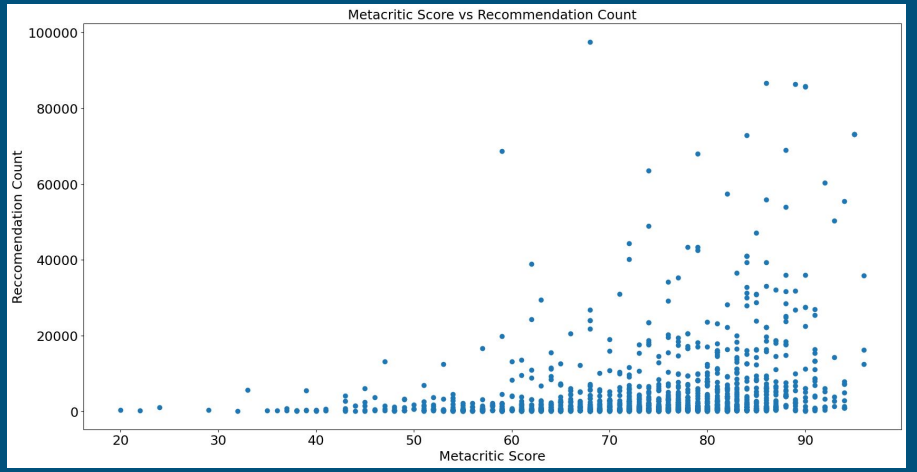
Steam Game Genre Analysis

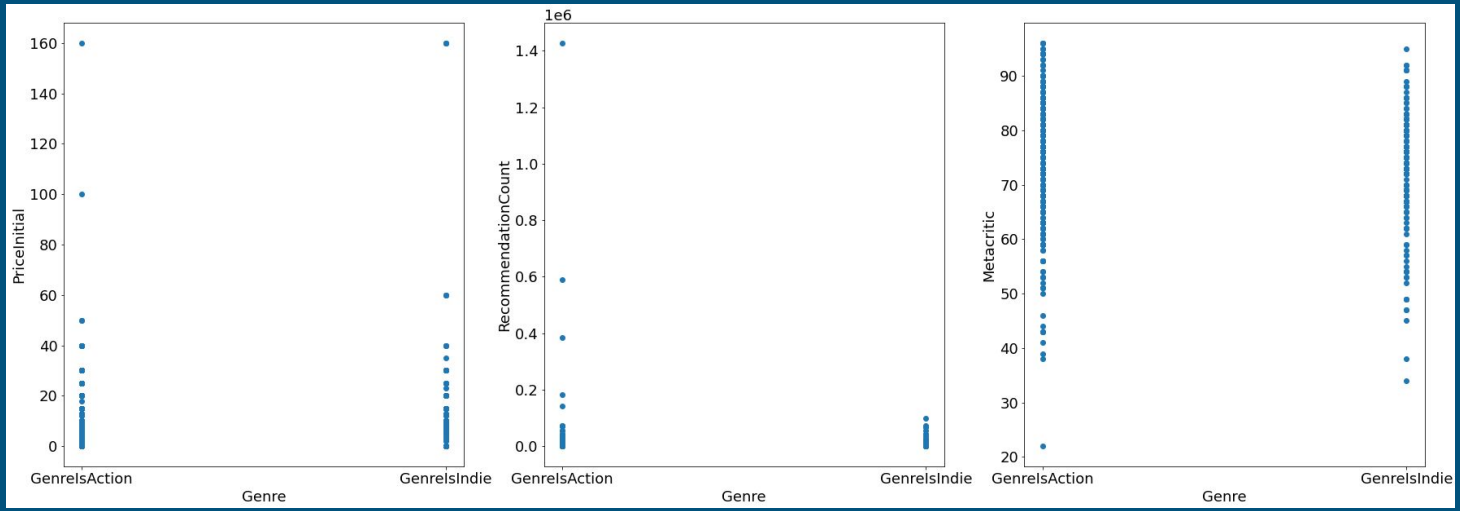
Matt Porter and Grant Ballard

Dataset

The dataset contains data on the Metacritic score, price, recommendation count, and genres for thousands of Steam games



It was collected using the Steam API and it was accessed via Kaggle

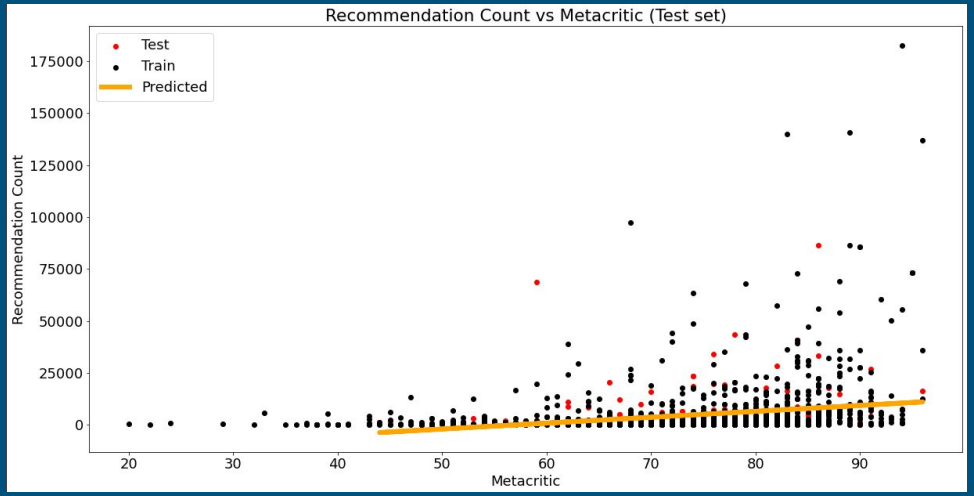


Machine Learning Methods

- The machine learning approaches we are using are linear regression and random forests
- Linear Regression
  - We are attempting to find a correlation between the metacritic score and the recommendation count of Steam games
  - We are still researching the hyperparameters for this approach
- Random Forests
  - We are trying to predict the genre of a game using recommendation count, metacritic score, and price as features
  - We are currently working with different values for n_estimators and max_depth

Results

Our initial results for linear regression were very bad, so we have switched our focus to random forests for now



Using Random Trees and attempting to predict whether the genre is action or indie yielded the following results:

	Training	Testing
Genre Accuracy	70%	66.6%