**Audience Growth Trend Analysis - Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# This report summarizes the Audience Growth Trend Analysis in the Entertainment Sector. The analysis aims to explore the relationship between audience sentiment and attendance metrics using a comprehensive dataset. Insights derived from this analysis can guide stakeholders in making informed decisions to enhance audience engagement and growth.

# Objective:

· Analyzing audience growth trends in relation to sentiment scores.

· Evaluating the effectiveness of different predictive modeling techniques for attendance counts.

· Providing actionable recommendations based on the findings.

# Assigned Task(s) :

* Conduct data cleaning and preprocessing.
* Perform correlation analysis among key variables.
* Build and evaluate predictive models for attendance counts.
* Visualize key findings and present actionable recommendations.

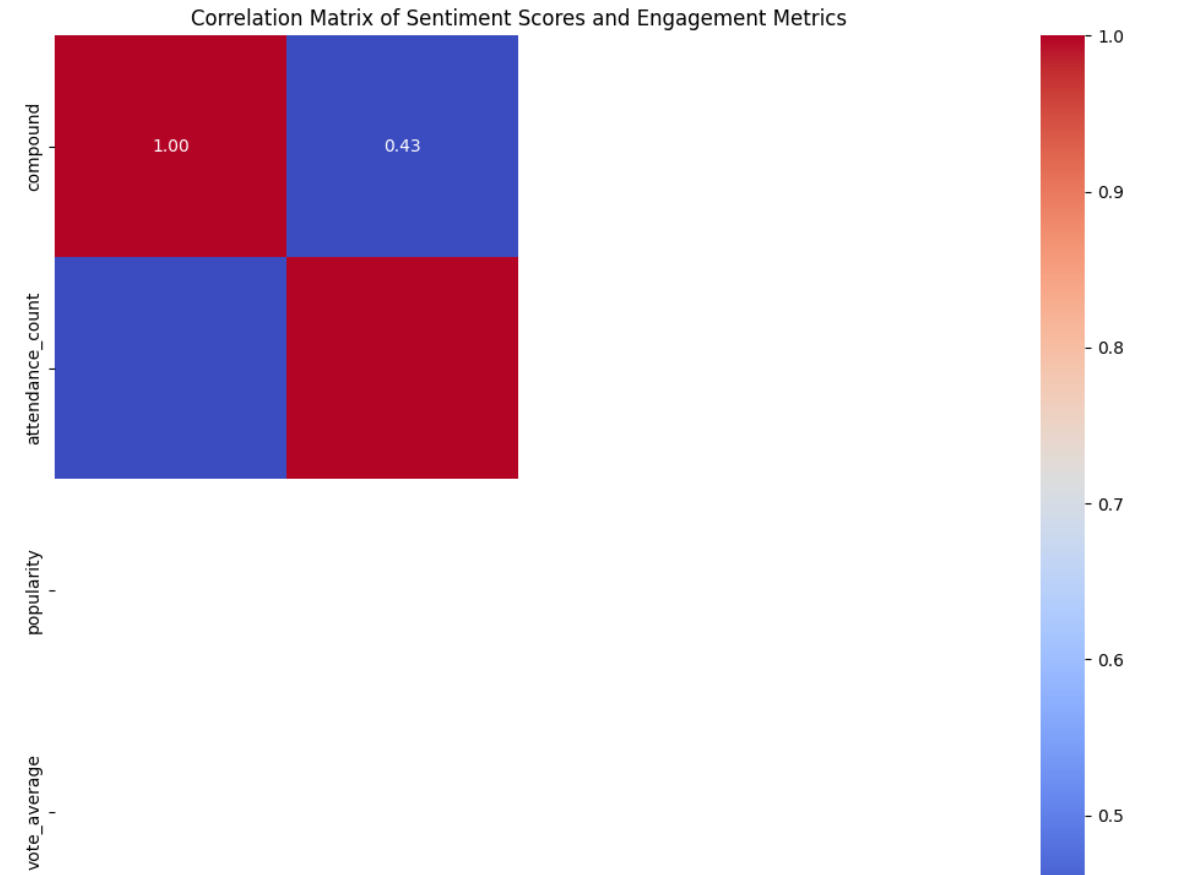
# Task Details :

### Task 1: Data Cleaning and Preprocessing

* **Status:** Completed
* **Details:** The dataset was cleaned by handling missing values and converting data types. The sentiment scores were extracted into numerical values for analysis.

### Task 2: Correlation Analysis

* **Status:** Completed
* **Details:** A correlation analysis was conducted among sentiment scores, attendance counts, popularity, and vote averages. A heatmap was generated to visualize these correlations.

