Lab 12. Analysis of Amazon Prime Dataset

i) Create a Donut chart to show the percentage of movie and tv shows.
ii) Create an area chart to shows by release year and type.
iii) Create a horizontal bar chart to show Top 10 genre.
iv) Create a map to display total shows by country.
v) Create a text sheet to show the description of any movie/movies.
vi) Build an interactive Dashboard.

1. Introduction to the Program.

- **Amazon Prime**, a popular streaming platform, offers a diverse range of movies, TV shows, and original content. This analysis demonstrates how visualizations can help uncover trends, provide insights into content distribution, and aid in decision-making processes. In this program, we analyze an **Amazon Prime Dataset** using Power BI, focusing on visualizing and understanding the patterns in the data.
- The dataset represents the details of movies and TV shows available on Amazon Prime. It includes key information such as titles, types (movie or TV show), genres, release years, countries, and descriptions. Analyzing such a dataset can provide actionable insights for content developers, marketers, and platform managers to understand content preferences and improve user engagement.

2. Introduction to the Dataset

The dataset is a simulated representation of Amazon Prime's content library. It consists of 50 records, each detailing attributes of a specific movie or TV show.

Key fields in the dataset include:

- **Title**: The name of the movie or TV show.
- **Type**: Specifies whether it is a movie or a TV show.
- **Genre**: The primary genre (Theme) of the content (e.g., Drama, Comedy, Action, Thriller…).
- **Release Year**: The year in which the movie or TV show was released.
- **Country**: The primary country associated with the production of the content.
- **Description**: A brief synopsis or description of the content.

3. Objectives of the Program

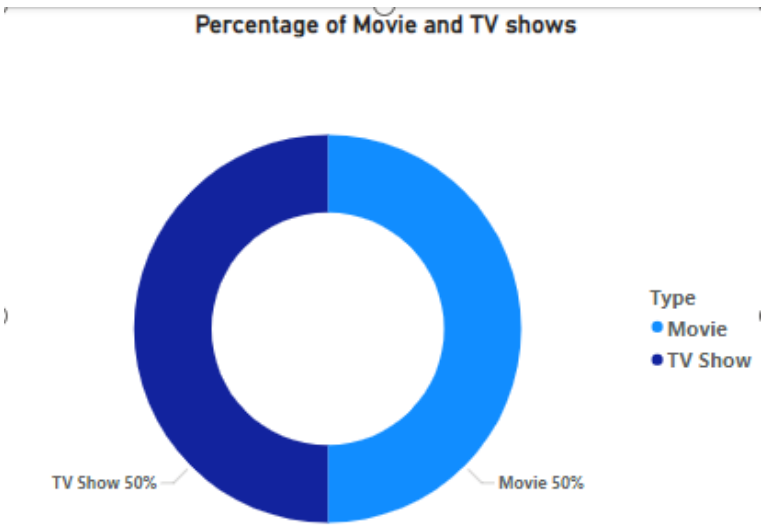
The objectives of this program are as follow-

1. To understand how to analyze a structured dataset in Power BI and create meaningful insights.
2. To visualize the distribution of movies and TV shows using a **Donut Chart**.
3. To explore the trends in content release over time using an **Area Chart** categorized by type.
4. To identify the most popular genres by creating a **Horizontal Bar Chart** for the top 10 genres.
5. To map the geographic spread of content availability using a **Map Visualization**.
6. To display detailed textual information for specific movies or TV shows using a **Text Sheet**.
7. To combine multiple visualizations into a cohesive **Interactive Dashboard** for comprehensive analysis.

i) Create a Donut chart to show the percentage of movie and TV shows.

1. **Create Donut Chart**:
    - Go to the **Report View**, Select the **Donut Chart** visual from the Visualizations pane.
    - Drag **`Type`** to **Legend**.
    - Drag **Title** with the count of `Title` to **Values**.

2. **Format the Chart**:
    - Customize colors for `Movie` and `TV Show`.
    - Display percentage in labels by selecting Format-> Visuals->Label Contents->Select **Category, Percentage of total**



ii) Create an Area Chart to shows by release year and type.

