ISE-5970: Energy Analytics

Homework 2

Due: Thursday September 19th, 11:59 p.m.

Before start, please read the following.

1. The questions in this homework allow you to practice your R skills for time series benchmark forecasting models.
2. For all questions, you must submit 1) the source file that contains the R commands, and 2) the snapshot of what R outputs after you run your R program.
3. I strongly prefer if you **electronically submit** your homework through Canvas by putting all files in a zip folder.
4. Please assign numbers for each solutions, so it would be easy for me to read the answers.

Good Luck! ☺

Question 1 (25 credits):

visnights contains quarterly visitor nights (in millions) from 1998-2015 for eight regions of Australia.

a) Use window() to create three training sets for visnights[,"QLDMetro"], omitting the last 1, 2 and 3 years; call these train1, train2, and train3, respectively. For example, `train1 <- window(visnights[, "QLDMetro"], end = c(2015, 4))`.
b) Compute one year of forecasts for each training set using the snaive() method. Call these fc1 (trained by train1 set), fc2 (trained by train2 set), and fc3 (trained by train3 set), respectively.
c) Use accuracy() to compare the MAPE over the three test sets. Comment on these.
Create three test sets for visnights[,"QLDMetro"] considering the last 1, 2, 3 years; call these test1, test2, test3. Use test1 with fc1, test2 with fc2, and test3 with fc3.
d) How sensitive are the accuracy measures to the training/test split?

Question 2 (25 credits):

Use the Dow Jones index (data set dowjones) to do the following:

a) Produce a time plot of the series.
b) Produce forecasts using the drift method and plot them.
c) Show that the forecasts are identical to extending the line drawn between the first and last observations.
d) Try using some of the other benchmark functions to forecast the same data set. Which do you think is best? Why?

Question 3 (25 credits):

Consider the daily closing IBM stock prices (data set ibmclose).

a) Produce some plots of the data in order to become familiar with it.
b) Split the data into a training set of 300 observations and a test set of 69 observations.
c) Try using various benchmark methods to forecast the training set and compare the results on the test set. Which method did best?
d) Check the residuals of your preferred method. Do they resemble white noise?

Question 4 (25 credits):

Consider the sales of new one-family houses in the USA, Jan 1973 – Nov 1995 (data set hsales).

a) Produce some plots of the data in order to become familiar with it.
b) Split the hsales data set into a training set and a test set, where the test set is the last two years of data.
c) Try using various benchmark methods to forecast the training set and compare the results on the test set. Which method did best?
d) Check the residuals of your preferred method. Do they resemble white noise?