



Computing Vision Business Proposal

August 2023.



Overview

Presentation Agenda

- Business Goals
 - Overview and Understanding
- Data
 - Usage and Analysis
 - Statistical Inference and Understanding
- Recommendations and Next Steps
 - Most Popular Genres
 - Profit per Release Month
 - First Movie to Create
- Concluding Thoughts



We seek to analyze extensive data from past films, aiming to understand what specific factors lead to the **greatest probability of success** in the creation of future movies

Data by the Numbers

How the team handled and analyzed the data

01

Start with Reliable Data

We used data from five sources:

- *Box Office Mojo*
- *IMDB*
- *Rotten Tomatoes*
- *TheMovieDB*
- *The Numbers*

02

Filter and Combine Data

We merged data sources for larger, more useful data by combining on filters such as name or ID. This allowed for us to later filter genres, dates, and more.

03

Analyze and Understand the Data

We finally searched through the data, looking for trends, leading our recommendations to follow. We then we able to create visualizations and tests to support our analysis

A Glimpse into our Data:

73,856

Movies

\$91.5 Million

Average Gross

\$31.5 Million

Average Budget

93 Minutes

Median Runtime

6.33

Average Rating

Documentary

News

Biography

Music

History

Sport

War

Reality TV

Musical

Drama

Family

Animation

Romance

25
Genres

Adventure

Crime

Comedy

Mystery

Fantasy

Action

Western

Game-Show

Thriller

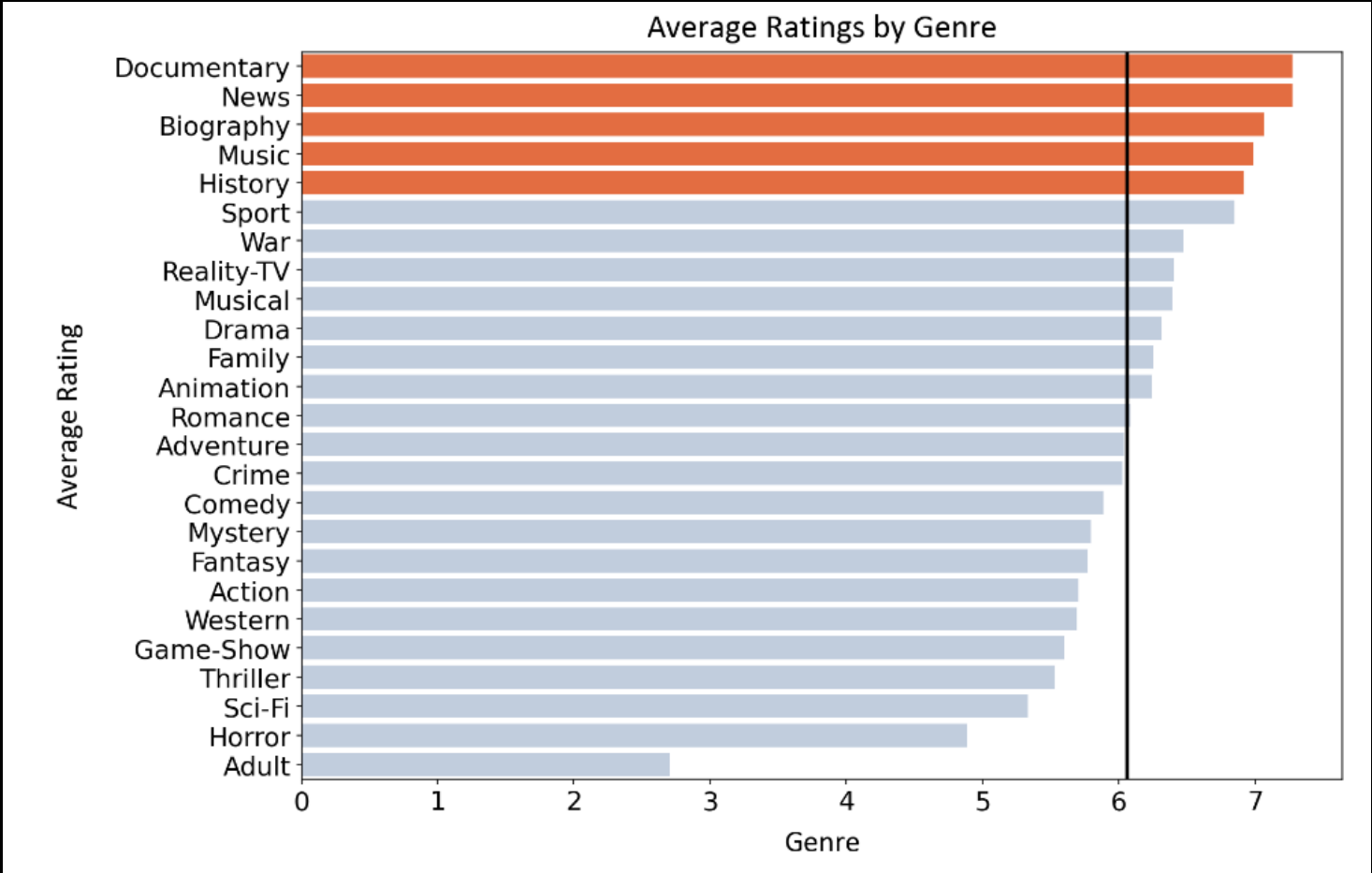
Sci-Fi

Horror

Adult

Ratings per Genre

Recommendation 1

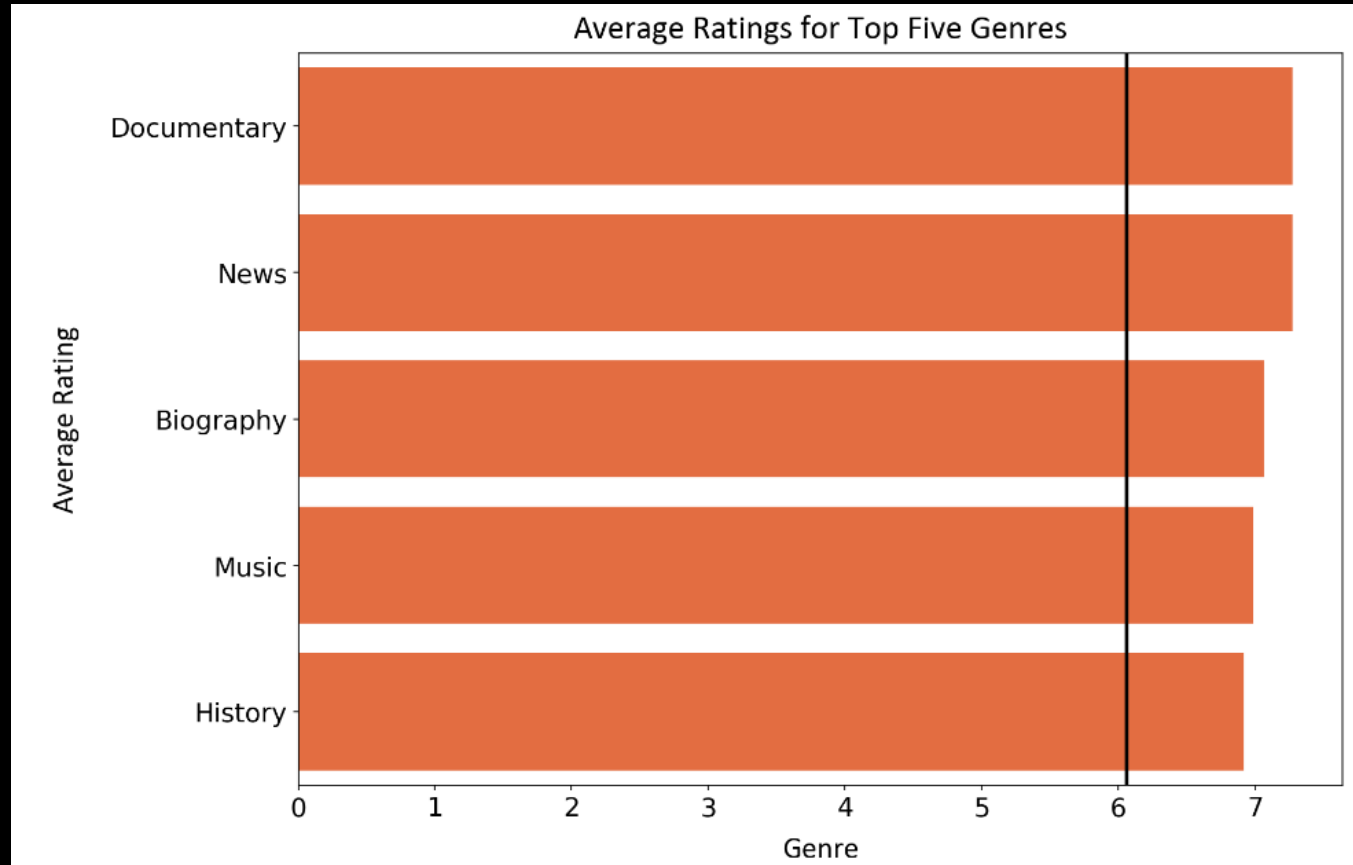


Genres impact rating

- Aim to make a film with average ratings above the average of all genres for the greatest likelihood of positive reviews among viewers

Recommended Genres

Recommendation 1 Continued

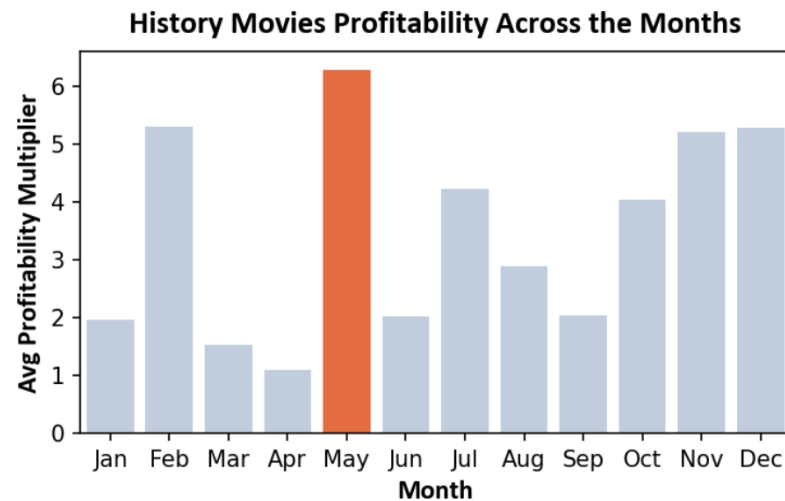
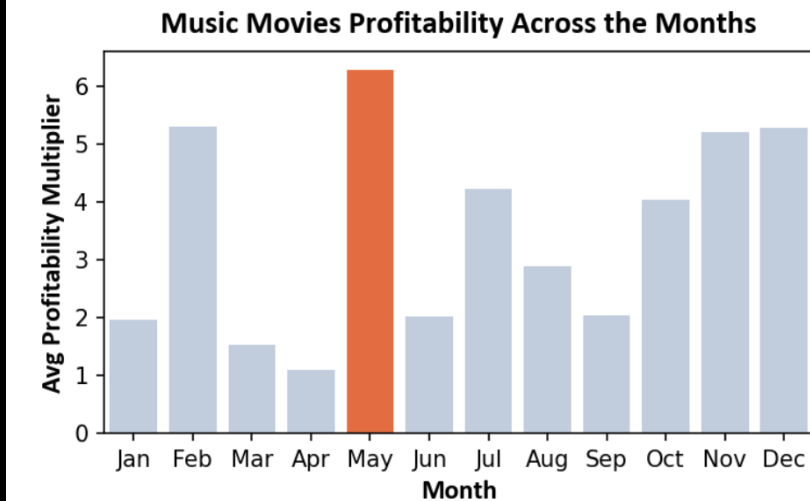
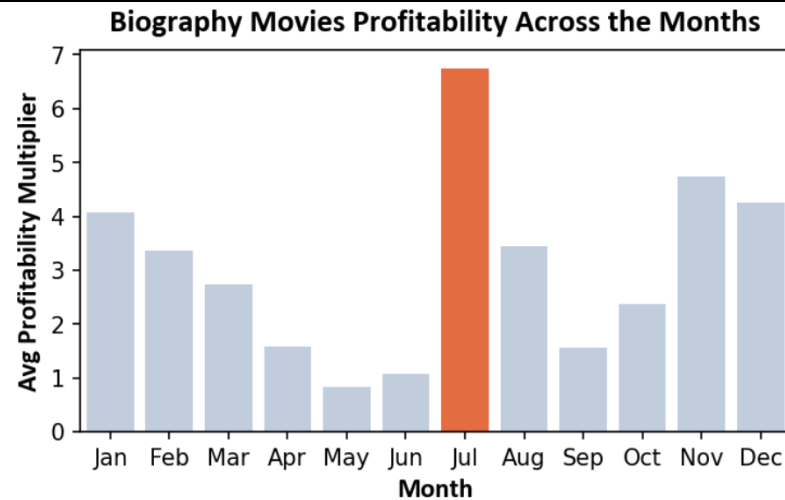
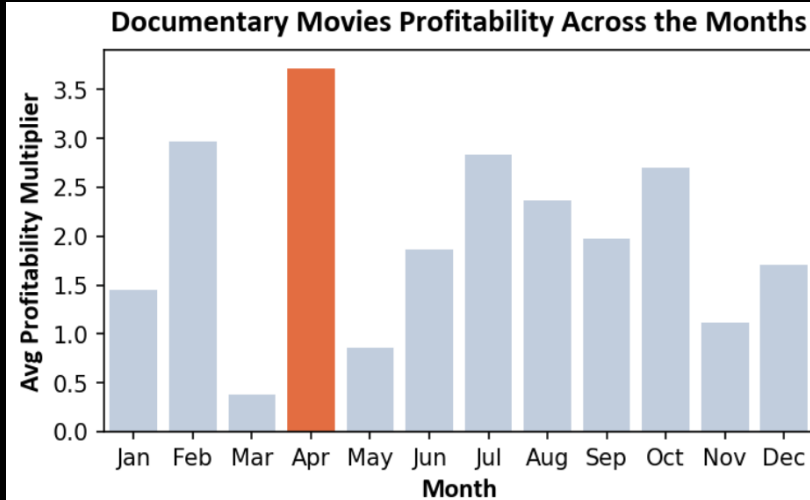


