MSIS 670 FINAL PROJECT REPORT

TITLE: "CINEMATIC INSIGHTS:

EXPLORING ENGLISH MOVIES

AND THEIR CASTS THROUGH

IMDB DATA"

BY:

MEGHANA ARAVALA

02083168

Table of Contents:

CONTENTS	PAGE NUMBER
Introduction	Page 3
Objective of project	Page 4-5
Targeted audience	Page 5
Brief summary of project	Page 6
Visualizations	Page7-21
Dashboards	Page 22-26
User guidance	Page 27
Values of project	Page 28
Conclusion	Page 28
Tableau workbook Link	Page 29

INTRODUCTION:

Global audiences are fascinated with the world of films and television shows, and knowing the dynamics that underpin these creations can provide insightful information. I use a large dataset from the Internet Movie Database (IMDb) to explore the world of movies in our project.

Significant data about each film is included in the dataset, including titles, production firms, IMDb ratings, genres, runtime, and whether or not they won an Oscar. It also offers deep insights on the performers, their parts, and the characters they played in these shows. My goal for this business intelligence project is to use this dataset's power to extract important trends and information. My goal is to find patterns, inclinations, and connections in the film industry by using data visualization and analytic tools like as Tableau. I intend to gain practical knowledge from my investigation that will be useful to different entertainment industry stakeholders.

With the help of real-world data, this project offers a fascinating opportunity to investigate the complex world of movies and their casts. Using innovative statistical methods, i hope to derive insightful findings that may impact decision-making and provide insightful viewpoints in the entertainment and film industries.

In order to give a thorough and perceptive summary of the film industry throughout this period, i will present our research results, graphic depictions, and interpretations based on the IMDb dataset in this report.

OBJECTIVES OF THE PROJECT

The primary goal of my project is to use a sizable dataset from IMDb to fully study and comprehend the world of movies. I'm concentrating on English-language films and TV shows that were released between 2010 and 2022. My objective is to unearth fascinating information regarding various facets of these films.

In order to determine any patterns or changes in the industry throughout time, I would need to know the number of films that were released annually during this time frame. I also want to know what categories of films, or genres, are typically nominated for and winning major Oscars like Best Picture. This, in my opinion, may tell me a lot about the kinds of films that are well-liked in the business.

Comprehending the ways in which audiences interact with these films is equally crucial. I'm examining IMDb ratings, which can tell me how much or how little people thought of a particular film. In order to observe the variety of content, I'm also looking into the age ratings assigned to these films. In addition, I'm looking into how production businesses perform as well as the parts that actors and actresses play in movie casts. This will enable me to recognize prosperous businesses and comprehend the parts that various actors play in films.

My objective is to compile all of this data, do a thorough analysis, and then display it in an understandable manner. By doing this, I intend to offer insightful information that will help streaming services, production businesses, filmmakers, and anybody else who wants to learn

more about the movie industry. In the end, I want to provide a thorough overview of the film world over these years, illuminating global cinematic trends, industry norms, and audience preferences.

TARGETED AUDIENCE & IMPORTANCE OF THESE QUESTIONS:

Production Companies and Filmmakers: Making educated judgments about production, content creation, and market targeting can be aided by insights on trends, genre preferences, and audience engagement.

Film distributors and streaming platforms: These might benefit from knowing audience preferences, age ratings, and country-specific IMDb ratings in order to better customize distribution plans and content recommendations.

Performers and artists: Knowledge of cast composition and character diversity can help them understand their parts in films and may even have an impact on their career choices.

Scholars and Researchers: This project can be used as a resource for scholarly research on audience behavior analysis, entertainment industry dynamics, and movies.

General Audience and Movie enthusiasts: By providing a deeper understanding of movie trends, awards, and industry dynamics, the project's visualizations and insights can pique the interest of casual movie aficionados.

BRIEF SUMMARY OF MY PROJECT JOURNEY:

Data Collection: The research project started by obtaining a large dataset from IMDb that included English-language films and television shows that were released between 2010 and 2022. I was able to acquire this dataset, which included a wealth of movie-related data, such as titles, genres, ratings, cast bios, production firms, and more.

