Movie Ratings Analysis

Group 5

Data Description

▶Before we embark on our journey of movie ratings analysis, let's take a moment to understand the data that forms the backbone of our exploration. The MovieLens dataset contains 1,000,209 anonymous ratings of approximately 3,900 movies, made by 6,040 MovieLens users who joined in 2000.

▶**Ratings File Description:** The 'ratings.dat' file contains all ratings, with details on UserIDs, MovieIDs, the 5-star scale ratings, and timestamps. Each user has provided at least 20 ratings.

▶**Users File Description:** Demographic information can be found in the 'users.dat' file, including Gender, Age, Occupation, and Zip-code. Users provided this information voluntarily, and only those with demographic details are included in the dataset.

▶**Movies File Description:** Movie information is stored in the 'movies.dat' file, complete with MovieIDs, titles matching IMDB entries, and genres separated by pipes. There are various genres available for categorization, from Action and Comedy to Film-Noir and Sci-Fi.
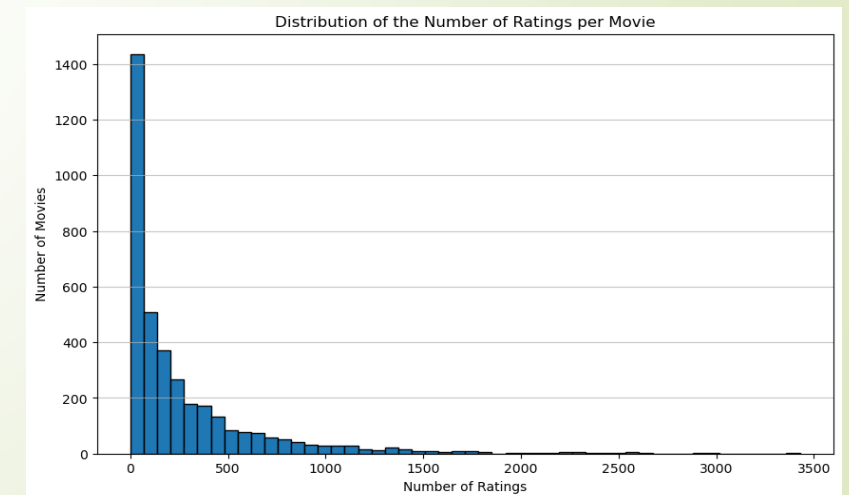
Motivation for Movie Ratings Analysis:

➡️In the era of data-driven decisions, understanding user preferences and behaviour is critical for businesses, especially in the entertainment industry.

➡️This topic is fascinating because it allows us to uncover valuable insights about how various factors such as occupation, gender, age groups, and movie genres impact movie ratings.

➡️By analysing this data, we can make more informed decisions about content creation, marketing strategies, and audience engagement.

➡️Understanding user preferences in the realm of movie ratings is like having a window into the minds of our audience. It's not only academically intriguing but also has real-world applications that can shape the future of the entertainment industry.

➡️This analysis enables us to create content that resonates with our audience, enhances their viewing experience, and ultimately drives success in the highly competitive world of entertainment.

Communicating Results with Figures and Tables

▶Distribution of Movie Ratings: To give you a clear picture of how users rate movies, we've created a histogram showing the distribution of ratings. Most ratings are concentrated around 3.0 to 4.0, indicating that users tend to give moderate to high ratings to movies. This figure helps us understand the overall sentiment of our audience.
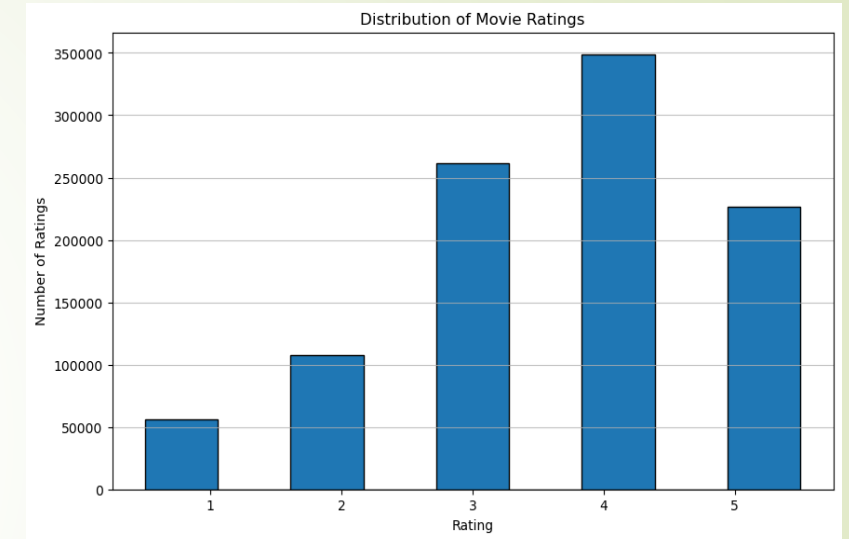
▶Distribution of the Number of Ratings per Movie: For content creators and decision-makers, it's important to know how frequently movies are rated. We illustrate this with a histogram that showcases the distribution of the number of ratings per movie. You'll see that most movies have a relatively small number of ratings, while a few have received a substantial number.

