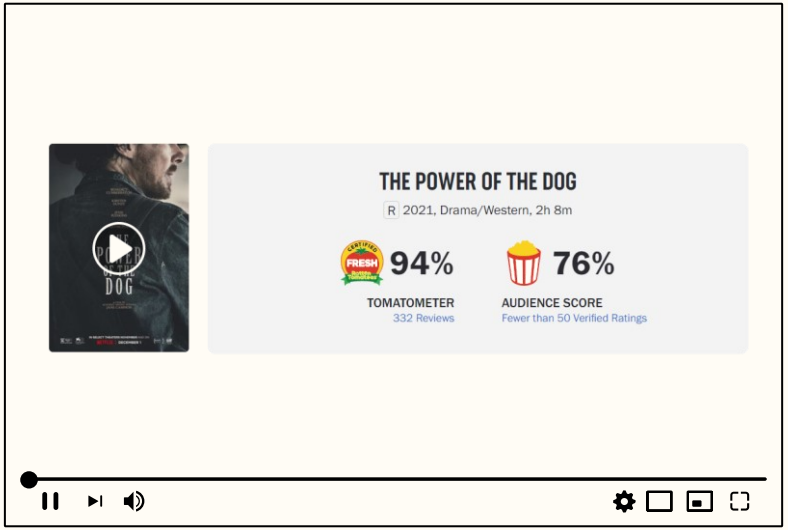
Movies Rating Distribution Analysis

Zimu Su/ Metis Business project

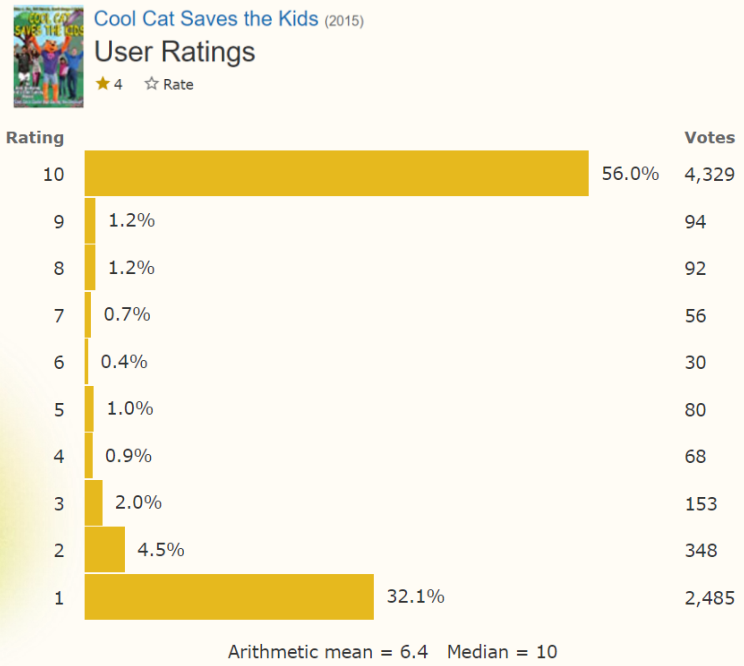


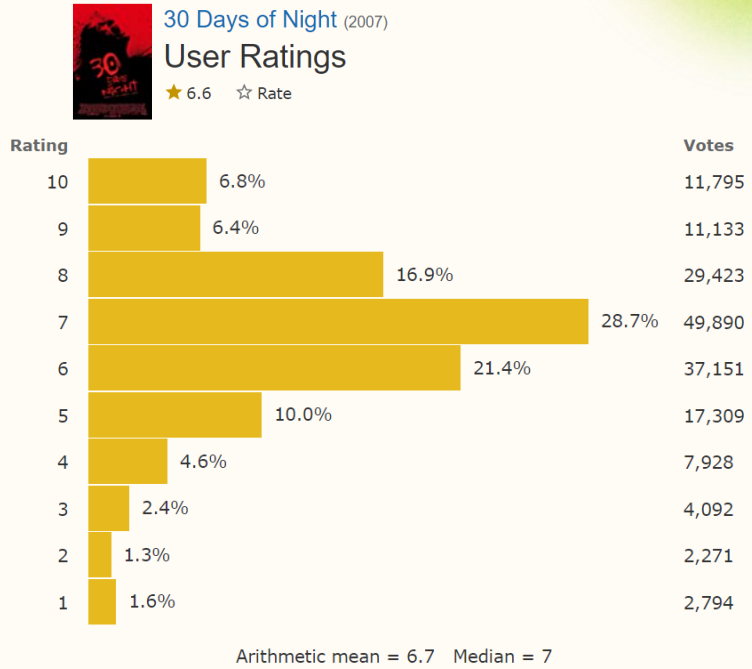
Ineligible impact of movie user score

- woven into the ticket-buying behavior.
- affect box-office performance.
- influence the decision to make a movie.
- important in recommendation system

....

Problem: Average scores removes distribution information.





