

A romantic scene from the movie Titanic, featuring Leonardo DiCaprio and Kate Winslet on the ship's deck at sunset. They are embracing, with their arms outstretched towards the horizon. The sky is a mix of orange, pink, and blue, and the ocean is visible in the background. The image is used as a background for a presentation slide.

What Can Make a Movie Successful ?

Hakob Antonyan
Jun Yoo
Mohan Xing
Vicky Zhang

Introduction:

- The United States is one of the most prolific producers of films in the world today.
- A movie's success at the box office is of utmost importance to movie production companies and investors.
- The prediction of the success of a movie is essential to the film industry nowadays.
- In this project, we give our detailed analysis of 10,000 movies from 2009 to 2019 based on The Movie Database (TMDB) and Open Movie Database(OMDB) data to figure out what can make a movie successful.



Part 1 General Explore

- 1. Release Date vs. Movie Genre**
- 2. Release Date vs. Movie Runtime**



Part 2 Popularity and IMDB Rating

- 1. Movie Run time vs. Popularity**
- 2. Movie Run Time vs. IMDB Rating**
- 3. Genre vs. Popularity & IMDB Rating**
- 4. Rated vs. IMDB Rating**

Part 3 Budget, Revenue, and Profit

- 1. Genre vs. Budget, Revenue & Profit**
- 2. IMDB Rating vs. Budget/Revenue**
- 3. Rated vs. Budget, Revenue**

Data limitations

2009 - 2019

United States

.DropNA()

```
[5]: clean_df.sort_values('omdb_runtime', ascending=True).head()
```

[5]:	omdb_title	Genres	omdb_released	omdb_runtime	omdb_country
7882	Code 8	['Science Fiction', 'Action']	22 Mar 2016	10 min	USA, Canada
4374	The Punisher: Dirty Laundry	['Action', 'Crime', 'Drama']	15 Jul 2012	10 min	USA
8363	Tron: The Next Day	['Science Fiction', 'Thriller']	14 Mar 2011	10 min	USA
6957	Under the Sea: A Descendants Story	['Music', 'TV Movie', 'Fantasy']	28 Sep 2018	10 min	NaN

Technical Difficulty Coding Skills

```
genre_df = clean_df[['Release Date', 'Genres']]
new_df = pd.DataFrame(genre_df.Genres.str.split(",").tolist(), index=genre_df['Release Date']).stack()
new_df = new_df.str.replace(r'^[a-zA-Z]\s?', '', regex=True)
```

APIs:

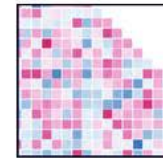
- The Movie Database (TMD)
- The Open Movie Database(OMDB)



