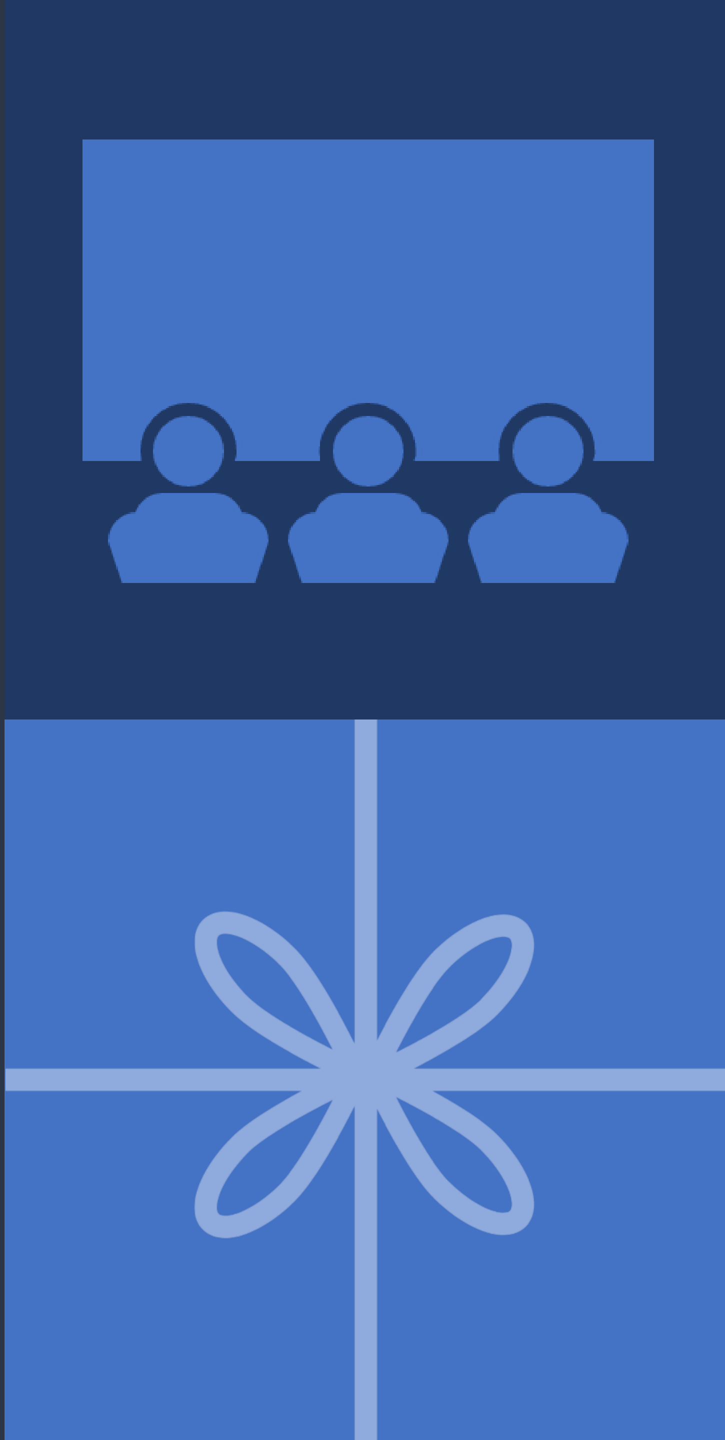
# Movie Recommendations system project

# Business Understanding

- To build a recommendation system for personalized movie suggestions on a streaming platform.
- Enhancing user engagement and retention by providing relevant movie recommendations based on past ratings.

# Objective

1. Predict which movies users will likely enjoy based on their past preferences and ratings.
2. Use collaborative filtering to understand patterns in user ratings.
3. Provide personalized movie recommendations based on user and item (movie) similarities.



