AE6819 – Trimester 1, AY 2023-2024

Time Series Modelling and Forecasting

Homework 1: Due at end of the day on 06 Sept 2023 (If you need an extension, you have to let me know by noon 06 Sept 2023 how much additional time you will need). Please submit via NTULearn.

After the suggested answers are uploaded at NTULearn, late submissions will receive a zero.

For the homework, please write the codes as below, followed by generated results or graphs and then if required answer the question. Marks will be benchmarked vis-à-vis the best answers in terms of clarity, detail and completeness.

Question 1 and Question 2 are from Chapter 2 of the online text. Question 3 and Question 4 are from Chapter 3 of the online text.

Question 1 (Exercise 2 of Chapter 2, section 2.10 of the online text).

Download the file tute1.csv from the book website (http://otexts.com/fpp2/extrafiles/tute1.csv) or from NTULearn, open it in Excel (or some other spreadsheet application), and review its contents. You should find four columns of information. Columns B through D each contain a quarterly series, labelled Sales, AdBudget and GDP. Sales contains the quarterly sales for a small company over the period 1981-2005. AdBudget is the advertising budget and GDP is the gross domestic product. All series have been adjusted for inflation.

Import the dataset into R Studio global environment as discussed in Lecture 1 (Using R). In doing so, please make sure you set the correct working directory from the pull down directory "session"

```
autoplot(mytimeseries, facets=TRUE)
```

Remember to call on the package fpp2 when you begin a new session either by typing **library(fpp2)** or checking on 'fpp2' in the lower right window of R-Studio

The graph will appear on the lower right window of R-Studio. Use "export" to save the graph as pdf or as an image and then copy the graph in Word.

Check what happens when you don't include facets=TRUE.

Question 2 (Exercise 3 of Chapter 2, section 2.10 of the online text).

Follow the instructions as above in importing the dataset and in exporting the resulting graphs. In importing the dataset, please note that you will import an Excel file for this question so make sure that you choose Excel file in the pop up window in RStudio.

Download some monthly Australian retail data from the book website. (https://otexts.com/fpp2/extrafiles/retail.xlsx) or from NTULearn. These represent retail

sales in various categories for different Australian states, and are stored in a MS-Excel file.

Move the file retail.xlsx into your working directory in R. You may do this in R Studio under the pulldown menu "Session" as discussed in Lecture 1. Alternatively, you may use the function setwd("")

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

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

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

Alternatively, you may also use the "Import Dataset" pull down window on the upper right window of R Studio in "Environment" as discussed in Lecture 1

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

In the example below, the authors choose "A3349873A".

You may choose a different series.

```
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 3 and Question 4 are from Chapter 3 of the online text

Question 3 (Exercise 3 of Chapter 3 in Section 3.7)

What Box-Cox transformation would you select for your retail data (from Question 2 above which is from Exercise 3 in Section 2.10)?

Question 4 (Exercise 8 of Chapter 3 in Section 3.7)

For your retail time series (from Question 2 above which is also Exercise 8 in Section 2.10):

a. Split the data into two parts using

```
myts.train <- window(myts, end=c(2010,12))
myts.test <- window(myts, start=2011)
```

b. Check that your data have been split appropriately by producing the following plot.

```
autoplot(myts) +
  autolayer(myts.train, series="Training") +
```

```
autolayer(myts.test, series="Test")
```

c. Calculate forecasts using `snaive` applied to `myts.train`.

```
fc <- snaive(myts.train)
```

d. Compare the accuracy of your forecasts against the actual values stored in `myts.test`.

```
accuracy(fc,myts.test)
```

e. Check the residuals.

```
checkresiduals(fc)
```

Do the residuals appear to be uncorrelated and normally distributed?

f. How sensitive are the accuracy measures to the training/test split?