Movie Characters Personality

MBTI Types and Their Reasons and Influence

Team 8: Xiaotong Ma, Sai Nandini Peesapati, Xuhui Zhan, Yitian Long

1. Abstract

In this exploratory data analysis project, we investigate the relationship between Myers-Briggs Type Indicator (MBTI) personality types and movie features. Our methodology encompasses comprehensive data cleaning, merging diverse datasets, and rigorous analysis to explore correlations between character personality traits and movie genres, trends in MBTI type prevalence over time and across regions, and the impact of these personality types on IMDb movie scores and overall movie success. While our findings reveal intriguing connections between MBTI types and genre preferences, as well as shifting trends in type representations reflecting societal contexts, the overall impact of MBTI types on a movie's critical and commercial success appears minimal. This study highlights the complexity of film success, influenced more significantly by a myriad of factors beyond character personality types.

2. Introduction

1) IMDB

IMDb, which stands for Internet Movie Database, is an online database of information related to films, television programs, home videos, video games, and streaming content online – including cast, production crew, personal biographies, plot summaries, trivia, ratings, and fan and critical reviews. In addition to its comprehensive database, IMDb offers a platform for users to rate movies and television shows and submit new material for approval by the site's volunteers and editors. It serves as a valuable resource for movie enthusiasts, industry professionals, and researchers looking to access a wealth of information about the entertainment industry.

2) MBTI

The Myers-Briggs Type Indicator (MBTI) is a psychological assessment tool that classifies individuals into 16 distinct personality types based on their preferences in perceiving the world and making decisions. These personality types are determined using four dimensions, each consisting of two opposite preferences:

- Extraversion (E) versus Introversion (I): This metric indicates where individuals prefer to focus their attention and derive energy—from the external world of activity, people, and things (Extraversion) or the internal world of thoughts and reflections (Introversion).
- Sensing (S) versus Intuition (N): This metric shows how individuals prefer to take in information—through direct sensory experience and present realities (Sensing) or by interpreting and adding meaning, focusing on patterns and possibilities (Intuition).
- Thinking (T) versus Feeling (F): This metric demonstrates the preferred approach to making decisions—objectively, using logic and consistent rules (Thinking), or subjectively, based on personal values and the impact on others (Feeling).
- Judging (J) versus Perceiving (P): This metric describes how individuals prefer to deal with the outside world—in a planned, orderly way, seeking closure (Judging), or in a flexible, spontaneous way, keeping options open (Perceiving).

These dimensions, when combined, form a four-letter code (like INFP or ESTJ) that provides a framework for understanding individual differences and guiding personal growth, communication, and career choices.

The Myers-Briggs Type Indicator (MBTI) classifies personalities into 16 distinct types, grouped into four roles: Analysts (Intuitive and Thinking), known for their strategic and logical thinking; Diplomats (Intuitive and Feeling), empathetic and principled, excelling in harmonious interactions; Sentinels (Observant and Judging), pragmatic and organized, upholding traditions and order; and Explorers (Observant and Perceiving), spontaneous and adaptable, thriving on freedom and hands-on experiences. Each role offers a unique perspective, enriching the tapestry of human behavior and interaction.

3. Pre-process

1) Data Clean

In preparation for an exploratory data analysis focusing on the relationship between MBTI of characters and various features of movies, a comprehensive data cleaning process was undertaken. The objective was to refine three distinct datasets: mbti.csv, imdb_movies.csv, and imdb_db.csv, ensuring they were free from inconsistencies, irrelevant information, formatting issues, and ready to be merged. This report outlines the key steps taken in the data cleaning process.

- MBTI Dataset Cleaning (mbti.csv)

The MBTI dataset, initially containing columns such as "stat," "enneagram," and "img_url," was streamlined by removing these unrelated columns. Duplicated rows were identified and eliminated to ensure data uniqueness. A critical step involved verifying the integrity of the "mbti" column; non-MBTI types labeled as "XXXX" were removed. Furthermore, a standard format was applied to the "role" and "movie" columns, involving capitalization adjustments and whitespace trimming. An intricate part of this process was the extraction and cleaning of movie release years using regular expressions, which involved separating the year from the movie title and converting it to a datetime format. The dataset was then restructured for consistency by renaming relevant columns and saving the cleaned data to a new file.

- IMDb Movies Dataset Cleaning (imdb_movies.csv)

For the IMDb movies dataset, a similar approach was employed. Columns not pertinent to the analysis, such as "overview," "crew," and others, were dropped. The focus was on ensuring the quality of the "names" column, which involved capitalization corrections and whitespace removal. The "date_x" column was converted to a datetime format to extract the release year of the movies. The dataset was then restructured to align with the other datasets, emphasizing consistency in column naming and data formats.

- IMDb Database Cleaning (imdb_db.csv)

The IMDb database underwent extensive cleaning to align it with the project's objectives. Several unrelated columns were removed, including those detailing votes, duration, and series information. The remaining data was standardized, particularly in the "Movie Name" and "Movie Type" columns, with specific attention to formatting and removal of unnecessary characters like square brackets and single quotes. The columns were renamed to maintain consistency across all datasets.

2) Data Merge

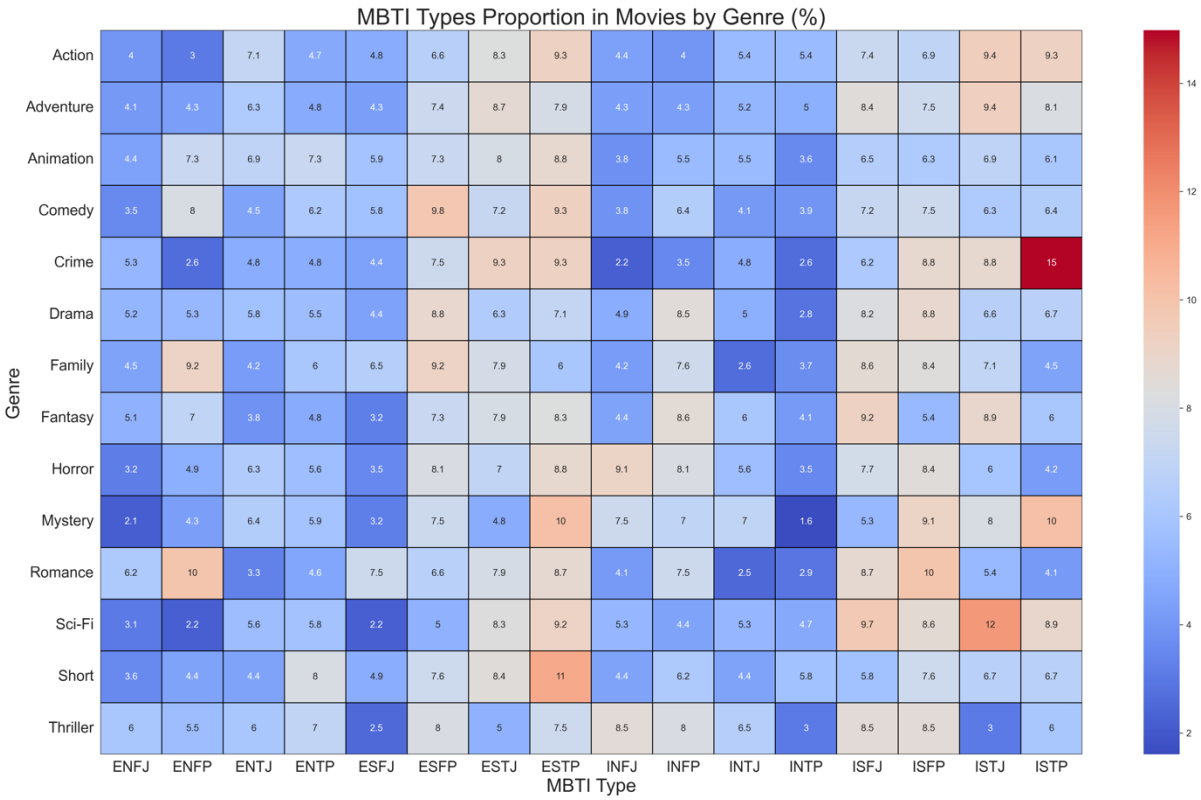
In the data merging phase of the MBTI and movie features analysis project, we carefully segmented the MBTI dataset into two based on the presence or absence of 'release_year' data. For entries with a release year, a precise merge was conducted with the IMDb database, using both 'movie_name' and 'release_year' as keys, followed by the removal of duplicates. Conversely, for MBTI entries lacking release year data, a merge was performed solely based on 'movie_name', after dropping the 'release_year' column. These two distinct merged datasets were then combined, creating a comprehensive and detailed dataset. This final dataset, sorted by 'movie_name' and 'mbti', is instrumental for an in-depth analysis of the relationship between character MBTI types and various movie features, accommodating a wide range of data while maintaining accuracy and relevance.

4. Experiment and Results

1) MBTI and Genres – Xiaotong ‘Brandon’ Ma

We utilized the dataset that records the MBTI types of characters alongside the movie genres their movie belongs to. Given that movies often span multiple genres, the dataset was pre-processed to separate and categorize the genres accordingly. Following this, we performed a grouping operation to count the occurrence of each MBTI type within each movie genre. To ensure the robustness of our results, only genres with at least 100 associated data points were included in the final analysis. We then computed the proportional preference of each MBTI type for each genre with normalization by genre and proceeded to visualize this data through two distinct graphical representations—a heatmap and a line plot—for ease of interpretation.

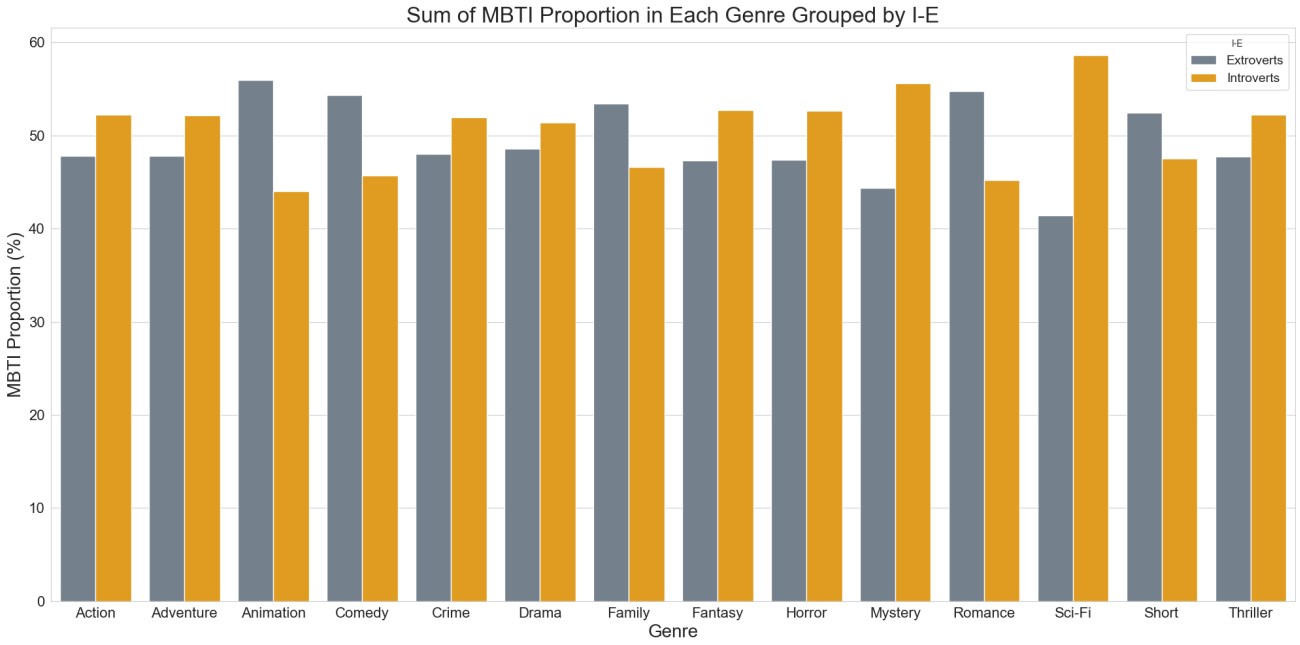
1. Heatmap – MBTI Types Proportion in Movies by Genre



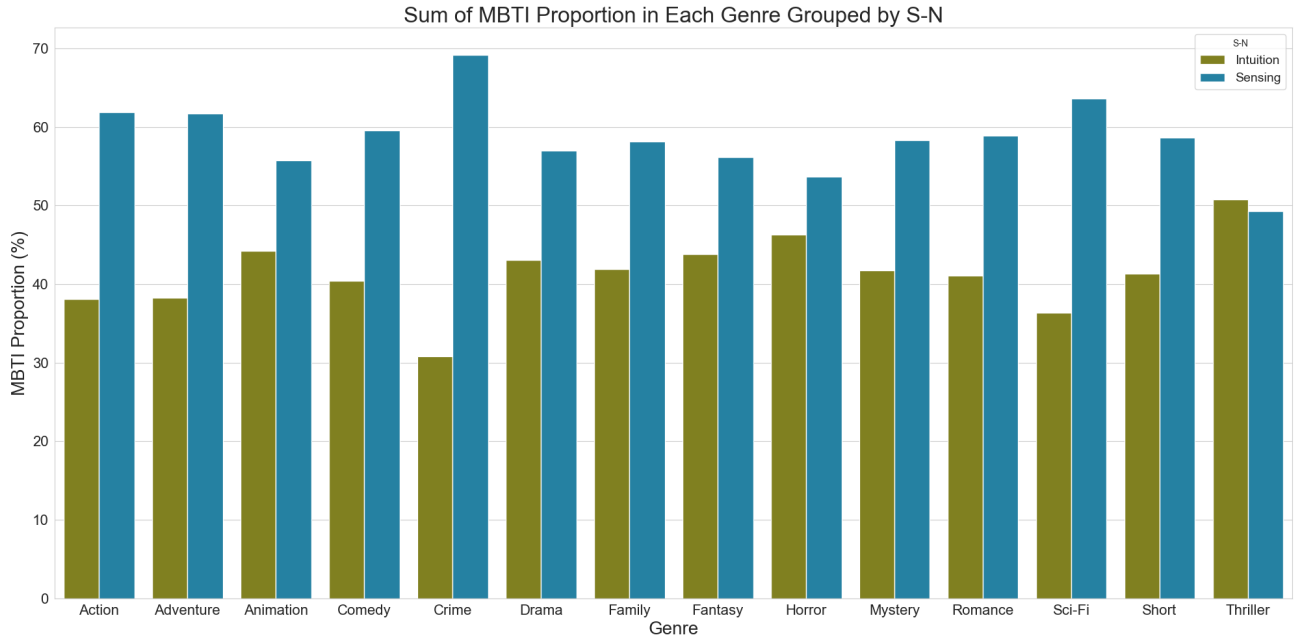
This is a heatmap of proportions of MTBI types across movie genres. The horizontal axis represents different MBTI types as well as the vertical axis shows different movie genres. Each block represents the proportion of the corresponding MBTI type in corresponding genre. Redder shades indicate higher proportions, while bluer shades indicate lower proportions. The percentages range from low (2%) to high (15%), with various shades indicating the degrees in between. It provides a general perspective of the distribution of MBTI types given the relationship between movie genres and MBTI types.

