FE-582 – Assignment 3

Problem 1

Follow the example in Lecture 5 R code to analyze the pair trading strategy for several pairs taken from the following list: PEP, KO, DPS for a period of at least 10 years. You may download stock data from WRDS (you can register and create an account as Stevens student at https://wrds-web.wharton.upenn.edu/wrds/) or other free data sources such that Yahoo. Include transaction costs of 0.1% (or 10 basis points) for each transaction. Discuss the results.

Problem 2

Use the data set Default.csv which has 7,000 observations on the following 4 variables:

- default - A factor with levels No and Yes indicating whether the customer defaulted on their debt
- student - A factor with levels No and Yes indicating whether the customer is a student
- balance - The average balance that the customer has remaining on their credit card after making their monthly payment
- income - Income of customer

Apply logistic regression, linear discriminant analysis, quadratic discriminant analysis and K-nearest neighbor classification methods to predict customers that are likely to default in DefaultPredict.csv dataset. Please use several values of K in the KNN classification method such that you can minimize the errors. Compare the errors for all the methods and draw conclusions.