:

The Eda, Visualisations are Done successfully :

https://github.com/Sairam-K26/Kovai.Co_Tasks/blob/main/Task_1/Task_1.ipynb