**On Demand Content Performance Analysis - Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# The On Demand Content Performance Analysis for the Entertainment Sector focuses on evaluating key performance metrics such as attendance count, popularity, rating, and valuation. The goal of this analysis is to gain insights into how on-demand content performs over time and develop a predictive model to estimate future performance based on available data.

# Objective:

The primary goal of today's tasks was to perform a time series analysis and predictive modeling to analyze content performance over time and predict future attendance using key factors such as popularity, valuation, and ratings.

# Assigned Task(s) :

* Add a release\_date column to the dataset for time series analysis.
* Conduct time series analysis to visualize trends in attendance count and popularity over time.
* Build a predictive model using Linear Regression to forecast attendance count based on available features.

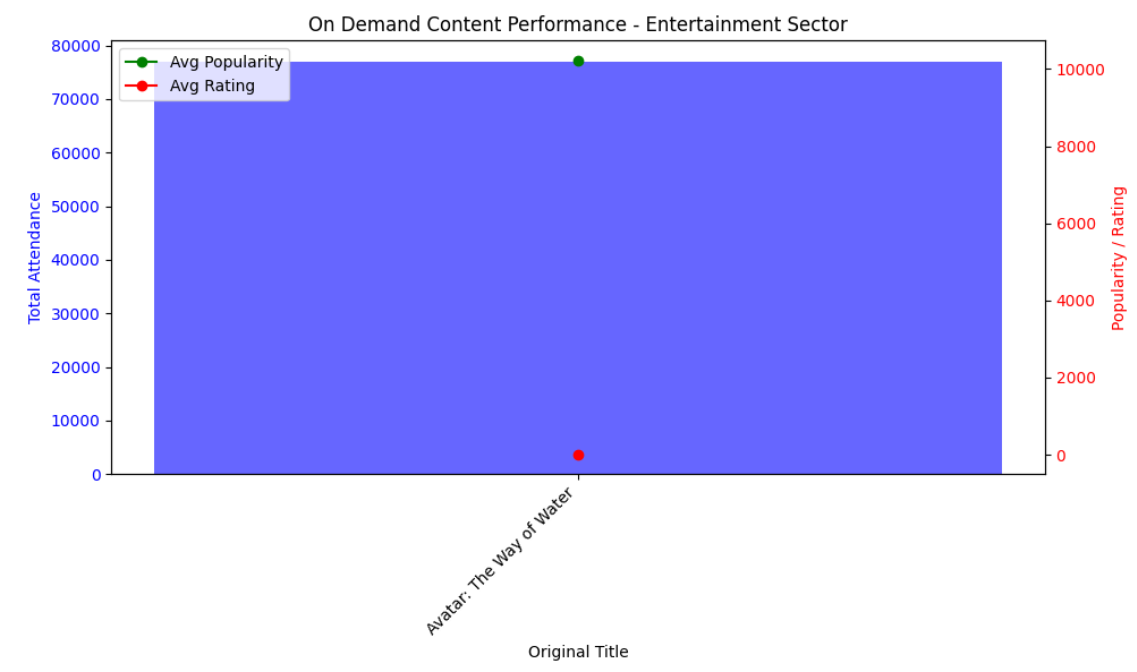
# Task Details :

#### Task 1: ****Add Release Date for Time Series Analysis****

* **Status**: Completed
* **Details**: A release\_date column was added to the dataset. Synthetic dates were generated, assuming each content was released weekly, starting from January 1, 2020. This column is essential for time series analysis, enabling us to observe how key metrics evolve over time.

