

# Analyzing Movie Ratings and Trends Using EDA

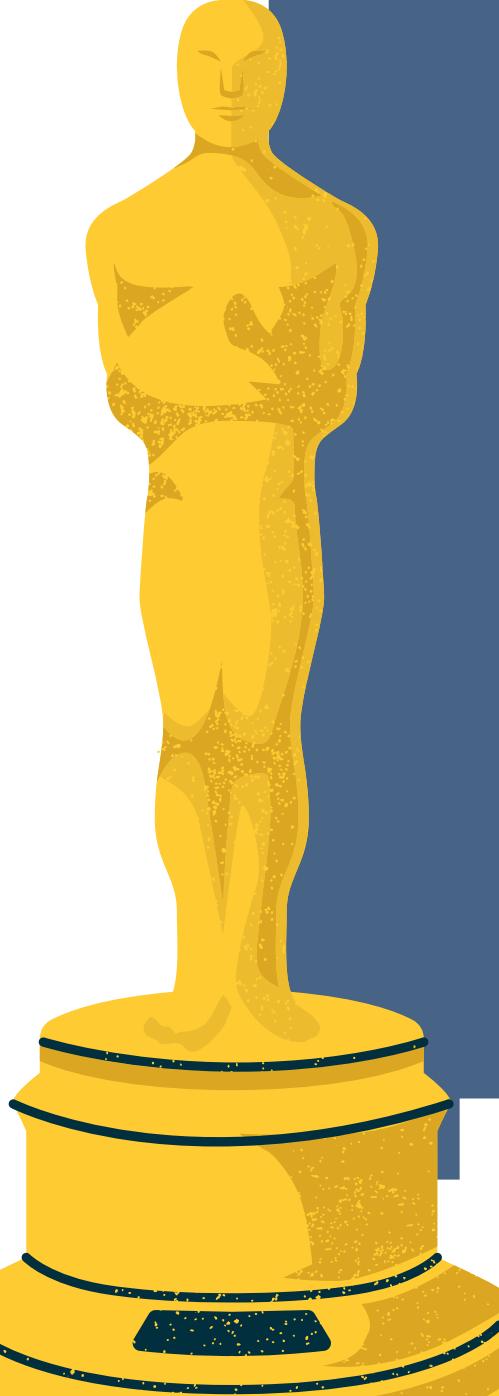
A project presented by Shashikumar.D, submitted towards MCA - 2nd Semester,  
Data Visualization Project (DAT203) for the academic year 2024-2025.

# Introduction

- 1. Video streaming platforms are transforming media consumption.**
- 2. Amazon Prime Video offers over 9,600 titles worldwide.**
- 3. Metadata includes title, director, genre, rating, etc.**
- 4. Analyzing metadata reveals content patterns and trends.**
- 5. Objective: Use EDA to uncover insights and support strategy.**

