**METHODOLOGIES**

Since movies and tv shows are a big part of the entertainment business, it was logical to analyze the dataset containing information about the movies and tv shows and get valuable insights and knowledge out of it. The main objective of this project is to find the factors that determine the success of a movie. There can be many predictor variables in the case of a movie or a tv show that would determine its success. A movie can be called successful in two contexts. One where the gross profits earned by a movie are more than the amount invested in producing a movie, and the other one is where the movie does not make any profits, but it gets good reviews from critiques. This project will demonstrate the relation of various predictor variables with both the type of successes related to a movie or a tv show.

There are various variables available within our dataset ranging from rudimentary variables like show ratings from companies, money invested, gross profit, name of the director, and actors to details like the social media statistics of every actor and director. We strongly feel that the social media influence of an actor can also be an important factor ensuring the success of the movie.

Since the project aims at finding the relationship between multiple variables related to movies and tv shows, multiple linear regression seems to be the statistical concept that serves the purpose. Along with multiple linear regression, we would also employ basic statistical concepts like mean, median, and mode to perform exploratory data analysis and extract as much knowledge and insights as possible from the dataset.

We plan to use python as our scripting language for this project. Python is widely popular for its use in data science, ease of learning, and vast data science libraries like pandas and NumPy. A python script will be used for analyzing data and designing regression models for the project. We plan to use matplotlib, a data visualization library supported by python, for visualizing the insights within the dataset. Along with matplotlib, we are also looking to use tableau to create an interactive dashboard for movies and tv shows.