

Leveraging EDA Insights for a Data-Driven Movie Studio Launch

Phase Two Project- Group 3

Endalkachew Dessalegne

Austine Otieno

Myrajoy Kiganane

Danton Kipngeno

Introduction

- A company is launching a new movie studio but lacks expertise in filmmaking.
- The goal is to analyze successful films and extract actionable insights.
- Findings will guide the studio in making Informed data-driven production decisions.

Key Business Questions

- Which movie genres have the highest average rating?
- Which movie genres receive the highest number of votes?
- What impact does runtime have on ratings and votes?
- How does language influence ratings and votes?
- How does region influence ratings and votes?

Data Overview

- Data sourced from IMDb (<https://www.imdb.com/>).
- Key tables used:
 - movie_basics: Contains genres & runtime
 - movie_ratings: Contains average ratings & number of votes
 - movie_akas: Contains region & language details

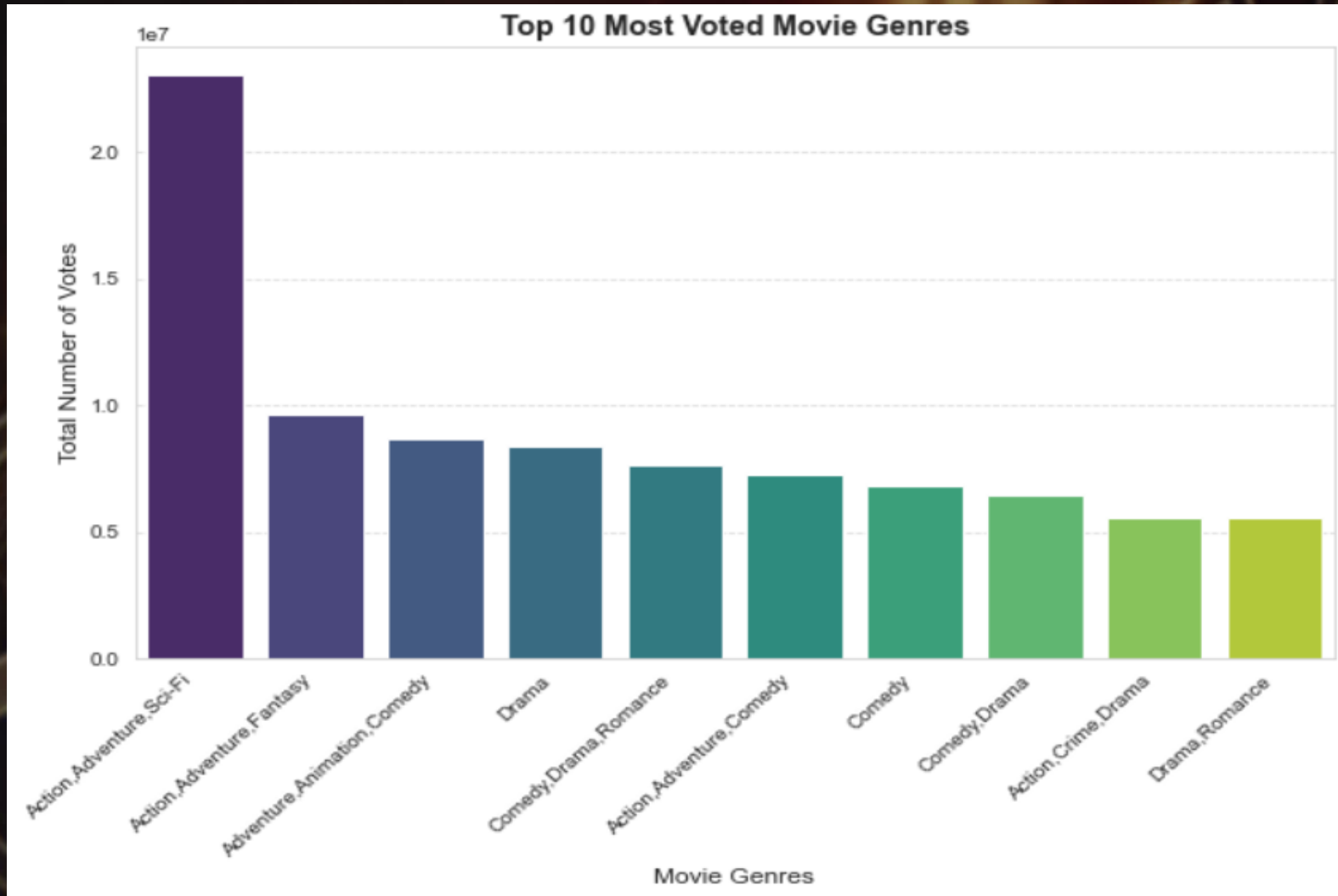
Methodology

- **Data Collection:** Extracted relevant movie data from IMDb.
- **Data Preprocessing:** Cleaned and merged tables using SQL.
- **Exploratory Data Analysis (EDA):** Uncovered trends in ratings, votes, genres, language, and region.
- **Visualization:** Used Python libraries for data representation.

Findings & Insights

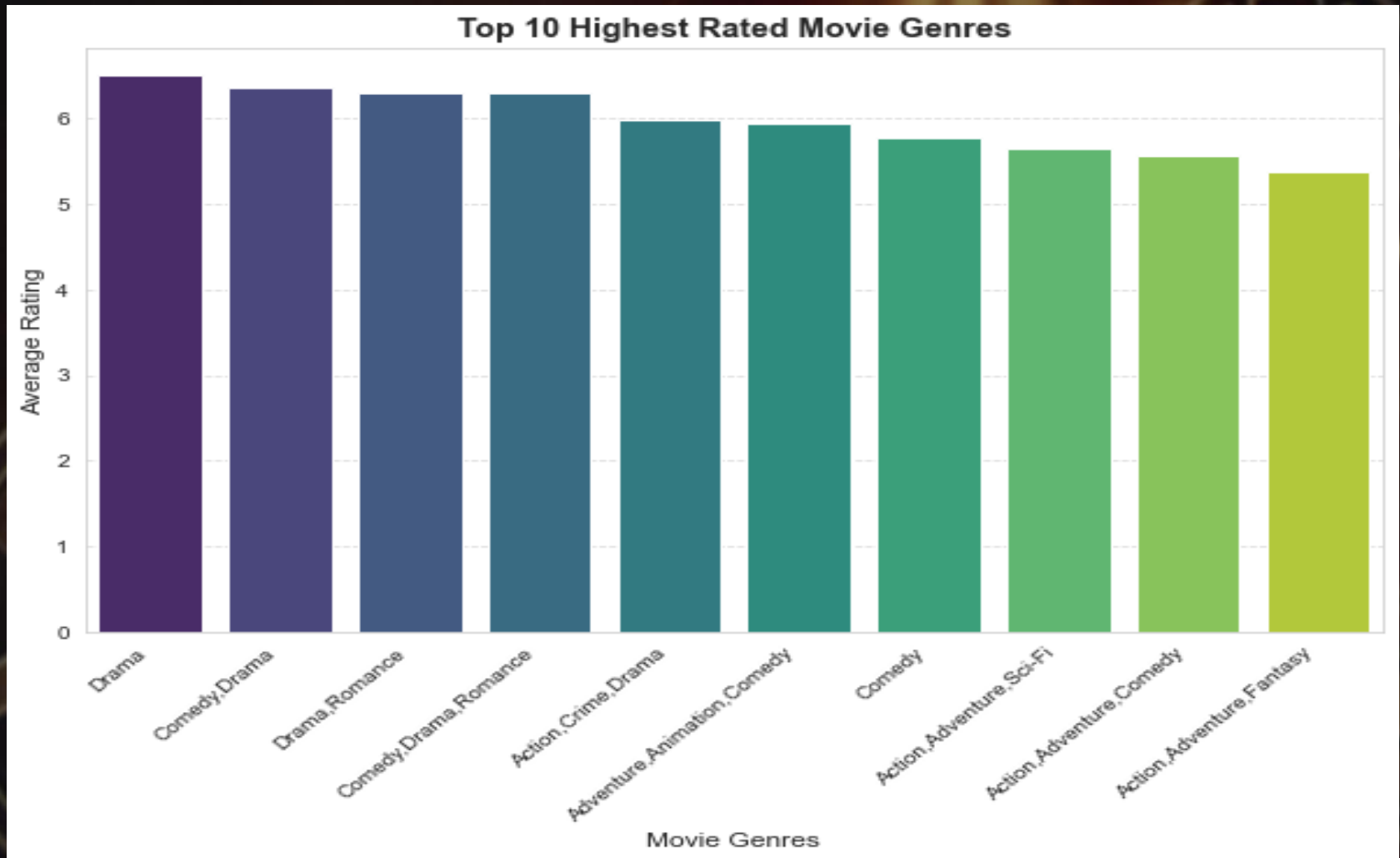
- **Certain genres consistently receive higher ratings.**
- **Movies with moderate runtime tend to perform better.**
- **Language and region impact audience engagement.**
- **Number of votes correlates with movie popularity and marketing reach.**

