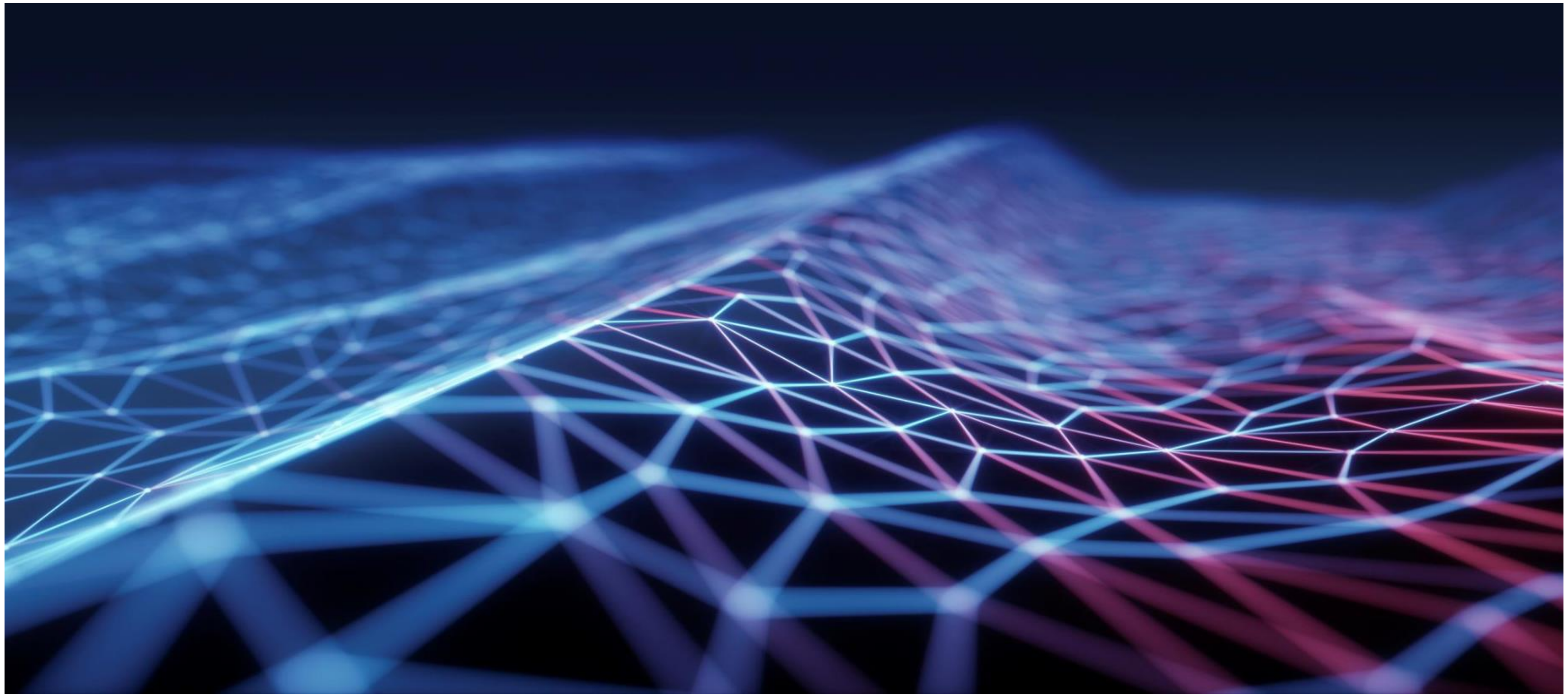
MOVIES RECOMMENDATION SYSTEM.



GROUP 10:

- Taylor Musa.
- Edna Wanjiku.
- Ian Macharia.
- Dorine Langat.
- Brian Nyagah.
- Samwel Muiru

PROJECT LAYOUT

- Business understanding
- Data understanding
- Data preparation
- Exploratory data analysis
- Modeling evaluation

BUSINESS UNDERSTANDING

The movie industry is vast and fast evolving , with countless movies and movie sequels released each year. This can be a challenge for the users to navigate through the vast amount of content and get to know which movies align with their preferences.