2. Four Different Dimension (I-E, S-N, T-F, J-P) Analysis:

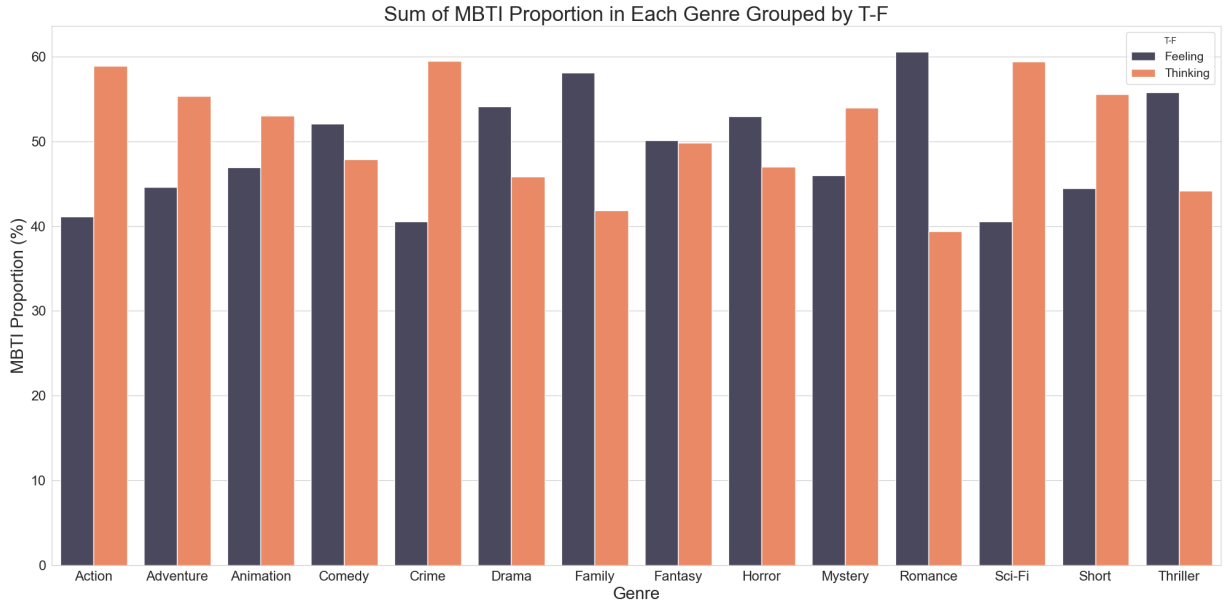
There are four groups of bar plots. Each plot shows the proportion of different dimension in MBTI types across different genres. The horizontal axis represents different genres while the vertical axis shows the proportion of dimensions of MBTI types.



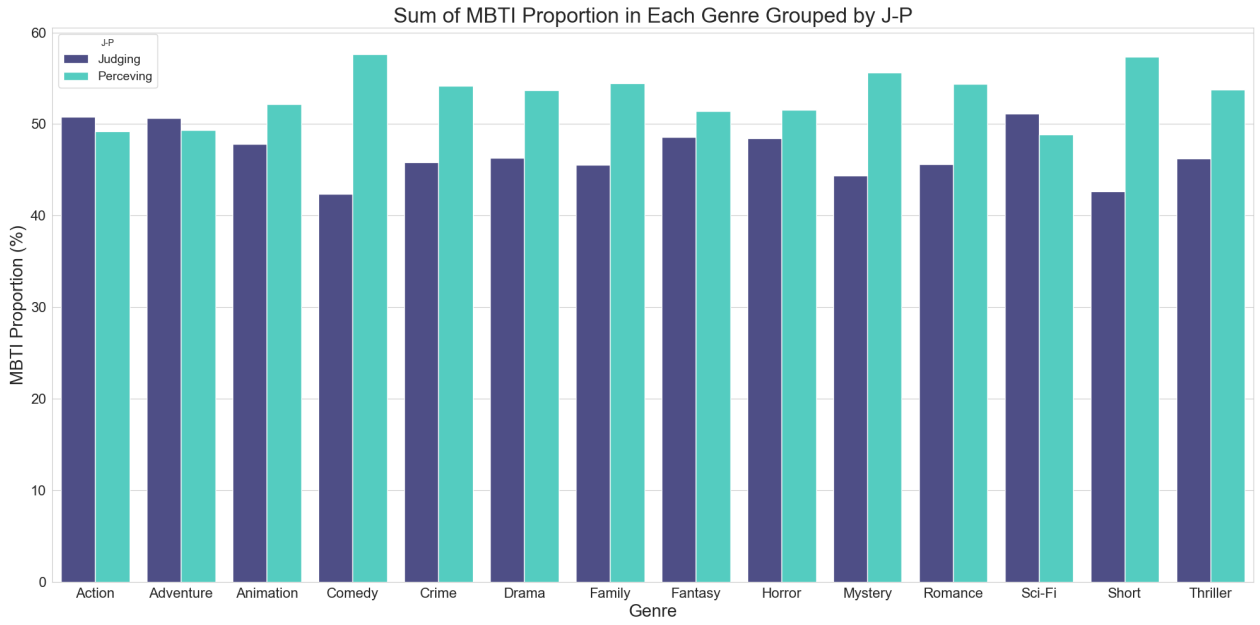
Introversion-Extraversion (I-E): The distribution of Introverts and Extroverts across various film genres shows a nuanced balance, reflecting broader societal trends. In dynamic and high-energy genres like Action and Adventure, there is a noticeable tilt towards Extroverted characters, who thrive in these stimulating environments. Conversely, genres like Short films, which often focus on deeper, more reflective themes, attract more Introverted characters. This suggests that the nature and thematic content of a genre can influence the type of personalities it attracts or portrays.



Sensing-Intuition (S-N): Sensing characters, who excel in dealing with concrete details and practical matters, are prominently featured in genres such as Comedy and Drama. These genres often revolve around relatable, real-life situations where practicality is key. Despite being less common in the general population, Intuitive types are surprisingly well-represented in Thriller movies. This indicates that Intuitive personalities are particularly drawn to the complex, abstract thinking often required in these suspenseful and intellectually stimulating narratives.



Thinking-Feeling (T-F): In Romance films, characters with a Feeling preference are predominant, navigating their worlds through a lens of personal values, emotional connections, and deep relationship dynamics. This contrasts with genres like Sci-Fi and Crime, where Thinking characters are more common. These characters approach problems with logic and critical thinking, tackling scientific, technological, or detective challenges with a rational and factual mindset. Their analytical approach is key in driving the narrative, especially in scenarios that demand objective problem-solving and strategic planning.



Judging-Perceiving (J-P): The Judging versus Perceiving personality dimension plays a significant role in character portrayal across film genres. Action and Adventure films often feature Judging characters, who bring a sense of order, decisiveness, and strategic thinking critical for the fast-paced and high-stakes plots in these genres. On the other hand, Perceiving characters, with their flexibility and openness to new experiences, are a natural fit for Family and Short films. These genres often have more relaxed, exploratory storylines where adaptability and a focus on the journey itself are more emphasized than a fixed outcome or resolution.

2) MBTI and Time-series – Sai Nandini Peesapati

The MBTI character in movies has been used for a very long time. From our dataset, we were curious to find the relationship between the MBTI characters, genre of the movie, the years and so on. So, we found some interesting relationships and formed our analysis based on a few visualizations.

1. The frequency of different Myers-Briggs Type Indicator (MBTI) personality types over a range of years from around 1920 to 2020.

MBTI Types of Analysts Over Years:

Variations in the frequency of Analyst personality types are depicted in this chart. All things considered, the frequency has increased over time, peaking in the 1980s, mid-2000s, and 2020.The significant increases point to times when Analyst-type personalities would have been especially esteemed or common.

MBTI Types of Diplomats Over Time:

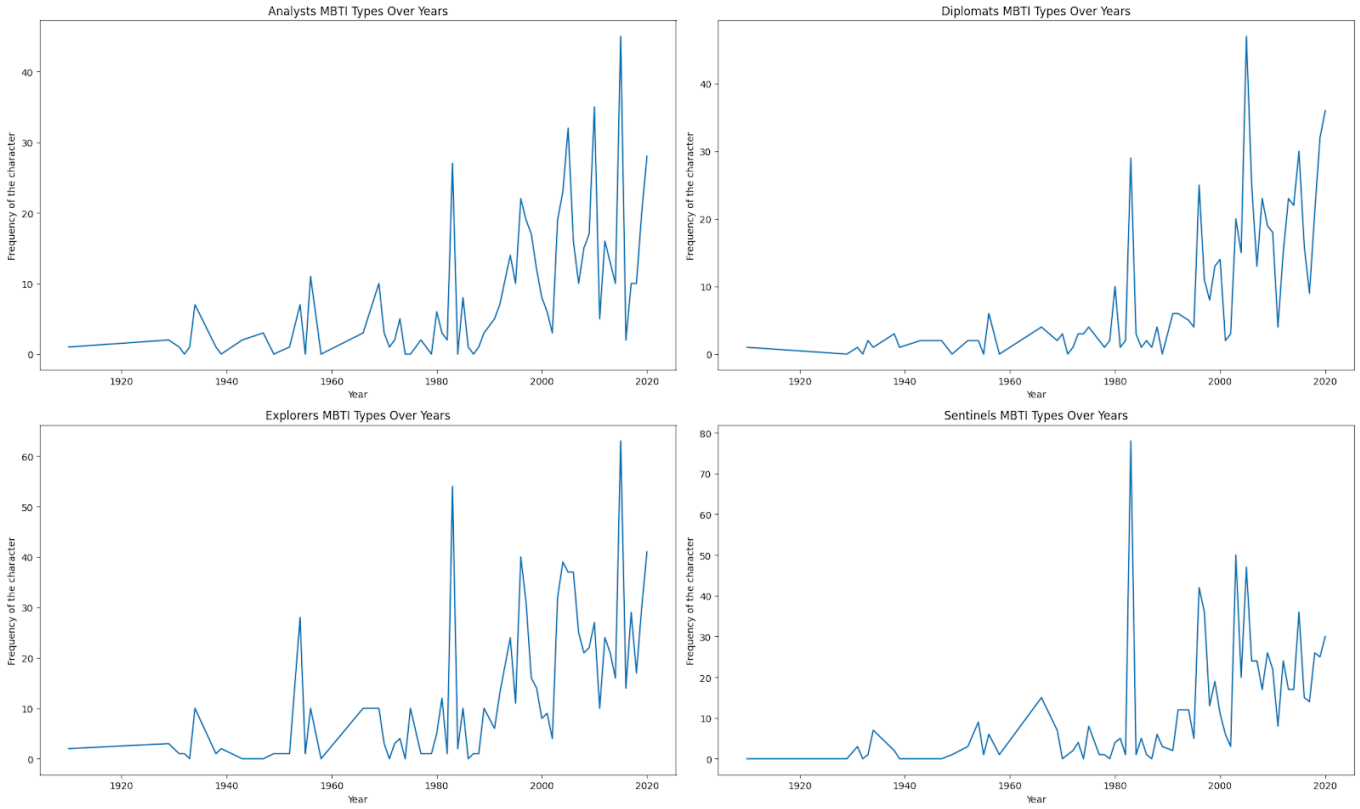
Diplomat personality types are becoming more common throughout time, with significant upticks occurring in the 1980s, early 2000s, and 2020. The pattern implies that there can be times when diplomat kinds are highly recognized or prevalent.

Explorers MBTI Types Throughout Time:

The graph indicates that Explorer personality types are quite uncommon until notable peaks appear in the 1940s, late 1970s, and 2000s. The erratic peaks might point to periods when there was a greater demand for Explorer traits or a higher number of people who identified with them.

