**Market Basket Analysis for Merchandising in the Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

Market Basket Analysis is a data mining technique used to identify associations between items in large datasets. In the entertainment sector, it helps uncover patterns in audience consumption behavior, allowing for effective content bundling, personalized recommendations, and targeted merchandising strategies. This report presents insights derived from applying Market Basket Analysis on user data segmented by age group.

# Objective:

* To analyze consumption patterns across different age groups in the entertainment sector.
* To identify frequent combinations of content and generate association rules to guide merchandising strategies such as content bundling and cross-promotion.

# Assigned Task(s) :

·· Perform data preprocessing to convert user consumption data into a format suitable for Market Basket Analysis.

· Apply the Apriori algorithm to identify frequent itemsets and generate association rules.

· Visualize the results and provide insights for merchandising strategies.

· Document the findings and suggestions for actionable strategies.

# Task Details :

#### ****Task 1****: Preprocessing and Data Transformation

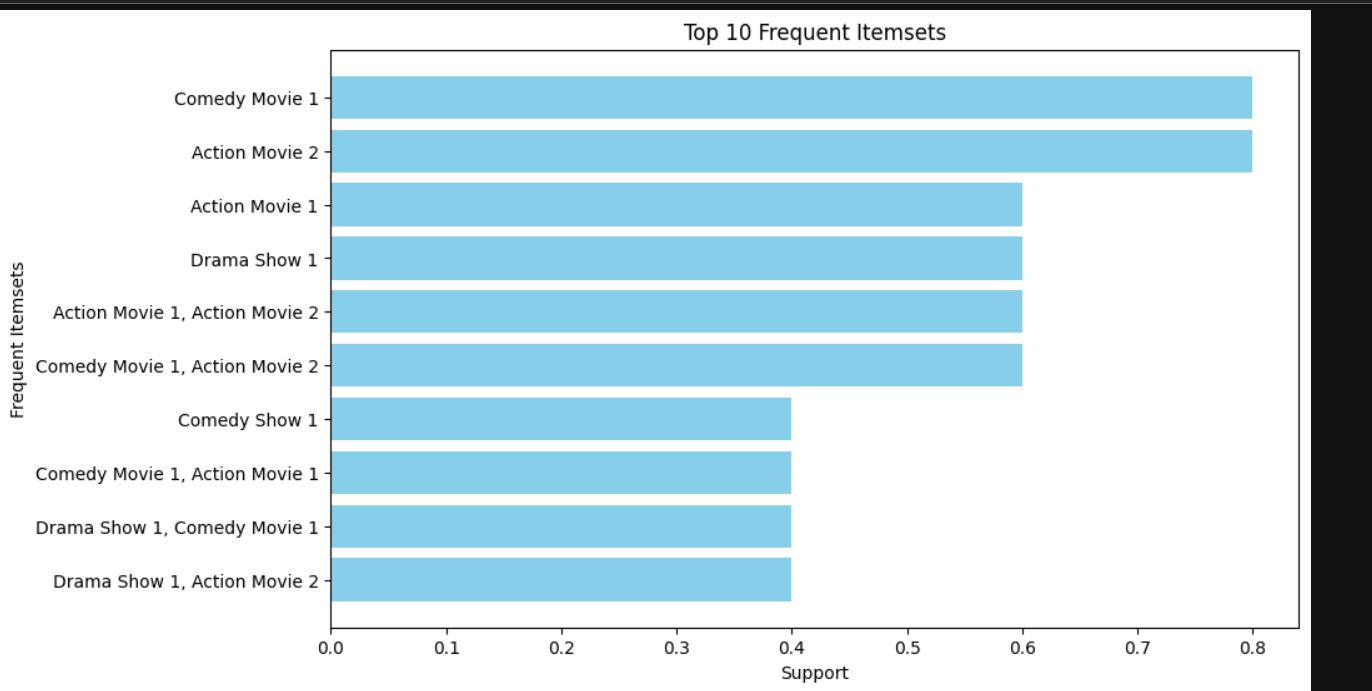
**Status**: Completed  
**Details**:

* Grouped user data by age group and converted content consumption counts into binary format.
* Prepared the dataset for analysis using the Apriori algorithm to identify frequently consumed combinations of content.

