# Homework 3

Due on 03/23/2021

This questions will be answered using the `Weekly` data set, which is part of the `ISLR` package. This data is similar in nature to the `Smarket` data on the textbook (ISL, Chapter 4.6) except that it contains 1,089 weekly returns for 21 years, from the beginning of 1990 to the end of 2010. A description of the data can be found by typing `?Weekly` in the Console. (Note that the column `Today` is not a predictor.)

(a) Produce some graphical summaries of the `Weekly` data.

(b) Use the data from 1990 to 2008 as the training data and the held-out data as the test data. Perform a logistic regression with `Direction` as the response and the five `Lag` variables plus `Volume` as predictors. Do any of the predictors appear to be statistically significant? If so, which ones? Compute the confusion matrix and overall fraction of correct predictions using the test data. Briefly explain what the confusion matrix is telling you.

(c) Now fit the logistic regression model using the training data period from 1990 to 2008, with `Lag1` and `Lag2` as the predictors. Plot the ROC curve using the test data and report the AUC.

(d) Repeat (c) using LDA and QDA.

(e) Repeat (c) using KNN. Briefly discuss your results in (c)-(e).