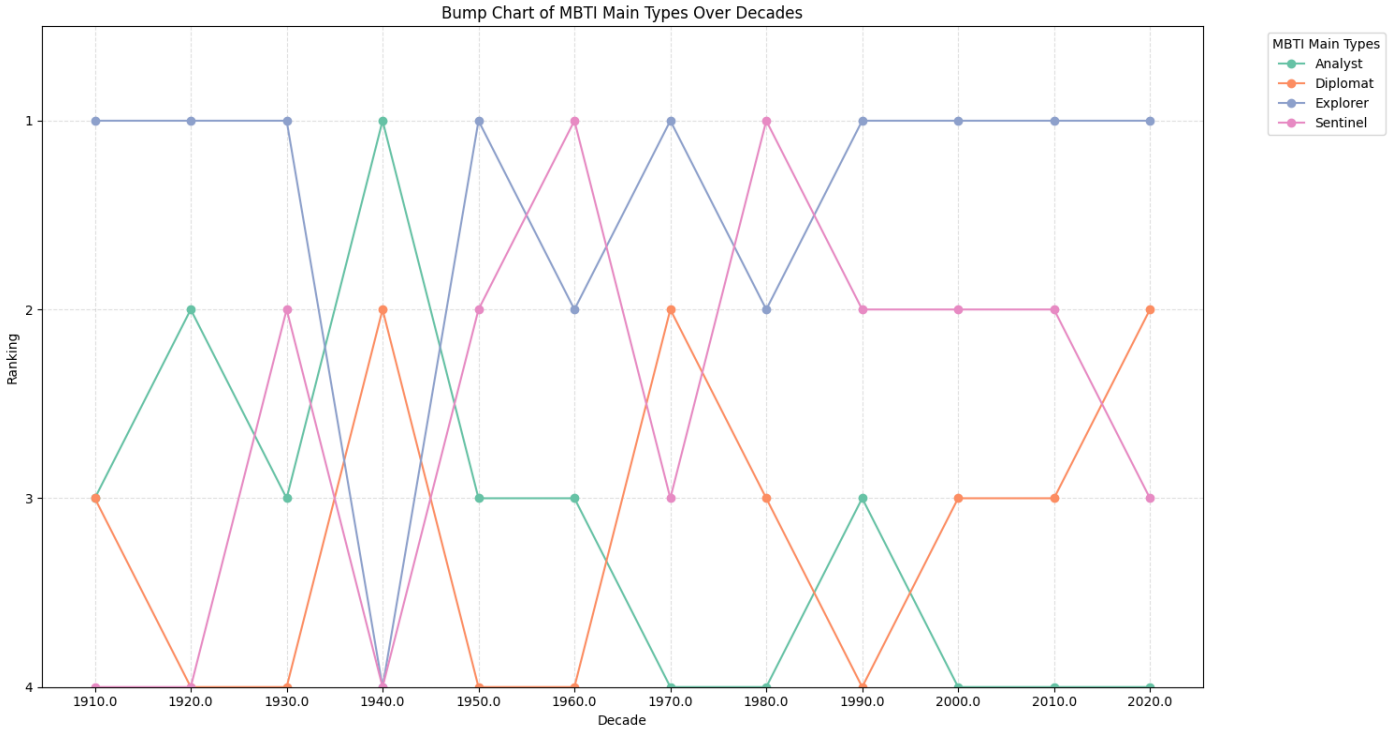
Sentinels MBTI Types Throughout Time:

Sentinel personality types are more common than expected; there was a notable rise in their frequency during the 1960s, and subsequent peaks were less significant. The significant surge might point to a time when Sentinel qualities were in great demand or very widespread.



The four charts show variations throughout time in the identification or prevalence of the various MBTI personality types. There are times when the prevalence of particular personality types spikes, which may be related to historical, societal, or economic circumstances that encouraged those characteristics. The maxima in all kinds around the 2000s can point to a rise in awareness or interest in MBTI personality types at that time. Changes in occupational responsibilities, societal upheavals, or cultural trends favoring certain personality qualities at different times could all be reflected in the data. To conduct a thorough analysis, it'll be necessary to take into account outside variables that can affect these frequencies, like fluctuations in the acceptance of MBTI tests in society, the role of psychology in management and education, and personality traits.

2. Bump Chart of MBTI Main Types Over Decades- comparison



Analyst : This line primarily floats near the top of the graph, showing that types of Analysts are continuously ranked first or second over the course of decades.

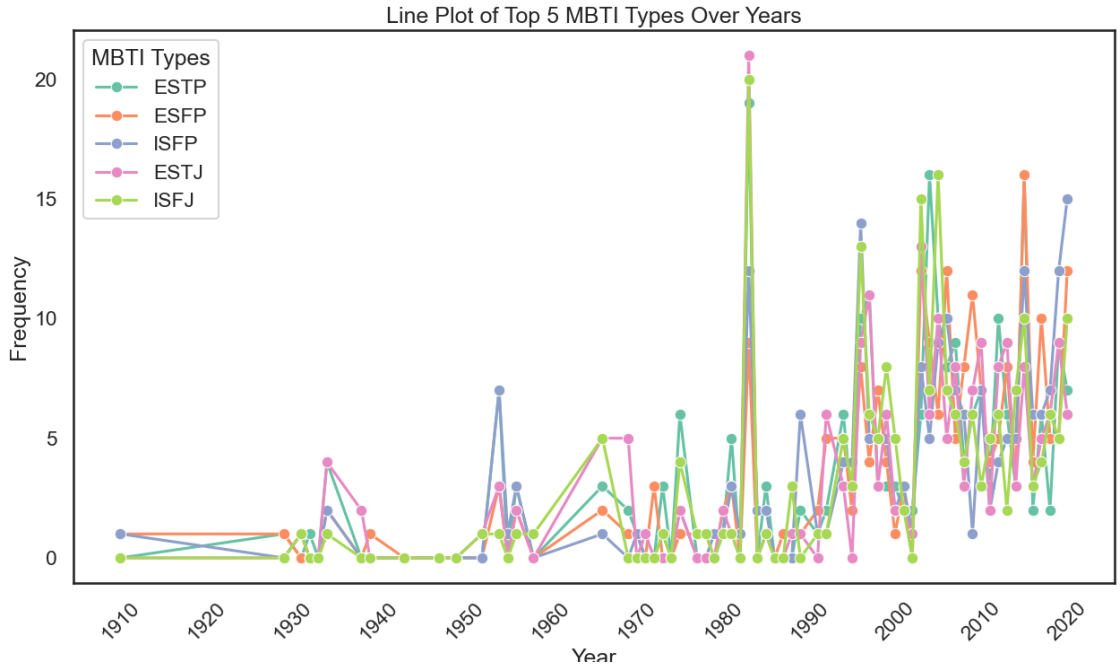
Diplomat : There is more variance in the Diplomat line, which shifts between all four ranks, indicating a changing preference or prominence for Diplomat personality types over the years.

Explorer : The Explorer line fluctuates as well, but it usually hovers around the lower half of the rankings, primarily between three and four, with a few ascents to the top two or first.

Sentinel : At some times, the Sentinel line crosses over the others and occupies every rank, exhibiting large fluctuations.

The Analyst personality type consistently ranks highest and exhibits a great degree of consistency. This may point to a consistent need or appreciation for the qualities of analysts, such as problem-solving and strategic thinking. The most variable personality types are those related to diplomacy, indicating that the traits linked to these individuals may fluctuate in popularity or significance throughout society based on the decade. Explorers are generally ranked lower, which may indicate that others don't always value or acknowledge the impulsive and daring qualities that make them famous. Sentinels show a very dynamic pattern, indicating that the practical and trustworthy characteristics of different Sentinel kinds are not always recognized or prevalent.

3. The frequency of five specific MBTI types over time



It is difficult to determine what is causing changes in frequency in the absence of more context. Is it because society values have changed, the MBTI has undergone modifications, or people are becoming more aware of and using it? The data sample itself raises certain concerns. Where is the data coming from? Does it reflect a certain demographic, a particular industry, or the whole public? The graph just illustrates correlation; it does not prove causality. Without more details, it is impossible to say with certainty why some varieties peaked. It could be challenging to identify distinct patterns or conclusions due to the noisy nature of the chart, which has many overlapping lines and variations. The coexistence of introverted and extroverted personality types suggests a varied spectrum of personalities that have been prominent at various points in time, maybe indicating a balanced respect for various features. Peaks may line up with historical or economic developments that placed a higher importance on the characteristics of particular MBTI types than others. Given that peaks appear to be more prominent in more recent decades, the data may possibly represent the increased interest in MBTI throughout time.

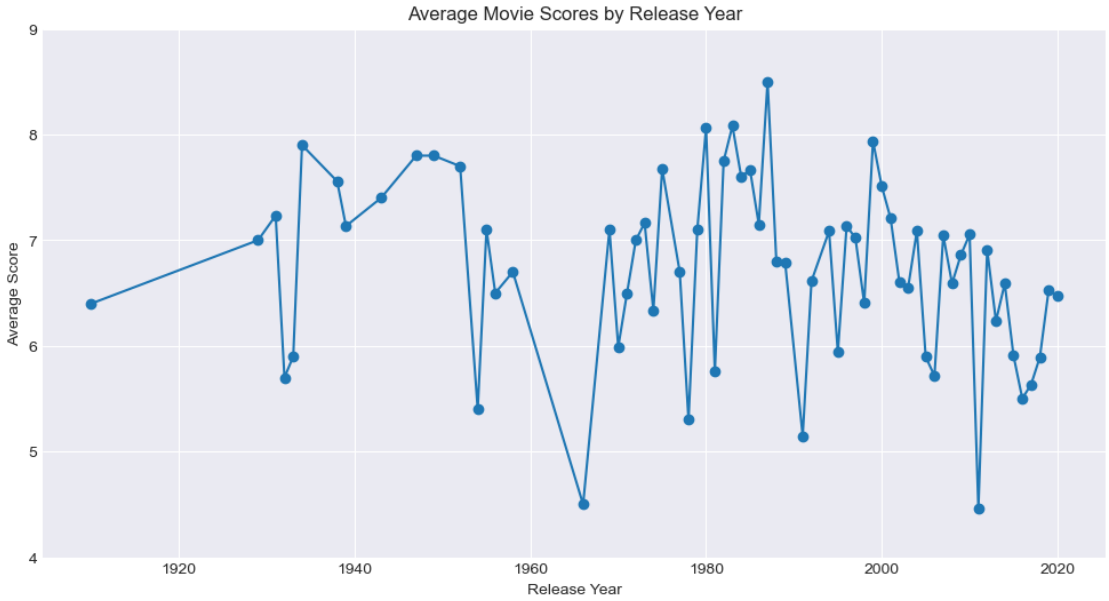
3) MBTI and IMDB Movie Scores – Yitian ‘Ewan’ Long

The main purpose of this section is to conduct a comprehensive analysis of the Myers-Briggs Type Indicator (MBTI) personality types of movie characters and their IMDb scores. The aim is to find potential correlations between the personalities of movie characters and their IMDb scores through data analysis, and to provide reasonable explanations. We

will start with an overall overview, trying to identify some unique findings through the density distribution of IMDb scores for MBTI personality types, and then delve into further analysis.

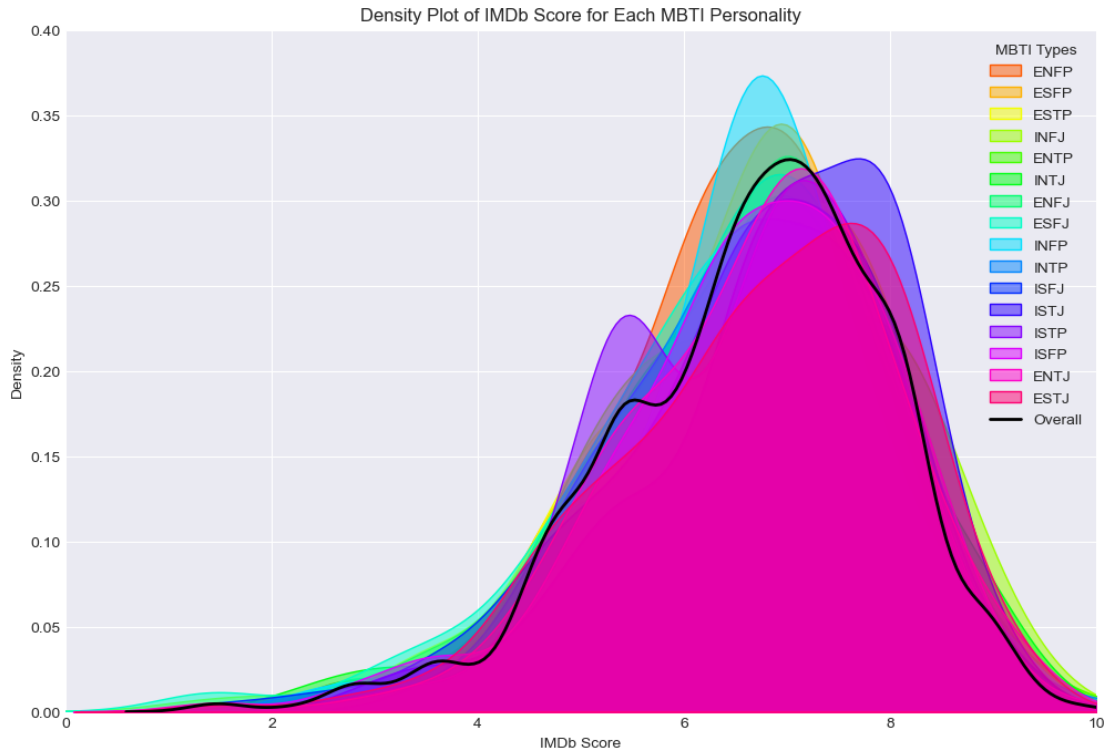
1. Overview of the Relationship between MBTI Personality Types and IMDb Scores

a. Average Movie Scores by Release Year



First , let us take a look to the line graph which represents the average IMDb scores of movies over a span of roughly a century, from around 1920 to the 2020s. the x axis represents the release year of the movies and y axis denotes the average IMDb Score, which ranges from about 4 to 9. The graph indicates early volatility in IMDb scores, likely due to fewer films influencing the average. The 1940s and 1960s saw peaks in ratings, possibly indicating periods of exceptional cinematic quality. Post-1960s, scores normalized, reflecting a greater number of movies being rated. Recently, scores have stabilized around 6 to 8, hinting at consistent film quality and rating behaviors. Overall, the IMDb Scores are quite reflective of the general situation and indicators of the movies, and they provide reference value for the audience.

b. Density Plot of IMDb Scores for Each MBTI Personality



People with different personalities appear in different movies and scenarios, bringing varied feelings to the audience. Could it be possible that they also have a certain impact on the audience's feelings, and consequently affect the IMDb scores?