Highest Voted Movie Genres



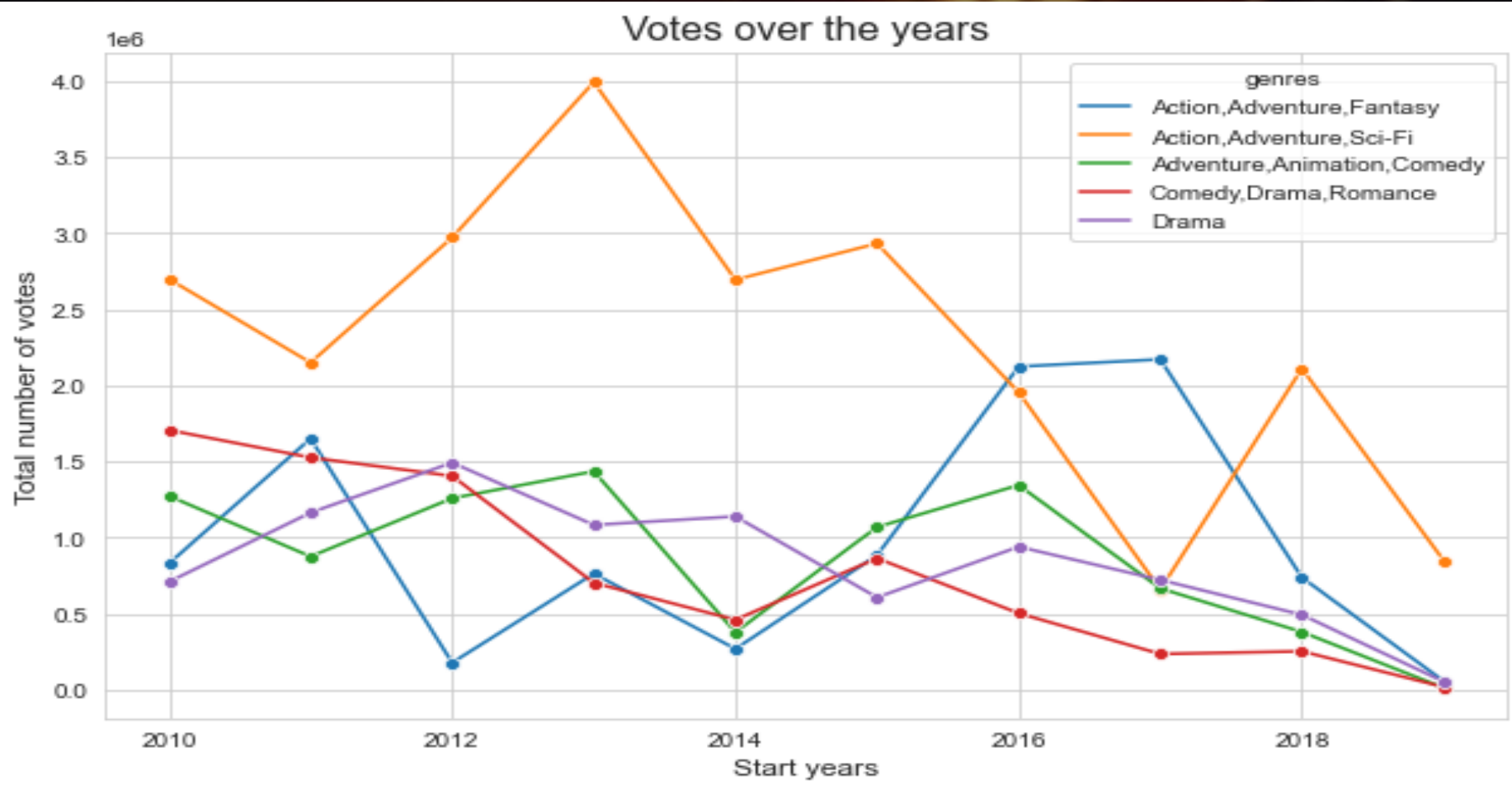
- Action, Adventure, Sci-Fi genre has the highest number of votes followed by Action, Adventure, Fantasy and Adventure, Animation, Comedy genres.
- Drama, Romance genre has the lowest votes among the top 10 highest voted.

