**Project- Phase 3: Dashboard Implementation**

**Title: CineMetrics: A Data Exploration into Movie Dynamics**

**Rishi Tadavarthi(1220383593)**

**Shirisha Nampally(1227792971)**

**Angelina Telidevara(1228354935)**

**Aditya Raj(1226878577)**

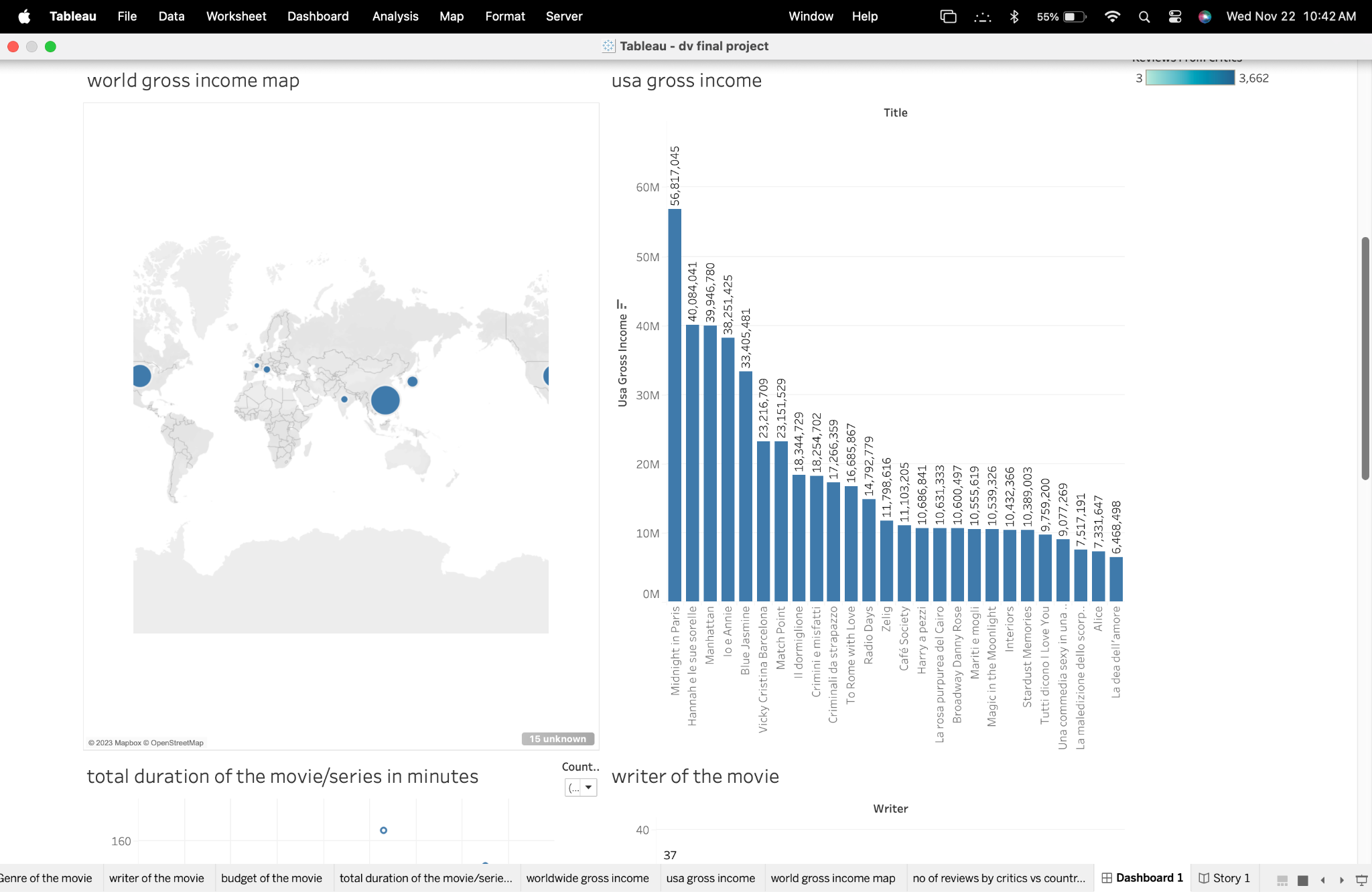
**Arizona State university**

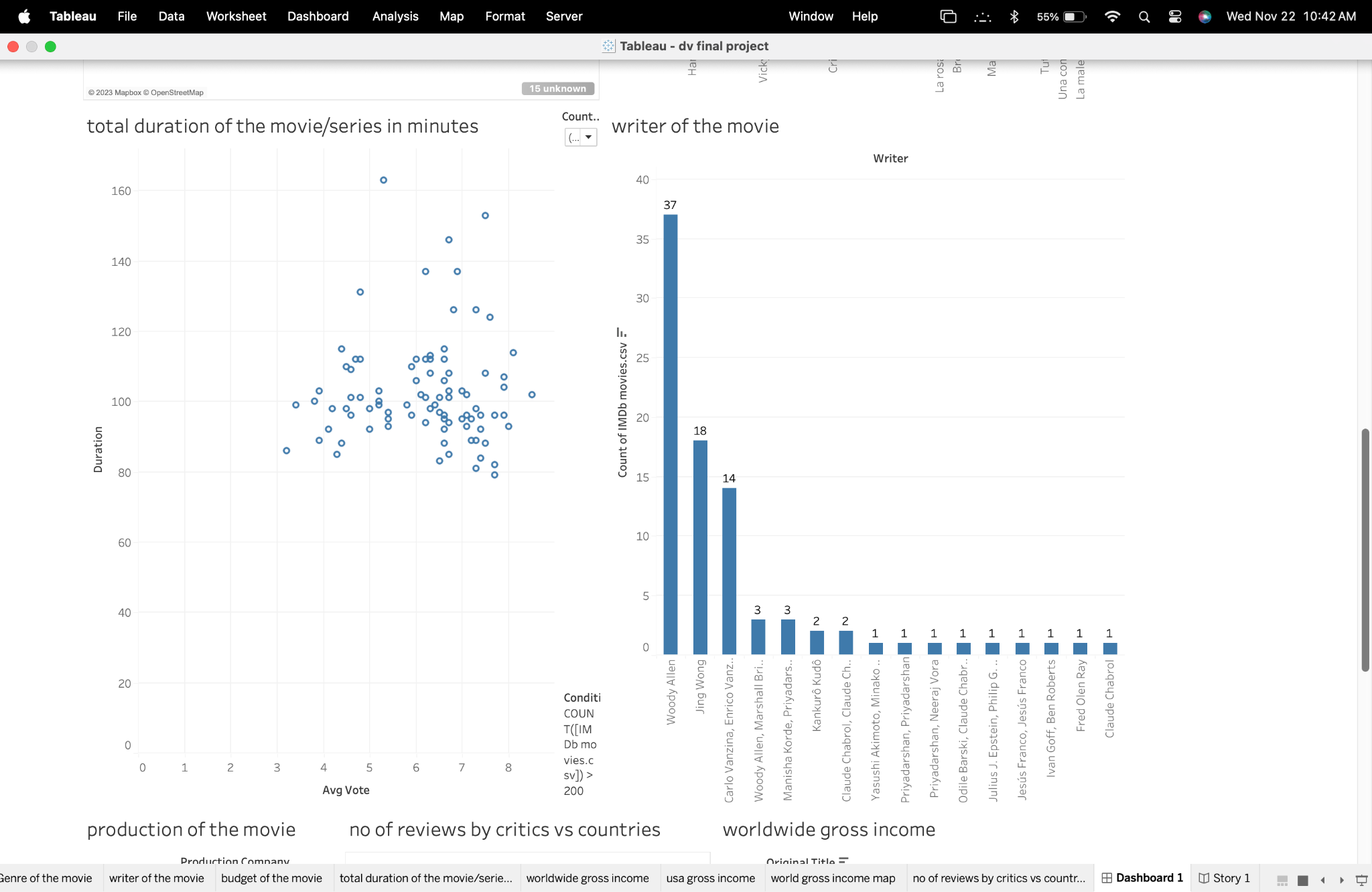
**IFT 533 Data Visualization & Reporting for IT**

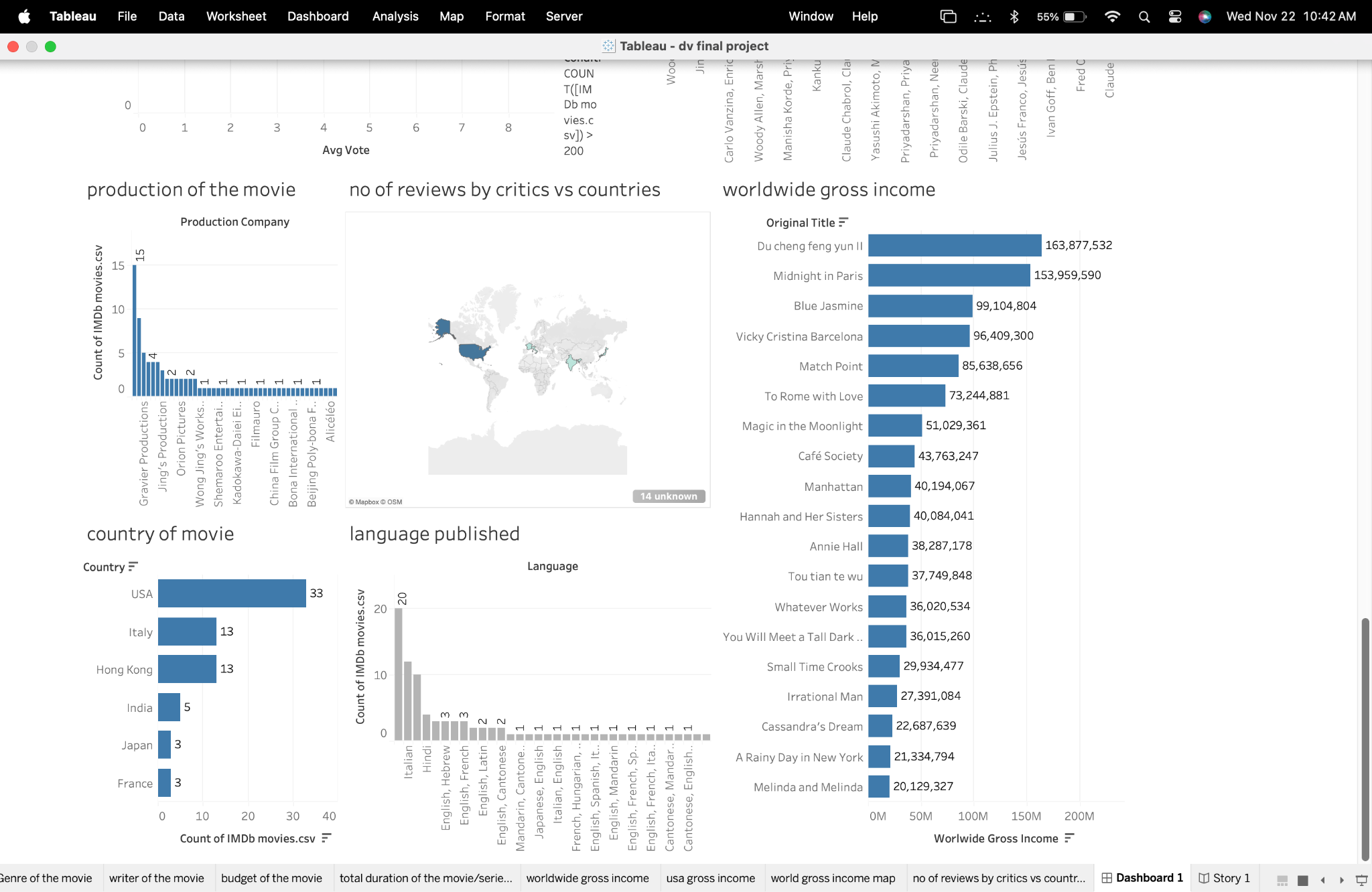
**Dr. Asmaa Elbadrawy**

**November 26th, 2023**

# **Section-1: Dashboard**

****

****

****

**A Brief explanation of Dashboard usage:**

The "Movie Dashboard" is a comprehensive tool designed to provide insights into the vast dataset of movies. It caters to users interested in analyzing and understanding trends, patterns, and key metrics within the film industry.