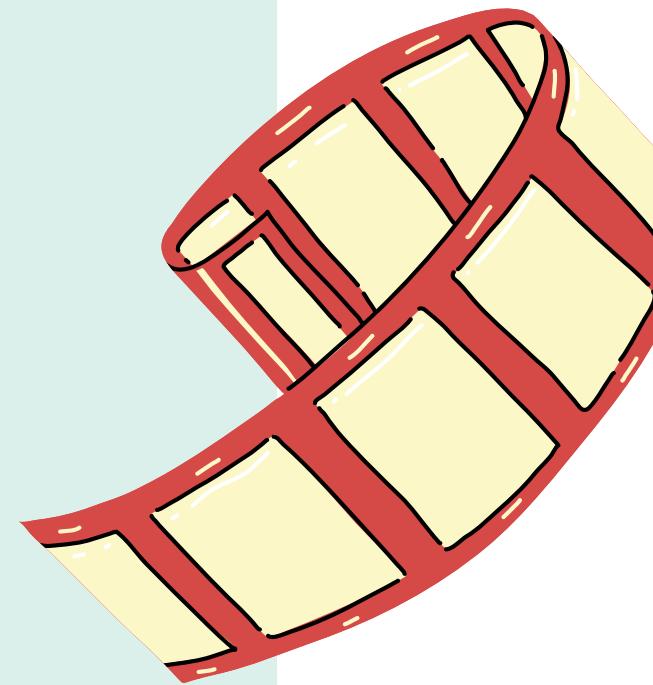
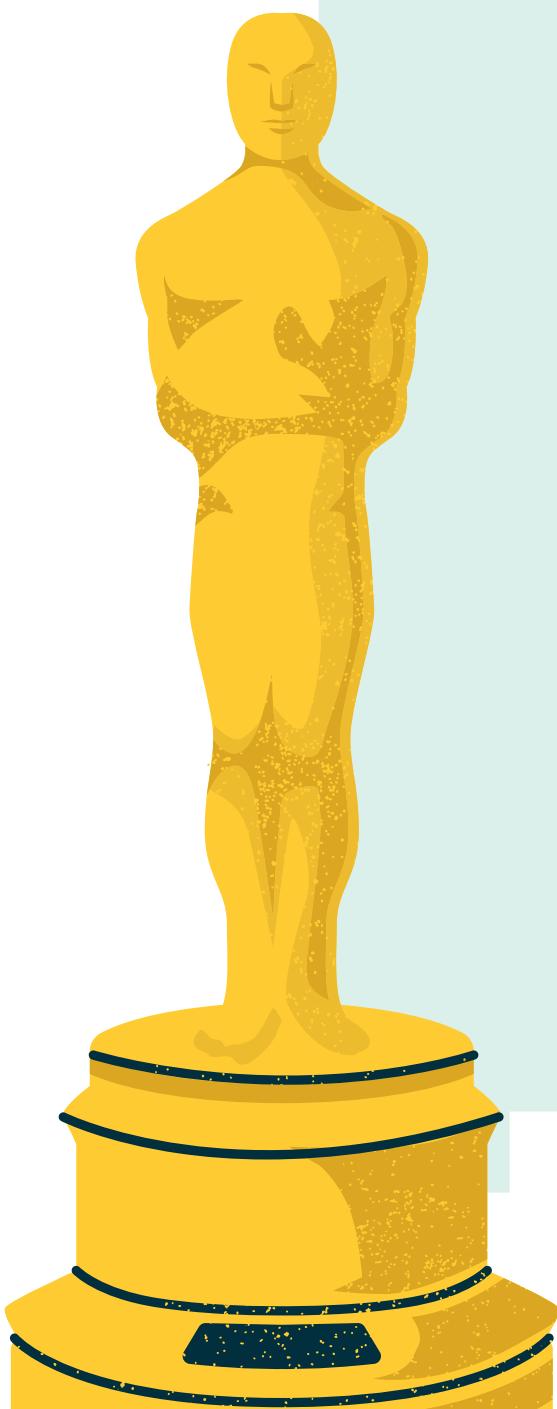


# About the Dataset

- 
- 
- 1. Amazon Prime Titles Dataset includes Movies & TV Shows.**
  - 2. Key columns: title, type, cast, country, release\_year, rating.**
  - 3. Structured data enables temporal, categorical, and text-based analysis.**
  - 4. Challenges: missing data, unstructured text fields.**
  - 5. Used for EDA, NLP, recommendation systems, and more.**

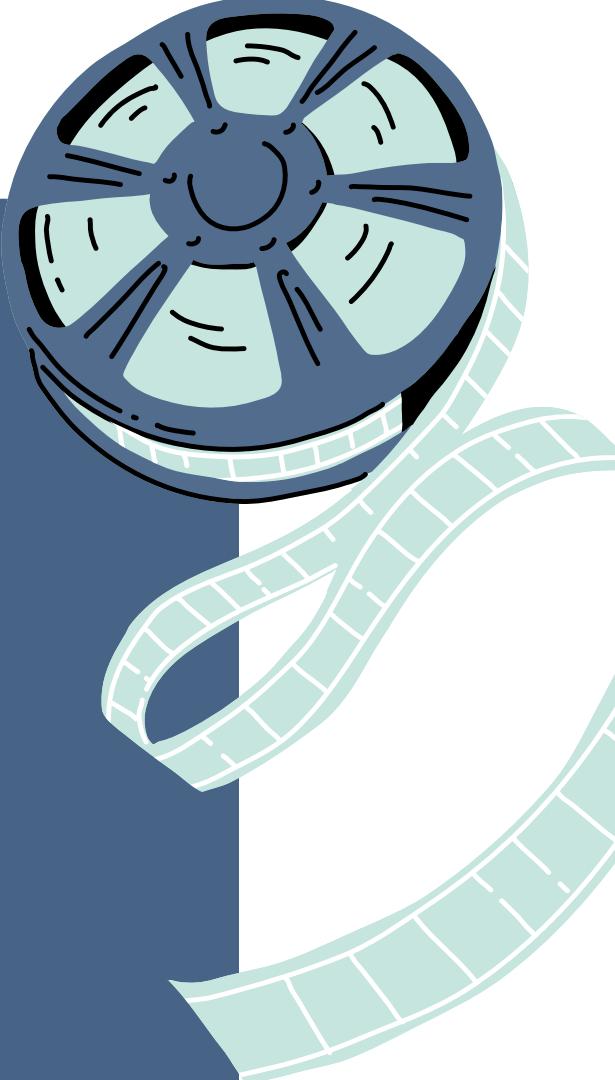
# Problem Statement

- 1. Users struggle to find content matching preferences.**
- 2. Unstructured fields and missing values hinder analysis.**
- 3. No user interaction data; rely solely on metadata.**
- 4. Key questions: content trends, genre dominance, country insights.**
- 5. Goal: Extract patterns to guide recommendations and decisions.**