Highest Rated Movie Genres



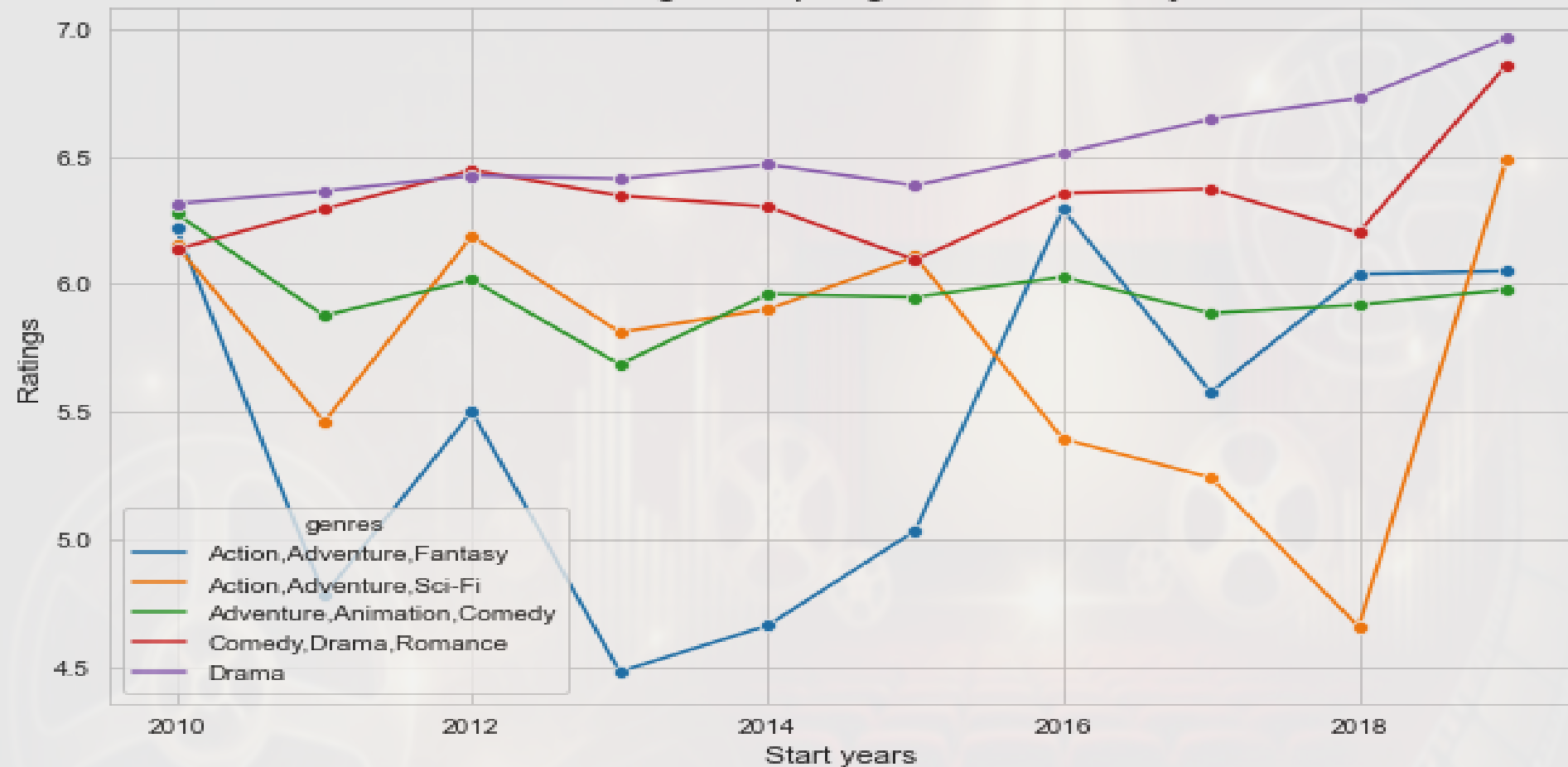
- Drama genre has the highest rating based on the top ten genres with the highest number of votes followed by Comedy, Drama and Drama, Romance.
- Action ,Adventure, Fantasy genre has the lowest rating among the top 10 highest rated genres.
- All top 10 genres have ratings in the range between 5.4 and 6.5.

