Name: Smitha Patil
      Indulekha Siddabathuni
      Mohiddin Bacha Shaik
Date: 02/25/2022
DATS 6101
Team SIM

Topic Proposal

With the rise of online streaming services, understanding what makes a movie or TV show popular on IMDB has become increasingly important for filmmakers and producers. While some research has been done on this topic, there is still much to learn about which factors are most strongly associated with a movie or TV show's popularity on IMDB, and how these factors differ between the two types of media.

We propose to use the dataset "IMDB Most Popular Films and Series" to explore this question. Our SMART question is:

> *Which factors are most strongly associated with a movie or TV show's popularity on IMDB, and how do these factors differ between movies and TV shows?*

To answer this question, we will use the following independent variables: genre, duration, certificate, and the presence of nudity, violence, profanity, and alcohol in the movie or TV show. The dependent variable will be the popularity rating on IMDB.

Our dataset includes approximately 6,000 observations of popular movies and TV shows on IMDB, with data on their respective ratings, votes, and characteristics. We plan to conduct a multivariate regression analysis to determine which factors are most strongly associated with popularity, while controlling for other potential confounding variables.

We anticipate that the results of our analysis will provide valuable insights into the factors that contribute to the popularity of movies and TV shows on IMDB, and how these factors differ between the two types of media. These findings could be useful for filmmakers and producers looking to make more popular movies and TV shows, as well as for researchers interested in understanding the drivers of media consumption.

Source of dataset: [Imdb Most popular Films and series](#)

GitHub repo: [DATS_6101_Group_SIM](#)