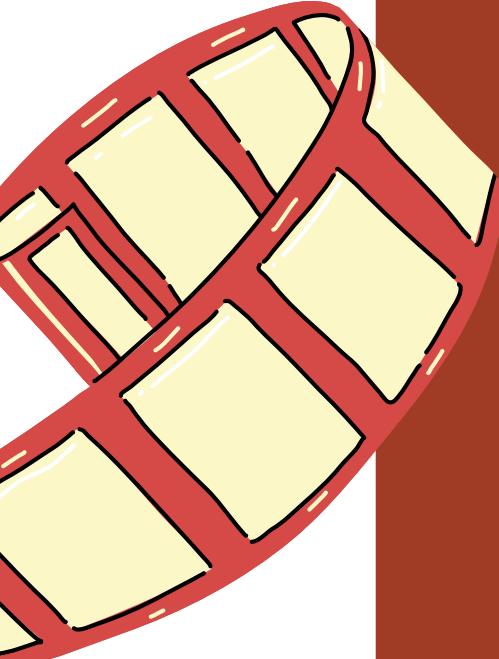
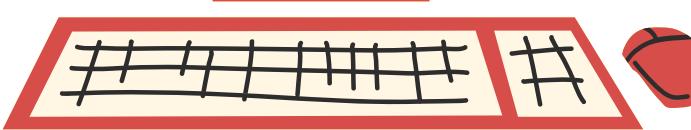


## Methodology - Data Cleaning

- 
- 
- 1. Handled missing values in director, country, cast, etc.**
  - 2. Converted dates and standardized fields.**
  - 3. Extracted new features: title length, word count.**
  - 4. Grouped ratings into Kids vs Adult for segmentation.**
  - 5. Prepared data for visual analysis and insights.**