OMDb API
The Open Movie Database

Tools

- Excel
- Pandas
- Numpy
- Matplotlib
- Seaborn



Seaborn

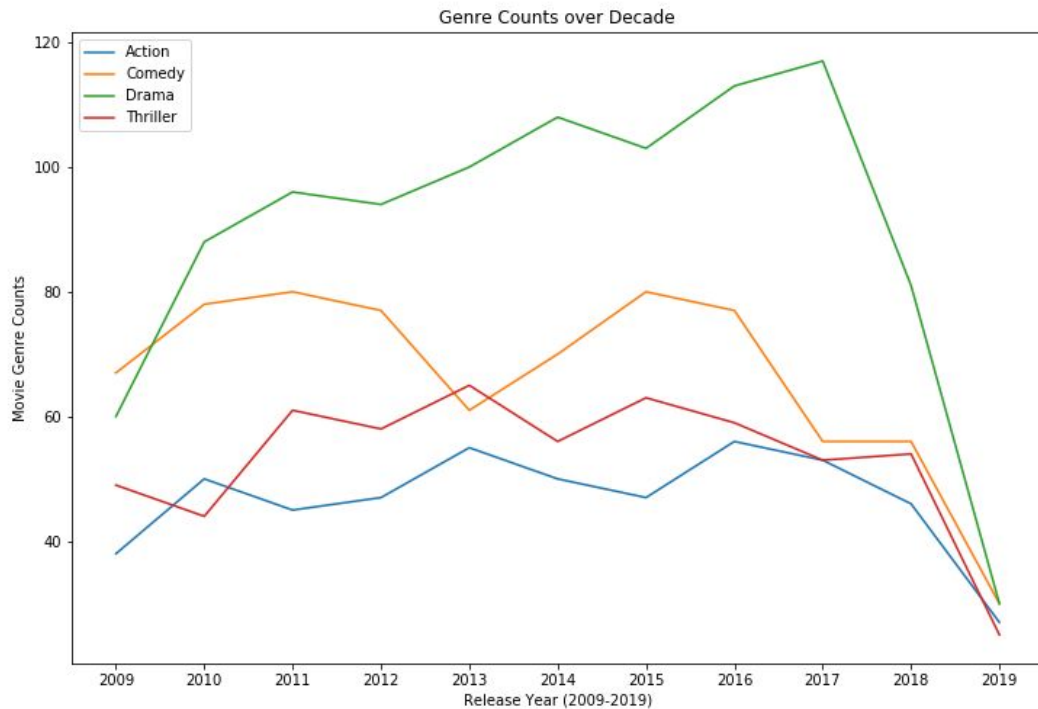


Research Part I:

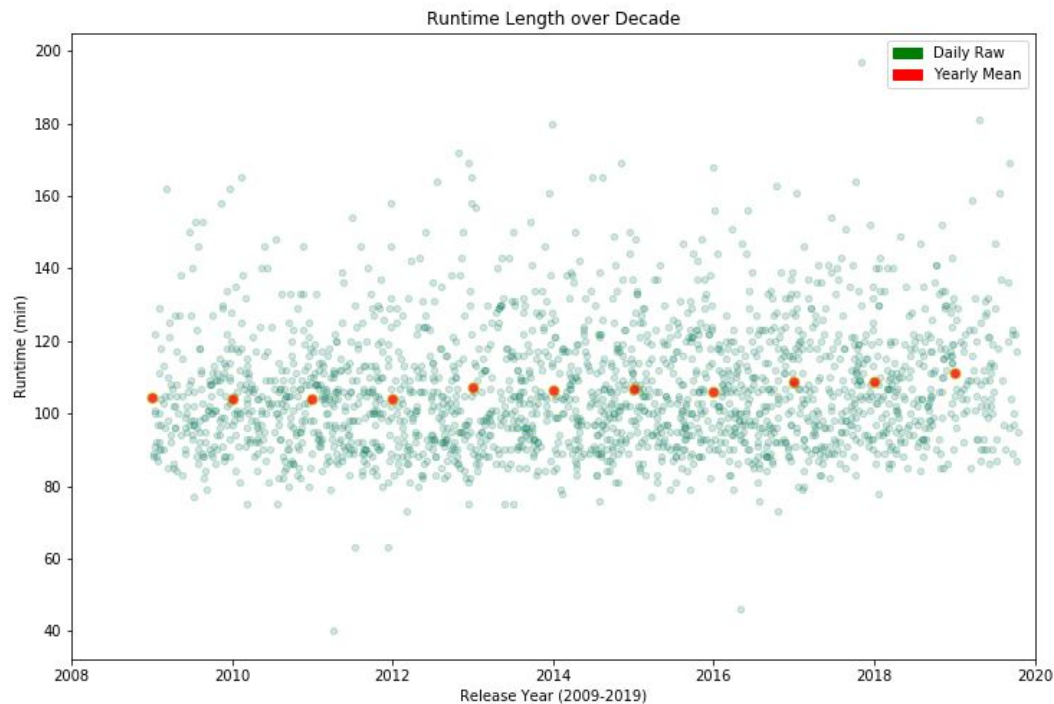
- Top 4
1. Action
 2. Comedy
 3. Drama
 4. Thriller