The density plot displays the distribution of IMDb scores for movies characterized by the MBTI personality types of their lead characters. The x-axis ranges from 0 to 10, representing IMDb scores, while the y-axis shows the score density for each personality type. Each color corresponds to one of the 16 MBTI personality profiles assigned to movie characters, with the black line representing the aggregate distribution for all personality types.

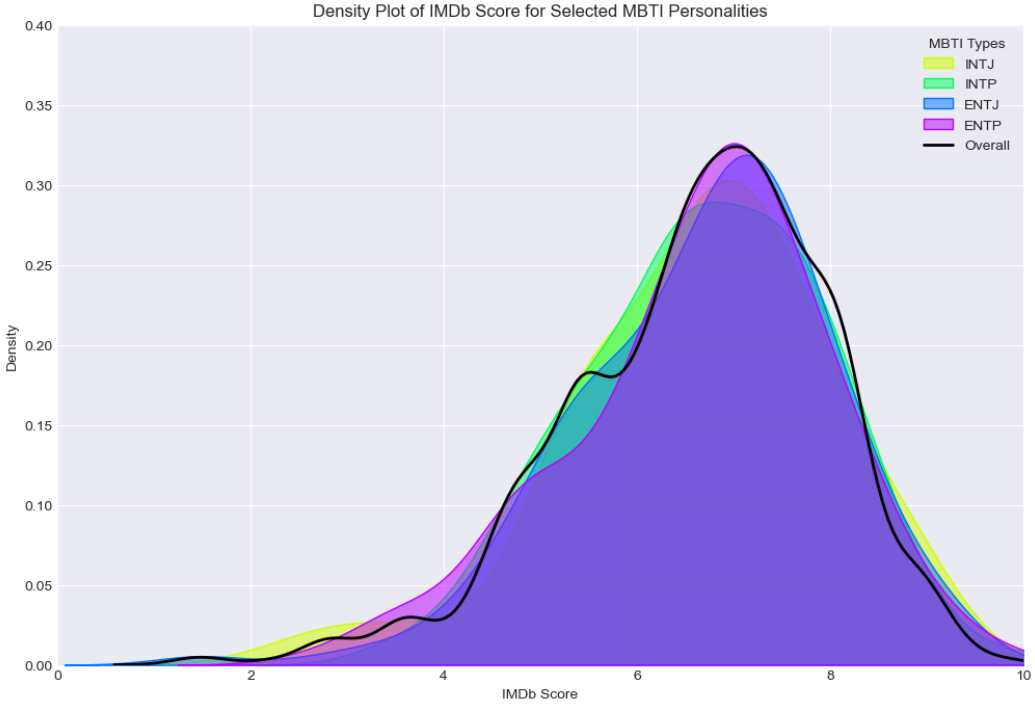
Most curves peak around the 6 to 8 score range, suggesting that movies with characters of different MBTI types generally receive similar ratings. The concentration around the 7 to 8 score mark across multiple personality types may reflect a broader audience appreciation for movies, regardless of the characters' personality types.

The overall distribution shows the collective trend across all character personality types, with a prominent peak suggesting a consensus on movie quality. The similarity in the distribution of ratings across various MBTI types indicates that while the personality of movie characters can influence audience ratings, the impact is nuanced, and the general audience ratings are not heavily polarized by character MBTI type.

Although the score distributions of characters with different MBTI types are generally similar, there are still relatively significant differences for some MBTI types. This may be because different types of characters give the audience different impressions.

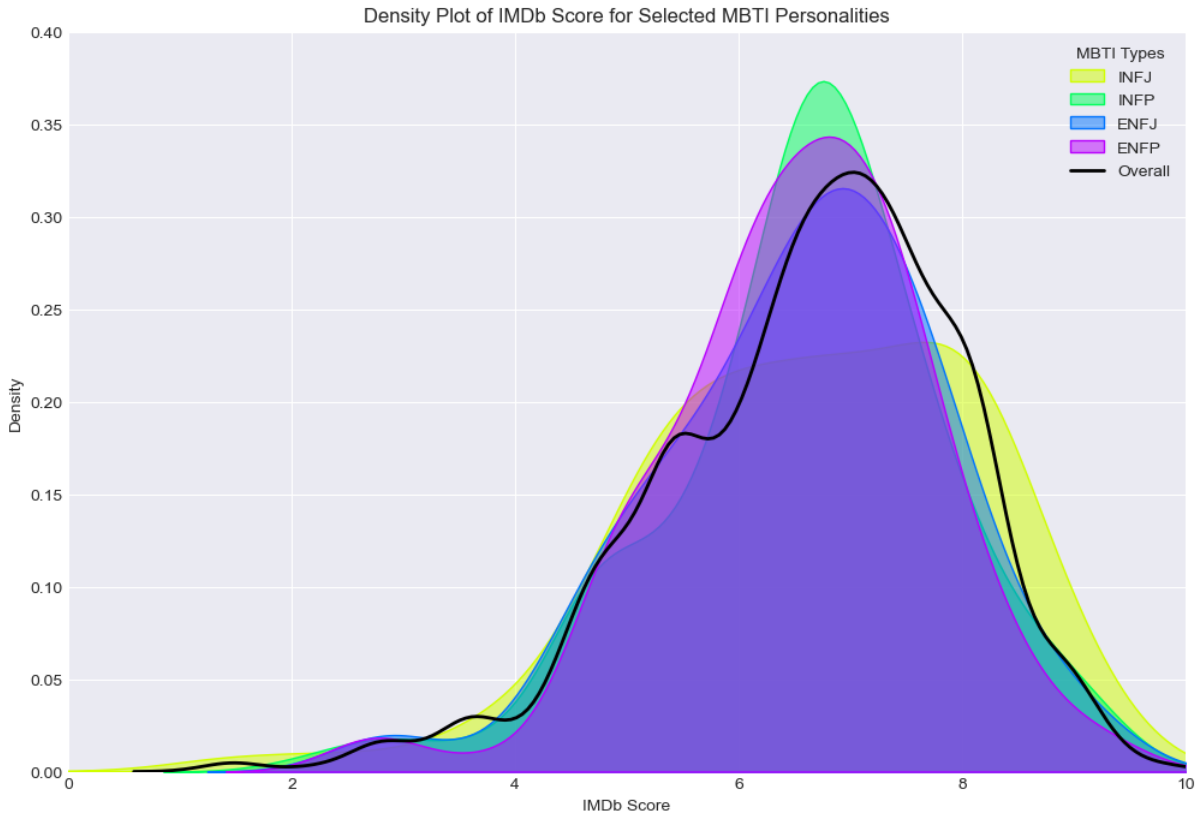
According to 16personalities.com, MBTI personalities can be divided into four categories. MBTIs within the same category have more in common. To better understand the differences between different MBTI categories, we can conduct more detailed analysis on the different groupings.

c. Density Plot of IMDb Scores for Each MBTI Personality Type Combination



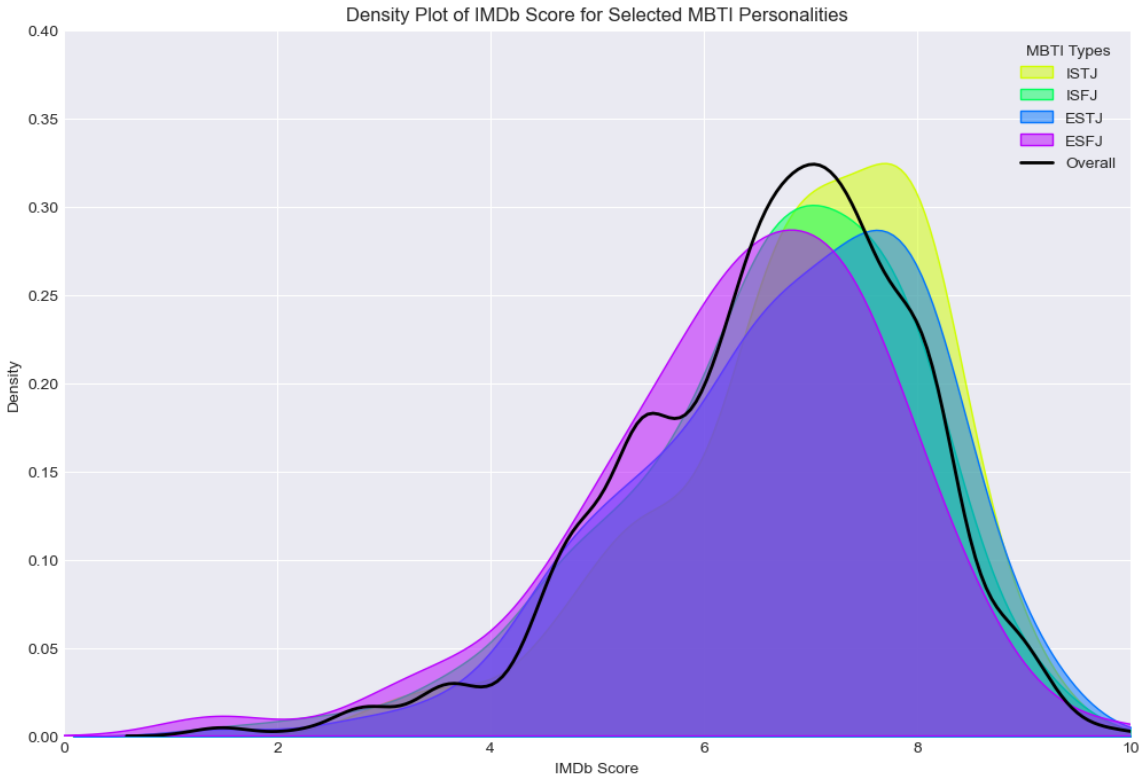
The chart density plot shows the distribution of IMDb scores for movies featuring characters with four specific MBTI personality types: INTJ, INTP, ENTJ, and ENTP. These types are often categorized together as "Analysts" in MBTI taxonomy due to their shared use of intuitive (N) and thinking (T) traits.

Overall, there is no particularly significant difference in the distribution of different MBTI types within the Analysts category. This suggests that these types of characters do not have particularly obvious preferences in the eyes of the audience and are close to the overall distribution.



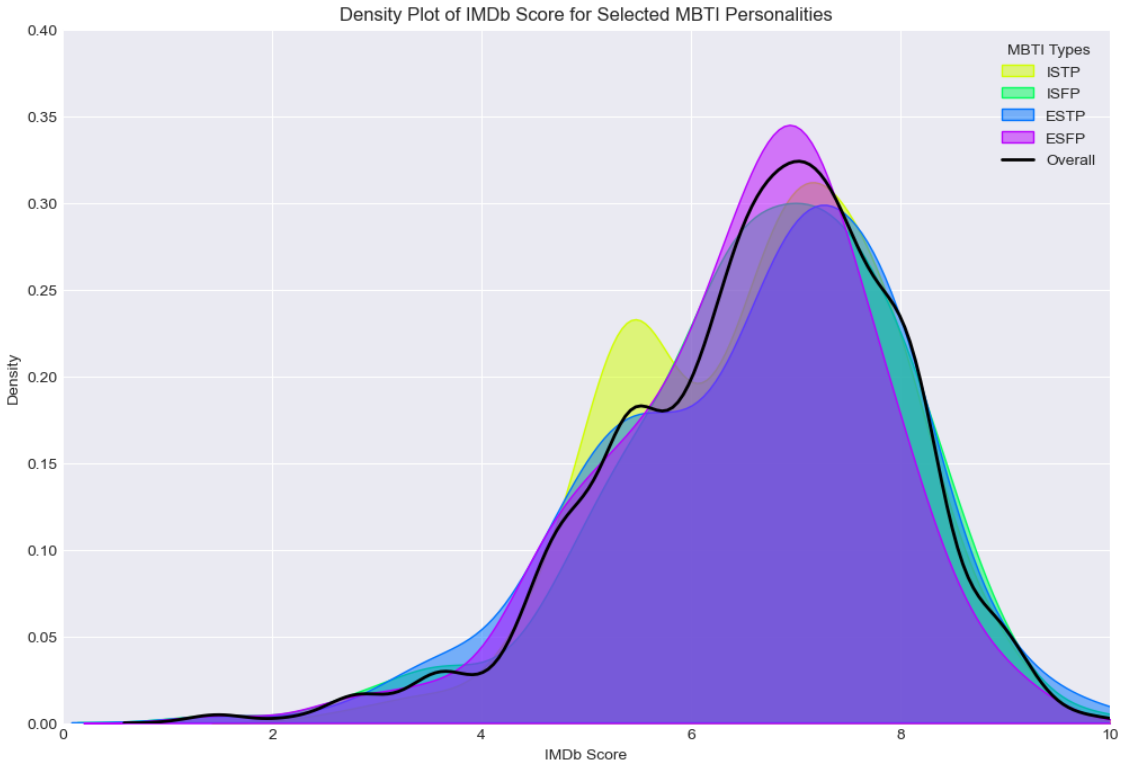
The chart is a density plot illustrating the distribution of IMDb scores for movies featuring characters with the MBTI personality types INFJ, INFP, ENFJ, and ENFP. These types are often classified as "Diplomats" in the MBTI framework, reflecting their shared intuitive (N) and feeling (F) traits.

Although all Diplomats do not have particularly noticeable differences in their score distributions, the score distribution for INFJ is relatively more right-skewed, indicating that audiences generally have a higher preference for movies featuring INFJ characters.



The density plot illustrates the distribution of IMDb scores for movies featuring characters with four different MBTI personality types: ISTJ, ISFJ, ESTJ, and ESFJ. These types are often grouped together as "Sentinels" according to the MBTI framework, known for their shared observant (S) and judging (J) traits.