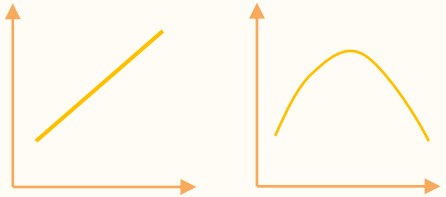
- The above two movies with completely different distribution give the almost same average scores.
- Distribution-based score analysis can provide elaborate audience's feedback.

Overview of distribution-based model

Movie Feature

Genres ,Release date, Title, Budget, Cast...

Rating distribution
(Mediator)



- Investigate/ Predict box office performance, popularity.
- Improve recommendation system.
- Investigate/predict audience type for popularity/unpopularity ratings.

....

Solution path

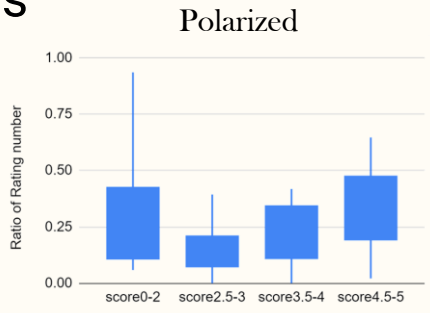
Identify rating distribution (rating info from Grouplens)

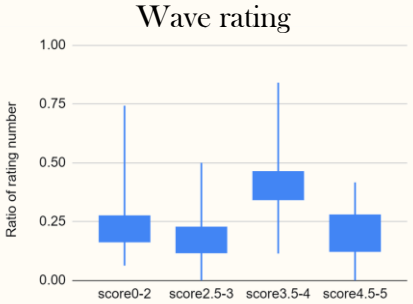
Identify relationship between movies features and distribution pattern.

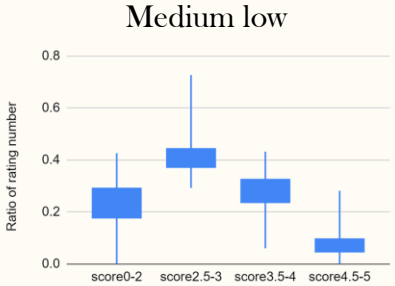
Use probability distribution function to characterize the shape.

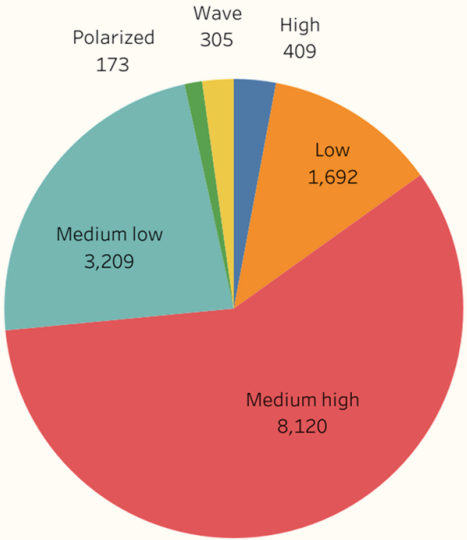
Regression analysis between features and distribution pattern.