Release Year vs Movie Genre Counts

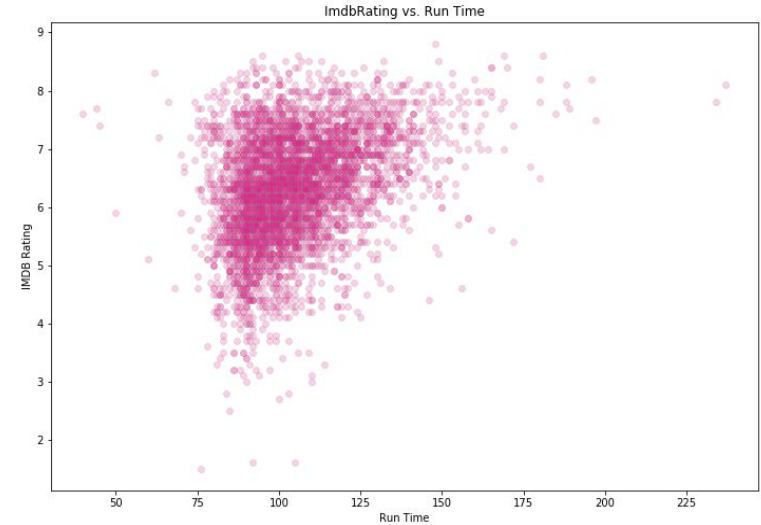
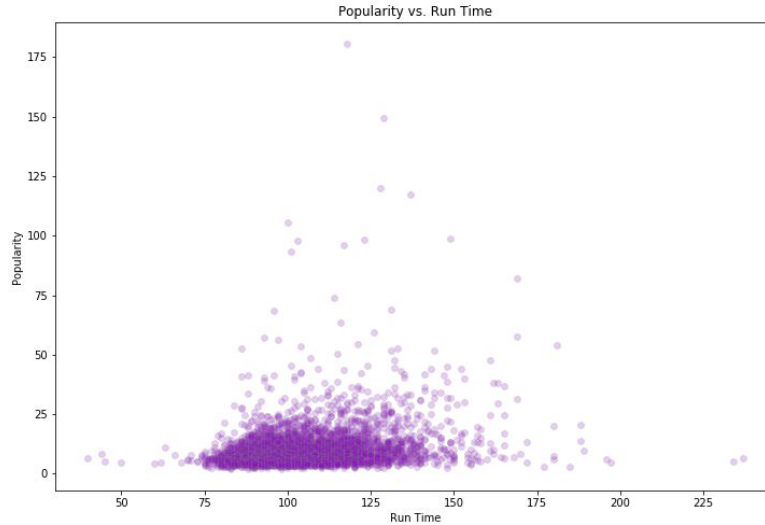


Research Part I: Release Year vs Runtime (min)

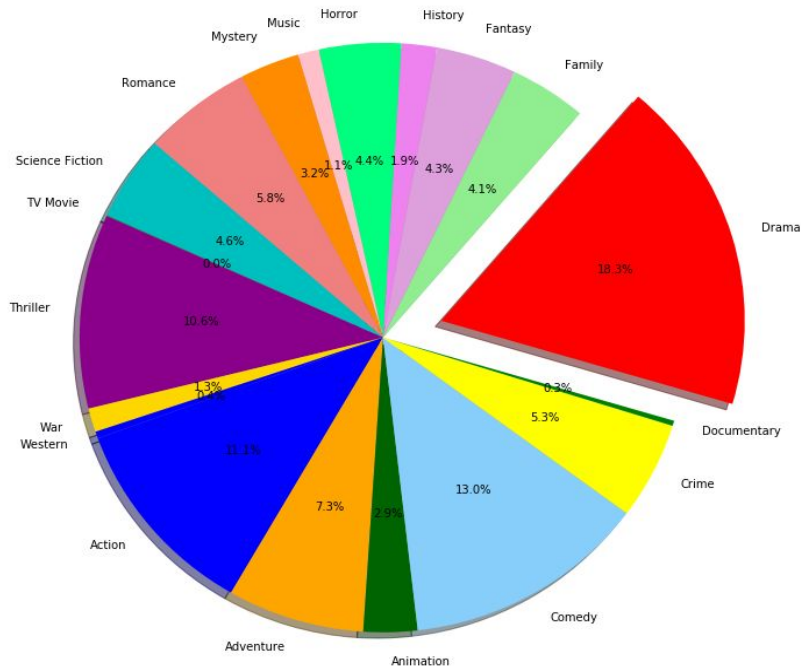


- Year-Month-Day Release Date
- Yearly Mean Release Date
- Steady

Research Part 2: Relationship between movie run time and popularity/IMDB rating

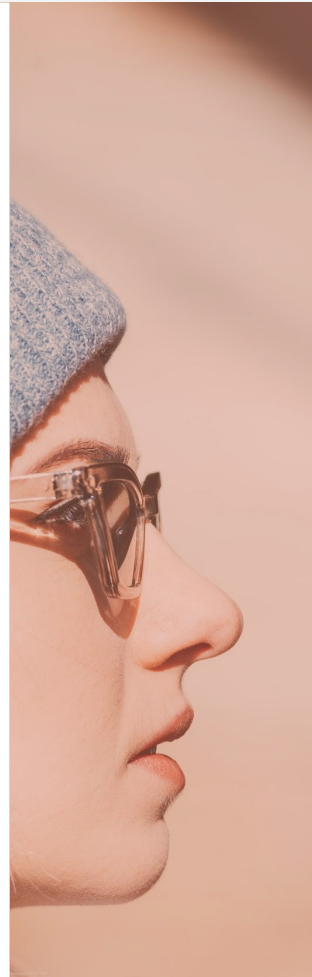


- Most of Movie run time is between 90 min to 120 min
- Movie has shorter/longer runtime has lower popularity but higher IMDB rating
- Audience prefer movie run time between 90 min to 120 min

