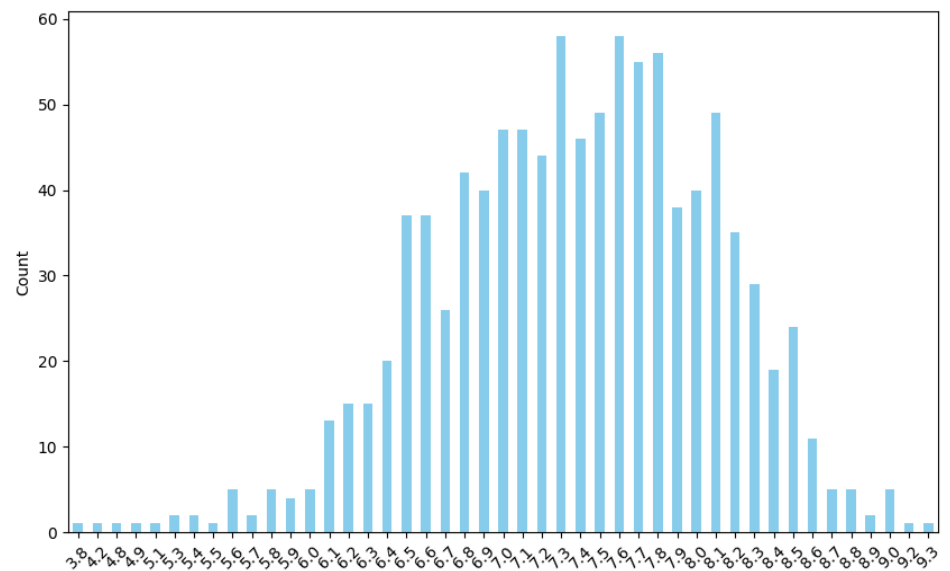
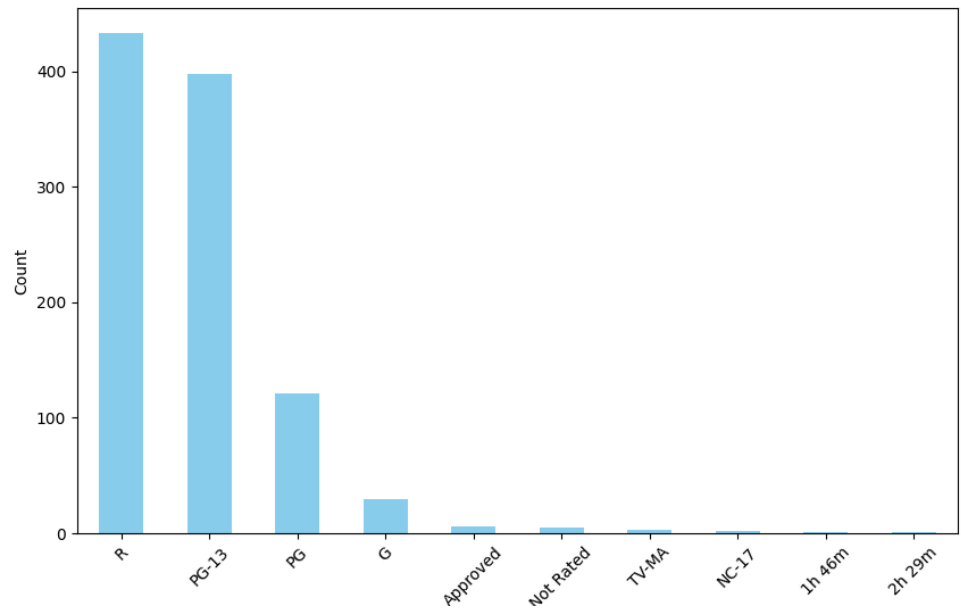