**Key Visualizations:**

Country-wise Movie Production (Map/Bar Chart): This visualization offers a geographical representation or bar chart to showcase the distribution of movies based on countries. It helps identify the top movie-producing countries, highlighting the global diversity of cinema contributions.

Annual Movie Production Trends (Line Chart): This line chart illustrates the evolution of annual movie production over the years. It captures significant trends and fluctuations in the film industry, offering valuable insights into its growth or changes.

Language Distribution (Bar Chart): The bar chart provides insights into the distribution of movies across different languages in the dataset. It sheds light on the number of films produced in each language, offering a glimpse into the linguistic diversity of cinematic storytelling.

Genre Distribution (Pie Chart/Bar Chart): This visualization presents the distribution of movie genres, offering information on the number of films within each genre. It provides insights into the diverse storytelling approaches prevalent in the dataset.

Top Movie Budgets (Bar Chart): This visualization showcases the top movies with the highest budgets. It allows users to quickly identify and analyze movies with significant financial investments.

Global Box Office Performance (Bar Chart): This chart displays the worldwide gross income of movies, enabling users to compare and rank films based on their global box office earnings.

Production Company Impact (Bar Chart): This visualization breaks down the number of movies produced by various production companies. It sheds light on the impact and influence of each company within the cinematic landscape.

**Dashboard Interactivity:**

Users can interact with the dashboard through filters, allowing them to dynamically explore the data based on specific criteria such as release year, genre, or language. The interactivity enhances the user experience, providing a customizable and insightful exploration of the movie dataset.

**Dashboard Strengths:**

The "Movie Dashboard" excels in offering a holistic view of the dataset, allowing users to explore diverse facets of the film industry. Its strength lies in the combination of informative visualizations and interactive features, enabling users to derive meaningful insights from the rich movie data.

# **Section 2: The Dataset**

The dataset comprises comprehensive information about a diverse array of movies, offering insights into the world of cinema. Each entry includes details such as the movie's title, original title, release year, publication date, genre, duration, production country, language, director, writer, production company, key actors, a concise movie description, budget, box office earnings in the USA and worldwide, metascore, and the number of reviews from both users and critics.

With a wide range of attributes, the dataset allows for in-depth exploration and analysis of trends in film production, financial performance, audience engagement, and critical reception. The rich and varied nature of the data makes it a valuable resource for understanding the dynamics of the film industry across different dimensions.

The IMDb dataset encompasses a comprehensive collection of information about movies, spanning various genres, languages, and production years. It includes details such as titles, original titles, release years, genres, durations, countries of origin, languages, directors, writers, production companies, cast members, movie descriptions, average viewer ratings, votes, budget information, box office earnings, and critical reviews.

This dataset offers valuable insights into the world of cinema, enabling analyses related to movie trends, viewer preferences, and industry performance.

**Datatype of the columns present in IMDB dataset:**

title: Categorical

original\_title: Categorical

year: Ordinal (assuming it represents the release year)

date\_published: Categorical

genre: Categorical

duration: Ratio

country: Categorical

language: Categorical

director: Categorical

writer: Categorical

production\_company: Categorical

actors: Categorical

movie description: Text (qualitative)

budget: Ratio

usa\_gross\_income: Ratio

worldwide\_gross\_income: Ratio

reviews\_from\_users: Ratio

reviews\_from\_critics: Ratio

Data Attributes:

| **Column Name** | **Description** |
| --- | --- |
| imdb\_title\_id | Unique identifier for each title on IMDb |
| title | The title of the movie in English |
| original\_title | The original title of the movie |
| year | The year of the movie's release |
| date\_published | The date when the movie was published |
| genre | The genre(s) of the movie |
| duration | The duration of the movie in minutes |
| country | The country or countries where the movie was produced |
| language | The language(s) of the movie |
| director | The director(s) of the movie |
| writer | The writer(s) of the movie |
| production\_company | The production company behind the movie |
| actors | The main actors in the movie |
| description | A brief description of the movie |
| avg\_vote | The average vote or rating the movie received |
| votes | The number of votes the movie received |
| budget | The budget of the movie |
| usa\_gross\_income | The gross income of the movie in the USA |
| worldwide\_gross\_income | The worldwide gross income of the movie |
| metascore | The metascore of the movie |
| reviews\_from\_users | The number of reviews from users |
| reviews\_from\_critics | The number of reviews from critics |

# **Section 3: Prospective Dashboard user’s**

| **User Type** | **Purpose** | **Data Needs** |
| --- | --- | --- |
| Film Industry Analysts | Conduct in-depth analyses of market trends, financial success, and audience preferences. Analysts can understand performance metrics across time, regions, and genres to predict future trends and provide strategic insights to stakeholders. | Financial data, trend analysis, genre popularity, historical performance. |
| Movie Producers and Studios | To understand strategic decision-making related to new productions, such as which genres or talents to invest in. | Genre trends, actor and director track records, and budget-to-revenue ratios. |
| Marketing and Distribution Teams | To determine effective ways to market and distribute movies. Getting insights into audience demographics, successful release strategies, and the impact of marketing campaigns to optimize reach and engagement. | Audience demographics, release timing, regional preferences. |
| Film Historians and Educators | To study the evolution of film as an art form and its cultural significance and interested in the historical context of filmmaking trends, the impact of societal changes on cinema, and educational material for academic purposes. | Historical data, genre evolution, cultural impact analysis. |
| Critics | Write informed articles and reviews. | Trending topics, critical reception, industry news. |
| Streaming Services Executives | To make content strategy, including acquisitions and original productions. They can understand viewer preferences, and content performance, and identify gaps in their offerings to stay competitive and satisfy subscriber demand. | Viewer preferences, competitive analysis, content gap analysis. |
| Data Scientists and Analysts | To build predictive models and analyze complex datasets to uncover hidden patterns, correlations, and insights that can help in decision-making processes. | Comprehensive metadata, ratings data, user reviews. |

# **Section 4: List of user requirements and potential questions**

Country-wise Movie Production Visualization:

* Users should be able to view a map or bar chart depicting the distribution of movies based on countries.
* The visualization should help identify the top movie-producing countries.

Genre-wise Movie Trends:

* Users require a chart showcasing trends in movie production over the years for specific genres.
* The chart should provide insights into the popularity of different genres.

Director's Impact on Ratings:

* Users want to explore the correlation between movie ratings and specific directors.
* A scatter plot or line chart should visualize how different directors influence the average rating of their movies.

Budget vs. Box Office Performance:

* Users need to understand the relationship between the budget allocated for a movie and its box office performance.
* A scatter plot or bar chart can illustrate this comparison.

Language Preferences and Global Reach:

* Users are interested in a visualization that displays language preferences and the global reach of movies.
* A world map or bar chart can represent the languages of movies and their distribution across countries.

Temporal Trends in Movie Production:

* Users want to explore how the number of movie releases has evolved over the years.
* A line chart or bar chart showing the temporal trends in movie production is required.

Comparison of Streaming Platform Libraries:

* Users are interested in a comparison of movie libraries on streaming platforms like Netflix and Amazon Prime.
* A bar chart or stacked bar chart can represent the number of movies available on each platform.

User Reviews and Critics' Ratings Analysis:

* Users need insights into how user reviews and critics' ratings align or differ.
* A scatter plot or line chart can visualize this comparison over different movies.

Exploration of Movie Descriptions:

* Users want a word cloud or textual analysis to explore common themes in movie descriptions.
* This should help identify recurring topics or trends in movie storytelling.

Budget Distribution and Outliers:

* Users need a visualization that highlights the distribution of movie budgets and identifies any significant outliers.
* Box plots or violin plots can effectively represent budget distribution.

**Potential Questions**

| **No.** | **Potential Questions** |
| --- | --- |
| 1 | How has the average length (duration) of movies changed over the years? |
| 2 | Which genres are most prevalent in the dataset and how has their popularity changed over time? |
| 3 | What is the distribution of movies across different countries of production? |
| 4 | Can we identify a correlation between a movie’s budget and its average vote (rating)? |
| 5 | What trends can be observed in the volume of movie production per year? |
| 6 | How do user and critic review counts compare for movies across different average vote brackets? |
| 7 | Are there any notable differences in the average votes of movies when looking at language or country-specific productions? |
| 8 | What is the relationship between a movie’s average vote and its commercial success (worldwide gross income)? |
| 9 | How do movies with high meta scores compare with their IMDb average votes? |
| 10 | Which directors have the highest average votes for their movies? |
| 11 | What are the most common themes in movie descriptions, and do these themes correlate with higher ratings? |
| 12 | Can we predict the metascore of a movie based on other available data points, like average vote, review counts, or genre? |
| 13 | Which production companies are associated with the highest-grossing movies? |
| 14 | What factors contribute to a movie achieving a rating of 9 or higher on IMDb? |
| 15 | How frequently do movies with high budgets result in high worldwide gross income? |
| 16 | Do movies with a higher number of user reviews tend to have a higher average vote? |
| 17 | Are movies based on true stories or historical events rated higher than fictional stories? |
| 18 | What is the relationship between the duration of a movie and its success in terms of average vote and gross income? |
| 19 | How does the presence of award-winning actors or directors influence a movie's rating and financial performance? |
| 20 | Which genres tend to have the highest budget and do they correspond to higher worldwide gross income? |

1. Which director has the highest number of movies?
2. (a)What countries are prominently featured in the movies you've analyzed, and could you provide a breakdown of the movie count for each country, shedding light on the diversity of global cinema contributions?