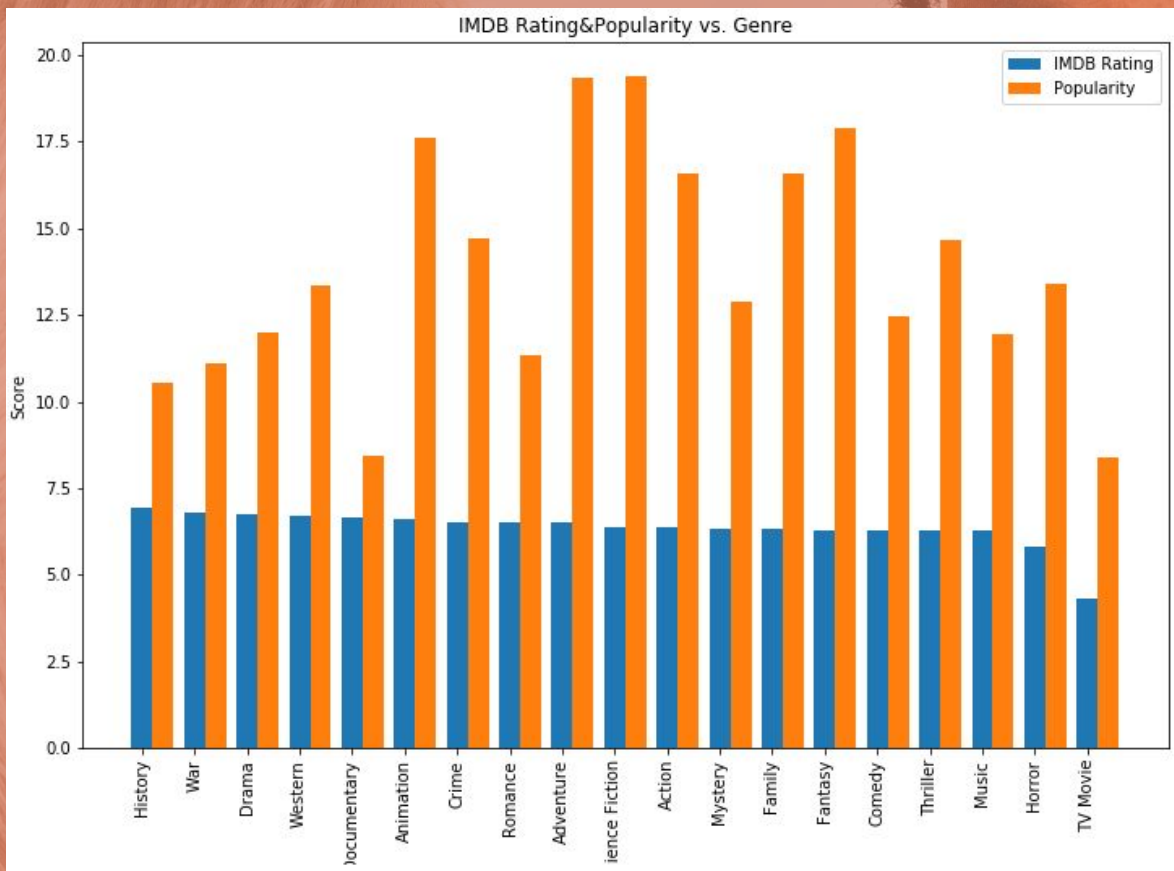


- Drama has the highest production numbers
- Genre vs. Popularity & Rating?

Research Part 2:
Find the
relationship
between genre and
popularity/ IMDB
rating



Research Part 2: Genre of Movie IMDB Rating and Popularity



Rank	Average IMDB Rating	Average Popularity
1	History	Science Fiction
2	War	Adventure
3	Drama	Fantasy
4	Western	Animation
5	Documentary	Family