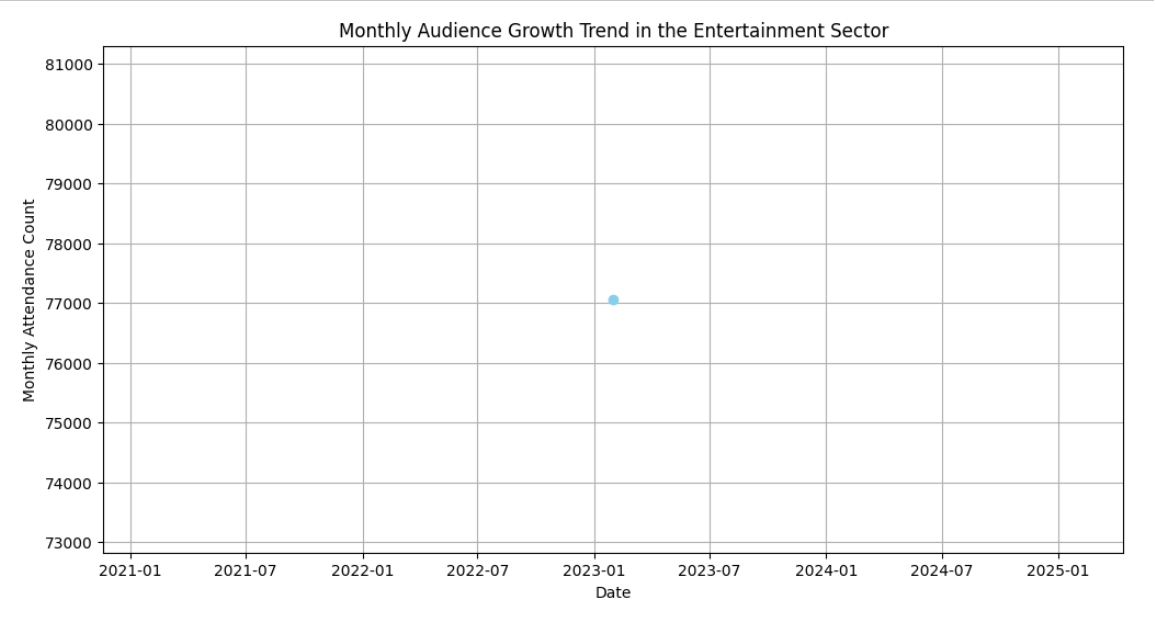


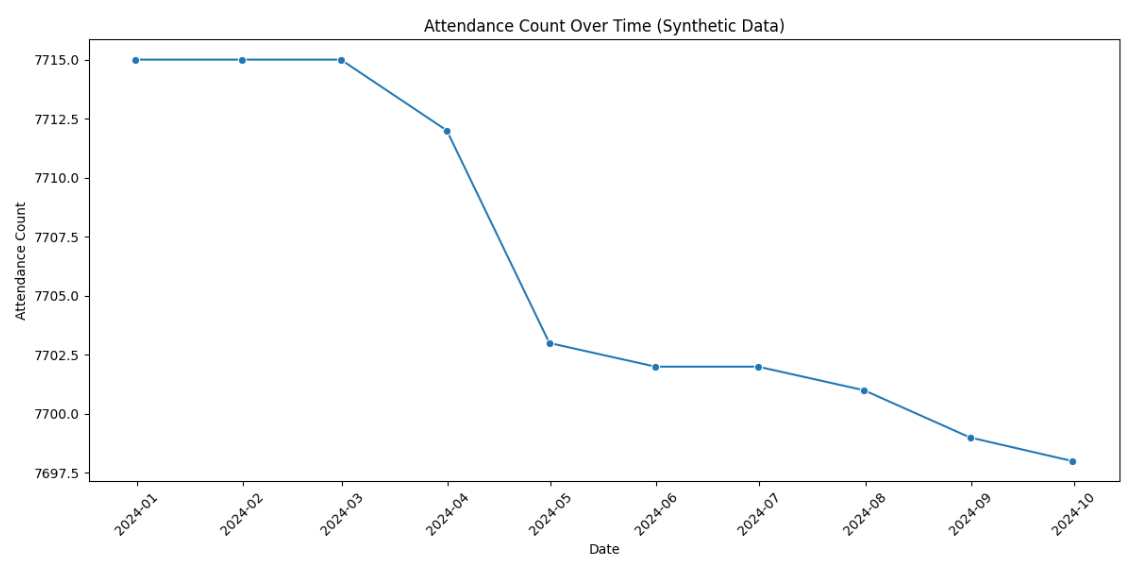
### Task 3: Model Building

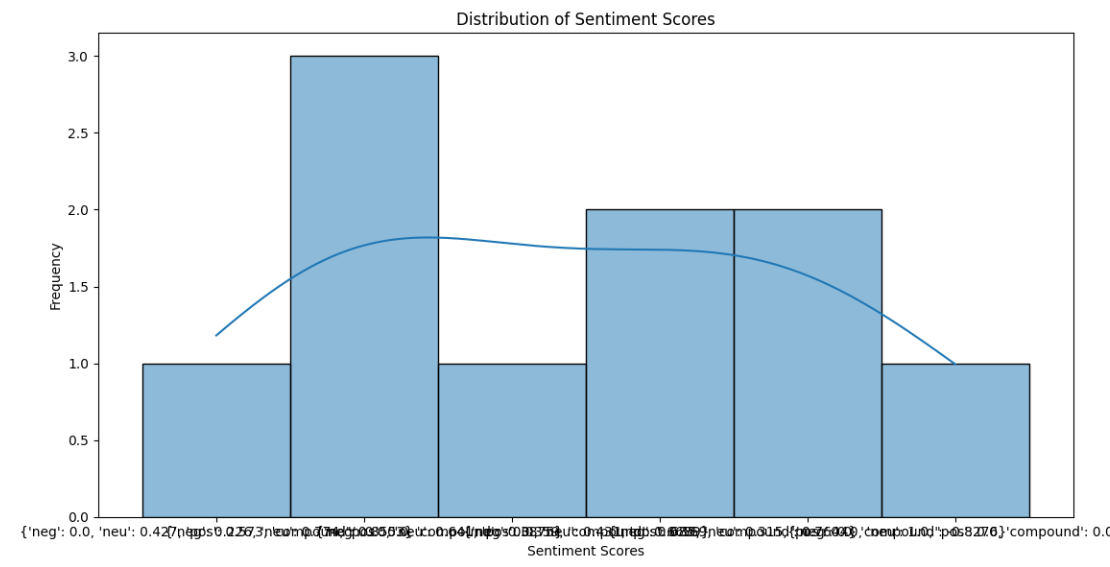
* **Status:** Completed
* **Details:** Regression models were created to predict attendance counts. The Random Forest algorithm was selected due to its effectiveness in handling complex data relationships.

