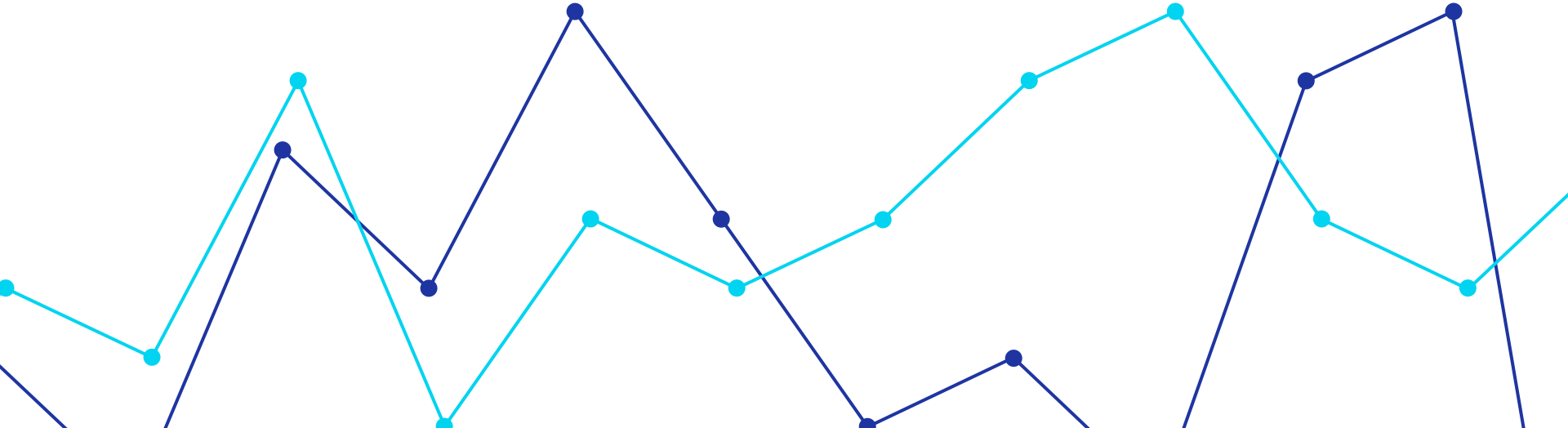
What Makes Movies Successful?

Team Green Bean
Nashville Software School
11/22/2025



Research Project Overview

This is a Movie Data Analysis & Presentation conducted by Grant Alan, Anitha Doddala and Sivaraja Vaithiyalingam.

Goal:

- Use real world movie data to understand what drives financial success.

Approach:

- Gather data from TMDB, Wikipedia and Federal Reserve Bank.
- Clean and Merge collected data.
- Preform EDA and exploratory analysis.
- Create sample statistics.
- Run hypothesis tests.
- Build linear regression models.
- Present Insights clearly.

Data Gathering

Websites and Methods:

- The Movie Database - Access API to collect Top 100 movies from 2014-2025.
- Wikipedia Best Picture Winners - Scrape website for Award Winning films.
- Federal Reserve Bank - Extract Adjusted CPI values to manipulate data.

Data Cleaning Process

Data adjustments:

- Merge dataframes
- Drop duplicate values.
- 2024 CPI adjusted revenue and budget.
- Expand genre variables for analysis.
- Rename data for simplicity and better readability.