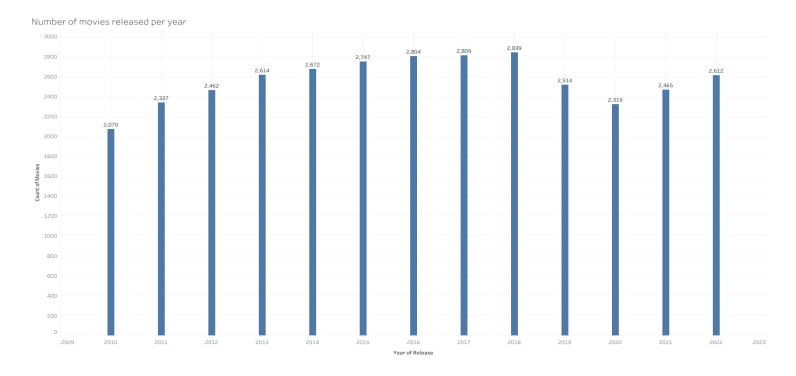
Analyses and Project Schedule: Different levels of analysis were considered when creating the project schedule. I started with exploratory analysis, looking at annual trends in the release of movies, looking into prizes based on genre, closely reviewing measures related to audience involvement such as IMDb ratings and age rating distributions, and evaluating the work of production businesses. Every stage of the research was methodically planned in order to derive valuable conclusions from the data.

Highlights and obstacles: During the project, I had a number of obstacles, major among them being managing substantial amounts of data and guaranteeing data integrity during the cleaning and preparation stages. It was crucial to make sure the dataset was accurate and consistent. Nonetheless, these difficulties helped me hone my problem-solving and data-handling skills.

Highlights of the project included identifying patterns in the release of movies over time, learning unique insights into audience participation through IMDb ratings, and finding fascinating relationships between genre preferences and award nominations. These discoveries greatly influenced the project's storyline and offered insightful information about the film industry.

VISUALIZATIONS:

1. NUMBER OF MOVIES RELEASED PER YEAR



- Yearly Film Releases: This graph illustrates the variations or patterns in the total amount of films released in various years. Abrupt increases or decreases signify noteworthy occurrences or modifications within the sector.
- We can observe a steady growth of releases till the year 2018 and there are few ups and downs from them till the year 2022.

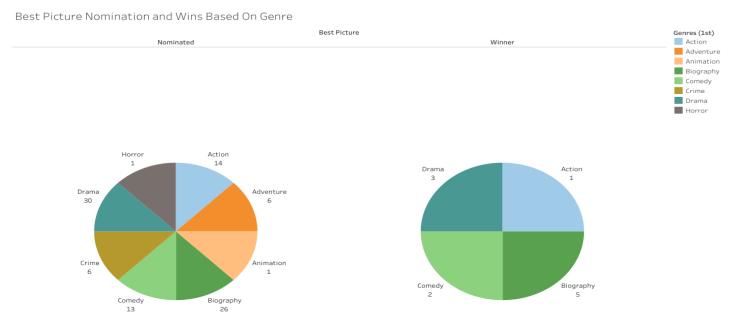
Importance and Insights:

• **Planning Strategically:** This data is essential for strategic planning for industry players, particularly for distributors, streaming platforms, and production companies. Recognizing

trends can help with better resource allocation, marketing strategy optimization, and release scheduling.

 Comparative Analysis: Stakeholders can assess their own performance in relation to industry norms by comparing trends over time, which enables them to establish reasonable objectives and benchmarks.

2. BEST PICTURE NOMINATION AND WINS BASED ON GENRE



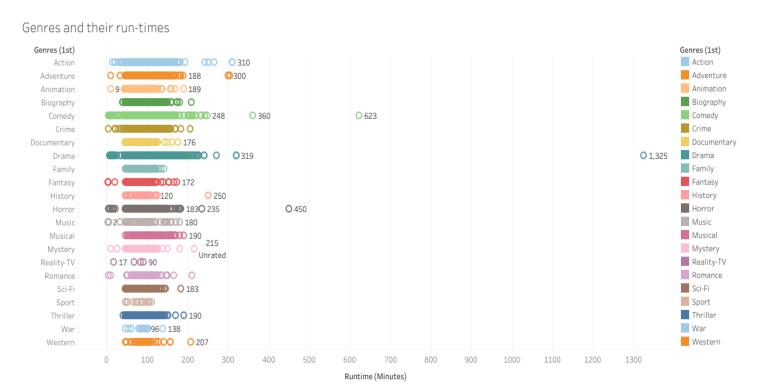
Distribution of Nominations and Wins Based on Genre: The pie chart shows the percentage of nominations and actual wins for Best Picture in each of the major genres.
It shows which genres are more likely to win awards and acquire more nominations.