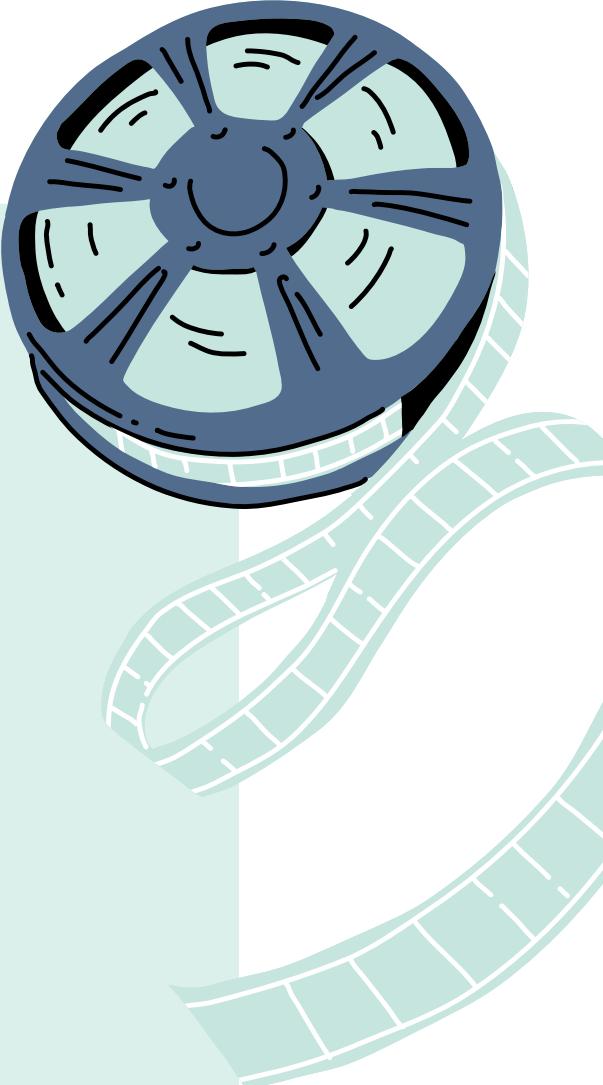
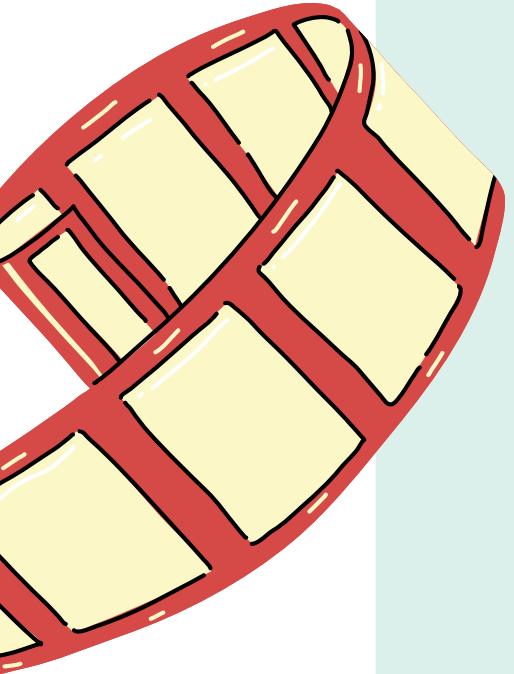
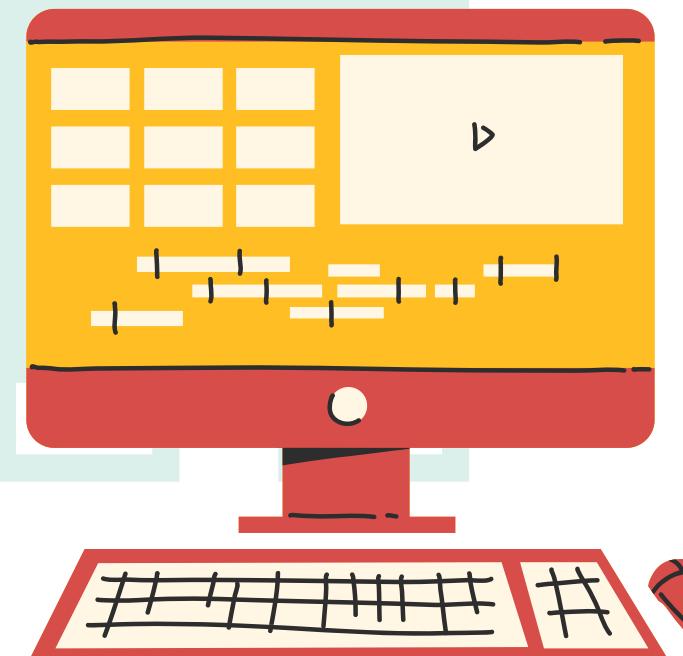


## **Methodology - Metrics & Visualization**

- 
- 
- 
- 
- 1. Analyzed content type (movies vs TV shows).**
  - 2. Top 10 content ratings and genre distribution.**
  - 3. Release trends across years (2000–2023).**
  - 4. Country-wise production volume and duration types.**
  - 5. Visualized using bar charts, pie charts, histograms, heatmaps.**
- 

## Insights - Content & Audience

- 1. Movies dominate over TV shows → focus on one-time viewership.**
- 2. Majority of content is for adult/general audience.**
- 3. TV-MA, R, and TV-14 are most frequent ratings.**
- 4. Need for more family-friendly and children's content.**
- 5. Title naming tends toward short, catchy words.**





## Insights - Genres and Countries

- 
- 
- 1. Drama, Comedy, and Documentary are top genres.**
  - 2. Opportunities in niche content: Sci-Fi, Anime, Regional films.**
  - 3. Most content originates from a few dominant countries.**
  - 4. Potential to diversify international content offerings.**
  - 5. Genre analysis helps target specific audience segments.**

## Insights - Trends Over Time

- 1. Content surged during 2016-2020, especially 2018-2020.**
- 2. Aligns with Amazon's expansion in originals and global licensing.**
- 3. Possible plateau or drop post-2021 due to market saturation.**
- 4. Temporal trends aid in forecasting content planning.**
- 5. Seasonal or year-based additions can be identified.**