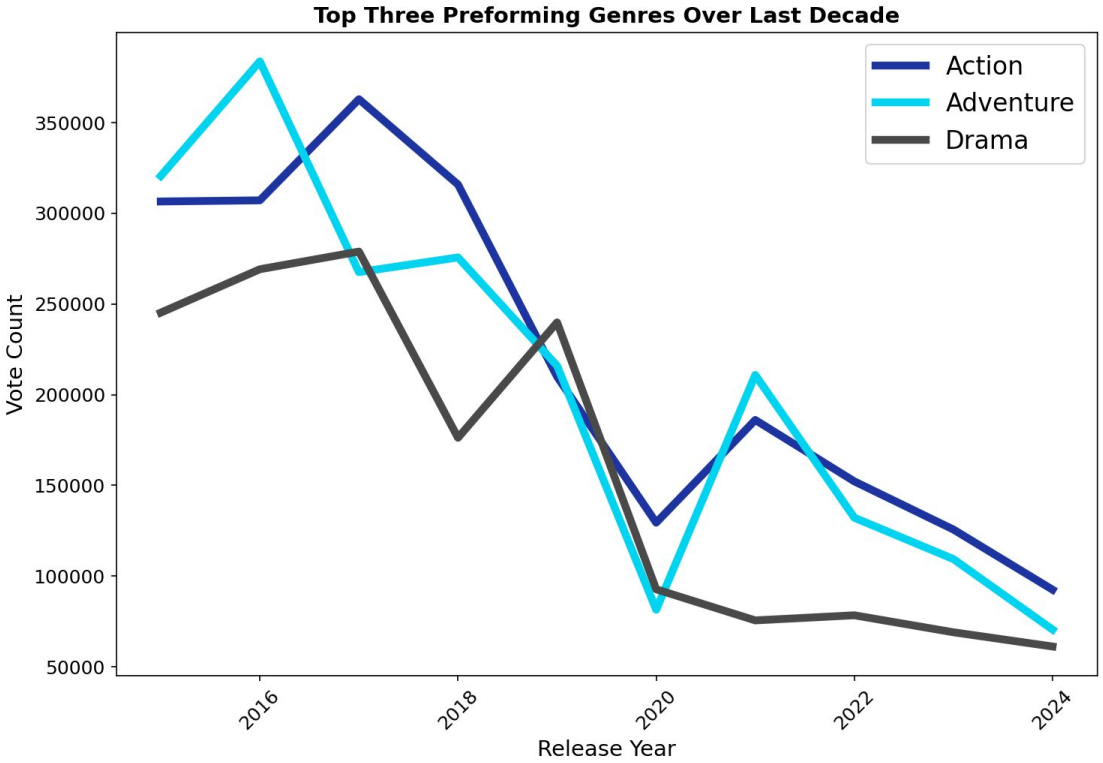
Why?

- We can execute more precise analysis with more data points. More Data = More Precise.
- Duplicate values can throw off our data and give unrealistic estimates.
- The value of the US dollar has changed drastically since this data was collected.
- We want to see the difference in budget and revenue by today's standard.
- When working on a group data project its important to keep names and files cohesive so multiple people can work on the same thing seamlessly. This can save a team a lot of time.

Exploratory Data Analysis

(EDA)

Movie Genre Popularity



Action

Maintains global appeal throughout last decade.

Adventure

This genre has huge spikes due to franchised releases.

Drama

Lower peaks but maintains a more consistent audience.