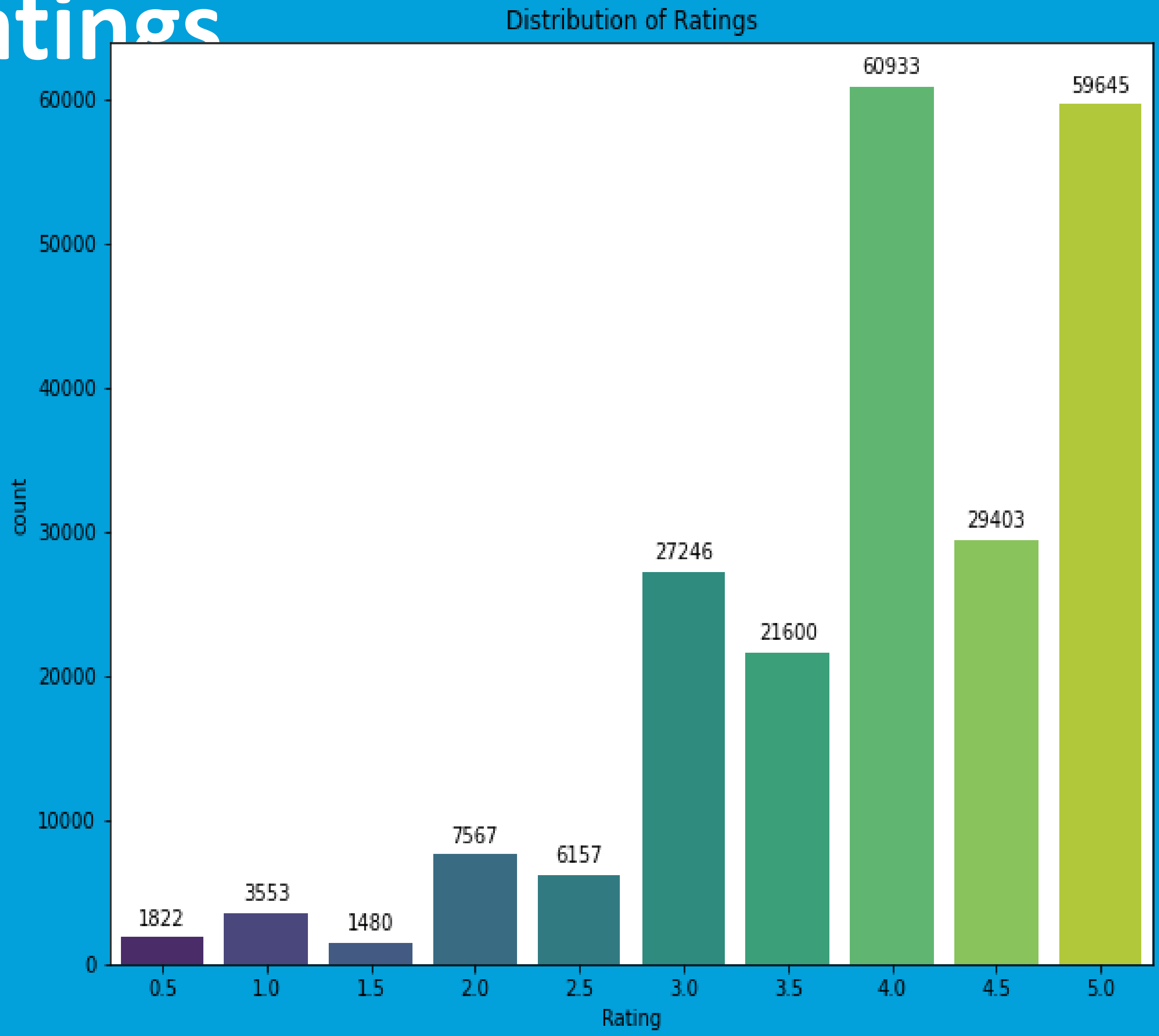
# Data understanding

**Dataset: Movie Lens dataset.**

**Key Features**:

- ✓User ID: Represents unique users.
- ✓Movie ID: Unique identifiers for movies.
- ✓Rating: User ratings of movies (on a scale from 1 to 5).
- ✓Tags: Keywords or phrases describing movie themes (e.g., “Action,” “Adventure”).
- ✓Timestamp: Records when users rated the movie.

