

Recommendation Engine for Media and Entertainment Platform

Title: - Social Suggest

By: -Team 2

Under Guidance of Abinash Sahoo

TEAM MEMBERS (Team 2)

- Ashutosh
- Lahari
- Varshini
- Mayank
- Mohana
- Pratish
- Sai
- Sanika
- Veera
- Vivek
- Vaishnavi
- Pratibha
- Sahil

1. MODELING TEAM

The aim of this part is to correctly model the data using certain domains based on different users hobbies, habits and likings while surfing over media and entertainment platforms. This will help in building the recommendation system on media and entertainment platform hook audience by suggesting them content based on their liking and demographic location.

Library imported :-

The libraries that were imported were numpy , pandas, sklearn.matrix.pairwise (Calculation of cosine similarity) and finally sklearn.preprocessing for (StandardScaler, OneHotEncoder). Import necessary libraries for data manipulation, similarity calculations, and preprocessing.

Loading and Preprocessing Data :-

We load the dataset from a CSV file. This dataset contains information about users, their ratings for different items, and other relevant details.

It utilizes the `read_csv` function from the Pandas library to import the dataset.

The file path of the dataset is specified as

“C: \Users \sai \Downloads \EntertainmentRecommendation.csv”, and the encoding parameter is set to ‘**latin1**’ to handle potential character encoding issues. This operation reads the dataset into a Pandas DataFrame named **data**, allowing for further manipulation and analysis.

In simpler terms, this code segment reads a CSV file containing entertainment recommendation data into a Pandas DataFrame, enabling subsequent analysis and processing of the dataset.

- **Extract user demographics:**

To extract user demographics from the dataset, the code selects specific columns - 'User ID', 'Age', and 'Gender' - from the provided data. This selection is achieved using the `[['User ID' , 'Age' , 'Gender']]` indexing method applied to the DataFrame named 'data'.

- Furthermore, the `drop_duplicates()` function is employed to ensure that each user is represented only once, eliminating any duplicate entries based on the combination of 'User ID', 'Age', and 'Gender'. This step helps in maintaining the uniqueness of users within the dataset, facilitating subsequent analyses and processing.

The extracted user demographics, comprising unique combinations of 'User ID', 'Age', and 'Gender', are stored in a new DataFrame named

'user_demographics', ready for further preprocessing and analysis.

- ***One-hot encode Gender:***

During the preprocessing stage, gender information from the user demographics is transformed into a numerical representation suitable for machine learning algorithms using a process called one-hot encoding. This is achieved through the following code:

encoder_gender = OneHotEncoder() : This line initializes a OneHotEncoder object, which will be used to perform one-hot encoding on the gender feature.

- **gender_encoded = encoder_gender.fit_transform(user_demographics [[' Gender']]).toarray() :** This line applies the one-hot encoding transformation to the 'Gender' column of the user_demographics DataFrame. The **fit_transform()** method first fits the encoder to the 'Gender' data to learn the categories, and then transforms the data into a one-hot encoded format. Finally, **.toarray()** converts the encoded sparse matrix into a dense array for easier handling.

In essence, this code segment converts categorical gender data into a numerical format suitable for further analysis and modeling, ensuring that the machine learning algorithm can effectively utilize this feature during training.

- ***Standardize Age:***

To standardize the age feature, the **StandardScaler()** function from the **scikit-learn library** is utilized.

This function initializes a scaler object named '**scaler_age**'.

Then, the **fit_transform()** method is applied to the '**Age**' column of the **user_demographics** DataFrame.

This method calculates the mean and standard deviation of the age data and standardizes it accordingly, resulting in '**age_scaled**', which contains the standardized age values.

Standardization involves transforming the data such that it has a mean of zero and a standard deviation of one, making it more suitable for certain machine learning algorithms.

Creating User-Item Matrix :-

The user-item matrix represents the ratings given by users to various items (e.g., movies, games).

We use the **pd.pivot_table** function to transform the original dataset into a matrix format where:

Rows (index): Each row represents a unique user (User ID).

Columns (columns): Each column represents a unique item (Item ID).

Values (values): The cell values are the ratings that users have given to items. If a user hasn't rated an item, the fill_value=0 parameter fills the cell with 0.

The user-item matrix is pivotal for representing the interaction between users and items. Each cell in this matrix contains the rating a user has given to an item, or zero if no rating has been given.

This matrix format is essential for applying collaborative filtering techniques, where the goal is to predict a user's rating for an item based on the ratings of similar users.

Aligning User Features with User-Item Matrix :-

Aligning User Features with User-Item Matrix is done to align the user demographic data (user_features) with the user-item rating data (user_item_matrix).

Intersection: common_users finds the intersection of user IDs present in both user_features and user_item_matrix. This ensures we only consider users who have both demographic information and rating data.

Subset Selection:

user_features = user_features.loc[common_users] selects only the rows in user_features corresponding to common_users.

user_item_matrix = user_item_matrix.loc[common_users] does the same for user_item_matrix.

User demographic features and user-item ratings are initially sourced from different parts of the dataset. This step ensures that the subsequent computations consider only those users who have both demographic data and rating data. Ensuring both matrices have the same set of users is crucial for accurate similarity computation and recommendation generation. It avoids mismatches and errors that could arise from having differing sets of users in the two matrices.

Computing User Similarities :-

The user similarities is calculated to find the similarity between users based on their demographic features.

`cosine_similarity(user_features)` computes the cosine similarity between all pairs of users. Cosine similarity is a measure that calculates the cosine of the angle between two vectors, which in this case are the demographic feature vectors of the users.

The resulting `user_similarities` is a matrix where each entry (i, j) represents the similarity between user i and user j .

The similarity matrix derived from user demographic features is used to find users who are demographically similar. This similarity information is then used to predict ratings and generate recommendations.

The computed similarities are used in the `predict_ratings` function to generate predictions based on weighted averages of ratings from similar users. This is crucial for generating recommendations in `generate_recommendations_for_demo` function.

Predicting Ratings :-

```
def predict_ratings(user_item_matrix, user_similarities):
    # Calculate the sum of similarities for each user
    denominator = np.sum(user_similarities, axis=1)

    # Calculate the weighted sum of item ratings
    numerator = np.dot(user_similarities, user_item_matrix)

    # Reshape the denominator to facilitate division
    denominator_resaped = denominator[:, np.newaxis]

    # Perform element-wise division, handling division by zero
    predicted_ratings = np.divide(numerator, denominator_resaped,
    out=np.zeros_like(numerator), where=(denominator_resaped != 0))

    return predicted_ratings
```

The `user_item_matrix` is a 2D NumPy array in which the elements show the ratings that users have assigned to various objects. Each row in the matrix represents a user, and each column is an item. This second 2D NumPy array is `user_similarities`, and each element (i, j) indicates how similar users i and j are to one another.

Initially `predict_ratings` computes the denominator using the formula **`denominator = np.sum(user_similarities, axis=1)`** in order to determine the expected ratings and adds up the similarities for every user, creating a 1D array that is used to normalize the projected ratings. Each member in the array indicates the total similarities for a particular user.

Then follows with **`numerator = np.dot(user_similarities, user_item_matrix)`** to generate the numerator, which determines the weighted total of the item ratings by utilizing the user similarities. The outcome is a 2D array with each member denoting the weighted total of ratings for a pair of users and items.

The denominator is then reshaped using **`denominator_reshaped = denominator[:, np.newaxis]`** to create a 2D column vector that will help with broadcasting in the division step.

In order to determine the predicted ratings, **`predicted_ratings = np.divide(numerator, denominator_reshaped, out=np.zeros_like(numerator), where=(denominator_reshaped!= 0))`** is used. This method divides the numerator by the reshaped denominator and handles division by zero by utilizing `where=(denominator_reshaped!= 0)` and `out=np.zeros_like(numerator)` to guarantee that no invalid operations take place. For every user-item pair, the function outputs a 2D array of anticipated ratings.

The influence of each user's rating corresponds with their similarity, the anticipated ratings are normalized by the total of the user similarities. Using `np.divide` with the `where` parameter ensures that division by zero is handled gracefully, avoiding unexpected runtime errors or erroneous results. The usage of matrix operations (`np.dot` and broadcasting) improves code efficiency and leverages NumPy's enhanced performance for huge datasets. The method is predicated on the idea that people who have comparable rating habits will probably rate things similarly. This is a widely used method in recommendation systems for collaborative filtering.

Generating Recommendation :-

The recommendation system uses user demographic data and user-item interactions to generate personalized suggestions. First, the system loads and preprocesses user demographic data, including age and gender. Gender is one-hot encoded, and age is standardized to ensure consistency. These

processed features are combined into a single DataFrame representing user demographics. Simultaneously, a user-item matrix is created from user ratings, filling in missing values with zeros. To ensure alignment, only users present in both the user features and the user-item matrix are considered.

Next, the system calculates the cosine similarity between users based on their demographic features, creating a user similarity matrix. This matrix quantifies how similar each pair of users is. Using this similarity information, the system predicts ratings for items that users haven't rated yet. This is done by taking a weighted average of ratings from similar users, ensuring that more similar users have a greater influence on the predicted ratings.

To generate recommendations, the system focuses on a specific target demographic profile. This profile is compared against existing user profiles to find similarities. Based on these similarities, the system predicts ratings for items for the target demographic. The predicted ratings are then sorted in descending order. To tailor the recommendations further, items are filtered by the desired type, such as 'Game', and the top-N items are selected for recommendation. This method ensures that the recommendations are both personalized and relevant, leveraging the similarities between users' demographic profiles and their past interactions with items.

Example Usage :-

To illustrate how the recommendation system works, let's consider an example usage scenario. Suppose we have a target demographic of a 30-year-old male user who is interested in receiving recommendations for 'Games'. The demographic features of this target user are first encoded and standardized to match the format of the existing user features in the dataset. Python code :-

```
# Example target demographic: age=30, gender='Male'
target_demo = np.array([1, 0, 30]) # Assuming 0 is Female and 1 is Male
after one-hot encoding

# Standardize the age feature
target_demo[2] = scaler_age.transform([[target_demo[2]]])[0, 0]

desired_type = 'Game'
```

In this step, the gender is one-hot encoded (1 for male and 0 for female), and the age is standardized using the previously fitted StandardScaler.

Next, the system calculates the cosine similarity between the target demographic and the existing user profiles. This allows the system to identify users with similar demographic characteristics.

The target user is interested in 'Games', so we set the desired type to 'Game'. Using the calculated similarities, the system predicts ratings for all items in the user-item matrix, focusing specifically on items categorized as 'Games'.

The `generate_recommendations_for_demo` function is called with the target demographic profile, user features, user-item matrix, and the desired item type. This function calculates predicted ratings for items by averaging the ratings from users who are similar to the target demographic. The items with the highest predicted ratings are then filtered to include only those of the desired type ('Game'), and the top-N items are selected.

Finally, the recommended items are printed out, showing the title, genre, and type for each recommended item. This provides the target user with a list of personalized game recommendations based on their demographic profile and the preferences of similar users. This process ensures that the recommendations are relevant and tailored to the user's interests, enhancing their overall experience with the recommendation system.