An **Area Chart** is used to show trends over time, highlighting the magnitude of changes across categories. In this case, we will create an Area Chart to display the count of Movies and TV Shows for each release year, categorized by type.

1. **Select the Area Chart** from the Visualizations pane.
2. Drag Release_Year to **X-Axis**,

   This will set the x-axis of the chart, showing the timeline of release years.

3. Drag Type to **Legend**:

- Drag the Type column (which categorizes content as either Movie or TV Show) to the **Legend** field in the Visualizations pane.
- This will create separate areas in the chart for each type, allowing us to compare their trends over time.

4. Drag Count to **Values** (ensure it aggregates properly). Drag the Title column to the Y-Axis Values field.

   To show the count of titles (Movies and TV Shows) for each year:

   - Drag the Title column to the Y-Axis **Values** field.
   - Power BI will automatically aggregate the values. By default, it will calculate the **Count** of titles.

   Ensure the aggregation is correct:

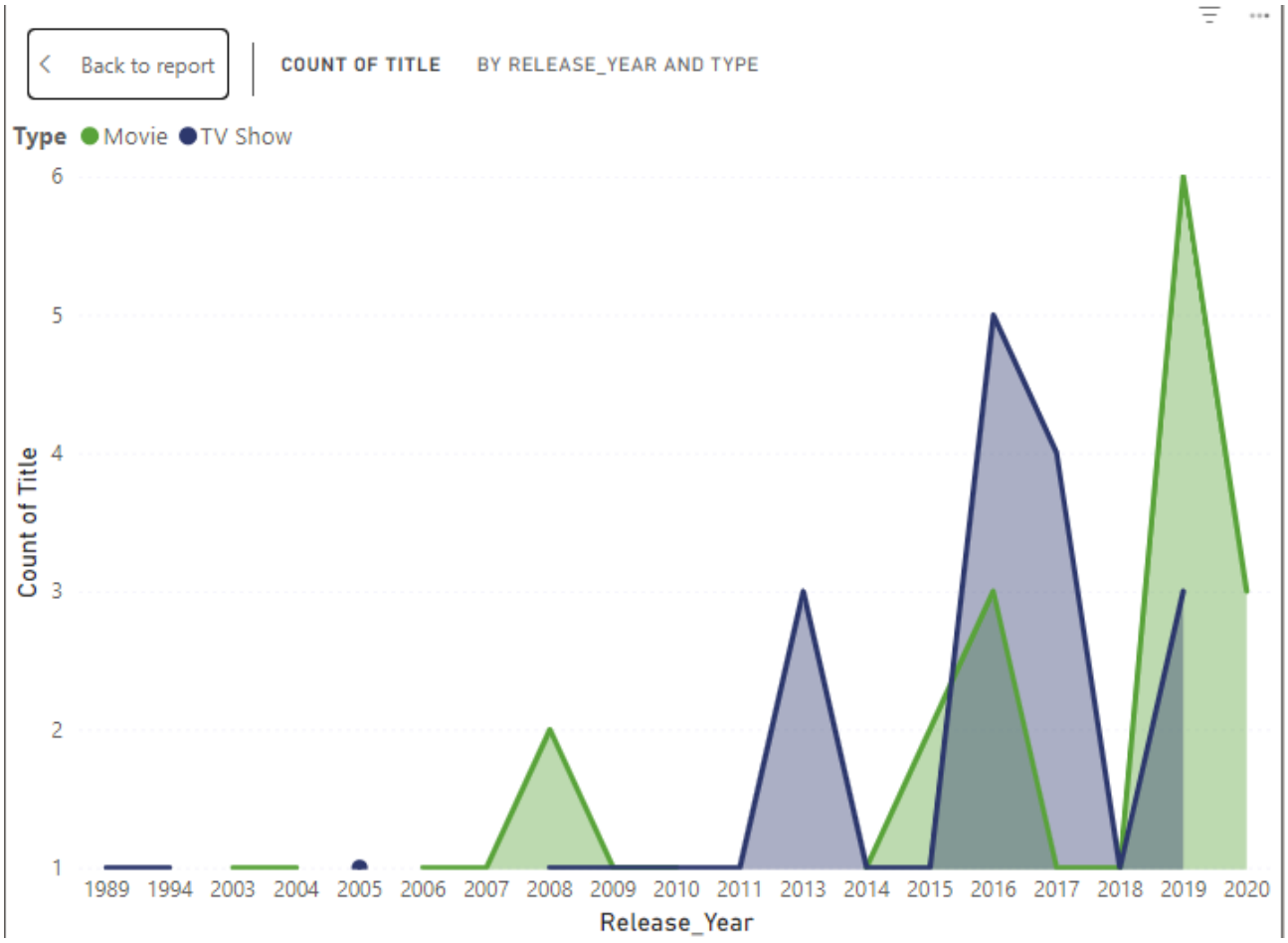
   - Click the dropdown next to the field in **Values**.
   - Select **Count** or **Count (Distinct)** to avoid summing numerical values unnecessarily.