PROJECT OBJECTIVES

- Develop a recommendation system that leverages user data and movie information to provide personalized movie recommendations.
- Implement different recommendation techniques to ensure a diverse and accurate set of movie recommendations.
- To develop a movie recommendation system based on movie attributes, user ratings, and user interactions.

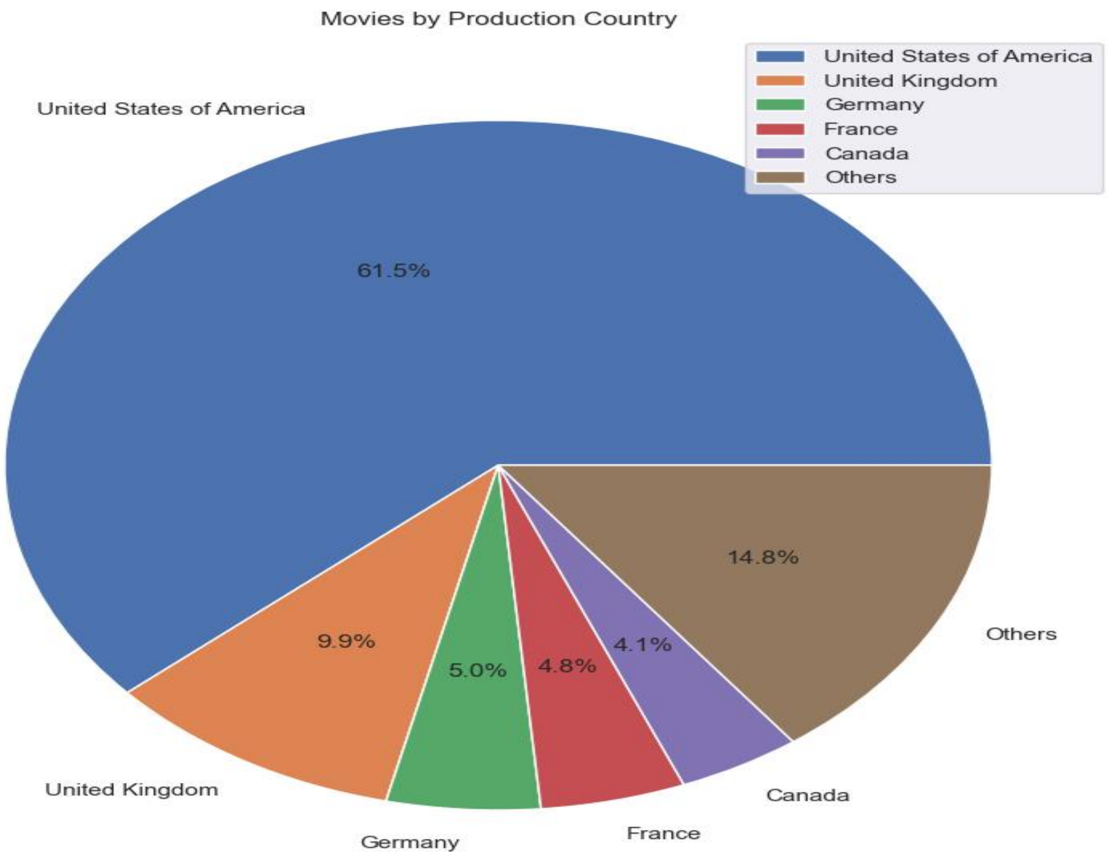
DATA UNDERSTANDING

TMDB is a popular database that provides comprehensive information about movies, and it contains the following: titles, release dates, genres, cast and crew information.

