LP3 2021 Internship

IP_LP3_Data_Science_Naschwin_Dias_4904.
The data is calculated and visualized here in a more Readable format

Problem Statement:

Analysis Tasks to be performed: Explore the datasets using visual representations (graphs or tables), also include your comments on the following:

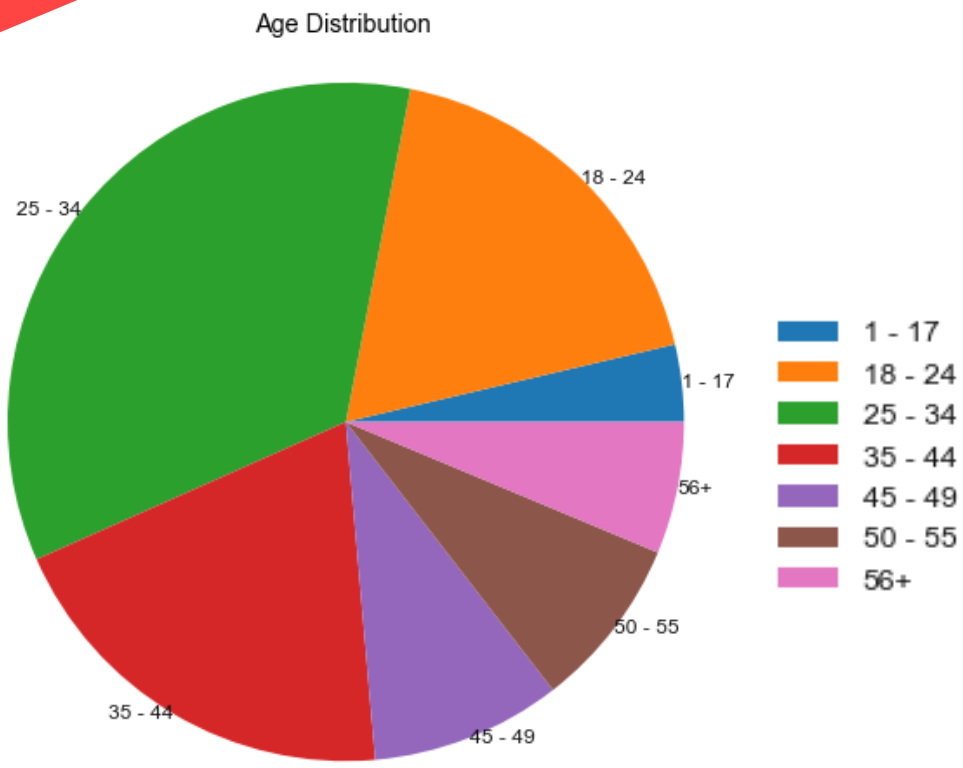
- User Age Distribution
- User rating of the movie “Toy Story”
- Top 25 movies by viewership rating
- Find the ratings for all the movies reviewed by a particular user of user-id = 2696
- There should the different graphs used for visualizing the dataset.

Use column genres:

