In-Class Notes 3-30-15

Car Price Walkthrough (Tiff's Treats Competition)

- Create a model using all of the available variables in the training data set and use the step function to parse out unnecessary variables
- If applicable, try using variables to account for possible nonlinearity and interaction terms
- Test the accuracy of your model by finding the mean of squared errors using the testing data set
    - This will basically tell you the difference between predicted price and actual price given all of the variables in your model
    - You can edit your model by using different interaction terms to see which of them has the lowest MSE
        - $MSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
- Rerunning the step function using interaction terms is not a good option – this operation is too much for any laptop to handle

Step 1: Create an initial model and use the step function to eliminate unneeded terms

```
lm_all=lm(MMRAcquisitionAuctionCleanPrice~., data=carauction_train)
lm_step=step(lm_all, direction='backward', trace=0, data=carauction_train)
```

Step 2: Add relevant interaction terms

```
lm_new=lm(MMRAcquisitionAuctionCleanPrice~Auction+VehYear+VehicleAge+Make+Col
or+Transmission+WheelType+VehOdo+Size+AUCGUART+VNST+IsOnlineSale+WarrantyCost
+ModelName+Make:Size+Make:VehOdo, data=carauction_train)
```

Step 3: Find the mean of squared errors

Formula: $MSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

R-Script:

```
yhat_test=predict(lm_new, newdata=carauction_test)
sqrt(mean((yhat_test - carauction_test$MMRAcquisitionAuctionCleanPrice)^2))
```

Step 4: Revise the model to try to lower MSE

```
lmfinal=lm(MMRAcquisitionAuctionCleanPrice ~ Auction + VehYear +
  VehicleAge + Make + Color + Transmission + WheelType + VehOdo +
  Size + AUCGUART + VNST + IsOnlineSale + WarrantyCost + ModelName +
Make:Size + Make:VehOdo + ModelName:Color,
  data = carauction_train)
yhat_test=predict(lmfinal, newdata=carauction_test)
sqrt(mean((yhat_test - carauction_test$MMRAcquisitionAuctionCleanPrice)^2))
```