**Part 1: Read and examine the data.**  
Read the data from the file, planttype.csv. (attached to this assignment)  
Use the “str” and “head” functions to look at the types of variables in the data frame and to view the first few lines of data to make sure it was read correctly.  
Take a look at a scatterplot matrix of variables x1-x8, using different colors to differentiate the plant types (use the "pairs" function with the “col” argument).

**Part 2: Logistic Regression**

Fit a logistic regression model to the data with response variable, type, and predictor variables, x1-x8.  
Add a fitted logistic regression probability curve to a plot of plant type versus predicted values on the logit scale. Note: “fitted()” calculates fitted probabilities, whereas “predict()” calculates predicted values on the logit scale.  
Use the summary command to extract the estimated regression coefficients.  
Use these coefficients to calculate the predicted value (on the logit scale) for a plant having values: x1=6, x2=7, x3=8, x4=7, x5=6, x6=6x6=6, x7=6, x8=7.  
Using the inverse of the logit transformation, convert this prediction to a probability of plant type 1 on the original scale.  
Calculate the likelihood-based 95% confidence interval for the logistic regression coefficients.  
Reclassify the plants using the logistic regression probabilities (predict type 1 if the probability is greater than 0.5, 0 otherwise) and tabulate the results against the actual types (i.e., determine how many actual type 1 plants were predicted as type 1 plants by the model, etc.).

**Part 3: Linear discriminant analysis**  
Carry out a linear discriminant analysis to find the variable combination of x1-x8 that best discriminates plants of different types.  
Draw two histograms of the linear discriminant function, one for type 0, and the other for type 1.  
Reclassify the plants using the linear discriminant function and tabulate results.  
Which model performs better, the logistic regression model or the linear discriminant analysis?

**Part 4: Principal Components Analysis**  
Examine associations among variables x1-x8 using graphs (ignore type).  
Carry out a principal components analysis on the plant measurements, x1-x8. Do NOT use type.  
Examine the proportion of variance explained by each principal component.  
Create a scree plot to visualize the magnitudes of the eigenvalues.  
Create a biplot to visualize the contribution of variables to the first two principal components.  
Examine and interpret the eigenvectors for the first two principal components.  
Examine the eigenvectors for the third, fourth and fifth principal components.  
Save the scores for the first two principal components and create a scatterplot using them.  
Replot the first two principal components, but this time use different colors/symbols for the different plant types. Add a legend to identify the colors/symbols.  
Would you expect the principal components analysis to necessarily help with predicting plant type? Does it help in this case?

the csv of the second project http://www.filedropper.com/finalprojectpart3planttype