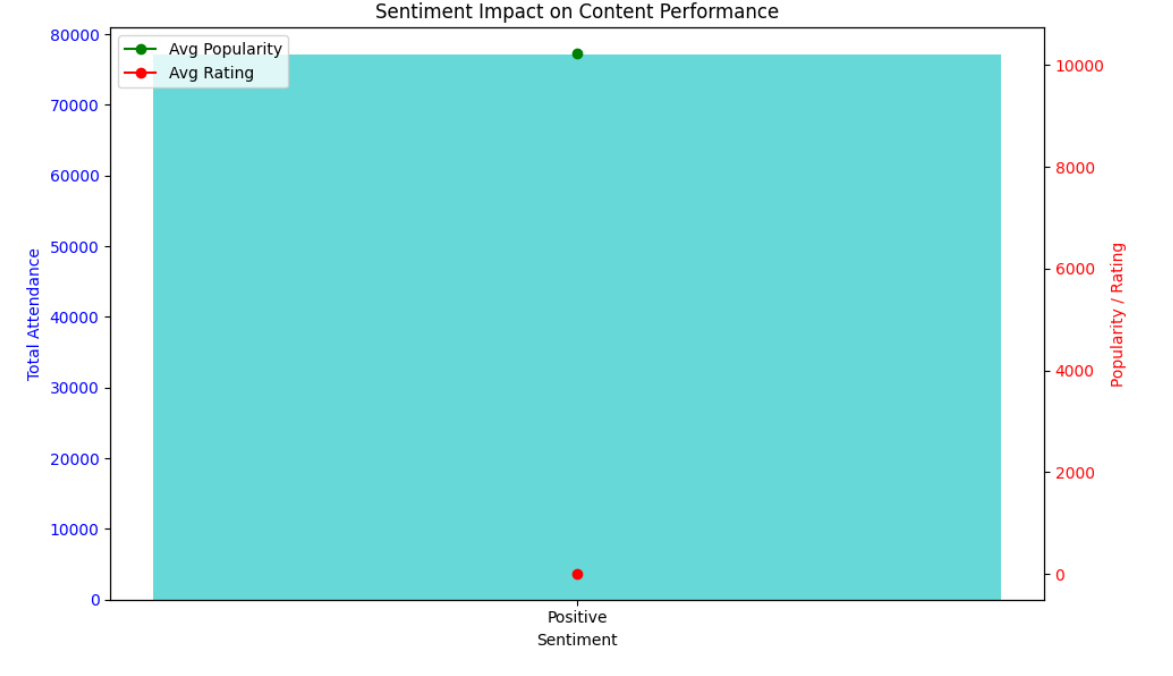
#### Task 2: ****Time Series Analysis****

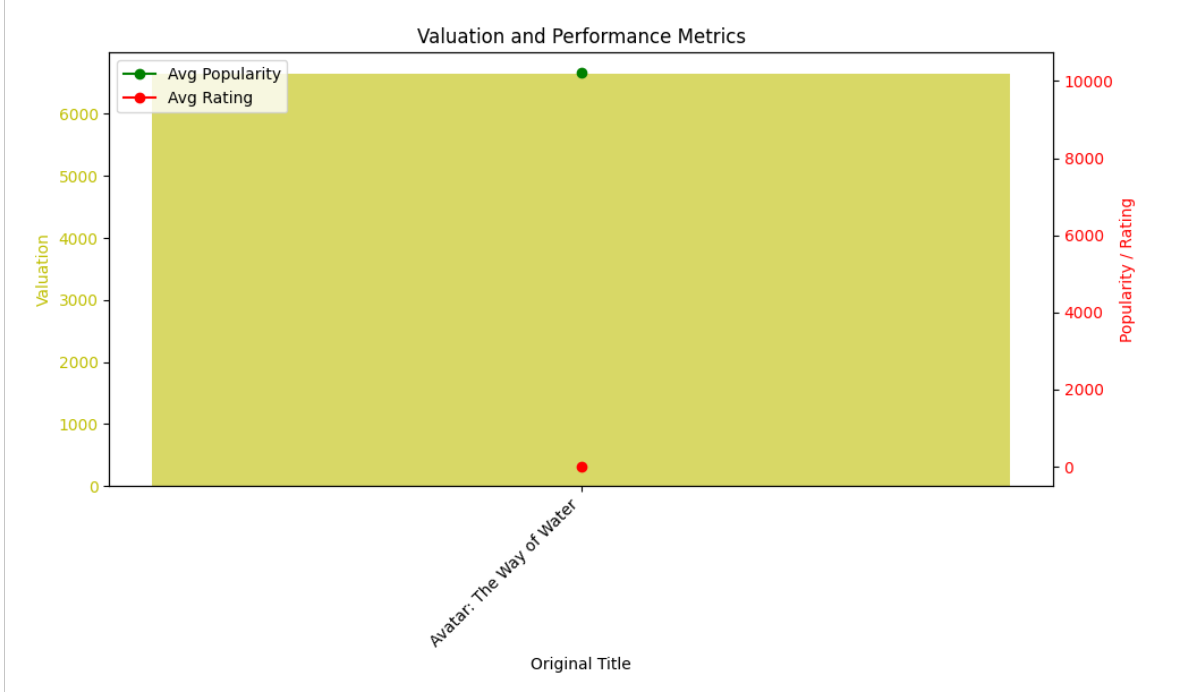
* **Status**: Completed
* **Details**: Plots for **attendance count** and **popularity** were generated over time. These visualizations helped identify trends and fluctuations in content performance based on the release timeline. The analysis revealed that attendance tends to spike soon after release and then stabilizes over time, while popularity follows a similar trend.