Even though the curves for all four Sentinel types' score distributions cluster closely together, peaking around the 6 to 8 score region, it's evident that they have greater distinction compared to Analysts. This could be because, although Sentinel types have similarities, there is greater variation in the role of these characters in movies, which can more significantly impact the emotions and preferences of the audience.



This density plot illustrates the distribution of IMDb scores for movies featuring characters with the MBTI personality types ISTP, ISFP, ESTP, and ESFP. These types are often classified as "Explorers" in the MBTI framework, sharing perceptive (P) and sensing (S) traits.

The score distribution of ESFP and ISFP in the Explorers category is the closest to the overall distribution, their sharing of three personality indicators explains their high similarity. However, we can see that ISTP and ESTP have strikingly different score distributions, with ISTP's peak score concentrated around 5-6, while ESTP's peak score is around 7-8. This may indicate that the corresponding roles can provide very different viewing experiences to the audience.

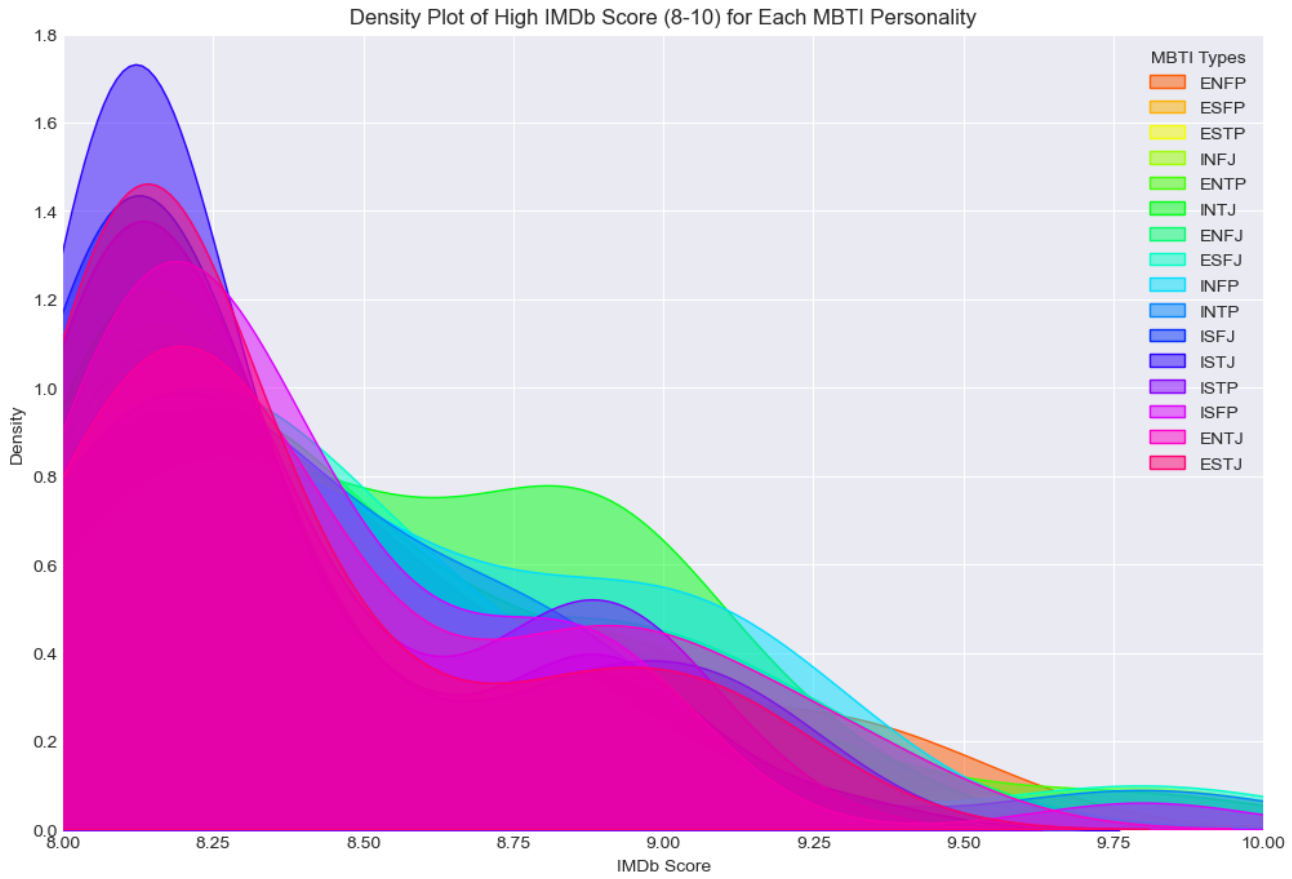
2. MBTI Personality Density in High IMDb Scores and Yearly Average Score Trends for High-Scoring Personalities

a. Density Plot of High IMDb Score (8-10) for Each MBTI Personality

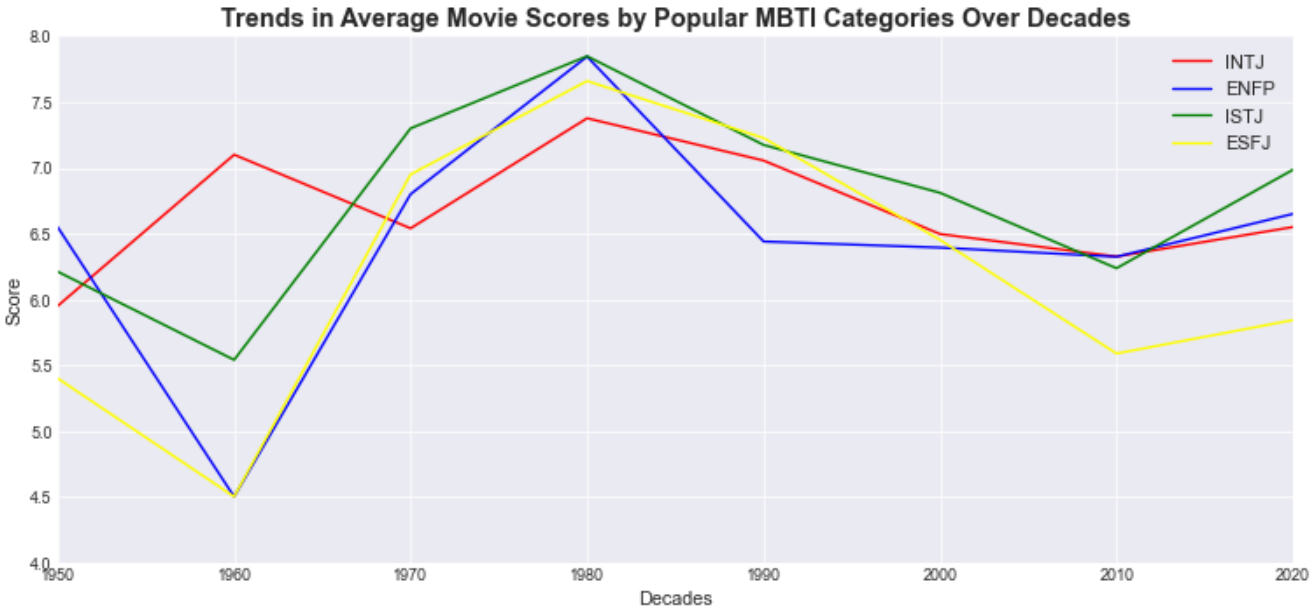
Next, let's focus on whether there are any noticeable differences in the distribution of different MBTI Types under the high score range (8-10) of IMDb Scores.

The plot highlights how each personality type is distributed within the high-scoring bracket. The density for all types decreases as the scores approach 10, which is expected since fewer movies achieve near-perfect ratings.

Certain types, like ISTJ, show higher density peaks closer to the 8 score mark, suggesting that characters with these personalities are in a larger number of highly-rated movies. Conversely, types like INTJ and ESFJ have their peaks further to the right, which could indicate a tendency for characters with these personalities to be in movies with the very highest IMDb scores, albeit less frequently. This makes me wonder, for these categories with relatively higher ratings, has their rating advantage been consistent historically?



b. Trends in Average Movie Scores by Popular MBTI Categories Over Decades

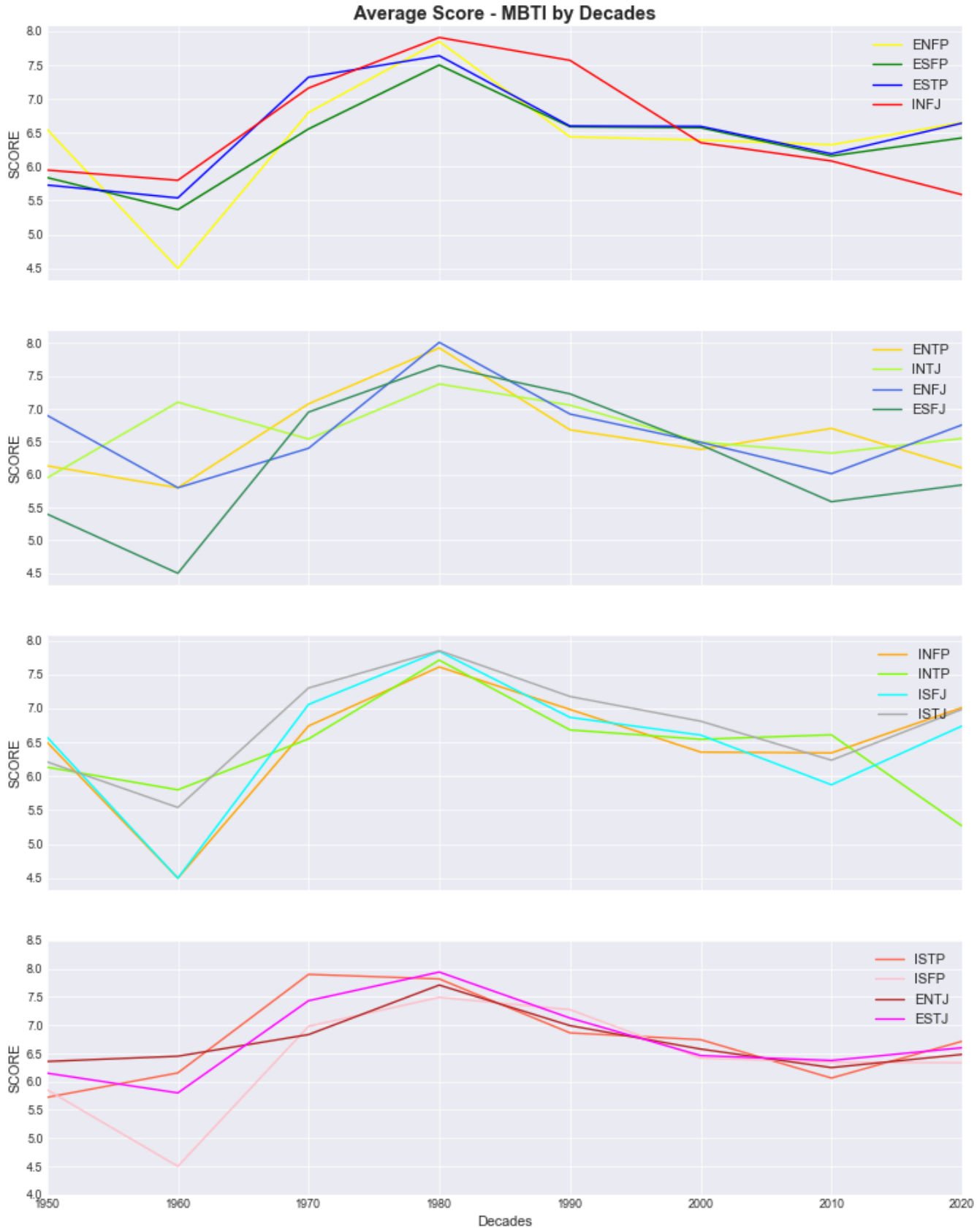


Then, I plotted a line chart illustrates the trends in average movie scores by these popular MBTI categories over the decades, from the 1950s to the 2010s. Four MBTI personality types are highlighted: INTJ, ENFP, ISTJ, and ESFJ.

However, all MBTI types show fluctuations in average movie scores across the decades. Although this might suggest that the audience's preference for a certain MBTI type is not consistent, more often, I believe this could be due to the overall change in movie ratings, rather than just the impact of MBTI types on the scores. Based on this consideration, I have drawn the score trend charts for all MBTI types according to their categories in the next part.

3. Trends in Average Movie Scores by MBTI Categories Over Decades

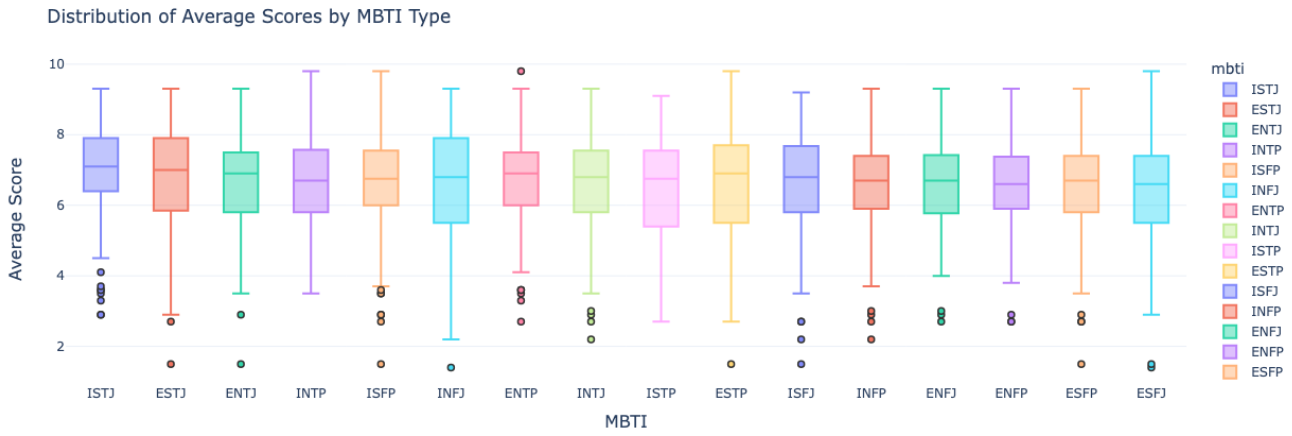
a. Trends in Average Movie Scores by MBTI Personality Type Combination Over Decades



The series of line charts depict average movie scores segmented by different groups of MBTI personality types within the MBTI framework across several decades, beginning from the 1950s up to the 2010s, each line represents one MBTI type.