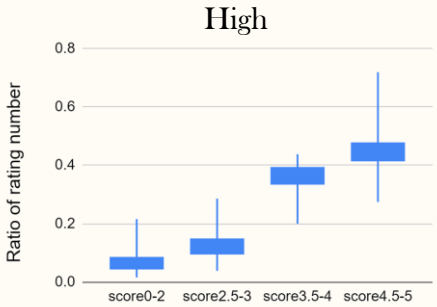
Distribution analysis

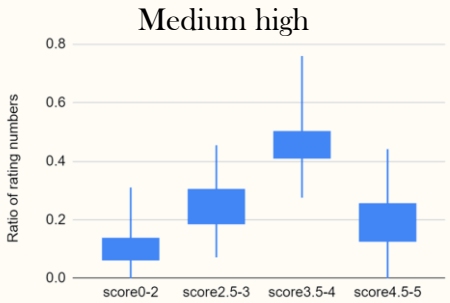


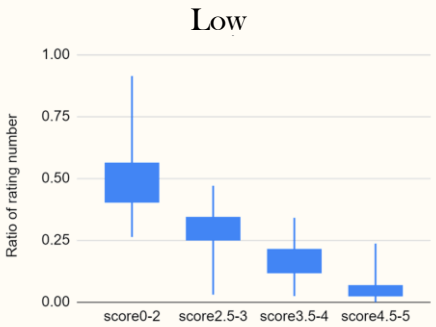




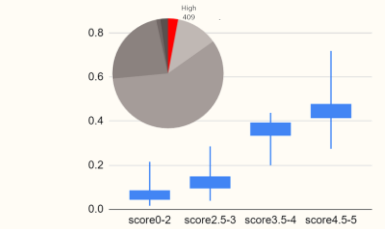


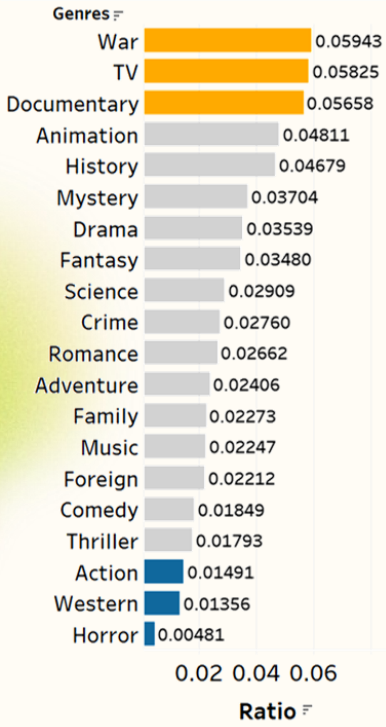


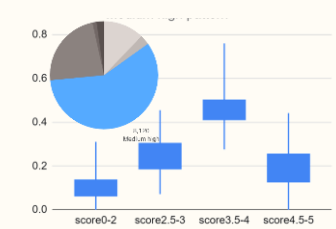


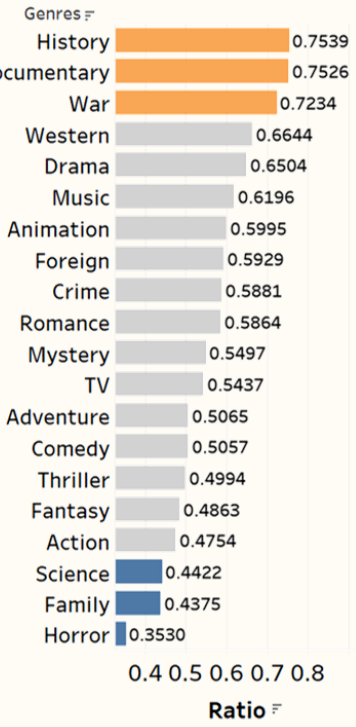


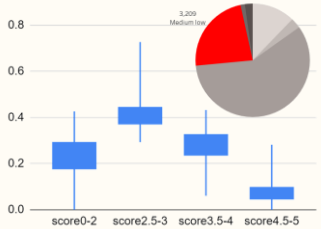
Genre analysis

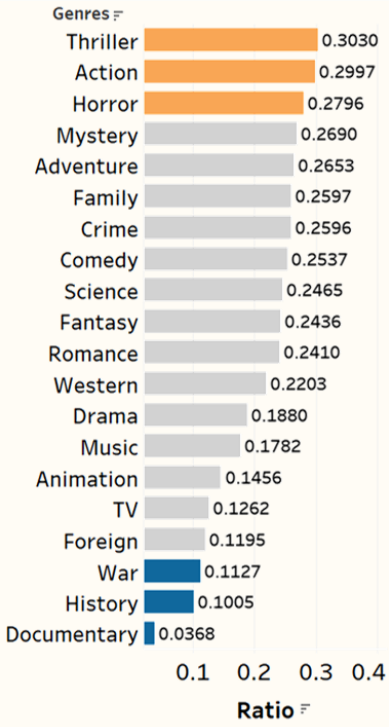