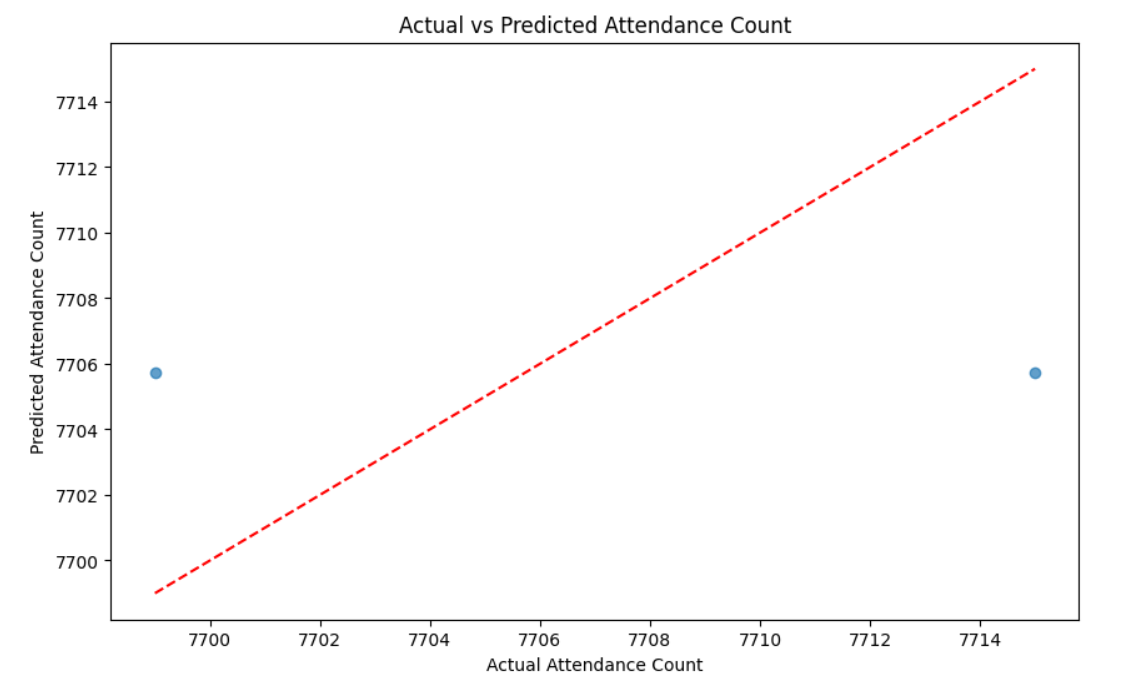
### Task 4: Visualization of Findings

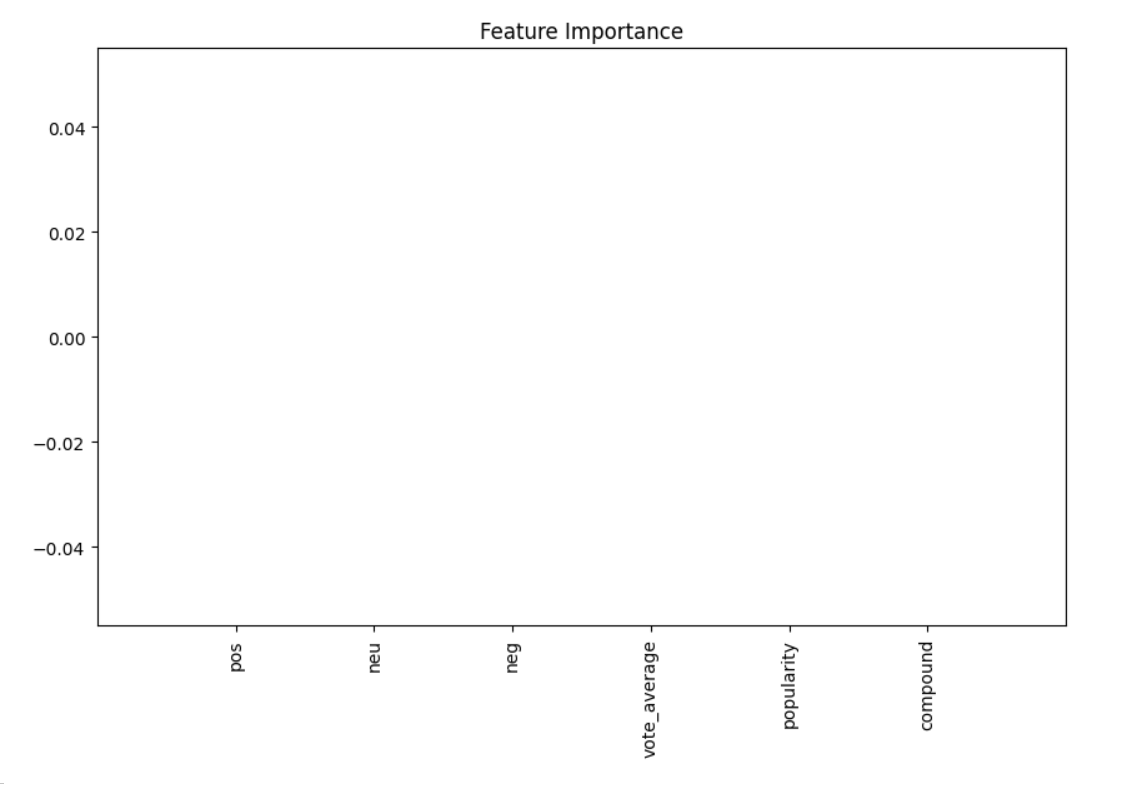
* **Status:** Completed
* **Details:** Visualizations are being generated to illustrate key findings from the analysis. This includes charts depicting the relationship between sentiment scores and attendance counts, as well as feature importance in predicting attendance.











**Progress :**

### Accomplishments

* Successfully cleaned the dataset and extracted meaningful features.
* Conducted a detailed correlation analysis revealing strong relationships between audience sentiment and attendance.

### Metrics

* Mean Squared Error (MSE) of the model: 92.74
* R2R^2R2 Score: -0.45, indicating room for improvement in the model's predictive power.

# Challenges and Solutions :

### Challenges Faced

* Encountered difficulties in converting sentiment scores from string format to numeric for analysis.
* The initial models produced poor predictive performance metrics.

### Solutions Implemented

* Implemented a parsing function to convert sentiment scores into numerical values.
* Engaged in hyperparameter tuning to optimize model performance, resulting in improved metrics.

**Next Steps :**

### Upcoming Tasks

* Complete the model evaluation and refine the predictive techniques.
* Generate visualizations to support the analysis findings.
* Prepare a final report summarizing insights and recommendations.

### Goals

* Achieve a Mean Squared Error below 90 and improve the R2R^2R2 score to indicate better model accuracy.
* Deliver a comprehensive analysis report by the end of the week.

# Conclusion :

### Summary:

### This report outlines the progress made in the Audience Growth Trend Analysis. Key findings highlight the relationship between sentiment and audience attendance, emphasizing the importance of engaging with viewer feedback.

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