Distribution of Movie Ratings

Distribution of the Number of Ratings per Movie

Distribution of Average Ratings per Movie

►Distribution of Average Ratings per Movie: Another aspect of our analysis is the diversity in movie ratings. We present a histogram that highlights the distribution of average ratings per movie. This information is vital for understanding how different movies are received by our audience.

►Analysis of Popular Movies: For a deeper dive into how the number of ratings impacts movie ratings, we've separated popular movies (those with 100 or more total ratings) from the entire dataset. The distribution of average ratings for popular movies differs significantly from all movies. Popular movies tend to have a more concentrated distribution of higher average ratings, suggesting that movies with a substantial number of ratings are more likely to receive consistent and favorable reviews.

Average Ratings for All Movies

Average Ratings for Movies Rated > 100 Times
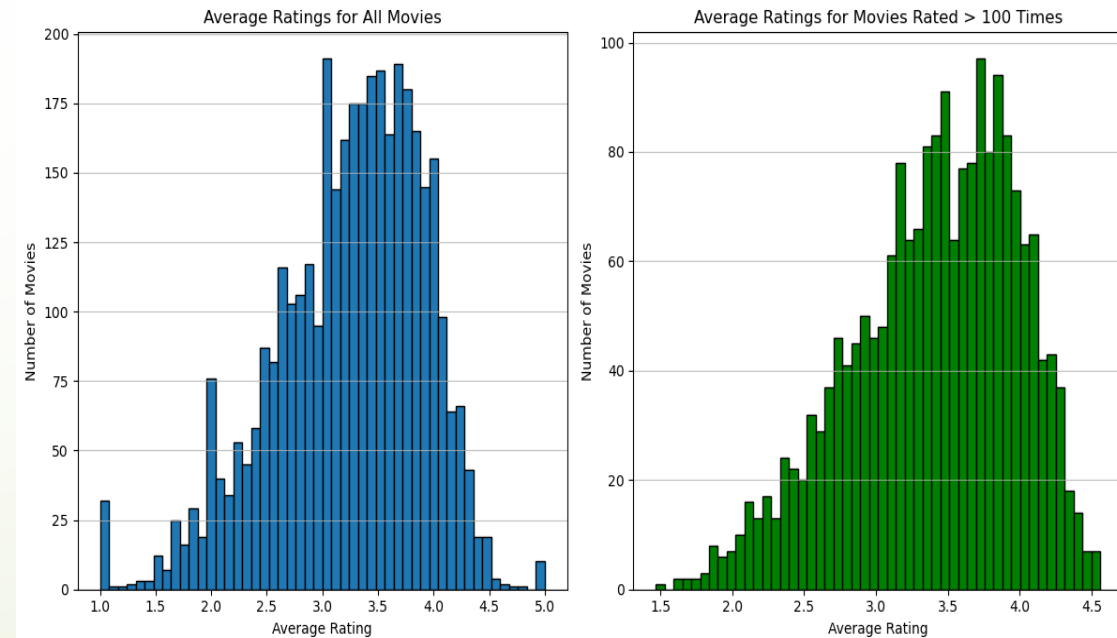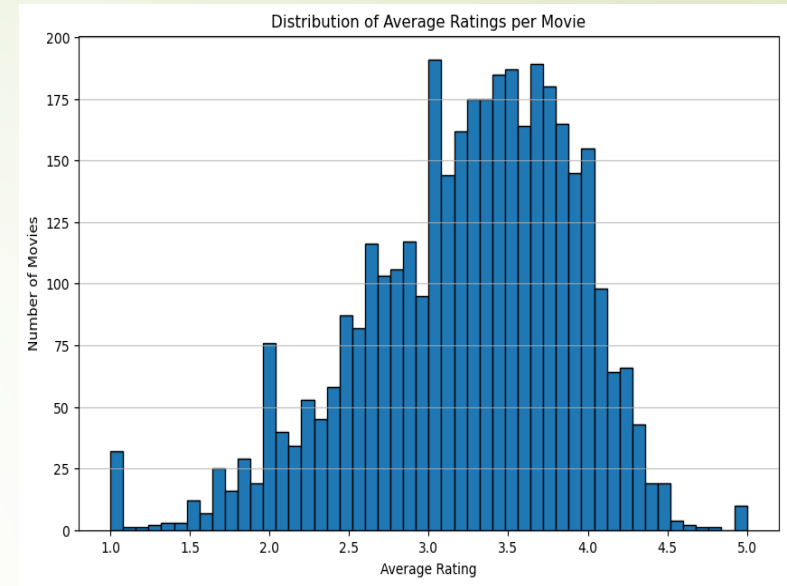
Storytelling

