

A silhouette of a person stands on a grassy hill, looking out over a vast mountain range under a sunset sky. The scene is framed by a large white L-shaped graphic on the left and right sides.

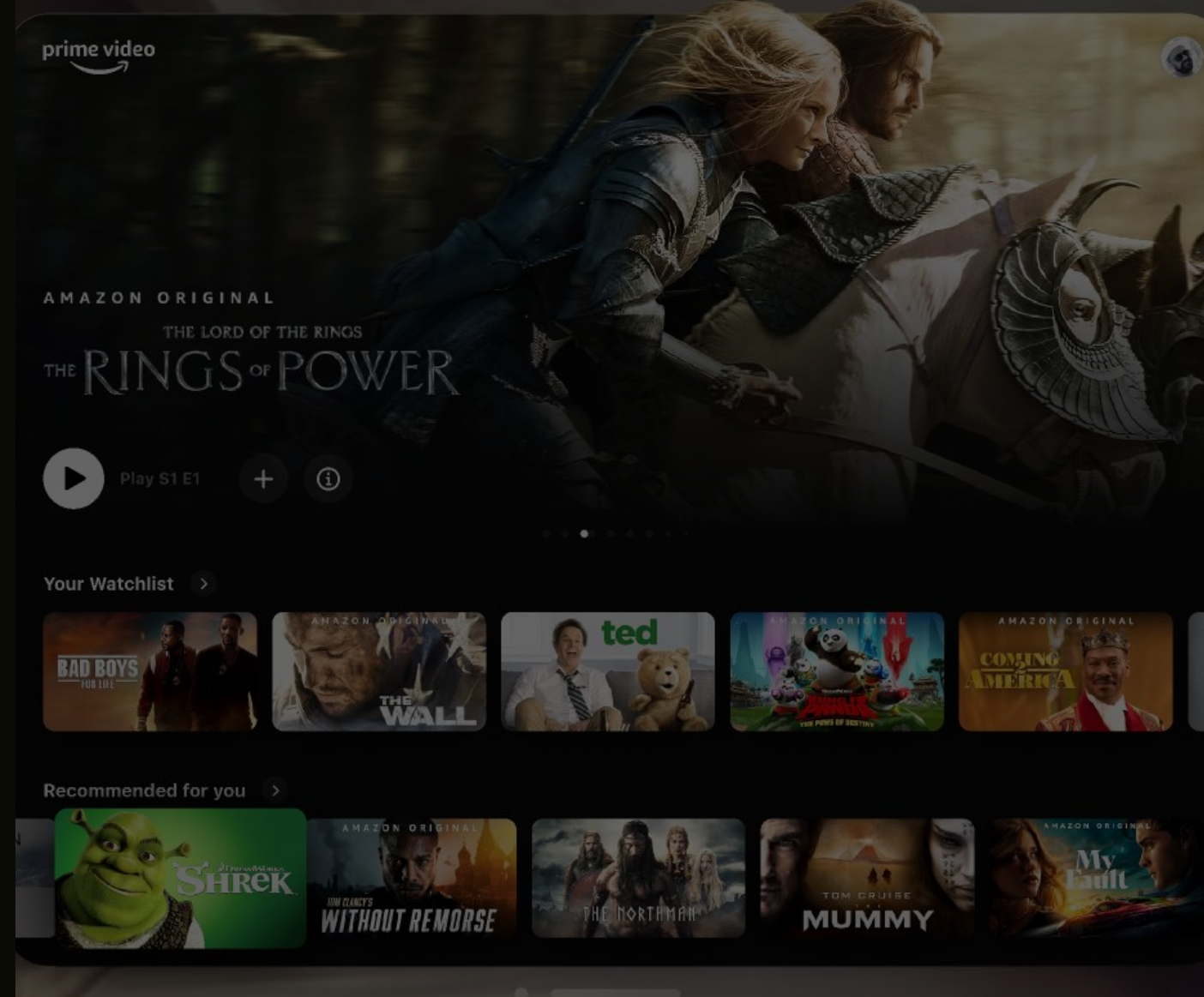
# MOVIE INDUSTRY ANALYSIS

Discovering Data Driven Insights To succeed in the  
Movie Industry

# Overview:

**Goal :** To provide actionable insights for our new movie studio by analyzing current trends in the movie industry.

**Objective:** Identify the most successful movie genres, optimal budget allocations, and best release times for movies.



# Business Understanding:

## Stakeholders:

1. CEO
2. New Head of the Movie Studio
3. Investors and Financiers
4. Data Analysts and Researchers.

## Key Questions:

1. What types of films are performing best at the box office?
2. How do IMDb ratings correlate with box office revenue?
3. When is the best time to release a movie?