VOTES OVER YEARS-LINE GRAPH



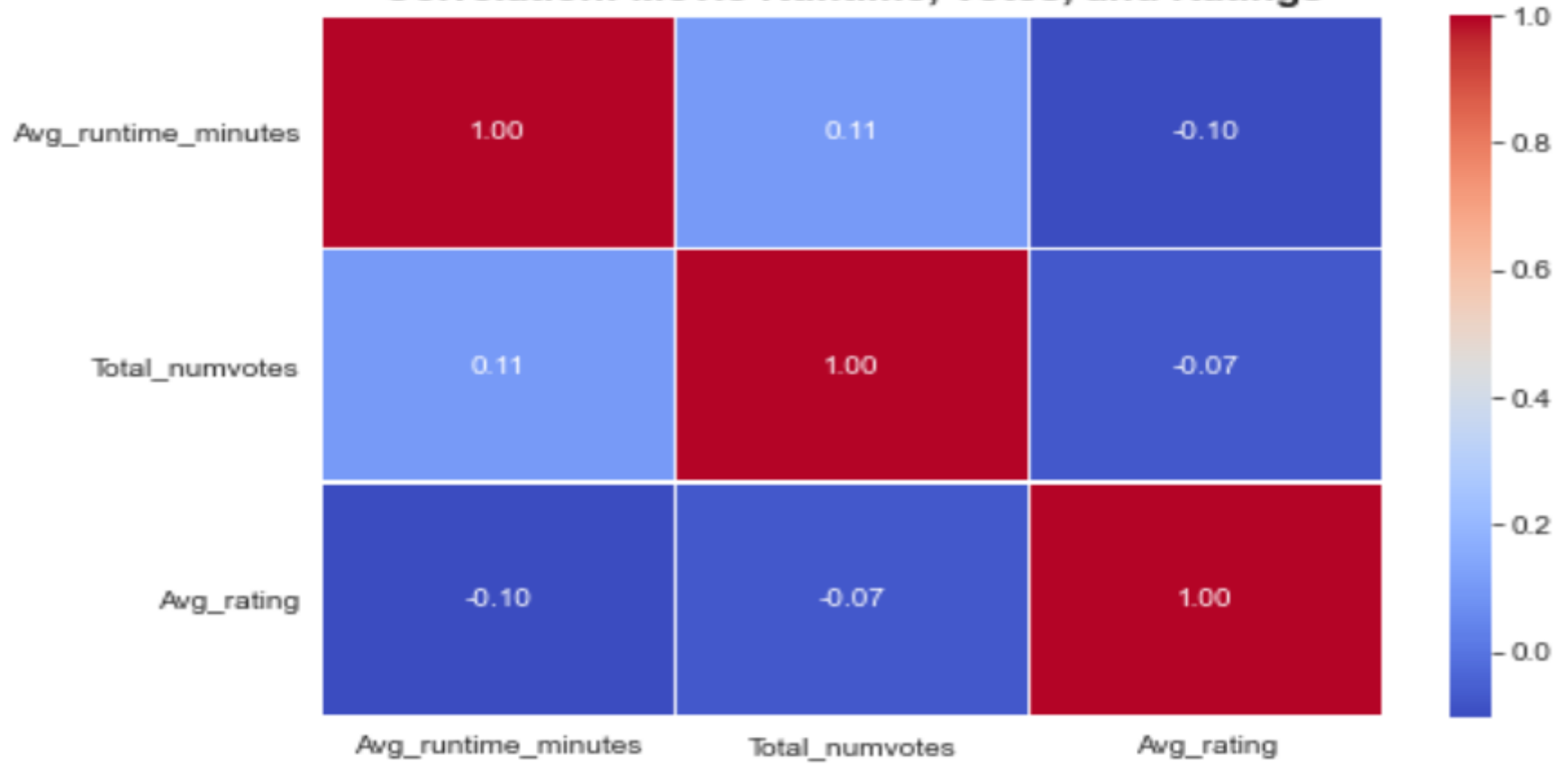
- This displays the total number of votes over the years from the year 2010 to the year 2019.
- Number of votes have been decreasing gradually over the years.
- Action,Adventure,Sci-Fi genre is with the highest number of votes over the years.