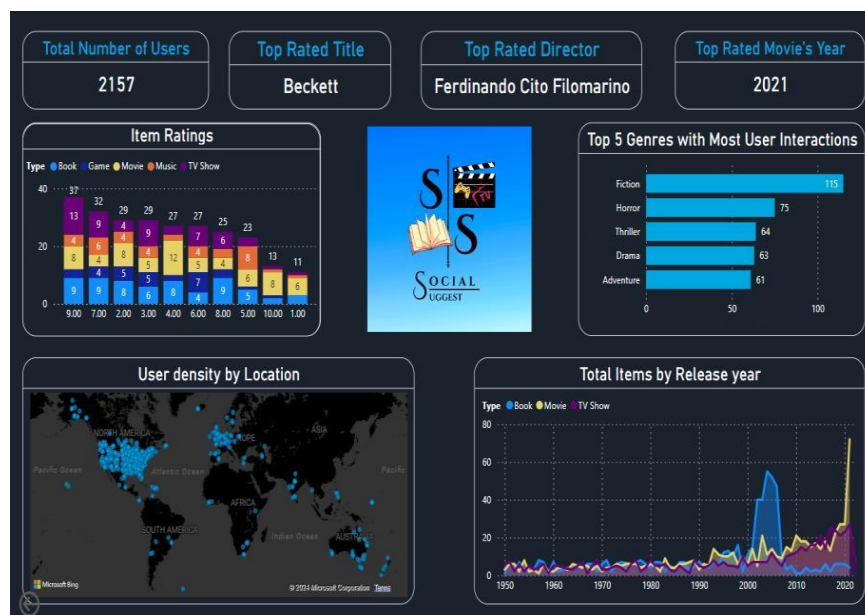
Enhancements and Additional Features :-

1. Improved Recommendation Logic: Implement more sophisticated recommendation algorithms such as collaborative filtering or matrix factorization.
2. Additional User Features: Include more demographic features like location, occupation, etc.
3. Hybrid Recommendations: Combine demographic-based recommendations with content-based or collaborative filtering.
4. User Feedback Integration: Incorporate user feedback to improve recommendation accuracy over time.
5. Performance Optimization: Optimize computation for large datasets, potentially using distributed computing techniques.
6. Evaluation Metrics: Implement metrics like RMSE, MAE, or Precision to evaluate the recommendation system's performance.

DATA VISUALIZATION REPORT

The aim of this part is to correctly visualize certain domains based on different users hobbies, habits and likings while surfing over media and entertainment platforms. This will help in building the recommendation system on media and entertainment platform hook audience by suggesting them content based on their liking. For this, we created a dashboard with multiple pages. Each page has its own importance. Let us see about each page in detail.

1. OVERVIEW



Why this page is needed:

- The Overview page serves as a comprehensive dashboard that provides a high-level summary of key metrics and insights derived from the dataset.
- This page is designed to offer users a clear and concise visualization of the most important aspects of the data, including top-rated content, user engagement metrics, and detailed visualizations.
- The Overview page integrates various elements to offer a holistic view of the dataset, combining high-level metrics with detailed visualizations.
- It effectively summarizes user engagement and content popularity, providing valuable insights for decision-making and strategy development.

Social Suggest Logo:



The logo for our project, "Social Suggest," symbolizes the core mission and vision of the platform. The design elements and colours were carefully chosen to reflect the project's emphasis on community-driven recommendations and user engagement. Social Suggest aims to leverage the collective wisdom of its user base to provide personalized and highly reliable recommendations for movies. The platform is designed to help users discover new content based on the preferences and ratings of a large, diverse community. By focusing on community-driven insights, Social Suggest aims to revolutionize the way users discover and enjoy new content.

Cards:

Total Number of Users	Top Rated Title	Top Rated Director	Top Rated Movie's Year
2157	Beckett	Ferdinando Cito Filomarino	2021

Why cards are needed:

- Card visualizations are a crucial element in data dashboards and reports, providing a clear and concise way to present key metrics and highlights at a glance. Card visualizations are an essential component of data dashboards, providing a quick, clear, and user-friendly way to present key metrics.

1. Total Number of Users: Card Visualization

This card displays the total number of users who have interacted with the model. This metric is crucial for understanding the size and scope of the dataset, as well as the breadth of user engagement. This number reflects the

engagement level with the platform or application, indicating how many unique users have rated, reviewed, or interacted.

2. Top Rated Title: Card Visualization

This card shows the title with the highest average rating given by users. It represents the most well-received title according to user reviews and ratings. The top rated title is a key indicator of user preferences and trends within the dataset. It highlights the most appreciated content and can be used for recommendations or promotional purposes.

3. Top Rated Director: Card Visualization

This card highlights the director whose titles have received the highest average ratings from users. It identifies the most favoured director based on user feedback. The top rated director card showcases the filmmaker whose work resonates the most with the audience. It reflects the director's ability to consistently produce highly rated content. Often, the top rated movie is directed by the top rated director, indicating a strong correlation between the director's influence and the movie's success.

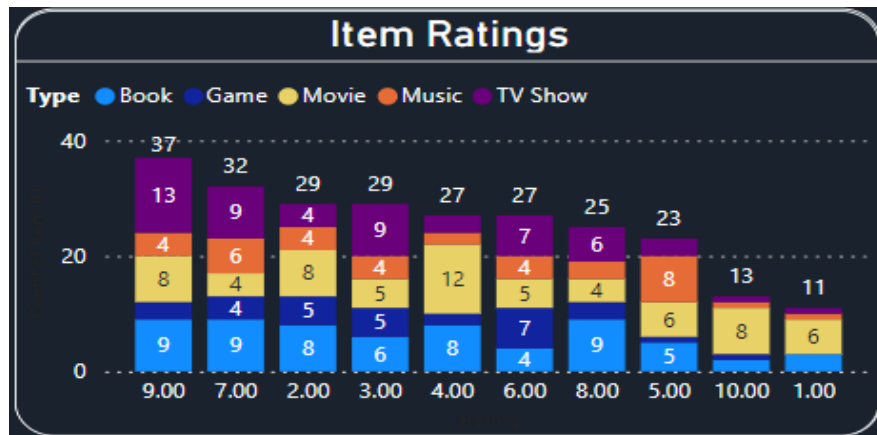
4. Top Rated Movie's Year: Card Visualization

This card displays the year in which the top rated movie was released. It provides context to the top rated title by situating it within a specific timeframe. Knowing the year of release helps understand the movie's reception in its contemporary context and its enduring popularity. This can be useful for analyzing trends over time, such as which years produced the most highly rated content and how audience preferences have evolved.

Together, these cards offer a comprehensive view of the highest quality content in the dataset, from the individual movie to the creative mind behind it, and the era it belongs to.

Item Ratings: Stacked Column Chart Visualization

- The "Item Ratings" visualization presents a stacked column chart that illustrates the distribution of ratings across different types of items, including books, games, movies, music, and TV shows. Each column represents a specific rating value on the x-axis, while the y-axis displays the count of unique item IDs.



Why This Chart is needed:

- The stacked column chart enables users to analyze the distribution of ratings across different types of items. Users can identify trends and patterns in user preferences for each item type and rating category. Comparing the total sum of item IDs across rating categories allows for insights into overall user engagement and satisfaction levels.

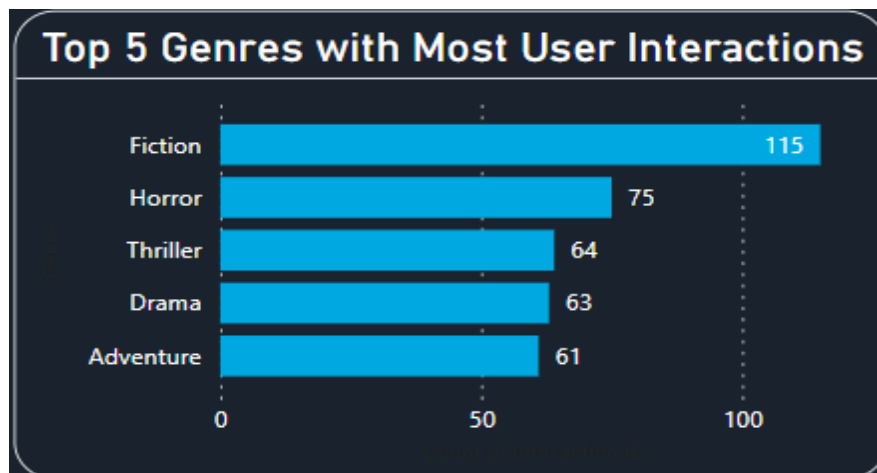
Details:

- X-axis:** Rating (that contains 1,2,3,4,5,6,7,8,9,10)
- Y-axis:** Count of unique Item IDs
- Legend:** Item Types (such as books, games, movies, music, and TV shows)
- Insights:** It provides a comprehensive visualization of user ratings across various types of items.

Result: The count of Item ID's with the rating 9 has the highest having the value as 37. It is divided into several types as for the TV Show we are having 13 , Music with 4, Movie with 8, Game with 3, Book with 9.

Top 5 Genres with Most User Interactions: Stacked Bar Chart Visualization

- The "Top 5 Genres with Most User Interactions" visualization presents a stacked bar chart that displays the top five genres with the highest number of user interactions. Each bar represents a specific genre on the y-axis, while the x-axis shows the count of interaction IDs.



Why This Chart is Needed:

- The stacked bar chart enables users to understand which genres are most popular based on user interactions. By analyzing the distribution of interactions across genres, users can identify trends in content preferences and engagement. This insight is valuable for content creators, marketers, and platform administrators to tailor content strategies and improve user experience.

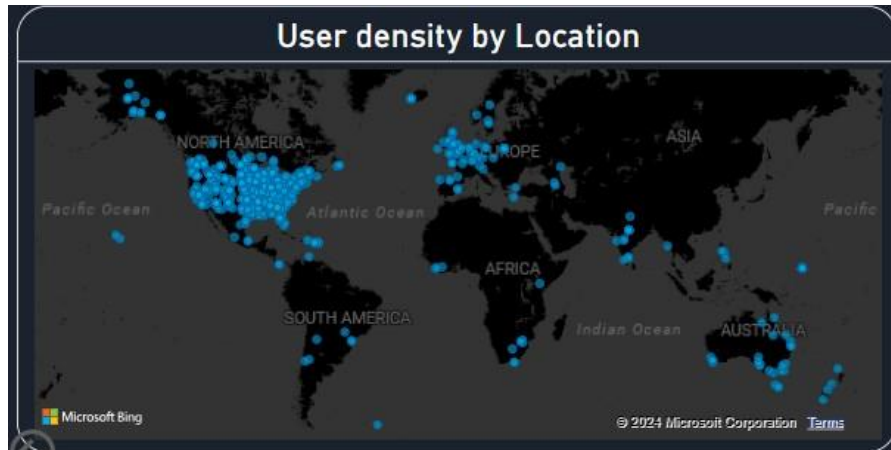
Details:

- **X-axis:** Count of Interaction IDs
- **Y-axis:** Genres (e.g., Action, Drama, Comedy, Sci-Fi, Romance)
- **Insights:** This visualization provides a clear comparison of user engagement across different genres.