Although there are slight differences in the lines of different MBTIs, overall, the trend within each MBTI category remains consistent. They collectively rise with the arrival of the 1980s movie wave and decline with the over-commercialization of movies. Perhaps to some extent, the changes in scores reflect the differences in MBTI personalities, but more so, they still reflect the overall trend in movie ratings.

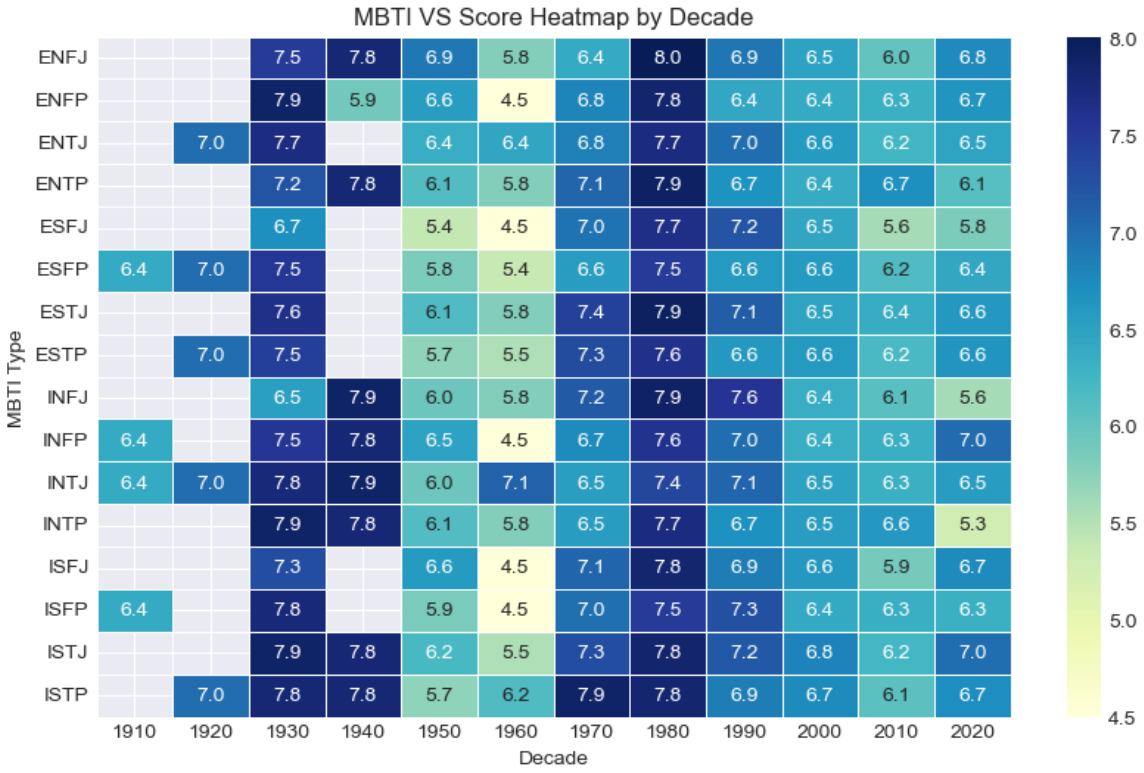
b. Distribution of Average Scores by MBTI Type



The box plot illustrates the distribution of average scores for movies, categorized by MBTI personality types. Each box represents the interquartile range (IQR) of scores for a particular MBTI type, with the median score indicated by the line within the box. The "whiskers" extending from the boxes denote the range excluding outliers, and the individual points represent outlier scores that fall outside this range.

From this box plot, we can more clearly see that although the box plots of average scores for different personalities are slightly different, overall, the range of medians remains stable around 7 points. This stability suggests that while the average scores may vary slightly from one personality type to another, there is a common baseline of quality or appeal that movies featuring these various MBTI-typed characters share. In order to see it more clearly, I create a heatmap to show the trend of MBTI and Score in decades.

c. MBTI VS Score Heatmap by Decade



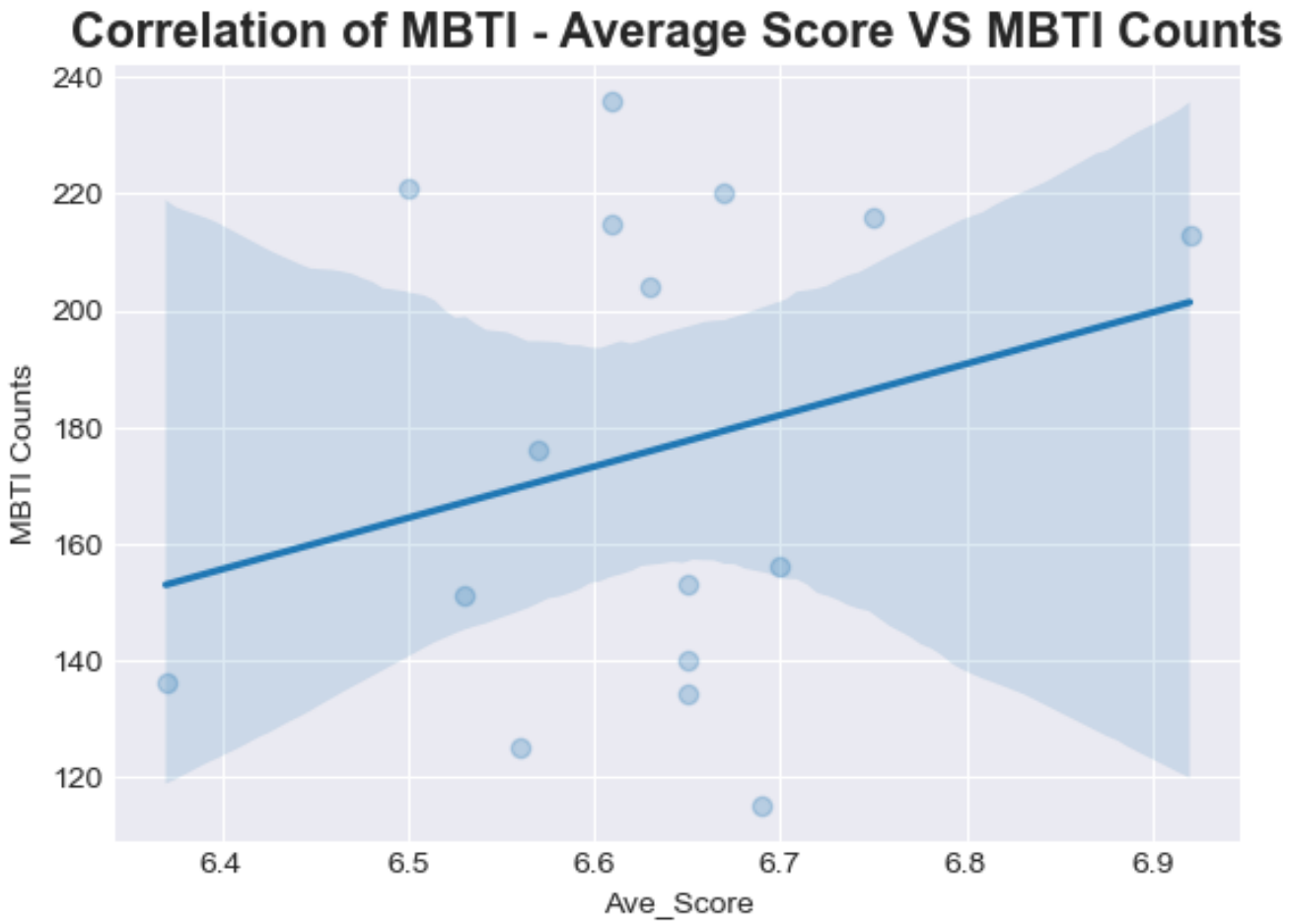
The heatmap visualizes the average IMDb scores by MBTI personality type across various decades, from 1910 through 2020. Each cell within the heatmap represents the average score for a given MBTI type within a specific decade, with the color intensity corresponding to the score—darker shades indicate higher scores, and lighter shades represent lower scores.

The average scores for most MBTI types appear to fluctuate over the decades, without showing a clear, consistent trend of increase or decrease. This suggests that the appeal of certain personality types in movies may cycle with cultural and cinematic trends. Some decades, like the 1930s and 2000s, show higher overall scores for many MBTI types, possibly reflecting periods when movie characters with distinct personality traits were particularly popular or well-received.

Even though specific MBTI characters receive higher ratings during certain periods, overall, it is difficult for me to draw a definitive trend conclusion to explain this. It seems more likely due to the presence of outstanding actors, cultural influences of specific periods, or changes in audience preferences.

Following this train of thought, I chose to undertake a correlation study between MBTI personality types and IMDb scores to determine if there are genuine correlations between IMDb scores and MBTI personality types.

d. Correlation of MBTI – Average Score VS MBTI Counts



The scatter plot with a trend line visualizes the correlation between average movie scores and the frequency of MBTI personality types featured in movies. The x axis (Ave_Score) ranges from approximately 6.3 to 7, indicating the average score of the movies, while the y axis (MBTI Counts) ranges from about 120 to 240, showing how often these MBTI types appear in the dataset.

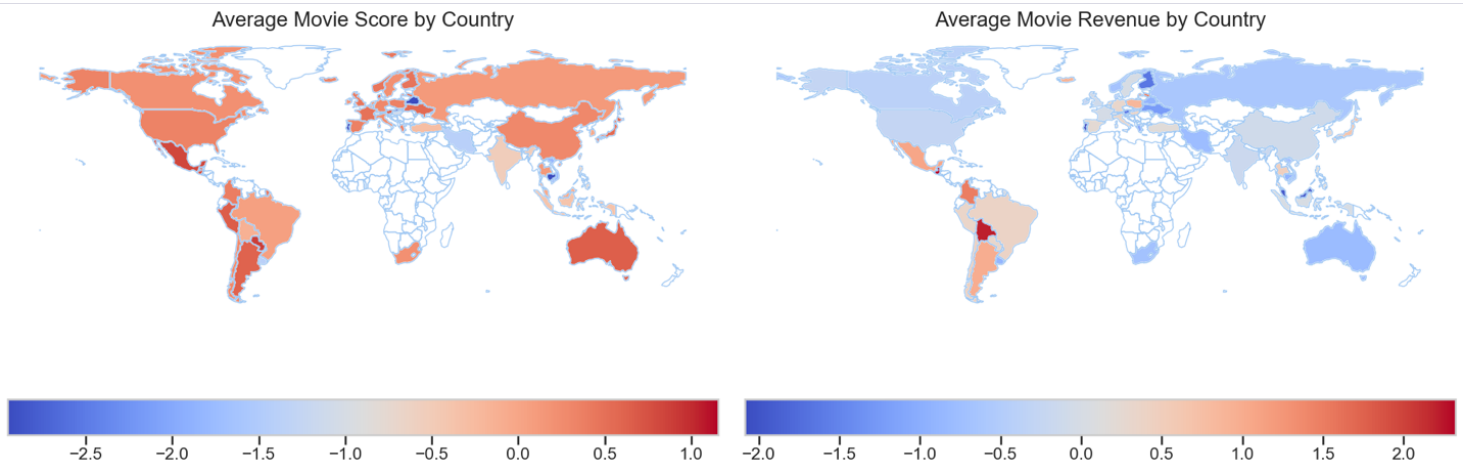
The trend line's positive slope suggests that there is a positive correlation between the average scores of movies and the prevalence of MBTI types within them, which means, as the average movie score increases, the frequency of MBTI types also increases. However, the provided R-squared value ($R^2 = 0.066$) is quite low, indicating that only a small fraction of the variability in MBTI counts can be explained by the average movie score.

This implies that while there is a positive relationship, it is weak, and other factors may play a more significant role in determining the frequency of MBTI types in movies. The shaded area around the trend line represents the confidence interval, which seems quite broad, this confirms my hypothesis and further indicates that the relationship between a movie character's MBTI type and the Score is very weak. It shows that the Score is determined by a multitude of factors, and the MBTI type is just a small part of the influence.

