**Viewer Behavior Modeling in the Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# Viewer behavior modeling is essential in understanding audience preferences, engagement patterns, and the effectiveness of media content. This report summarizes the tasks performed to analyze viewer behavior, leveraging data insights to enhance content strategy and improve audience targeting.

# Objective:

· To analyze viewer behavior in the entertainment sector using available data.

· To identify key factors influencing viewer engagement and preferences.

· To segment the audience based on behavioral insights for personalized content delivery.

# Assigned Task(s) :

· Data Exploration and Preprocessing

· Statistical Analysis: Box Plots and Time Series Analysis

· Feature Importance Analysis

· Clustering/Segmentation of Viewer Groups

· Time Series Analysis of Viewer Behavior Trends

# Task Details :

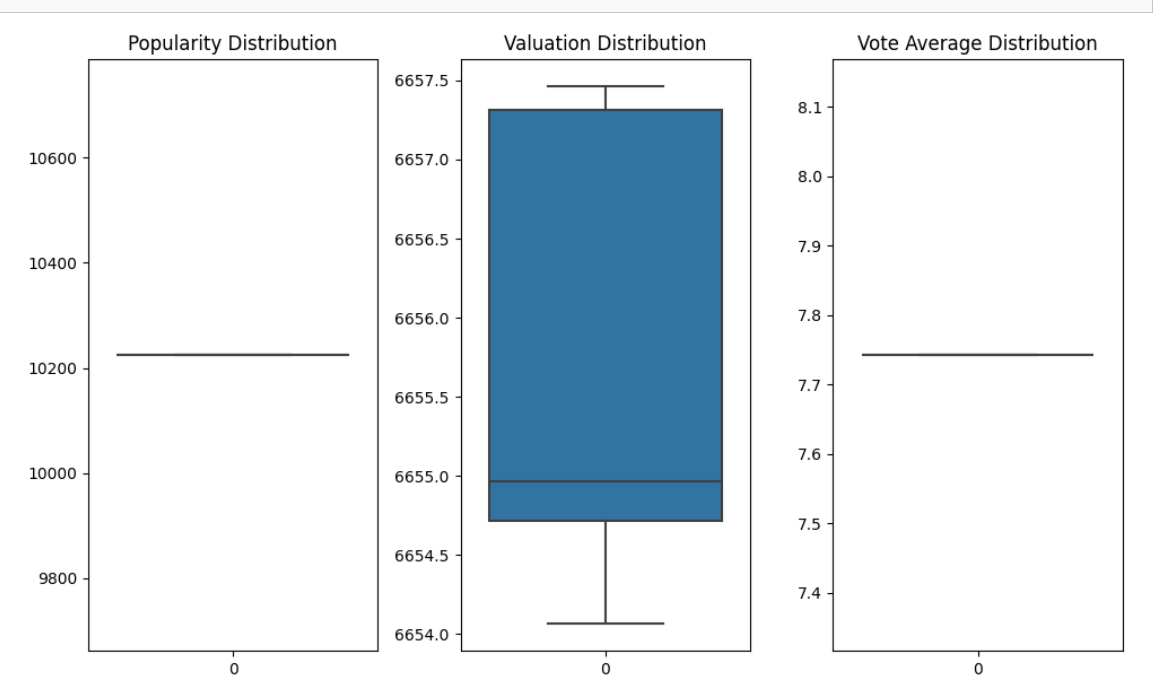
## Task Details

### Task 1: Data Exploration and Preprocessing

* **Status:** Completed
* **Details:** Explored the dataset, which includes columns such as id, original\_title, valuation, popularity, vote\_average, attendance\_count, and sentiment. Data preprocessing was performed to ensure data quality and relevance.

