Case study 3- Group 10

What is the relationship between this topic and Business Intelligence?

Analysing the MovieLens dataset is a pivotal activity that aligns with core Business Intelligence (BI) objectives. This dataset is a compilation of movie ratings that, when examined thoroughly, provides deep insights into consumer preferences - a fundamental aspect for any content-driven enterprise.

In the BI context, the MovieLens data serves as a prime example of how granular consumer data can be leveraged. It aids in discerning intricate patterns in user behaviour and preferences. Such analysis is crucial for content providers to customise offerings and enhance user engagement through sophisticated recommendation engines.

Moreover, this dataset is instrumental in discerning broader market dynamics, enabling entertainment businesses to adapt their strategies in line with evolving trends. The predictive capabilities derived from this analysis are invaluable for forward-planning, guiding investment in content genres that are likely to resonate with audiences.

Additionally, by dissecting the MovieLens data, companies can measure the impact of their content portfolios and refine their marketing tactics. It also supports the segmentation of the viewer base into distinct groups for targeted outreach, thereby optimising promotional efforts and maximising return on investment.

In essence, the insights derived from the MovieLens dataset through BI methodologies are integral to decision-making processes, ensuring that businesses stay aligned with consumer demand and maintain a competitive stance in the marketplace.

What business decision do you think this data could help answer? Why?

The movie rating insights across demographic groups could inform several strategic business decisions:

- Content and Marketing - Knowing which types of movies perform well with different audiences can guide what content is acquired and how it is promoted. For example, the data shows classic dramas are favorited by women over 30, so a streaming platform may licence more of those films to appeal to that segment.
- Recommendation Algorithms - Rating biases by demographic can be incorporated to improve recommendation systems. If older viewers consistently rate higher, their ratings could be adjusted to account for this bias.
- Targeted Advertising - Marketing campaigns can be designed targeting specific groups based on the movie preferences observed in the data. This allows for more effective ad spending tailored to key segments.
- Distribution Priorities - Film distributors could leverage the data to determine which movies to prioritise for promotion and distribution to different regions/ages, improving box office revenues.

Findings and Conjectures (Part1):

What are the ten most popular movies?

```
Popularity by number of ratings:

title
American Beauty (1999)                                      3428
Star Wars: Episode IV - A New Hope (1977)                   2991
Star Wars: Episode V - The Empire Strikes Back (1980)       2990
Star Wars: Episode VI - Return of the Jedi (1983)           2883
Jurassic Park (1993)                                        2672
Saving Private Ryan (1998)                                  2653
Terminator 2: Judgment Day (1991)                           2649
Matrix, The (1999)                                          2590
Back to the Future (1985)                                   2583
Silence of the Lambs, The (1991)                            2578
Name: rating, dtype: int64
```

I personally believe that the popularity of a movie can be determined by the total interaction users have with the movie, and in this analysis, I've chosen the total number of ratings as the parameter to gauge a movie's popularity. This assumption relies on the presumption that the data was collected randomly, and that ratings reflect the collective preferences of a diverse group of users

Conjecture about Gender:

The data indicates that women tend to give higher movie ratings than men. Specifically, there are more movies with an average rating above 4.5 from women (51 movies) compared to men (23 movies). A similar pattern emerges when looking at median ratings from those over 30 years old - significantly more movies have a median rating over 4.5 from women (149 movies) versus men (86 movies) in this age group. One interpretation of this discrepancy is that women may be more inclined to give very positive ratings or that their preferences in movies may be easier to satisfy based on this dataset. However, more investigation would be needed to definitively conclude why this difference in rating behaviour between genders exists.

In summary, the core ideas expressed are:

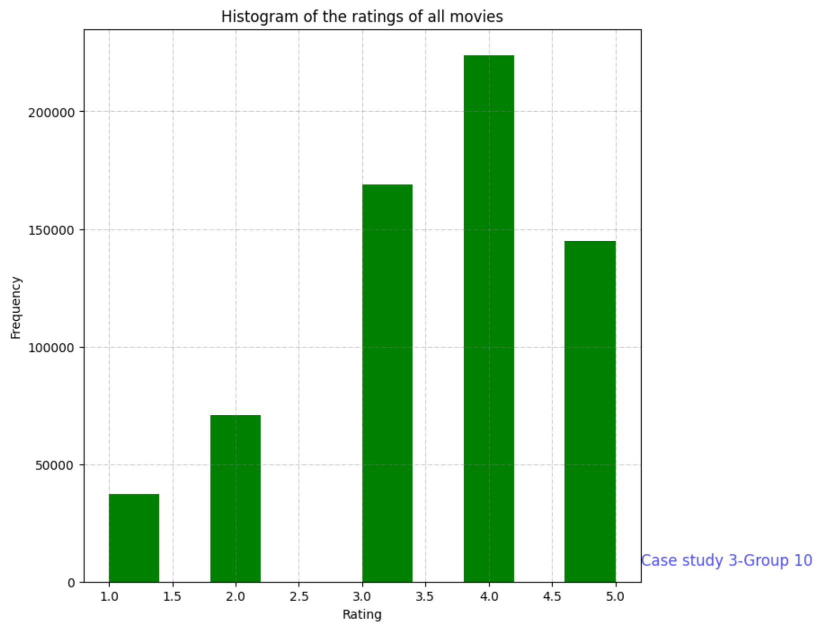
- Women give higher average and median movie ratings compared to men based on the data
- More movies have high average or median ratings (over 4.5) from women than men

- A hypothesis raised is women may rate more positively or have different preferences that are easier to satisfy
- More analysis would be required to fully understand the reasons behind the observed rating difference by gender

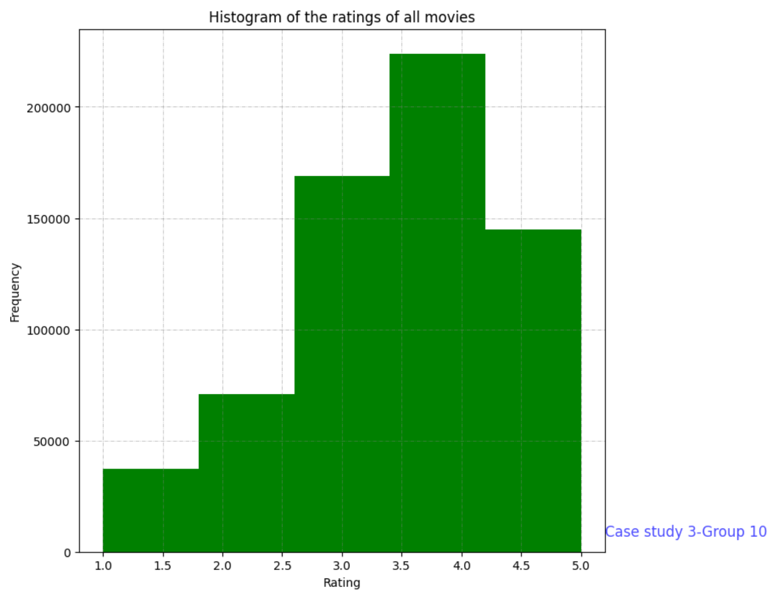
Findings and Conjectures (Part2):

Plot a histogram of the ratings of all movies.

There exist 5 distinct count of rating values. We select 10 and 5 as bin sizes for different visuals. The x-axis has the ratings while the y axis has the count corresponding to each rating.



Here, we've set the bin size as 10.

