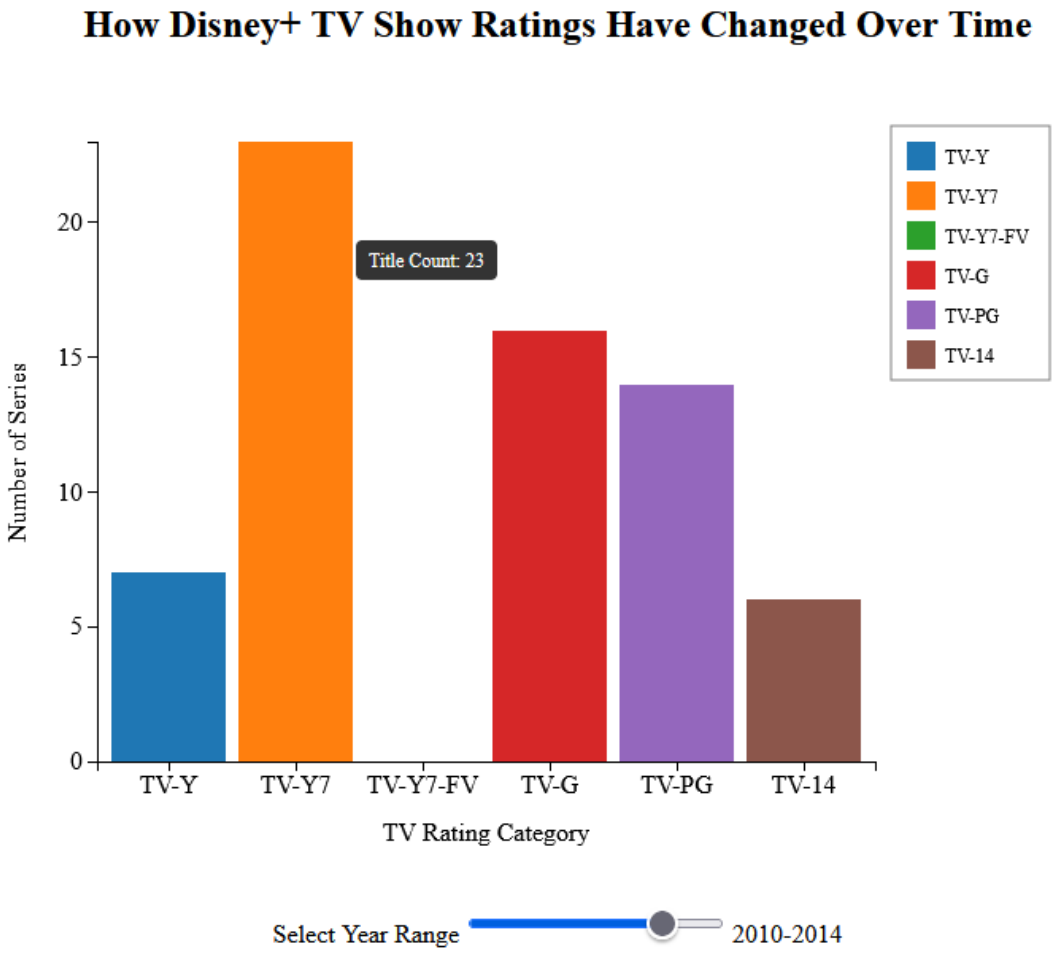
Rationale On Disney+ Data Visualization

Dataset Description:

The dataset used for this project is the Disney+ Movies and TV Shows dataset created by Shivam Bansal on Kaggle. It contains metadata for titles available on Disney+, including both movies and TV shows. Each entry includes information such as title, genre, duration, cast, and ratings. For this visualization, I focused specifically on TV shows and their age ratings across the years, the variables used to create the chart are release year, age rating, and a derived variable representing the total count of titles within each category.

The original plan for the visualization was to use a line chart or stacked bar plot instead of the side-by-side bar plot below. However, I realized that the best way to visualize the distribution in a clear and understandable way was to keep each rating in its own bin and avoid any designs that could be misunderstood.



Design Choices & Process:

I explored the dataset to refine my question and ensure it could be illustrated clearly in a visualization. I chose to focus on TV shows, because I noticed that there was much more titles and ratings classified under TV compared to movies. The second variable selected for the chart was rating, which contains the age ratings for all titles. Rating is an ordinal categorical variable with clearly defined categories such as TV-PG, TV-G, and TV-Y7. Additionally, to quantify the number of series produced, I created a third variable to count the total number of TV shows within each age rating. To compare age ratings over time, release year was introduced as the forth variable and grouped into five-year intervals to avoid clutter.

In order to visualize the distribution of TV shows across rating categories, I used a bar chart that clearly displays the counts within each five-year interval. The bars were ordered from the youngest to the oldest age rating, and this order remained consistent across all intervals. Additionally, each rating category was assigned a distinct color to clearly visualize the data, and a legend was placed on the right side of the chart with their associated colors.

After creating the bar chart, a slider was placed under the graph to allow users to switch between intervals. This interactive feature enables viewers to explore the data across different years without overwhelming them with multiple charts. In addition, a tooltip was incorporated to give users the ability to hover over a bar and see the exact number of TV shows in each rating category, providing precise information without cluttering the visualization.

Findings:

Using the interactive chart, I was able to examine how Disney+ TV show ratings shifted over time. In the first five intervals, from 1970 to 1994, very few TV shows were produced, and they mainly focused on a younger audience. By the early 2010s (2010-2014), production increased significantly, displaying a total count of 23 shows rated TV-Y7. In the following years, there was a clear rise in the number of shows aimed at older audiences, such as TV-PG and TV-14. This trend suggests that Disney expanded their target audience over time. Overall, the visualization can help viewers explore the distribution of age ratings across every five-year interval for Disney+ TV series.