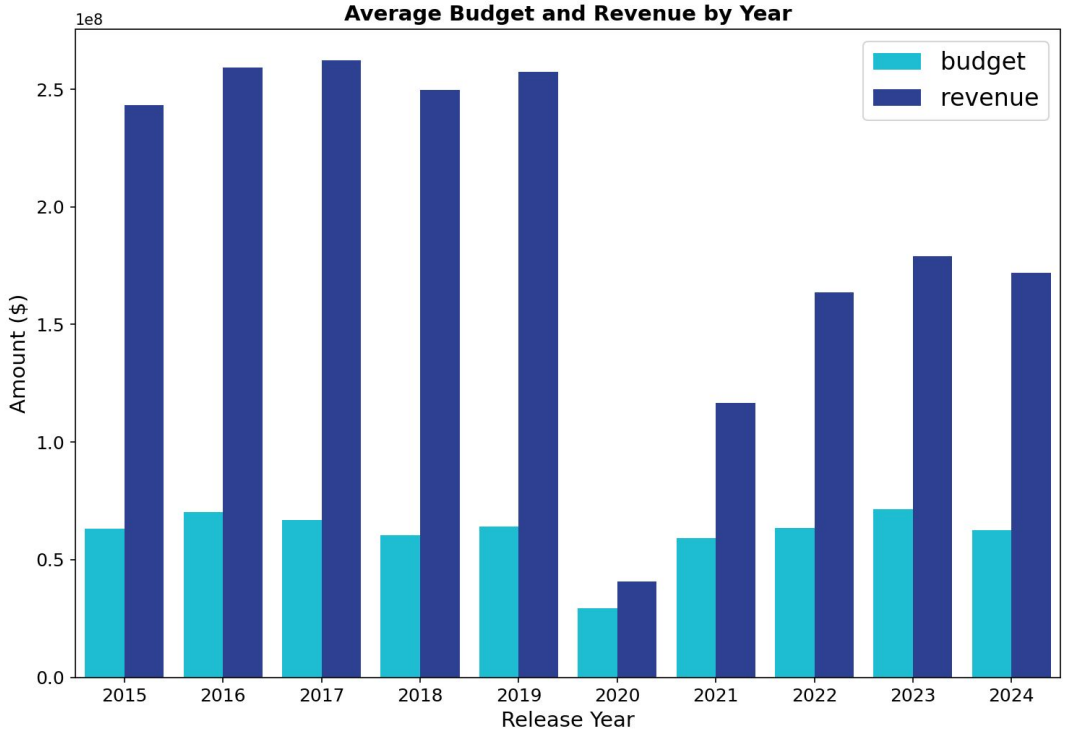
Movie Budget & Revenue by Year

Budget

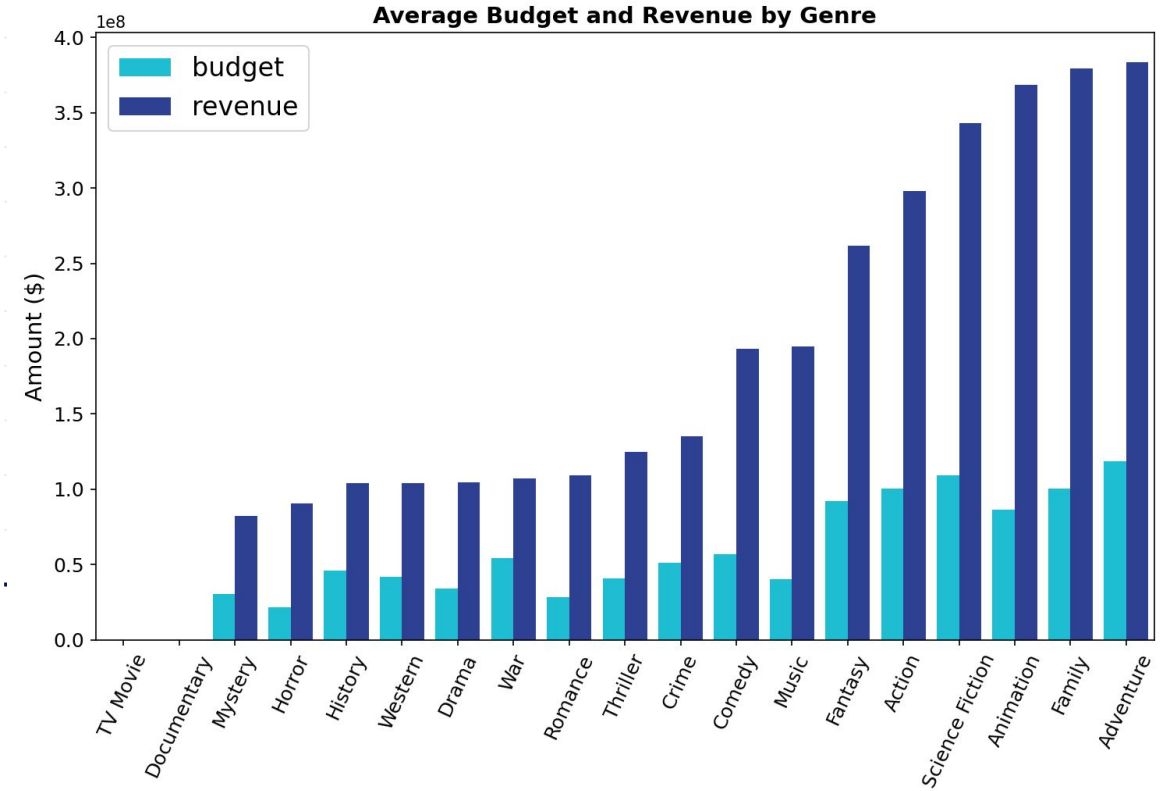
Movies with higher budgets tend to earn higher revenue.

Revenue

Revenues vary, big budgets do not guarantee big success.



Movie Budget & Revenue by Genre



Animation

Often bears lower production costs and is most profitable.

Music

Smaller budget while maintaining strong returns.