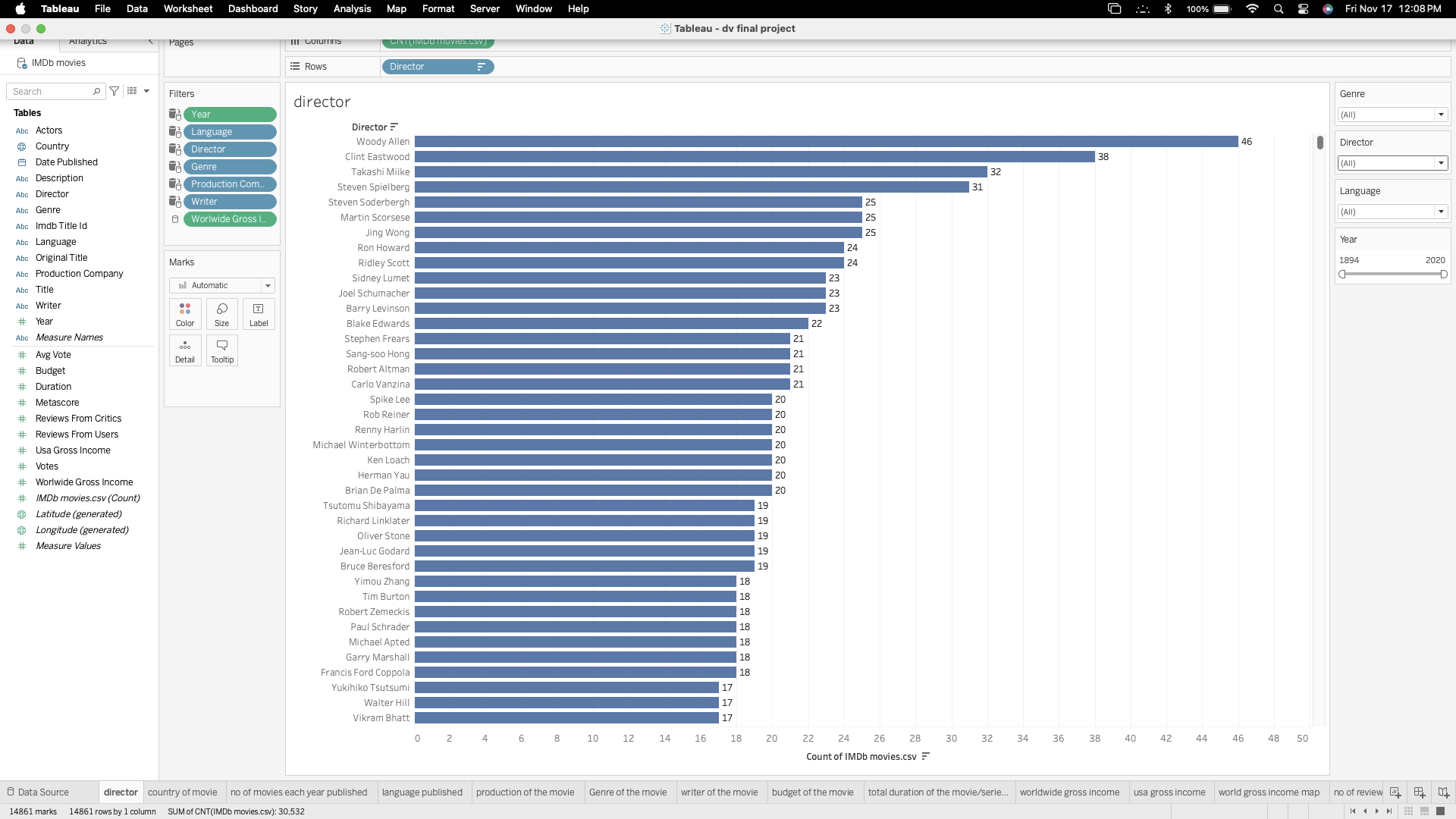
(b)Which country has the highest number of movies from Genre “action, scifi, adventure, comedy, crime, fantasy, history, drama”?

1. How has the annual production of movies evolved over the years, highlighting any significant trends or fluctuations in the film industry?
2. What is the distribution of movies across different languages in the dataset, and can you provide insights into the number of films produced in each language?
3. What is the distribution of movie genres in the dataset, and could you provide information on the number of films within each genre, offering insights into the diversity of cinematic storytelling?
4. Which production companies have been prolific in the film industry, and can you provide a breakdown of the number of movies produced by each company, shedding light on their impact and influence within the cinematic landscape?
5. Can you list the number of reviews by critics vs countries?
6. What were the top 20 movies with the highest budget?
7. How long is the duration of the movie/series in minutes, and does the runtime contribute to the overall storytelling and viewer experience?
8. What is the worldwide gross income of and how does it rank among other films in terms of global box office earnings?

# **Section 5: The plots**

# Plot 1:

Question 1: Which director has the highest number of movies?



From the above visualization we can see that the director “Woody Aaron '' has the highest number of movies(46) from the year 1894 to 2020 across all the languages and all the Genres in this IMDB movie dataset.

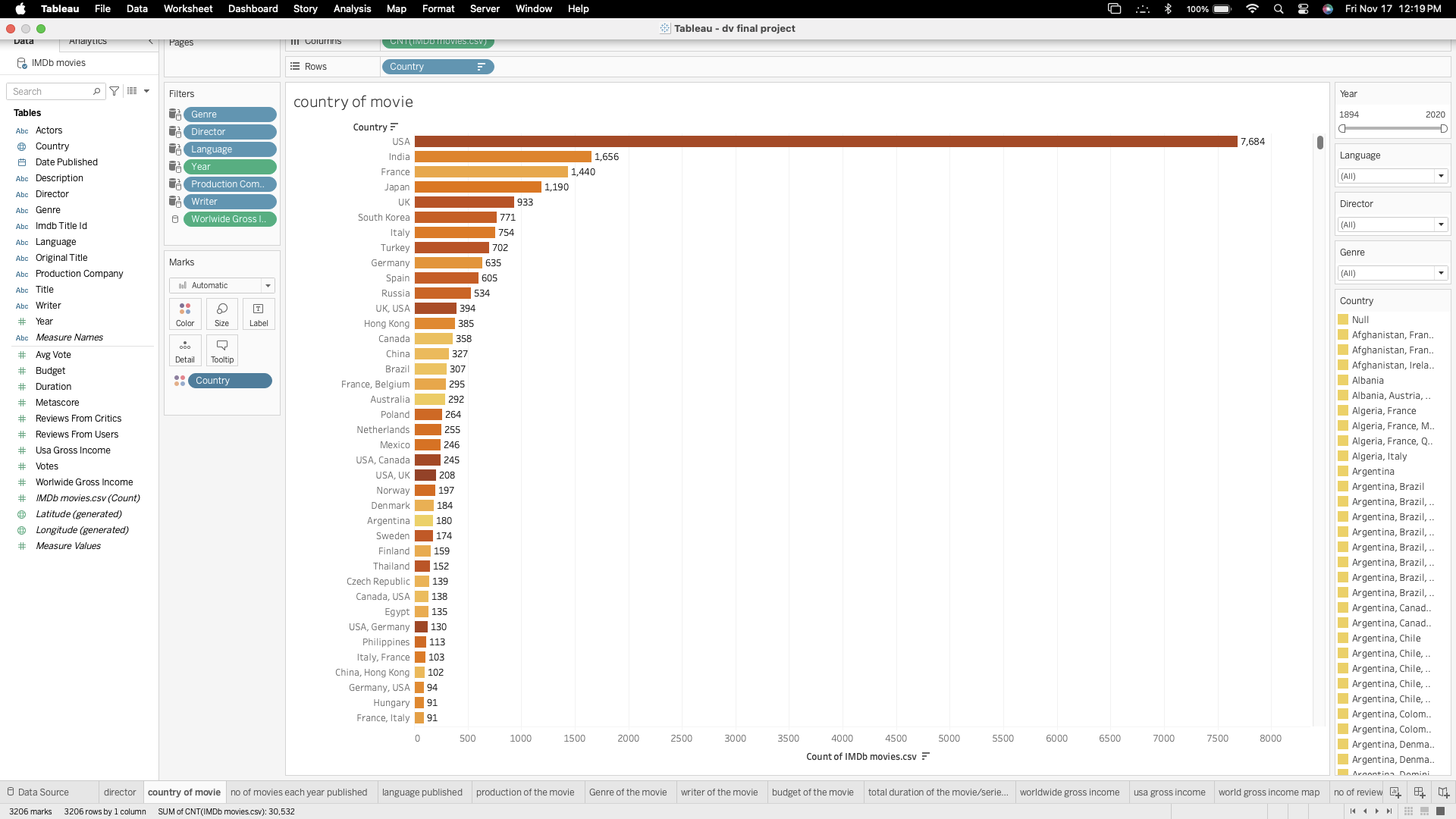
The provided bar graph effectively shows the directors with their movie count from 1894 to 2020. Each bar in this graph represents a director and their movie count. Director Clint Eastwood has the second highest count of movies under his belt with 38 movies and so on.

**Pre-attentive attributes**:

* Length: the length of the bar indicates the number of movies a director has directed movies. Longer the bar indicates the more movies for that particular director.

# **Plot 2:**

Question 2(a): What countries are prominently featured in the movies you've analyzed, and could you provide a breakdown of the movie count for each country, shedding light on the diversity of global cinema contributions?



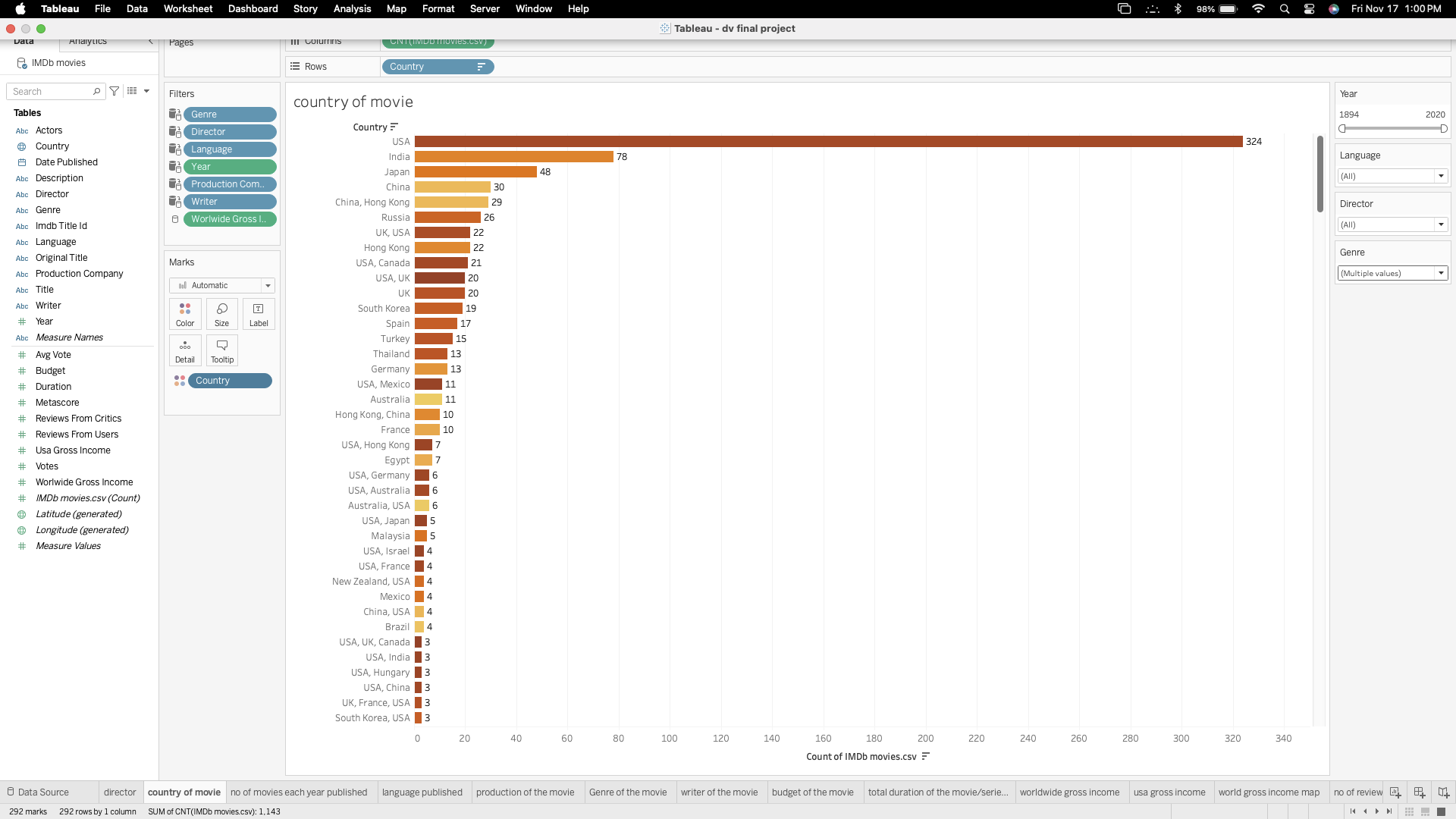
From the above visualization, the most prominent country with the highest number of movies released in “USA” with 7500+ movies from the year 1894 to 2020. India has second highest movies with 1600+ movies.