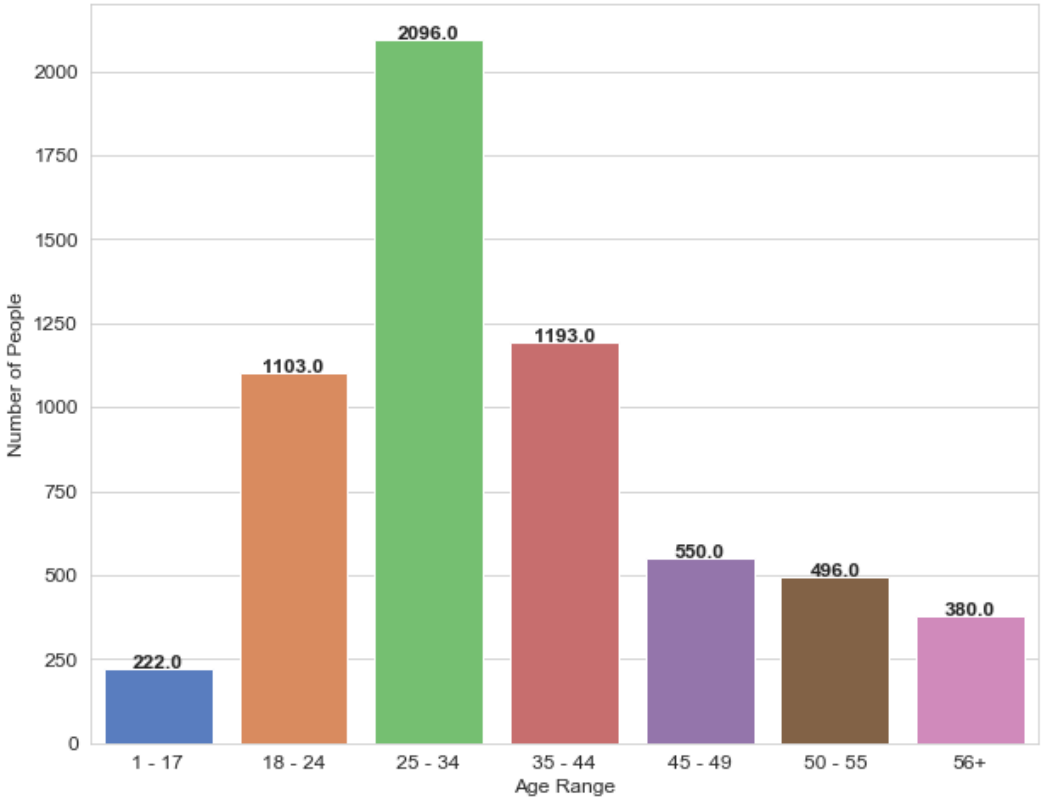
- Find out all the unique genres (Hint: split the data in column genre making a list and then process the data to find out only the unique categories of genres)
- Create a separate column for each genre category with a one-hot encoding (1 and 0) whether or not the movie belongs to that genre.
- Determine the features affecting the ratings of any particular movie

Date: 29/07/21

User Age Distribution



Pie chart - Age Distribution



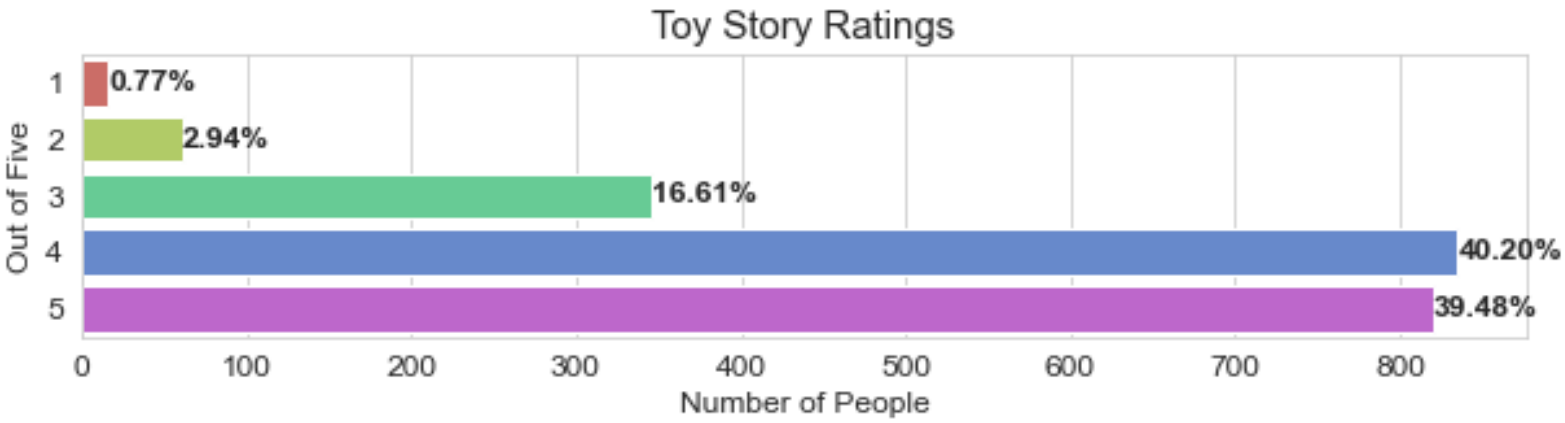
Report:

Looking at the data we can summarise most people who watch movies range from 18 to 44. With young adults being the most in number.