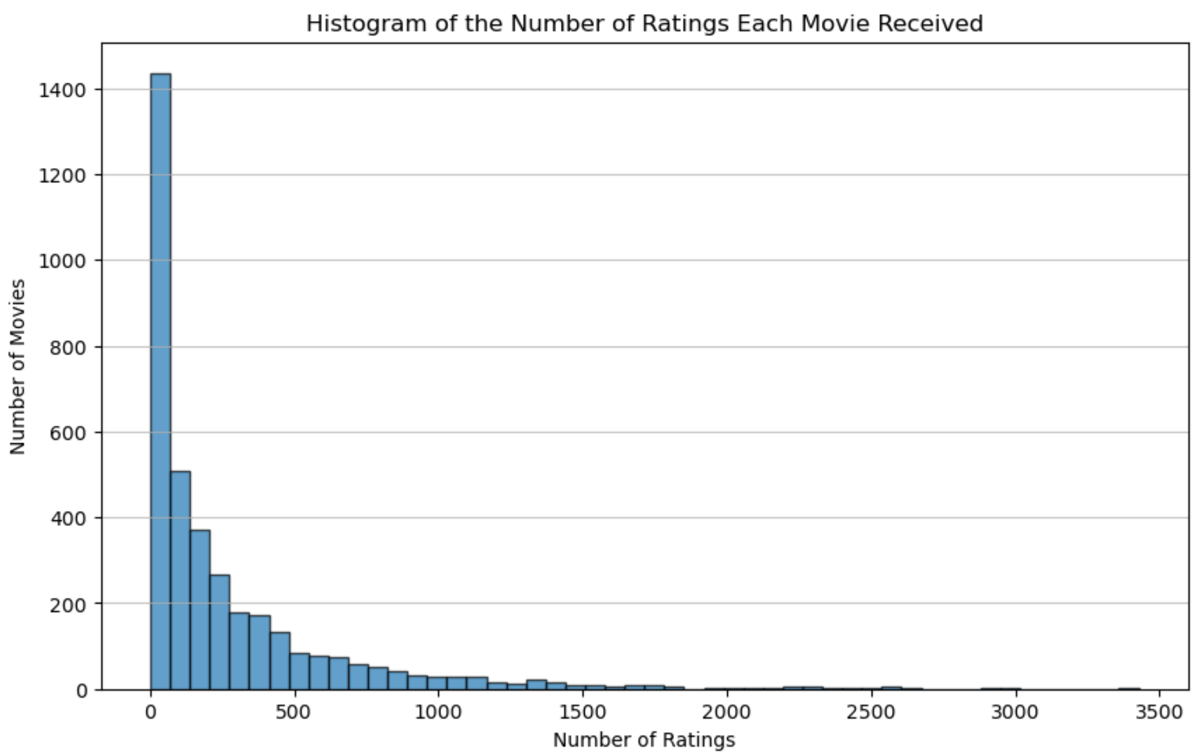


Here, the bin size is 10

We observe 3 and 4 to have maximum count, i.e people are less likely to give extreme ratings

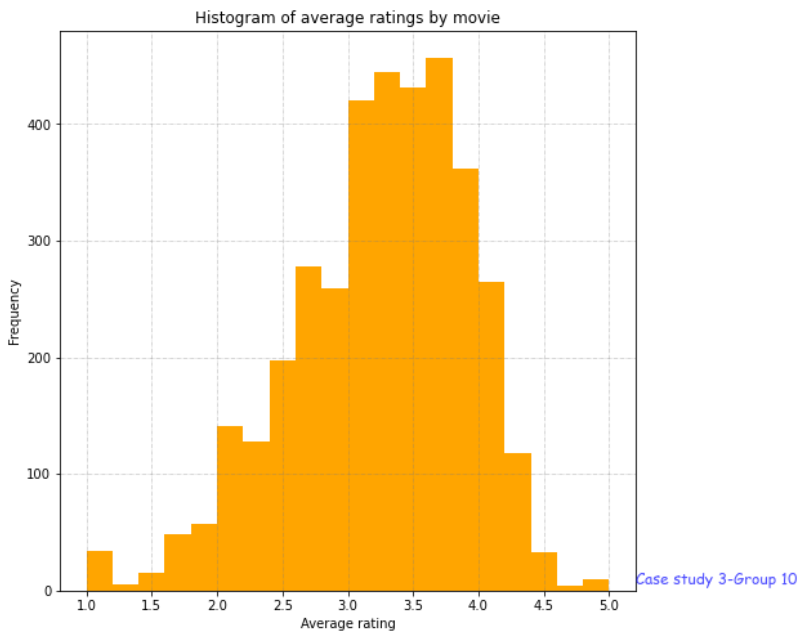
Histogram of the *number* of ratings each movie received.

The distribution of how many ratings each movie received, shown in the histogram, demonstrates that most movies have just a small number of ratings. As the count of ratings per movie increases, the number of movies with that rating count drops off sharply. This pattern indicates there is a large quantity of movies that have only a few ratings, and progressively fewer movies have higher and higher counts of ratings. In other words, the histogram reveals that movies receiving many ratings are far less common compared to the predominant number of movies with low ratings counts. The data is heavily skewed towards movies having small rating set sizes rather than large ones.



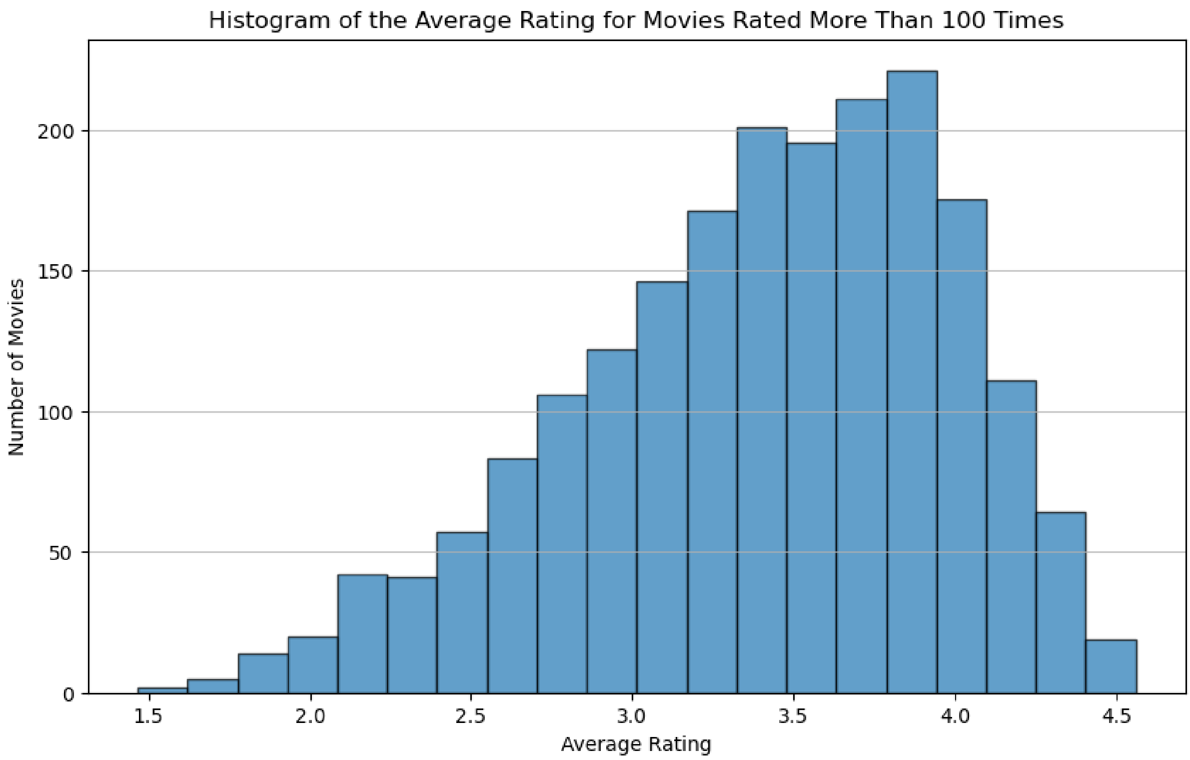
Histogram of the *average rating* for each movie.

We've grouped the data by movie and computed the average rating for each movie, post which we plot the same on a histogram with average ratings on the x axis and the frequency on the y axis (20 bins)



Histogram of the *average rating* for movies which are rated more than 100 times.

The procedure is similar to that mentioned above but here we get the count of ratings for each movie and only consider those with a count greater than 100.



On comparing tails of the above 2 histograms, we notice the first one (considering movies with less than a 100 ratings as well) has relatively higher frequency of extreme ratings, especially the low ones. This could be due to personal bias strongly affecting average movie ratings with lesser reviews to a greater extent.

I would trust highly rated movies with over a hundred ratings more than a movie with similar rating but less than 100 reviews, as the second case would be a smaller sample set, who could have tastes different from mine or even a matter of chance of that group liking the movie since the sample set is smaller. From the central limit theorem, we can conclude that the sample estimate gets closer to that of the population as the sample size increases, making it more trustworthy.