The dataset encompasses a wide array of countries, each contributing a varying number of movies. This breakdown not only showcases the dominance of certain film-producing nations but also emphasizes the rich and diverse landscape of global cinema. The movie count for each country serves as an insightful metric, reflecting the global reach and influence of the film industry across different regions.

**Pre-attentive attributes:**

* Length: The length of the bar indicates the number of movies released in a particular country. So, the larger the length of the bar indicates the higher number of movies from that country.
* Color: The sequential coloring in this visualization indicates that degree of prevalence for each country and its movies count.

**Question 2(b):** which country has the highest number of movies from Genre “action, scifi, adventure, comedy, crime, fantasy, history, drama”?



From the above visualization, the most prominent country with the highest number of movies under the specified Genre was released in “USA” with 324 movies from the year 1894 to 2020. India has the second highest movies with 78 movies from Genre “action, scifi, adventure, comedy, crime, fantasy, history, drama”.

# **Plot 3:**

**Question 3:** How has the annual production of movies evolved over the years, highlighting any significant trends or fluctuations in the film industry?



We can see that the annual production of movies across the years from 1875 to 2020 see a significant rise with each year from 1979 to 2020 being the most number of movies produced. And in the 2018 year 1691 movies have been produced which is the highest in a single year across the dataset.

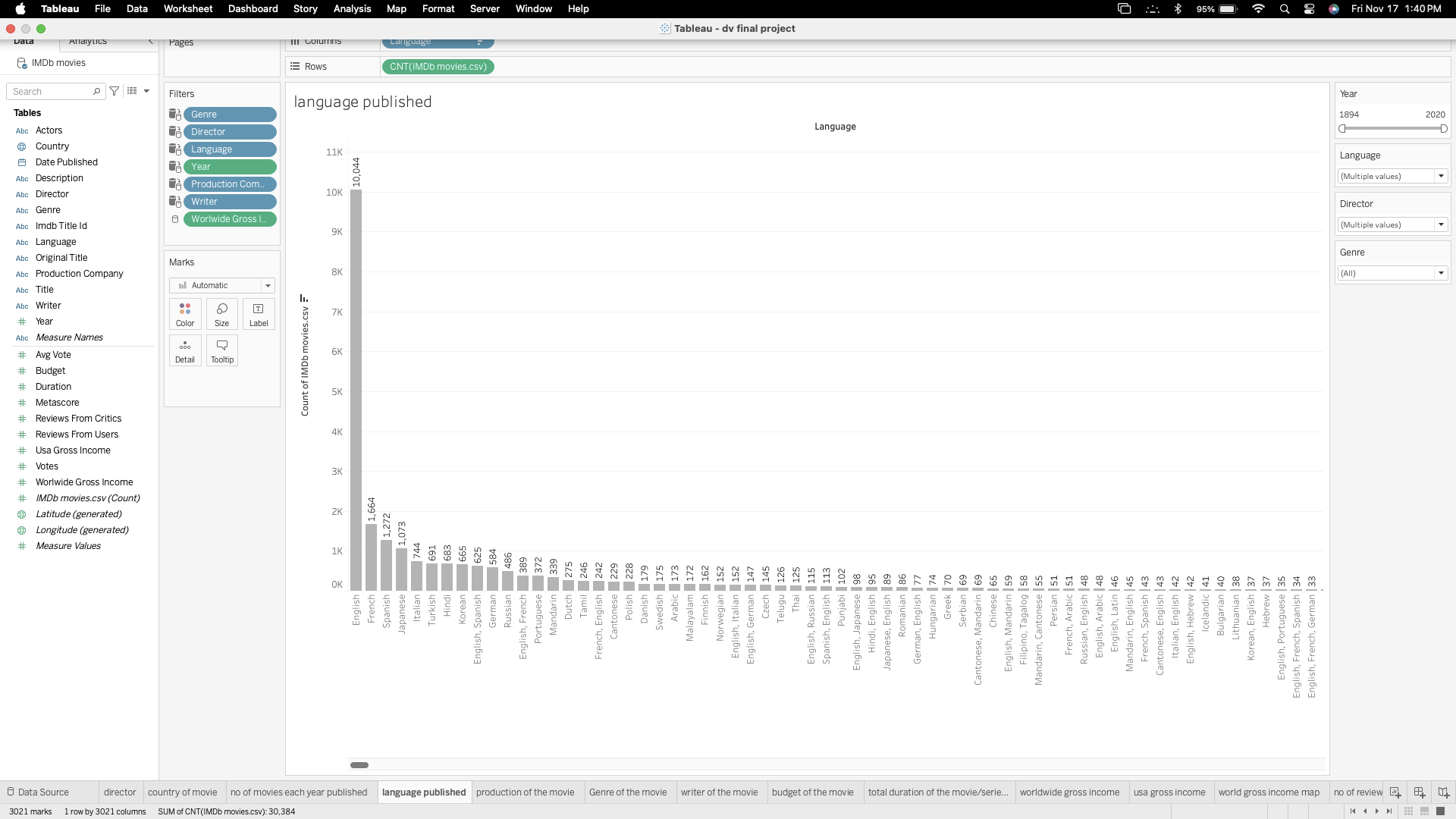
So from the visualization we can conclude that people started admiring the movies from the year 1978 where we can see the significant rise in the bar chart from 1978 to 2020.

**Pre-attentive attributes used**:

* Length: The length of the bar indicates the number of movies released in a particular year. So, the larger the length of the bar indicates the higher number of movies from that country.

# **Plot 4:**

Question 4: What is the distribution of movies across different languages in the dataset, and can you provide insights into the number of films produced in each language?



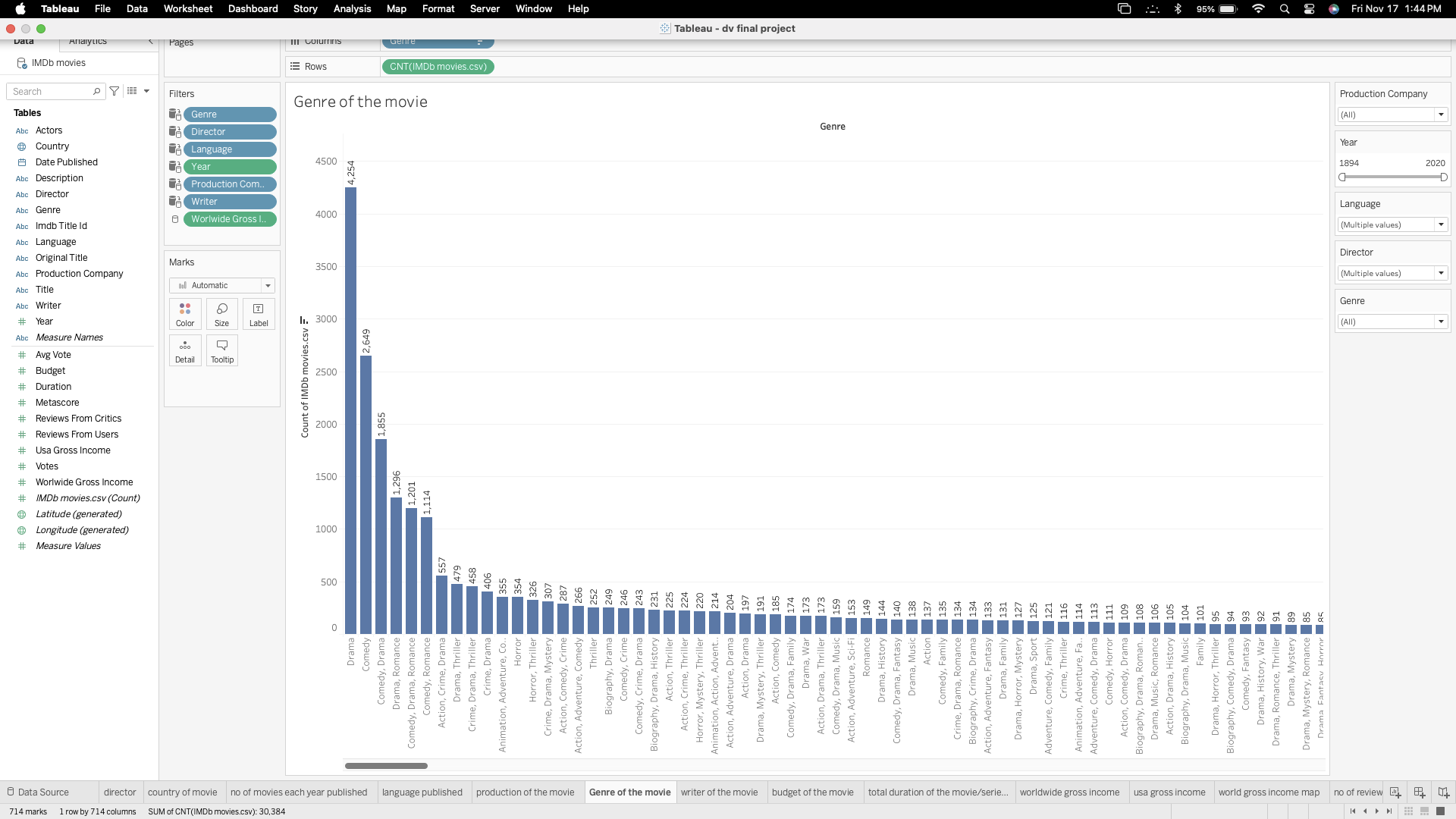
* The visualization used is a bar chart, which is effective for comparing the frequency of categorical data—in this case, the number of films produced in each language.This bar chart allows viewers to quickly assess the distribution of movies across languages, with English prominently having the most number of movies, followed by other languages like French, Spanish, and Japanese. The chart highlights the dominance of English in the dataset, as well as the significant representation of European and Asian languages.It should be noted that there are languages with combined entries, such as "English, Spanish" and "English, French," indicating movies that have multiple languages. Additionally, there is a category for "Unknown," which represents movies where the language data is missing or not available.
* The visualization simplifies complex data and allows for immediate insights into the production of films across languages, illustrating the linguistic diversity present in the global film industry as captured by the IMDb dataset.

**Pre-attentive Attributes used:**

* Length: The varying lengths of the bars allow quick comparison among the categories.
* Position along a common scale: The aligned start of the bars along the y-axis helps in judging the length more accurately.
* Color: If there is color coding, it can draw attention to specific bars or help in grouping languages.

# **Plot 5:**

Question 5: What is the distribution of movie genres in the dataset, and could you provide information on the number of films within each genre, offering insights into the diversity of cinematic storytelling?



This kind of visualization is particularly effective for showing the frequency of categorical data, such as the count of films within each genre. Here’s a breakdown of the pre-attentive attributes and insights that can be derived from such a visualization:

* The bar chart allows viewers to quickly discern which genres are most prevalent in the dataset, with genres like Drama, Comedy, and Thriller likely appearing at the top due to their popularity in film production.
* The visualization also highlights the diversity of cinematic storytelling by showing a broad array of genres, from mainstream to more niche ones.
* This genre distribution can provide insights into industry trends, audience preferences, and the historical evolution of film genres over time.
* One can also observe the long tail in the distribution, where a large number of genres have a relatively small number of films, indicating the specialized nature of certain genres.