5. Adjust the X-axis to display all years.

By default, Power BI may group or compress the x-axis if there are too many years.
To fix this:

- Select the chart.
- Go to the **Format pane** (paint roller icon).

- Expand the **X-axis** section, Set **Type** to **Categorical** (if it is Continuous by default).
- Ensure all release years are displayed in order. Add data labels for clarity.



iii) Create a horizontal bar chart to show Top 10 genre.

Step 1: Select the Horizontal Bar Chart

In the **Visualizations pane**, select the **Horizontal Bar Chart** icon. This icon looks like a bar chart with horizontal bars. This will add a blank bar chart to your report canvas.

Step 2: Prepare the Dataset for the Bar Chart

We need to display the **Genres** and the number of times each genre appears. The data should include a **Genre** field and a way to count the number of items (e.g., movies or TV shows) in each genre.

1. **Y-axis**: In a **horizontal bar chart**, the **Y-axis** represents the categories (in this case, the **genres**), and each genre will have a corresponding horizontal bar.

2. **X-axis**: The **X-axis** will represent the values associated with each category (in this case, the count of titles or movies/TV shows in each genre), which will determine the length of the bars.

Note: We can also add a measure that shows the **count of Titles** or **the number of shows in each genre**. This will serve as the bar's length (how many movies/TV shows exist in each genre).

Step 3: Add a Measure for Count

Create a measure that counts the number of shows or movies in each genre.

To create the count:

1. select **New Measure**.
2. Enter this DAX formula:

```
Count of Titles = COUNTROWS('AP')
```

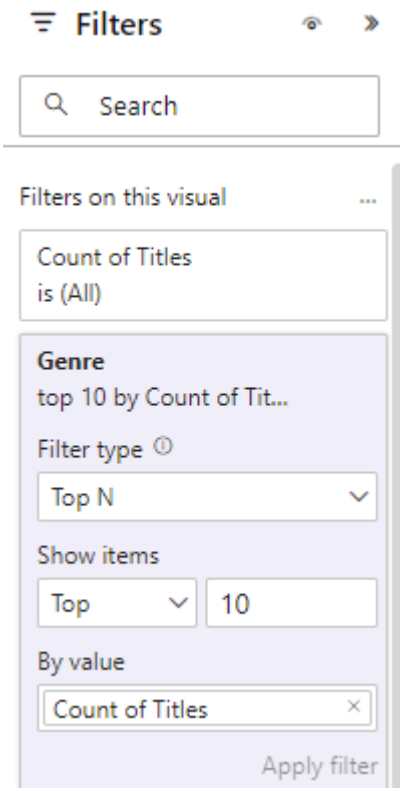
This formula counts the number of rows (shows/movies) in your dataset.