Creating movies catered for young adults is more profitable

Date: 29/07/21

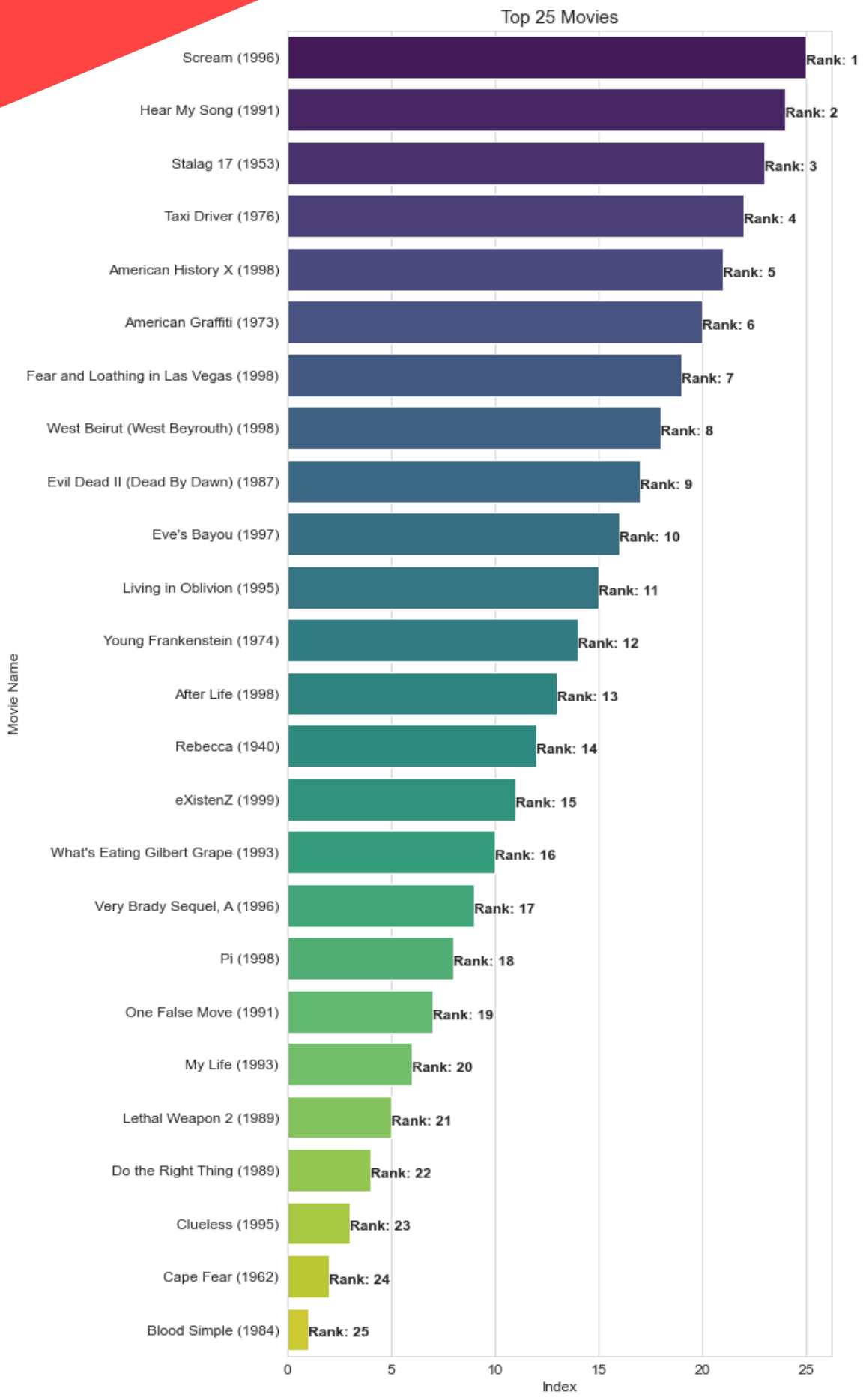
Toy Story Ratings



Report:
From all the ratings, as you can see toy story is positively liked by most people, With about 80% of the voters voting more than 3/5.