◆ Our story begins with the MovieLens dataset, a treasure trove of movie ratings. This raw data forms the foundation of our narrative. It contains a diverse range of information, from user demographics and movie genres to occupation and age groups.

◆ The heart of our story lies in the analysis phase. Here, we apply statistical techniques, visualization, and conjecture testing to delve deeper into the data. We ask questions, make hypotheses, and seek evidence to understand how different variables influence movie ratings.

◆ As our analysis unfolds, we uncover insights about how occupation, gender, age groups, and movie genres impact ratings. These results provide the plot twists, the 'aha' moments, and the fascinating discoveries in our data story.

Business Insights

In the realm of the entertainment industry, understanding user preferences is a strategic imperative. Our journey into movie ratings analysis has unearthed a wealth of insights with profound implications for businesses in this field.

▶**Genre Preferences Influence Content Creation:** Our analysis has revealed that specific age groups exhibit distinct genre preferences. For instance, younger viewers tend to favour certain genres while middle-aged individuals have different inclinations. This insight is a golden nugget for content creators as it guides them in tailoring content to cater to specific audience segments.

▶**Age and Rating Extremes:** It's fascinating to note that young adults in their 20s and 30s are more likely to give extreme ratings (1 or 5) due to their passion for movies. This passion becomes a valuable aspect to consider when creating and marketing movies for this demographic.

▶**Children's Preferences:** Children, in the age group 1-10, are more likely to rate movies with a 5. Their less critical and more enjoyment-oriented approach makes them a target audience for family-friendly and delightful content.

• **Occupation Matters:** Our analysis also highlights the influence of occupation on rating patterns. For example, 'critics' or 'students' tend to be more critical and harder to please, while other occupations might be more lenient.

• **Gender-Based Insights:** Gender differences play a crucial role in movie ratings. Our data suggests that women tend to be more critical when it comes to movie ratings compared to men. This knowledge can shape promotional tactics and content promotion strategies.

• **Popular Movies and Consistent Ratings:** When we focused on popular movies, we discovered that they tend to have a more concentrated distribution of higher average ratings. This insight can influence content creators' decisions regarding which movies to promote more aggressively.