By examining such a visualization, stakeholders can make informed decisions about content creation, marketing strategies, and understanding audience interests. It encapsulates complex data into an easily digestible format, providing a comprehensive overview of the genre landscape within the dataset.

**Pre-Attentive Attributes used:**

1. Length: The lengths of the horizontal bars enable quick comparisons of the quantities they represent, in this case, the number of films within each genre. This is a fundamental pre-attentive attribute that allows for immediate visual assessment.

2. Width: All bars have a uniform width, which standardizes the comparison across different genres and allows the length alone to convey the information.

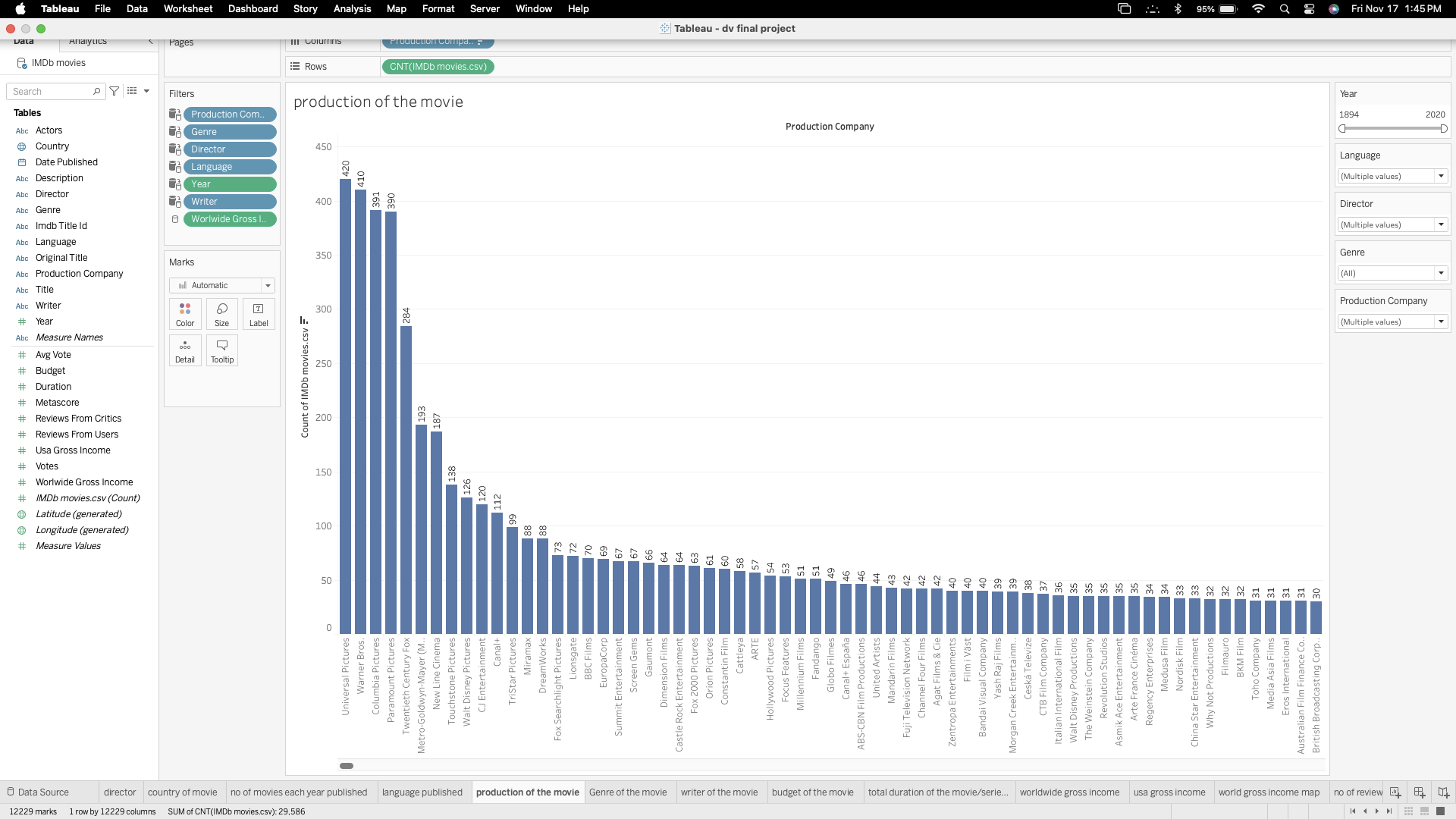
3. Position: The positioning of the bars along the x-axis correlates to the genre and along the y-axis to the number of films. By arranging the bars in descending order, the chart immediately communicates the genres with the most and least number of films.

4. Color: A color gradient may be applied to the bars to represent different values, but in this visualization, a single color appears to be used for all bars, keeping the focus on the length.

5. Text Labels: The genres are clearly labeled along the x-axis, and the exact counts are likely indicated at the end of each bar or on the y-axis. This textual information complements the pre-attentive attribute of length by providing precise data.

# **Plot 6:**

Question 6: Which production companies have been prolific in the film industry, and can you provide a breakdown of the number of movies produced by each company, shedding light on their impact and influence within the cinematic landscape?



This visualization helps to identify which production companies are the most prolific in the film industry. Let's discuss the visualization and its pre-attentive attributes:

* The chart quickly communicates which production companies have produced the most films, indicating their impact and influence in the cinematic landscape.
* The leading companies can be seen to have a significant number of productions, potentially reflecting their market share and the resources at their disposal.
* The visualization also shows the long tail of the production landscape, with many companies producing a smaller number of films, which can reflect the presence of independent and smaller studios in the industry.
* This distribution can reveal industry dynamics, such as consolidation of production power or diversity in film production sources.

In summary, this bar chart condenses complex production data into a simple format, allowing viewers to assess the prolificacy of production companies in the film industry. It can be used by analysts, filmmakers, and industry professionals to understand market dynamics and the competitive environment within the cinematic landscape.

**Pre-Attentive Attributes used:**

1. Length: The lengths of the bars represent the number of movies produced by each production company. This visual attribute allows for an immediate comparison of output between companies.

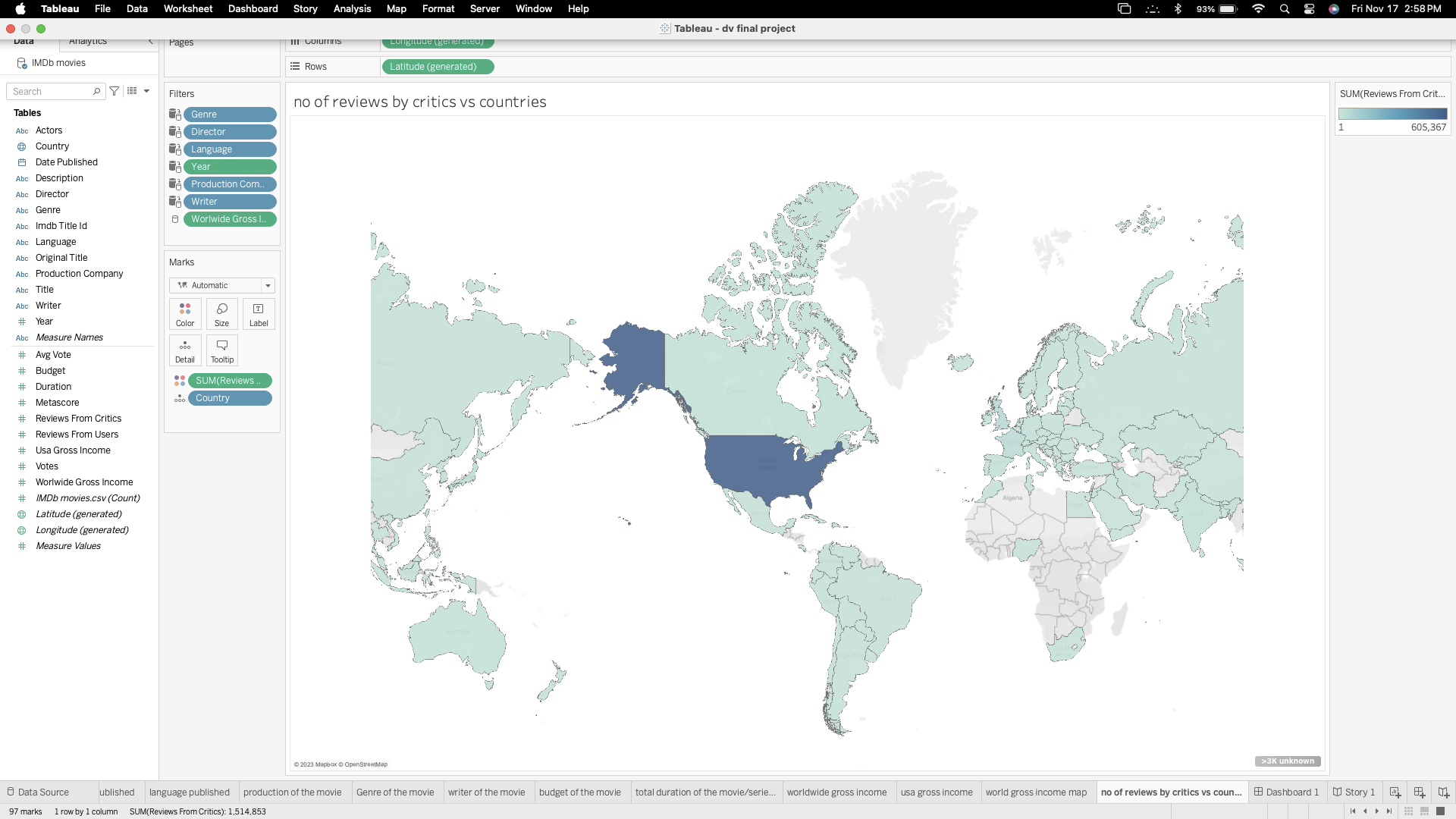
2. Positioning: The vertical positioning of the bars corresponds to the production company, with the chart likely sorted in descending order to highlight the most productive companies at the top.

3. Color: A single color across all bars focuses the viewer's attention on the length of the bars rather than differentiating between the companies based on color.

4. Alignment: The bars are aligned along a common baseline, which aids in comparing their lengths effectively.

# **Plot 7:**

Question 7: Can you list the number of reviews by critics vs countries?



Choropleth maps are effective for displaying how a measurement varies across a geographic area. This map allows viewers to see which countries have a higher number of critic reviews, potentially correlating to a larger film industry or more active critical engagement. Countries with darker shades will stand out, indicating a significant concentration of critic reviews, which could suggest a robust cinematic culture or a larger audience for movies. The map also shows global participation in film criticism, highlighting areas where the film industry may be subject to more critical analysis.

Such a visualization can be instrumental for film distributors or producers looking to understand the critical reception in various markets and to strategize their releases and promotional activities accordingly. This visualization type is particularly useful for presenting data related to physical locations, offering an immediate sense of geographic distribution that can be crucial for decision-making processes in industries like film, where the market is inherently global.