# Data Understanding

Data Sources are the IMDB and the BOX Office Mojo

## IMDB :

- The IMDB data is a database that include a number of tables the most relevant tables being:
  - 1. Movie basics
  - 2. Movie ratings
  - 3. Movie akas
- They provide data such as movie ratings, genres, runtime minutes, and number of votes.

## BOX OFFICE MOJO:

- The BOM data is a csv file that has necessary columns on movie data such as:
  - 1. domestic and foreign gross
  - 2. studios producing the movies.



# Data Analysis Methodology

## 1. Technology Used:

- Used Pandas a python library for cleaning and analyzing the data,
- Used Matplotlib and seaborn which are also python libraries for creating visualizations of the data.
- Used sqlite3 to analyze the IMDB database.

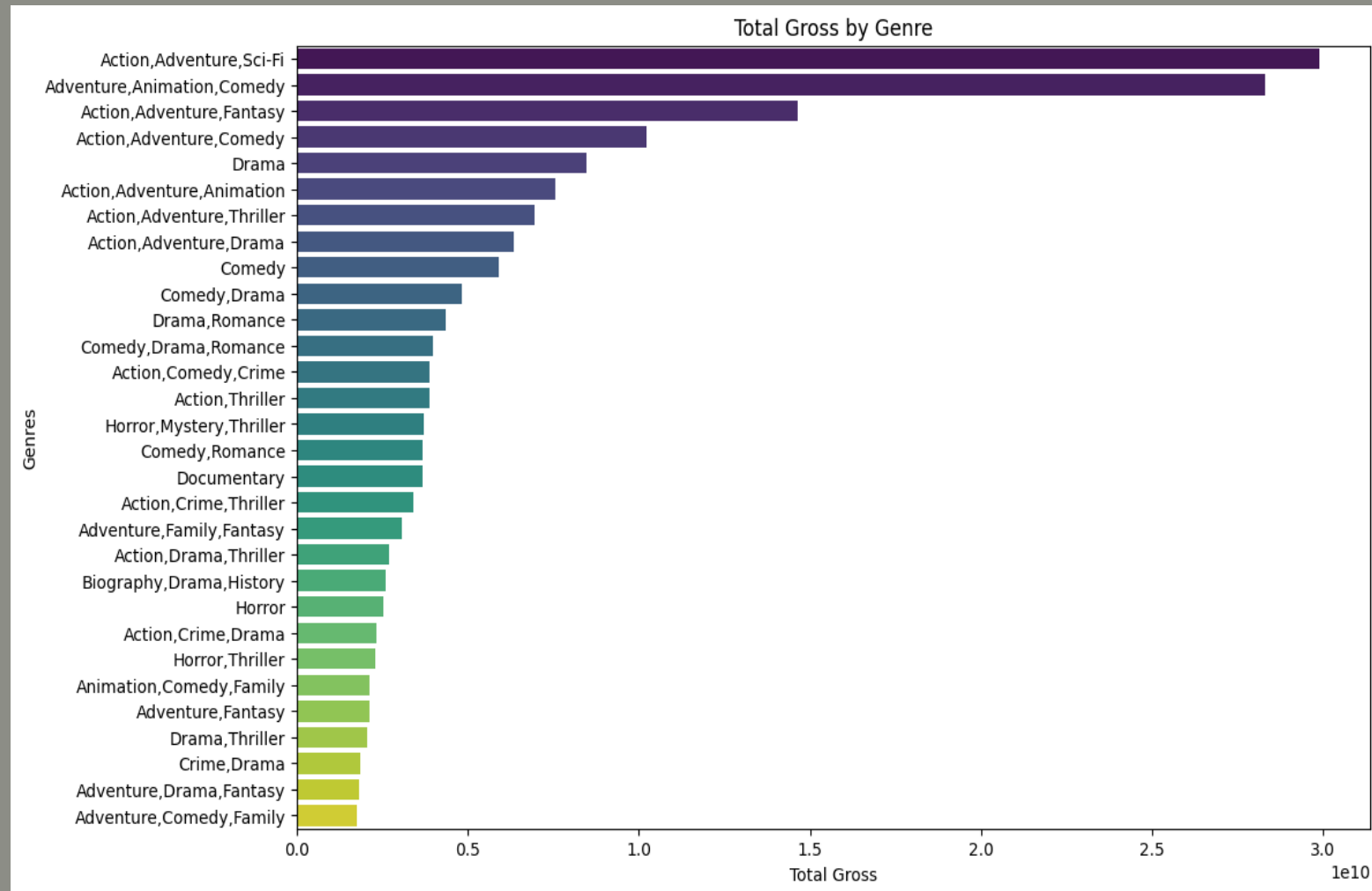
## 2. Data Cleaning:

- Analyzed the data and handled missing values and duplicated values. This also involved extracting only the necessary columns and dropping the unnecessary columns.