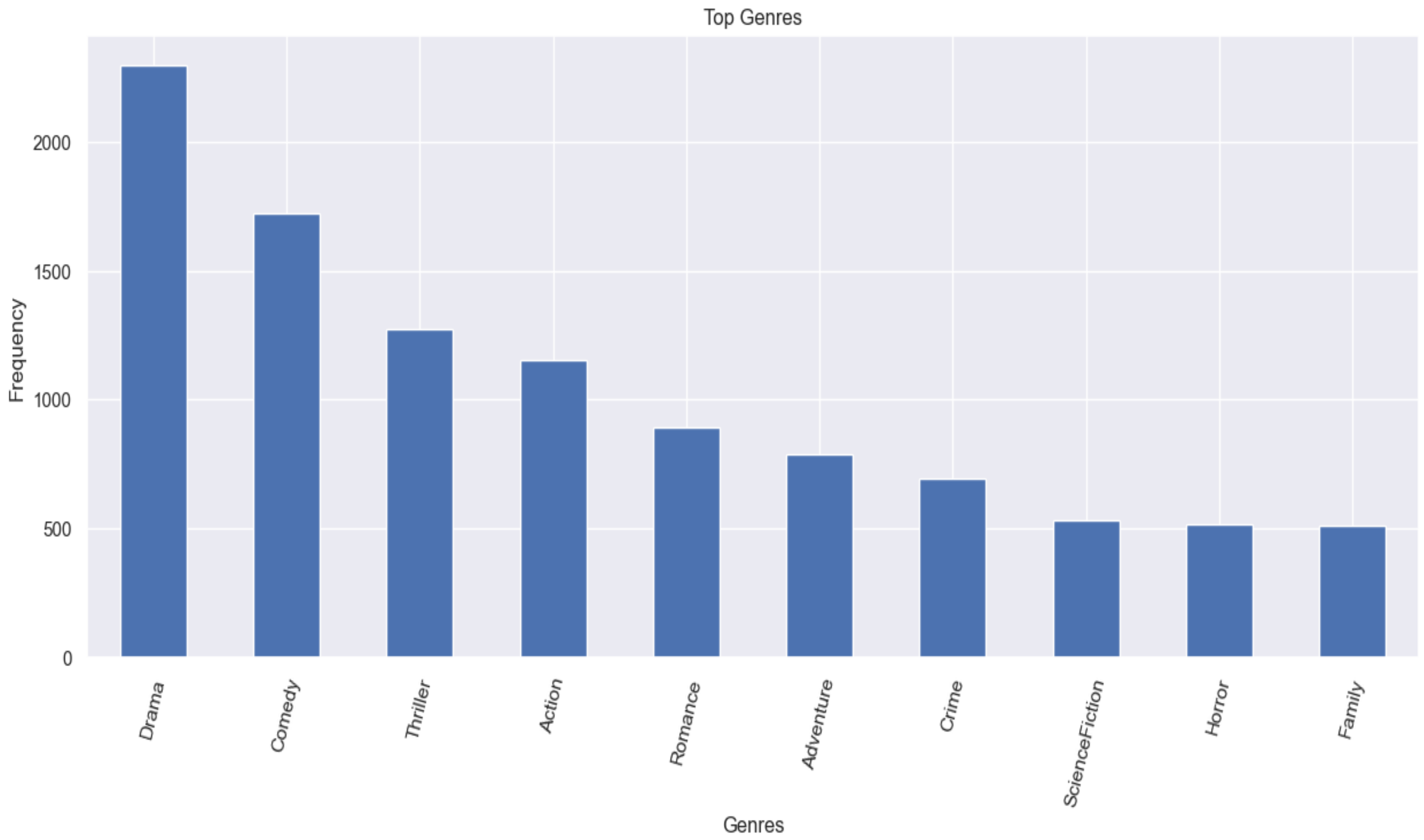
Credit information is given as well about the cast and crew information. With the combination of the datasets, we gain valuable insights and perform various analyses related to the movie industry.

Distribution of Movies based on Country of Production

Based on this analysis, we conclude that majority of the movies (61.5%) were produced in the United States of America.



