Movie Recommendation Insights:
An Advanced Tableau Dashboard Case Study

Prasang Biyani

Introduction

As a Data Scientist with more than 5 years of experience in machine learning and analytics, I developed an advanced dashboard using Tableau Public to analyze movie preferences and provide recommendation insights. Leveraging the MovieLens 1M dataset, this project simulates a real-world application for a streaming platform—identifying top movies, trending genres, and personalized recommendations. The goal was to blend data preprocessing, predictive modeling, and interactive visualization into a portfolio piece that demonstrates my ability to derive actionable insights from complex data.

Data Overview

The MovieLens 1M dataset contains 1 million ratings from ~6,000 users across ~4,000 movies, with user demographics (age, gender) and movie genres. I preprocessed the raw .dat files using Python (Pandas), merging ratings, movies, and users into a single dataset (movielens_data.csv) for simplicity, and creating an expanded version (movielens_merged_expanded.csv) by splitting multi-genre fields into individual rows for granular analysis. This resulted in ~1M rows (merged) and ~2.8M rows (expanded), with timestamps converted to datetime for time-series analysis.

Methodology

Data Visualization: Using Tableau Public, I created five key visuals on separate sheets:

1. *Top-Rated Movies (Bar Chart)*: Ranked movies by a custom "Popularity Score" (rating count × average rating).
2. *Popular Genres Trending Over Time (Line Chart)*: Tracked genre popularity (rating count) across years, filtered to top 5 genres.
3. *Ratings by Genre and Age (Heatmap)*: Analyzed average ratings across genres and age groups.
4. *Movie Recommendations (Table)*: Displayed ML-driven recommendations.

Machine Learning: I implemented a collaborative filtering model in Python using cosine similarity on a user-movie ratings matrix, generating top-5 movie recommendations for a subset of titles. These were merged with movie titles and imported into Tableau.

Dashboard Assembly: Combined visuals into an interactive dashboard with filters for gender, age, genres, and movie selection, published via Tableau Public.

Key Insights

1. Top Movies: Classics like "The Shawshank Redemption" topped the Popularity Score, balancing high ratings with broad appeal.
2. Genre Trends: Action and Comedy genres saw consistent popularity in the early 2000s, with Drama peaking around specific years, suggesting cyclical viewer preferences.
3. Demographic Patterns: Younger users (18-24) rated Action movies higher, while older groups (35+) favored Drama, as seen in the heatmap.
4. Recommendations: For "Star Wars: Episode IV," the model suggested related sci-fi classics (e.g., "The Empire Strikes Back"), validating its relevance.

5. Interactivity: Filters revealed how preferences shift by gender and age, offering a dynamic tool for exploration.

Technical Highlights

1. Preprocessing: Handled large-scale data merging and genre expansion using Python, optimizing for Tableau.
2. Custom Metrics: Developed a Popularity Score (COUNT × AVG) in Tableau to refine movie rankings.
3. ML Integration: Built and deployed a cosine similarity-based recommendation system, bridging Data Science and BI.
4. Visualization: Designed an interactive, multi-layered dashboard, demonstrating both technical and storytelling skills.