- A higher percentage of nominations for a certain genre suggests that award committees
 have given it more attention. A genre with a larger slice in the "Wins" category is more
 likely to have won the Best Picture Oscar.
- We can see 'Drama' and "biography' genre movies are mostly nominated with the movie count of 30 and 26 respectively and 3 from drama and 5 from biography has won at the Oscar level.

IMPORTANCE AND INSIGHTS

The pie chart makes it easier to see which genres receive greater recognition or
preference from major prizes. For filmmakers and other industry professionals hoping for
critical acclaim and award recognition, this information is invaluable. Knowing which
genres are more likely to be nominated or win can help production companies and studios
make decisions about their investment strategies and content creation schedules.

3. GENRES AND THEIR RUN TIMES

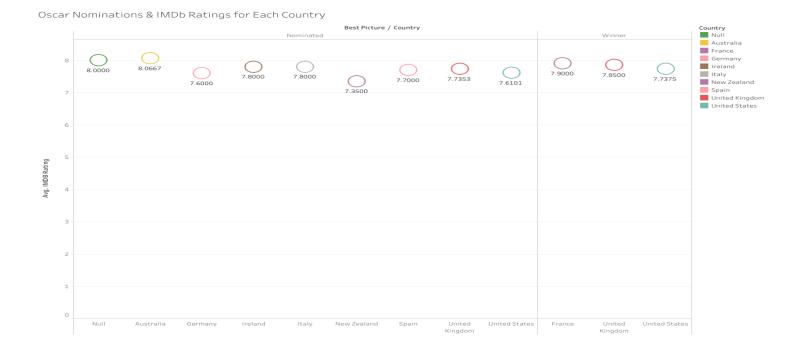


- Scatter Distribution: The scatter plot shows the distribution of each film's runtime
 among its various categories. It displays the variety and dispersion of runtimes in every
 genre.
 - The color filter helps in differentiating each genre and helps in better understanding.
 - Runtime(minutes) is represented on the x axis and the genre types on y-axis. In the
 results we can see 'Drama' genre with the highest runtime with '1325 min'. Dramas
 typically have longer run times, which may provide directors more creative freedom
 to depict complex themes that appeal to viewers looking for poignant and
 contemplative stories.

IMPORTANCE AND INSIGHTS:

- Understanding runtime patterns unique to a genre is essential to controlling audience expectations. It's possible that some genres have runtimes that are normally shorter or longer; this knowledge helps cater to audience preferences for marketing campaigns and content presentation on streaming platforms
- Understanding genre-based runtime conventions helps producers and filmmakers make innovative choices about pacing, editing, and screenplay for particular genres.

4. OSCAR NOMINATIONS AND IMDB RATINGS OF EACH COUNTRY



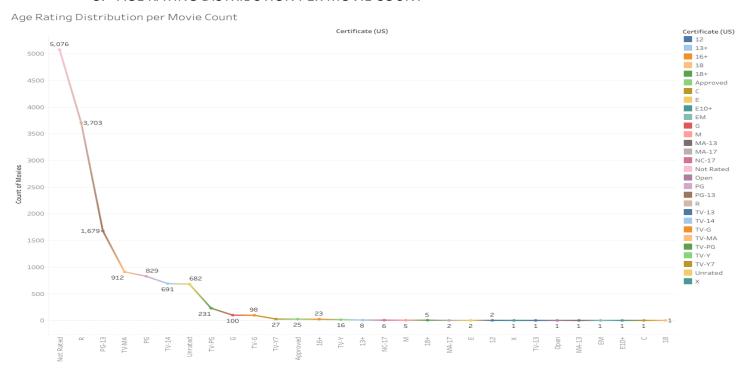
- Examining the relationship between IMDb ratings and Oscar nominations provides
 information on whether films with high ratings typically get nominated, suggesting a possible
 relationship between industry recognition and critical praise.
- The graphic aid facilitates a comparative evaluation of nations according to their box office
 receipts and nominations, emphasizing which countries routinely generate highly acclaimed
 films honored with esteemed accolades.
- We can see 'Australia' has the highest IMDB rated movie of '8.6' when compared to other states and was nominated but failed to bag an Oscar Award. But the movie which was rated less than this has managed to achieve the award

 By this analysis, we can come to an understanding that IMDB rating does not help in winning an Oscar.

