

# MovieLens

---

Team Members: Justin Yeung, Swati Mudhelli, Mohan Ganesan,  
Neeraj Kumar, Wei Zhang

Group 7

# What We Did and Found

- Our team investigated different aspects of movies and users that impact movie ratings.
- Key Findings:
  - a. Certain genres tended to correlate with how high a movie rating is.
  - b. Users who rate a lot of movies, tend to rate movies lower.
  - c. Ratings and revenues were largely correlated. We used the small dataset for ratings, and dataset from kaggle which listed movie revenues for 2200 movies.

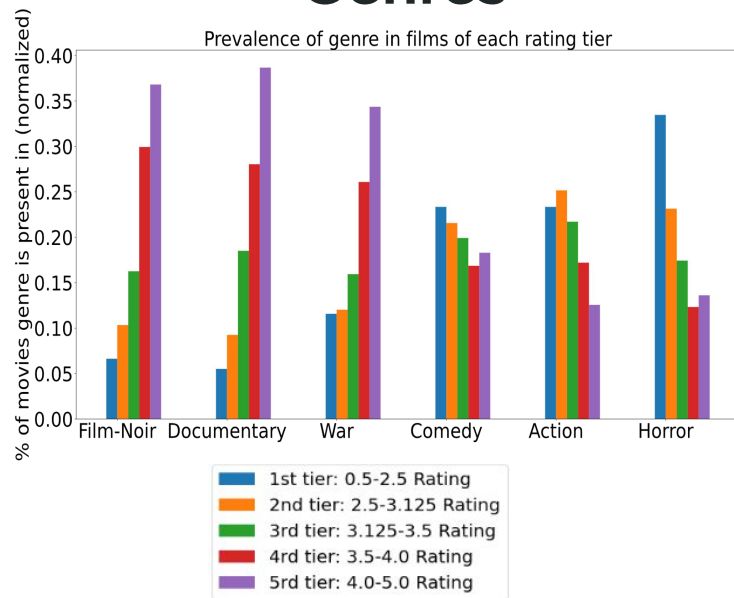
## **DataSets Used:**

MovieLens Latest Datasets - <https://grouplens.org/datasets/movielens/latest/>

<https://www.kaggle.com/c/tmdb-box-office-prediction>

# Movie Aspects

## Genres



### Higher Ratings

- **War, film noir, documentary**
- More high-brow & intellectual

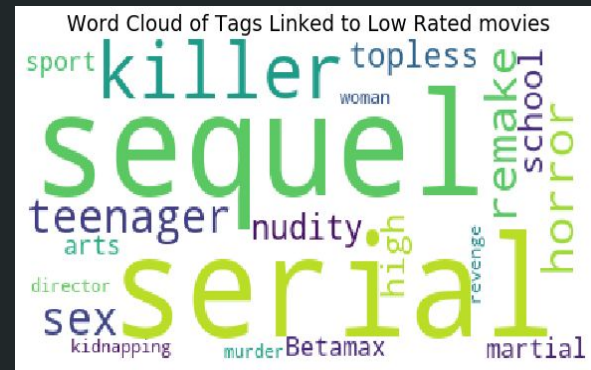
### Lower Ratings

- **Horror, action and comedy**
- Low on creative substance

## Tags

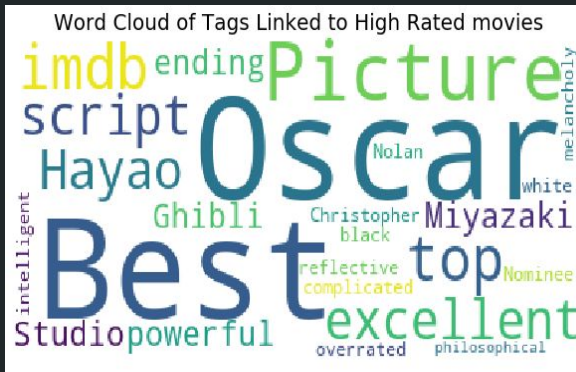
### Low-rated movies

- Cheap thrills (**violence, nudity**)
- More common in **horror/action**

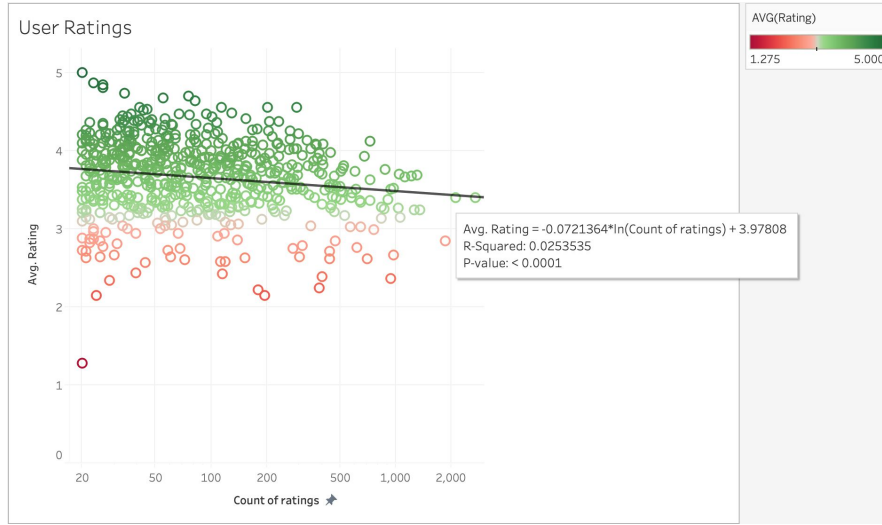