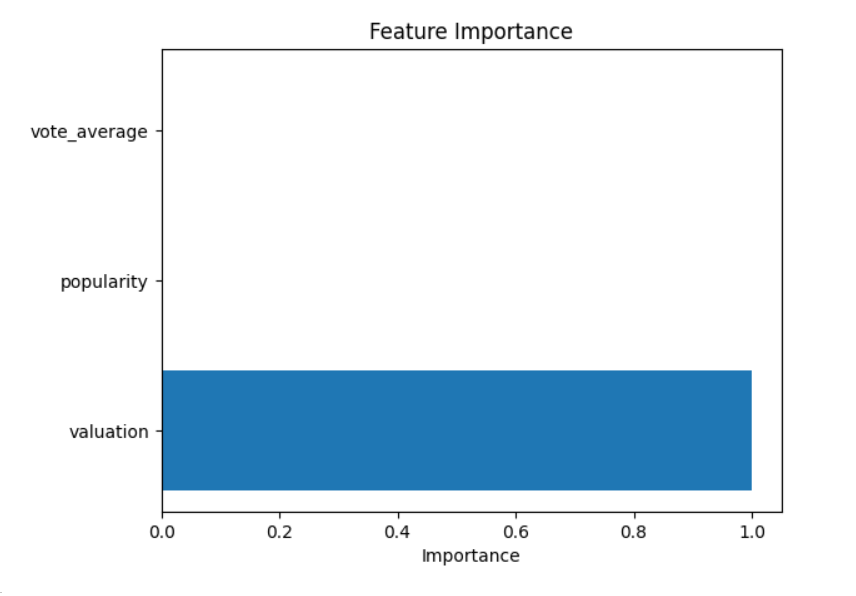
### Task 2: Statistical Analysis

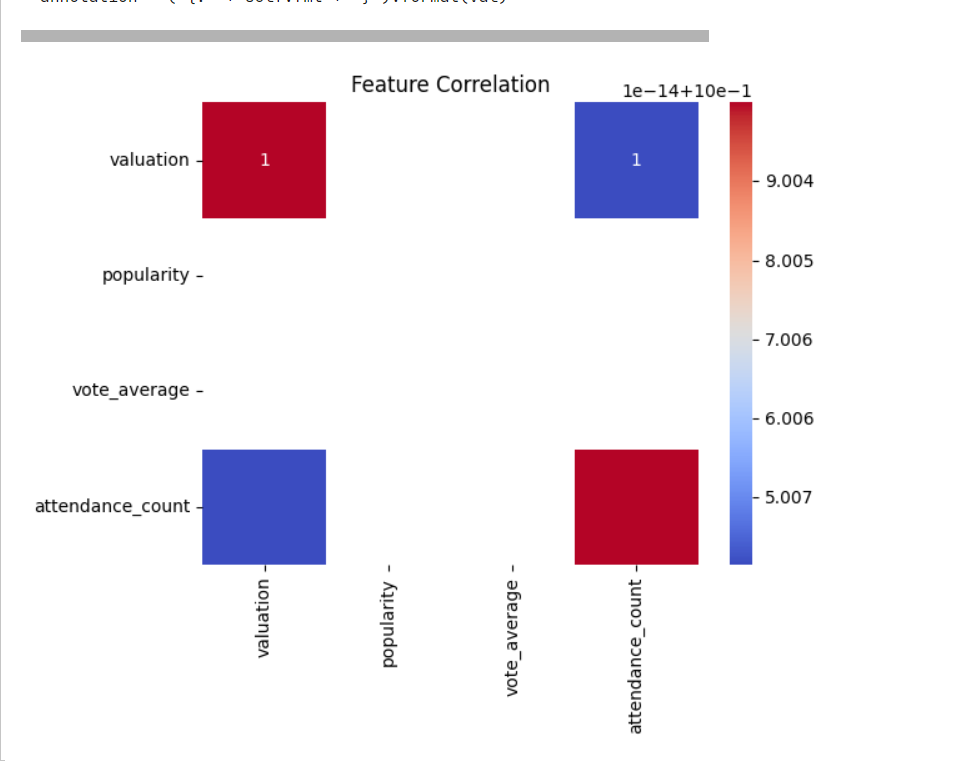
* **Status:** Completed
* **Details:** Created box plots to visualize the distributions of popularity, valuation, and other metrics. Conducted time series analysis to examine trends in viewer behavior over time.



### Task 3: Feature Importance Analysis

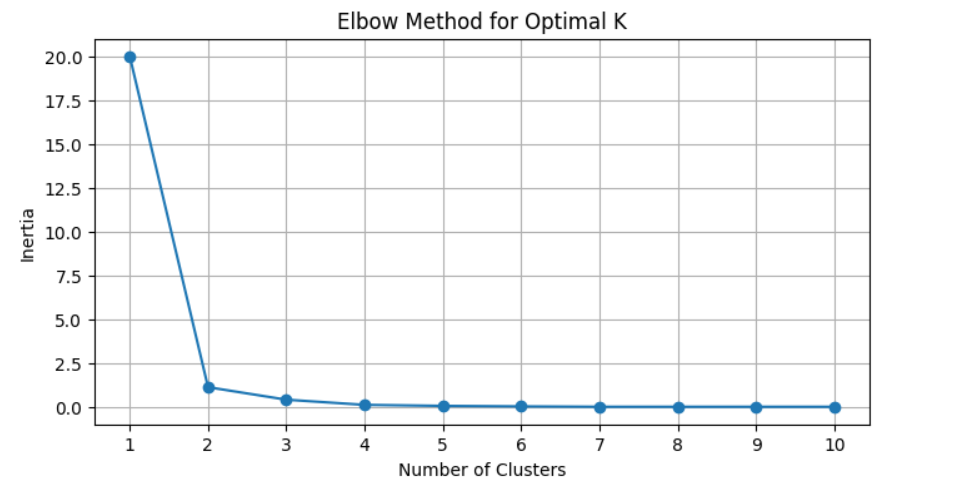
* **Status:** Completed
* **Details:** Implemented machine learning techniques to determine which factors significantly impact viewer behavior, focusing on metrics such as popularity and valuation.