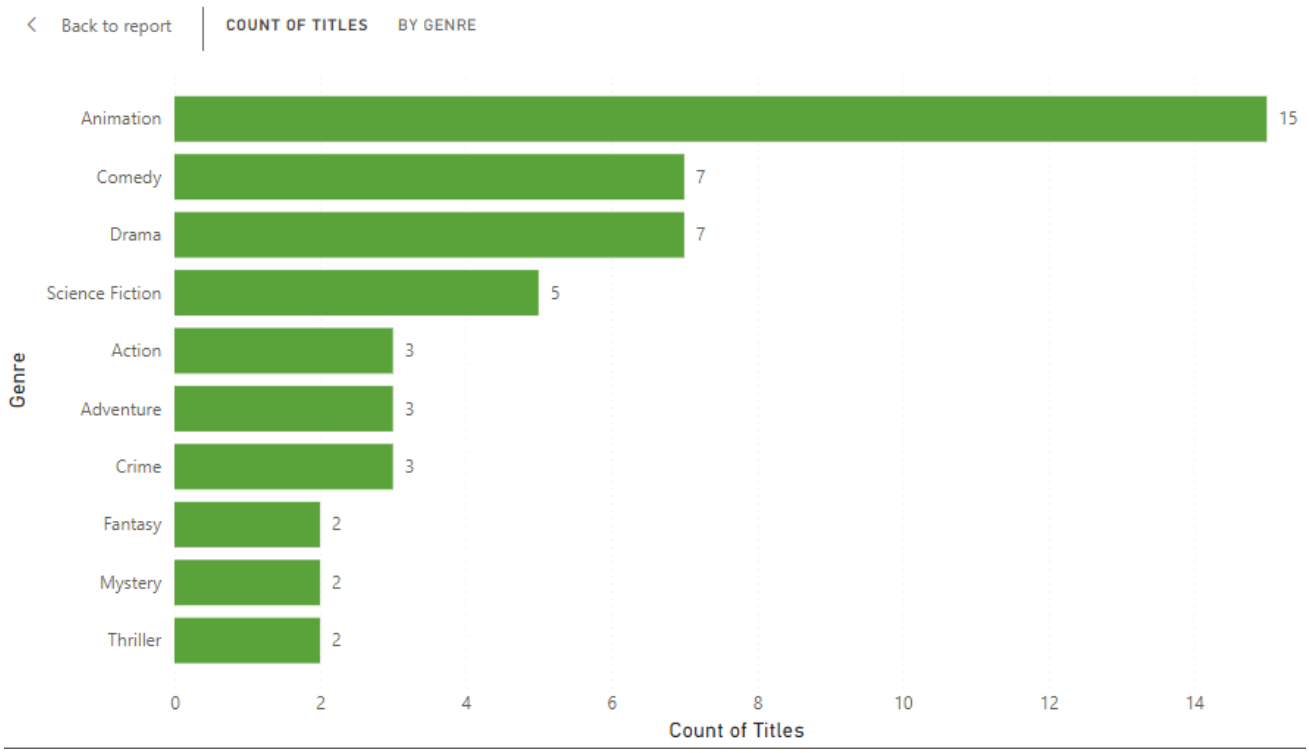
3. Now, drag this new measure (Count of Titles) to the **Values** field well of the chart. This will represent the count of titles for each genre and will determine the length of each bar.

Step 4: Sort and Filter the Top 10 Genres

To display only the **Top 10 Genres**, we need to apply a filter.

1. **Go to the Filters pane** (on the right side of the screen).
2. Find the **Genre** field under **Filters on this visual**.
3. Click on the **Genre** field dropdown and select **Top N**.
4. In the **Top N filter** settings:
    - Select **Top**.
    - Enter **10** as the number of items.
    - For the **By value**, drag the Count of Titles measure into the **By value** field.
    - Set the **Sort by** to **Count of Titles** measure to ensure that the genres with the highest counts are shown first.


Filters
Search
Filters on this visual
Count of Titles
is (All)
Genre
top 10 by Count of Tit...
Filter type
Top N
Show items
Top
10
By value
Count of Titles
Apply filter


Back to report
COUNT OF TITLES
BY GENRE
Genre
Animation
Comedy
Drama
Science Fiction
Action
Adventure
Crime
Fantasy
Mystery
Thriller
15
7
7
5
3
3
3
2
2
2
0
2
4
6
8
10
12
14
Count of Titles

iv) Create a map to display total shows by country.

In Power BI, the **Map visual** allows us to plot data on a geographical map, where different areas (such as countries, regions, or cities) are represented based on data points.