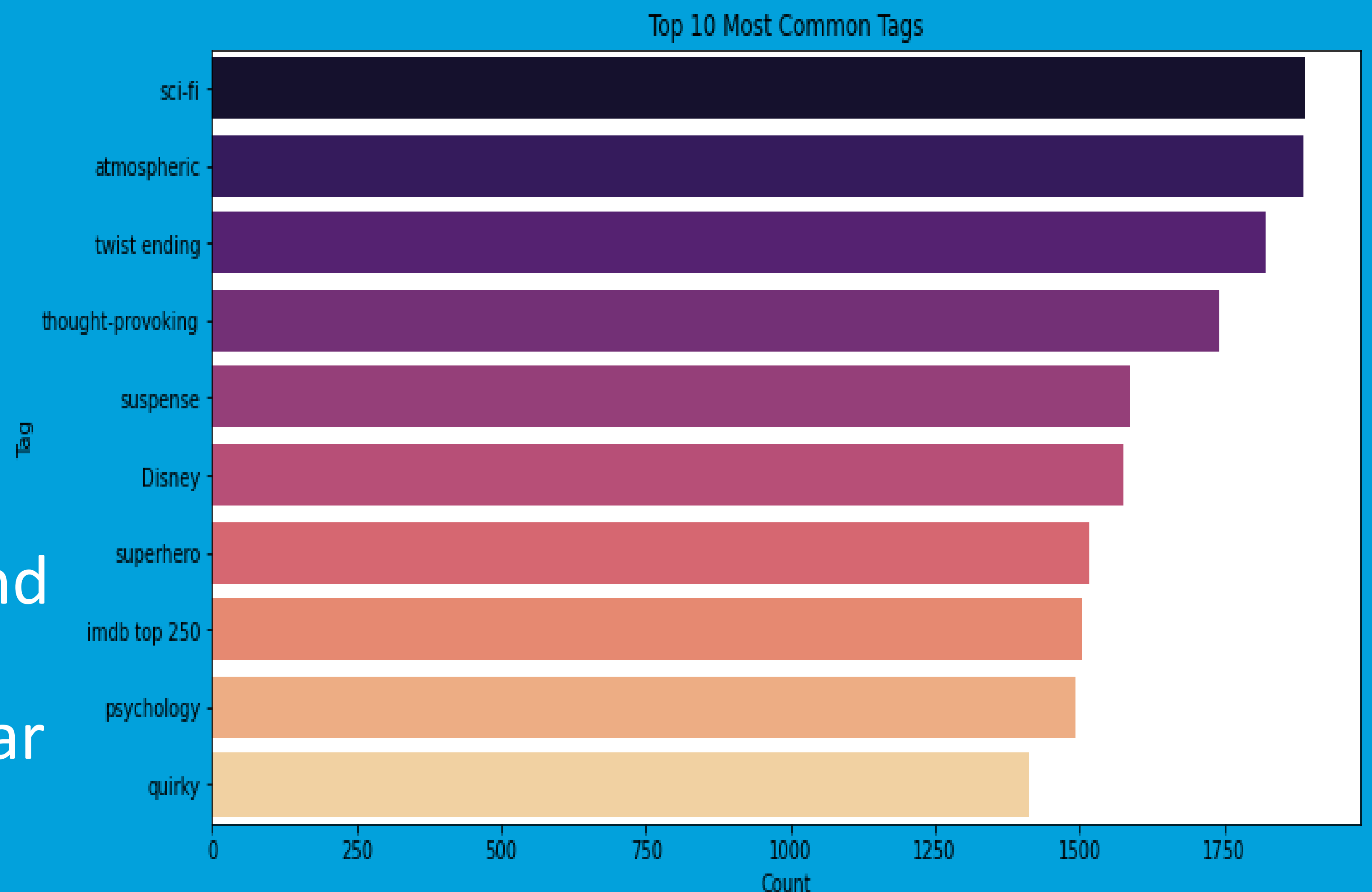
# 4: EDA (Explanatory Data Analysis)
# Distribution of Movie Ratings