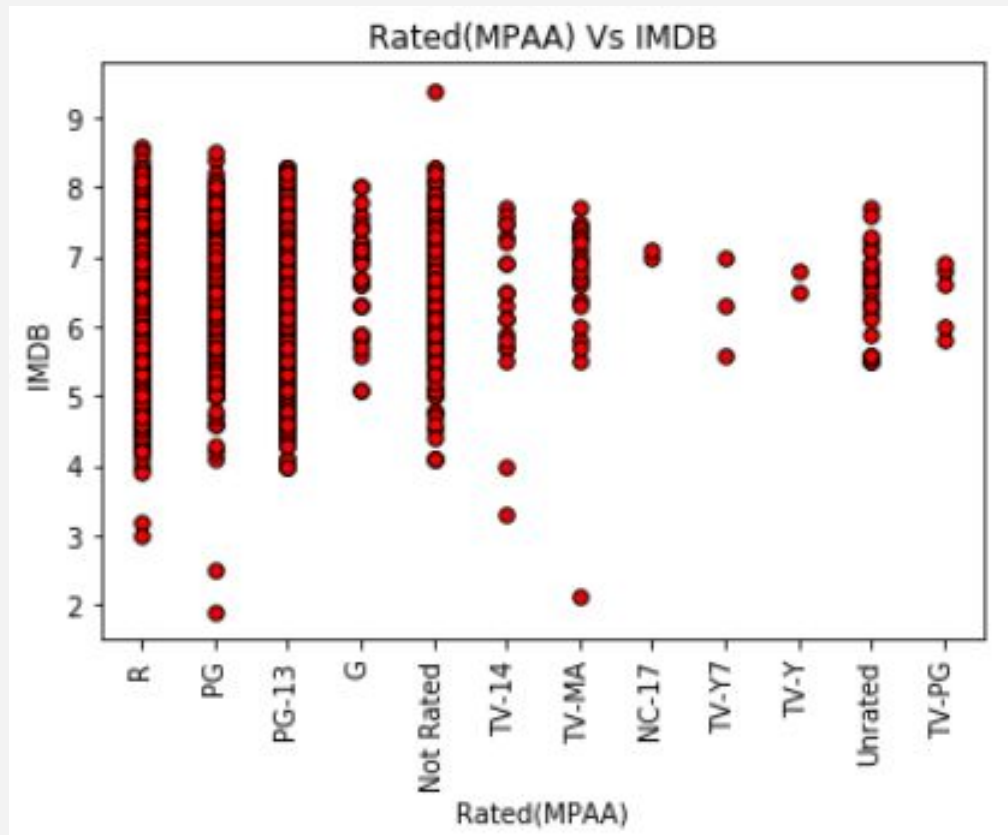
- The Popularity and IMDB Rating have no correlation

Research Part 2: Relationship Between IMDB and Rated (MPAA)

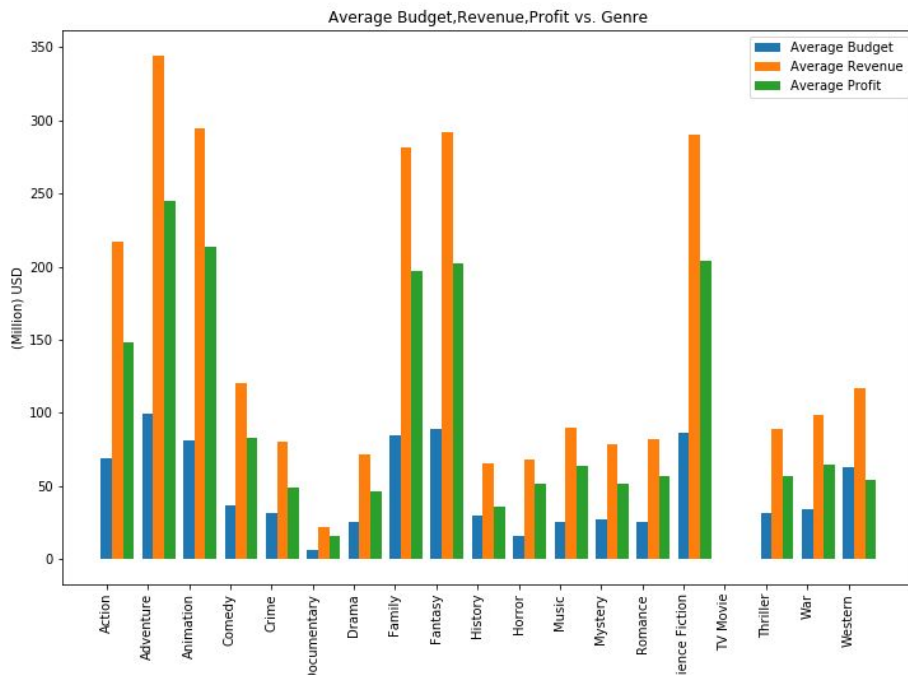
R	2575
Not Rated	1679
PG-13	1216
PG	521
TV-MA	426
TV-14	302
Unrated	203
TV-G	199
TV-PG	166
G	130
NOT RATED	30
TV-Y7	23
TV-Y	19
NC-17	4
TV-Y7-FV	3
M	3
Approved	1
TV-13	1
UNRATED	1

The mean for IMDB Rating:

- PG - 13 mean is 6.33
- PG - mean is 6.4,
- R - mean is 6.0
- Not Rated - mean is 6.2
- G - mean is 6.5