IMPORTANCE AND INSIGHTS:

 Knowing that there is no correlation between the Oscar and the Imdb rating allows filmmakers to focus entirely on the process rather than worrying about reviews and criticism.

5. AGE RATING DISTRIBUTION PER MOVIE COUNT



 Age Rating Frequency: The line graph indicates which ratings are more common in the dataset by displaying the frequency or count of movies connected to each age rating group.

- Each color represents each rating certificate.
- It provides an overview of the content's acceptability for various audience age groups by highlighting the percentage of movies categorized under particular age ratings.
- We can observe majority of movies are 'R' and 'PG-13' rated movies when compared to all other.

IMPORTANCE AND INSIGHTS:

- Understanding the age rating distribution facilitates content filtering and parental guidance by assisting parents and content platforms in determining which movies are appropriate for certain age groups.
- Using this information, streaming services and distributors can build their own content libraries that suit a wide range of audience tastes by suggesting films according to age-appropriate categories.

6. OSCAR NOMINATIONS AND WINS BY CONTINENT



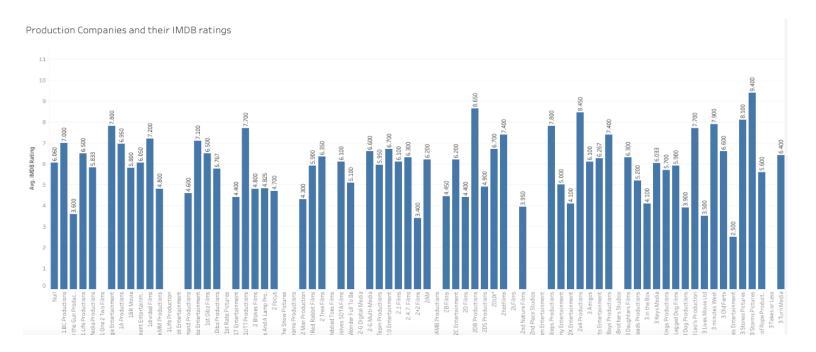
Identification of Geography: The placement of each nation on a map, with symbols standing in for Oscar nominations and victories, offers a graphic depiction of the distribution of recognition in the film industry throughout the world.

 Each color represents a continent and as we can observe the Oceania, America, and Europe continent have been nominated but the Americas and Europe has declared as the winner.

IMPORTANCE AND INSIGHTS:

- The map acknowledges the variety of cinematic achievements made across the globe
 by offering a visual depiction of the cultural diversity and contributions of many
 nations to the global film landscape.
- By providing insights into the areas recognized for receiving critical praise and recognition, the symbol map aids in the identification of nations or locations that have made major contributions to the cinema industry.

7. PRODUCTION COMPANIES AND THEIR IMDB RATINGS



- Bar Heights: The height of each bar, which stands for a production firm, is correlated with the
 average IMDb rating of the films the company is involved with. It makes rating comparisons
 between several production firms possible.
- Taller bars represent firms with higher average IMDb ratings, indicating that their films often get higher viewer ratings.

IMPORTANCE AND INSIGHTS:

 The bar graph makes it easier to evaluate the effectiveness and caliber of films made by different production companies, which helps determine which ones regularly generate highly regarded films.

- Based on the track record of performance of production businesses, filmmakers, distributors, or investors can use this information to make well-informed judgments about investments or collaborations.
- By showcasing businesses with higher-grossing films, it offers insights into the competitive environment of the film industry and may have an impact on industry positioning and strategy.

