

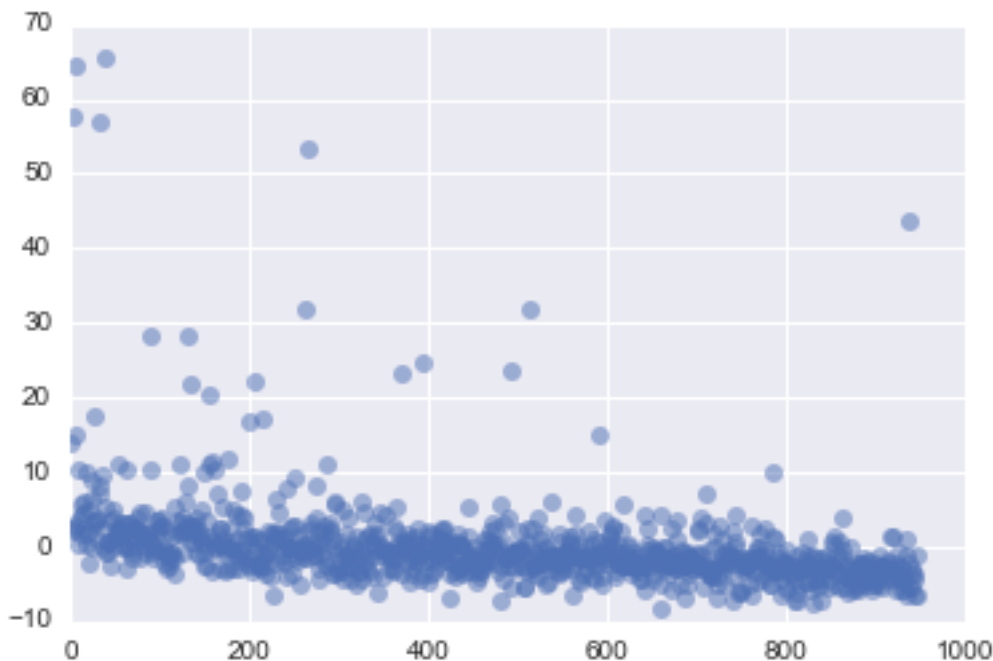
Project Luther – MVP Nils

Goal: Build a model that advises studios what to spend their money on when producing a movie by setting up a regression with target Return on investment $(= (\text{Worldwide Total adj gross} - \text{adj Budget}) / \text{adj. Budget})$ and the following features:

- Budget
- Genre (Documentary, Foreign Language, Animation, SciFi/ Fantasy, Comedy, Romance/Drama, Action/ Adventure)
- 3D technology
- Reputation of Actors (via imbd list)
- Reputation of Director (via imbd list)
- MPAA Rating
- Season (winter, summer, spring, autumn)
- Runtime

Then evaluate the effect of each feature towards return on investment.

(Preliminary) Result: All features mentioned above are built into the model. However, I haven't found major correlation and only got an R-squared of 0.1. Plot of residuals:



Next Steps: Clean up analysis and then include more features and look for correlations. Prioritized list of features to include:

- Is the movie an adaption of some book? (dummy variable)
- Is the movie based on a true story? (dummy variable)
- Tenure of actors/ directors