Adventure

High budgets and fandom maximizes worldwide appeal.

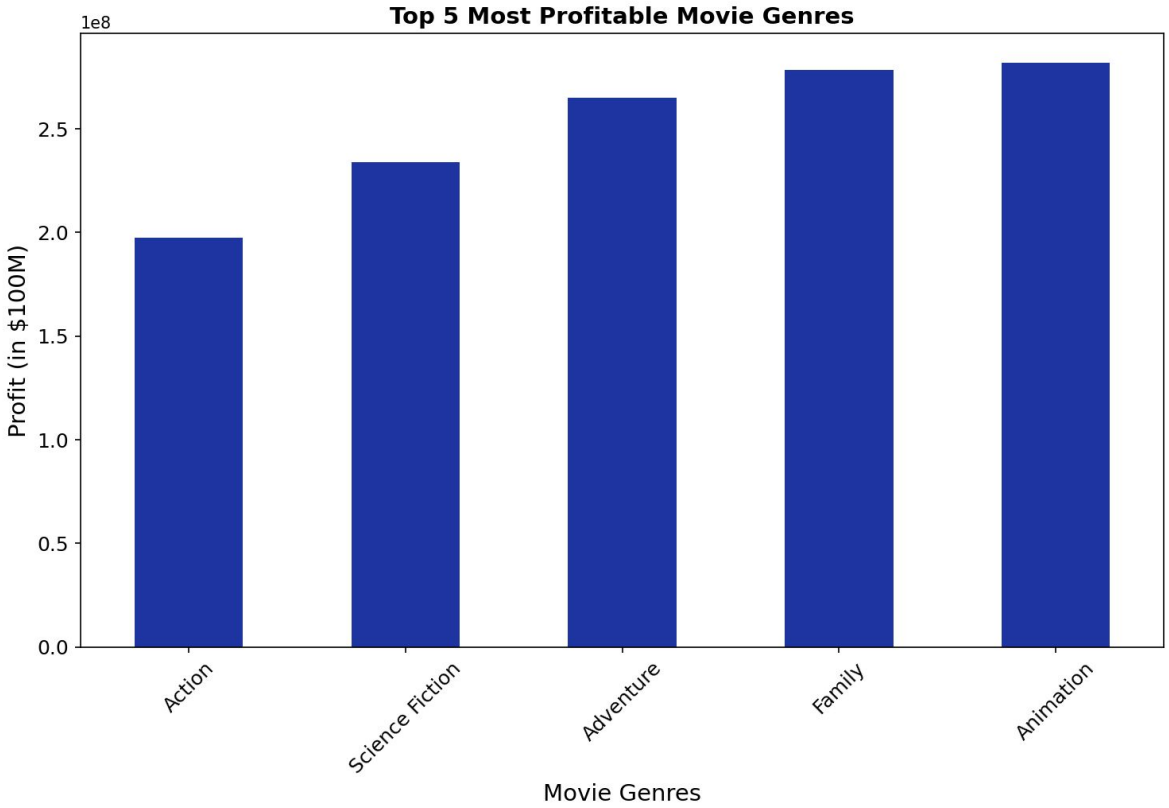
Profitable Genre Highlights

Animation

Lower budgets and Highest overall profit in our dataset.

Family / Adventure

Consistent worldwide appeal to a large demographic.



Performance Comparison & Awards

Revenue Comparison

Winners earn **1.2%** more revenue on average.