Movies Ratings of top 5 genres over the years



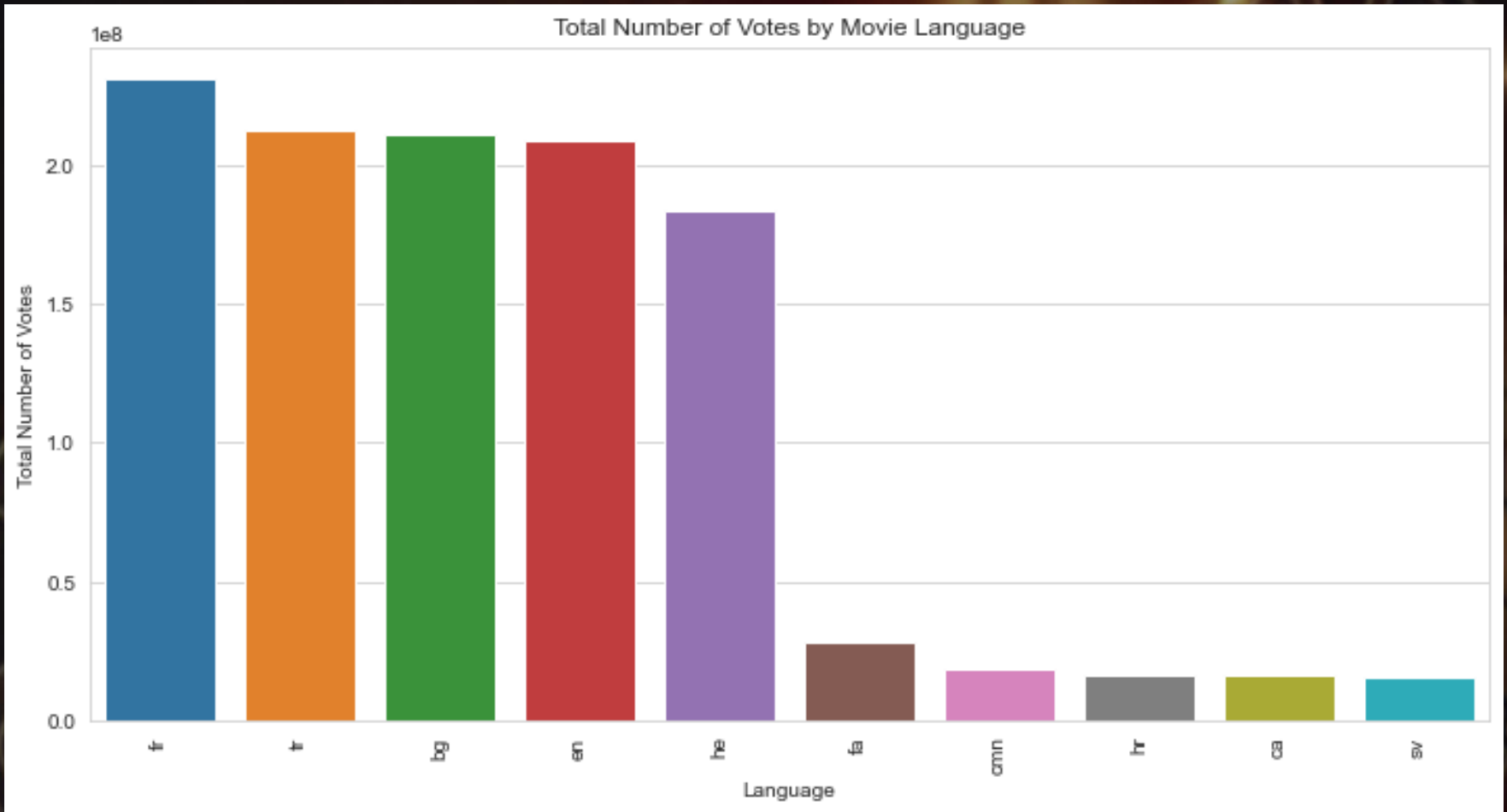
- There is a general trend of increase in the rating of most of the genres.
- Drama genre has the highest and increased rating over the years followed by comedy, Drama, Romance and Adventure, Animation, comedy genres.
- Action, Adventure, Sci-Fi and Action, Adventure, Fantasy genres had fluctuating ratings over the years.