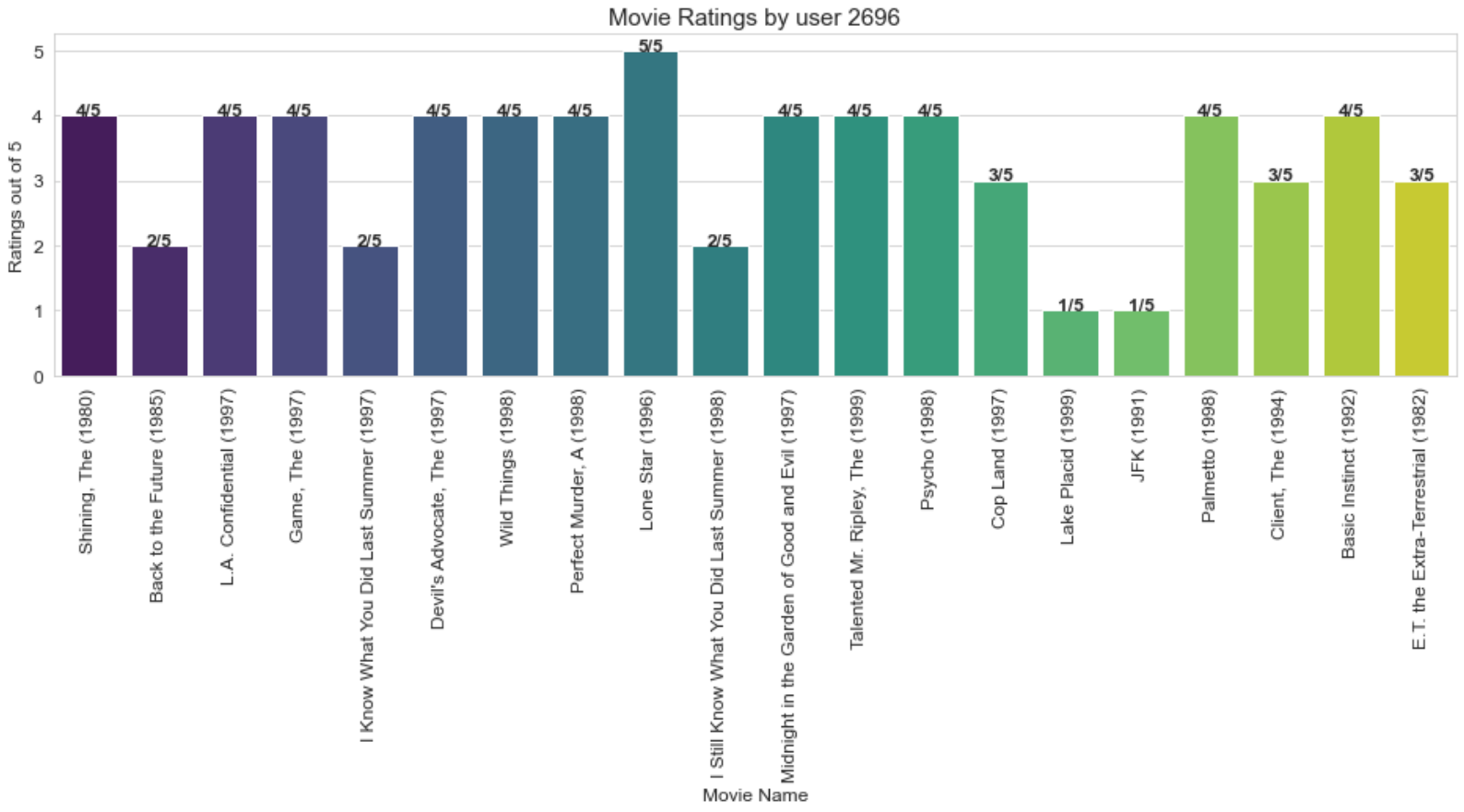
Date: 29/07/21

Top 25 Movies



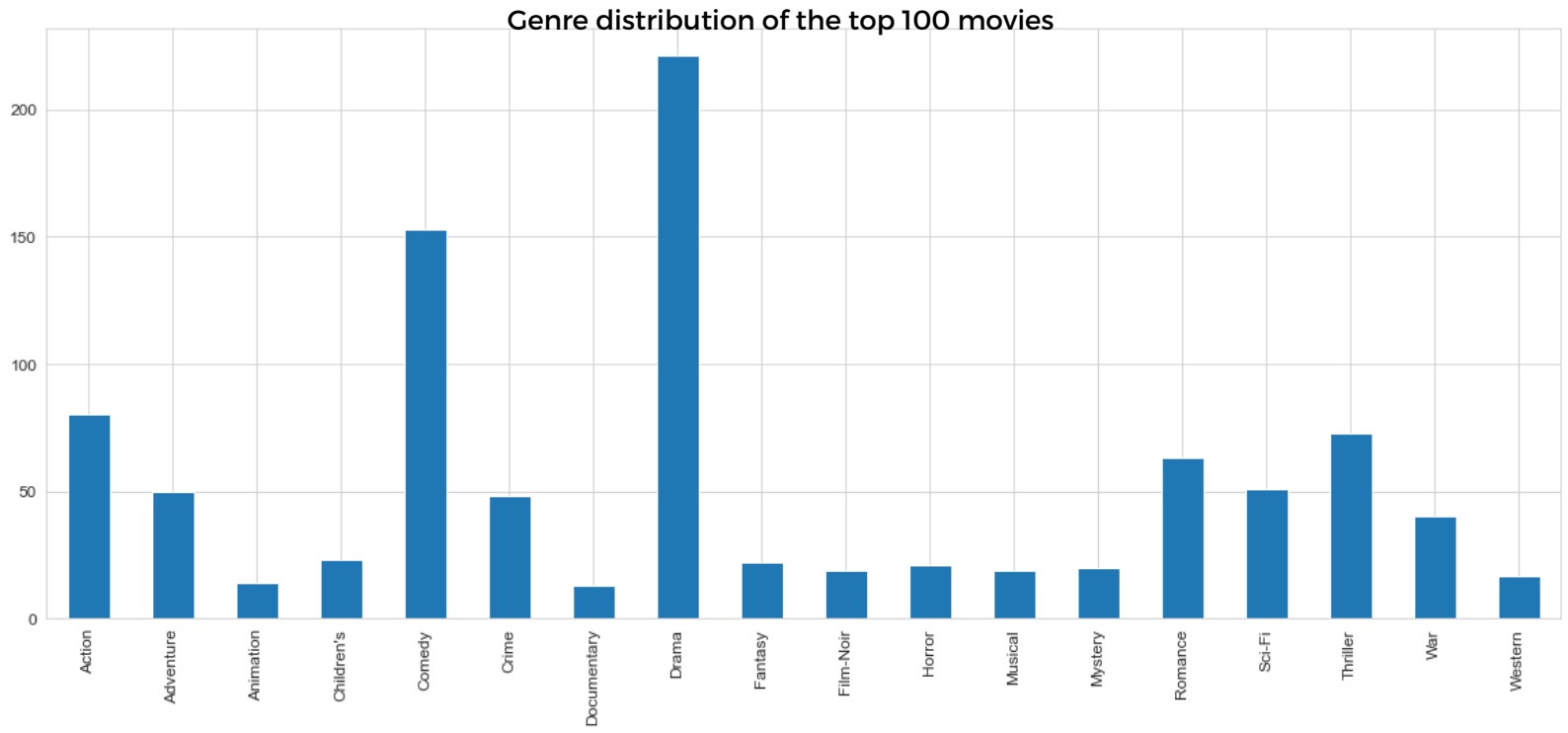