## 3. Data Merging:

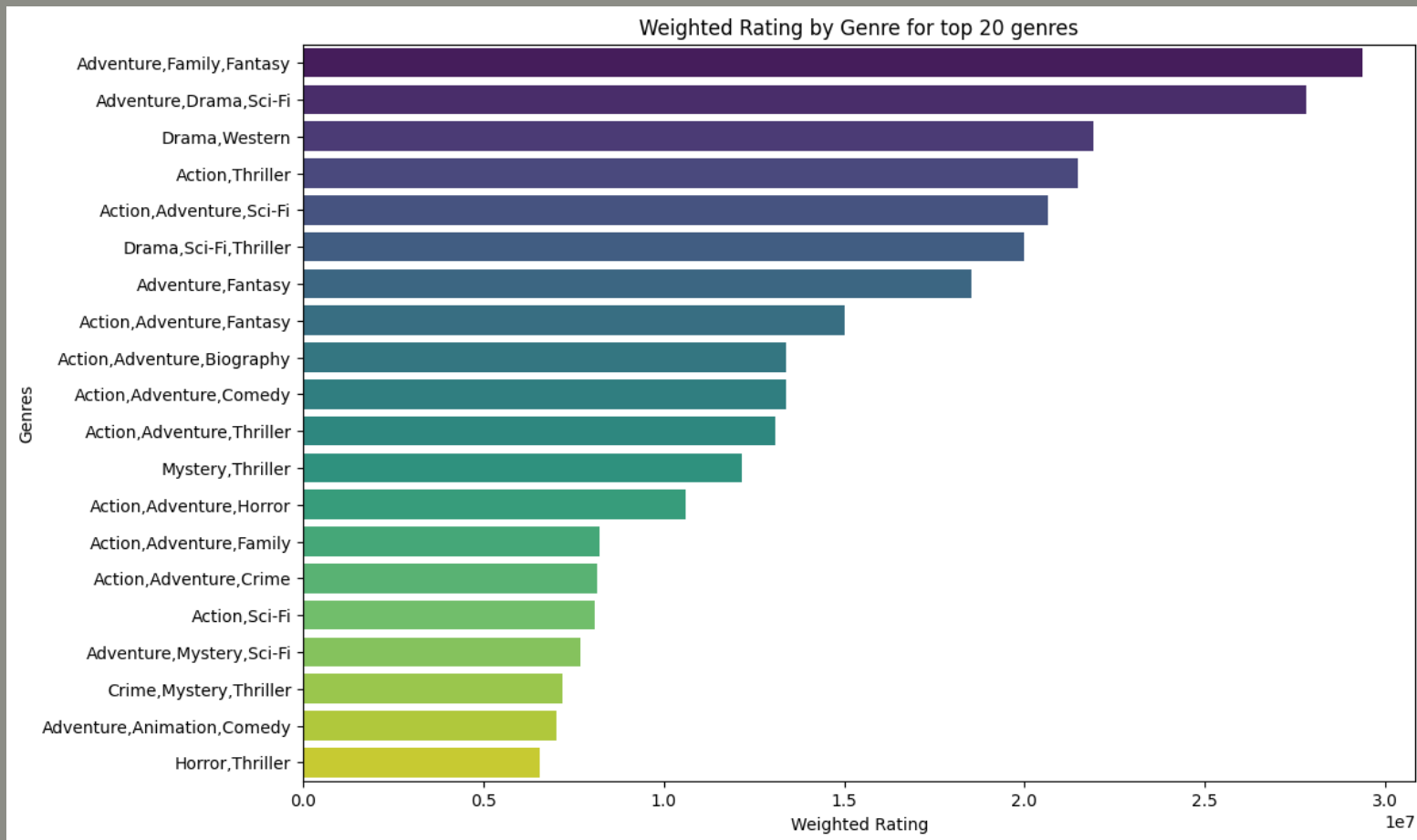
- Merged relevant tables such as the movie basics, movie ratings and movie akas in IMDB with the BOM data.
- Merging made it easier to visualize the data and discover insights of the data.

# Visualization 1: Total Box Office by Genre



- Bar chart showing total box office revenue for each genre.
- **Key Insight:** Genres like Action, Adventure, Sci-Fi, Drama and Fantasy generate the highest revenues.