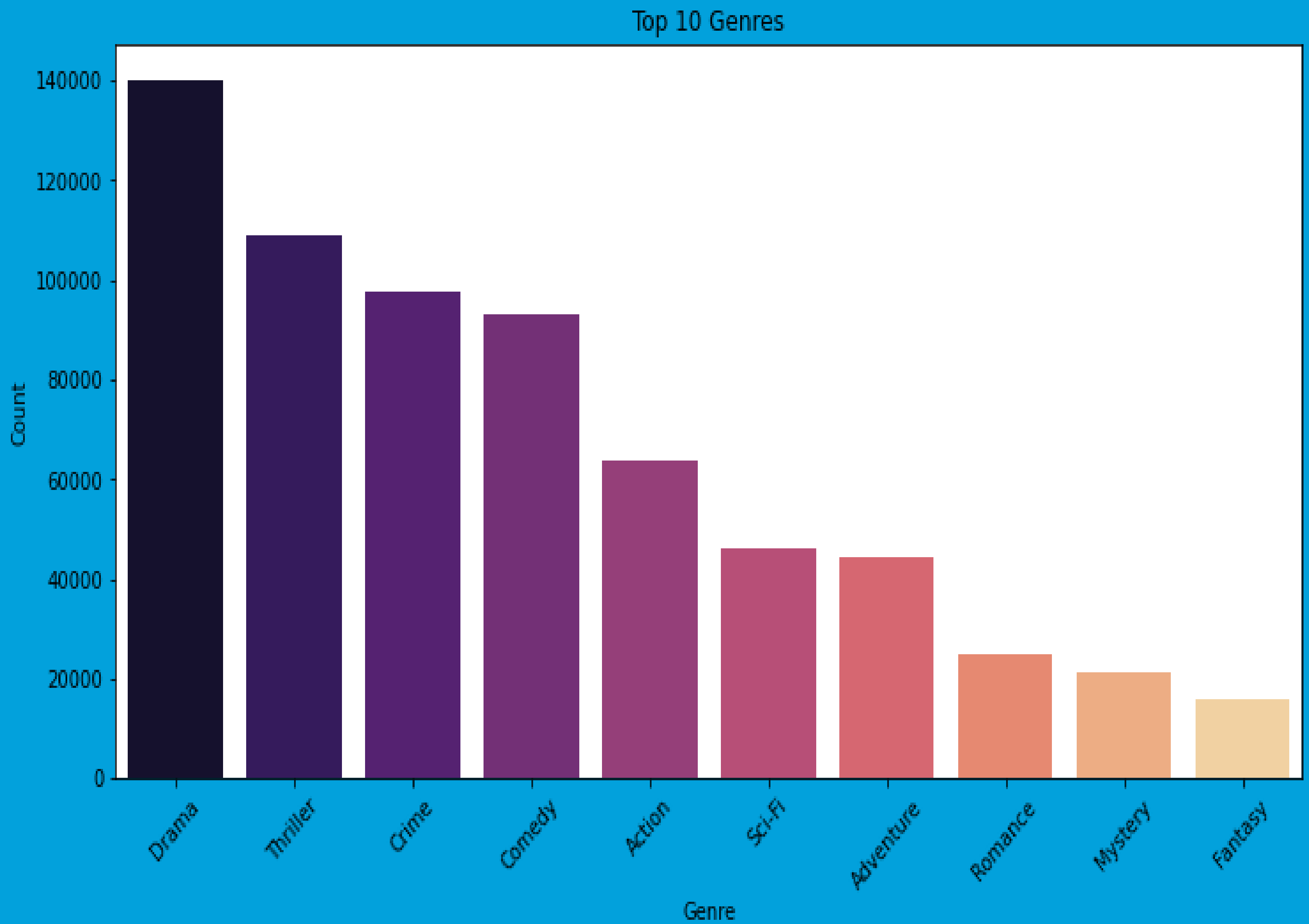

- ✓ It helps to understand how users rate movies, which can reveal rating patterns and user tendencies.
- ✓ Most ratings are concentrated between 3 to 5 stars, showing a bias towards positive ratings.
- ✓ 4 and 5 had highest user ratings

# Top 10 Common movie TagsPurpose

- ✓Identify the most frequently used movie tags, which highlights dominant themes or genres.Insight
- ✓Common tags such as ‘Adventure,’ ‘Comedy,’ and ‘Action’ dominate the dataset, reflecting popular genres.



# Top 10 Genres




- ✓ This bar chart showcases the most common movie genres in the dataset, giving a sense of the overall genre distribution.
- ✓ By analyzing the top genres, we can identify trends in user preferences and which genres are the most frequent on the platform.
- ✓ Genres like **Adventure, Action, and Comedy** likely dominate, indicating user interests in these types of movies. The most frequent genres will guide content acquisition and recommendation strategies.

# Modelling

## K-Nearest Neighbors (KNN)

- Finds similar users or movies based on their ratings and recommends items that similar users liked.
- It is intuitive and simple and works well with dense datasets.
- Evaluation Metric: Accuracy, Precision.

## Collaborative Filtering

- Matrix factorization decomposes the user-item interaction matrix into two smaller matrices (latent factors), predicting missing values (unrated movies).
- It Learns hidden patterns from both users and items, leading to accurate recommendations.
- Evaluation Metric: RMSE, Precision, Recall.

# Modelling (cont.)

## Content-Bases Filtering

- Recommends movies based on the user’s past interactions, using movie features such as genres, directors, and actors.
- It Works well when user interaction data is sparse; does not require extensive user history.
- Evaluation Metric: RMSE, Precision.