4) MBTI and Geographical Landscape – Xuhui 'Daniel' Zhan

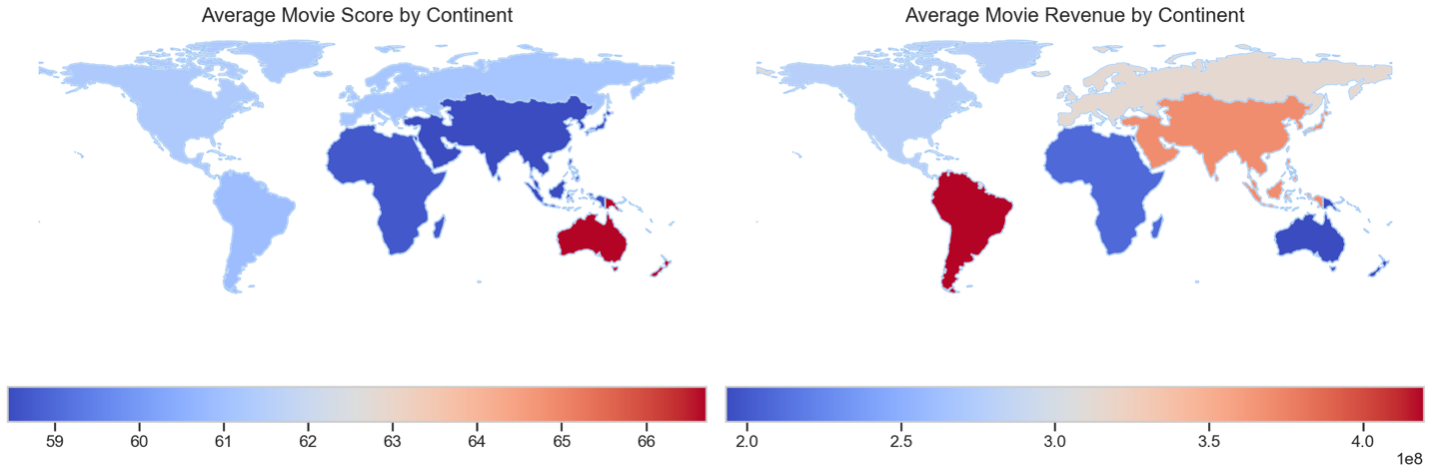
This part aims to investigate the relationship between geographic distribution and the performance of movies, both in terms of revenue and critical scores. By dissecting global data and focusing on key metrics, we seek to understand how different regions contribute to the film industry's overall landscape. The analysis progresses from a granular country-level examination to a broader continental view, eventually identifying the leading countries in terms of movie production volume.

1. Revenue and Scores by Country: A Detailed Perspective



Our initial analysis begins with a comprehensive review of movie performance metrics across various nations. Utilizing advanced data visualization techniques, we've constructed a map delineating the average movie scores and revenues by country. This map serves as a pivotal tool for identifying geographical patterns in movie success. It is observed that certain regions consistently yield higher revenues and scores, suggesting the presence of robust film industries and favorable market conditions. Conversely, other regions display notably lower averages, which may indicate market challenges or differences in cinematic output and reception.

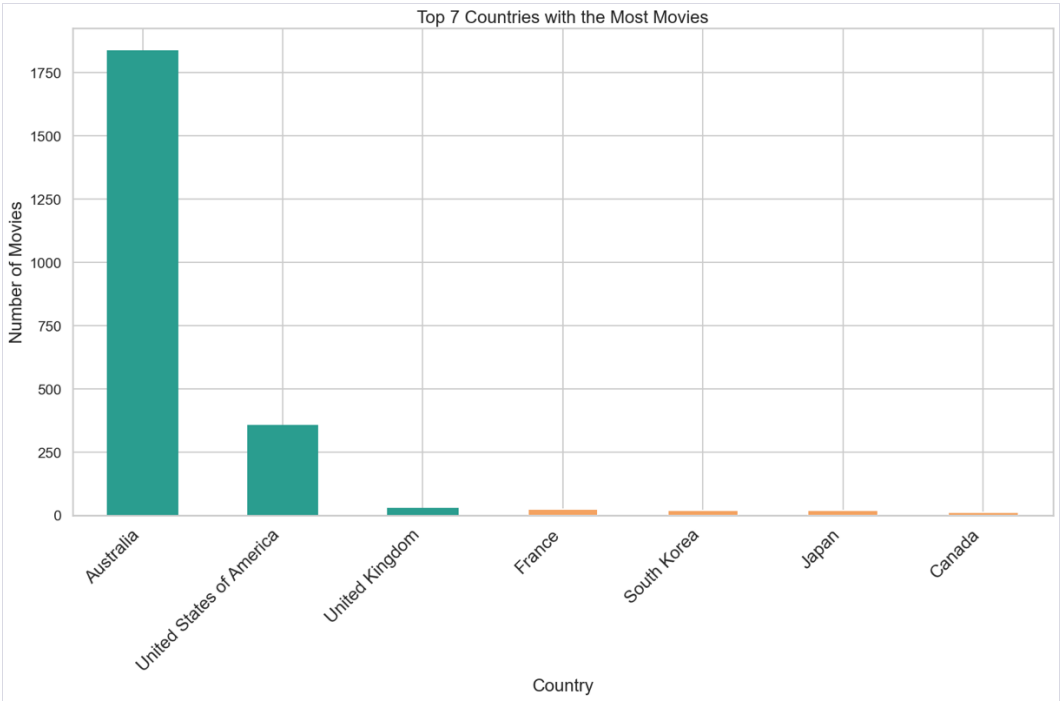
2. Continental Analysis: Understanding Regional Dynamics



Expanding our scope, we transition from country-specific data to a continental overview, aggregating the previously examined metrics to a higher level. This broader perspective reveals regional trends that offer insights into the diverse nature of the global film industry. For instance, the aggregated data points to a trend where African and Australian markets typically experience lower movie revenues, while Asia stands out with comparatively higher revenue generation. Similarly, when it

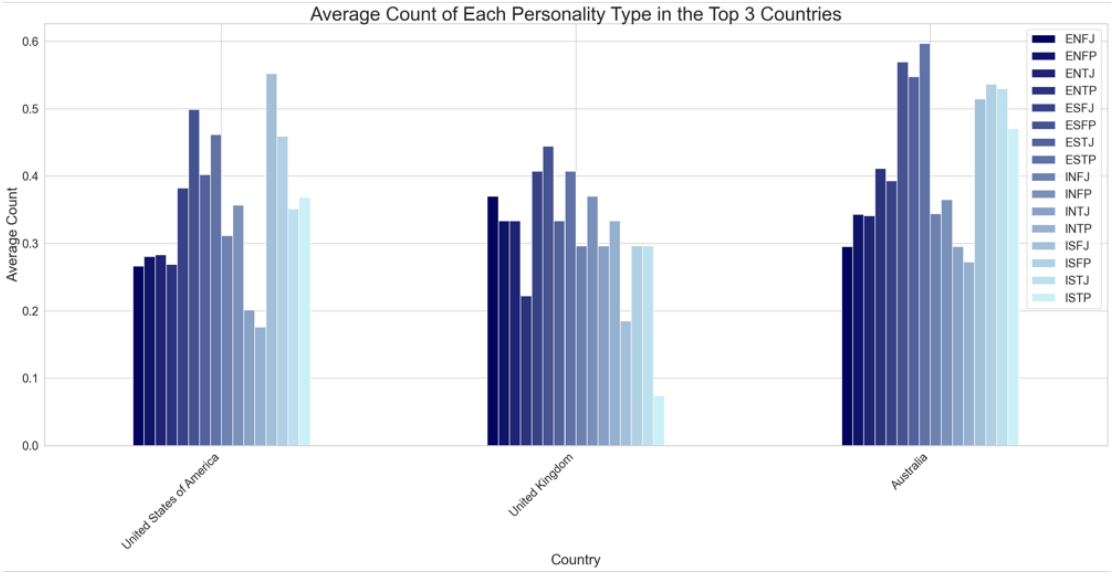
comes to critical reception, Africa and Asia are observed to have lower average movie scores. This continental analysis underscores the varying degrees of commercial success and critical acclaim, which could be attributed to numerous factors, including economic conditions, cultural preferences, and industry infrastructure.

3. Production Volume: Identifying Leading Contributors



To further our understanding of global movie production, we analyze the volume of movies produced by each country. This aspect of the report zeroes in on the number of films created, which serves as an indicator of the industry's size and output capacity. Our findings highlight Australia, the United States, and the United Kingdom as the top contributors, with the highest number of movies produced. The high production volume in these countries not only reflects their dominant positions in the global film market but also suggests a correlation between output frequency and market maturity.

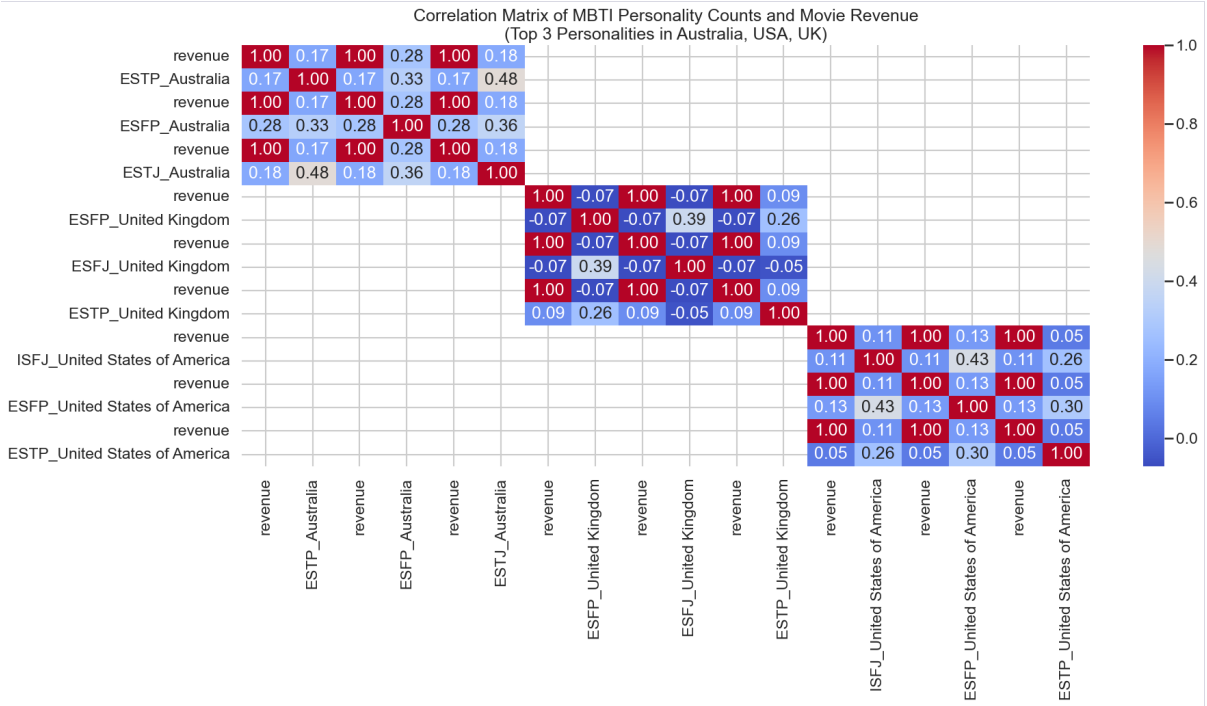
4. Personality Profiles in Film: MBTI Analysis Across Top Producing Countries



Building upon our foundational analysis of movie production and success, we now turn our attention to the actors and actresses who bring these films to life. Specifically, we examine the distribution of the 16 Myers-Briggs Type Indicator (MBTI) personality types among actors in Australia, the United States, and the United Kingdom. This segment of the report presents a colorful tapestry of personalities that contribute to the film industry's vibrancy. Our data visualization reveals a

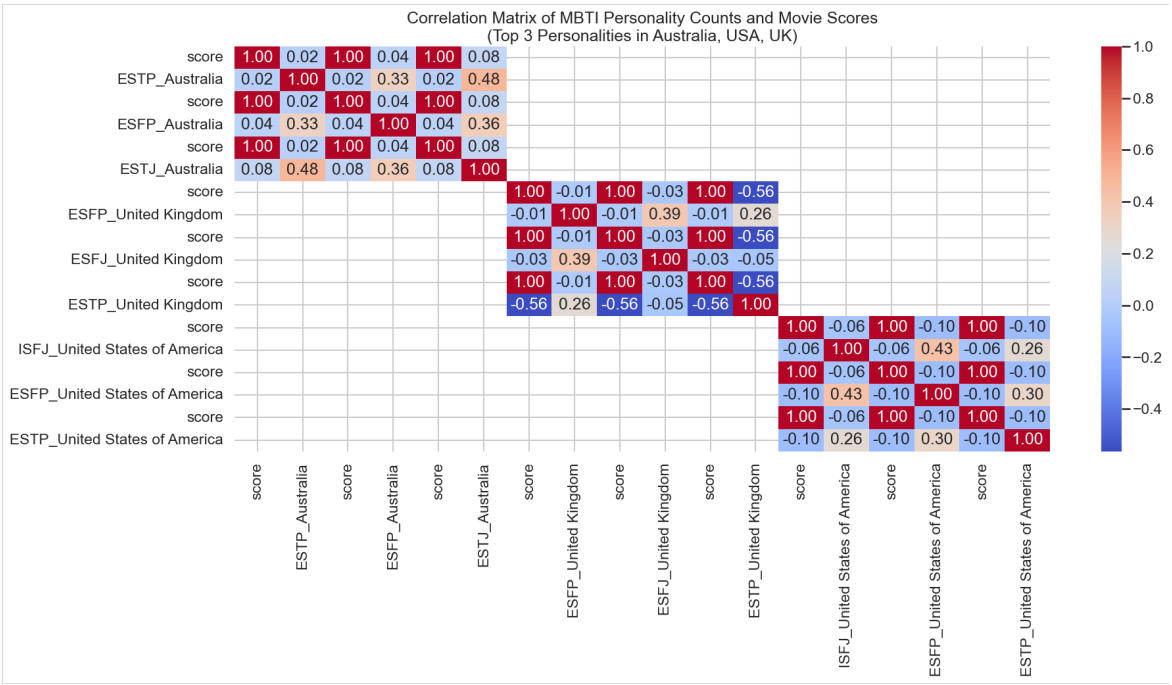
diverse array of MBTI types, with no single personality type dominating the landscape. This section argues for the richness that a variety of personality types brings to cinematic storytelling, highlighting the industry's multifaceted nature.