• **Genre Impact on Ratings:** The genre of a movie also plays a significant role in its ratings. Some genres receive higher average ratings than others. This finding can influence the type of content produced and recommended to users.
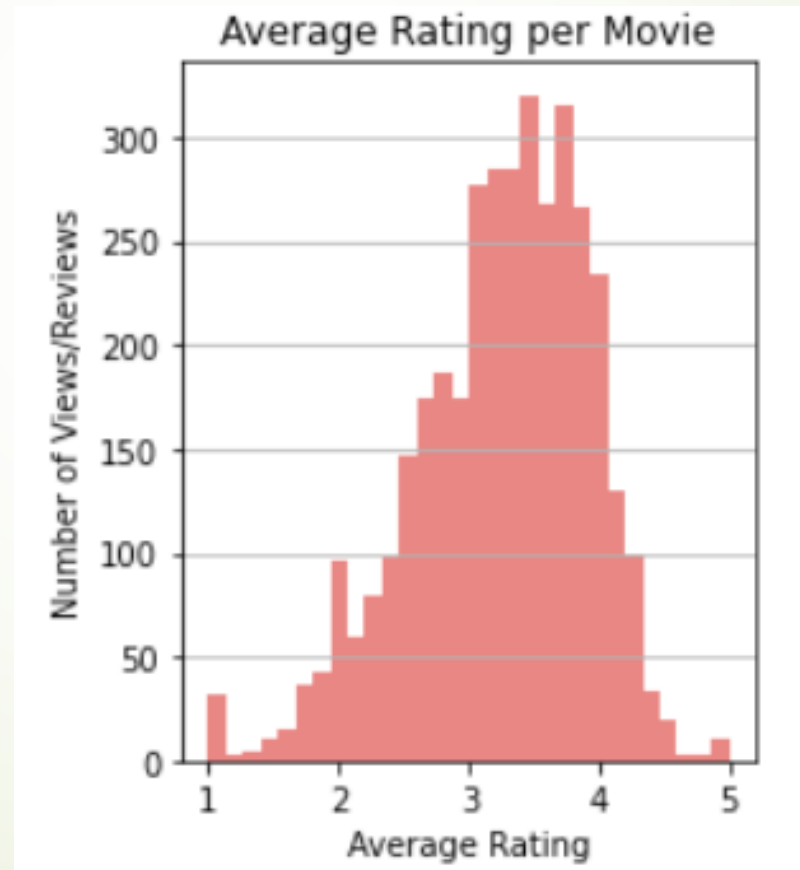
Potential Business Questions

- Do movies get more viewers only if they get a good rating?

- Do certain genres of movies tend to get more viewers than others?

- Do movies that belong to more than one genre get more viewers than movies that stick to one particular genre?

- What is the age group of the majority of the viewers/reviewers?

The Big Question:

- Should we direct our budget towards getting a critically acclaimed director and film crew or an extensive marketing campaign?
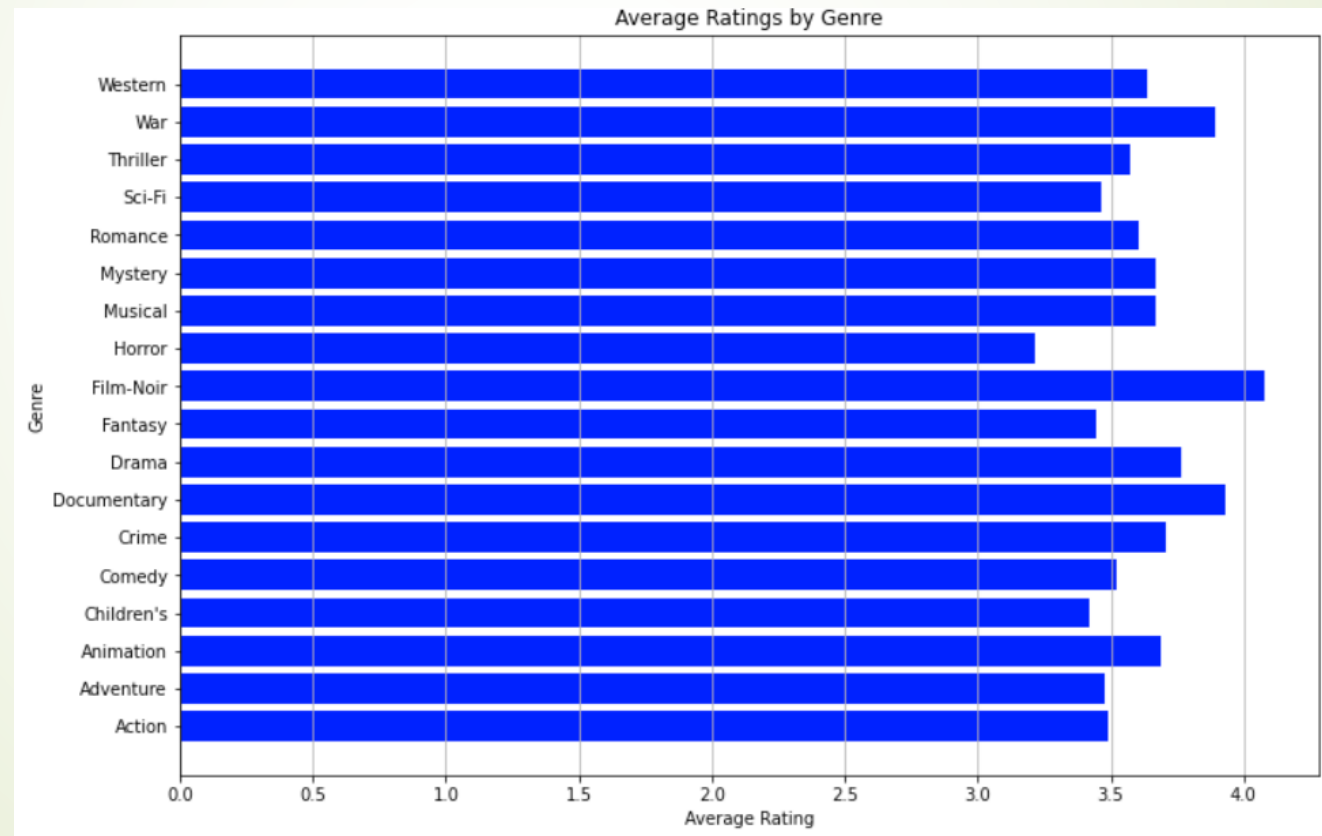
Potential Business Questions