Correlation: Movie Runtime, Votes, and Ratings



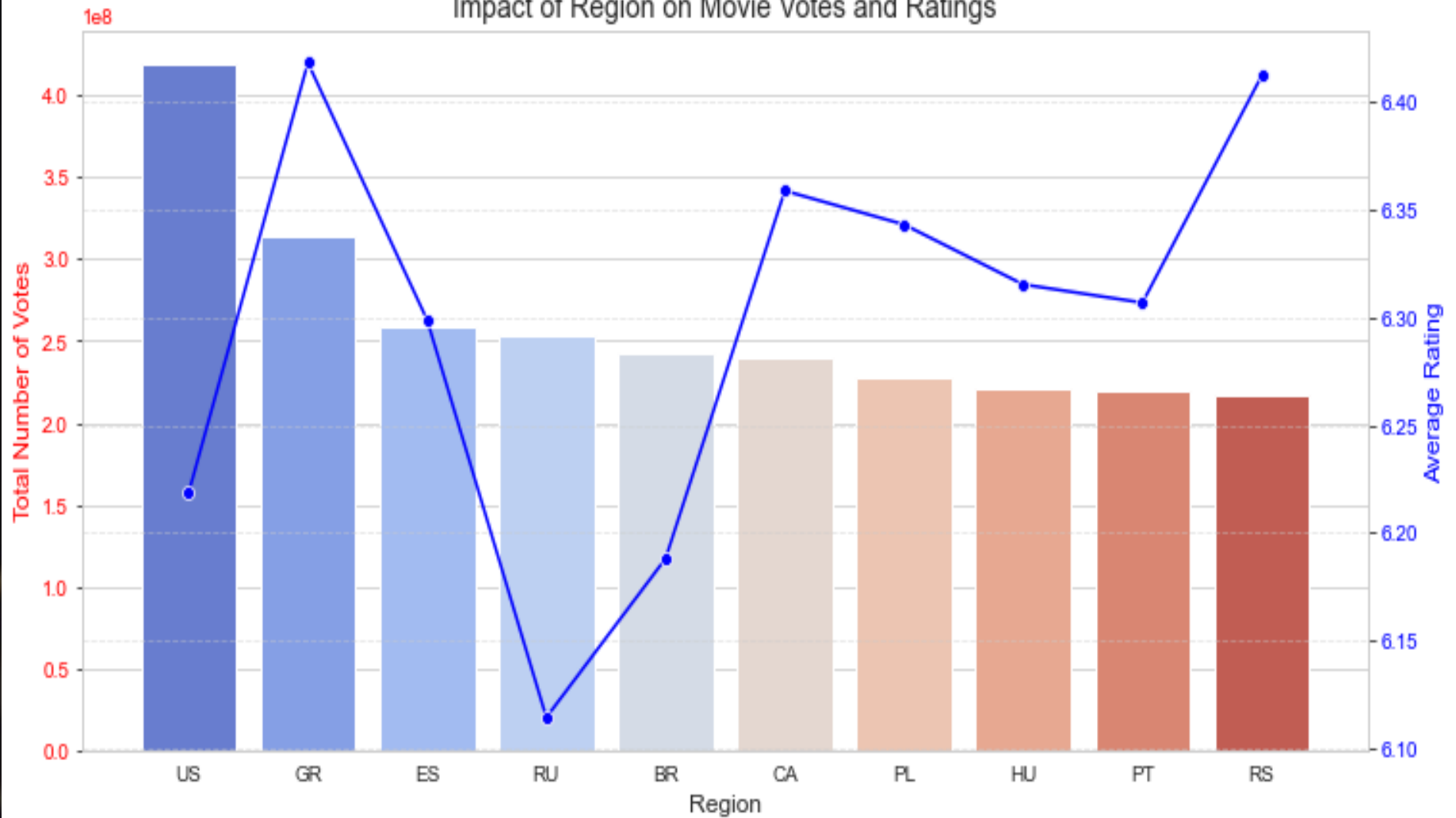
- Correlation of 0.11 between the total number of votes and average runtime which indicates a weak positive relationship.
- Correlation of -0.07 between the total number of votes and average rating which indicates a weak negative relationship
- Correlation of -0.10 between the average rating and average runtime which indicates a weak negative relationship.
- Generally, there is a weak correlation between movie runtime and number of votes and ratings.