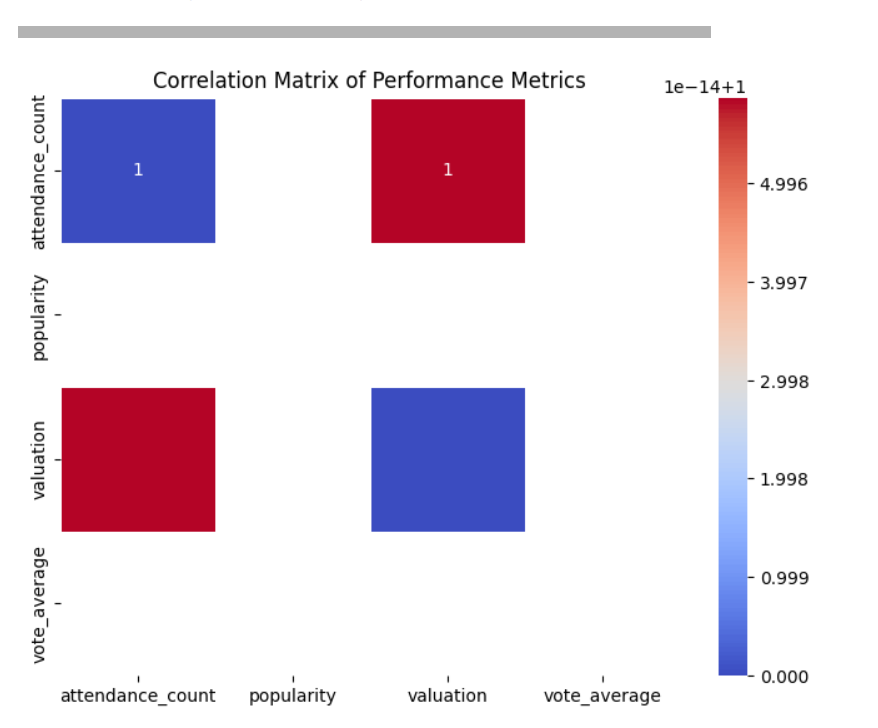


**Task 3 : Visualization of Key Performance Metrics**

* **Status**: Completed
* **Details**: Generated time series plots to visualize the **attendance count** and **popularity** over time, highlighting trends and performance fluctuations post-release. These visualizations provided insights into content dynamics throughout the release timeline.