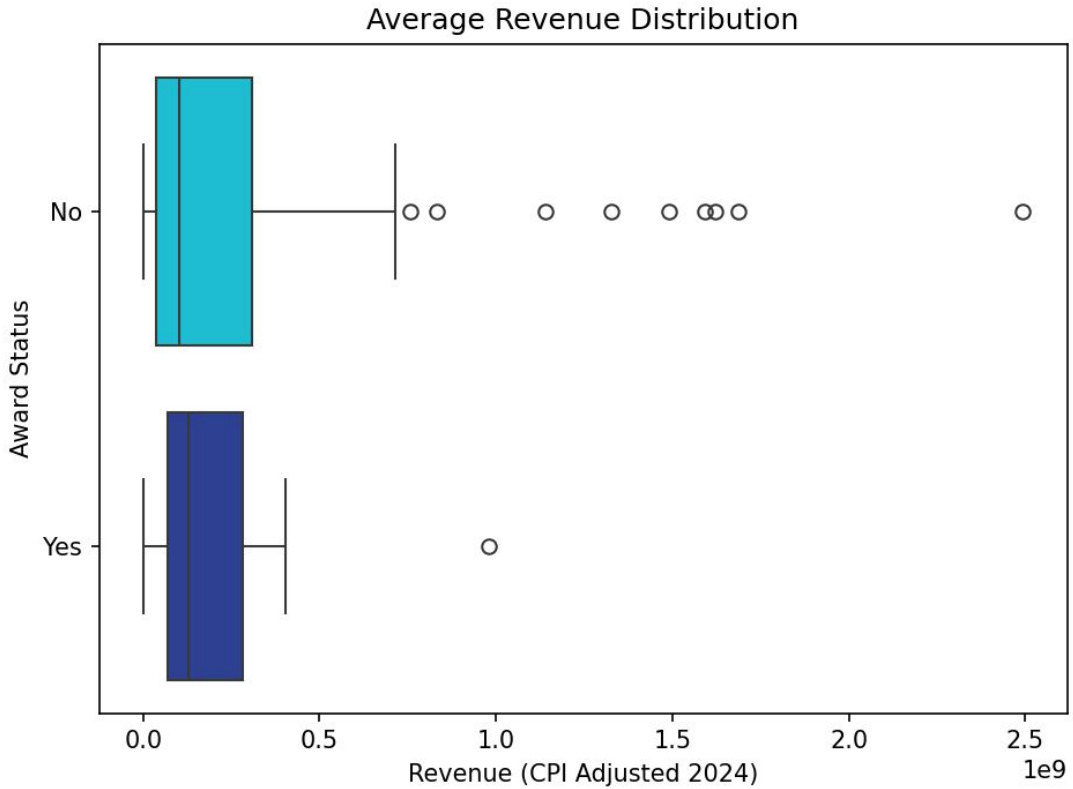
(This observed effect is not statistically significant)

Non-Winners

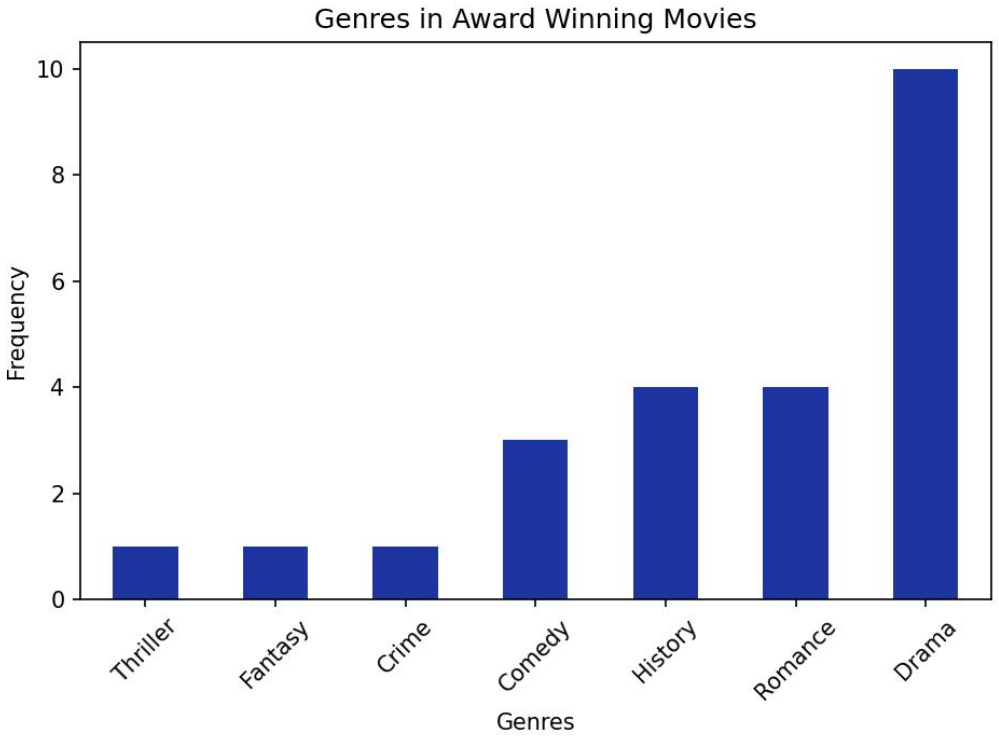
Higher revenue outliers with slightly lower average.