## Hybrid Model

- Combines collaborative filtering and content-based filtering to create a more robust recommendation system.
- It  Uses both user behavior and movie attributes for more accurate.
- Evaluation Metric: RMSE, Precision, Recall.

# Modelling (Evaluation Results)Collaborative

## Collaborative Filtering

- Providers the most accurate recommendations, as measured by RMSE and Precision@K.
- It Learns from user-item interaction to predict unknown ratings effectively.

## KNN

- Works well with smaller datasets but suffers from performance issues in sparse matrices.Strength: Simple and effective for highly-rated or niche movies.

# 7: Conclusion

- Collaborative Filtering proved to be the most accurate model for movie recommendations.
- Content-Based Filtering is useful for new users with fewer ratings.
- Hybrid Models offer the best balance, combining the strengths of both approaches
- The recommendation system can significantly enhance user experience by providing personalized suggestions based on historical data.

# 8: Recommendations

- Implement the collaborative filtering model on the platform to provide real-time recommendations.
- Regularly update the system with new user data for continuous improvement.
- Continuous Learning:Add new user ratings and movie information over time to improve model accuracy.
- Additional Data:Incorporate user feedback and social media interactions to further refine recommendations

# 9: Further Improvements

- ✓Explore user segments (e.g., age groups, regions) to deliver more personalized recommendations.
- ✓Address the issue of recommending for new users with no ratings by leveraging content-based filtering and hybrid approaches.
- ✓A/B Testing: Continuously monitor and improve the system by conducting A/B tests on recommendation quality and user engagement.


Thank you!


NEXT
STEP