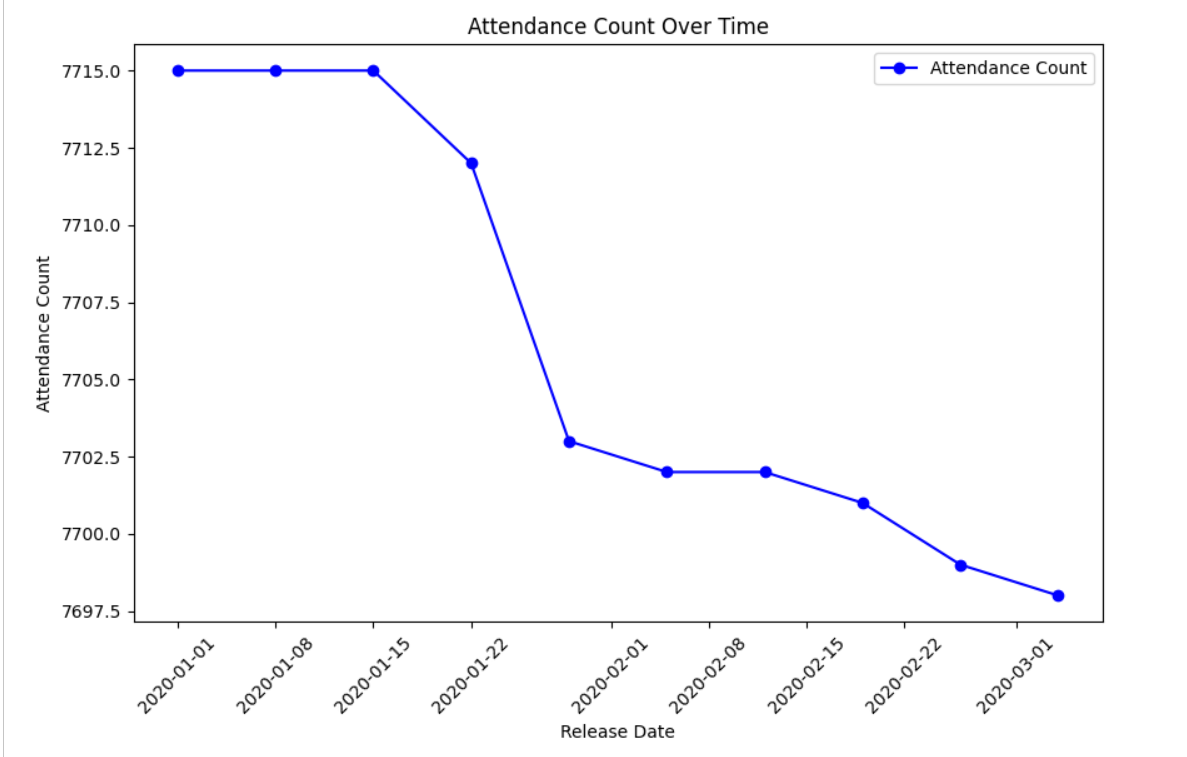


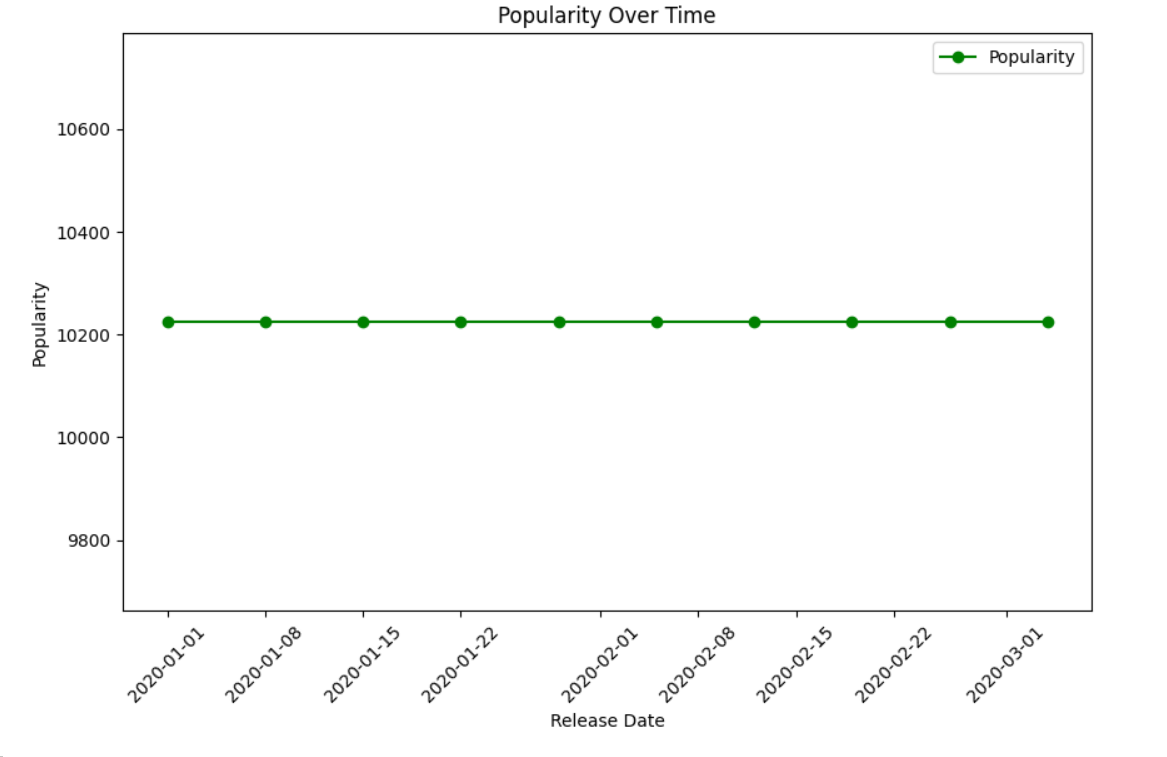




#### Task 4: ****Predictive Modeling****

* **Status**: Completed
* **Details**: A Linear Regression model was developed to predict attendance count. Features such as valuation, popularity, vote\_average, and days\_since\_release were used to train the model. The model’s performance was evaluated using **Mean Squared Error (MSE)**, which helped assess the accuracy of predictions.





**Progress :**

· **Accomplishments**:

* Successfully added a release\_date column and performed time series analysis.
* Built a predictive model to estimate future attendance based on key metrics.

· **Metrics**:

* The **MSE** for the predictive model was calculated, indicating the error rate for predictions. Further tuning may be required to improve model performance.

# Challenges and Solutions :

* **Challenges Faced**:
  + Initially, the dataset lacked a release\_date column, which is crucial for time series analysis.
* **Solutions Implemented**:
  + A synthetic release\_date was generated, assuming regular content releases. This allowed for the successful completion of the time series analysis.

# Next Steps :

* **Upcoming Tasks**:
  + Refine the predictive model by testing other algorithms like Random Forest or Gradient Boosting.
  + Analyze real-time performance by incorporating actual release dates if available.
* **Goals**:
  + Improve the accuracy of the predictive model and incorporate advanced techniques for better forecasting.

# Conclusion :

**Summary**: Today's tasks were focused on performing time series analysis and predictive modeling for on-demand content performance in the entertainment sector. Key trends in attendance and popularity were identified, and a predictive model was built to forecast future performance.

# **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.