**Game Plan for Housing Prices Prediction**

1. Wayne - read and understand existing code
2. Chris - Create scripts to separate out test and train sets (75% vs 25%)
3. Wayne and Chris - learn github workflows:
   1. <https://guides.github.com/introduction/flow/>
   2. <https://datasift.github.io/gitflow/IntroducingGitFlow.html>
   3. <https://stackoverflow.com/questions/18188492/what-are-the-pros-and-cons-of-git-flow-vs-github-flow>
   4. <https://learngitbranching.js.org/>
4. Come up with a tech stack / methodology for coding collaboration
   1. There seems to be multiple workflow methods for git and coding collaboration. Initial research makes it seem like there’s gitflow and github flow.

Scripts we need:

Training Data Prep + Cleanse (DONE)

Training Data Imputation (DONE)

Linear Model Training (DONE)

LASSO Model Training (DONE)

Random Forest Regressor Training (DONE)

Ensemble Training

Exploratory Data Analysis

Training Set Held Out Prep

Build Scripts to calculate RMSE on the 25% train set

* LASSO
* RF
* Linear

Scoring Data Prep + Cleanse (DONE)

Scoring Data Imputation (DONE)

Master Scoring Script

* Linear Model Scoring
* LASSO Model Scoring
* Random Forest Regressor Scoring

Game Plan:

1. Create 75 / 25 training script to compare different models and generate errors for validation script
2. Using the information from step 1) we’ll generate a full training script + full validation script