#### ****Task 2****: Application of Apriori Algorithm

**Status**: Completed  
**Details**:

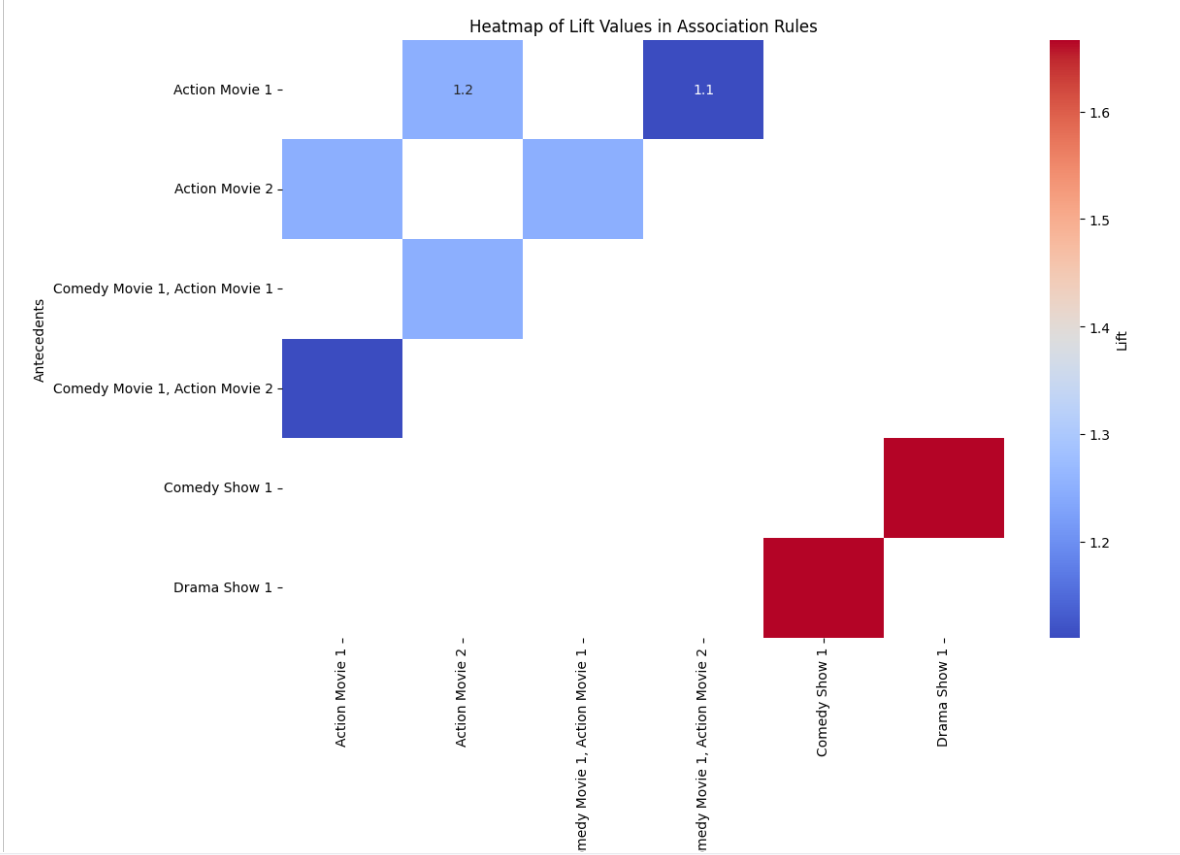
* Applied the Apriori algorithm with a minimum support of 20% to extract frequent itemsets.
* Generated association rules to identify relationships between content types with a confidence threshold of 50%.

