Juan Barrera  
  
Professor Jeffrey Seaman  
  
09/03/2025  
  
Data Analysis and Visualization  
  
Module 1 Summary  
  
  
  
This week I worked with my classmate, Toby Wright, and together we analyzed a dataset we found on Kaggle about movies and TV shows. The dataset was titled Unveiling Modern Cinema: 100 Must-See Films of the 21st Century. It came as a text file containing detailed information about the top 100 films, including the title, director, year of release, score, genre, cast, duration, country, budget, sales, number of Oscars, and availability on platforms like Apple TV, Amazon Prime, HBO Max, and Netflix.

For our analysis, we focused on organizing and interpreting some of these categories, especially genre and streaming availability. Using a simple code, we looped through the dataset, counted the occurrences of different genres, and formatted the results so that they could be displayed on an HTML page.

To represent the data visually, I created several charts: a bar column chart to show the distribution of genres, a doughnut chart to represent the countries, a column chart to display the number of Oscars won, a line graph to track release years, and a pie chart to highlight which movies are available on Netflix. These visualizations made the data easier to understand and allowed us to see clear trends across the different categories. Overall, the exercise helped us practice data organization, sorting, and visualization while working with real-world information.