Result: The genre with the highest user interactions is fiction, with a count of 115, followed by horror, which has 75 user interactions.

User Density By Location: Map Visualization

- The "User Density By Location" visualization presents a map that displays user demographics and density based on their geographical locations. This interactive map allows users to explore detailed information about user distribution, including user IDs, gender, and age, at various locations. When users hover their cursor over any point on the map, a tooltip appears, providing all relevant details for that location.



Why This Chart is Needed:

- This map visualization is essential for understanding the geographical spread and density of the user base. It provides insights into where users are concentrated and helps identify regional trends in user demographics. This information is valuable for targeted marketing, regional content customization, and improving user experience based on location.

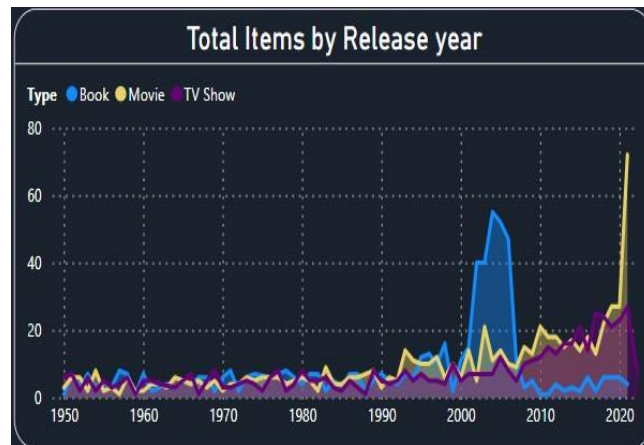
Details:

- **Location:** Location
- **Tooltips:**
 1. **User ID:** The unique identifier for each user.
 2. **Gender:** The gender of users at that location.
 3. **Age:** The age range or specific ages of users at that location.

Result: The map offers a visual representation of user density, making it easy to identify regions with high or low user concentrations. Placing the cursor on any location it displays the No. of users, ID's , Gender.

Total Items by Release Year: Area Chart Visualization

- The "Total Items by Release Year" visualization is an area chart that shows the number of items released each year, focusing on books, movies, and TV shows. The x-axis represents the release years, while the y-axis displays the count of item IDs. The chart includes a legend for item types, providing a breakdown of the number of each type released annually.



Why This Chart is Needed:

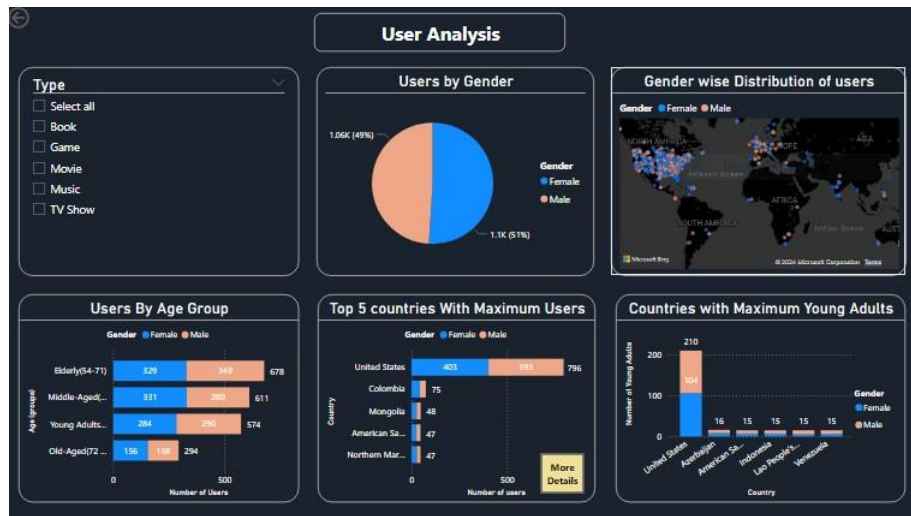
- The area chart helps users understand trends in the release of different types of items over time. It highlights peaks and troughs in content production and can assist in identifying periods of high or low activity in content creation.

Details:

- **X-axis:** Release Year
- **Y-axis:** Count of Item IDs
- **Legend:** Item Types (Books, Movies, TV Shows)
- **Insights:** The visualization provides a comprehensive view of the number of items released each year, broken down by type. It helps users track the evolution of content production and identify trends over time.

Result: For example, the year 2021 saw a peak in the release of TV shows, with a total of 27, while the number of books released was 4 and movies were 72.

2. USER ANALYSIS

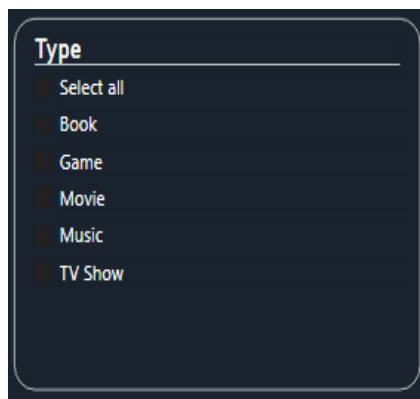


Why User Analysis is needed:

- User analysis is critical in refining and enhancing a recommendation engine by understanding the demographics, preferences, and behaviors of users.
- It gives more information about user demographics including gender, age group, location and distribution across different media categories.

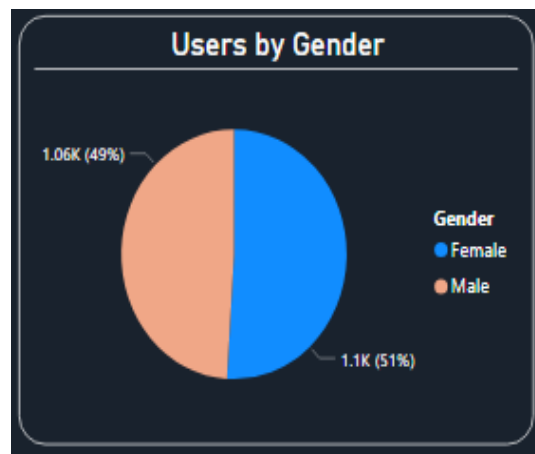
Type Selection [Slicer] or Why a Slicer was needed?

- The section involves a filter for selecting different types of content, including books, games, movies, music, and TV shows. It allows users to filter the analysis based on their interest in these specific types. This filter can help in understanding user preferences and engagement with different content types, which is crucial for targeted recommendations.



Users by Gender: Pie Chart

- A pie chart is ideal for displaying parts of a whole, making it perfect for showcasing how the user base is divided between Genders (Female and Male). A pie chart shows the distribution of users by gender. The pie chart reveals that 51% of users are Male (1.1K) and 49% are Female (1.06K).



Why pie chart is needed:

- Pie charts are ideal for displaying the proportions of a whole. In this case, it shows the distribution of users between two categories: Male and Female. The pie chart allows viewers to quickly understand the percentage split between genders.

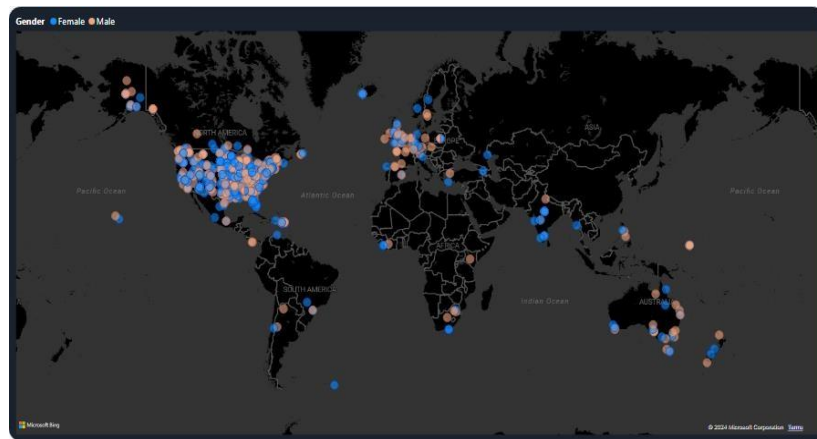
Details:

- **Legend:** Gender: Female and Male, it is indicated using Blue for Female and Orange for Male along with their corresponding percentages.
- **Values:** Count of User ID
- **Insights:** It provides an understanding of number of users by gender.

Result: The pie chart conveys a near even split between genders, with a slight majority being Male 51% as compared to Females 49%.

Gender-wise Distribution of Users: Map visualization

- A world map indicates the geographical distribution of users based on gender. Each dot represents the location of a user, color-coded by gender. Blue dots represent female users. Orange dots represent male users.



Why map is needed:

- A map is used to provide a geographical representation of data. This visual shows where users are located around the world and uses color or markers to indicate the gender distribution in different regions. It is effective for understanding geographical trends and concentrations.

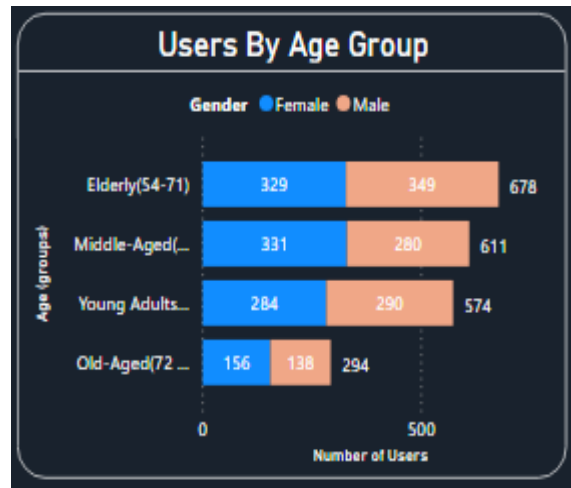
Details:

- **Location:** Location
- **Legend:** Gender (Male: orange color and Female)
- **Tooltips:** When hovering over a dot on the map, a tooltip provides detailed information about that specific data point. This typically includes:
 - ✓ Location: The geographic location of the users.
 - ✓ Gender: The gender of the users (male or female).
 - ✓ Count of User ID: The number of users at that specific location.
 - ✓ Average of Age: The average age of the users at that location.

Result: There is a high concentration of users in North America, particularly in the United States, with a mix of both male and female users. There are fewer users in South America compared to North America, with a notable concentration in Brazil. Europe has a significant distribution of users spread across various countries. There are scattered users across Asia, including India, China, and Southeast Asia. Australia has a noticeable number of users distributed across the country. There are sparse dots representing users in Africa, with a few scattered across the continent.

Users by Age Group: Stacked Bar Chart

- A stacked bar chart is well-suited for comparing discrete categories, which in this case are the Age (groups). A stacked bar chart breaks down users by Age (groups) and Gender and the number of users falling into each Age (groups). The Age (groups) include Young Adults (18-35), Middle-Aged (36-53), Elderly (54-71), and Old-Aged (72 and above).



Why stacked bar chart is needed:

- Stacked bar charts are useful for comparing quantities across different categories. Here, it shows the number of users in different age groups, separated by gender. The bars make it easy to compare the sizes of these groups and see the distribution of users by age.

