CAPSTONE 1: FILM SEASONALITY

Main Idea/Question(s):

Within the film industry, there are common rules of thumb about the seasonality of film performance and, by extension, consumer behaviors. There is substantial resistance to the idea that a film can perform well outside of the traditional 'peak' times of year (mostly considered to be the early parts of the summer and the holiday season). The primary question I would like to ask is; is this belief warranted? Does season/time of year play a substantial role in a film's performance? If so, does seasonality vary by genre, studio, actor/actress, etc.? Has seasonality changed over time (i.e. is it as big a factor in 2019 as it was in 2009, 1999, or 1979)? Additionally, can we identify any other important factors, such as the competitive set of a film (the movies released on the same weekend or the weekend before, etc.)?

Dataset:

To examine the above, I will start with the IMDB extensive movies dataset from Kaggle (found here. It contains over 81,000 films (most of which will need to be excluded from the analysis due to missing data and/or not being produced and distributed by a major studio). This dataset contains 4 different tables, housing data about the films, as well as the actors/producers/studios who made them.

The data has some categorical features (like genre, release date (when classified to peak and off-peak), rating, production company, cast, etc.) and some numerical data (budget, gross revenue (domestic and worldwide), duration, etc.).

MVP:

The MVP for this project is a thorough analysis of the relationship between time of year and film performance, as well as at least some consideration of how seasonality might vary by genre (or some other categorical feature).. MVP+ would include an examination of the relationship between competitive set and film performance. MVP++ would build on the above and include additional data sources (Rotten Tomatoes' reviews, NYT reviews, and/or wikipedia page view data) to further the understanding of seasonal and competitive effects. MVP+++ would bring all of this together to make accurate predictions about film success early enough to affect choices made by the studios when it comes to marketing/advertising spend, slate planning, etc.