Step 1: Select the Map Visual 

In the **Visualizations pane**, find and click on the **Map** visual icon. The **Map visual** is represented by a globe icon with a location pin.

Step 2: Assign Data to the Map

For this map to display the total number of shows by country, we need to provide data fields for **Location** (to specify where the shows are located) and **Values** (to represent the total number of shows in each location).

2.1. Drag the Country Field to Location

- **Drag the Country field** and drop it into the **Location** well under the **Map** visual in the **Visualizations** pane. The location field defines the geographical points on the map. Since we are mapping by country, **Country** should be placed here.
- **Power BI automatically recognizes** common geographical names like countries, states, or cities, and maps them accordingly.

2.2. ***Drag the Count Column to Bubble Size***

Next, we need to display the data on the map by using the **Size** fields.

- **Size**: The **Size** field determines the size of the markers (dots or circles) displayed on the map. Larger markers will represent higher values.
- Drag the **Count of Titles** measure into the **Bubble Size** field. This will make the size of the markers on the map correspond to the number of shows in each country.

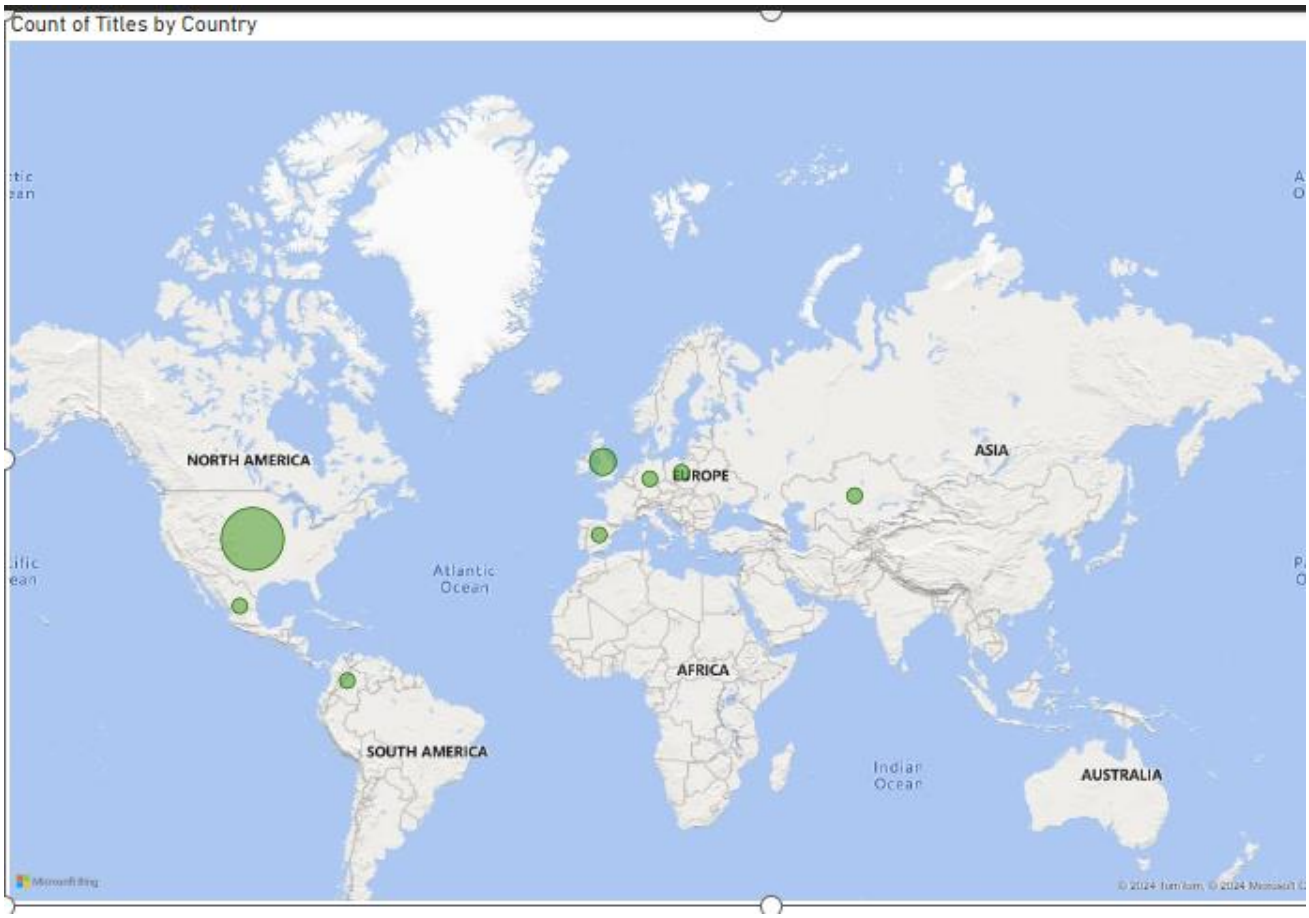
Step 3: Review and Interpret the Map

After following these steps, your map will show the **total number of shows by country** using either varying marker sizes.

- **Larger markers** represent countries with more shows.
- **Smaller markers** represent countries with fewer shows.

The map will provide a geographical distribution of the shows, making it easy to spot which countries have the most shows and which ones have fewer.

Map Visual



v) Create a text sheet to show the description of any movie/movies.

Text Sheet in Power BI

Is a **Table** or **Card Visual** that allows you to display textual data, such as the description of movies in this case.

Slicer in Power BI

A **Slicer** in Power BI is a visual tool used to filter data interactively. When applied to a text sheet or other visuals, a slicer allows users to select specific values to dynamically filter and display related information in the associated visuals. This enhances interactivity and usability by enabling focused exploration of data.

Steps to Create a Text Sheet

1. Add a Table or Card Visual

- In the **Visualizations pane**, select either:
    - **Table**: For a list of movie descriptions.
    - **Slicer** Users can drill down into specific details for a single movie description (useful for displaying description of one movie at a time).