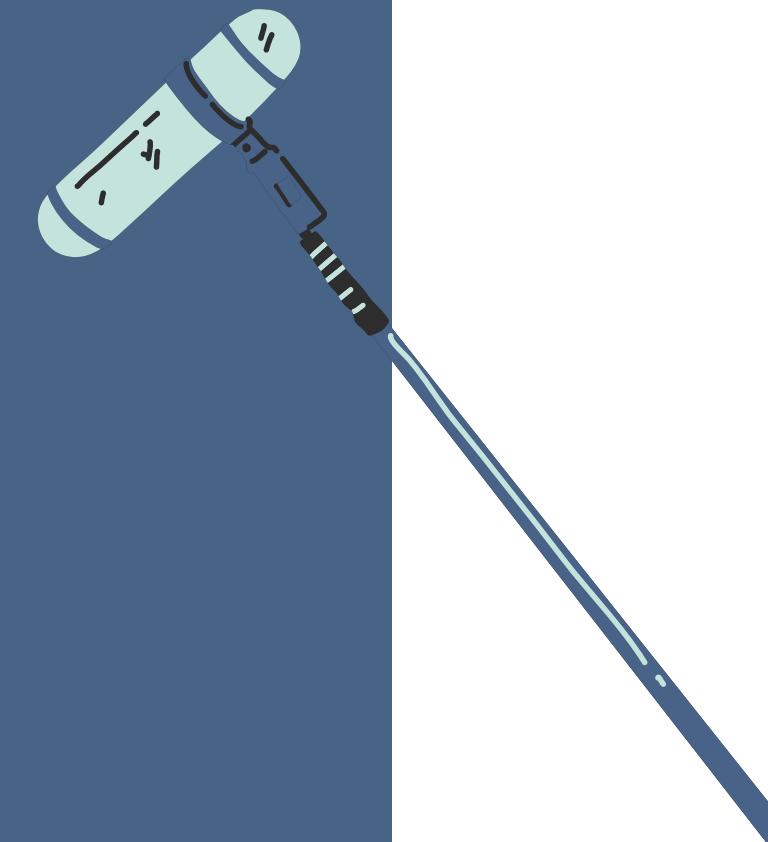
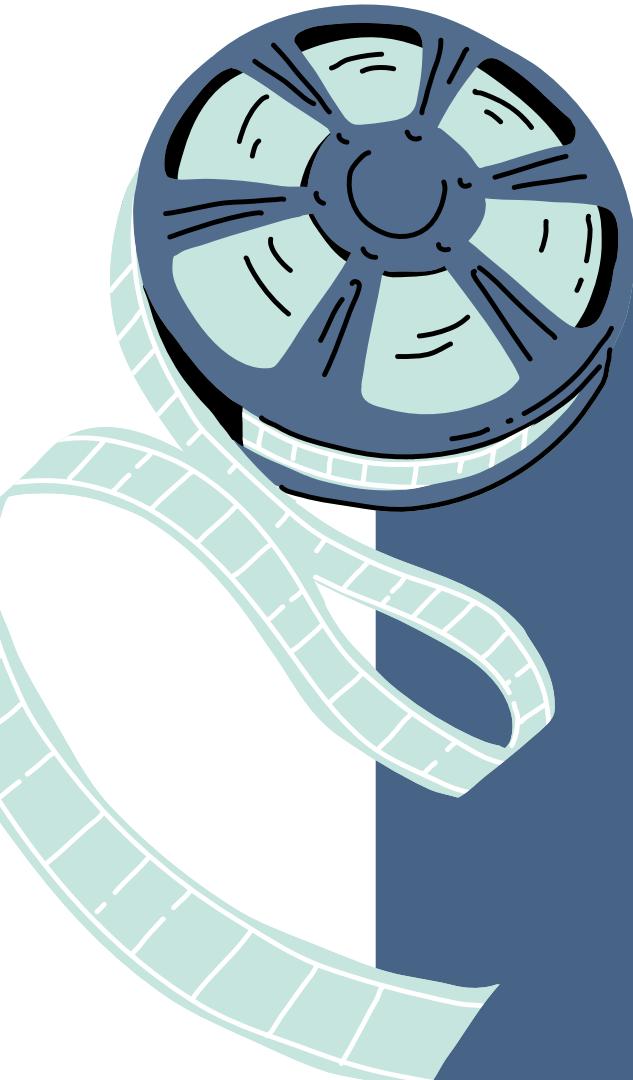


## Business Implications

- 
- 
- 1. Invest in original TV shows to increase viewer retention.**
  - 2. Add kid-safe content to improve brand inclusivity.**
  - 3. Target regional and international markets with tailored content.**
  - 4. Use metadata-driven insights to guide acquisition and promotion.**
  - 5. Support ML applications like recommendation systems and segmentation.**

## Conclusion

- 1. Amazon Prime is movie-focused with adult-targeted content.**
- 2. Drama and Comedy dominate; scope for niche genre expansion.**
- 3. Content growth peaked around 2018-2020.**
- 4. Short titles and detailed descriptions reflect marketing strategy.**
- 5. Future focus: expand TV shows, kid-friendly content, global reach.**



# THANK YOU!

Presented By :

Shashikumar.D

Shashi Kumar