Details:

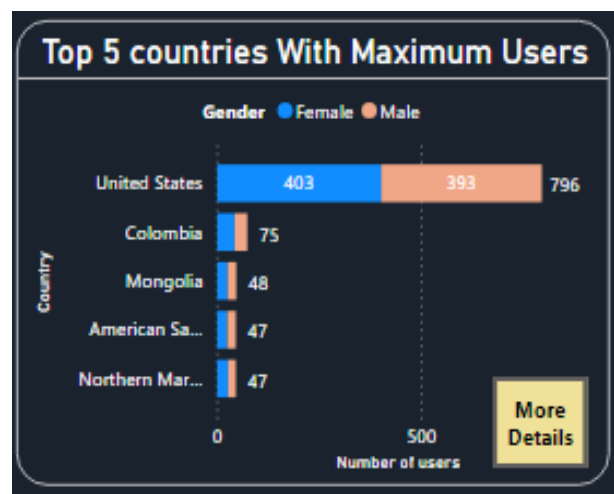
- **X-axis:** Number of Users (representing the number of users in each age group)
- **Y-axis:** Age (groups), specific labels for each age range Young Adult: 18-24, Middle-aged: 25-53, Elderly: 54-71, Old Age:72 and above.
- **Legend:** Gender (Female in blue, Male in orange)
- **Insights:** The chart provides a clear overview of number of users falling under each age group.

Filter applied: Clicking on male or female users will enable the chart to show the number of users by gender falling under each Age group.

Result: The largest user groups are the Elderly (678) further dividing into 329 females and 349 males and the lowest number of users are in Old Aged (294) further dividing into 156 females and 138 males. The largest user groups are Elderly (678) followed by Middle-aged (611), Young adults (374) and Old Aged (294). The number of Male users are high in Elderly while in other age-groups the number of Female users are high as compared to Males.

Top 5 Countries with Maximum Users: Stacked Bar Chart

- A stacked bar chart effectively highlights number of users across different categories country wise. This stacked bar chart lists the top five countries with the highest Number of users, segmented by Gender. The chart describes the “Top 5 countries with Maximum Users”. The chart describes the Number of users within a particular country segmented by gender. The Top 5 Countries are filtered by using Top N feature that filters the Top 5 countries on the basis of Number of users.



Why stacked bar chart is needed:

- Bar charts are effective for ranking and comparing the sizes of different groups. This chart displays the top five countries with the most users, making it easy to see which countries have the largest user bases and compare the gender distribution within those countries.

Details:

- **X-axis:** Number of Users

- **Y-axis:** Country
- **Legend:** Gender, the legend distinguishes between different types of Gender such as Blue for Females and Orange for Males.
- **Insights:** The chart provides a clear visual representation of gender distribution within each country.

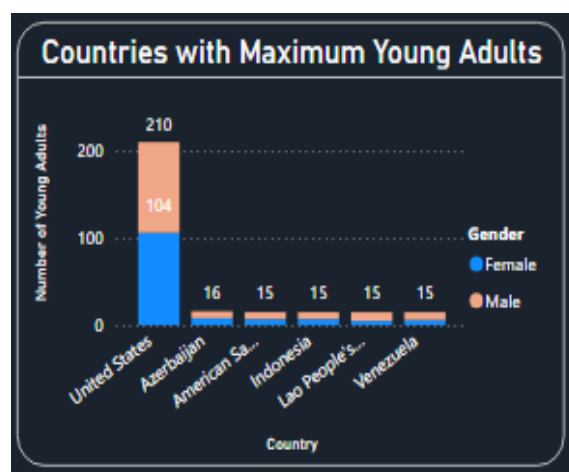
Button feature: A button is included in this “More details” for the purpose of navigating to “Detailed Analysis” page to give us more insights about the number of users in other countries also.

Filter Applied: Clicking on a segment, such as the male users in the United States, will highlight the specific segment within the chart. For instance, clicking on the orange part representing males will show the 393 males out of the total 796 users in the USA.

Result: The United States has the highest user base, suggesting a strong market presence. The United States leads with 796 users, followed by Colombia-75 users, Mongolia 48 users, American Samoa and Northern Mariana Islands with same number of users i.e.,47 users. The number of Female users are high in these countries as compared to Male users.

Countries with Maximum Young Adults: Stacked Column Chart

- A stacked column chart titled “Countries with Maximum Young Adults” displays countries with the highest Number of Young Adult users (18-35 years), segmented by gender.



Why stacked column chart is needed:

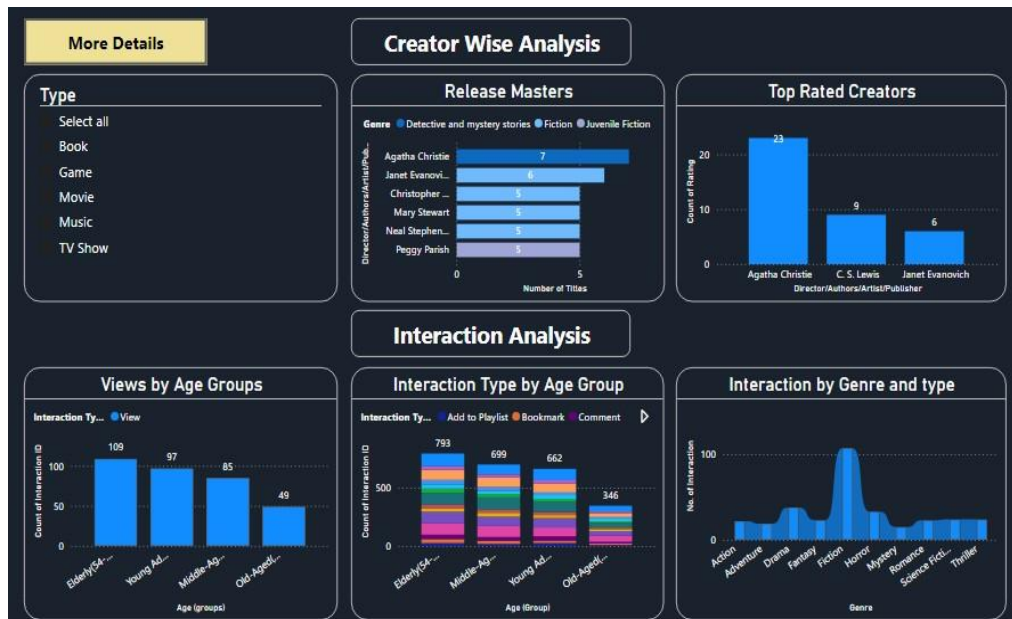
- It allows for easy comparison of the total number of young adults between countries. By stacking the columns, it shows how the total is divided between males and females within each country. It consolidates data into a single column per country, making efficient use of space and allowing for more countries to be displayed without cluttering the chart.

Details:

- **X-axis:** Country
- **Y-axis:** Number of Users
- **Legend:** Gender (Male: Orange color and Female: Blue color)
- **Insights:** The chart highlights the concentration of young adult users within each country.

Result: The United States has the highest count (210), followed by several countries with much smaller Young Adult populations. The United States not only has the most users overall but also the Young Adult users followed by several countries with small number of users Azerbaijan with 16 users, and American Samoa, Indonesia, Lao People's Democratic Republic, Venezuela with same number of users (15 users).

3. Creator wise & Interaction Analysis



Creator Wise Analysis

- The Creator Wise Analysis is a crucial tool in understanding the performance and impact of various content creators within a given platform and to identify which Creators (Authors, Directors, Artists, Publishers etc.) are most prolific and highly rated.

Why it was needed:

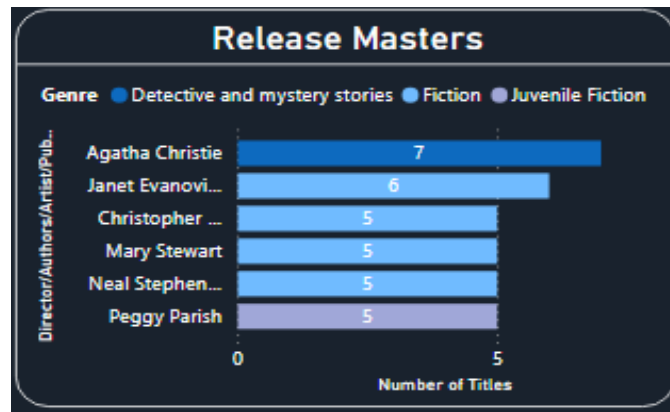
- Understanding which Creators have the most titles released helps in recognizing prolific content producers.
- Identifying top-rated creators highlights who is most appreciated by the Users.

More Details

Button feature: The Creator wise analysis contains a button, which will navigate us to “Detailed Analysis” page after clicking on it. This button was necessary so as to get more details about the “Release Masters” as the current visuals only includes those Director/Author/Artists/Publisher who have the count of titles greater than or equal to 5.

Release Masters: Stacked Bar Chart

- The above is stacked bar chart which visualizes the Number of Titles released by various Directors/Authors/Artist/Publishers across different genres. The visualization only shows those values which has Number of titles greater than or equal to 5 corresponding to a particular Director/Authors/Artist/Publisher.



Why this chart is needed:

- A stacked bar chart is ideal here because it allows for easy comparison of the number of titles across creators. It accommodates long names and ensures that the viewer can quickly have a look which creators have the most titles.

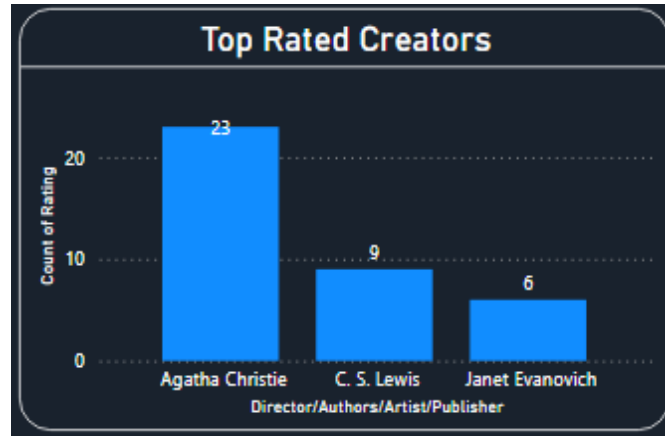
Details:

- **X-axis:** Number of Titles
- **Y-axis:** Director/Authors/Artists/Publisher
- **Legend:** Genre (Detective and mystery stories, Fiction, Juvenile Fiction). Different shades of Blue color are used to indicate different Genre.
- **Insights:** It provides an overview of number of titles released by each director/artist/publisher/author.

Results: Agatha Christie leads with 7 titles in Detective and Mystery Stories. Janet Evanovich follows with 6 titles in Fiction. Christopher Paolini, Mary Stewart, and Neal Stephenson each have 5 titles in Fiction. Peggy Parish has 5 titles, mainly in Juvenile Fiction.

Top Rated Creators

- The above mentioned is a stacked column chart titled “Top Rated Creators” which displays the Count of Rating received by Top Creators.



Why this chart is needed:

- The stacked column chart type is effective for showing the magnitude of ratings for each creator, making it easy to compare the popularity.

Details:

- **X-axis:** Director/Authors/Artist/Publisher
- **Y-axis:** Count of Rating
- **Insights:** This visualization provides a clear understanding of Top Rated Creators.