Date: 29/07/21

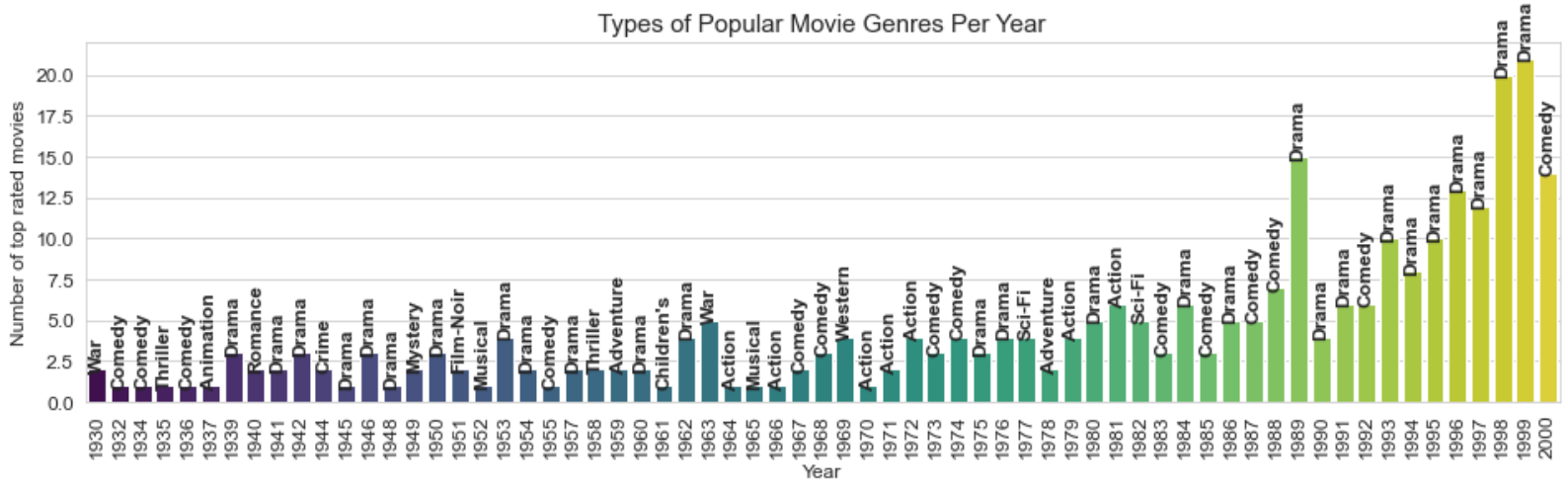
Movies rated by the user - 2696