Top 5 Genres

- **Documentary:** Interpreting factual events for education and entertainment
- **News:** Relating to recent events often with drama
- **Biography:** Story centered around an individual and their experience/s
- **Music:** Song or dance interwoven into the narrative
- **History:** Narrative related to past events set in a historical period

Profitability per Month

Recommendation 2



Months impact Profit

- **Documentary**: Release in April
- **Biography**: Release in July
- **Music**: Release in May
- **History**: Release in May

**Profitability Multiplier = Worldwide Gross / Production Budget*

First Movie to Create

Recommendation 3

Confidence Interval for Profitability Multiplier	
Lower Bound	5.14
Upper Bound	5.26

Confidence Interval for Average Rating	
Lower Bound	7.02
Upper Bound	7.26

First Movie Created: Documentary

- Likely to be highly profitable and highly rated

Next Steps



1 Look to Implement

Bucket movies based on production budget & gross
Sentiment Analysis on individual reviews
Machine learning models to predict revenue & profit



2 Find more recent and reliable data

Web Scraping Film-Related Reviews (Twitter, YouTube, Reddit, etc.)
Film Scripts




3 Additional Insights

Fuzzy word matching (by similarity) for merging existing datasets
Process the Rotten Tomatoes datasets and try to find useful information


Thank You!

If you have any questions, thoughts, or would like to follow up, kindly reach out to any of the team below:



Campbell, Hadley
hacampbell@deloitte.com 



Rodriguez, Brian
briarodriguez@deloitte.com 



Stauble, Payton
pstauble@deloitte.com 



Wang, Pengwei
pengwwang@deloitte.com 