Total Number Of Votes By Movie Language



- Movies in French language have the highest votes followed by Turkish and Bulgarian.
- We have movies with swedish language with the lowest number of votes.

Impact of Region on Movie Votes and Ratings



- From the top 10 Regions, United states has the highest number of votes and Average rating .
- Russia has the lowest average rating.
- Republic of Serbia has the lowest number of votes overall.

Key Findings

Genre Trends:

- Drama movies have the highest rating among the top ten genres with the highest number of votes, followed by Comedy, Drama, and Drama, Romance.
- The Action, Adventure, Sci-Fi genre has the highest number of votes, followed by Action, Adventure, Fantasy and Adventure, Animation, Comedy. These genres have a strong audience base, indicating their popularity in mainstream cinema.

Runtime and Popularity:

- There is a very weak positive correlation (0.10) between the total number of votes and average runtime, and very weak negative correlation (-0.07) between average rating and average runtime.
- This suggests that runtime of movies has little or no effect on ratings or number of votes.

Language Preference:

- Movies in the French language have received the highest number of votes (231,092,513), followed by Turkish and Bulgarian films. This suggests that French-language films have a wider global reach or a highly engaged audience.

Regional Popularity:

- Movies from United States have the highest number of votes and ratings, reflecting its dominance in global cinema.
- Greece ranks second, indicating a strong movie-watching culture or high engagement from Greek audiences.

Conclusion & Next Steps

- **Data-driven decision-making is crucial for movie success.**
- **Future steps: Deeper analysis of budgets, revenues, and marketing strategies.**
- **Implement findings to maximize profitability and audience reach.**

Recommendations

1. Genre Selection for High Engagement

- Focus on producing or promoting Drama, Comedy, and Romance films, as they receive the highest ratings with high number of votes.
- Action, Adventure, Sci-Fi, Action, Adventure, Fantasy and Adventure, Animation, Comedy genres attract the most votes, which in turn means the highest global audience, making them ideal for blockbuster-style films targeting a broad audience.
- Consider blending elements of popular genres (e.g., Action-Drama or Sci-Fi-Romance) to appeal to both high-rating and high-vote segments.

2. Language and Market Expansion

- Given the high number of votes for French, Turkish, and Bulgarian films, production companies should consider dubbing or subtitling in these languages to increase global reach.
- Investing in regional language content can help tap into underrepresented but highly engaged audiences.

3. Target High-Engagement Markets

- The United States leads in votes, making it a key market for movie releases.



THANK YOU !