Date: 29/07/21

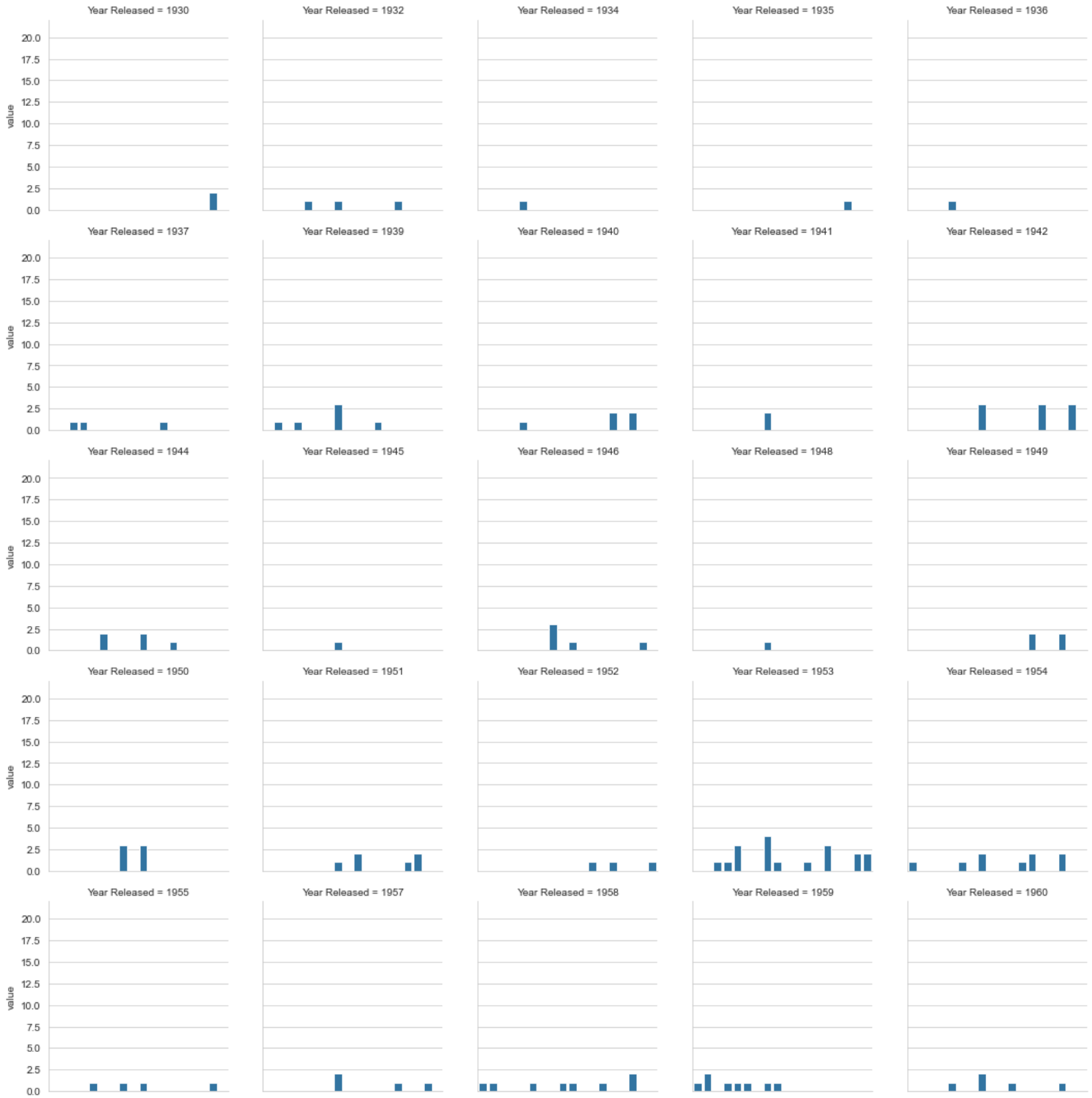
Possible features affecting movie ratings