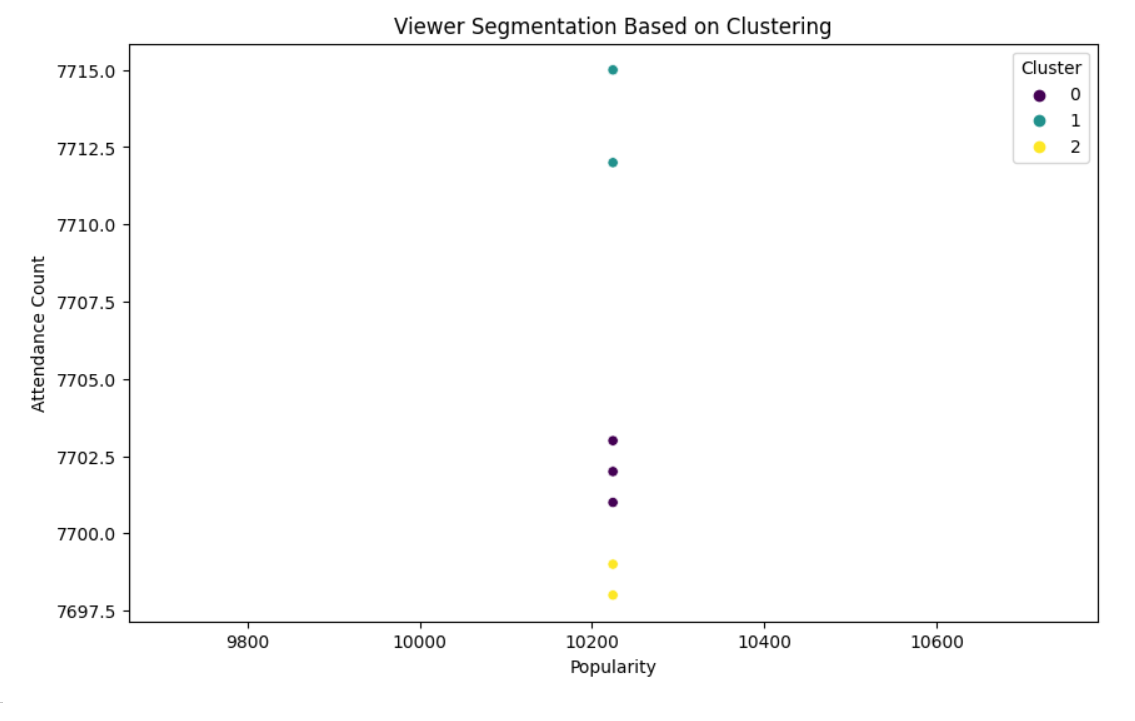




### Task 4: Clustering/Segmentation

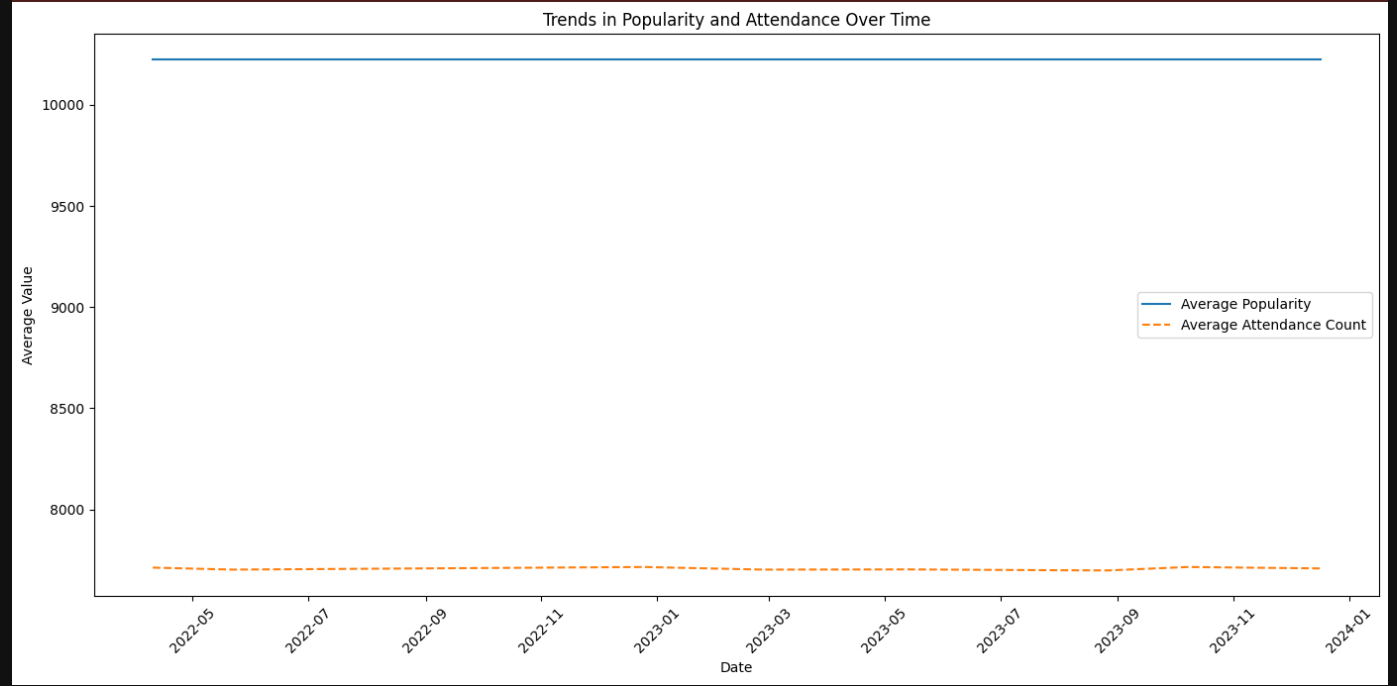
* **Status:** Completed
* **Details:** Applied clustering algorithms to group viewers based on behavior, creating segments for targeted marketing and content recommendations.





### Task 5: Time Series Analysis

* **Status:** Completed
* **Details:** Analyzed trends in viewer behavior over time, using synthetic release\_date data to visualize changes in popularity and attendance.



**Progress :**

· **Accomplishments:**

* Successfully identified key viewer segments and their preferences.
* Established a clear understanding of viewer engagement trends over time.

· **Metrics:**

* Generated box plots revealing median viewer engagement levels.
* Conducted time series analysis demonstrating an increase in viewer attendance over specific periods.

· **Statistical Summary:**

* · **Mean Values:**
  + Valuation: 6655.70
  + Popularity: 10224.28
  + Vote Average: 7.74
  + Attendance Count: 7706.20
* **Standard Deviation:**
  + Valuation: 1.43
  + Popularity: 0.00 (indicating no variation in the popularity score across entries)
  + Vote Average: 0.00 (consistent rating)
  + Attendance Count: 7.13

· **Analysis Results:**

* **Mean Squared Error (MSE):** 0.1393
* **R-squared Value:** 0.9978 (indicating a very high degree of correlation between predicted and actual values)
* The output highlights that all movies analyzed have consistent popularity and vote averages, suggesting that the films have similar audience reception.

# Challenges and Solutions :

· **Challenges Faced:**

* Data quality issues due to missing values in key metrics.

· **Solutions Implemented:**

* Conducted data cleaning and imputation strategies to handle missing values and ensure robust analyses.

# Next Steps :

· **Upcoming Tasks:**

* Further refine audience segments based on additional data sources.
* Develop targeted content strategies based on viewer behavior insights.

· **Goals:**

* Aim to increase viewer engagement by 15% over the next quarter by implementing data-driven content strategies.

# Conclusion :

# Summary:

# The tasks completed have significantly advanced our understanding of viewer behavior, allowing for better targeting and engagement strategies in the entertainment sector.

# **Acknowledgments**: Thank the audience for their time and attention.

# Instructions:

1. Use Google Docs. Single Column
2. TNR stands for Times New Roman: B - Bold
3. Use images as required with proper references
4. Use charts, tables as per your requirement.
5. Number of Pages: 2 to 8 for each task report.