**Pre-Attentive Attributes used:**

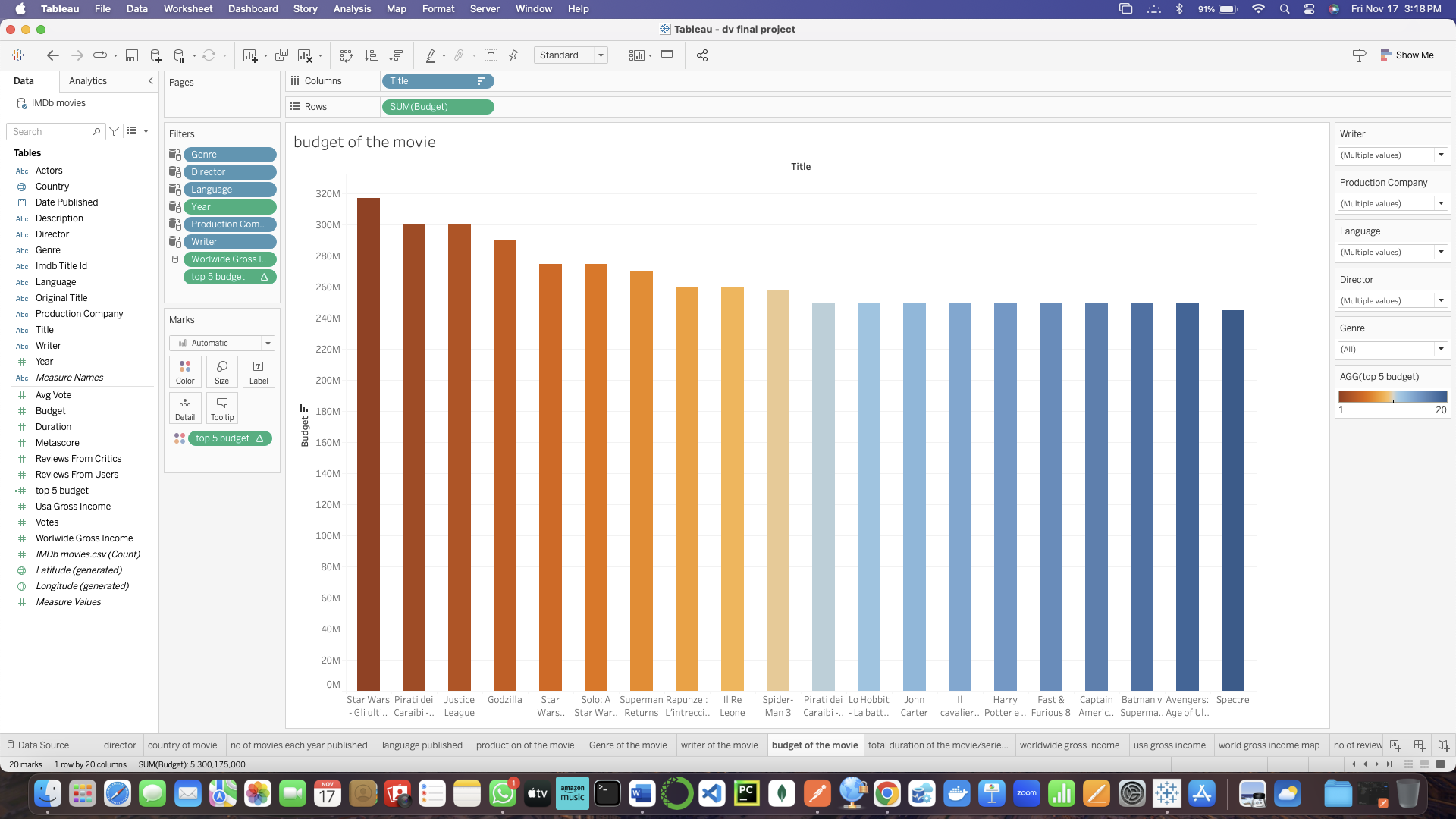
1. Color Hue and Intensity: Different shades of color indicate the density or quantity of the data point, which in this case is the number of critic reviews. Darker shades typically represent higher values, allowing viewers to quickly identify areas with more critic reviews.

2. Spatial Position: The geographic placement of the countries on the map corresponds to their real-world location, giving context to the data.

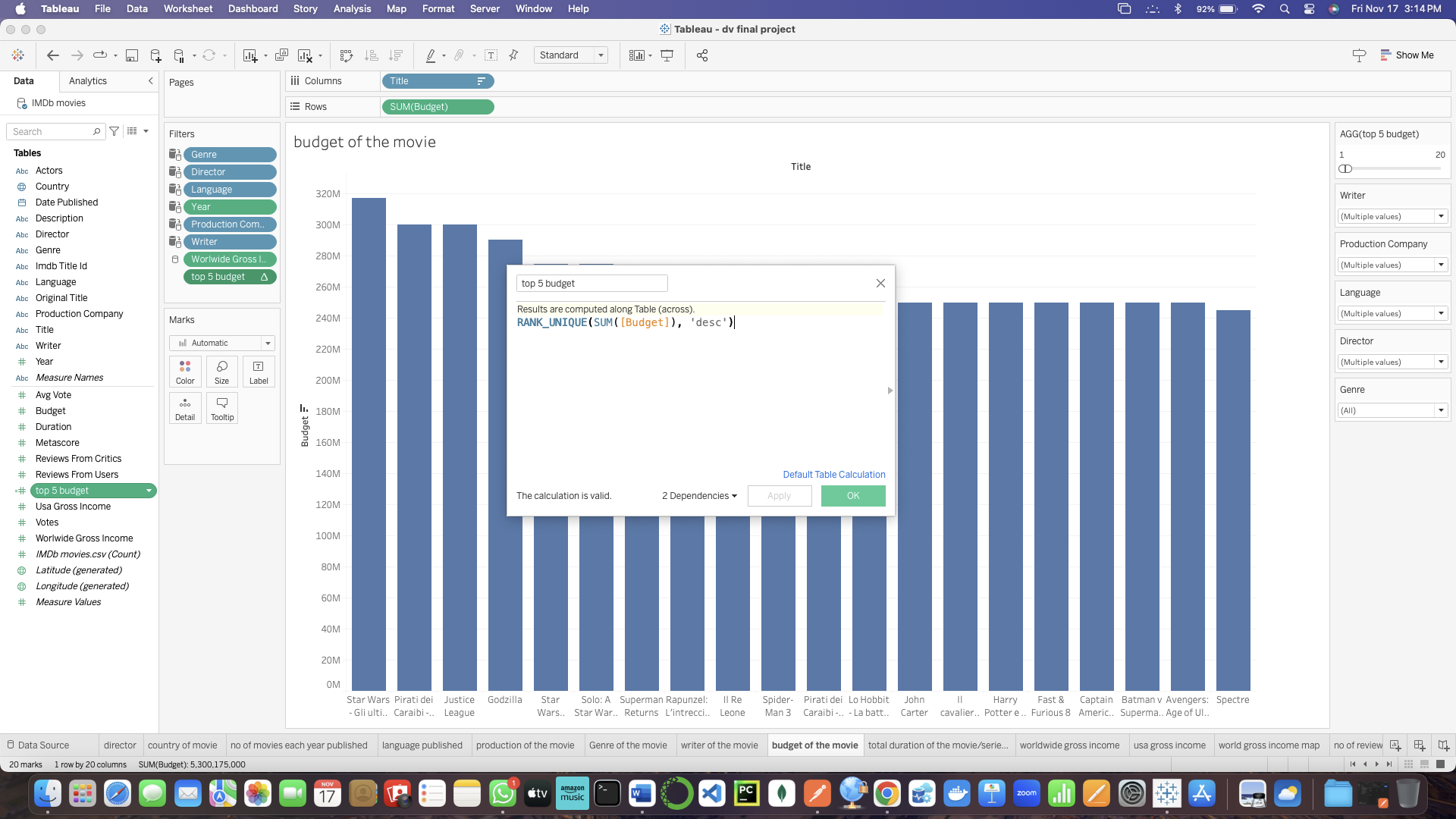
3. Size: Although not a direct size comparison as with bar or bubble charts, the size of each country on the map provides a subconscious cue about the possible market size and influence.

# **Plot 8:**

Question 8: What were the top 20 movies with the highest budget?



**Calculated field:**



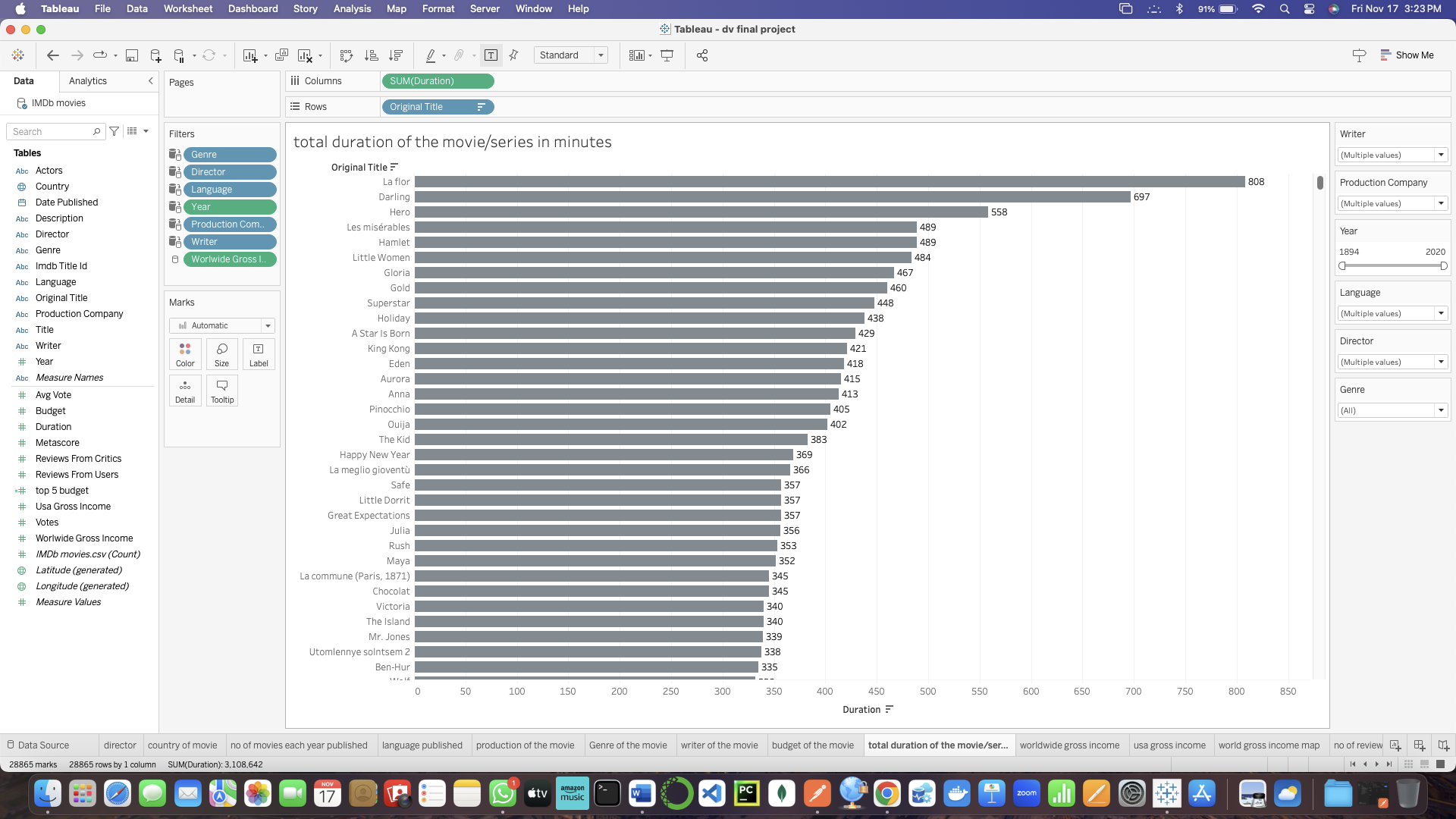
We have created a calculated field which will rank the movies based on the budget as shown above SQL code. Using that calculated field we have sorted the top 20 movies with highest budget which resulted in the above visualization in which the movie “star wars” is the movie with highest budget of around $320M.