Results: Agatha Christie is the top-rated creator with 23 ratings. C.S. Lewis follows with 9 ratings. Janet Evanovich has less count of ratings which is 6.

Interaction Analysis

- The Interaction Analysis page provides a detailed look into user engagement with various content types across different demographics. This analysis helps in understanding user behavior and preferences, allowing for better content targeting and personalized recommendations.

Slicer: This section has a slicer for selecting different Types of content, including Books, Games, Movies, Music, and TV shows. It allows users to filter the analysis based on their interest in these specific types.

Type

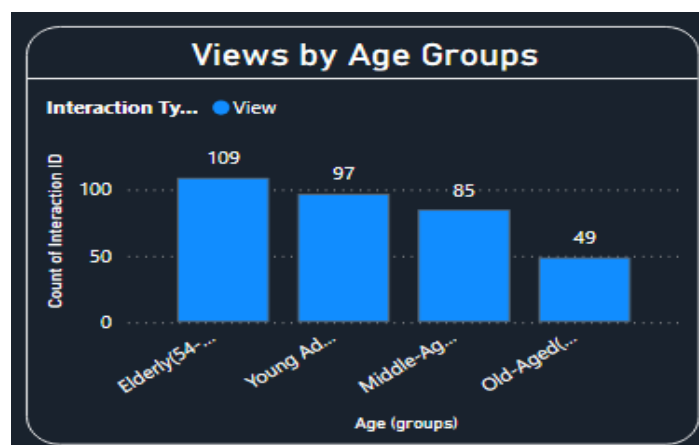
☐ Select all
 ☐ Book
 ☐ Game
 ☐ Movie
 ☐ Music
 ☐ TV Show

Why Interaction Analysis was necessary:

- To understand how different age groups interact with content and the types of interactions they engage in.
- To know more about which age groups are viewing content the most helps in tailoring content offerings.
- To reveal the preferences for different types of interactions (e.g., adding to playlists, bookmarking, commenting and many more) among various age groups. It provides insights into user behavior and engagement strategies.
- To analyze interactions by Genre and Type helps in understanding which Genres are most engaging and the preferred ways users interact with them.

Views by Age Groups: Stacked column chart

- The above chart titled “Views by Age groups” visualizes using a stacked column chart which shows the Count of Interaction ID (Views) by different Age (groups).



Why stacked column chart is needed:

- This chart illustrates the number of interactions (views) across different age groups. Stacked column charts are effective for categorical comparisons, and they make it simple to compare the engagement levels between age groups.

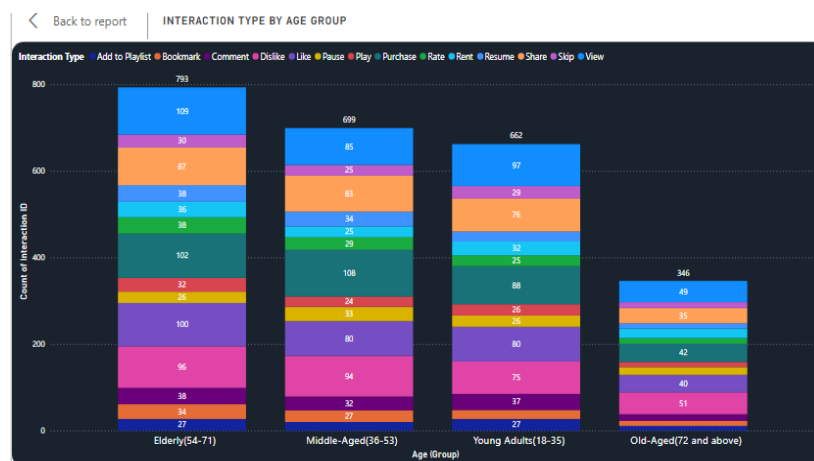
Details:

- **X-axis:** Age (groups), (Elderly, Young Adult, Middle-Aged, Old-Aged)
- **Y-axis:** Count of Interaction ID
- **Legend:** Interaction Type (View) represented by a Blue color.
- **Insights:** This visualization provides a clear understanding of count of Interaction ID across different age groups.

Result: Count of Interaction ID are highest for Elderly (109), followed by Young Adults (97), Middle Aged (85) and lowest being the Old Aged (49).

Interaction Type by Age Group: Stacked column chart

- The Stacked column chart shows the various types of interactions (Add to Playlist, Bookmark, Comment, Dislike, Like, Pause, Purchase, Rate, Rent, Return, Share, Skip, View) across different age (Group) along with the count of interactions.



Why stacked column chart is needed:

- A stacked bar chart allows for the comparison of total interactions per age group while also breaking down the contribution of each interaction type

within those groups. This provides a comprehensive view of both overall and specific interaction trends.

Details:

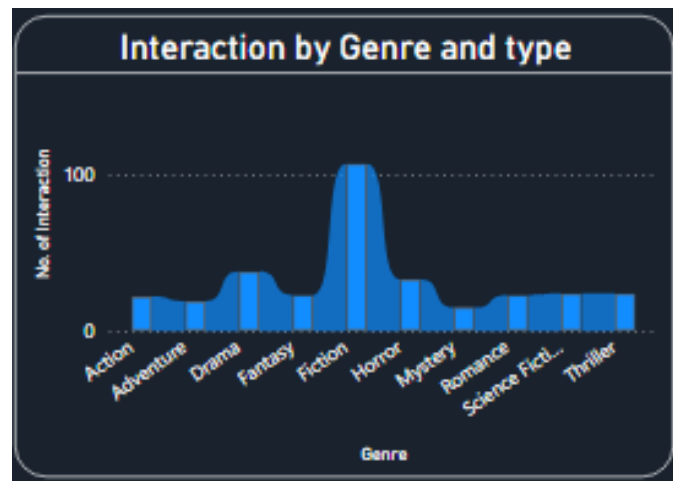
- **X-axis:** Age (Group)
- **Y-axis:** Count of Interaction ID
- **Legend:** Interaction Type (Add to Playlist, Bookmark, Comment, Dislike, Like, Pause, Play, Purchase, Rate, Rent, Resume, Share, Skip, View)
- **Insights:** It provides a clear understanding of total interactions per age group while also knowing the interaction type.

Result:

- **Elderly:** Highest interactions with a total of 793 interactions and the Interaction Type highest in Elderly is View (109) followed by Purchase (102), Like (100), Dislike (96).
- **Middle Aged:** Second highest with 699 interactions, the Interaction Type highest in Middle-Aged is Purchase (108) followed by Dislike (94), Share (83), Like (80).
- **Young Adults:** Close behind Middle Aged with 662 interactions, the Interaction Type highest is View (97) followed by Purchase (88), Like (80), Share (76).
- **Old Aged:** Significantly lower interactions with 346 interactions, the Interaction Type highest is Dislike (51) followed by View (49), Purchase (42), Like (35).

Interaction by Genre and Type: Ribbon chart

- The Ribbon chart visualizes the number of interactions across different genres. According to slicer we can view it for different types and get to know about Number of Interaction by Genre based on type selected.



Why Ribbon chart is needed:

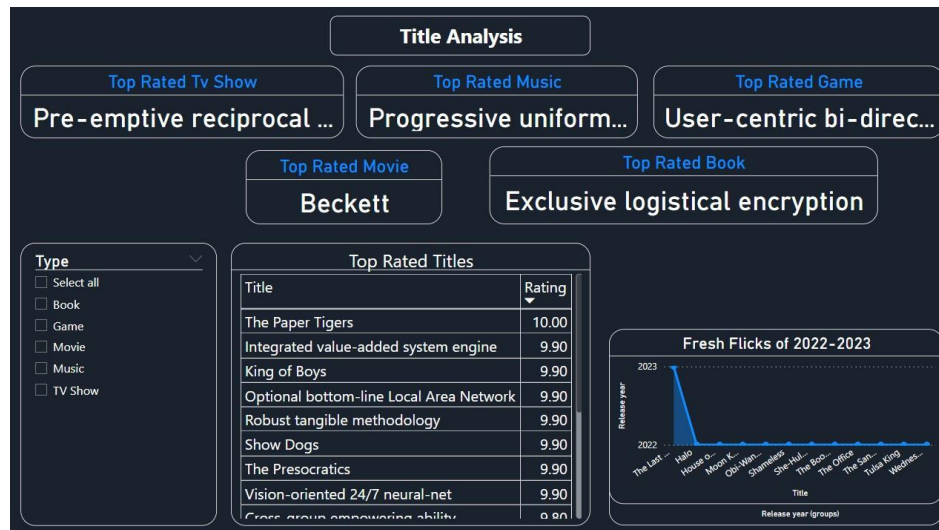
- A ribbon chart shows how interactions vary by genre over time or across different types. It highlights changes and trends, making it easy to see which genres are gaining or losing popularity and how they rank relative to each other over the period. The ribbons show the transitions smoothly, helping to compare interactions between genres in a visually appealing way.

Details:

- **X-axis:** Genre, Type
- **Y-axis:** Number of Interaction
- **Genres:** Top 10 Genres are filtered based on Rating for each type, Includes Action, Adventure, Drama, Fantasy, Fiction, Horror, Mystery, Romance, Science-Fiction, Thriller, etc.
- **Insights:** This visualization provides a clear understanding of how interactions vary by genre over time or across different types.

Result: Interaction patterns vary significantly by genre, highest is Fiction with number of interactions (107) and Drama (38) being more popular, followed by Horror (33) and Mystery being the lowest (15).

4. TITLE ANALYSIS



Why this page is needed:

- Title analysis is a crucial aspect of understanding user engagement, content performance, and trends in media consumption.
- It involves examining various attributes of titles such as ratings, release periods, genres, and user interactions to derive meaningful insights.
- This analysis is essential for content creators, marketers, and platform administrators to make informed decisions.
- By incorporating title analysis into the dashboard, we can provide stakeholders with the insights needed to make data-driven decisions and improve overall platform performance.

Cards:



Why cards are needed:

- Card visualizations are a crucial element in data dashboards and reports, providing a clear and concise way to present key metrics and highlights at a

glance. Card visualizations are an essential component of data dashboards, providing a quick, clear, and user-friendly way to present key metrics.

Components of Each Card:

1. **Title:** Displays the name of the top-rated item.
2. **Rating:** Shows the highest rating received by the item.
3. **Type:** Indicates the type of content (e.g., TV show, movie, music, game, book).
4. **Description or Additional Details:** May include a brief description or relevant details about the item.

Individual Cards Description:

1. Top Rated TV Show: Card Visualization

- Title: Name of the highest-rated TV show.
- Rating: Maximum rating received by the TV show.
- Type: TV show.
- Description: Brief synopsis or notable details about the TV show.

2. Top Rated Movie: Card Visualization

- Title: Name of the highest-rated movie.
- Rating: Maximum rating received by the movie.
- Type: Movie.
- Description: Brief synopsis or notable details about the movie.

3. Top Rated Music: Card Visualization

- Title: Name of the highest-rated music track or album.
- Rating: Maximum rating received by the music.
- Type: Music.
- Description: Brief description or notable details about the music track or album.