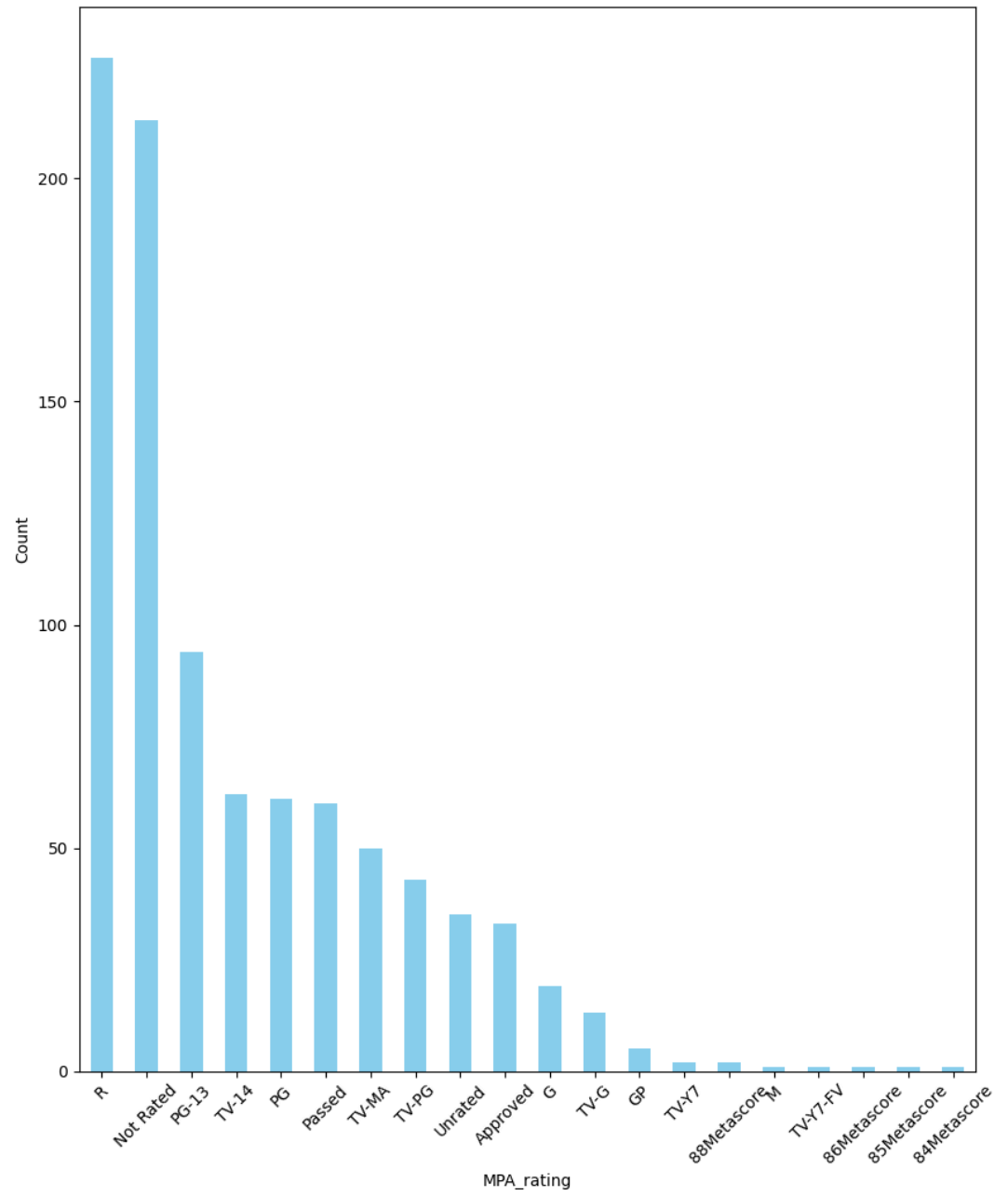