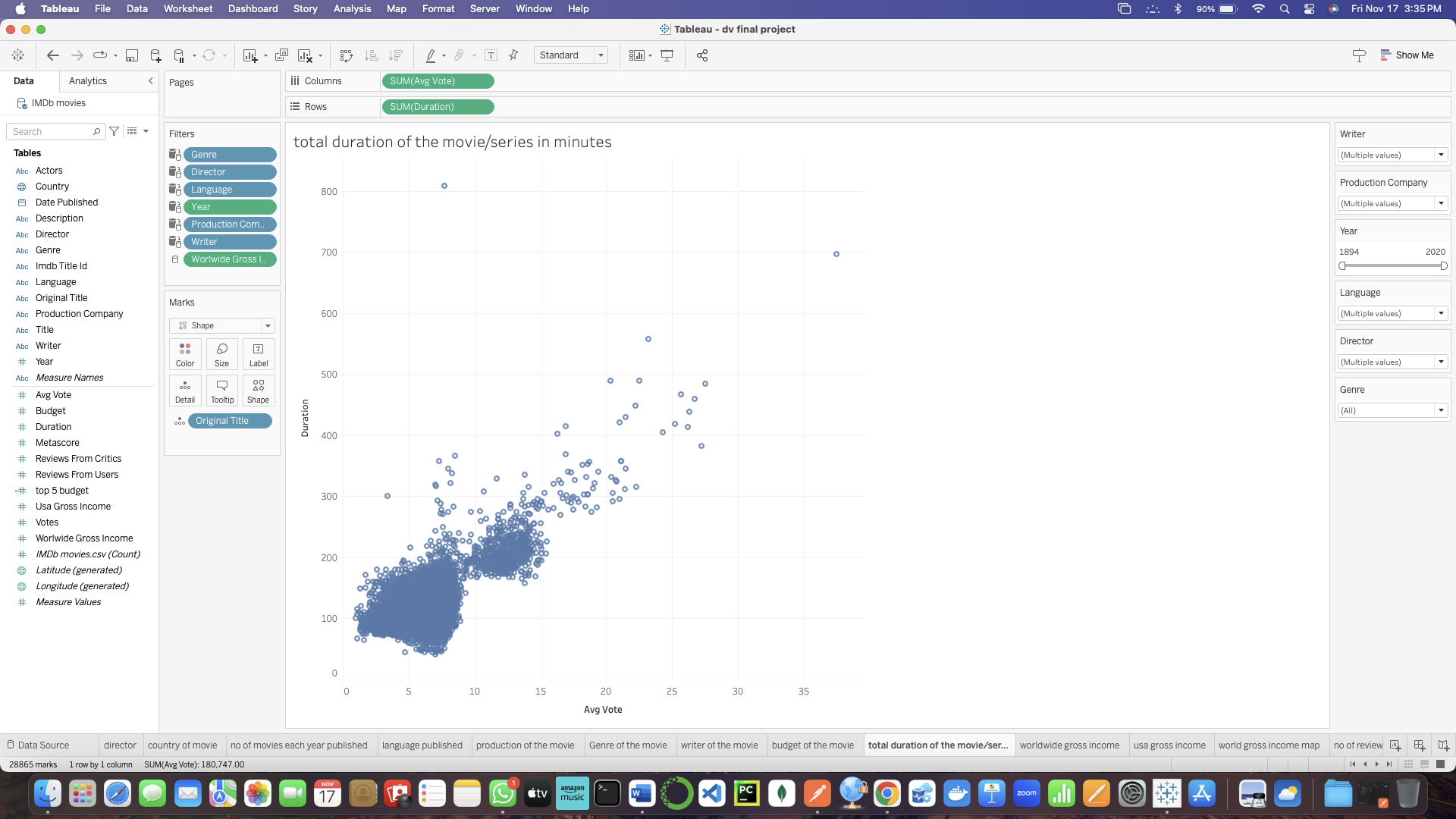
**Pre-attentive attributes:**

* Length: The length of the bar indicates the budget of the movie. So, the larger the length of the bar indicates the higher the budget of the movie
* Color: We have used sequential coloring for the above visualization which indicates the Orange with high budget and blue with less budget compared to orange.

# **Plot 9:**

Question 9: How long is the duration of the movie/series in minutes, and does the runtime contribute to the overall storytelling and viewer experience?





So from the above visualization we can see that the movie/series in the range of 100- 200 minutes has a higher number of average votes compared to the movies with more duration.

Based on the analysis of the IMDb movies dataset, we find the following about the duration of movies/series:

* The average (mean) duration of a movie/series is approximately 100 minutes.
* The median duration is 96 minutes, indicating that half of the movies/series are shorter than this, and half are longer.
* There is a wide range of durations, with a minimum of 41 minutes and a maximum of 808 minutes.
* The majority of movies/series (25th to 75th percentile) range between 88 and 108 minutes in duration.

The histogram visualization, which we can relate to the first image conceptually, uses the following pre-attentive attributes:

1. Height of Bars: Indicates the frequency of movies/series within specific duration intervals.

2. Width of Bars: Each bar has a uniform width, representing a fixed range of durations.

3. Color Intensity: If applied, could indicate the frequency of movies within the intervals, but here we use a single color for simplicity.

From the first image description, we can infer that a bar chart was used to display the total duration of movies/series in minutes across different categories. The pre-attentive attributes would be similar to those listed above for the histogram, with length of bars being the primary visual cue for duration.

For the second image description, which may be conceptualized as a scatter plot, we can speculate on pre-attentive attributes like:

Regarding the impact of duration on storytelling and viewer experience, the runtime of a movie/series is a critical factor. It influences the depth of the narrative, character development, and pacing of the story. While longer movies have the capacity to explore complex plots and character arcs in depth, they also risk losing the audience's attention if not well-edited. Shorter films need to convey their story more concisely, which can be challenging but also keeps the narrative tight and potentially more engaging. The optimal duration often depends on the genre, target audience, and the story being told.

**Pre-attentive attributes used:**

1. Position on Axis: Each point's position along the x-axis and y-axis represents two variables, such as the number of reviews and duration.

2. Point Size: If used, could represent a third variable, such as budget or box office return.

3. Color: May categorize points by genre, rating, or another categorical variable.

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

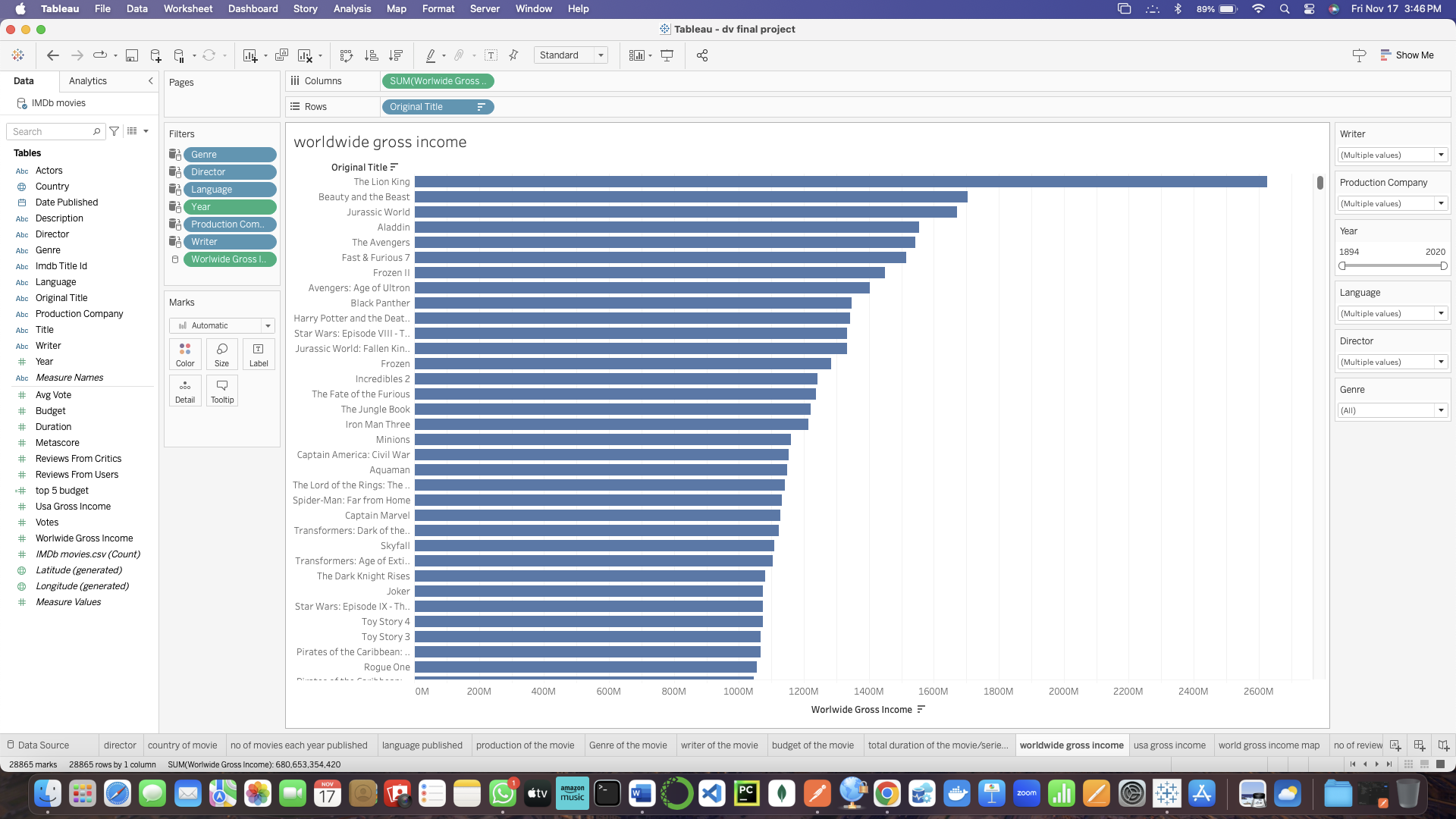
# 

# 

# 

# **Plot 10:**

Question 10: What is the worldwide gross income of and how does it rank among other films in terms of global box office earnings?

 This type of visualization is adept at providing a clear ranking of films based on their global box office earnings. Here's how we can interpret the visualization based on the description:

* The chart clearly shows which films have the highest worldwide gross which is The lion king, highlighting the most commercially successful films.
* Films with longer bars dominate in terms of box office earnings, and their position at the top indicates their rank among other films.
* The differences in bar length also give an idea of how close the competition is between top-grossing films.

**Pre-Attentive Attributes used:**

1. Length: The length of each horizontal bar represents the gross income of a film. This allows for a quick comparison of financial performance across films.

2. Ordering: The bars are typically sorted in descending order, which places the highest-grossing films at the top of the chart, making it easy to see the rank order.

3. Color: While color here may not serve a differentiating purpose since the metric is consistent across all bars (worldwide gross income), it does help to distinguish individual bars from one another.

4. Alignment: All the bars align at a common baseline (the y-axis), which is essential for comparing their lengths accurately.

# 

# **Section 6: Interactivity**

# **Plot-1:**

**List of what they will be used for:**

**Directors:** The relationship between the movies and the directors will be visualized using the year, language, genre, production company, writer, and worldwide gross income as a filter.

**Movies:** The relationship between the directors and the movies will be visualized based on the number of movies that are being published each year with the highest number of movies in different countries.