▶ Do movies get more viewers only if they get a good rating?

Potential Business Questions

- Do certain genres of movies tend to get more viewers than others?

Average Ratings by Genre

Potential Business Questions

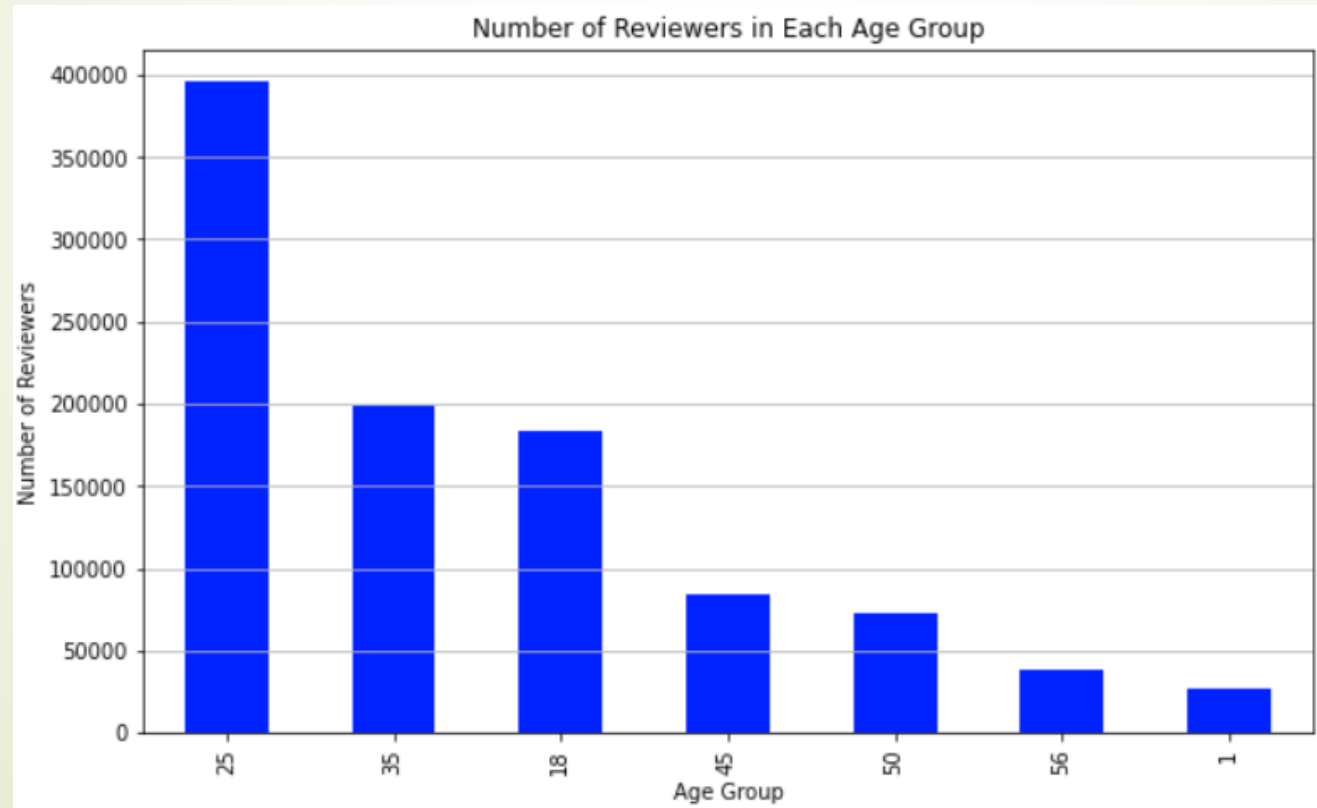- Do movies that belong to more than one