8. CHARACTERS PLAYED BY ACTORS/ACTRESSES

	W	What did they do?
erson Name	actor	actress
Chainz	Freddie 6.00	
D Na'Tee		Daisy 13.00
the General	Boneman 8.00	
2Rainer		Ada Wong 4.00
1 Savage	21 Savage 14.00	
2nd Street Pete	Red 7.00	
0 Cent	SO Cent Deon Enson Levoux Hush Jordan Mains Lamar Lionel Malo Pimp Clate Johnson Rich Sonny The Pharmacy Tino 19.00 1.00 4.00 6.00 5.00 5.00 5.00 1.00 16.00 1.00 1.00 1.00 7.00 3.00	
70 Shake		Nadia 11.00
00lbs of Guwap	300lbs of Guwap 13.00	
60	Jimmy 5.00	
Big Boo' Brett Nitchell	Tractor Mechanic 11.00	
Big' Jack Little	Damien 10.00	
Big' LeRoy Mobley	Jamican 13.00	
Bra' Jeffrey Idlwazilwana	Bra Z 19.00	
Cowboy' Matt	Jim	
opewell -Gotti' Eric	1.00	
ohnson	Murda 8.00	
Freeway' Ricky	Reggie	

The portrayal of roles In brief:

 Actor/Actress List: This table gives a summary of the many parts and contributions played by actors and actresses in different movies by listing them next to the characters they have played.

IMPORTANCE AND INSIGHTS:

- By presenting the variety of characters that actors and actresses have played in various films, it helps to comprehend their adaptability and range.
- Using this data, casting directors and filmmakers can determine whether an actor is a
 good fit for a specific role, how much experience they have playing a specific character
 type, or whether there is any possibility of typecasting.
- This table can be used as a resource for enthusiasts or fans to delve deeper into the range
 of roles portrayed by their preferred actors or actresses, leading to a greater appreciation
 for their performances.

9. TITLES AND THEIR PLOTS

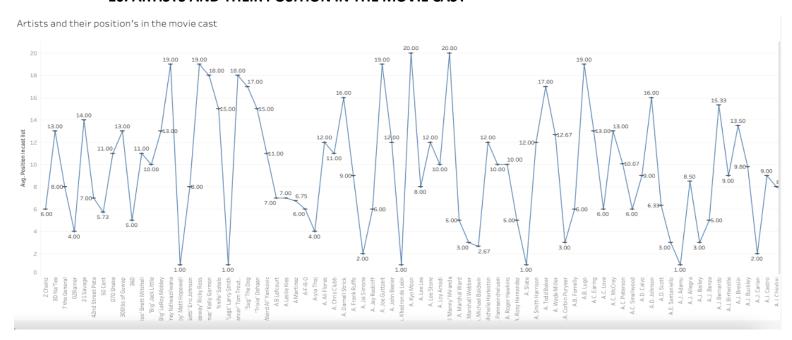
itie s c	and th	heir Pl	ots																						
44 0.	.357	0s & 1s	1	1+1=	1 Angry Black M		1 Chance 2 Dance	1 Interro	1 Message	1 Mile to You		1 Night in San D			1 Out of 100	1 Versus		1-800-H	1-Jan	1-Oct	1,000 Times	1,001 Ways t	1.7 Alpha	1BR	1
small Or www.isia.com. or for for for for for for for for for	on her wedding ight, ade wakes	James lives easy Life 2.0. But when his	Early morning silence is broken	A Science Fiction feature about a	Mike	Shifting from one pocket to another, from one man's drama to another, a lowly dollar bill, 'one buck' takes us on an odyssey throug	When a	Detecti	Tells the story of Becca, a beautiful 28 year-old woman whose perfect life is brought to an abrupt halt after she acciden		Past and present collide when two couples explore love over the course of one	A former reality star and her best friend are invited to San Diego for	Six lives change when	Runaway teen Lexi find herself alone and pregna		After Malia's father mysteri	Brian	When 13-year	Maybe	As	Four friends pay the ultimate	In the near future, where	With her social life seemin	Sarah tries to start	S V (i

• **Title and their Plot:** This text table links film titles to corresponding plot summaries, providing a handy, one-stop shop for quick access to summaries of various films.

IMPORTANCE AND INSIGHTS:

- For those looking for rapid access to plot synopses of particular films, it acts as a catalog or reference, facilitating content exploration and selection.
- This table can be used by scholars, analysts, or critics to examine narrative structures,
 storytelling approaches, or themes in various films.
- Using plot summaries, viewers looking for details about a film's idea or story can use this table to help them choose which movies to see.