Some conjectures about the distribution of ratings

1. **What age range do you think has more extreme ratings? Do you think children are more or less likely to rate a movie 1 or 5?**

Conjecture- children are more likely to rate a movie 1 or 5

Proportion of children providing extreme ratings (1 or 5)= 0.3322185880710007

Proportion of adults providing extreme ratings (1 or 5)= 0.27844658767616426

Hence, we can conclude that children are more likely to rate a movie 1 or 5 (extremely) as a greater proportion of children in the dataset have extreme ratings. This might be due to less nuanced opinions as they have not fully matured.

2. **Conjecture- Not much difference in men and women**

We group the data by gender and compute the mean for the 2 groups.

The mean of ratings given by women is 3.620366

The mean of ratings given by men is 3.568879

Though there isn’t a significant difference between the ratings provided by women, the mean of ratings provided by women is slightly higher.

3. **Conjecture- average ratings go down with age as one would get more critical**

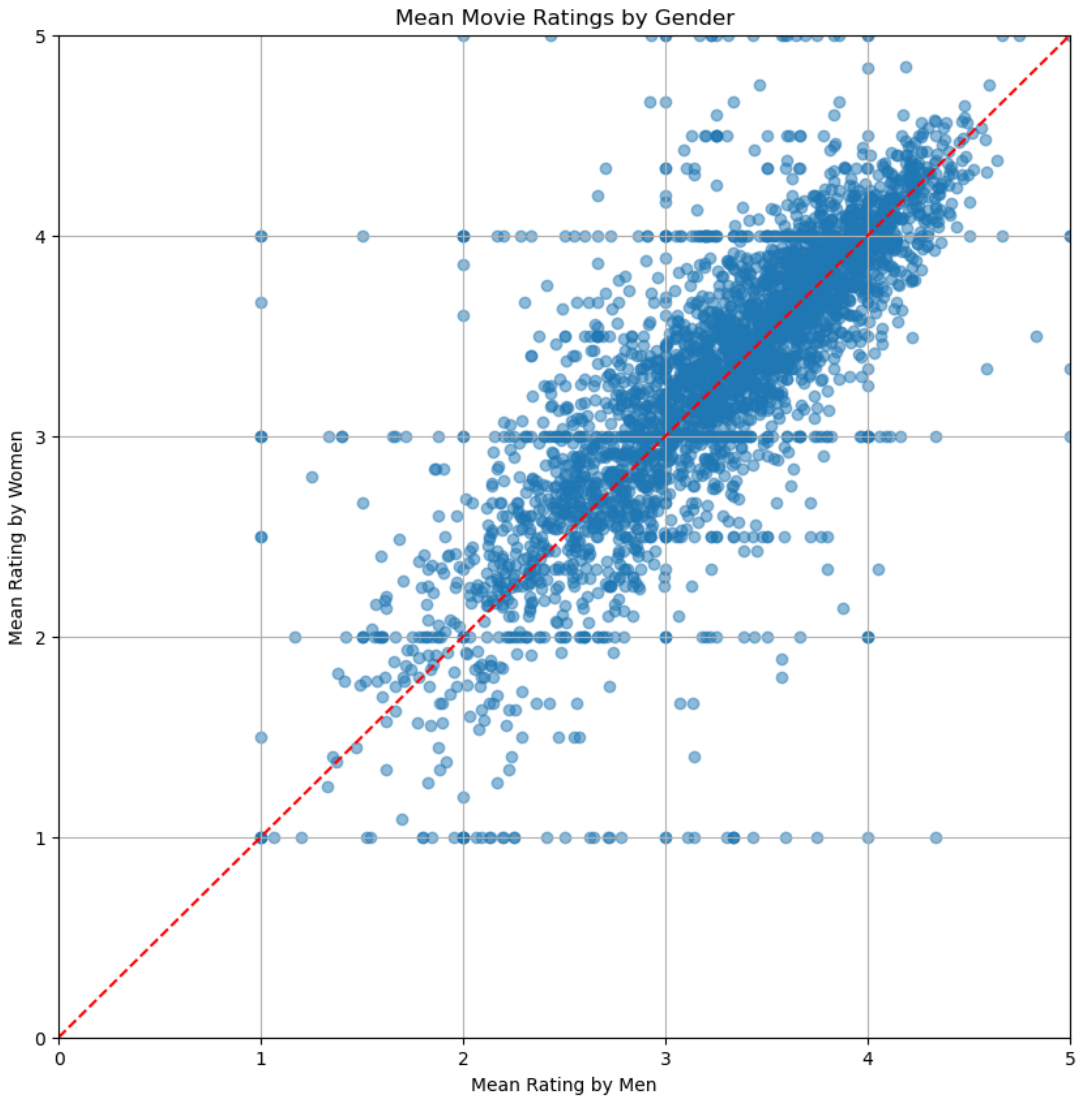
We group the data based on age (each value representing different age bins)

age	user_id	movie_id	rating	timestamp	occupation	count_of_ratings_for_movie	count_of_ratings_received_by_movie
1	2717.664437	1869.380875	3.549520	9.759793e+08	8.765977	821.032009	821.032009
18	2789.483818	1815.417967	3.507573	9.736918e+08	6.725596	840.614354	840.614354
25	3121.309319	1849.590119	3.545235	9.720568e+08	7.887353	823.190752	823.190752
35	3040.853475	1895.714251	3.618162	9.715593e+08	8.816138	809.906630	809.906630
45	3178.966138	1903.871127	3.638062	9.708210e+08	8.531321	781.893045	781.893045
50	3016.087543	1917.925962	3.714512	9.722278e+08	8.359098	783.051511	783.051511
56	2963.619082	1927.316194	3.766632	9.712856e+08	9.569856	795.117586	795.117586

The above conjecture is wrong, from the 7 age groups included in the dataset, we observe a gradual increase in ratings from younger to older (1 seems too young an age to consider, more prone to extreme ratings as seen earlier)

Findings and Conjectures (Part3):

Make a scatter plot of men versus women and their mean rating for every movie:

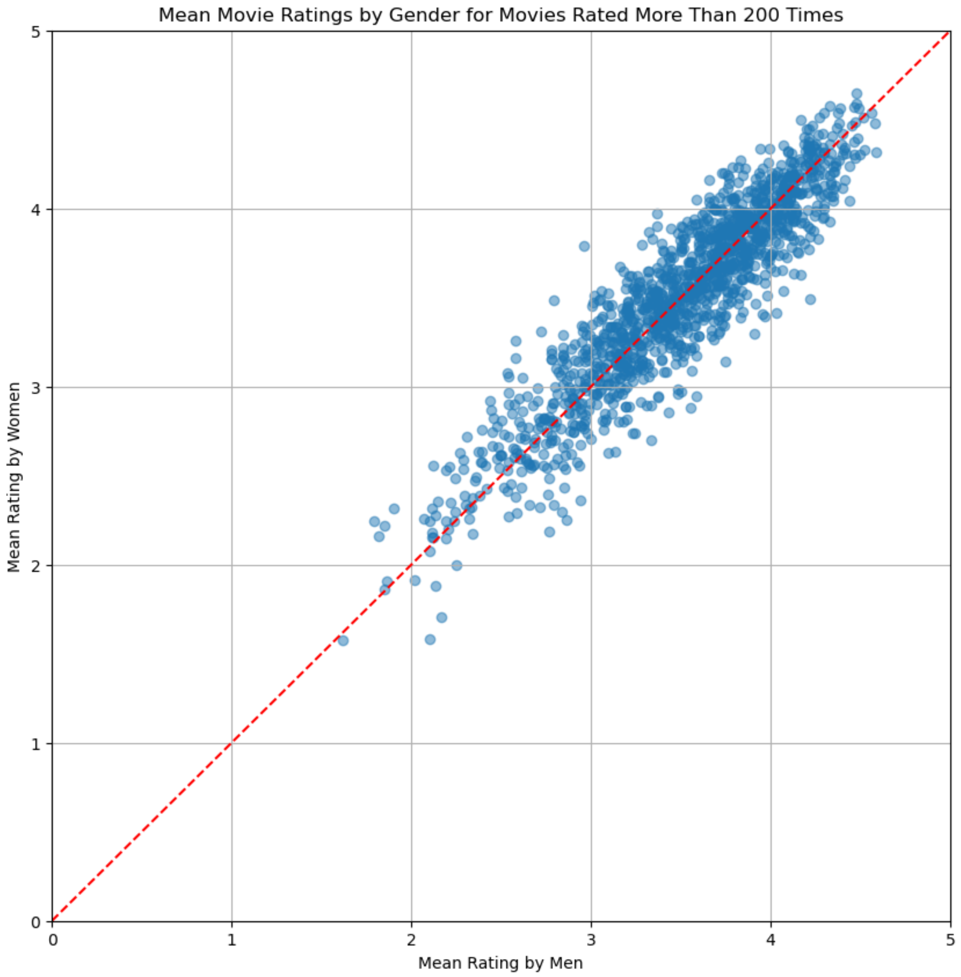


- The x-axis represents the mean ratings given by men.
- The y-axis represents the mean ratings given by women.
- The red dashed line indicates the line of equality (i.e., where ratings by men and women are equal).