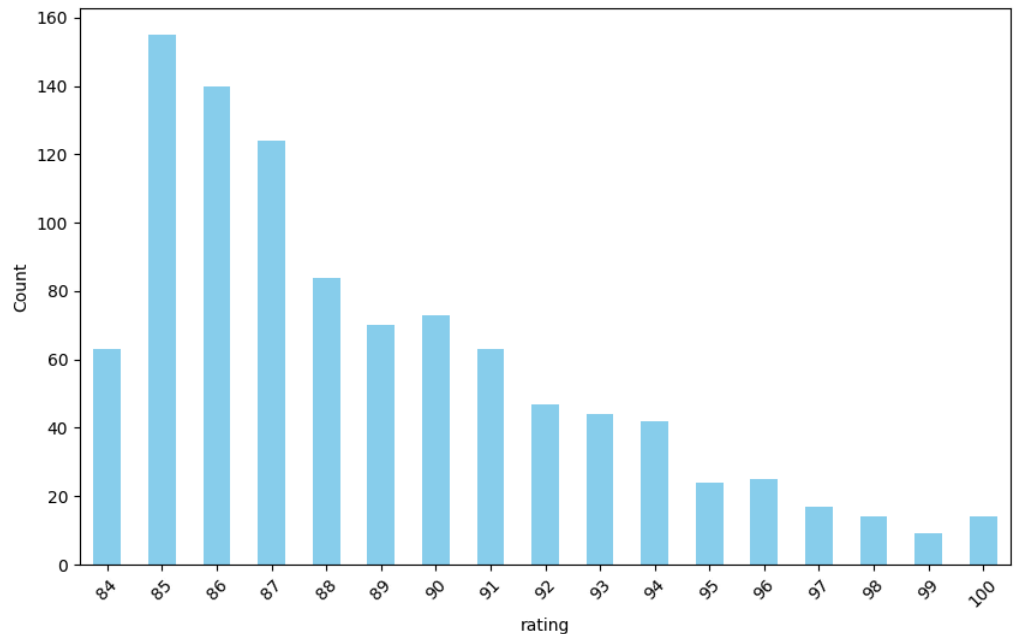
- Ratings

- On table B there was a 8% data loss the other table there was no data loss
- I will probably just assign a null value to those attributes. This is because the attribute is not necessary for identifying what that record is.
- The attribute is a categorical string.
- There are some outliers from table B from the ratings because there is a decreased number of 84 that is out of place from the rest
- Histograms for table A

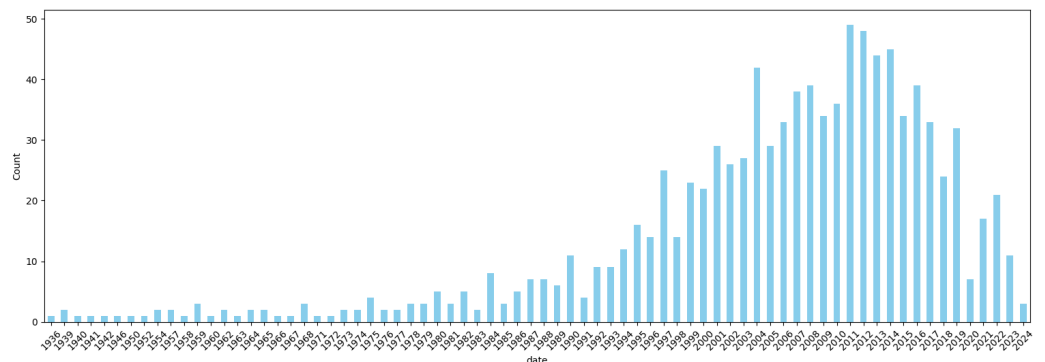


- Histograms for B





- There is no format for the two rating systems
  - There may be some synonyms in the rating system, however this is due to the fact that they use different rating systems for movies and TV where movies are rated by the Motion Picture Association film rating system, and TV is rated by The TV Parental Guidelines.
  - There is some sprinkling problems where the metascore rating is being used instead of the standard MPA rating.
- Date and title
  - No data loss
  - Title is going to be an identifier and date is numeric
  - There cant be any outliers for the title as they are all unique
  - Histogram for Table A year



- Histogram for Table B date
  - Because the date uses the full date and not just the year there are very few movies that came out on the same day.

- There is a format for the date on Table B and all the values follow the same format.
  - There is no values sprinkled over the items
- I used pandas, matplotlib.pyplot, and datetime to help fix some issues with my data.