Intro to the Topic

Team 5 presents Through the Lens - a look at films and what drives their success. Success can be defined in two ways for this project: critical acclaim and profit. This project looks at patterns in film, seeking to understand which factors contribute to a film's success.

Users will be able to predict a film's critical success based on certain criteria, interact with a dashboard to explore the financial success of different films, and see the way films are rated on Rotten Tomatoes.

Data sources

The Box Office Mojo data provides an insight into the financial success of films, alongside the actors, directors, genres, ratings, and runtimes.

The Rotten Tomatoes data provides information about audience and critic ratings.

Final Website
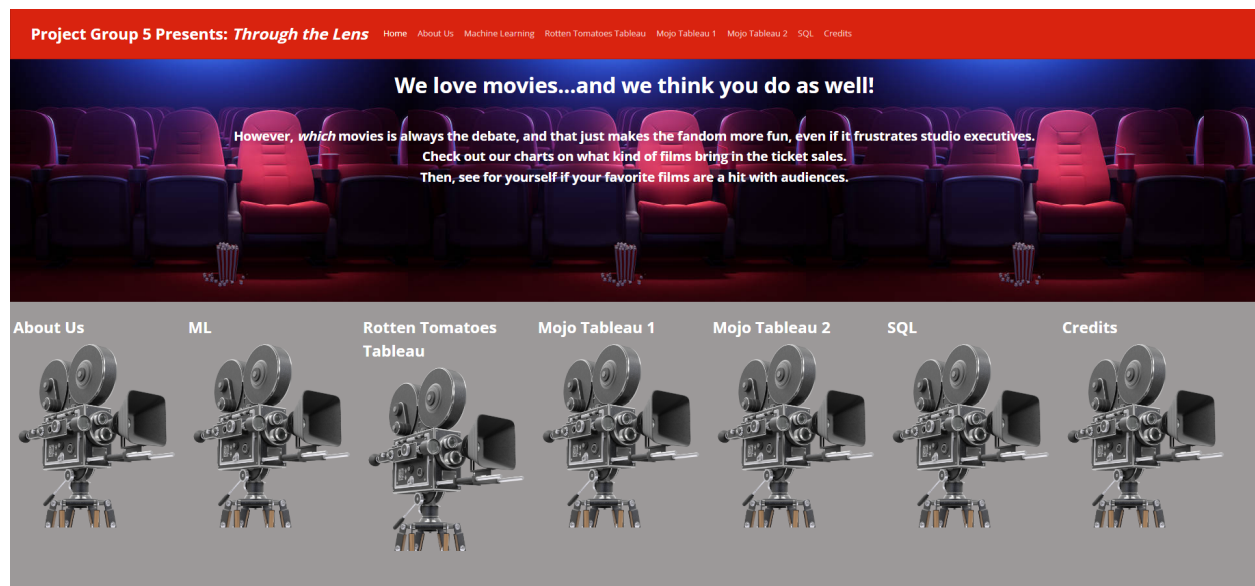
The final product of this project is the above website, which includes the following sections:
- About Us
  - Meet the team who worked on Through the Lens!
- Tomatometer Predictor
  - Use our machine learning model to predict the audience and critic Tomatometer (Rotten Tomatoes) scores from a list of 14,000+ movies. You can also see the film's actual score to see how well our machine predicts critical behavior.

- Database Query
  - Filter through the Box Office Mojo Data and explore the budget data.
- Rotten Tomatoes Tableau
  - See the evolution of Rotten Tomatoes, exploring the number of reviews by genre, studio, over time, and by content rating.
- Mojo Tableau 1
  - Start with the static dashboard exploring the average profit by top billed actor, director, content rating, runtime, and genre.
- Mojo Tableau 2
  - Explore the Box Office Mojo data further, exploring film budgets and profits over time with filters for content rating, runtime, and genre.
- Credits
  - See our resources.

Tools Utilized

- Python and Jupyter Notebook.
- Pandas, Numpy, Scikit-learn.
- Linear Regression, Random forest, LGB, and GB Regressor for model prediction.
- SQLite for Database
- Tableau for Dashboards
- HTML,CSS,Java Script for web page

Conclusions

Our analysis genuinely concluded that human behavior can be difficult to predict, especially when it comes to their reactions to art and that franchises / familiar properties tend to dominate at the box office.

To deepen the overall analysis, the first thing we would do is tie together our data, connecting the financial and Rotten Tomatoes data, and perhaps pull in some additional ratings data. Seeing how these metrics interact might help us understand audience and critic behavior on a more meaningful level.

The Box Office Mojo Tableau dashboards seem to highlight this long-standing issue within the movie industry: franchise films or familiar concepts seem to out-perform the rest. Now this begs the question - are data-driven filmmaking decisions poised to demolish artistry? Has it already? Perhaps the data could help figure out ways to make new media that has a lessened risk without the promise of a big box-office payout.