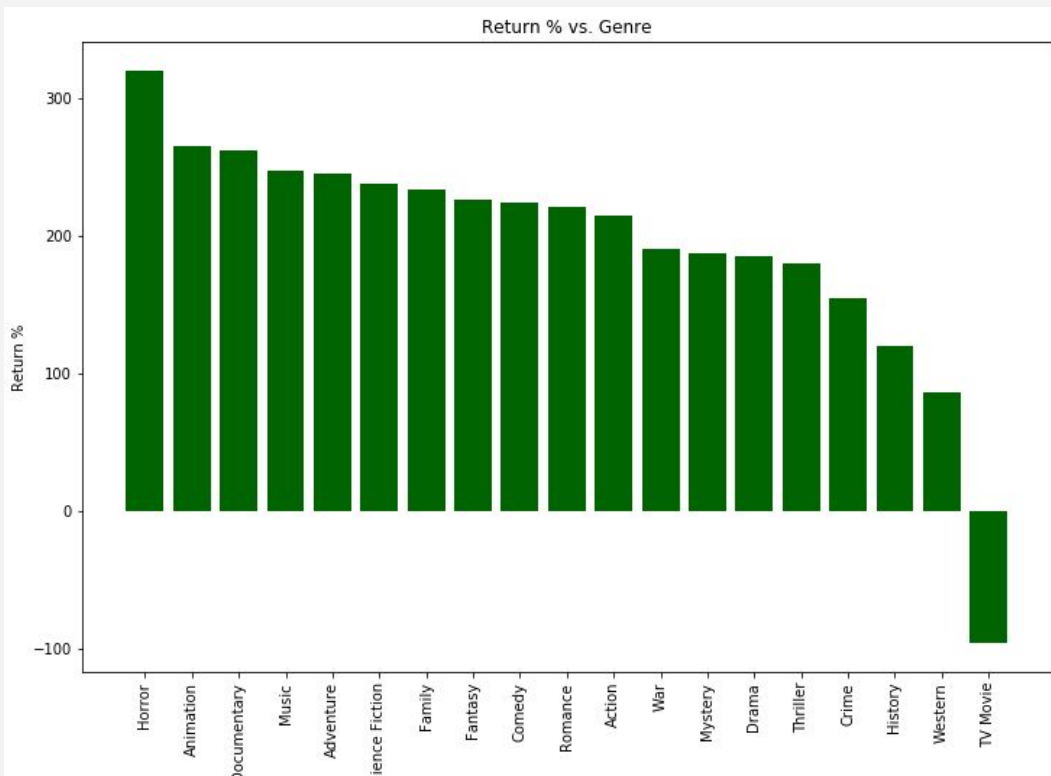
Research Part 3: Genre of Movie Budget, Revenue & Profit



Rank	Average Budget	Average Revenue	Average Profit
1	Adventure	Adventure	Adventure
2	Fantasy	Animation	Animation
3	Science Fiction	Fantasy	Science Fiction
4	Family	Science Fiction	Fantasy
5	Animation	Family	Family