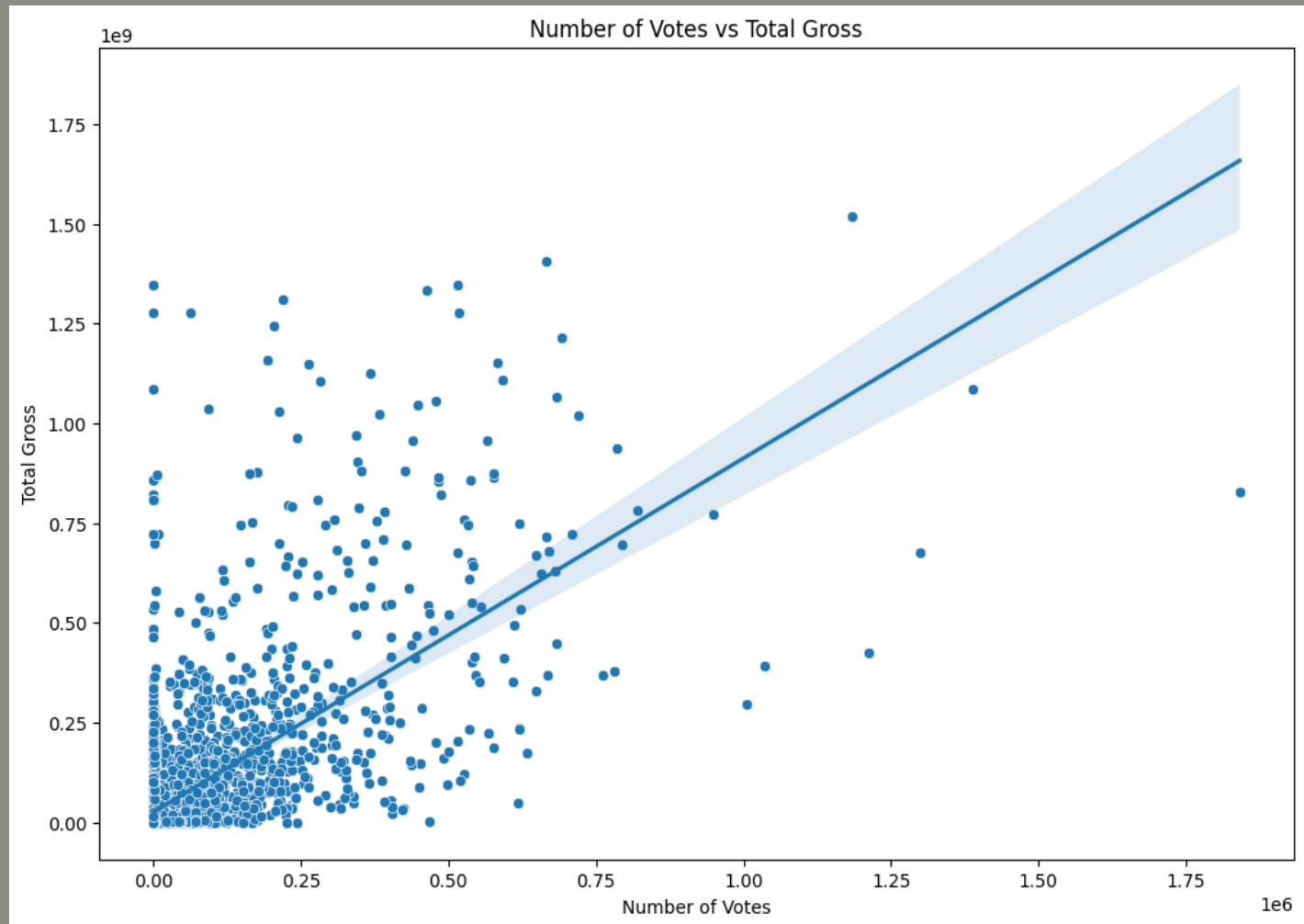
## Visualization 2: Average IMDb Rating by Genre



- Bar chart displaying average IMDb ratings for the 20 genres .
- Key Insight: High ratings are seen in genres like Action, Adventure, Sci-Fi, and Drama.

## Visualization 3: Number of Votes VS Total Gross of the Genres

---



- Scatter plot displaying the relationship between the number of votes and total revenue.
- Key Insight: The number of votes a movie receives often correlates with higher box office revenues.



## Recommendations:

## Next Steps:

### 1. Market Research:

Conduct further research to understand Specific audience preferences within high-performing genres.

### 2. Talent Acquisition:

- Hire experienced directors, writers, and
- actors who excel in the targeted genres.

### 3. Continuous Feedback and Improvement:

- Release films in phases and gather
- Audience feedback to refine future projects.



Source: Business Insider + Statista

Created by genuine impact

More charts: [genuineimpact.substack.com](https://genuineimpact.substack.com)



# Q & A

- Feel free to ask any question



A person wearing a red coat is standing next to a light-colored suitcase. Their hand is resting on a set of keys that are placed on top of the suitcase. The background is a soft-focus outdoor scene. The entire image is overlaid with a dark, semi-transparent layer, and framed by white L-shaped corner brackets on the left and right sides.

# THANK YOU

Presented by: FAITH NASOMBI WAFULA

[faithwafula68@gmail.com](mailto:faithwafula68@gmail.com)

[linkedin.com/in/faith-wafula-145715198](https://www.linkedin.com/in/faith-wafula-145715198)