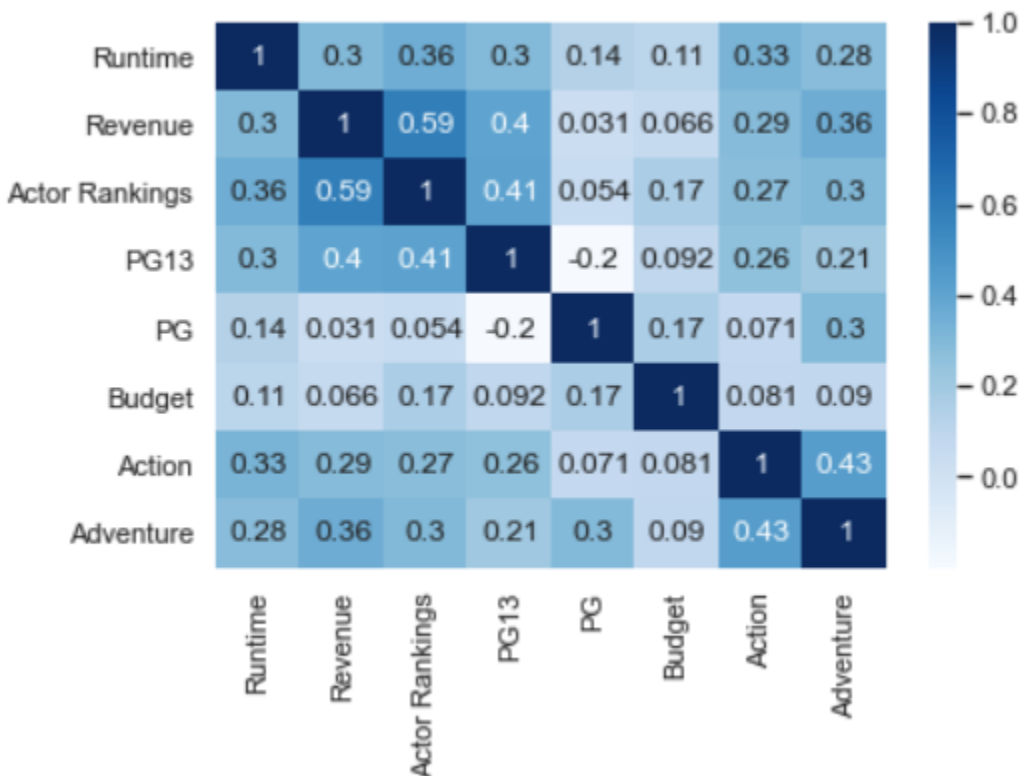


Juliana McCausland
Linear Regression Project MVP



My focus of this project is to determine the features that predict higher worldwide revenues for sci-fi films. I am looking at the length of the film (runtime), actor rankings (I scraped a list of the highest-grossing actors and counted how many of these actors was in each film), MPAA ratings, budget, and a few genres that co-occurred with sci-fi genres. This heatmap does not include all of the MPAA ratings or genres because it was illegible prior to cutting down on the number of features.

Nearly all of the features have low correlation values associated with revenue, however there does appear to be a relationship between actor rankings and revenue. There may also be a slight relationship between film MPAA ratings and revenue, as PG13 and revenue have a correlation value of 0.4, which is the second highest after actor rankings.

This does not feel incredibly insightful. I may add more features, like directors/production companies, if time permits.

(I only now realize that the position of Revenue and Runtime should be switched, since Revenue is my target. I will adjust this before the presentation!)