Notes from the Model:

* Decision trees don’t have to standardize your features

Visualizing Decision Trees:

* Careful: Decision Trees Tend to overfit
* The decision tree is still predicting all our participants as unemployed.
* But this is a powerful visualization for explanatory reasons

Bagged Trees:

* Multiple decision trees come together to build bagged trees
* Evaluation metric: for bagged tree regressor can use R^2
* Our R^2 score here is fairly low, about .03, so only about 3% explanatory power in our current model
* Here our score improves after a certain number of estimators

Random Forests

* Type of bagged tree model.
* Takes long to run, so only using a sample of the data (1000 out of 38000)
* Here: uses R^2 as a metric
* Can visualize individual decision trees
* Cool thing about Random Forest: gives relative importance of each feature