### High-rated movies

- Intellectual/realistic elements (**scripts, philosophies**)
- More common in **documentary and war**.

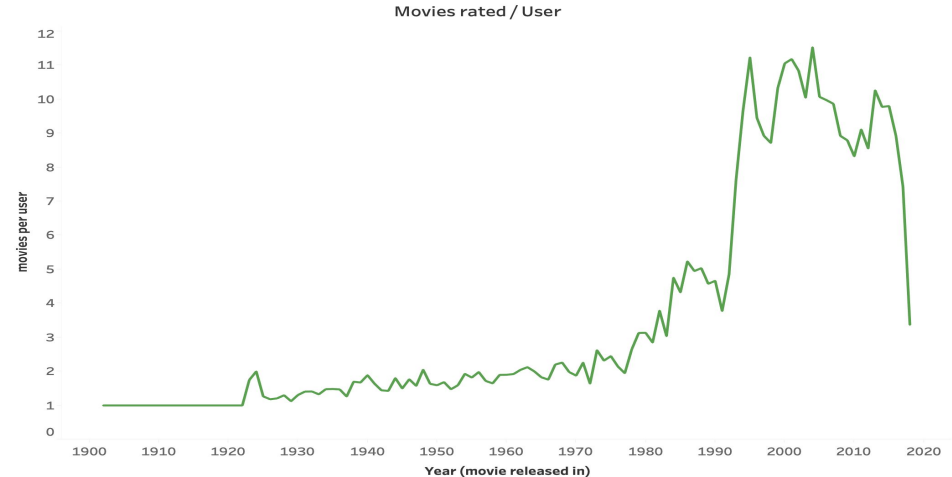


# User Behavior



- 3X spike in viewership and rating for movies made after 1980, compared to older movies
- iMDB members may tend to be from the millennial generation and after (40 or younger) and might not be aware of movies before 1980

- 78.03% of users had between 20 and 200 movie ratings but some 12 people rated over 1,000 movies
- A user with 1,000 movies rated left a rating that was on average, 0.166 lower than a user with 100 ratings
- Low p-value shows statistical significance in data



# Box-office & Conclusions

## Conclusions

- High brow (drama/documentary) films tended to be rated higher.
- User ratings and their number of reviews had a negative relationship and newer movies tended to be more viewed.
- In general, movies with higher rating tend to make more box office revenue. However, movies rated 0-1 also have a higher avg revenue because of popular movies (franchises, high budget movies, popular actors/directors etc)

*\*<https://www.kaggle.com/c/tmdb-box-office-prediction>*

Avg Revenue (in millions) - by rating