genre get better ratings than movies that stick to one particular genre?

```
Average ratings by the number of genres:
num_genres
1      3.579876
2      3.585568
3      3.572580
4      3.569970
5      3.723008
6      3.380952
```
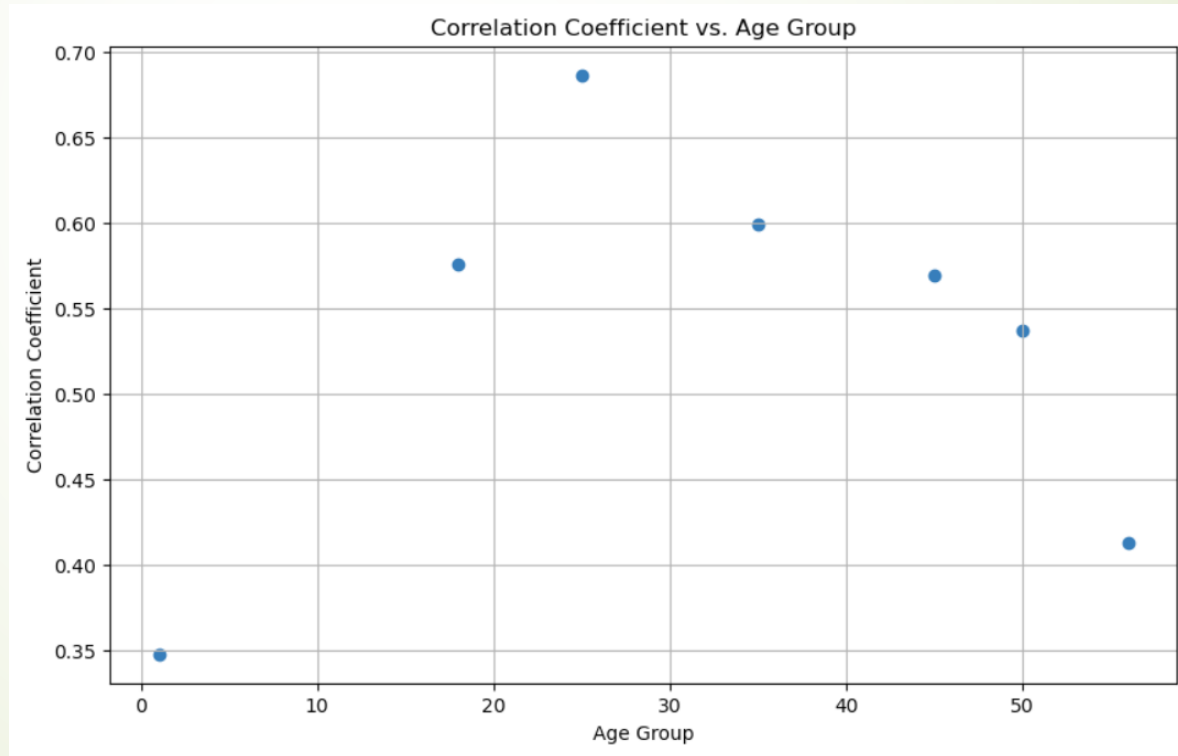
Potential Business Questions