Movies that lie on or close to this line have similar ratings from both men and women. Movies above the line are rated higher by women than by men, and movies below the line are rated higher by men than by women.

From a glance at the plot, many movies seem to receive comparable ratings from both genders, as many points are near the red line. However, there are some movies that have differing ratings between the genders.

Make a scatter plot of men versus women and their mean rating for movies rated more than 200 times.



- The x-axis represents the mean ratings given by men.
- The y-axis represents the mean ratings given by women.
- The red dashed line is the line of equality.

Movies that lie on or close to this line have similar ratings from both men and women. Movies above the line are rated higher by women than by men, and movies below the line are rated higher by men than by women.
By focusing on movies with a substantial number of ratings, the plot helps highlight trends in the preferences of more widely viewed movies.

Compute the *correlation coefficent* between the ratings of men and women.

- **What do you observe?**
- **Are the ratings similiar or not? Support your answer with data!**

For Movies >200 rated:

The Pearson's correlation coefficient between the ratings of men and women for movies rated more than 200 times is approximately 0.9180.918.

This value is close to 1, indicating a strong positive linear relationship between the ratings of men and women. In other words, when the ratings by men for a movie tend to be higher, the ratings by women also tend to be higher, and vice versa.
Observation: The ratings given by men and women for popular movies (rated more than 200 times) are quite similar, as evidenced by the high correlation coefficient. This suggests that for these movies, both genders generally agree in their assessment.
In summary, based on the data for popular movies, the ratings by men and women are closely aligned and exhibit a strong positive relationship.

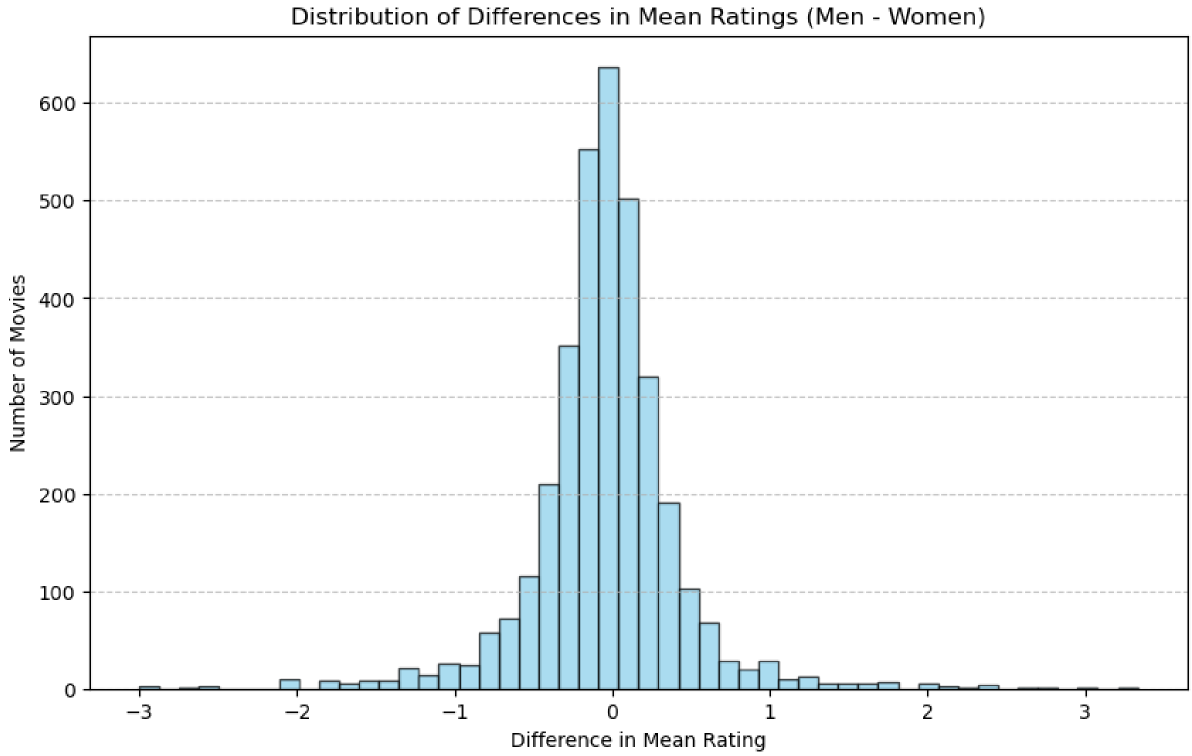
For all movies:

The Pearson's correlation coefficient between the ratings of men and women for all movies in the dataset is approximately 0.7630.763.
This value indicates a strong positive linear relationship between the ratings of men and women, albeit somewhat weaker than the correlation for just the popular movies.
Observation: The ratings given by men and women across all movies are quite similar, as suggested by the correlation coefficient of 0.7630.763. This value implies that there's a general agreement between the two genders in their assessment of movies. While the relationship isn't as strong as the one observed for popular movies (rated more than 200 times), it's still significant.
In conclusion, based on the data for all movies, the ratings by men and women are closely aligned and exhibit a strong positive relationship.

What do you observe?
Are the ratings similar or not? Support your answer with data!

1. **Correlation Coefficient**: The correlation coefficient of 0.7630.763 is close to 1, indicating a strong positive relationship. This means that when men rate a movie highly, women tend to rate it highly as well, and vice versa.
2. **Scatter Plot Analysis**: When we plotted the mean ratings of men against women for all movies, many of the data points were close to the line of equality (the red dashed line). This indicates that for a majority of movies, the mean ratings provided by both genders are closely aligned. Movies far from the line would represent significant differences in rating between the genders, but such movies were relatively few.
3. **Distribution of Differences**: Another way to gauge similarity is to look at the distribution of the differences in mean ratings between the genders. Let's compute the differences and visualize them using a histogram.

Observations from the Histogram:



1. **Centered around Zero**: A significant number of movies have differences close to zero, indicating that for these movies, the mean ratings given by men and women are almost identical.
2. **Symmetry**: The distribution appears roughly symmetric, meaning that there are similar numbers of movies where men rate slightly higher as there are where women rate slightly higher.
3. **Few Extreme Differences**: There are fewer movies as we move further from the center of the distribution. This means that large discrepancies in ratings between genders (either positive or negative) are rarer.

In summary, the data supports the conclusion that the ratings given by men and women are quite similar for a majority of movies. While there are some movies with noticeable differences in ratings between the genders, they are in the minority.

Conjectures:

1. **Age**: younger or older individuals from both genders might have more similar movie preferences. For instance, younger audiences might have more alignment in their tastes due to shared cultural experiences, while older audiences might diverge based on life experiences.

2. **Occupation**: People in similar professions might have comparable tastes regardless of gender, e.g., artists, scientists, or educators might value certain movie themes or genres similarly.
3. **Genres**: Some genres might show more alignment between genders in terms of ratings, while others might show divergence. For instance, romantic movies might (stereotypically) be assumed to be rated higher by women than men, while action movies might have the opposite trend.
4. **Rating Extremes**: Highly rated or poorly rated movies might show more agreement between genders. If a movie is universally acclaimed or panned, both genders might agree on its quality.

