Case Study 2

Sam Lee & Patric Platts

1

After standardizing the numerical covariates, fitting a lasso (or elastic selection model) and applying some level of cross-validation to select the best hyperparameters for the model of best fit, we would examine the covariates that are the largest in absolute value and are significant.

2

We would look at the adjusted R^2 statistic as an overall model fit. We would also compare how well our selected covariates from the lasso regression (or whatever model we end up deciding) compare to the full model fit by running an ANOVA test of the two models (comparing their f-statistics).

3

We would look at how well the (out-of-sample) RMSE improves (or doesn't) with our selected covariates compared to other models used with other covariates.