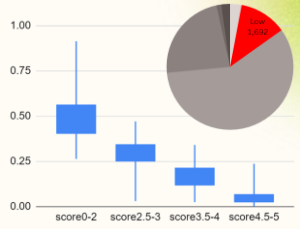


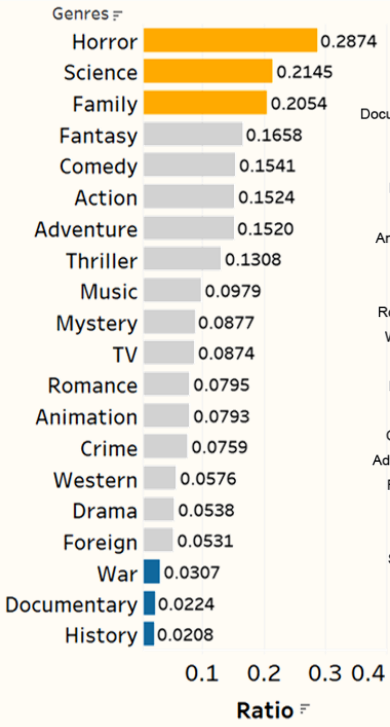


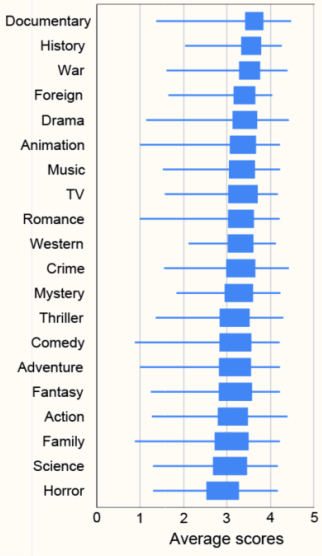






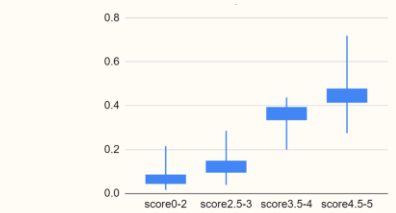




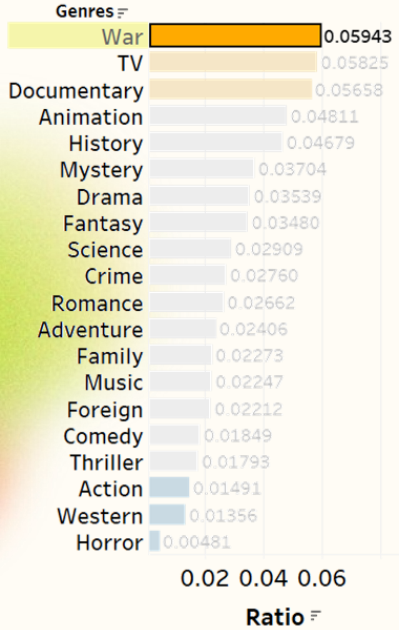


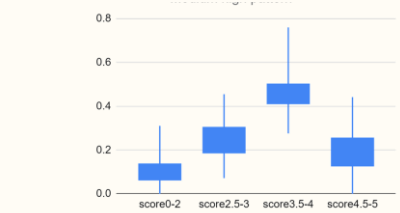
Consistent good rating

Other examples: Documentary, animation, history

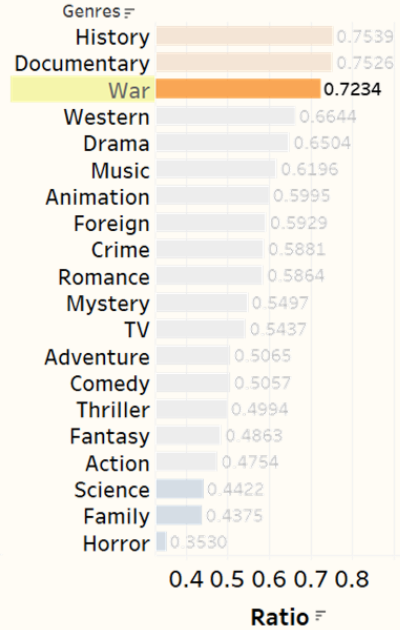


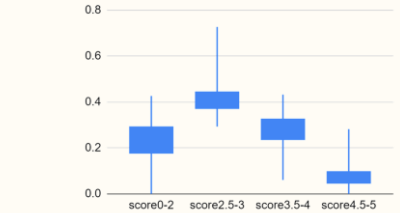
High rating



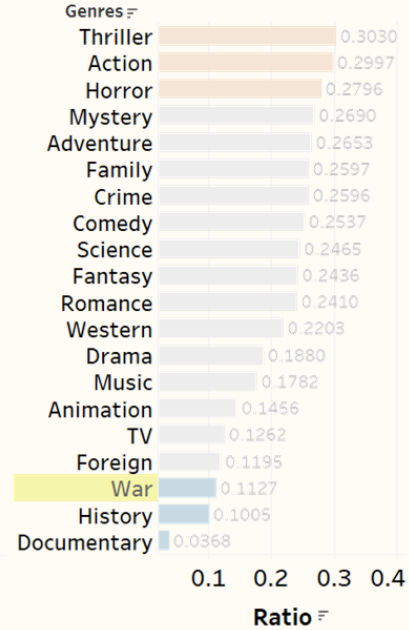