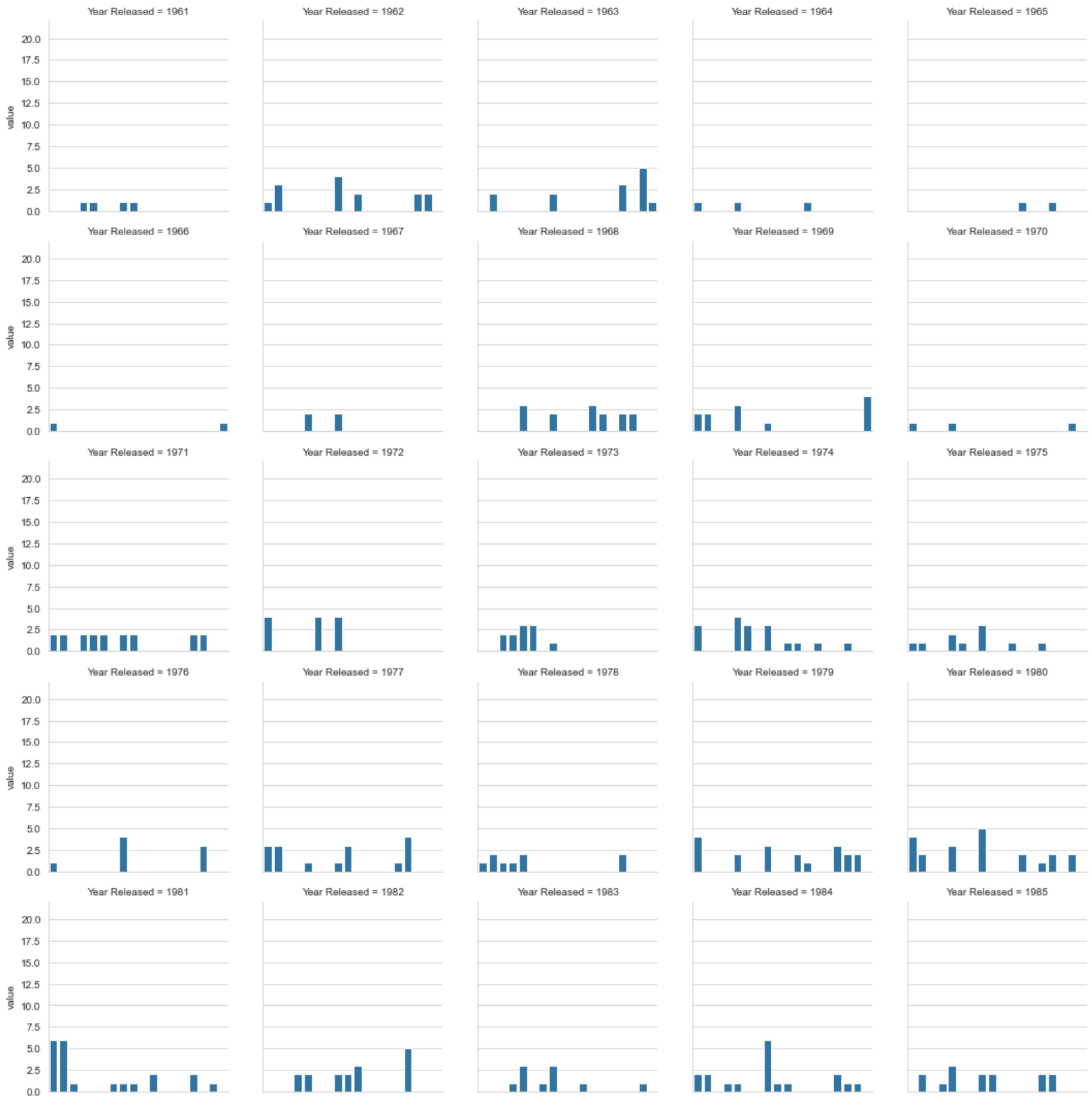


Date: 29/07/21

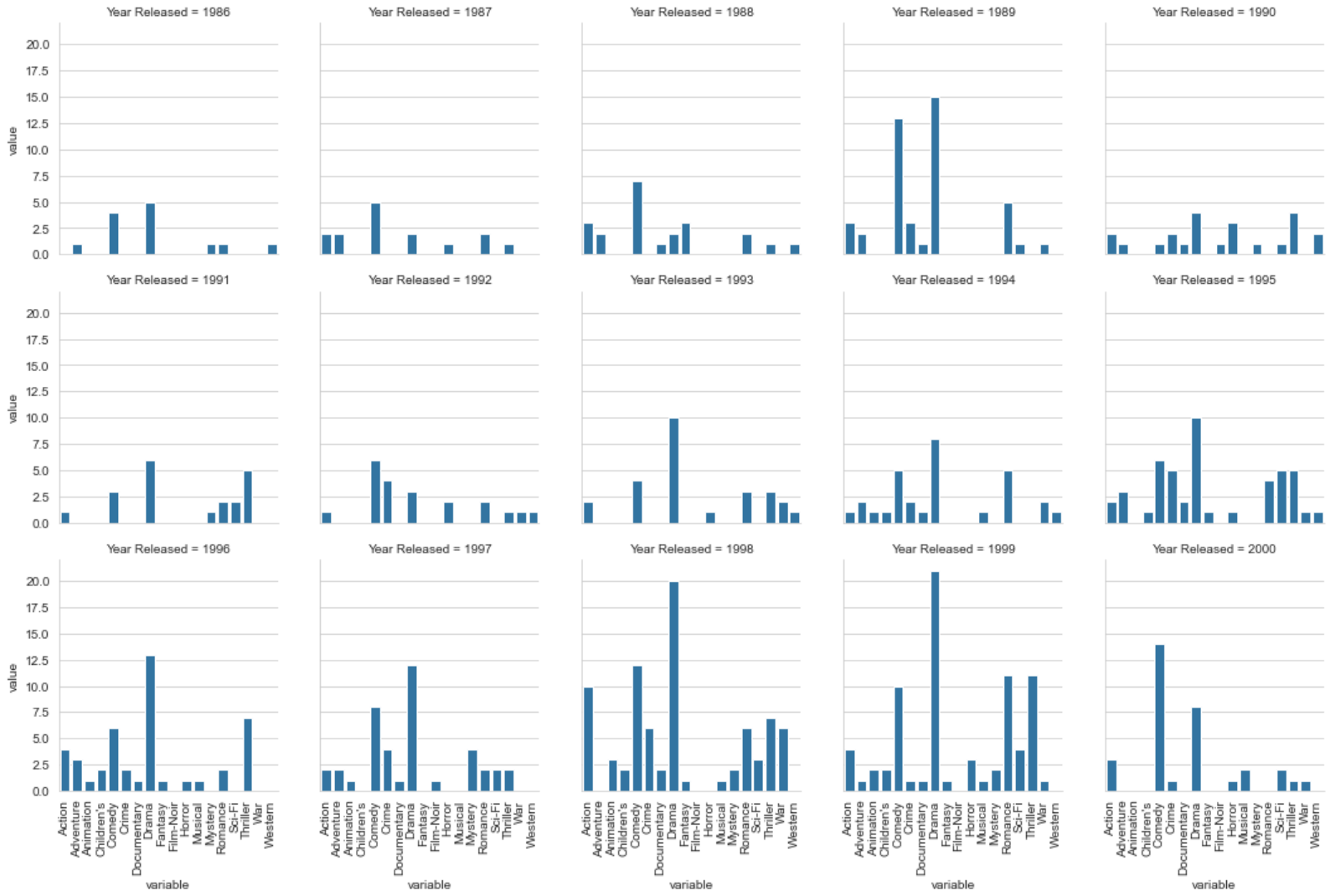
Mapping out The top 100 movies by each year from 1930 to 2000



Date: 29/07/21


Year Released = 1961
Year Released = 1962
Year Released = 1963
Year Released = 1964
Year Released = 1965
Year Released = 1966
Year Released = 1967
Year Released = 1968
Year Released = 1969
Year Released = 1970
Year Released = 1971
Year Released = 1972
Year Released = 1973
Year Released = 1974
Year Released = 1975
Year Released = 1976
Year Released = 1977
Year Released = 1978
Year Released = 1979
Year Released = 1980
Year Released = 1981
Year Released = 1982
Year Released = 1983
Year Released = 1984
Year Released = 1985
value
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

Date: 29/07/21



Conclusion:
Making Drama and/or Comedy movies does seem to work. Keep on doing that. Also, make movies targeted at young adults.

Date: 29/07/21