Winners

Consistent distribution of revenue with few outliers.



Genre Trends in Award Winners



Drama

This is the most common genre in award winning movies.

Romance / History

Equal strong appeal but not as common in award winning movies.

Statistical Modeling

Awards & Revenue Hypothesis Test

We observed that award winning movies make 1.2% more revenue on average.

In order to prove that this observed difference is not due to chance we must conduct a hypothesis test. We simulated the variables discovered in our dataset 10,000 times to fabricate this same environment. Then we can check to see how common this difference really is in our generated observations.

Hypothesis Test:

- Null: “Award winning movies and non-award winners have the same average revenue”.
- Alternative: “Award winning movies have higher average revenue”.
- Create a bootstrap resample of our dataset.
- We discover our confidence interval contains zero.
- Fail to reject the null hypothesis.

Insights:

- There is not statistically significant evidence to reject the null hypothesis.
- This observed difference likely due to random chance.

Linear Regression

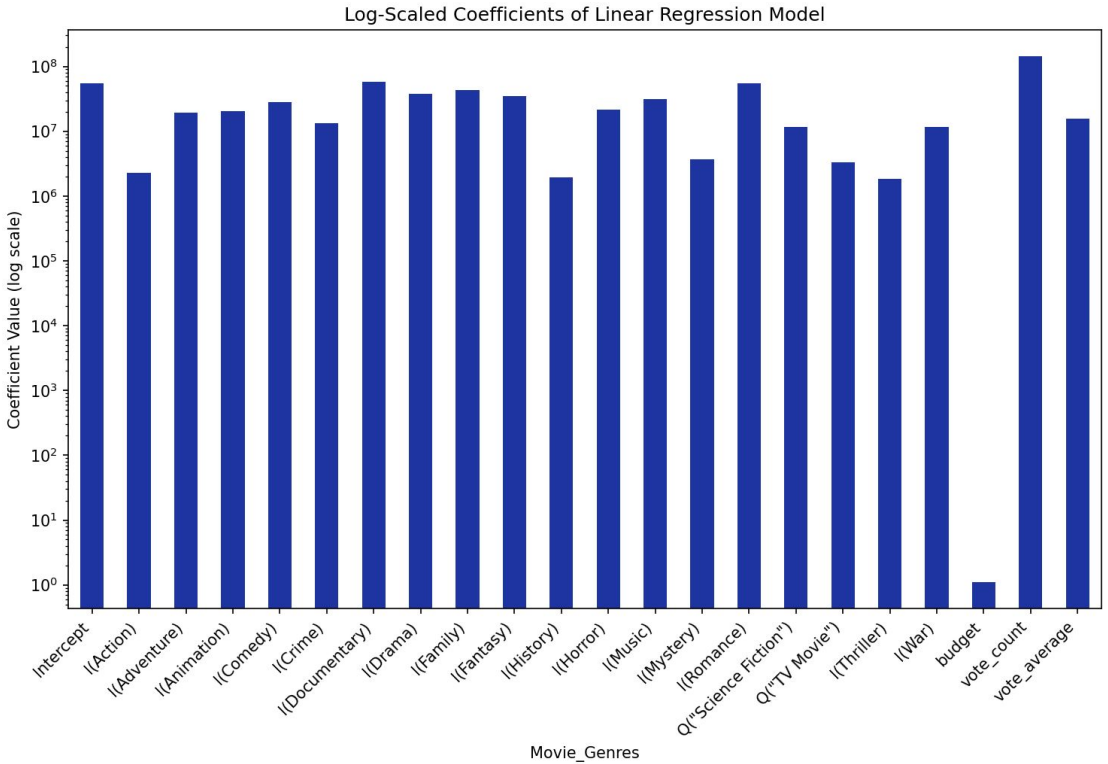
Key Findings from our Linear Regression Model

Profit Drivers

Budget is the strongest predictor of profit. Vote count and average also boost profit.

Genre

Genre coefficients appear large but are not statistically significant.



Thank You

Questions, comments and concerns are appreciated.