# Modeling Thoughts for Lane Predictions:

* For a given lane, see historically if the gross income (All in Charges- All in Costs) has a systematic bias when comparing gross income as calculated by (NormalizedCustomerLine –NormalizedCarrierLineHaul)
  + If so, we can model this as a departure from the model prediction which may vary by lane and seasonal specific factors.
  + The most natural prediction format is RPM (at least normalized RPM). This is what mike and camp want. We should add a premium/discount to this as a separate modeling feature to capture this departure and also make it easier to explain.
* I like the idea of ensemble modeling where we fit an overall model for all of the data (using a boosted tree) on the simple factors, enter that rate prediction set into the lane model as a predictor. This is kind of a Bayesian approach where we have a prior and update that as the likelihood from the current model.
  + Maybe we work with the deltas form the current model, or we put it right in… Not sure which I like best.
* Maybe a better idea is to have some prototypical lanes and then do a K(X,X) similarity measure between these “protypical” lanes
  + Basically, save 6 really detailed model images and do a similarity score between constructed lanes and the ones of interest
  + You can use a DAG pull weather etc…
    - Basically, your established model will be a linear combination of these central lanes that you fit a very expensive detailed model (boosted trees, svm, nueral net, gp etc…).