5. **Older movies (1980 or below) tend to be liked or disliked similarly in both genders.**

6. **People who live in the same state tend to have similar taste**

1. **Age:**

Here are the computed Pearson's correlation coefficients for the mean ratings of movies by men and women within different age brackets:

- **Under 18**: 0.3480.348
- **18-24**: 0.5760.576
- **25-34**: 0.6860.686
- **35-44**: 0.5990.599
- **45-49**: 0.5690.569
- **50-55**: 0.5370.537
- **56+**: 0.4130.413

Observations from Age Analysis:

1. The age group **25-34** shows the strongest alignment in movie preferences between men and women, with a correlation of 0.6860.686. This suggests that men and women in this age bracket have the most similar movie tastes.
2. The correlation is lowest for the **Under 18** and **56+** age brackets, suggesting that younger and older audiences might have more divergent movie preferences based on gender.
3. Overall, there seems to be a trend where the correlation is strongest in the mid-age ranges and weaker at the extremes.

Based on this data, one could conjecture that the ratings given by one gender could be more reliably used to predict the ratings given by the other gender for individuals in the age bracket of 25-34. On the other hand, this predictive power might be weaker for the youngest and oldest age brackets.

Occupation:

Here are the computed Pearson's correlation coefficients for the mean ratings of movies by men and women within different occupations:

- **Other**: 0.5790.579
- **Academic/Educator**: 0.6360.636
- **Artist**: 0.4720.472
- **Clerical/Admin**: 0.4390.439
- **College/Grad Student**: 0.5730.573
- **Customer Service**: 0.3300.330
- **Doctor/Health Care**: 0.5180.518
- **Executive/Managerial**: 0.5730.573
- **Farmer**: 0.2750.275
- **Homemaker**: 0.2770.277
- **K-12 Student**: 0.3310.331
- **Lawyer**: 0.3940.394
- **Programmer**: 0.4500.450
- **Retired**: 0.2940.294
- **Sales/Marketing**: 0.5340.534
- **Scientist**: 0.4800.480
- **Self-Employed**: 0.4690.469
- **Technician/Engineer**: 0.5790.579
- **Tradesman/Craftsman**: 0.2770.277
- **Unemployed**: 0.4080.408
- **Writer**: 0.6070.607

Observations from Occupation Analysis:

1. **Highest Correlation**: Academic/Educator (0.6360.636) and Writer (0.6070.607) occupations have the highest correlations, suggesting that men and women in these professions have quite similar movie tastes.
2. **Lowest Correlation**: Occupations like Farmer (0.2750.275), Homemaker (0.2770.277), and Tradesman/Craftsman (0.2770.277) have the lowest correlations, indicating more divergence in movie preferences based on gender within these professions.

Given this data, one could conjecture that in professions related to academia, education, or writing, the ratings given by one gender could be more reliably used to predict the ratings given by the other gender. However, in professions like farming or homemaking, there might be more variance in movie preferences between genders.

Observations from Genre Analysis:

1. **Largest Differences**: The genres "Children's" and "Musical" have the most significant differences in mean ratings between genders, with women rating them higher than men.
2. **Smallest Differences**: Genres such as "Drama", "War", "Action", and "Thriller" have almost negligible differences in mean ratings between men and women, indicating strong alignment in preferences for these genres across genders.
3. **Genres with Women Rating Higher**: Women tend to rate genres like "Children's", "Musical", "Romance", and "Fantasy" higher than men.
4. **Genres with Men Rating Higher**: Men tend to rate genres like "Western", "Film-Noir", "Crime", and "Sci-Fi" slightly higher than women.

From this analysis, one could conjecture:

- For genres like "Drama", "War", "Action", and "Thriller", the ratings given by one gender might be a good predictor for the ratings given by the other gender, as both genders rate these genres similarly.
- For genres like "Children's" and "Musical", one might expect more variance between genders in terms of movie preferences.

Conjecture: Women and Men have similar rating when it comes to very bad movie or very good movie (1,5)

Based on the analysis, when a man rates a movie as "1" (very low), there is approximately a 5.48% chance that a woman will also rate that movie as "1".

This suggests that a very low rating by a man does not strongly indicate a similar low rating by a woman. The likelihood is relatively low, indicating that the extreme ratings are not strongly correlated across genders on an individual rating basis.

Based on the analysis, when a man rates a movie as "5" (very high), there is approximately a 24.00% chance that a woman will also rate that movie as "5".

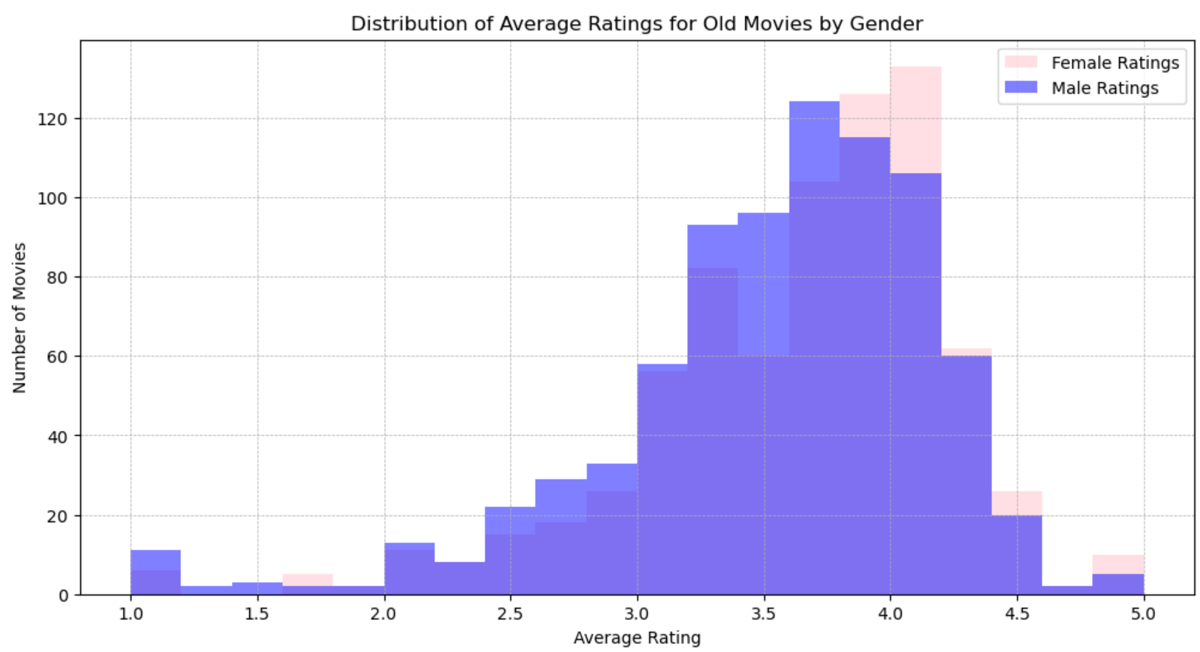
This suggests that a very high rating by a man somewhat increases the likelihood of a similar high rating by a woman compared to the low rating scenario. However, even with this increased likelihood, it's not a strong indication, as the proportion is still below 25%.

Conjecture: Older movies (1980 or below) are tend to be liked or disliked similarly in both genders.

For movies released before 1980 (which we're categorizing as "old"):

- The average rating given by females is approximately 3.643.64.
- The average rating given by males is approximately 3.543.54.

These ratings are moderately high, suggesting that both genders tend to like old movies. The ratings are also quite close, indicating that both genders have similar preferences for old movies.



The histogram displays the distribution of average ratings given by females and males for old movies (those released before 1980). Here's what we can infer:

- Both male and female raters have a spread of ratings for old movies, but there's a noticeable concentration towards higher ratings.
- The distributions are quite similar for both genders, suggesting that both genders have similar preferences when rating old movies.
- There isn't a clear concentration of ratings at the very low end, indicating that old movies are not generally disliked by either gender.

