ISE-5970: Energy Analytics

Homework 1

Before start, please read the following.

1. The questions in this homework allow you to practice your R skills for time series data visualization.

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.

4. **BONUS CREDIT (10 credits):** It is very helpful that you learn how to prepare documents in LaTeX, especially if you are a graduate student. If you are not familiar with this tool, this is good opportunity to learn! LaTeX is a set of libraries that allow you to prepare high-quality typesetting in PDF format especially when it comes to preparation of scientific documents. It is available for free to both Mac and PC. If you have not already installed LaTeX on your machine, please see MiKTex website (https://miktex.org) to download and install the LaTeX engine and many of its common libraries. You might also find easy-to-use editors such as TeXstudio (free) or WinEdt (free for evaluation for Windows users).

   Using LaTeX may not seem easy at the beginning, but the more frequently you use it, the more comfortable you will become with the syntax. For this homework, I would like you to prepare your answer to Question 4 in LaTex. I will post a sample LaTeX file to make life easier for you, but I strongly recommend that you start playing with LaTeX as soon as possible if this will be your first experience.

Question 1 (25 credits):

Download some monthly Australian retail data from the book website. These represent retail sales in various categories for different Australian states, and are stored in a MS-Excel file.

    a. You can read the data into R with the following script:

```
retaildata <- readxl::read_excel("data/retail.xlsx", skip=1)
```

       The second argument (skip=1) is required because the Excel sheet has two header rows.

    b. Select one of the time series as follows (but replace the column name with your own chosen column):

```
myts <- ts(retaildata[,"A3349873A"], frequency=12, start=c(1982,4))
```

    c. Explore your chosen retail time series using the following functions:
       Autoplot(), ggseasonplot(), ggsubseriesplot(), gglagplot(), ggAcf()

       Can you spot any seasonality, cyclicity and trend? What do you learn about the series?

Question 2 (25 credits):

       Use the following graphics functions: autoplot, ggseasonplot, ggsubseriesplot, gglagplot, ggAcf and explore features from the following time series: hsales, usdeaths, bricksq, sunspotarea, gasoline. Can you spot any seasonality, cyclicity and trend? What do you learn about the series?

Question 3 (25 credits):

       The pigs data shows the monthly total number of pigs slaughtered in Victoria, Australia, from Jan 1980 to Aug 1995. Use mypigs <- window(pigs, start=1990) to select the data starting from 1990. Use autoplot and ggAcf for mypigs series and compare these to white noise plots from Figures 2.17 and 2.18.

Question 4 (25 credits):

       dj contains 292 consecutive trading days of the Dow Jones Index. Use ddj <- diff(dj) to compute the daily changes in the index. Plot ddj and its ACF. Do the changes in the Dow Jones Index look like white noise?