5. Searching for Correlations: MBTI Types and Movie Revenue



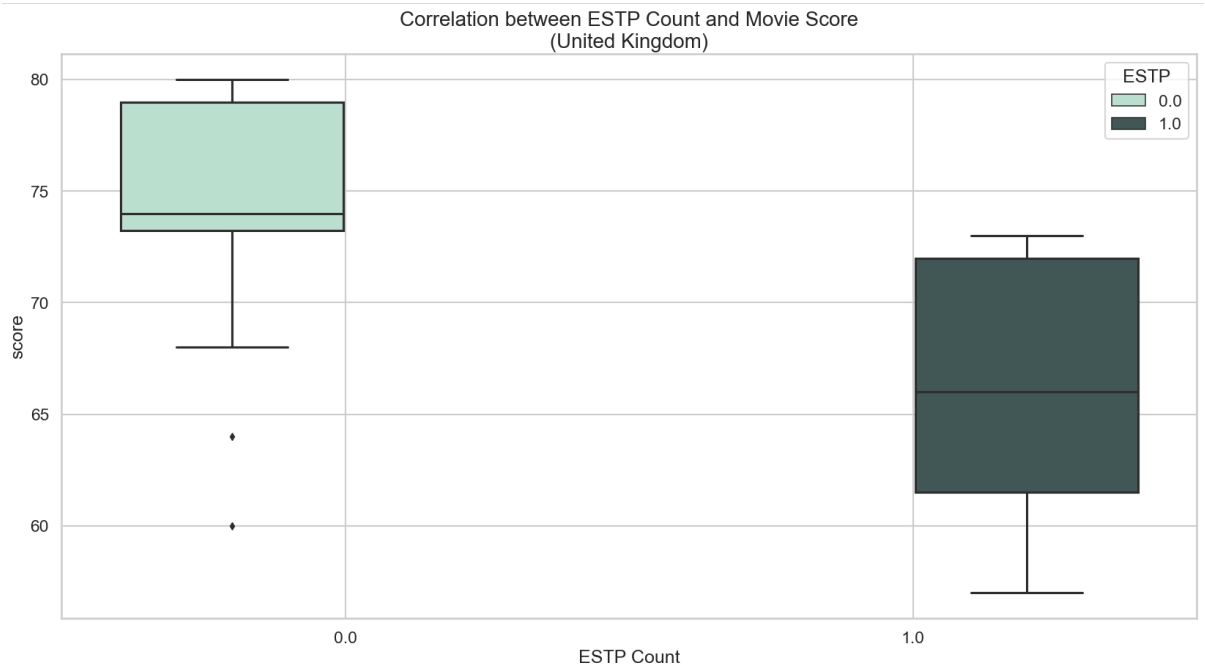
In our quest to decipher the complex formula behind a movie's financial performance, we have constructed a correlation matrix to measure the relationship between the prevalence of the top three MBTI personality types in each country and the revenue of movies produced therein. The analysis yields an unexpected finding: the absence of a significant correlation (no absolute values of scores higher than 0.5). This suggests that the MBTI personality types of actors are not reliable predictors of a movie's financial success. This portion of the report examines potential reasons for this lack of correlation, considering factors such as the diverse roles that different personality types may play in film production and the multitude of variables that contribute to a movie's revenue.

6. Beyond Revenue: MBTI Types and Critical Scores

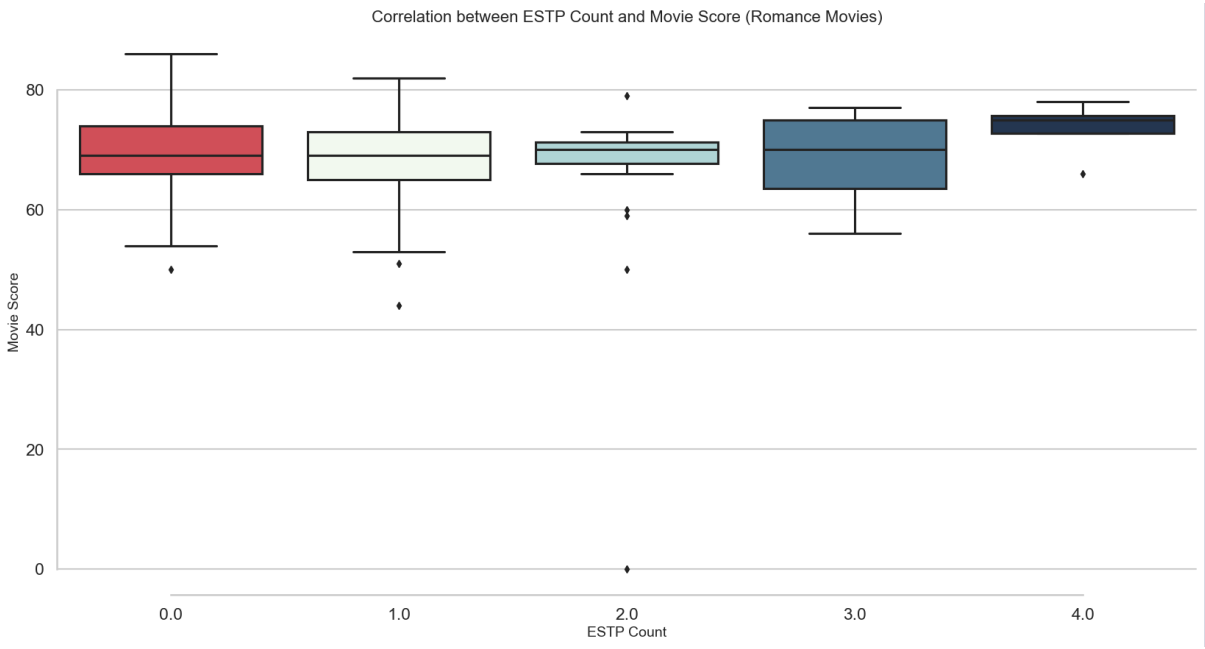


Shifting focus from financials to film criticism, we extend our correlation analysis to explore the relationship between MBTI types and movie scores. Here, we encounter a notable trend: a marked negative correlation (-0.56) between the ESTP personality type and the critical scores of movies, particularly within the UK. This finding is intriguing, as it hints at a potential mismatch between the characteristics of ESTP individuals and the preferences of film critics or audiences in the UK. However, we are careful to contextualize this within the limitations of the data, acknowledging the potential for outliers or biases due to sample size.

7. A Closer Look: ESTP Counts and Movie Scores in the UK



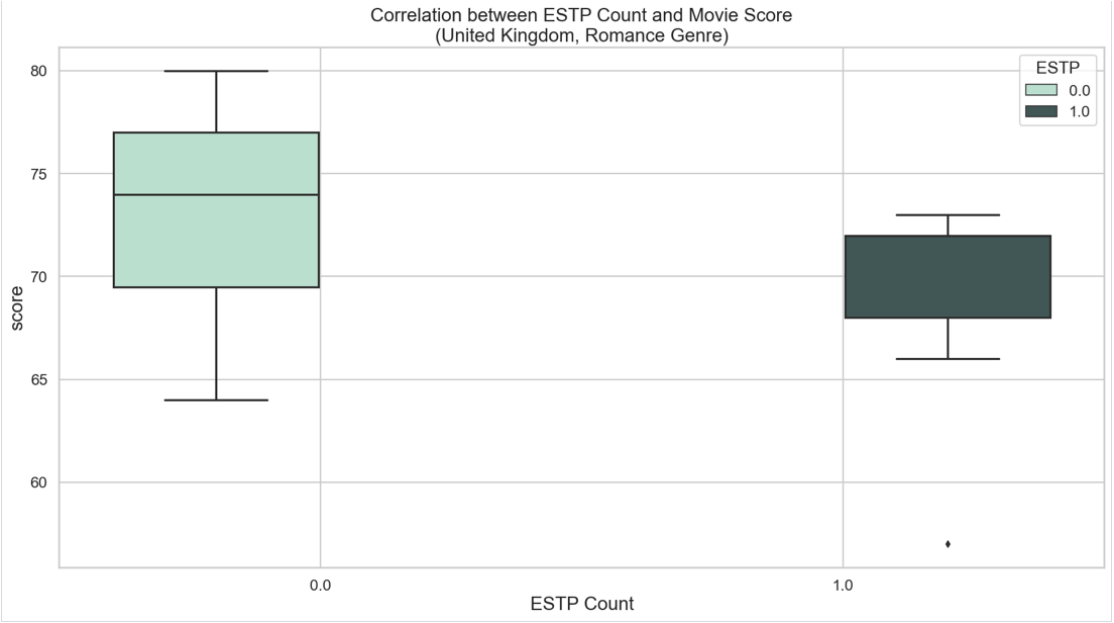
To provide a more nuanced understanding of the observed negative correlation, we employ boxplot analyses to visualize the distribution of movie scores against the ESTP count within the UK.



The boxplots of this relationship in the top 3 countries reveal a more complex picture, with no significant differences in the median scores between movies with varying counts of ESTP personalities. This suggests that the negative correlation observed earlier may not hold across a broader spectrum of films. The report discusses the implications of these findings, emphasizing the potential for overinterpretation of correlations in complex datasets and advocating for a cautious approach to drawing conclusions from such analyses.

8. Genre-Specific Analysis: The Impact of MBTI Types on Movie Scores

Diving deeper into the realm of genre-specific performance, our analysis endeavors to discern whether certain MBTI types, particularly ESTP, influence movie scores within specific genres that are prominent in the UK, such as drama and romance. We compare these genres because they are representative of a significant portion of the UK's cinematic output and are genres in which character development is often central to audience engagement and critical appraisal.



We presents a series of detailed visualizations, including boxplots, which dissect the distribution of movie scores against the ESTP count within these genres. Our investigation reveals that while the overall trend suggested a negative correlation between ESTP prevalence and movie scores, this correlation does not uniformly extend to these popular genres. For instance, in the romance genre, the presence of ESTP characters does not consistently align with lower scores. This suggests that factors inherent to the genre itself, such as narrative structure, thematic depth, and audience expectations, may play a more pivotal role in determining a film's critical success than the personality types of its actors.

Furthermore, when examining the drama genre, known for its character-driven plots and emphasis on emotional depth, the data does not support a significant relationship between ESTP presence and movie scores. This highlights the complex interplay between character portrayal, narrative context, and genre conventions in influencing critical reception. Furthermore, on the bigger landscape of the whole dataset, the correlation matrix of ESTP count and movie score has a score of 0 indicated there are no relationship between them. Which could address that negative relationship findings between ESTP count and scores in movies produced by UK is probably a biased finding due to the limitation of records in the dataset.

These findings underscore the multifaceted nature of storytelling and character utilization in films. They caution against attributing the success or failure of a movie to the MBTI types of the cast. Instead, they suggest that a successful film is more likely a product of how effectively it resonates with its audience, fulfills genre expectations, and presents its narrative and characters, regardless of the actors' personality types.

5. Conclusion

Character Traits and Genre Themes: The study found a clear link between MBTI personality types of characters and movie genres. For example, Action and Adventure movies often feature Extroverted characters, while Introverts are more common in introspective short films. Comedy and Drama frequently showcase Sensing characters, whereas Intuitive types are notable in complex Thrillers. Romance genres typically have Feeling characters due to their emotional depth, and Thinking characters are drawn to Sci-Fi and Crime for their strategic elements. Action and Adventure align with Judging personalities, while Perceiving types are often seen in Family and Short films.

Trends from 1920 to 2020: The analysis over a century shows how different MBTI types' popularity in movies correlates with societal changes. Analysts were prominent in the 1980s and 2000s, reflecting periods that valued their traits. Other types like Diplomats and Sentinels also had varied but significant representations, highlighting the evolving societal preferences.

MBTI Types and IMDb Scores: Films with high IMDb scores (8-10) often feature a diverse range of MBTI personalities, with types like INTJ being more common in these high-scoring movies. However, there's a minimal correlation between the prevalence of MBTI types and average IMDb scores. This indicates that while certain personality types might be appealing, other factors like cultural shifts, audience preferences, and industry influences play a larger role in a film's success.

Impact on Movie Success: The comprehensive analysis concludes that the MBTI personality types of characters do not significantly influence a movie's success in terms of revenue or critical acclaim. Initial correlations observed in specific genres or regions do not hold up under more detailed examination. This suggests that while MBTI types offer an interesting lens to view character design, they are not a decisive factor in a film's overall success.

6. Limitations

The analysis faces two primary limitations. First, the dataset merging process, particularly for entries without a release year, relied solely on 'movie_name'. This method may have led to data blurring, as some movies could have been overlooked due to title variations. Additionally, the process of removing duplicates, while necessary to maintain data integrity, might have excluded relevant entries. As a result, the final dataset is possibly not fully representative of all movies, impacting the overall accuracy and breadth of the MBTI character distribution across genres.

Secondly, the analysis is limited by the partial geographical data within the datasets. While some geographical information is included, it is not comprehensive, potentially skewing the analysis towards the trends and norms of overrepresented regions. This limitation is significant as movies from various geographical backgrounds often feature distinct cultural narratives and character archetypes, influencing the portrayal of MBTI types. The underrepresentation of certain regions can lead to biases in the findings, emphasizing the need for more inclusive and geographically diverse data to provide a more accurate and globally relevant analysis of the relationship between MBTI personality types and cinematic elements.

7. Future Research

Expanding the Dataset: Future studies could benefit from larger and more diverse datasets, including a broader range of genres and international markets.

Alternative Personality Assessments: Exploring other psychological frameworks or personality assessment tools could provide different insights into the relationship between actor characteristics and movie performance.

Audience Analysis: Researching audience demographics and preferences in relation to movie success could yield findings that are more actionable for filmmakers and marketers.

8. References

- https://www.personality-database.com/
- https://www.kaggle.com/datasets/subinium/movie-character-mbti-dataset
- https://www.16personalities.com/
- https://www.kaggle.com/datasets/harshitshankhdhar/imdb-dataset-of-top-1000-movies-and-tv-shows
- https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset?select=movies_metadata.csv
- https://chat.openai.com/
- https://en.wikipedia.org/wiki/Myers%E2%80%93Briggs_Type_Indicator