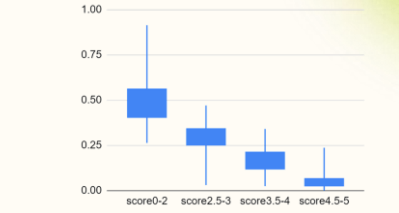
Medium high rating



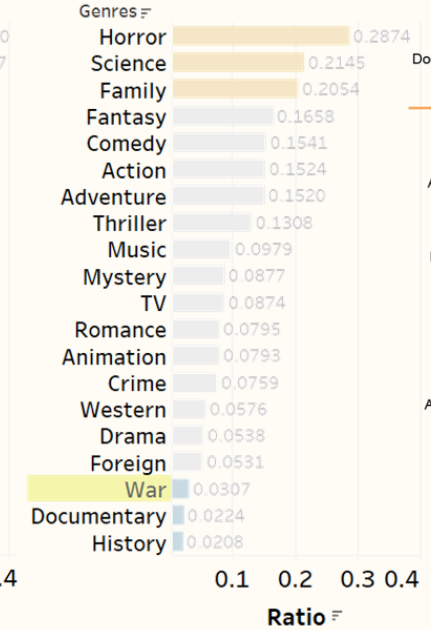


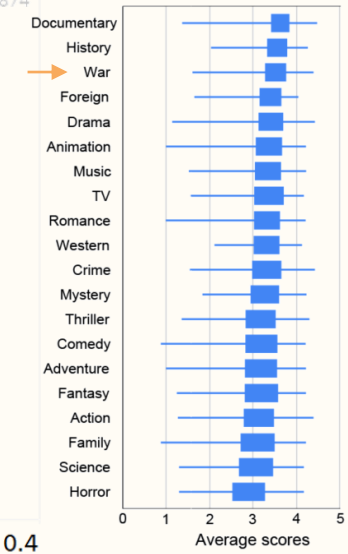
Medium low rating





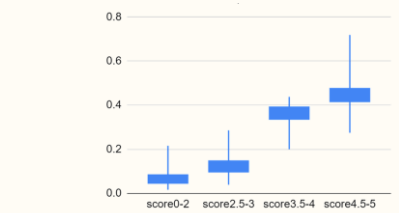
Low rating

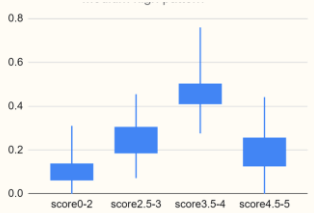


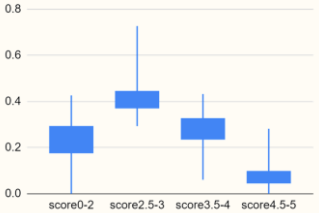


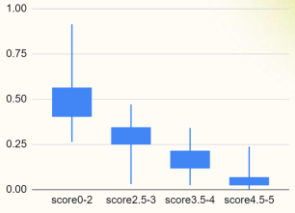
Consistent bad rating

Other examples: Horror, Thrill