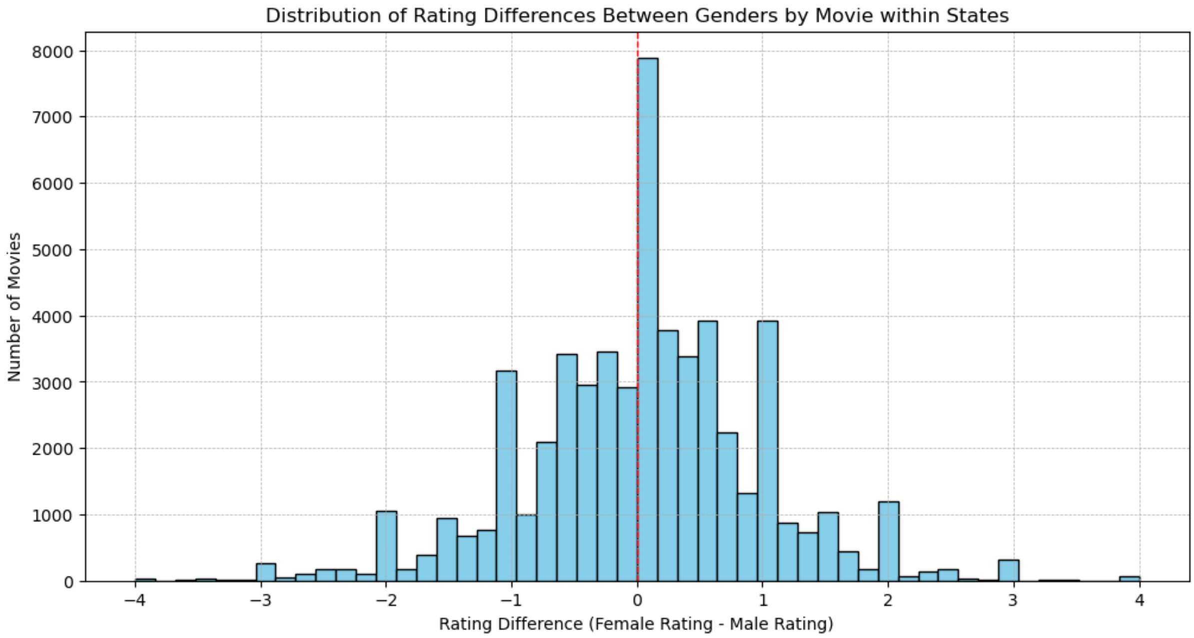
In summary, old movies tend to be liked by both males and females, with a similar distribution of ratings across genders.

Conjecture: Men and Women have similar preference when it comes to rating movies if they live in the same state

Analysis of the differences in movie ratings between men and women within states:

- **Mean Difference**: The average difference in ratings is approximately 0.0370.037. This suggests that, on average, women rate movies only slightly higher than men.
- **Standard Deviation**: The standard deviation is 0.9530.953, indicating that the differences in ratings can vary significantly across different movies and states.
- **Minimum Difference**: The most negative difference is −4−4, suggesting that there are some movies that men rate much higher than women.
- **Maximum Difference**: The maximum difference is 44, indicating that there are also movies that women rate much higher than men.
- **Median Difference**: The median difference is 00, suggesting that for many movies, the average rating by men and women is the same.

Given that the mean and median differences are close to 0, and considering the large standard deviation, it seems that while there are differences in how men and women rate movies within states, on average, the differences are minimal.



The histogram provides a visual representation of the distribution of rating differences between men and women for movies within states.

Key observations:

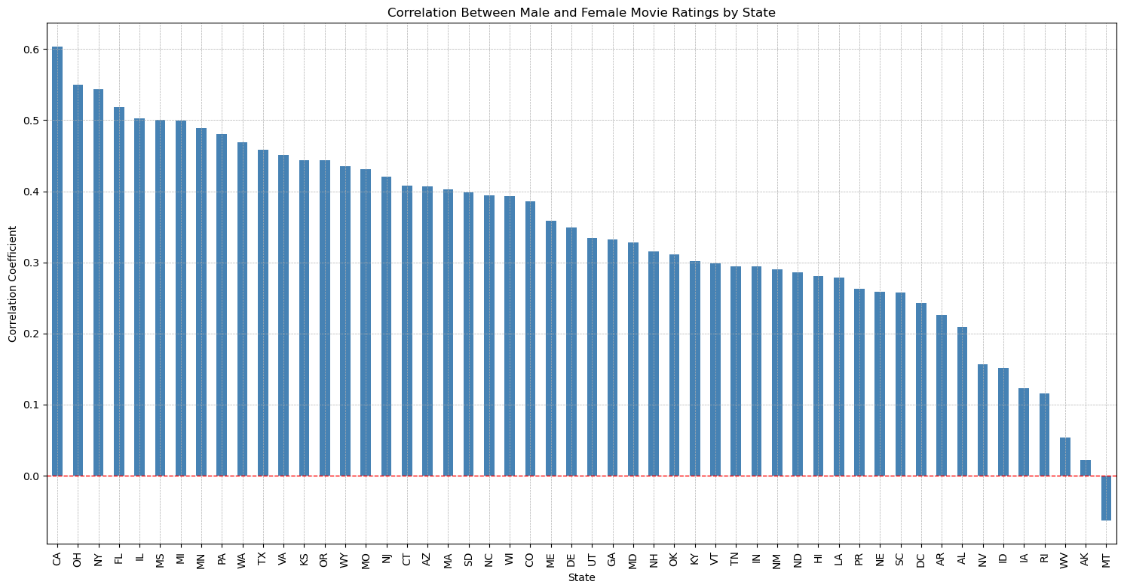
1. The distribution is approximately centered around a difference of 0, indicated by the red dashed line. This means that for many movies, the average ratings given by men and women are quite similar.
2. The distribution exhibits some spread, indicating that there are movies for which men and women have differing opinions. However, these are less common compared to movies where the ratings are similar.
3. The majority of movies fall within a small difference range, further suggesting that men and women generally have similar preferences.

The conjecture holds true for many movies. While there are some exceptions where there's a noticeable difference in ratings between genders, on average, the differences are minimal.

- **Highly Correlated States (e.g., > 0.5)**: Ratings in these states show a strong positive correlation between men and women. For instance, in California (CA), the correlation is approximately 0.6030.603. This means that in these states, the ratings given by one gender can be a good predictor of the ratings given by the other gender.

    - Examples: CA, OH, NY, FL, IL
- **Moderately Correlated States (e.g., between 0.3 and 0.5)**: Ratings in these states show a moderate positive correlation between men and women. The predictive power might be lower than in the highly correlated states, but there's still a reasonable relationship.
    - Examples: WA, TX, VA, KS, OR
- **Lowly Correlated States (e.g., < 0.3)**: Ratings in these states show a weak correlation between men and women. Using ratings from one gender to predict the other might not be as reliable in these states.
    - Examples: GA, MD, NH, OK, KY
- **Negatively Correlated States**: A negative correlation suggests that as the rating by one gender increases, the rating by the other gender decreases. This behavior is unexpected and might be due to a lack of sufficient data or other factors.
    - Example: MT

Histogram:



- The states on the left side of the chart (with higher bars) have a strong positive correlation, suggesting that ratings from one gender can be a good predictor of ratings from the other gender in these states.
- As we move towards the right side of the chart, the correlation decreases, indicating that the predictive power of one gender's ratings for the other's diminishes.
- The states with bars near or below the red dashed line have weak or negative correlations, implying that using ratings from one gender to predict the other's may not be reliable in these states.

So, California has the highest correlation, meaning that ratings from one gender can be a good predictor of ratings from the other gender. However, it's also essential to understand the underlying reasons for these correlations and to ensure there's sufficient data to make reliable predictions.

Part 4:

Do any of your conjectures in Problems 1, 2, and 3 provide insights that a movie company might be interested in?

Propose a business question that you think this data can answer.

Suppose you are a Data Scientist at a movie company. Convince your boss that your conjecture is correct!

