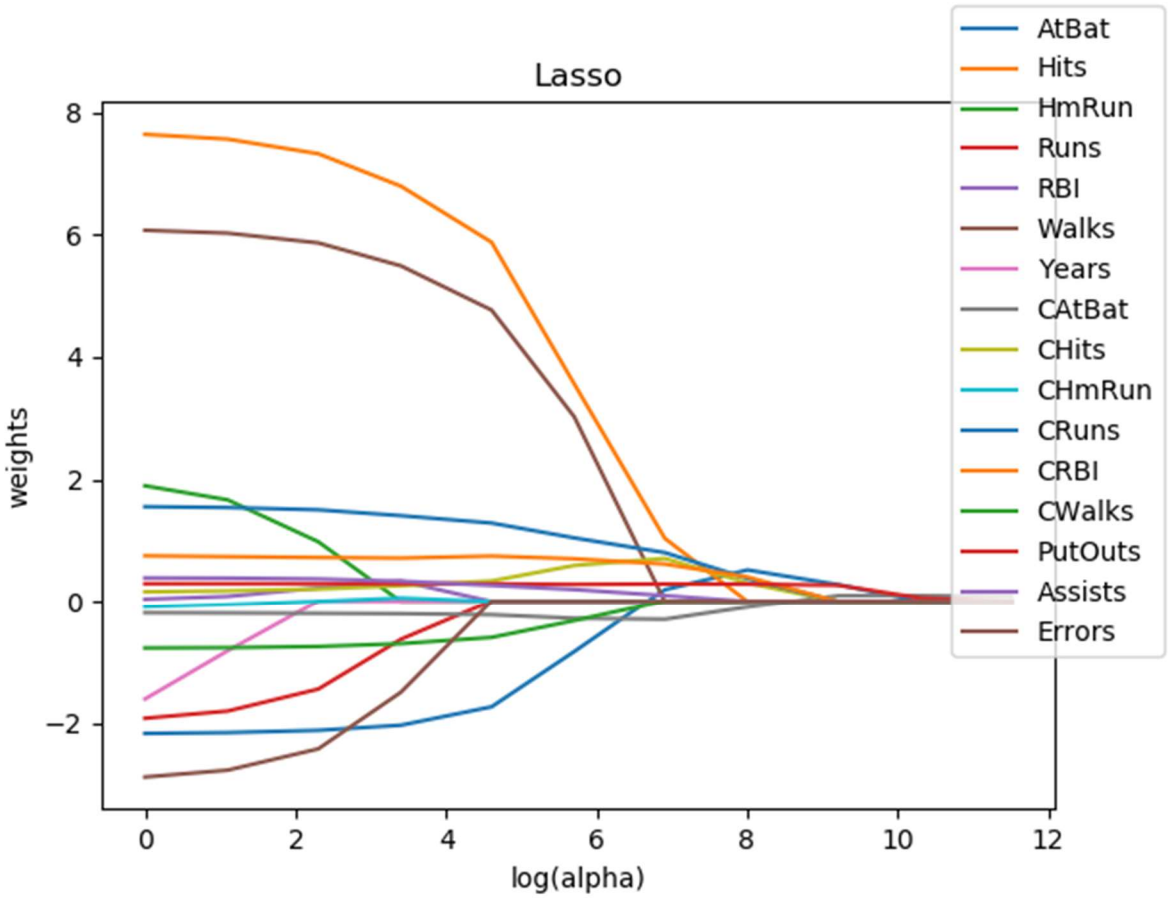
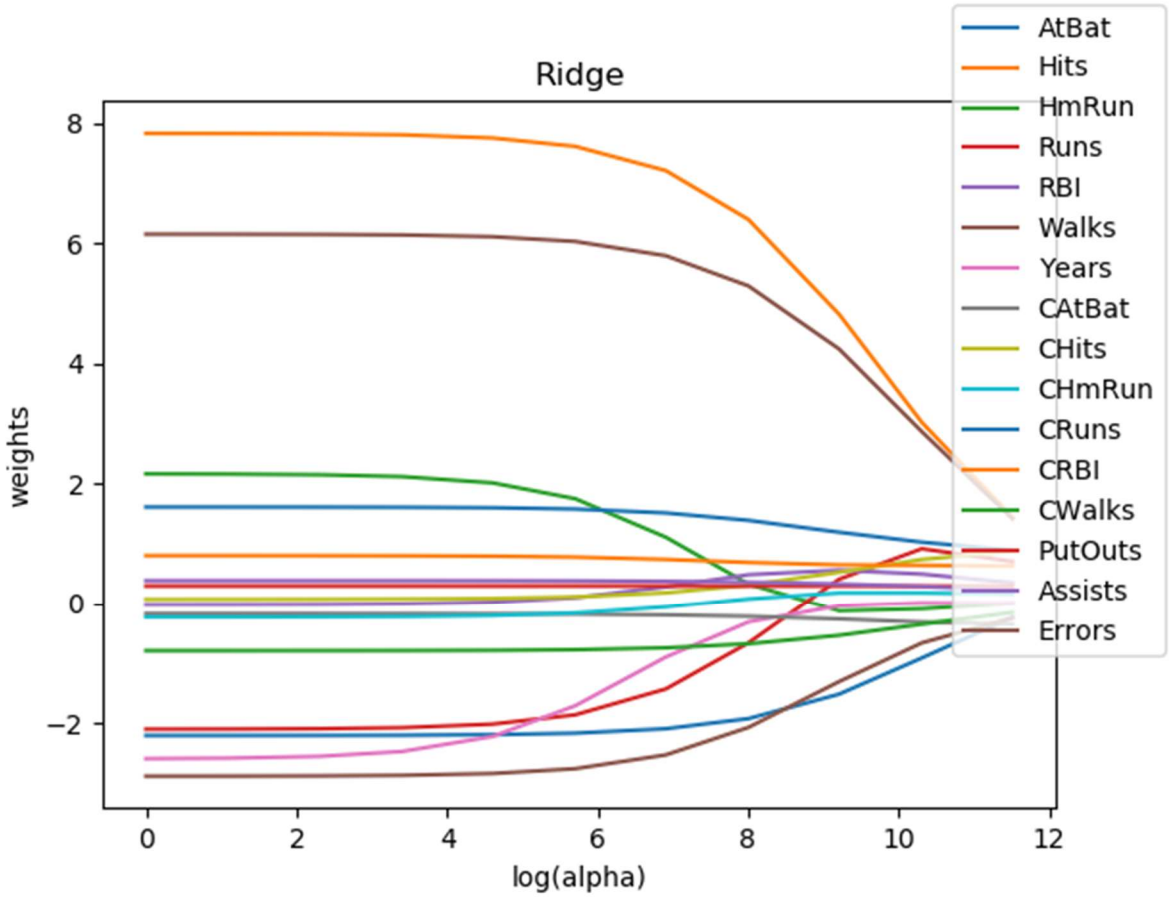
1.1 My data visualization of Lasso with different alphas is shown above. The optimal alpha found with cross-validation is **944.8**. It's difficult to tell from the plot, but the 3 features that remain the longest with nonzero coefficients are **Runs, CRuns, and CAtBat**. There are **8 predictors** left in the optimal model.



1.2 My data visualization of Ridge with different alphas is shown above. The optimal alpha found with cross-validation is **10.0**.



2 The bias-variance tradeoff is a mathematical way of describing the pros and cons of flexible statistical models. A highly flexible model has high variance and low bias, and tends to overfit to the data it is trained on. An inflexible model has high bias and low variance, and tends to underfit data. Regularization, such as Lasso (L1) or Ridge (L2) regularization, is a way of forcing a constraint on a model. For example, looking at the graphs above, the reader can see high-variance models on the left side of each graph where there is a low regularization weight *alpha.* In the high-variance models, the coefficients take on a much wider range of values. On the right side of each graph, the regularization weight constrains the models much more and results in low-variance models with a much lower range of coefficients. With an optimal regularization weight, we can train a model that is less likely to overfit or underfit.