Show diminishing returns of hyperparameter tuning via graph

Run gridsearchcv on random forest and see if there is a RMSE improvement

Do some EDA-

Nice way to show the signals via plot

Summary statistics showing why we drop certain columns that aren’t giving us any signal

Come up with metric for performance vs computer time

Use separate test data to see how close we got

Write up report

Write conclusion about looking into other ensemble methods maybe neural networks and then extending it to the multiple data sets