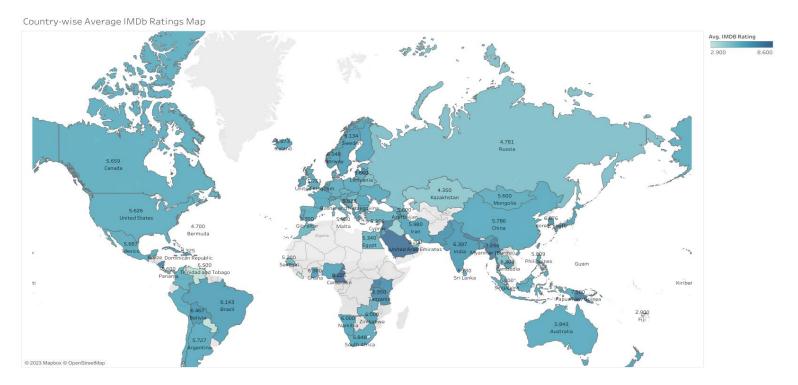
10. ARTISTS AND THEIR POSITION IN THE MOVIE CAST



 Positional Trends: The graph's lines represent each artist's place in the cast of different films. It displays whether they usually receive top billing or are ranked lower in the cast list, demonstrating the consistency or unpredictability of their billing.

- The X-axis is used to show films or productions.
- Y-axis: Shows an artist's place on the cast list (e.g., major roles, supporting roles, top billing).
- Line Paths: Each artist's line trajectory through multiple films indicates where they are in the cast and gives an indication of how consistently or inconsistently, they are cast.
- Determining which artist ranks highest among all the others is a difficult process due to the large dataset and plenty of artists.
- It helps to comprehend a performer's casting habits by disclosing if they routinely land major roles or are cast in supporting or smaller roles in films.
- In order to help artists make wise career decisions, casting directors, managers, and filmmakers can examine an artist's career trajectory and spot any changes or patterns in their cast positions over time.
- Filmmakers might choose the best actors or actresses for certain roles or billing positions
 by considering an artist's typical casting position.

11. COUNTRY WISE AVERAGE IMDB RATING MAP



- Visual Comparison: The map makes it easy to compare the IMDb ratings of various nations quickly and visually, showing which areas often have better or lower average ratings for particular films.
- Color Gradients: Each color on the map represents a range of IMDb ratings, and color gradients are used to show different hues or tones between nations. Darker hues could, for instance, correspond to greater average ratings, whereas lighter hues would correspond to lower ratings.
- Data Particular to Each Country: The average IMDb rating of each nation is shown as a tint on the map. Lower-rated countries are represented by lighter hues or shades, whereas higher-rated countries seem darker or with deeper shades.

Regional Trends: It provides information about which regions are known to produce films
with higher or lower IMDb ratings, providing insight into the caliber of films produced
there or how well-received they are by audiences.

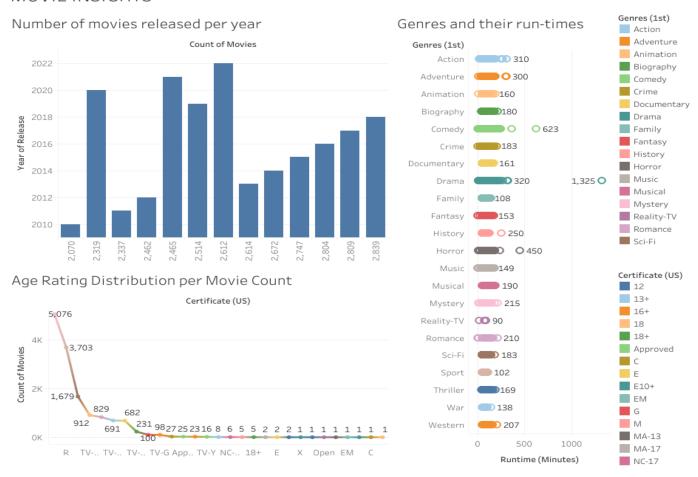
IMPORTANCE AND INSIGHTS:

- It helps stakeholders identify areas recognized for higher-rated films by providing insight into regional trends in movie quality perceptions.
- Based on average IMDb ratings, filmmakers, distributors, or streaming services can
 utilize this data to analyze audience preferences or market potential in various
 nations.
- By emphasizing areas with higher audience-rated films, the map helps with strategic decision-making related to content distribution, marketing, or investment.