2. Drag the Required Fields

- Drag the **Title** column from your dataset into the visual. This will display the names of the movies.
- Drag the **Description** column into the same visual. This will show the corresponding descriptions for the movies.

3. Filter to Show Specific Movies

- To display a specific movie or group of movies:

    - Add a **Slicer** from the **Visualizations pane**.
    - Drag the **Title** column into the slicer.
    - This allows users to select one or multiple movies, and the table or card visual will update dynamically to show only the descriptions for the selected movies (Hold Control key for multiple selection on the slicer).


Table: For a list of movie descriptions.
Title
Description
The Marvelous Mrs. Maisel
A 1950s housewife turned comedian.
Coco
A boy's journey with music and family.
Sherlock
A brilliant detective in modern times.
Breaking Bad
A chemistry teacher turned meth maker.
Mindhunter
A chilling look at a killer's mind.
Jack Ryan
A CIA analyst on a global mission.
Avengers: Endgame
A climactic superhero battle.
Seinfeld
A comedy about nothing and everything.
Narcos
A crime drama about drug cartels.
Fleabag
A dark comedy about a young woman.
Dark
A dark, mysterious sci fi drama.
The Irishman
A dramatic saga of mob families.
Sound of Metal
A drummer copes with hearing loss.
The Incredibles
A family of superheroes saves the day.
The Witcher
A fantasy epic of a monster hunter.
Money Heist
A heist drama with intricate plots.
The Crown
A historical drama about British royalty.
The Mandalorian
A journey through a galaxy far away.
The Lion King
A lion cubâ€™s coming-of-age story.
Rick and Morty
A mad scientist and his grandson.
Roma
A Mexican domestic worker's story.
Inception
A mind-bending tale of dreams.
The Grand Tour
A motoring show with a global twist.
Soul
A musician finds his purpose.
Moana
A Pacific islander finds her destiny.
House of Cards
A political drama of power and intrigue.
Parks and Recreation
A quirky look at small-town life.
Ratatouille
A rat dreams of becoming a chef.
The Big Sick
A romantic comedy based on a true story.
Borat
A satirical mockumentary.
Finding Dory
A story of friendship under the sea.
Manchester by the Sea
A story of grief and family bonds.
Up
A tale of adventure in the skies.
Slicer for Drilling Down
Title
Avatar
Avengers: Endgame
BoJack Horseman
Borat
Breaking Bad
Brooklyn Nine-Nine
Coco
Dark
Finding Dory
Fleabag
Friends
Frozen II
Game of Thrones
House of Cards
Inception
Inside Out
Interstellar
Jack Ryan
Manchester by the Sea
Mindhunter