1.Insights:

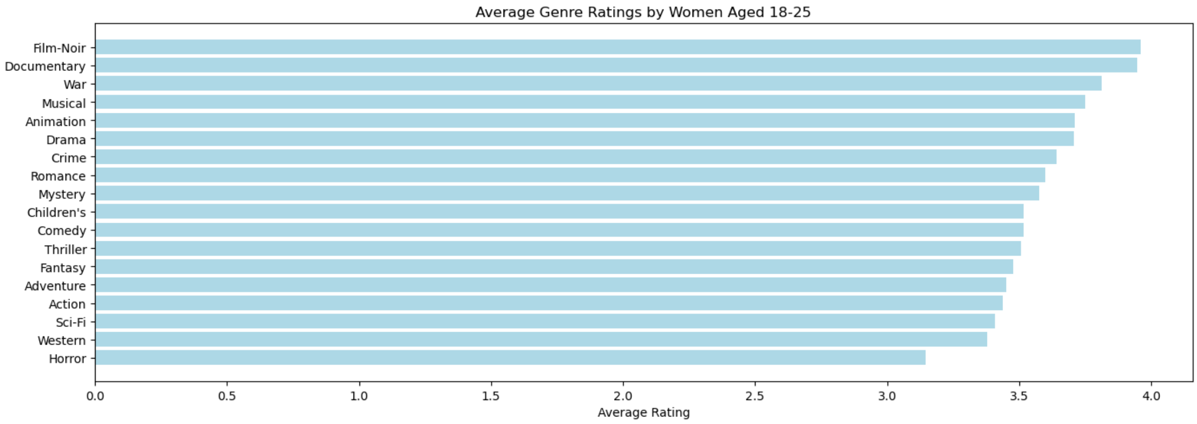
1. **Popularity and Ratings**: A production company would be keenly interested in the characteristics of movies that not only receive high ratings but also engage a significant portion of the audience. High ratings from a broader audience segment can indicate a movie's potential for success in the box office.

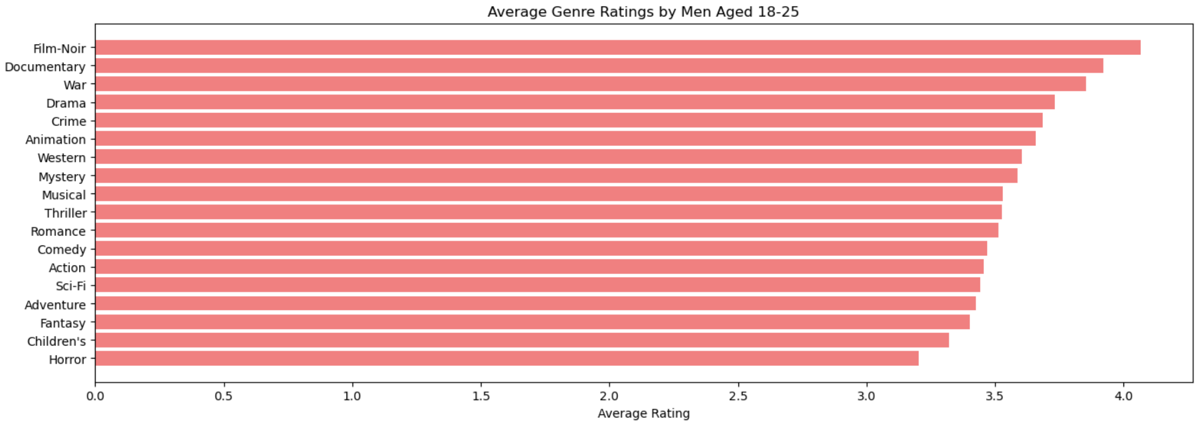
2. **Genres and Ratings**: Knowing which genres consistently receive higher ratings can guide the company's production decisions. If the company identifies that genres like "Film-Noir" or "Documentary" consistently receive high ratings, it might consider investing in such movies. However, it's crucial to balance this with the understanding of market size for these genres. A highly-rated niche film might not yield as much box office revenue as a moderately-rated blockbuster genre.

3. **Age and Ratings**: The demographics of movie-goers can vary significantly. If a company is producing movies aimed at younger audiences, they would want to ensure the content aligns with the tastes and preferences of that age group. Conversely, movies targeted towards older audiences might focus on entirely different themes and narratives.

4. **Gender and Ratings**: Gender-based preferences can be especially important for movies that are targeting a specific gender demographic. For instance, certain rom-coms might be

targeting a female audience, while action-packed war movies might be targeting a male audience. Understanding these nuances can help in both the production and marketing phases.





In conclusion, for a movie production company, the key would be to balance the insights from ratings with the potential market size and production costs. These insights can help a movie company tailor its production and marketing strategies to target specific demographics more effectively, choose genres that resonate with audiences, and understand the correlation between popularity and ratings.

2. **Propose a Business Question**:

"How can we optimize our movie production and marketing strategies based on audience preferences in terms of genre, age, and gender to maximize viewership and ratings?"

Genre Preference:

We've identified that certain genres, such as "Film-Noir" and "Documentary," tend to receive higher average ratings, while others like "Horror" and "Sci-Fi" might receive slightly lower ratings.

Understanding the genre preferences of the audience allows the company to invest in producing content that's more likely to be well-received.

Age-based Strategies:

Younger audiences (especially children) have shown a tendency to provide more extreme ratings. This insight suggests that content targeting younger audiences should be impactful and engaging to ensure positive feedback.

For older audiences, the preferences might lean towards genres or themes that they resonate with, based on cultural or generational experiences.

Gender-based Marketing:

While the ratings provided by men and women were generally similar, there might be specific genres or themes that resonate more with one gender over the other. Tailoring marketing campaigns based on these nuances can increase engagement.

Popularity and Ratings:

Movies that have been rated more frequently might provide a more reliable indication of their quality or appeal. Identifying characteristics of popular movies can guide future production decisions.

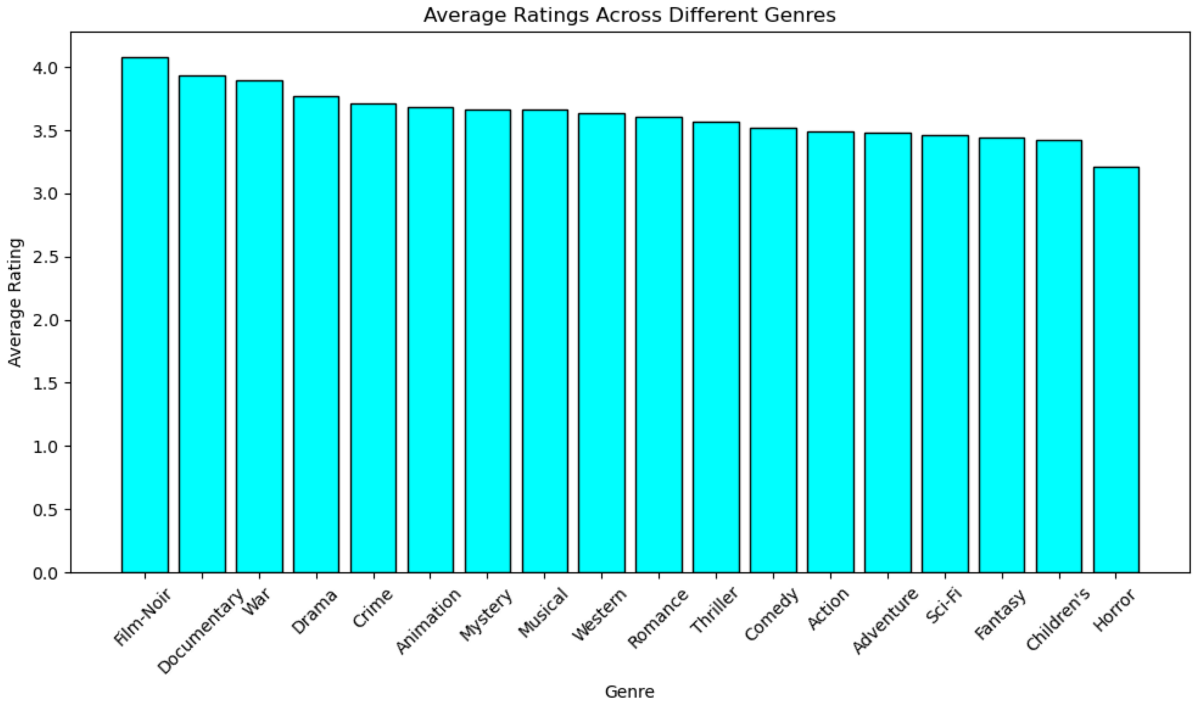
By addressing this business question, the movie company can make data-driven decisions in both the production and marketing phases, ensuring content resonates with the target audience and maximizes potential success.