Top Genres



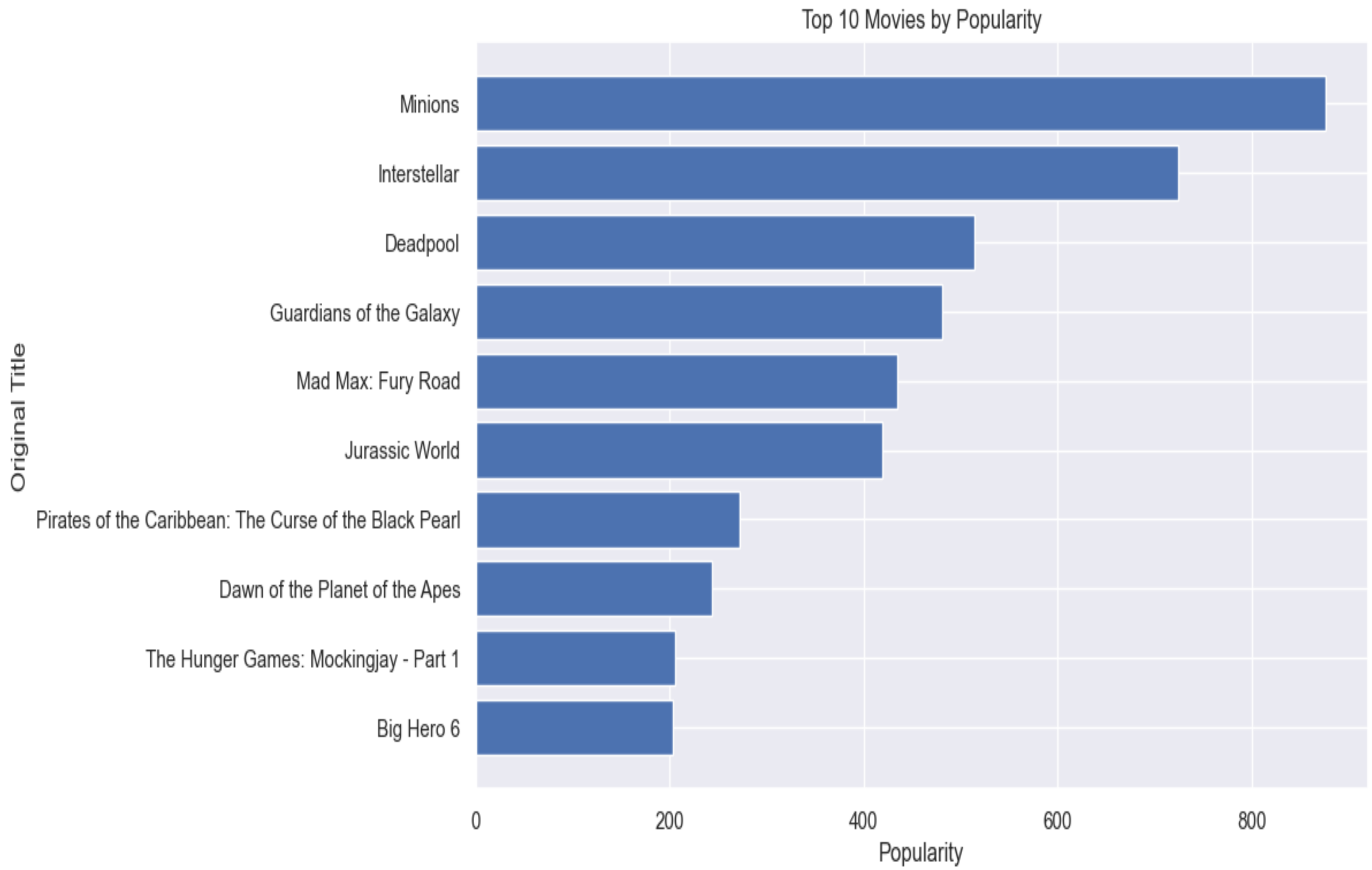
Based on this analysis, we found Drama to be the genre with the largest number of movies.