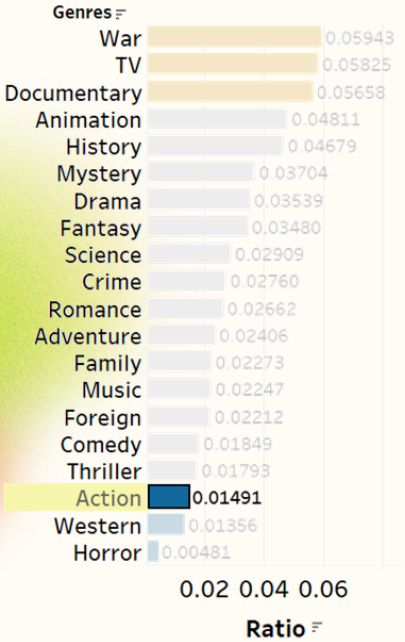




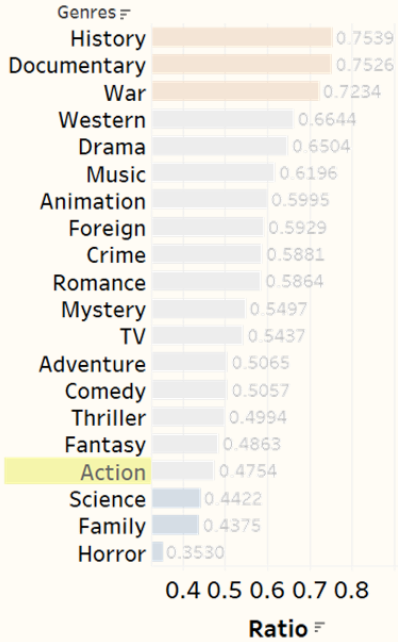




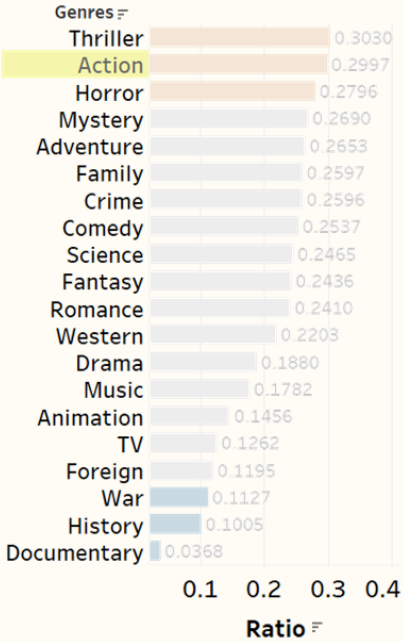
High rating



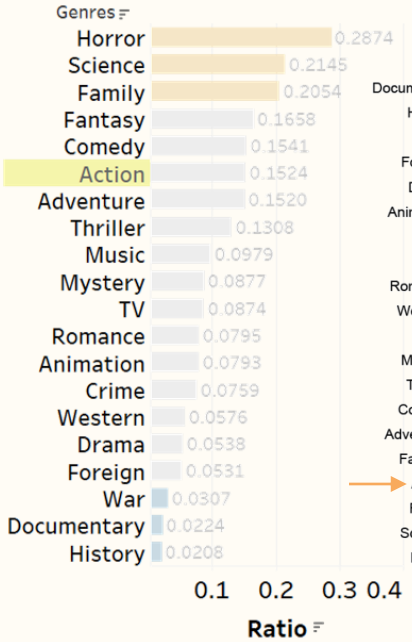
Medium high rating

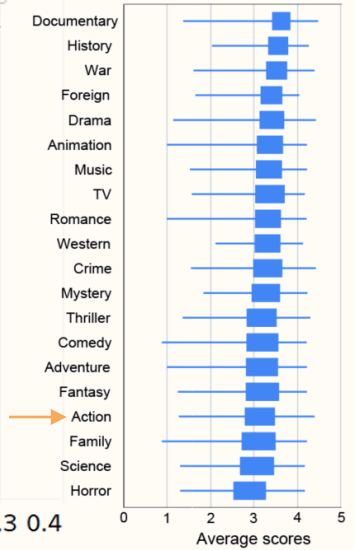


Medium low rating



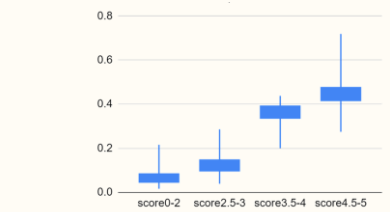
Low rating



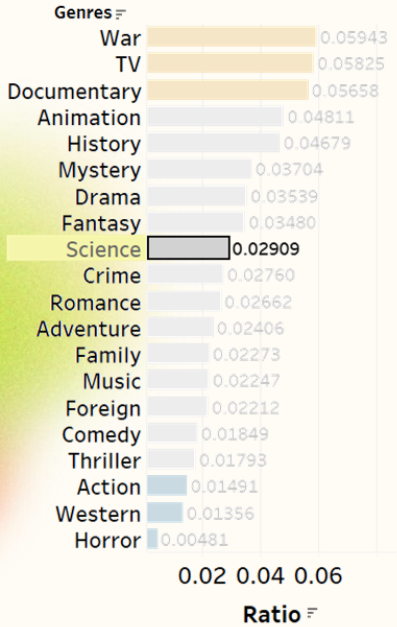


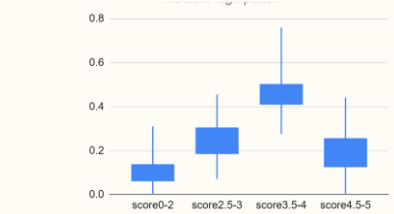
Some are excellent but most of them are bad