3. **Convincing the Boss**:

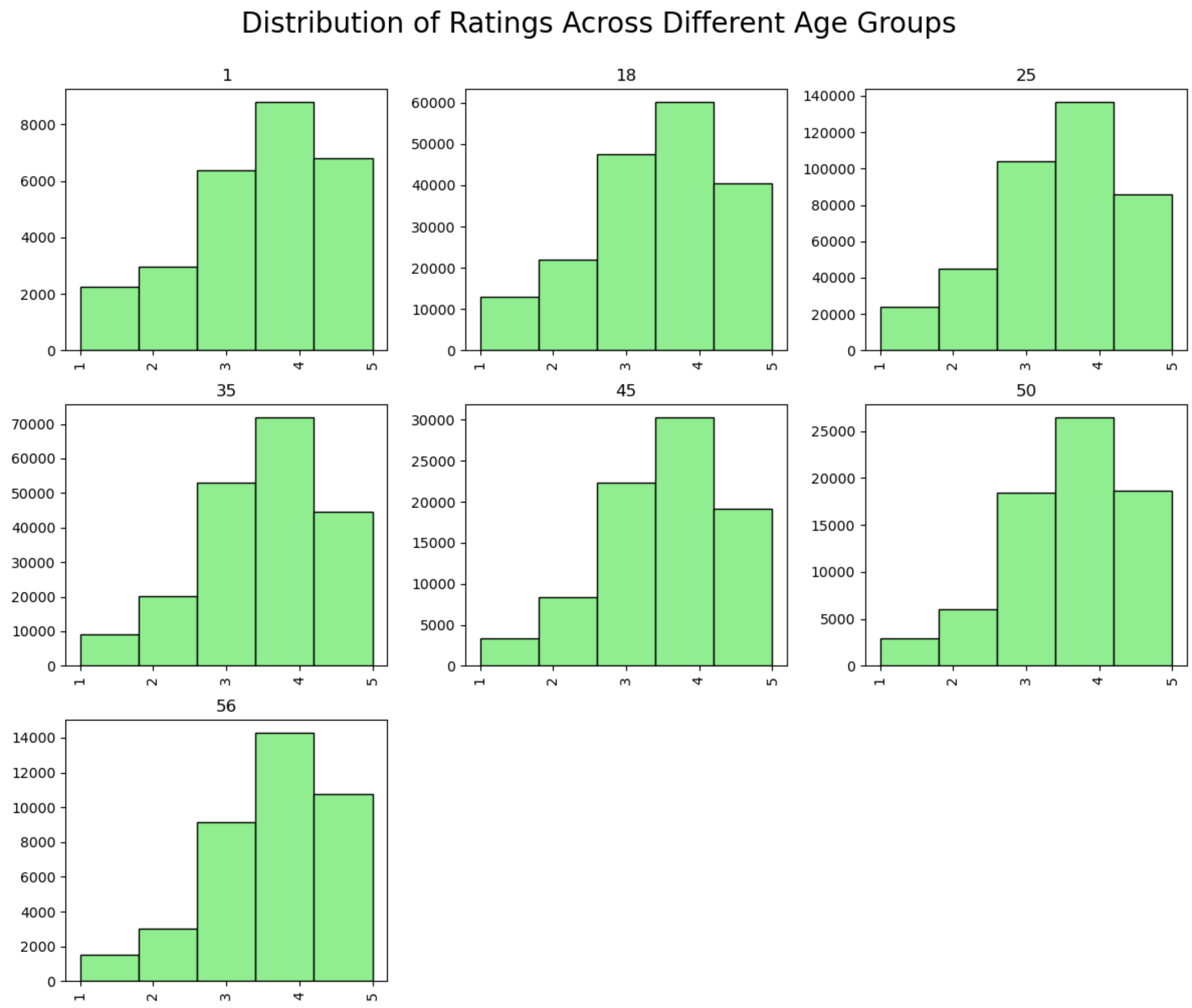
Considering we’re a data scientist in a meeting to convince our boss of the insights we discovered a short but concise presentation is done as follows:

Good morning! Today, I want to share some important insights we've derived and understood from our extensive analysis of the MovieLens dataset. These insights can significantly inform our future production and marketing strategies.

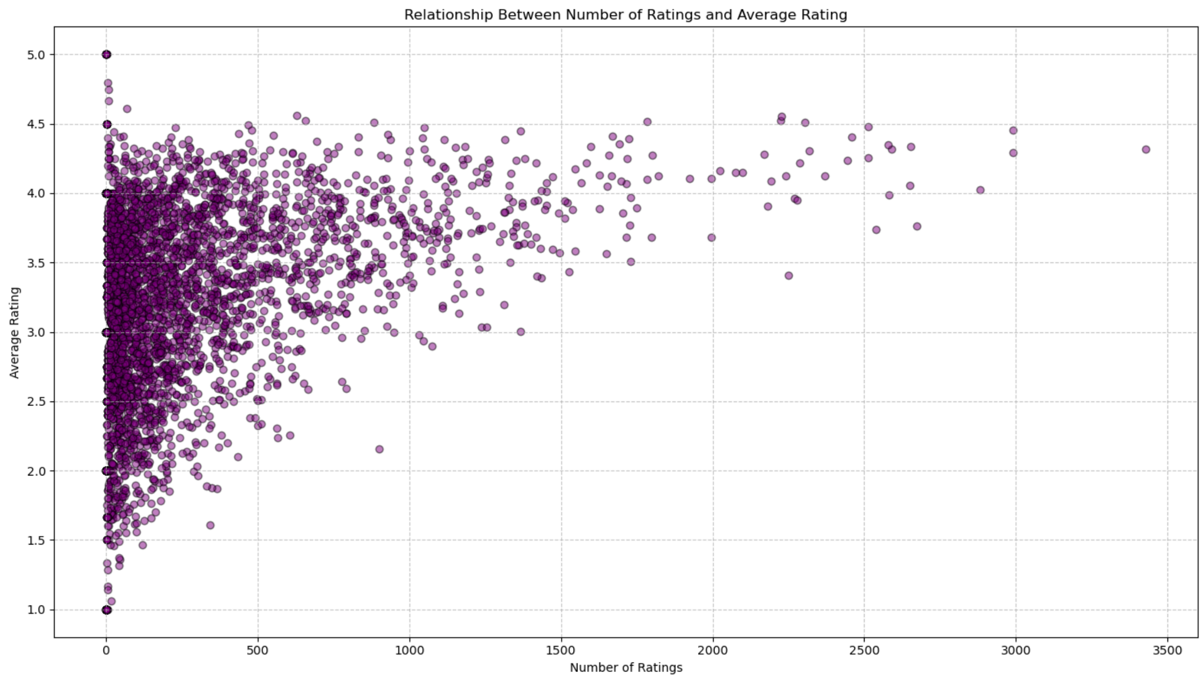
Let's start by looking at this chart titled 'Average Ratings Across Different Genres'. As you can see, genres like 'Film-Noir' and 'Documentary' consistently achieve higher ratings. This suggests that audiences have a deep appreciation for these genres. Investing in such genres could enhance our brand reputation for producing quality content.



Next, we can talk about age-based preferences. The histogram here titled 'Distribution of Ratings Across Different Age Groups' reveals a good trend. Younger audiences, especially children, have a strong tendency to provide extreme ratings. This could mean that when we get it right with this demographic, the appreciation is immense. However, if we miss the mark, the feedback can be equally strong in the opposite direction. The oldest age group again shows a strong preference for 4-star ratings but also has a noticeable count of 5-star ratings. This might suggest that this age group, similar to the youngest, has a tendency to give more extreme positive ratings.



Lastly, I want to emphasize the correlation between a movie's popularity and its ratings. The scatter plot 'Relationship Between Number of Ratings and Average Rating' indicates that popular movies generally cluster in the favorable rating range. This means movies that resonate with a broader audience not only receive more ratings but also tend to get positive feedback.



By leveraging these data-driven insights, we can make informed decisions in our production, from genre selection to targeted marketing strategies. The data suggests clear pathways to cater to audience preferences, and aligning our strategies with these insights can position us for greater success in the box office and beyond.

Story of the Group:

Part 1: Rohan Rana
Part 2 : Isha Jain
Part 3: Bashir Gulistani
Part 4: Ahmad Seayar