MIS 6357 **Homework 2** Due: Friday, January 27, 2017

1. Write an R function that takes two vectors, y and $\hat{y}$ and returns the RMSE.

   Questions 2 : 4 are all from the text. For some questions I have provided more specific directions. These directions supplement and clarify. They do not replace the questions in the text.

2. Complete exercise 3.1 in the text.
   (a) Display two meaningful, intersiting histograms. Create a visualization of the correlation matrix. The function `corrplot` is very useful here.
   (b) Use `apply` to compute the skewness of each predictor.
   (c) Perform at least one Box-Cox transformation and plot 2 histograms: one of the pre-transformed data and one of the post-transformed data.

3. Complete exercise 3.2 in the text.
   (a) Make use of the `apply` function. Also, remove any features with degenerate distributions.
   (b) Find the percent of missing data for each of the classes.
   (c) Implement the strategy you develop and produce a data set with no missing values.

4. Complete exercise 3.3 in the text.
   (a) Just do what it says.
   (b) Also, remove any features with degenerate distributions.
   (c) Perform PCA on the data and decide how many principal components to keep.