DASHBOARDS

1. MOVIE INSIGHTS DASHBOARD:

MOVIE INSIGHTS



HOW IS THIS HELPFUL:

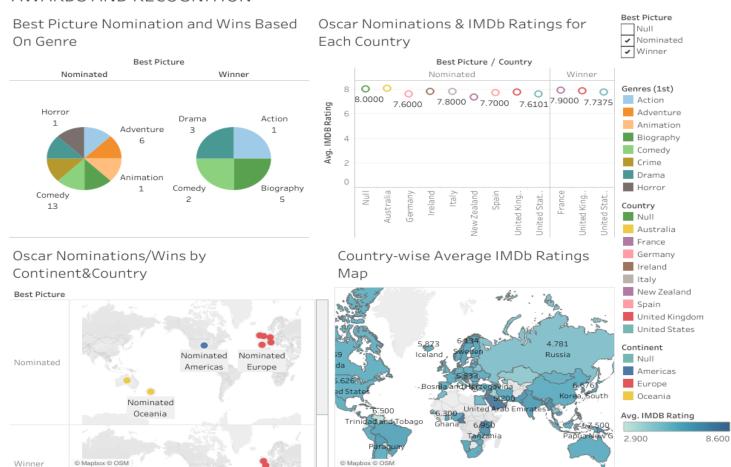
 This "Movie Insights" dashboard is beneficial for Film business insiders, filmmakers, distributors, and movie buffs, which displays yearly release patterns, age distribution by movie count, and genre runtimes.

- Industry professionals can benefit from knowledge regarding age appropriateness, genre-specific runtimes, and annual release patterns, which can help with audience targeting and strategic planning.
- For improved movie selection, viewers and enthusiasts may quickly grasp trends
 in movie releases, age-appropriate content, and runtime preferences specific to a
 particular genre.

All things considered, this dashboard is a helpful tool that helps industry participants make decisions and improves audience movie-watching experiences by providing crucial information on trends and features in the film business.

2. AWARDS AND RECOGNITION

AWARDS AND RECOGNITION



• The "Awards and Recognition" dashboard consists of four worksheets that show the Oscar nominations and wins by genre, the Oscar nominations and wins by continent and region, and a map of the average IMDb ratings by country. Users can now select a country and receive data about it from all three sheets by using an action that has been included especially for the "Country-wise Average IMDb Ratings Map" worksheet.

How does it work:

- Interactive Country Selection: On the "Country-wise Average IMDb Ratings Map" worksheet, users have the option to select a particular nation on the map.
- Data Filtering and Information Retrieval Across Sheets: Selecting a nation initiates data filtering, which provides country-specific information on the other three sheets.
- Consolidated Insights: Viewers obtain a thorough rundown of the accomplishments of the chosen nation, including Best Picture nominees, Oscar nominations and wins, IMDb ratings, and rating distribution by region.

IMPORTANCE OF THIS DASHBOARD

- Users are given a comprehensive overview of a nation's film industry accomplishments and accolades, which makes it easier to comprehend the contributions that nation has made to cinema.
- Comparative Analysis: Lets users assess how various nations have performed in terms of IMDb ratings, wins, and award nominations.

Strategic Decision-Making: business experts can use this function to evaluate the effects
that various countries have on the film business, which can help them make decisions
about partnerships, investments, and content purchases.

3. CAST AND CREW ANALYSIS

CAST AND CREW ANALYSIS



Production Companies (1st) 15 Avg. IMDB Rating 10 6.500 6.050 Bullet in the Gun Pro.. 1 Media Productions 1 One 2 Two Films tooge Entertainment 1A Productions 1BR Movie 1divided Films Lin MM Productions 1Shot Entertainment 1st Command Producti. 1st Dibz Entertainment 1st Dibz Productions 1st Rate Pictures Department Entertai.