Other examples: Fantasy, Mystery, Adventure

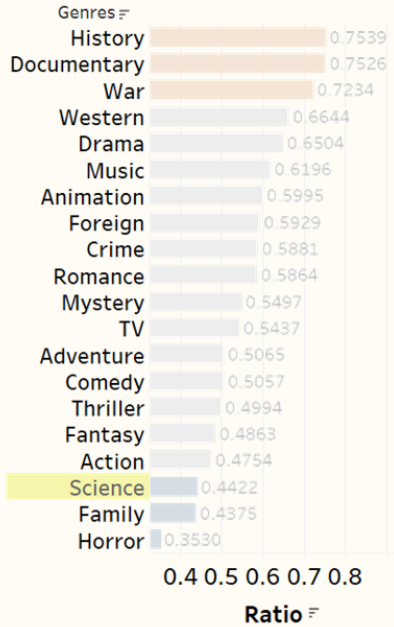


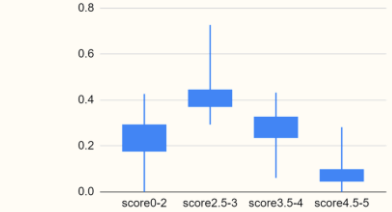
High rating



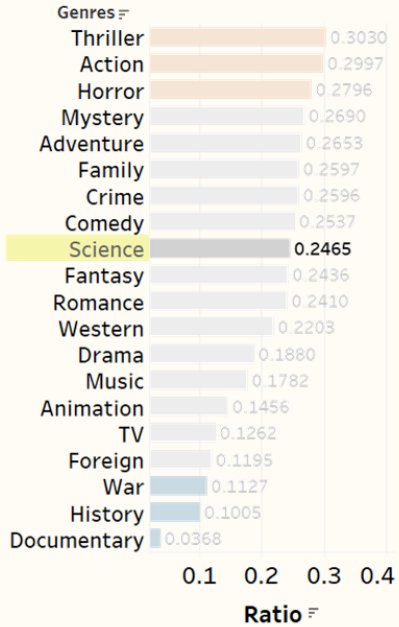


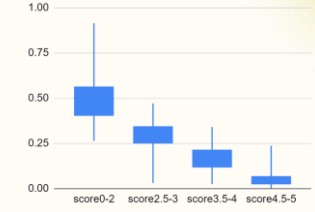
Medium high rating



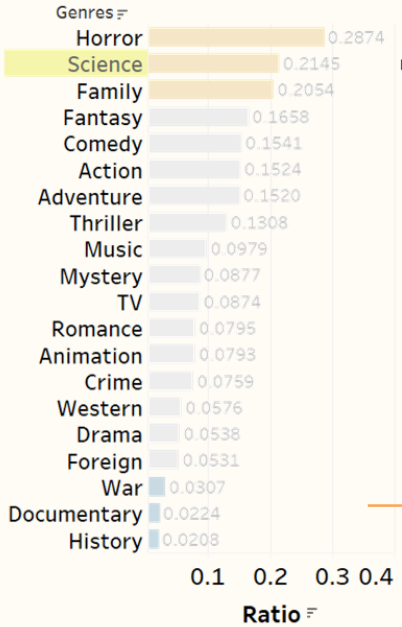


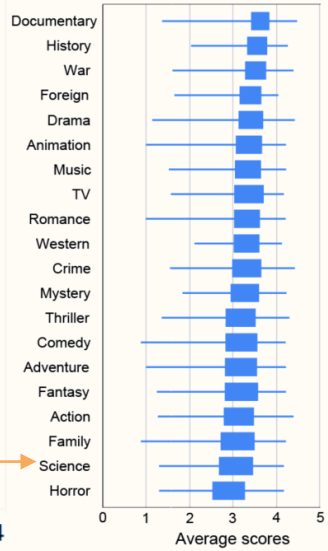
Medium low rating



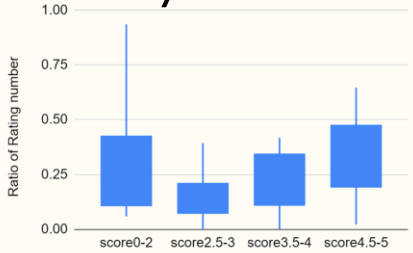


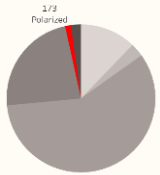
Low rating

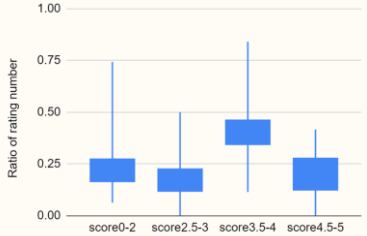