4. Top Rated Game: Card Visualization

- Title: Name of the highest-rated game.
- Rating: Maximum rating received by the game.
- Type: Game.

- **Description:** Brief synopsis or notable details about the game.

5. Top Rated Book: Card Visualization

- **Title:** Name of the highest-rated book.
- **Rating:** Maximum rating received by the book.
- **Type:** Book.
- **Description:** Brief synopsis or notable details about the book.

Selection Criteria:

- **Filtering:** The dataset is filtered to find the highest-rated item for each type.
- **Max Rating:** The item with the maximum rating within each category is selected as the top-rated item.
- **Specific Type:** Each card is focused on a specific type of content, ensuring clarity and relevance.

Top Rated Titles: Table Visualization

- The "Top Rated Titles" visualization presents a table that displays the top-rated titles viewed by users. The table includes filters to show only interactions of type "view" and ranks the titles by the highest ratings. The top five ratings with the titles across all genres are displayed, providing a detailed view of the most highly rated content in the dataset.

Top Rated Titles	
Title	Rating
The Paper Tigers	10.00
Integrated value-added system engine	9.90
King of Boys	9.90
Optional bottom-line Local Area Network	9.90
Robust tangible methodology	9.90
Show Dogs	9.90
The Presocratics	9.90
Vision-oriented 24/7 neural-net	9.90
Cross-group empowering ability	9.80

Why This Chart is Needed :

- The table visualization allows users to see a detailed list of the highest-rated titles based on user views. By presenting this information in a table format, it provides an easy-to-read and organized view of top-rated content. This helps

users quickly identify the most popular and highly rated items, aiding in content discovery and decision-making.

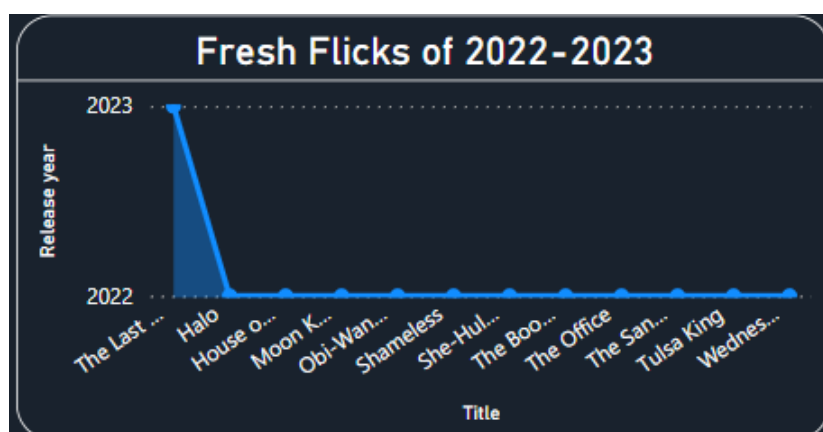
Details:

- **Filters:** Interaction Type (View), Top 5 by Max Rating, All Titles, All Types
- **Columns:** Titles, Ratings
- **Insights:** The table provides a clear and concise list of top-rated titles, helping users identify which items are most appreciated by the audience. This insight is useful for understanding user preferences and highlighting standout content.

Result: The number of titles with a rating of 10 is 1, which is "The Paper Tigers." There are 7 titles with a rating of 9.90, 2 titles with a rating of 9.80, 2 titles with a rating of 9.70, and 3 titles with a rating of 9.60.

Fresh Flicks of 2022-2023: Area Chart Visualization

- The "Fresh Flicks of 2022-2023" visualization is an area chart that displays all titles released between the years 2022 and 2023. The x-axis shows the titles, while the y-axis displays the release year, highlighting the distribution of new releases over this period.



Why This Graph is Needed :

- This area chart helps users identify new and recently released titles, focusing on content from the past two years. It provides insights into the latest trends

and popular new releases, which is valuable for users looking to discover current and relevant content.

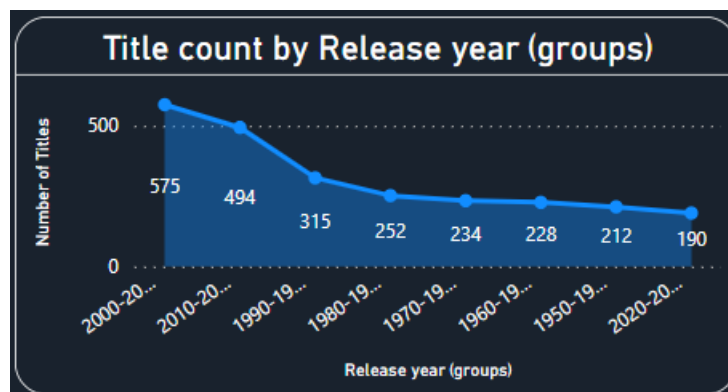
Details:

- **X-axis:** Titles
- **Y-axis:** Release Year
- **Insights:** The chart provides a visual representation of titles released in 2022 and 2023, helping users track the latest additions to the content library. This is particularly useful for users interested in exploring recent releases.

Result: The titles released in the specified period include The Last of Us (2023) , Halo (2022) , House of Cards (2022).

Title Count by Release Year (Groups): Stacked Area Chart Visualization

- The "Title Count by Release Year (Groups)" visualization is a stacked area chart that displays the count of titles released over different periods, divided into 10-year groups from 1950 to 2023. This chart helps visualize the distribution and trends of title releases over these grouped time periods.



Why This Chart is Needed :

- The stacked area chart allows users to understand the historical trends in title releases. By grouping the releases into 10-year intervals, it provides a clear view of how content production has evolved over time and highlights significant periods of activity.

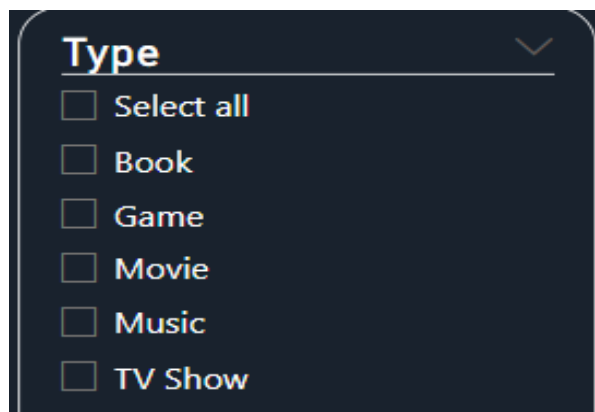
Details:

- **X-axis:** 10-year periods from 1950 to 2023
- **Y-axis:** Number of Titles
- **Legend:** Time Groups
- **Insights:** This visualization shows the distribution of titles released across different decades, helping users identify trends in content production over time.

Result: The highest number of titles, 575, were released during the period of 2000-2009, while the lowest number of titles, 190, were released during the period of 2020-2023.

Type: Slicer Visualization

- The slicer is an interactive filter that allows users to select a specific type of content (e.g., book, movie, TV show, game, music). By choosing a type in the slicer, all visualizations on the page update dynamically to reflect data related only to the selected type.



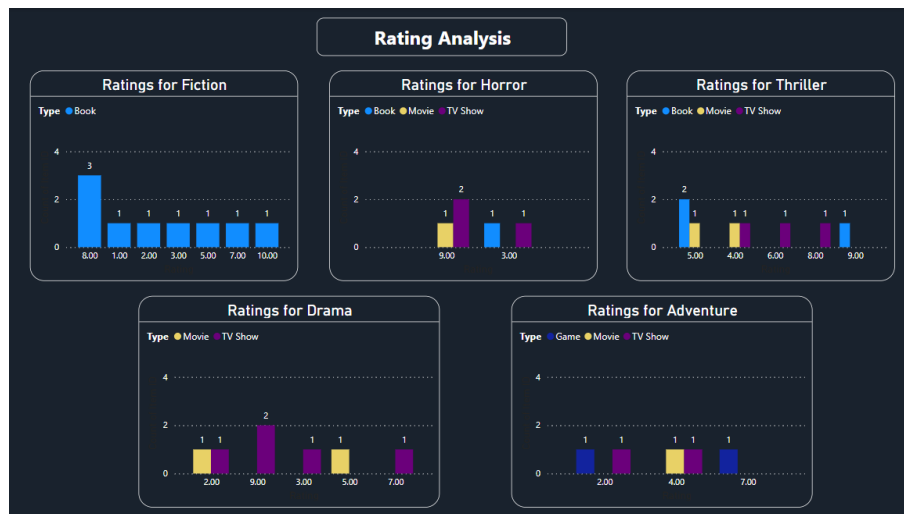
Why this is Needed:

- The slicer enhances the interactivity and user experience by providing customized views based on user preferences. It ensures that all visualizations on the page display relevant and specific information according to the user's selection.

Details:

- **Slicer Name:** Type
- **Options:** Book, Movie, TV Show, Game, Music

5.RATING ANALYSIS

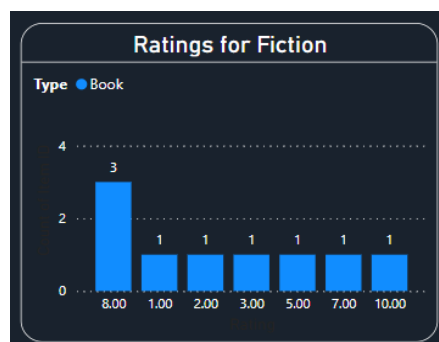


Why this page is needed:

- This page focuses on Rating Analysis, providing a detailed examination of ratings across the top 5 genres.
- Through chart type clustered column charts, it dives into user interactions and preferences within genres such as fiction, horror, thriller, drama, and adventure.
- The analysis offers valuable insights into trends and user engagement patterns, aiding in understanding audience preferences and content performance.

Ratings for Fiction

- This visualization analyses the distribution of ratings for fiction books using a clustered column chart. The chart was selected to clearly represent the count of items within each rating category, highlighting user preferences within the fiction genre.



Graph Type with Overview:

- **Graph Type:** Clustered Column Chart
- **Overview:** This visualization shows the ratings for the Fiction genre. The chart illustrates the count of items based on their ratings.

Details:

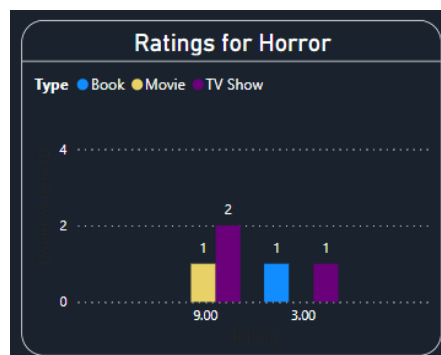
- **X-axis:** Rating
- **Y-axis:** Count of Items
- **Legend:** Type (All items are books in this genre)
- **Insights:** Fiction contains only book items, so the count of users with respect to ratings is shown.

Result of the Visualization:

- The result shows the distribution of user ratings, helping to determine the most and least favoured ratings within the fiction genre.
- The result shows that the rating of 8.00 has the highest count, with 3 books receiving this rating. All other ratings (1.00, 2.00, 3.00, 5.00, 7.00, 10.00) each have only 1 book.
- This indicates that 8.00 is a relatively common rating for fiction books in this dataset.