- In general higher budget always brings higher revenue and higher profit

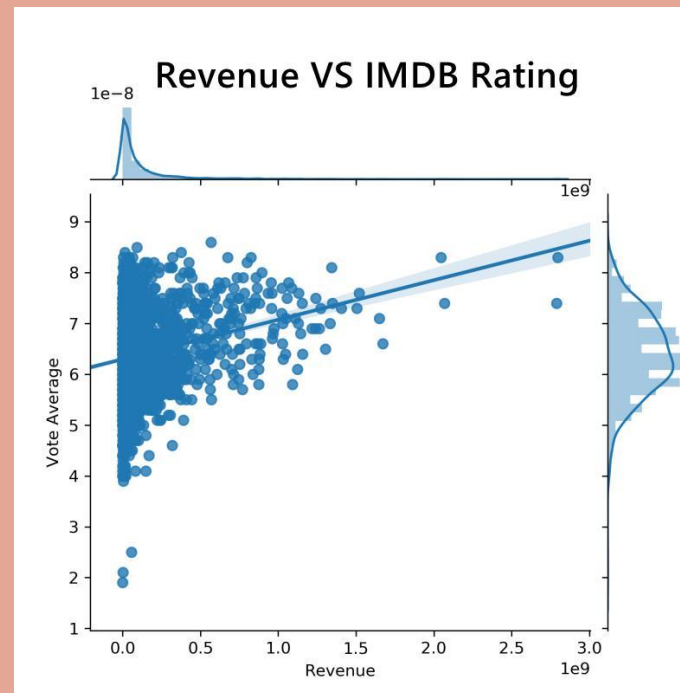
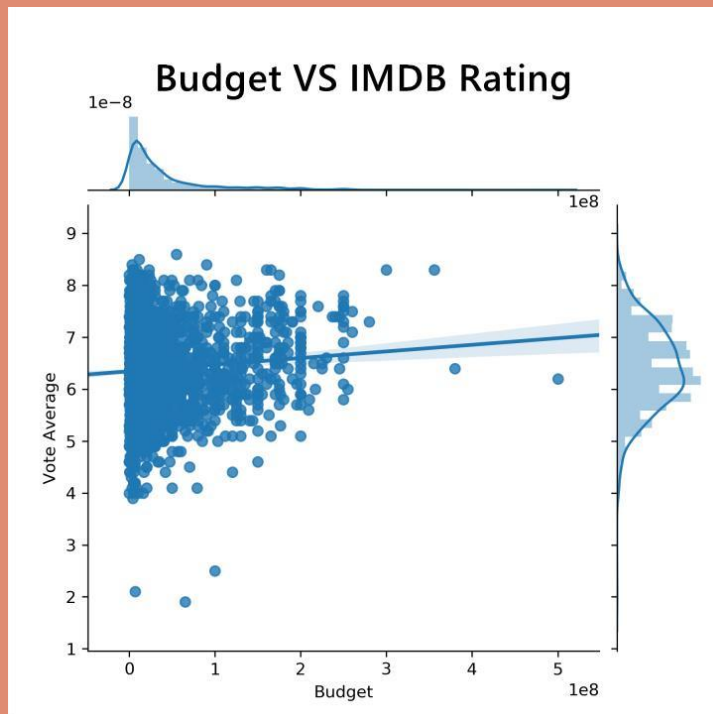
Research Part 3: Genre and Movie Profit



Rank	Return %
1	Horror
2	Animation
3	Documentary
4	Music
5	Adventure

- Horror type movie has a relative low budget, but it has the highest return rate than any other types of movie
- Higher budget movies do not have good return rates than lower budget movies

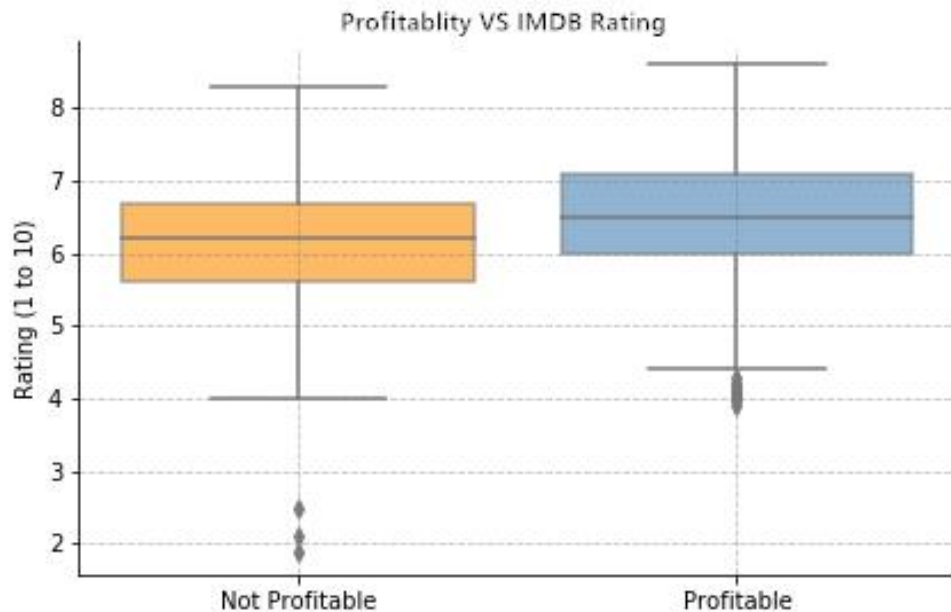
Research Part 3: Relationship between IMDB Rating and Budget/Revenue



On the assumption that the vote data is authentic:

- Most of Movie budget under \$10,000,000
- Most of Movie Revenue under \$100,000,000
- Most of Movie Review is between 5~8
- The higher is either of Budget or Revenue, less likely the movie rate is low

Research Part 3: Relationship between IMDB Rating and Budget/Revenue



On the assumption that the vote data is authentic:

