**Write up:**

We were able to Scrape the data from themoviedb.org to get the information for the movies and other related info like the year and the rating.

We scraped 3621 rows from the website and after doing the EDA we had 1871 rows to work on in the regression and our target is to predict the revenue.

Our starting goals were:

* Investigate the relationship between movie user rating and revenue. What is the expected revenue for a movie given a specific user rating?
* Can we predict user rating for a movie given the total budget and the number of genres which the movie belongs to?

Models that we used on this dataset:

* OLS
* Linear Regression with a score of 0.65.
* LASSO with a score of 0.67.
* Polynomial with a score of 0.74.
* CROSS Validation with a score of 0.63.

On the test portion of the dataset our models had these scores:

* Linear Regression with a score of 0.598
* LASSO with a score of 0.598
* Ridge with a score of 0.599
* Elastic Net with a score of 0.599
* Random Forrest with a score of 0.500
* Gradient Boosting with a score of 0.572