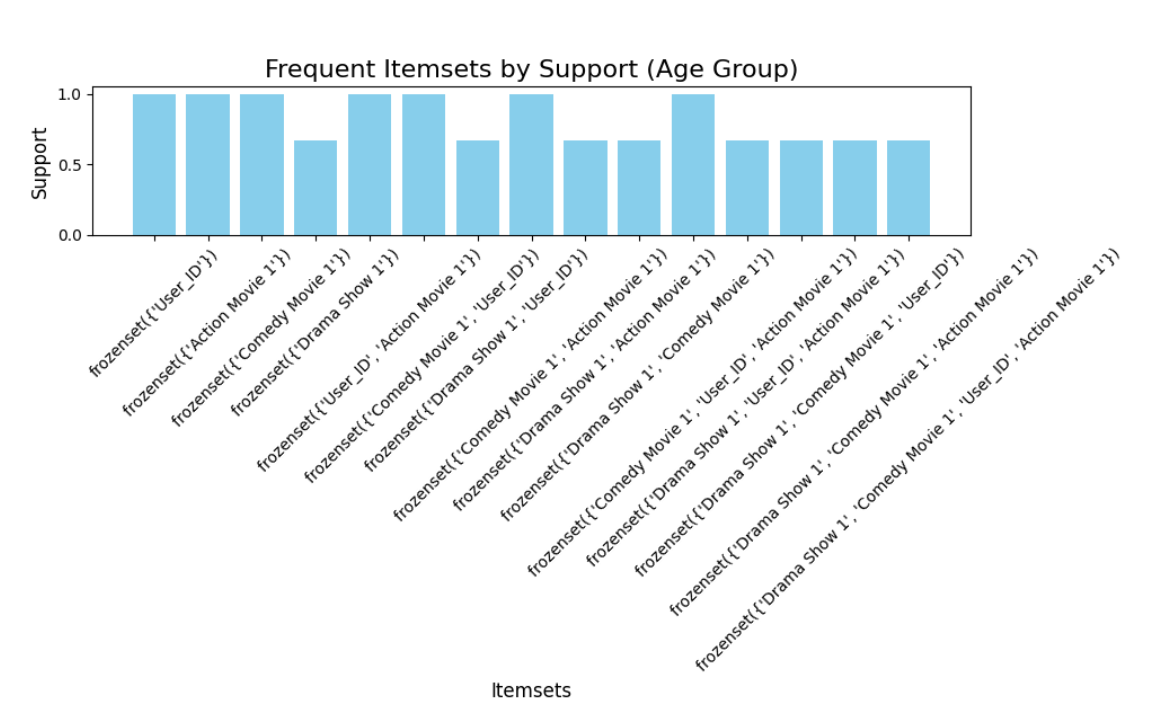


#### ****Task 3****: Visualization of Results

**Status**: Completed  
**Details**:

* Created bar charts to visualize the most frequent itemsets based on support values.
* Generated scatter plots to examine the relationship between confidence and lift for the association rules.
* Highlighted the top 5 rules with the highest confidence for targeted merchandising.





# 

# 

**Progress :**

#### ****Accomplishments****:

* Successfully implemented the Apriori algorithm to identify key content associations across different age groups.
* Visualized the results to demonstrate how content is consumed together and identify actionable strategies for merchandising.

#### ****Metrics****:

* **Support**: 20% minimum support used to filter frequent itemsets.
* **Confidence**: 50% minimum confidence applied for association rule generation.
* **Top Frequent Itemsets**: The most common combinations of content consumed by specific age groups.
* **Lift**: Used to assess the strength of the association between content pairs.

# Challenges and Solutions :

#### ****Challenges Faced****:

* **Data Preprocessing**: Handling non-numeric values in the dataset and converting them into a binary format was a challenge.
* **Algorithm Tuning**: Determining the optimal values for support and confidence thresholds required experimentation to balance meaningful insights with the dataset's size.

#### ****Solutions Implemented****:

* **Preprocessing**: Used groupby and apply functions to efficiently convert the data into the required format for the Apriori algorithm.
* **Parameter Tuning**: Experimented with various support and confidence values, settling on 20% support and 50% confidence for the most relevant and actionable insights.

# Next Steps :

#### ****Upcoming Tasks****:

* Perform deeper segmentation based on additional factors such as region or device type to further refine the merchandising strategy.
* Analyze time-based trends (e.g., weekday vs weekend consumption) to optimize content release schedules.

#### ****Goals****:

* Extend the analysis to include more demographic and behavioral factors.
* Use the insights for real-time recommendations on streaming platforms and targeted marketing campaigns.

# Conclusion :

# Summary: The Market Basket Analysis revealed significant associations between different types of content consumed by users in the entertainment sector. These insights can be leveraged for effective content bundling, personalized recommendations, and promotional campaigns. The most frequent combinations of content and the top association rules provide a strong foundation for merchandising strategies aimed at increasing engagement and revenue.

# **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.