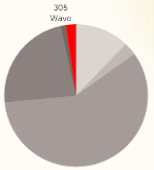


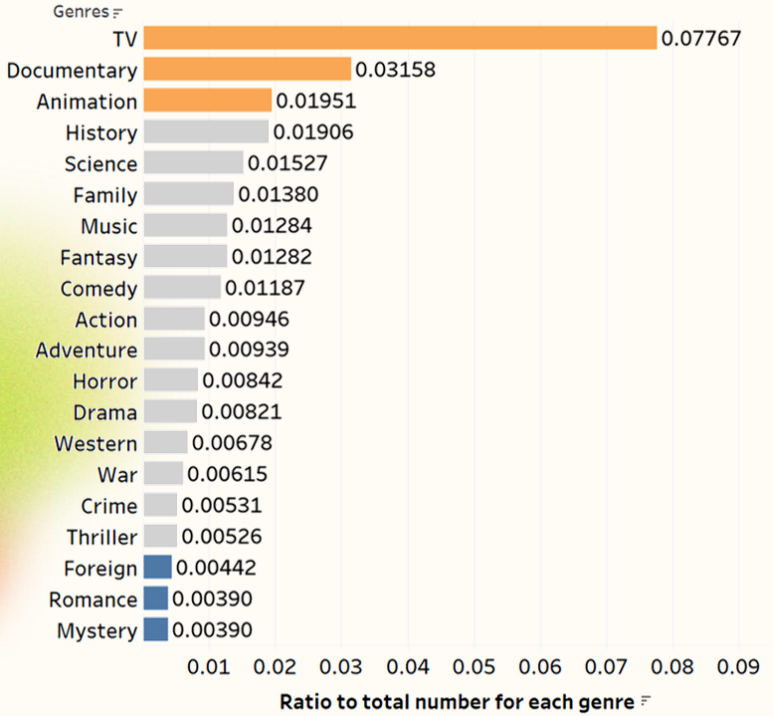
Genre analysis of controversial movies

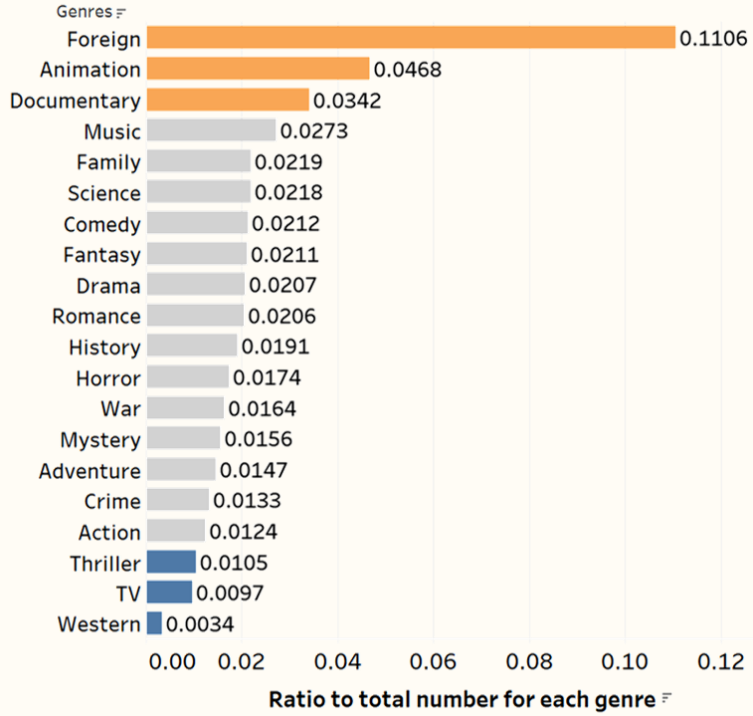


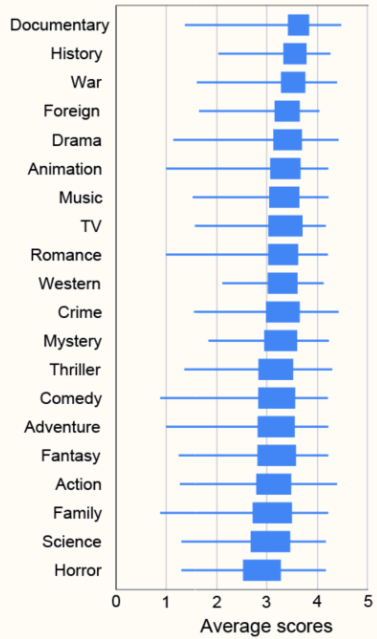




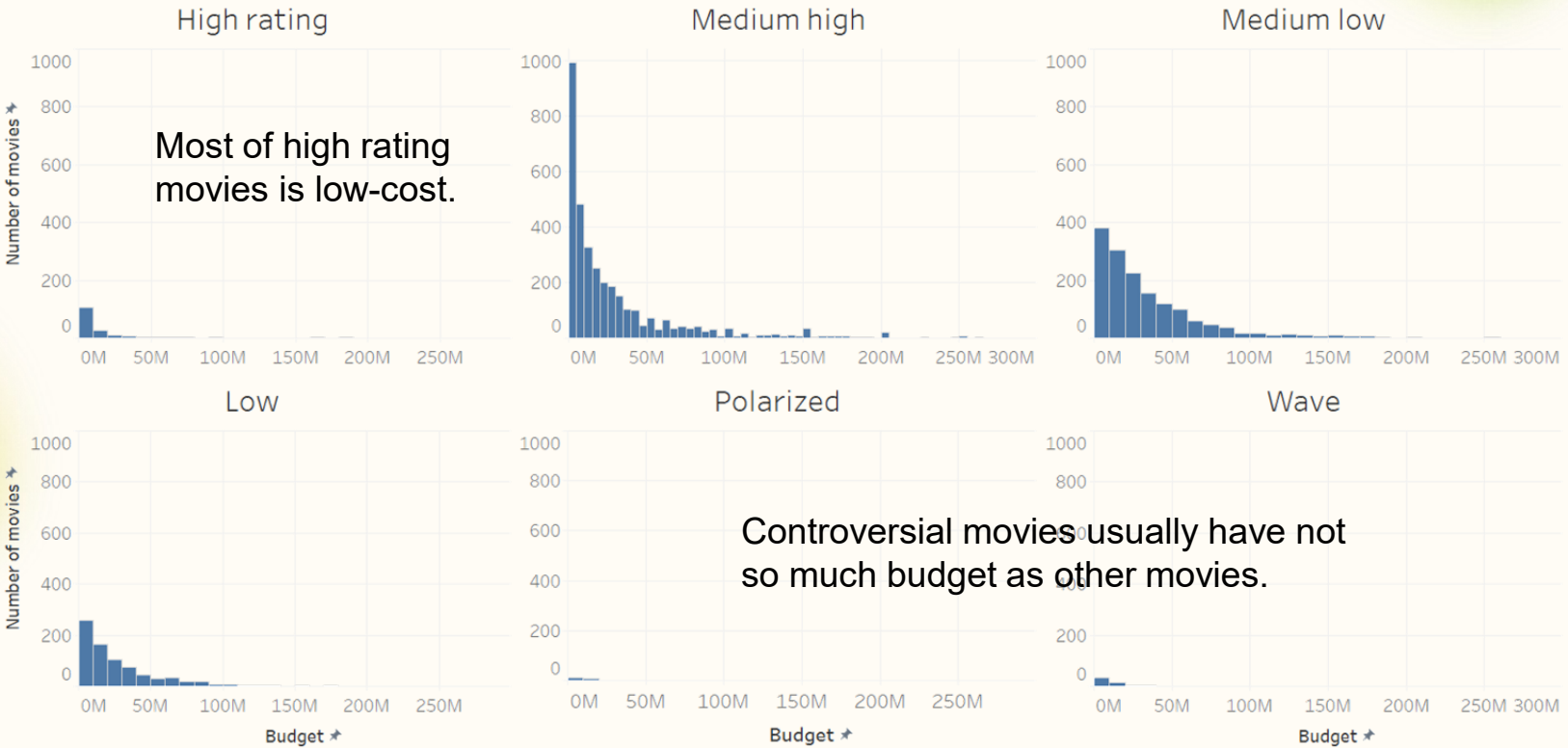








Budget analysis



Next step:

- Parameterize the distribution shape.
- Incorporate more features such as title, cast, director, etc.
- Build regression model between features and distribution/categories and parameters.

Thank you