Bombs Away!
Juan Morales & Yakir Forman

Motivation

The filmmaking industry is a massive muti-billion dollar industry, and thus the individuals involved in the process of making a film place high value on a predictive model which gives them a sense of how much they can expect to get back from their investment. We note that in reality the set up of how much each stakeholder is expected or wants to invest in the making of a film is very complex. For example, contracts are often written to describe precisely how revenue will be shared between the creators and the exhibitor. We further note that films can make profits beyond the confines of a movie theater, examples include BluRays, streaming services, Video on Demand, and TV, among others. For the purposes of our project we will limit our analysis to the confines of a movie theater.

Goals

The goals of our project are:

- Understand the Key Performance Indicators of a film's success at the box office.
- Construct a model that accurately predicts revenue generated at the box office.
- Test and evaluate the model/s.

Our Model

We develop a model to predict revenue. (The model actually predicts the logarithm of revenue, which more accurately reflects the distribution of movie revenues.) The model does modestly better than predicting solely on the basis of budget, and considers only six predictors but has similar performance to more complex models that consider a wider range of predictors.

Key Takeaways

- The single most important factor in predicting revenue is budget. A higher-budget film is likely to have higher revenue. Production companies should keep this in mind when determining budgets.
- In addition to budget, some other factors predict a higher revenue: longer runtime, presence of a tagline, and belonging to the Adventure and/or Family genres. These could all be considered when deciding which movie to back and/or what changes to make to a movie or its advertising.
- While considering only these factors performs similarly to considering a wider range of predictors (including the presence of specific actors and directors), and it does perform better than a baseline average, the improvement is only modest. There is still a lot of uncertainty involved.