Ratings for Horror

- This visualization explores the ratings for horror items using a clustered column chart. The chart was chosen to illustrate the count of items for different types within the horror genre based on user ratings.



Graph Type with Overview:

- **Graph Type:** Clustered Column Chart
- **Overview:** This visualization shows the ratings for the Horror genre, displaying the count of items of different types.

Details:

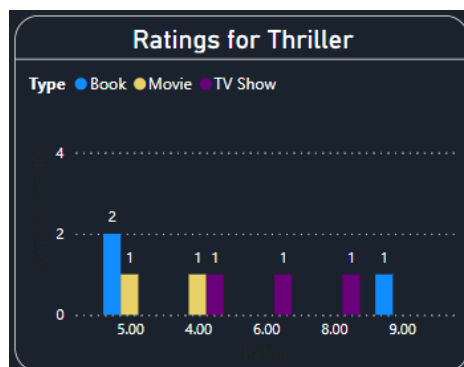
- **X-axis:** Rating
- **Y-axis:** Count of Items
- **Legend:** Type (Book, Movie, TV Show)
- **Insights:** The chart shows the ratings distribution for three different types of items in the Horror genre.

Result of the Visualization:

- The result shows the distribution of user ratings, helping to determine the most and least favored ratings within the horror genre.
- The rating of 9.00 is the most common for TV shows (2 counts), followed by movies (1 count).
- The rating of 3.00 has one book and one TV show.

Ratings for Thriller

- This visualization examines the ratings for thriller items using a clustered column chart. It was selected to show the count of items within the thriller genre across different rating categories.



Graph Type with Overview:

- **Graph Type:** Clustered Column Chart

- **Overview:** This visualization shows the ratings for the Thriller genre, displaying the count of items of different types.

Details:

- **X-axis:** Rating
- **Y-axis:** Count of Items
- **Legend:** Type (Book, Movie, TV Show)
- **Insights:** The chart displays the ratings distribution for books, movies, and TV shows in the Thriller genre.

Result of the Visualization:

- The result shows the distribution of user ratings, helping to determine the most and least favored ratings within the Thriller genre.
- The rating of 5.00 is the most common for books (2 counts), followed by movies (1 count).
- The rating of 4.00 has one book and one movie.
- The ratings of 6.00 and 8.00 are exclusively given to TV shows (1 count each).
- The rating of 9.00 is given to one book.

Ratings for Drama

- This visualization investigates the ratings for drama items using a clustered column chart. It was chosen to represent the count of items for different types within the drama genre based on user ratings.



Graph Type with Overview:

- **Graph Type:** Clustered Column Chart

- **Overview:** This visualization shows the ratings for the Drama genre, displaying the count of items of different types.

Details:

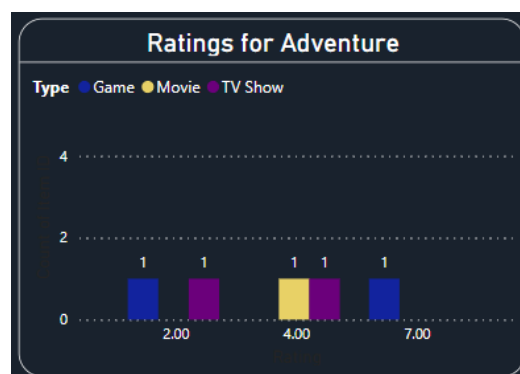
- **X-axis:** Rating
- **Y-axis:** Count of Items
- **Legend:** Type (Movie, TV Show)
- **Insights:** The chart shows the ratings distribution for movies and TV shows in the Drama genre.

Result of the Visualization:

- The result shows the distribution of user ratings, helping to determine the most and least favoured ratings within the Drama genre.
- The rating of 9.00 is the most common for TV shows (2 counts).
- The rating of 2.00 has one movie and one TV show.
- The ratings of 3.00 and 7.00 are exclusively given to TV shows (1 count each).
- The rating of 5.00 is given to one movie.

Ratings for Adventure

- This visualization studies the ratings for adventure items using a clustered column chart. The chart was selected to depict the count of items within the adventure genre across different rating categories.



Graph Type with Overview:

- **Graph Type:** Clustered Column Chart

- **Overview:** This visualization shows the ratings for the Adventure genre, displaying the count of items of different types.

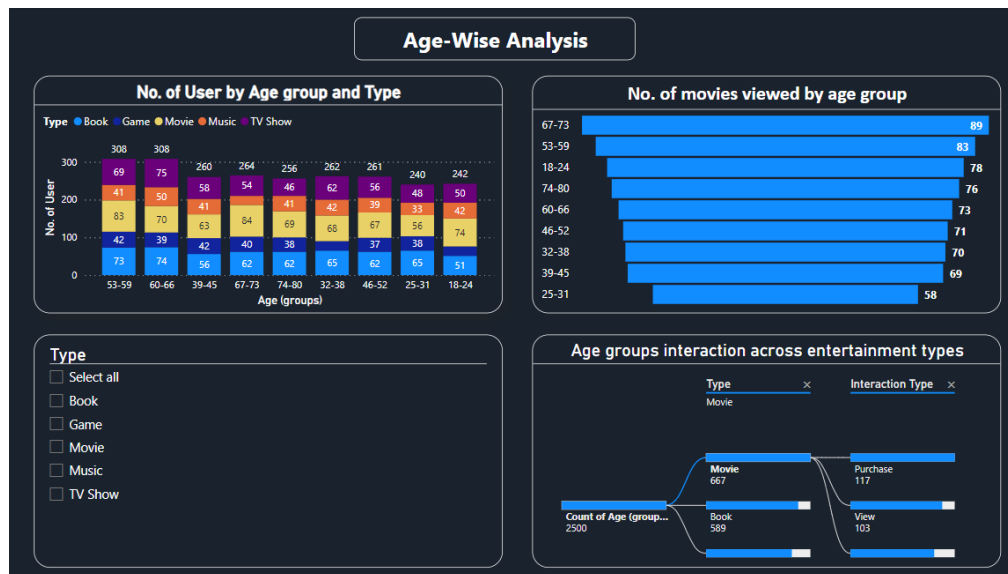
Details:

- X-axis: Rating
- Y-axis: Count of Items
- Legend: Type (Game, Movie, TV Show)
- Insights: The chart illustrates the ratings distribution for games, movies, and TV shows in the Adventure genre.

Result of the Visualization:

- The result shows the distribution of user ratings, helping to determine the most and least favoured ratings within the Adventure genre.
- There is one Game and one TV show with a rating of 2.00, one Movie and one TV show with a rating of 4.00, and one Game with a rating of 7.00.

6. Age-Wise Analysis

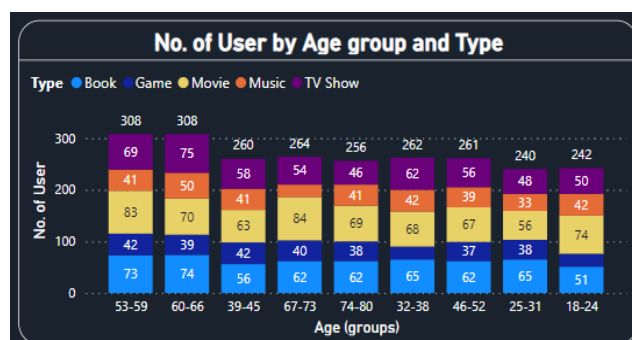


Why this page is needed:

- This page in the dashboard describes the important relationships between the age entity and other dimensions in the dataset.
- Ages are divided into nine groups with equal width, ranging from 18 to 80 years old. The groups are 18-24, 25-31, 32-38, 39-45, 46-52, 53-59, 60-66, 67-73, and 74-80.
- These groups help identify which user groups are more active, how interactions with different types of items occur, and which groups are least active. All age-related important analyses can be done using these groups.

Number of Users by Age Group and Type

- This visualization analyzes user demographics and engagement using a stacked column chart. It was selected to show the count of users for different item types within each age group.



Why this Chart is needed:

- This chart is needed to understand user demographics and preferences across age groups. The result shows which age groups are most and least active, helping to tailor content and marketing strategies accordingly.

Graph Type with Overview:

- **Graph Type:** Stacked Column Chart
- **Overview:** This visualization shows the total number of users belonging to each age group and the individual total count for each type of item.

Details:

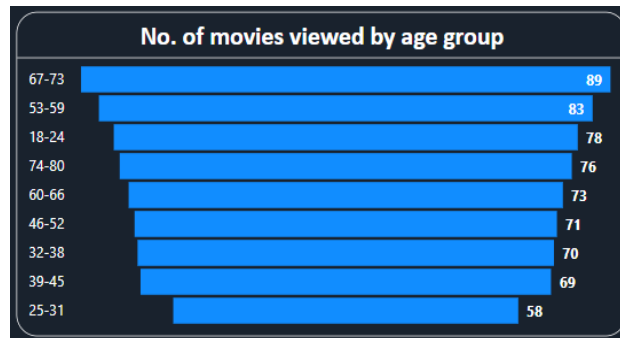
- **X-axis:** Age Groups
- **Y-axis:** Count of Users
- **Legend:** Type (Different types of items)
- **Insights:** The highest number of users (308) is in the age groups 53-59 and 60-66, while the least number of users (240) is in the age group 25-31.

Result of the Visualization:

- This chart helps in understanding the distribution of users across different age groups and their interactions with various item types.
- The highest number of users from the age group 53-59 interacted most with movies (83 users), while the age group 60-66 interacted most with TV shows (75 users).

Number of Movies Viewed by Age Group

- This visualization explores movie viewing patterns across age groups using a funnel chart. It was selected to show the count of movies viewed within each age group.



Why this Chart is needed:

- This chart helps to understand movie viewing behavior and preferences among different age demographics. It highlights trends in movie viewership, aiding in content recommendations and user engagement strategies.

Graph Type with Overview:

- **Graph Type:** Funnel Chart
- **Overview:** This visualization shows how many movies each age group has viewed, helping to identify the age groups with the highest and lowest movie viewership.

Details:

- **X-axis:** Age Groups
- **Y-axis:** Number of Movies
- **Insights:** This chart highlights the age groups that are most engaged in watching movies.

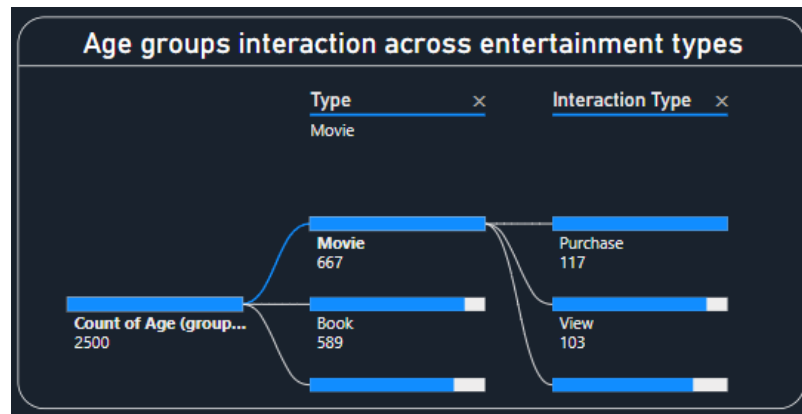
Filter Applied:

- **Type:** Movie
- **Interaction Type:** View

Result of the Visualization: The age group 67-73 has the highest movie viewership (89), while the age group 25-31 has the lowest(58).

