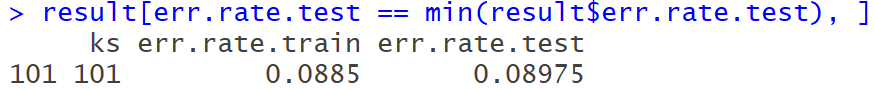
STAT 4360 Mini Project 1

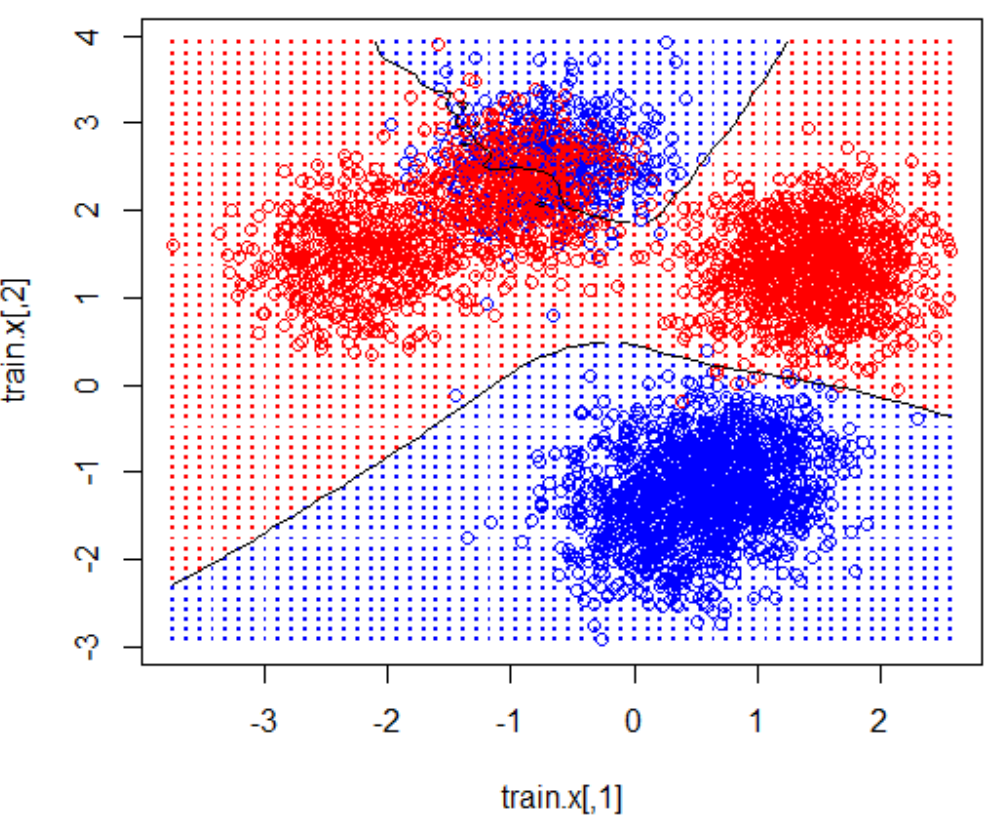
Name: Jaemin Lee

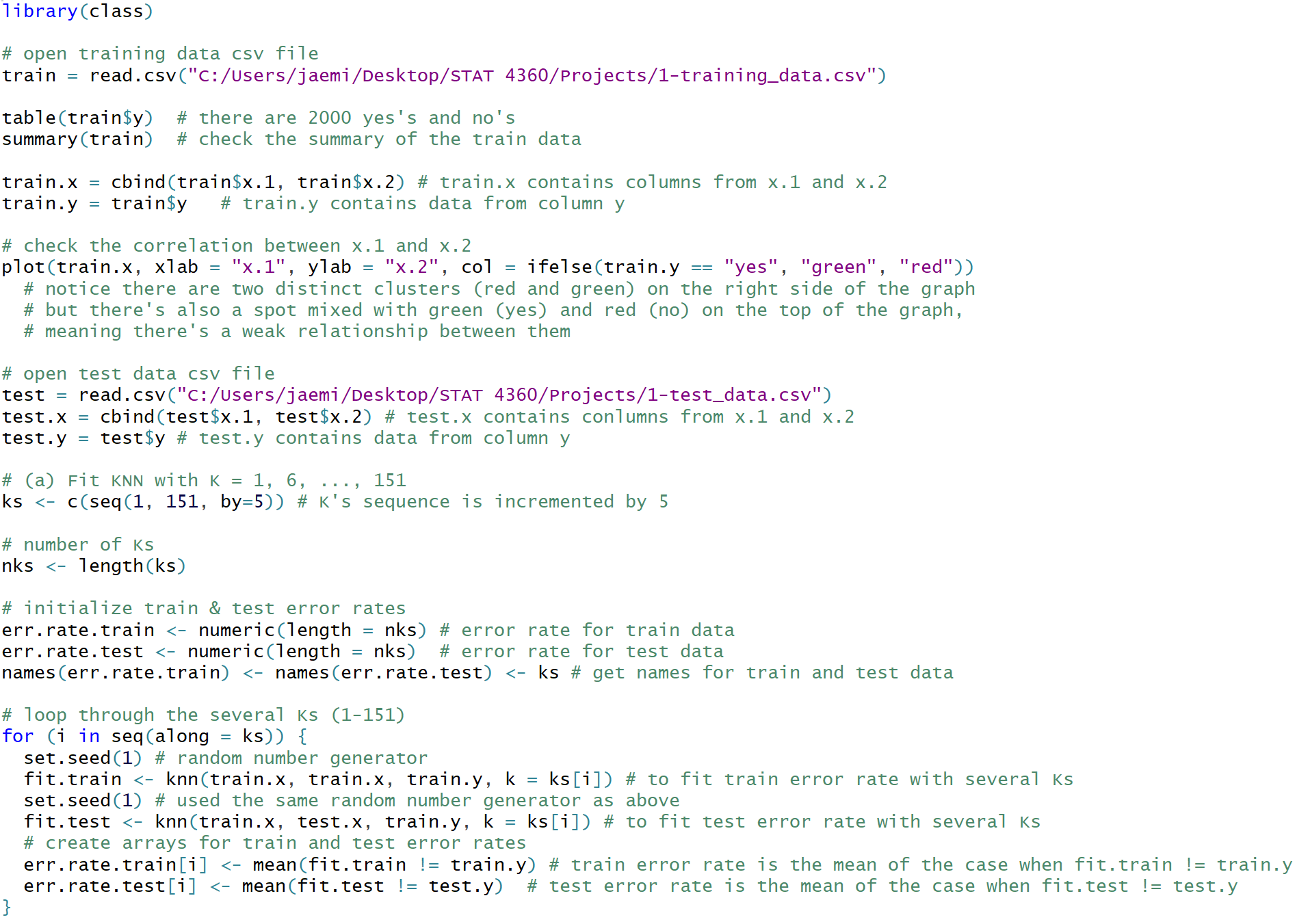
Section 1: Answers to the specific questions asked

1. Below is the graph of training and test error rates against K. Notice the red graph, test error rate, forms a U-shape. Train error rate increases as K increases. This is consistent with what we’ve discussed in class. This means that the graph is fitted correctly.
2. The optimal K is 101. The train error rate is 0.0885 (= 8.85%) and the test error rate is 0.08975 (=8.975%) when K = 101. Below shows how to find the optimal K, test and train error rates.



1. The graph below represents the plot of the training data that shows the decision boundary when K = 101. The decision boundaries (two black lines) look well smoothed, meaning that it is neither too over-fitted nor under-fitted. You can tell two classes (blue and red) are well separated by the decision boundary, thus the decision boundary does seem sensible when K = 101.



Section 2: R code

