4a. Conduct a single layer of validation on the supervised learning technique.  Tune at least one parameter or compare at least two subsets of predictor variables.  (Tuning multiple parameters may help improve your model’s performance.)

* + If your data set is moderately-sized, use 10-fold or 5-fold cross-validation.
  + If your data set is large enough that 5-fold CV is prohibitively time-consuming, use a validation set.

4b. Conduct an outer layer of validation, containing all of the supervised learning models.  Modify your code from the single layer of validation to use *traindata.out* rather than the full data set.  Assess the performance of your model selection process.

* + If your data set is moderately-sized, use a *for* loop to conduct 10-fold or 5-fold cross-validation.
  + If your data set is large enough that 5-fold CV is prohibitively time-consuming, use a validation set.
  + Include a comment in your code indicating what the best estimate of your model's future performance is, and which line of code computes this.  (It is OK if this value changes slightly when your instructors re-run your code.  However, you are welcome to set the random seed to ensure consistency.)
  + ~~If you earned less than 33 out of 35 points on the “technical accuracy” portion of the midterm project, then you must either use two layers of cross-validation here, or earn 10 out of 10 points on your Cross-Validation Resubmission from the midterm project.~~

4c. Based on your results from steps 4a and/or 4b, select one “best” model (including its tuning parameter values).  Fit the model on the entire data set.  (One way to do this is to extract the *$finalModel* component of the *caret* object from step a, when used to perform a single layer of CV on the entire data set.)  In step 5, you will be interpreting this final supervised model **and** the unsupervised model with the parameters you select in step 4d.

4d. Apply the unsupervised learning technique to your full data set.  Use graphs and tables to explore and interpret the results.  Adjust parameters (such as *k* for clustering or the minimum support for association rules) as needed to obtain informative, insightful interpretations.  See steps 5-7 for more about interpreting your model.