What is the age group of the majority of the viewers/reviewers?
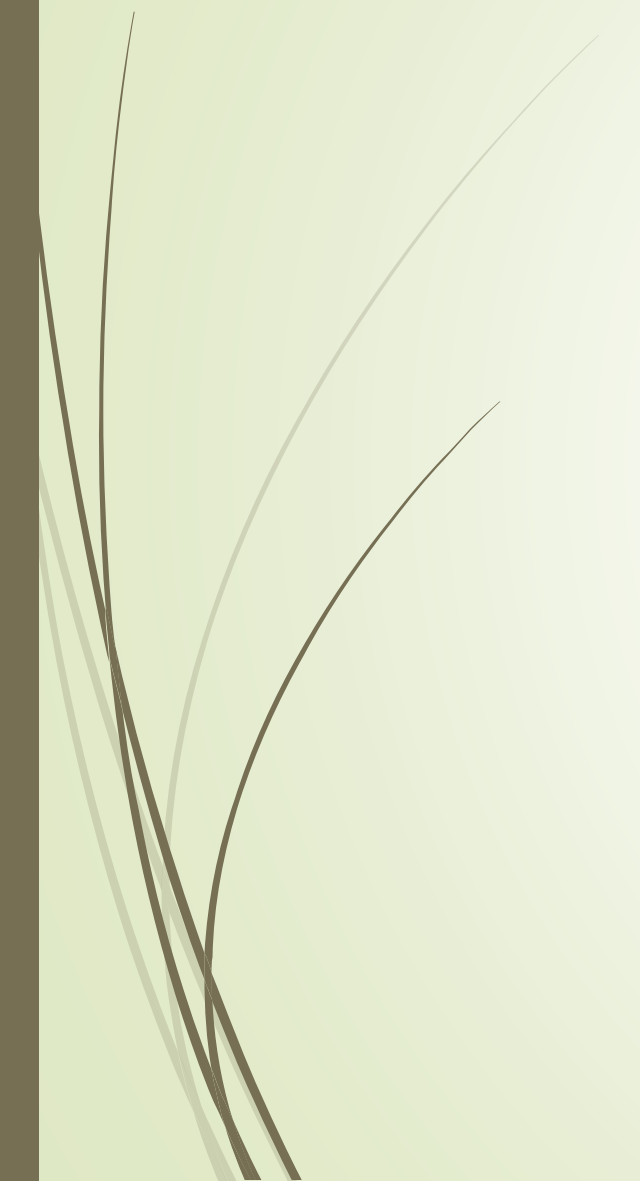
Potential Business Questions

➡ Do men and women of a particular age group have similar opinions/ratings for movies?

Key Assumptions

The main assumption in this analysis is that we are taking the number of reviewer as a representation of the number of viewers the movie has, as generally the more viewers the more reviewers a movie will get.

Potential Business Questions

The Big Question:

- Should we direct our budget towards getting a critically acclaimed director and film crew or an extensive marketing campaign?

- It was found that the rating of a movie does not always correlate to the number of viewers or reviewers. As long as the movie is fairly good (rating 3-4), it seems to be that it is more profitable to spend on good marketing than an extraordinarily good plot or direction to generate the most revenue.

- Based on the analysis, we can assume that in order to maximise the revenue of the next movie, we need to make a movie belonging to the War and Film-Noir genres and direct the majority of the marketing budget towards advertising the movie in mediums that have a stronger reach to the 25 and under age group.

Conclusion

▶ In conclusion, these business insights provide a roadmap for content creation, marketing, and audience engagement. They enable businesses in the entertainment industry to make informed decisions that resonate with their audience, enhance user experiences, and drive success in a highly competitive landscape.

▶ These insights are not just intriguing; they are actionable, providing a strategic advantage in the dynamic world of entertainment.

THANK YOU!