**Visualization plot:** The relationship between the directors and the movies with respect to the number of movies in different countries will be represented by the visualization plot, which could be a bar chart, line chart, scatter plot, or any other type of suitable plot.

**Plots connected to each one:**

Visualization Plot: The movies, and directors will be linked to the visualization plot. A graphical representation of the association between the movies and directors will be the factors for the number of movies that are being produced each year and also the director with the highest number of movies, as well as the pertinent data attributes from the IMDB movies dataset.

**Value range and data attribute:**

The chart clearly shows that the USA made a lot more movies than any other country, with over 7500 films from 1894 to 2020. This proves how much the USA dominates the movie world. India comes next with more than 1600 movies, showing it's also a big player in global movies.

This dataset includes movies from many different countries. The big differences in the number of movies made by each country show how strong some places are in making films. It also shows how movies from around the world tell different stories about different cultures. This breakdown helps us see how much movies affect the whole world and how they're made in many different places.

# **Plot-2:**

**List of what they will be used for**: This visualization aims to showcase the distribution of movies across different countries, emphasizing the diversity of global cinema contributions.

**Visualization Plot**: A horizontal bar chart showing the count of movies for each country.

**Plots Connected to Each One**: Linked bar chart showing the count of movies for a specific country when clicked or selected in the main visualization.

**Value Range and Data Attribute**:

The value range will be the count of movies, and the data attribute will be the countries.

**Dashboard Interactivity Implementation**:

Interactive filters or parameters can be added to allow users to specify a range of years or genres, influencing the movie count.

* **Filter by Year or Genre**: Adding interactive filters or parameters to enable users to choose specific years or genres, dynamically updating the movie count visualization based on their selections.

# **Plot 3:**

**List of what they will be used for**: This visualization aims to illustrate the trends and fluctuations in the film industry's annual movie production over the years.

Visualization Plot:

A bar chart showing the annual count of movie productions over the years.

**Plots Connected to Each One**: A bar chart showcasing the annual count of movies for individual countries, allowing users to compare global trends with specific countries.

**Value Range and Data Attribute**: The value range will be the count of movies, and the data attribute will be the years.

Interactive filters or parameters can be added to enable users to focus on specific genres or countries.

* **Time-Based Filtering**: Added interactive sliders or date range selectors to allow users to focus on specific periods, making it easier to identify trends.
* **Genre-Specific Trends**: Implemented interactivity to enable users to click on a specific genre in the main visualization, updating connected plots to show the genre's contribution to the overall trend over the years.

# **Plot 4:**

**List of what they will be used for**: This visualization aims to present no. of movies published in different languages over the years.

**Visualization Plot:**  
A bar chart showing the number of films produced in each language.

**Value Range and Data Attribute:**

Each bar represents a different language or a combination of languages, depending on whether the film is multilingual or not, with the x-axis representing the main categorical variable.

Dynamic Filtering: Added a dropdown menu with a search box that allows users to select specific languages to display in the bar chart.

**Dashboard Interactivity Implementation**:

**Time-Based Filtering:**

* **Period Selection:** Provide users with the ability to select a specific time period for analysis in order to reveal how the production in different languages has changed over time.

**Genre-Specific Trends:**

* **Clickable Genres:** Users should be able to select a genre in a legend or filter panel to see how films in each language contribute to genre-specific trends.

# **Plot 5:**

**List of what they will be used for:**

This visualization aims to showcase the distribution of movie genres, offering insights into the number of films within each genre and highlighting the diversity of cinematic storytelling.

**Visualization Plot:**

A bar chart displaying the count of movies for each genre.

Plots Connected to Each One:

**Value Range and Data Attribute**:

The value range will be the count of movies, and the data attribute will be the different genres.

**Dashboard Interactivity Implementation**:

* **Genre Selection**: Add interactivity to allow users to click on a specific genre in the main chart, updating connected plots to show details about the selected genre.
* **Filtering Options**: Include filters or parameters to enable users to focus on specific periods or countries within a selected genre.

# 

# 

# **Plot 6:**

**List of what they will be used for**:

This visualization aims to showcase prolific production companies, providing a breakdown of the number of movies produced by each company. It sheds light on the impact and influence of production companies within the cinematic landscape.

Visualization Plot:

A bar chart or horizontal bar chart displaying the count of movies for each production company.

**Plots Connected to Each One**:

A detailed summary or card view displaying additional information about a selected production company when clicked, including average ratings, top movies, and critical reviews.

**Value Range and Data Attribute**:

The value range will be the count of movies, and the data attribute will be the different production companies.

**Dashboard Interactivity Implementation**:

* **Company Selection**: Added interactivity to allow users to click on a specific production company in the main chart, updating connected plots to show detailed information about the selected company.
* **Time-Based Analysis**: Implement a time slider or range selector to enable users to explore how the production landscape has evolved over time. Users can observe trends, spot emerging companies, and analyze shifts in production patterns.
* **Genre Filter**: Introduce a genre filter that users can apply to see the production distribution of companies within a specific genre. This helps in understanding the specialization of production companies in different cinematic genres.

# **Plot 7:**

**List of what they will be used for**:

This visualization aims to show the writer name of the movie, providing a breakdown of the number of movies produced by each writer. This visualization illuminates the significance and sway of writer names within the cinematic realm, providing insight into their influence and impact on the film industry.

**Visualization Plot:**

A bar chart or horizontal bar chart displaying the count of movies for each writer.

**Plots Connected to Each One**:

Include a dropdown menu or search functionality that allows users to select or search for specific writers. Upon selection, the bar chart updates dynamically to showcase the count of movies attributed to the chosen writer.

**Value Range and Data Attribute**:

The value range will be the count of movies, and the data attribute will be the different writers for each movie.

**Dashboard Interactivity Implementation**:The Dashboard Interactivity Implementation for the visualization depicting the count of movies produced by each writer could encompass several interactive functionalities to enhance user engagement and exploration:

* **Interactive Bar Chart:** Display a bar chart (or horizontal bar chart) showing the count of movies for each writer. Enable interactivity such as hover tooltips to display specific counts when users hover over individual bars.
* **Click Interactivity:** Enable users to click on a writer's bar within the chart, triggering additional details or a pop-up window with more comprehensive information about the movies associated with that writer.
* **Filtering Options:** Incorporate filtering options based on movie genres, release years, or other relevant criteria. Users can filter the displayed data to explore the counts of movies for a writer within specific genres or timeframes.

# **Plot 8:**

**List of what they will be used for**:

This visualization aims to illustrate the relationship or distribution of movie budgets across different movie titles. This visualization allows for the comparison of budgets among various movies, showcasing how budgetary allocations vary across titles in the dataset.

Visualization Plot:

A line chart displaying the budget of movies for each title of the movie.

**Plots Connected to Each One**:

The detailed summary of elements within the line chart collectively facilitates the visual representation and comparison of movie budgets across different movie titles, allowing viewers to discern and analyze the variations in budget allocations among the titles presented in the dataset.

**Value Range and Data Attribute**:

The value range will be the budget of the movie, and the data attribute will be the title of the movie.

**Dashboard Interactivity Implementation**:

Implementing interactivity for a dashboard with filtering options and genre selection for a plot displaying movie budgets on the y-axis and movie titles on the x-axis involves several components:

* **Dashboard Layout:** The dashboard layout should include the visualization area where the plot will be displayed, along with dropdown menus, checkboxes, or input fields for filtering and genre selection.
* **Genre Selection:** A dropdown menu or another interactive element allows users to select a specific movie genre. This selection filters the data to display only movies belonging to the chosen genre.
* **Filtering Options:** Additional filtering options could include sliders for budget ranges, checkboxes for specific movie attributes (e.g., language, release year), or text input for searching movie titles.

# 

# **Plot 9:**

**List of What They Will Be Used For:**

* Assessing Film Duration: This visualization is designed to assess the total duration of movies or series. It allows users to understand the typical lengths of films and how they might vary across different categories or genres.
* Analyzing Viewing Experience: By providing insights into film lengths, the dashboard can be used to analyze potential viewer experience implications, as duration can affect audience engagement and content scheduling.

**Plots Connected to Each One:**

* Detailed Duration Breakdown: Upon selecting a specific category or movie within the bar chart, a detailed plot or table could display further information, such as average duration, comparisons with other films, or breakdowns by season or episode for series.

**Value Range and Data Attribute**:

* Duration in Minutes (Value Range): The primary data metric visualized is the duration in minutes, represented by the length of each bar in the horizontal bar chart.
* Film Categories or Individual Films (Data Attribute): The data attribute is likely film categories or individual film titles that classify the data displayed by each bar.

**Dashboard Interactivity Implementation:**

1. Category/Film Selection:

* Function: Allows users to interact with the chart by selecting a film category or an individual film.
* Result: Triggers an update in connected plots to show a detailed breakdown of duration or related metrics for the selected item.

2. Filter by Genre or Type:

* Control: Users could use a filter control to select specific genres or types of content, such as action, comedy, drama, or documentary.
* Function: Filters the main bar chart to display only the durations corresponding to the selected filter.
* Result: Provides a focused view on how film durations vary across different genres or content types.

3. Time Period Selector:

* Control: A dropdown or slider control that lets users choose a time period, like release year or decade.
* Function: Filters the data to include only movies or series released within the selected time period.
* Result: Users can explore how the average duration of movies or series has changed over time, offering insights into trends and industry standards.

# **Plot 10:**

**List of What They Will Be Used For:**

* Showcasing Box Office Performance: The visualization is designed to display the worldwide gross income of movies, providing a clear picture of their commercial success.
* Ranking of Films by Revenue: It also serves to rank films by their gross income, helping viewers see which movies have earned the most at the box office globally.

**Plots Connected to Each One:**

* Detail View of Film Financials: When a specific film is selected from the bar chart, a connected plot or detailed view could be updated to show more granular financial data, such as opening weekend income, domestic vs. international earnings, and budget versus gross comparisons.

**Value Range and Data Attribute:**

* Worldwide Gross Income (Value Range): The main metric displayed is the monetary value representing worldwide gross income, which is the length of the bars in the bar chart.
* Film Titles (Data Attribute): The films themselves, with their titles serving as the categorizing attribute for the income data.

**Dashboard Interactivity Implementation:**

1. Film Selection:

* Function: Enables the selection of a specific film within the chart to view more detailed financial data.
* Result: Updates a connected, detailed view with comprehensive financial metrics for the selected film.

2. Genre and Year Filters:

* Controls: Filters that allow users to select genres or a range of release years.
* Function: Filters the main bar chart to show gross income for films that match the selected criteria.
* Result: Tailors the displayed information, providing insights into how box office performance varies by genre or over time.

3. Comparison Feature:

* Control: A feature that allows users to select multiple films for a side-by-side comparison of their gross incomes.
* Function: Facilitates the direct comparison of box office earnings between selected films.
* Result: Enables an analysis of what factors might contribute to the financial success of films, such as release timing, marketing strategies, or star power.

# **Section 7: References**

**Published Dashboard:**

<https://public.tableau.com/views/DV-team18project/Dashboard1?:language=en-US&:display_count=n&:origin=viz_share_link>

**Dataset:** <https://www.kaggle.com/code/sujalbhagathansda/netflix-recommendation-system/input?select=IMDb+movies.csv>

**Mural**:

<https://app.mural.co/t/dvproject8825/m/dvproject8825/1699564214984/1e4d92fdcf1e4eb161161cc853ef0ad26e234d70?sender=ud6d0ef268bcc87ae80c21683>

# 

# 

# 

# 

# 

# 

# 

# 