NEWM-N328 Final Project Documentation

George David Adams

For my final project, I chose to work with a dataset that described many different attributes about movies and shows across Netflix. The dataset described many different features about each show or movie including the title, type, director, cast members, country it was made in, date it was released, data it was added to Netflix, rating, run time, genre, and a short description of the show or movie. All of those attributes provide valuable information about the show or movie and gives us many options to explore a lot of different aspects of the dataset. For my final project, I chose to explore the different runtimes between the different movies. I wanted to explore the different run times of different movies and figure out the average run time of the movies.

When I first thought of the idea to study the different run times of the movies, my first thought was to create a scatterplot to compare all the movies in the dataset. However, after looking at how the scatterplot would be laid out, I thought the chart would look clustered and made the cart look confusing for those who would be using the chart. After I reviewed the chart and determined that would not work, I chose to use a bar chart to display the data. Since there was so much information to compare, I chose to use a smaller portion of the dataset to compare with the bar chart. I used different colors for each bar in the chart and chose movies from different genres in order to compare. After completing the chart, I can see that the average run time for movies is around 90 minutes. There are some outliers with a movie at 140 minutes and two movies at 60 and 61 minutes, but the average time seems to hover right around 90 minutes.

A graph of different colored rectangular shapes

Description automatically generated with medium confidence