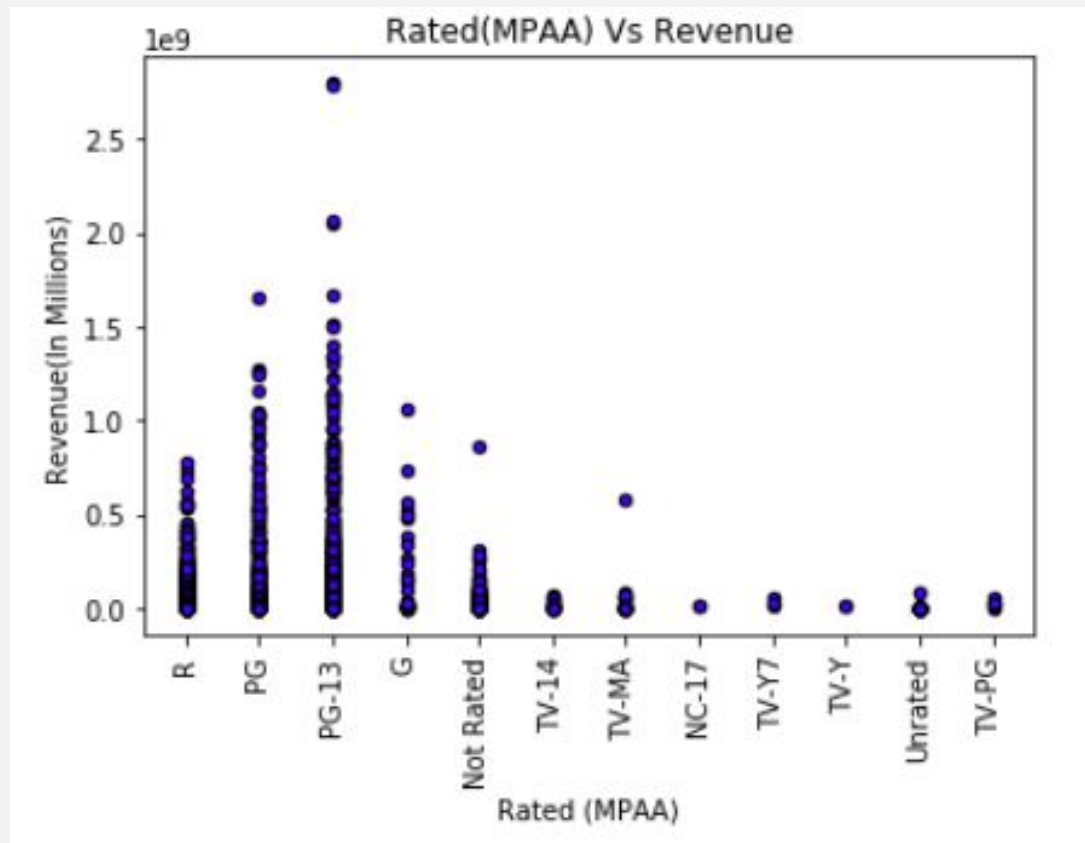
- On Average, The profitable movies' reviews are higher than non-profitable movies

Research Part 3: Relationship between Rated and Revenue

R	1224
PG-13	837
PG	310
Not Rated	304
G	30
Unrated	27
TV-MA	22
TV-14	20
TV-PG	5
TV-Y7	3
TV-Y	2
NC-17	2

The Revenue for Rated

- PG - 13
- PG
- R
- G
- Not Rated



Recap of the analysis

- Movies in drama genre is the highest production rate every year over the decade
- In course of decade long movie runtime history, there is definitely favored runtime
- There is no clear relationship between movie popularity and IMDB Rating
- Movie with higher IMDB rating does not seem to be having higher popularity
- Higher budget can always bring higher revenue
- Horror type movie has the highest profit return rate, but it always has low budget
- Either of the higher the budget or revenue, less likely the movie reviews is estimated low
- On average, profitable movie's ratings are higher than non-profitable movies
- Among the rated movies PG-13 and PG have the most revenue



Conclusion:

- If an investor wants to invest in making a movie, there are several factors that needs to be considered. We provided details analysis and visuals from movie database and compared different factors (mentioned in the rcap) to see which kind of movie would be successful.
- Based on our analysis we can predict that a movie in adventure/horror genre, with 90-120 minutes run time, rated PG/PG-13 with a high enough budget can have the potential to be a successful movie.

Challenges

- Finding The official IMDB
- Combine OMDb and TMDb
- Timing

Takeaway

- Teamwork and collaboration
- Learning from each other
- Improve the knowledge by doing a lot of research

A person wearing a dark suit and a light-colored shirt is holding a rectangular, light-colored sign with both hands. The sign has the word "QUESTIONS?" written on it in a bold, dark, serif font. The background is a solid, light-colored wall.

QUESTIONS?