RECOMMENDATION SYSTEMS

We created four recommendation systems:

- Demographic recommendation based on popularity.
- Content based.
- Collaborative based recommendation.
- Hybrid recommendation

Most popular movies



This analysis shows the top movies, based on popularity.

HYBRID RECOMMENDATION

```
`# Applying the function` is a comment in the code indicating that the function
#`hybrid_recommendations` is being called with the arguments `7` and `'The Conjuring 2'`
# Applying the function
hybrid_recommendations(7,'The Conjuring 2')
```

✓ 0.0s

index	title	movieId	vote_average	vote_count	estimated_rating
1011	The Tooth Fairy	53953	4.3	13	3.358947
2096	The Conjuring	138843	7.4	3092	3.076316
1627	Deliver Us from Evil	184346	5.9	690	3.076316
4644	Teeth and Blood	325123	3.0	1	3.076316
3554	Insidious: Chapter 2	91586	6.5	1211	3.076316
2430	Evil Dead	109428	6.4	1723	3.076316
3569	Paranormal Activity: The Marked Ones	227348	5.2	449	3.076316
4224	Insidious	49018	6.8	1737	3.076316
2477	Jennifer's Body	19994	5.3	837	3.076316
2379	Restoration	396152	5.3	11	3.076316

```
# A different genre of movie, some Science Fiction perhaps?
hybrid_recommendations(56,'Avatar')
```

✓ 0.0s

index	title	movieId	vote_average	vote_count	estimated_rating
466	The Time Machine	2135	5.8	631	3.945905
2403	Aliens	679	7.7	3220	3.921125
47	Star Trek Into Darkness	54138	7.4	4418	3.798053
61	Jupiter Ascending	76757	5.2	2768	3.798053
83	The Lovers	79698	4.8	34	3.798053
1201	Predators	34851	6.0	1206	3.798053
260	Ender's Game	80274	6.6	2303	3.798053
2372	Megaforce	27380	3.5	15	3.798053
71	The Mummy: Tomb of the Dragon Emperor	1735	5.2	1387	3.700681
2327	Predator	106	7.3	2093	3.652790

EVALUATION

Evaluating RMSE, MAE of algorithm SVD on 6 split(s).

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Mean	Std
RMSE (testset)	0.4213	0.4230	0.4200	0.3956	0.4144	0.3890	0.4105	0.0133
MAE (testset)	0.1629	0.1651	0.1621	0.1550	0.1621	0.1510	0.1597	0.0050
Fit time	0.91	1.38	0.98	0.95	0.91	0.81	0.99	0.18
Test time	0.11	0.09	0.15	0.08	0.08	0.08	0.10	0.03

Fit time represents the time taken by the algorithm to train on each fold of the data.

Test time represents the time taken by the algorithm to make predictions on the test data for each fold.

Overall, the SVD algorithm shows relatively low RMSE and MAE values, indicating good accuracy in predicting movie ratings. The algorithm has moderate fit and test times, suggesting efficient performance.

CONCLUSION

The recommendation system serves as a valuable tool in the movie industry to address the challenge of content navigation and provide personalized movie recommendations. By understanding user preferences, leveraging similarities between users, and utilizing movie features, the system aims to enhance the user experience, increase engagement, and ultimately contribute to user retention on the platform.

RECOMMENDATIONS

- Real-time Updates: Incorporate a mechanism to continuously update the movie database with the latest releases, ratings, and reviews.
- Contextual Factors: Consider contextual factors such as time of day, location, mood, and social trends to provide personalized recommendations that align with the user's current situation and preferences.
- User Feedback and Improvement Loop: Implement a feedback mechanism that allows users to rate and provide feedback on recommended movies.

NEXT STEPS

- Iterative Refinement: Continuously iterate and refine the recommendation system based on user feedback and performance evaluation results.
- Evaluation and Validation: Conduct extensive testing and evaluation of the recommendation system using real-world scenarios and user feedback.
- System Design and Implementation: Design and develop a scalable and robust recommendation system architecture that can handle a large volume of users and movies.

THANK YOU