Age Groups' Interaction across Entertainment Types

- This visualization examines user interactions across entertainment types using a decomposition tree. It was selected to provide a comprehensive view of user behaviour and preferences across different demographics.



Why this Chart is needed:

- This chart provides insights into how different age groups interact with various types of content. It helps identify trends in user behaviour, informing content strategy and user engagement efforts.

Graph Type with Overview:

- **Graph Type:** Decomposition Tree
- **Overview:** This visualization shows the count of users in each age group and their interactions with different types of items, broken down by type and interaction type.

Details:

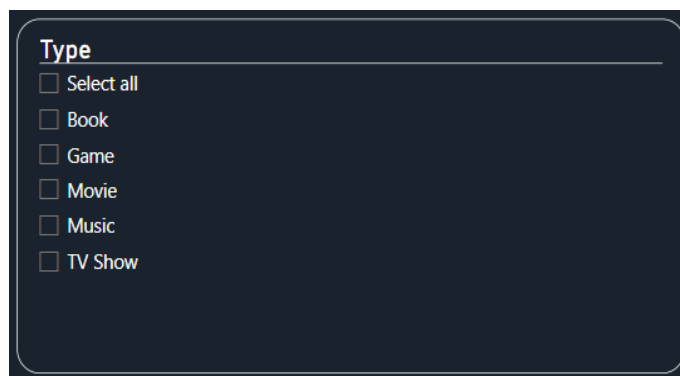
- **X-axis:** Age Groups
- **Y-axis:** Count of Users
- **Legend:** Type, Interaction Type
- **Insights:** This chart provides a detailed analysis of how users in different age groups interact with various types of entertainment items.

Result of the Visualization:

- The highest number of users is for the type "Movie" (667) and the least number of users is for "Game"(335).
- For example, among 2500 users, 667 interacted with movies: 117 purchased, 103 viewed, 94 disliked, etc.

Type Selector (Slider)

- This visualization enhances dashboard usability with a selector for focused analysis. It allows users to analyse specific user types, such as book readers or game enthusiasts.



Why this is needed:

- This visual is needed to provide users with a tool for detailed analysis, making the dashboard more interactive and user-friendly. It helps in understanding specific user behaviours and preferences for different content types.

Visual Type with Overview:

- **Visual Type:** Selector
- **Overview:** This special visualization allows users to analyze data for a specific type of item. For example, selecting "Book" shows data only for book users, "Game" for game users, and so on.
- **Insights:** This selector helps in focusing the analysis on a particular type of item, providing a more granular view of user interactions.

Filter Applied:

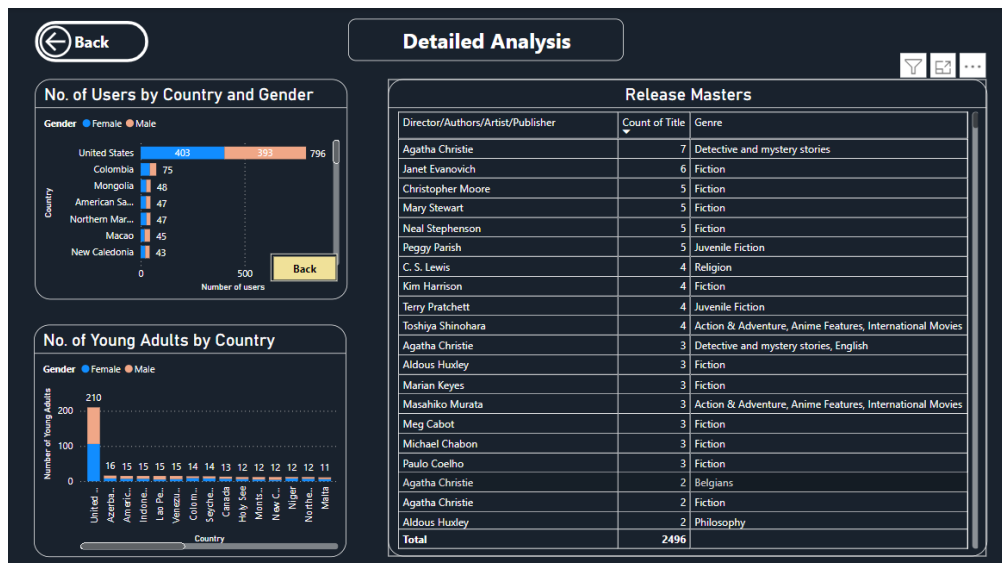
- User-selected item type

Selector: Enables focused analysis on specific item types, improving usability and flexibility.

Functionality: Allows users to filter and view data specific to a chosen item type.

Result of the Feature: This feature enhances the flexibility of the dashboard, allowing for targeted analysis of user interactions with specific item types.

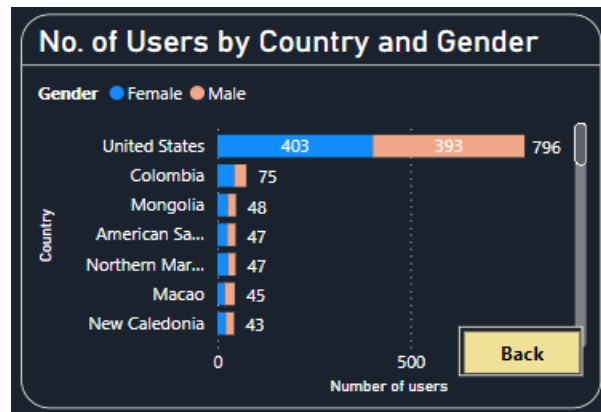
7. Detailed Analysis:



Why Detailed Analysis is needed:

- The Detailed Analysis Dashboard offers valuable insights into user demographics and the content preferences of users.
- It includes detailed visualizations and tables to help analysts understand user distribution by country and gender, as well as an inventory of content releases by various creators.
- This analysis is crucial for understanding the audience distribution and tailoring content recommendations to suit user preferences better.
- The page consists of three main sections: the number of users by country and gender, the number of young adults by country, and a detailed table of release masters.
- Each section presents data that is vital for making informed decisions regarding content preference, marketing strategies, and user engagement.

No. of Users by Country and Gender:



Graph Type: Stacked Bar Chart

- This chart presents the distribution of users by country and gender. It helps in understanding the geographic and gender demographics of the user base.

Why this Chart is Needed :

- This type of chart allows for straightforward comparison between countries as well as between genders within each country.
- It helps in identifying the geographic distribution of the user base and understanding the gender dynamics within different regions.
- It is essential for developing targeted marketing campaigns and improving user experience by recognizing regional preferences and gender-specific content needs.

Details:

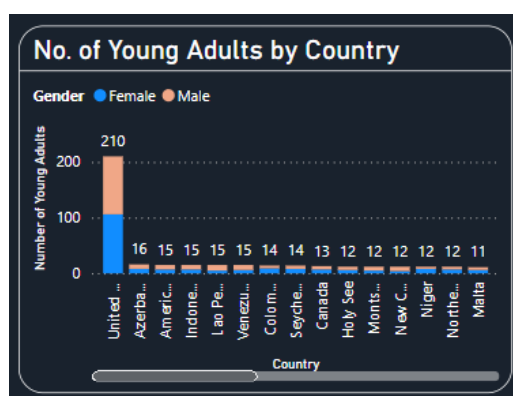
- **X-axis:** Number of users
- **Y-axis:** Country
- **Legend:** Gender (Female in blue, Male in orange)
- **Insights:** The chart provides a clear visual representation of gender distribution within each country.
 - The United States has the highest number of users, with 403 females and 393 males, totaling 796 users.

- Other countries like Colombia, Mongolia, and American Samoa have significantly fewer users.

Filter Applied:

- Clicking on a segment, such as the female users in the United States, will highlight the specific segment within the chart.
- For instance, clicking on the blue part representing females will show the 403 females out of the total 796 users in the USA.

No. of Young Adults by Country:



Graph Type: Stacked Column Chart

- This chart displays the number of young adults by country, broken down by gender. It is useful for understanding the younger demographic within the user base.

Why This Graph is needed:

- This visualization is crucial for understanding the young adult demographic, which is a key audience for entertainment recommendations.
- Knowing the distribution helps in curating content that appeals to this age group, thereby enhancing user engagement and retention.

Details:

- **X-axis:** Country
- **Y-axis:** Number of Young Adults
- **Legend:** Gender (Female in blue, Male in orange)
- **Insights:** This chart highlights the concentration of young adult users within each country.
 - The United States again leads with the highest number of young adults, with a significant skew towards females (210 females vs. 15 males).

- Other countries show a more balanced or different distribution but with much smaller numbers.

Filter Applied:

- Selecting a segment, such as young males in the United States, filters the "Release Masters" table accordingly.
- For example, selecting the orange part representing 104 young males out of 210 total young adults in the USA filters the release data to show titles based on the preferences and data of young male users from the selected country.

Release Masters Table:

Release Masters		
Director/Authors/Artist/Publisher	Count of Title	Genre
Agatha Christie	7	Detective and mystery stories
Janet Evanovich	6	Fiction
Christopher Moore	5	Fiction
Mary Stewart	5	Fiction
Neal Stephenson	5	Fiction
Peggy Parish	5	Juvenile Fiction
C. S. Lewis	4	Religion
Kim Harrison	4	Fiction
Terry Pratchett	4	Juvenile Fiction
Toshiya Shinohara	4	Action & Adventure, Anime Features, International Movies
Agatha Christie	3	Detective and mystery stories, English
Aldous Huxley	3	Fiction
Marian Keyes	3	Fiction
Masahiko Murata	3	Action & Adventure, Anime Features, International Movies
Meg Cabot	3	Fiction
Michael Chabon	3	Fiction
Paulo Coelho	3	Fiction
Agatha Christie	2	Belgians
Agatha Christie	2	Fiction
Aldous Huxley	2	Philosophy
Total	2496	

Visual Type: Table

- This table provides a detailed inventory of content available, categorized by the creator and genre.

Why this Table is Needed:

- Tables allow users to sort data by different columns (e.g., count of titles, genre), making it easy to focus on specific aspects of the data.
- The "Release Masters" table lists directors, authors, artists, and publishers along with the count of titles they have released and the genres of those titles.
- This table is essential for content management team to understand which creators are popular and which genres are in demand.
- It helps in making data-driven decisions for content preference, ensuring a diverse and appealing content library for users.

Details:

- **Columns:**

- **Director/Authors/Artist/Publisher:** The name of the content creator.
- **Count of Title:** Number of titles released by the creator.
- **Genre:** Genre of the titles.

- **Insights:**

- The table shows a diverse range of genres, indicating a wide variety of content available to users.
- Agatha Christie is a prominent figure with the highest count of titles (7) in the "Detective and mystery stories" genre.
- Other notable creators include Janet Evanovich, Christopher Moore, and Mary Stewart, each with multiple titles primarily in fiction.
- The total count of titles listed is 2496, reflecting a substantial library.