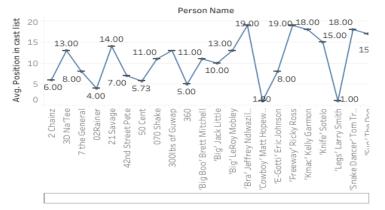
Characters Played by Actor/Actress

			What did	they do?
	Person Name	actor		actress
	2 Chainz	Freddie 6.00		
7 OU	3D Na'Tee			Daisy 13.00
	7 the General	Boneman 8.00		
2 Rinbac Eilme	02Rainer			Ada Wong 4.00
	21 Savage	21 Savage 14.00		
	42nd Street Pete	Red 7.00		
	50 Cent	50 Cent Deo 19.00 1.00	•••	
	070 Shake			Nadia

Title's and their Plots

Title 1+1= 1 Angry 1 Chance 0.44 0.357 0s & 1s 11 Black M.. 1 Buck 2 Dance Shifting When a Mike A small On her James Α wedding lives Science Anders.. from one town is easy Life riddled night, pocket 2.0. But feature with a Jade to another. series of wakes when his about a murders up comput.. young from one targeti.. holdina woman man's the hand drama to of her tormen.. another, dead a lowly husband dollar after she bill, 'one buck' ta. narrowl.

Artists and their position's in the movie cast



• The "Cast and Crew Analysis" dashboard includes four informative worksheets: the roles that artists play in the cast of movies, the plots of the titles, the characters that actors and actresses play, and the average IMDb ratings of production organizations.

HOW IS THIS HELPFUL

• FOR FILMMAKERS AND PRODUCTION HOUSES:

- Casting Decisions Information from "Artists and their positions in the movie cast" can be used to gauge an artist's level of notoriety, which can help with casting choices for roles in future projects.
- Script Development: Plot summaries are provided in the "Titles and their plots" worksheet, which aids in the development of scripts and serves as an inspiration for new ones.

• FOR MOVIE ENTHUSIASTS:

- Viewers learn about the roles that their favorite actors have played in a range of films as well as the casting decisions made for them.
- With the help of plot descriptions provided by "Titles and their plots," viewers can select movies according to their preferred storylines.

A wide range of users are served by the "Cast and Crew Analysis" dashboard, including researchers, industry experts, hobbyists, and filmmakers. It offers comprehensive insights on character portrayal, casting, storylines, and production quality, so facilitating research efforts, decision-making, and the enjoyment of film for enthusiasts.

USER GUIDANCE

Making Use of the Dashboard: Navigation

- By choosing from a variety of spreadsheets that reflect different analyses, users can traverse the dashboard.
- Each worksheet has buttons or tabs that make it simple to access various insights.
- When users access the "Country-wise Average IMDb Ratings Map," in the dashboard they
 can use the interactive feature to choose a desired country.
- The interactive function allows users to select a country from the map by clicking on it.
 and provide all the information related to that country.

Step-by-Step Exploration:

- Selection Action: By clicking on a country in the "Country-wise Average IMDb Ratings
 Map," in the "awards and recognition" dashboard users trigger a data filtering action.
- **2. Data Insights Across Worksheets:** The executed action dynamically filters data in that dashboard's linked worksheets after the user selects a country.
- 3. Detailed Country Analysis: Visitors may explore the accomplishments and contributions of a specific country in a number of areas, such as award nominations, IMDb ratings, continental representation, and more.
- **4. Comparative Analysis:** By choosing and examining different nations, users can evaluate and contrast how well each nation has performed and gained fame in the film business.

VALUES OF THE PROJECT:

The following are the study's values and possible conclusions or recommendations:

- Data-Driven Insights: Using the IMDb dataset, comprehensive insights about film releases, genres, reviews, actors, production firms, and national accomplishments are produced.
- 2. Comprehensive Analysis: Providing a detailed insight of audience preferences, industry trends, and movie performance indicators by presenting several perspectives through a variety of visuals.
- **3. Interactive Dashboards**: Personalized analyses catered to particular interests or queries are supported by dynamically navigating, exploring, and filtering information by users.
- **4.** Helping filmmakers, production companies, distributors, and analysts make well-informed decisions about casting, content creation, market targeting, and industry partnerships is known as Decision Support for Industry Professionals.

CONCLUSION:

In summary, the project's value is found in providing a wealth of information and resources for both industry participants and hobbyists. It acts as a creative platform for market research, content development, and decision-making, moving the film business in the direction of data-driven tactics, better content curation, and increased audience engagement. The project's findings place a strong emphasis on using data-driven insights to inform strategic choices and support ongoing adaptation in the always changing film business.

MY TABLEAU WORKBOOK LINK:

I am including a link to my tableau workbook, which has all of my visualizations, for your better understanding.

https://prod-useast-b.online.tableau.com/#/site/msis670fall2023/workbooks/1064564