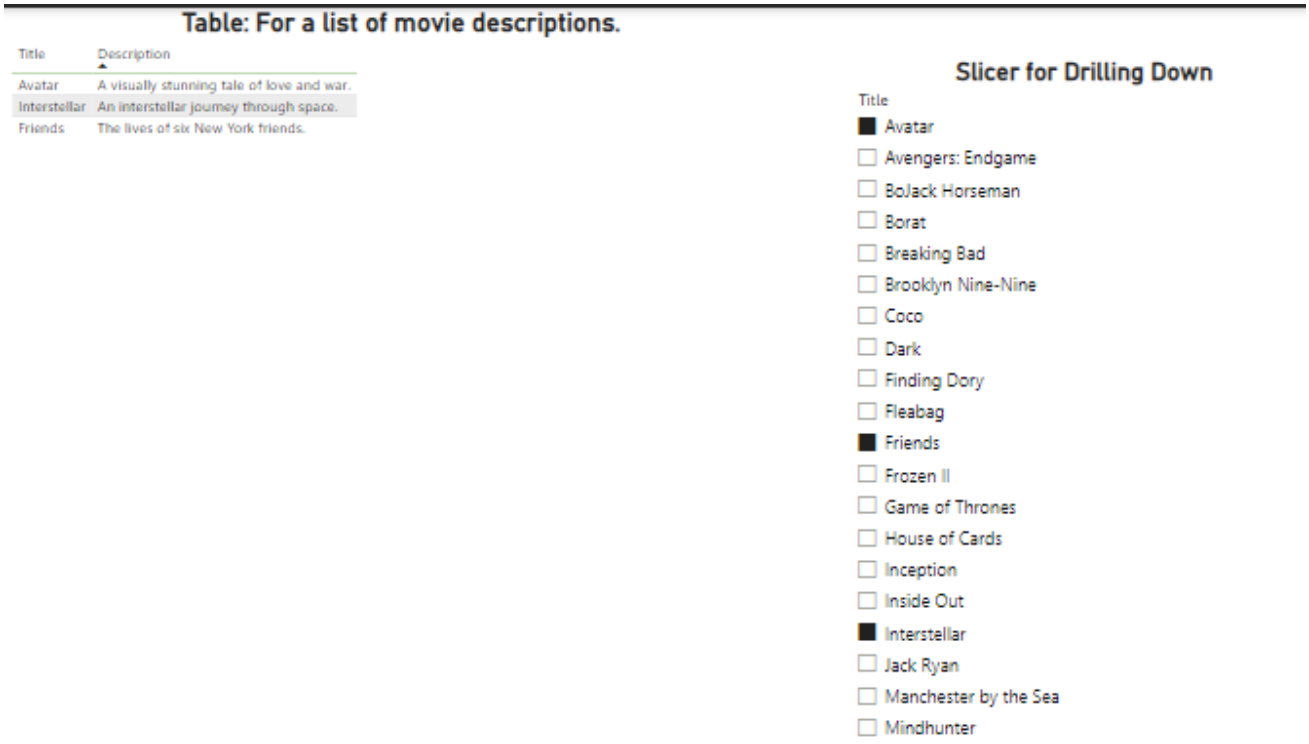


Table: For a list of movie descriptions.
Title
Description
Avatar
A visually stunning tale of love and war.
Interstellar
An interstellar journey through space.
Friends
The lives of six New York friends.
Slicer for Drilling Down
Title
Avatar
Avengers: Endgame
BoJack Horseman
Borat
Breaking Bad
Brooklyn Nine-Nine
Coco
Dark
Finding Dory
Fleabag
Friends
Frozen II
Game of Thrones
House of Cards
Inception
Inside Out
Interstellar
Jack Ryan
Manchester by the Sea
Mindhunter