ADS-506 Mod 3

Module 3 Assignment Exercises

The assignment for this module is a mixture of programming and written work. Complete this assignment in R Markdown. You will need to include the question and number that you are answering within your submitted assignment. For programming answers using R, answers should be written in **R Markdown and 'knitted' to a Word/PDF file.**

- Submit the written responses as a Word/PDF document. Answers must be clear and explicit.
- Submit the answer code as a PDF document. You can publish R code and output in R Markdown and convert it to a PDF file.

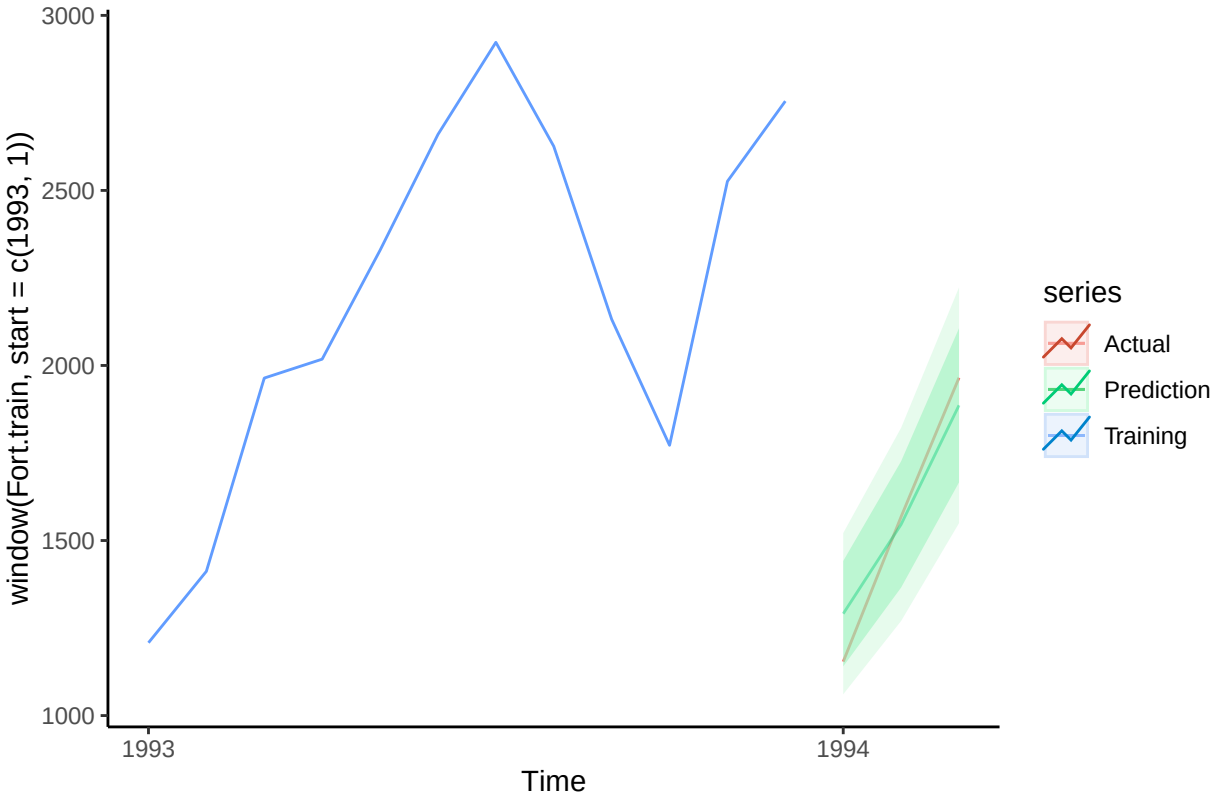
Textbook Exercises (Pages 113-116 & 141)
5.8
Forecasting Australian Wine Sales : Figure 5.13 shows time plots of monthly sales of six types of Australian wines (red, rose, sweet white, dry white, sparkling, and fortified) for 1980-1994. Data available in AustralianWines.xls. The units are thousands of liters. You are hired to obtain short-term forecasts (2-3 months ahead) for each of the six series, and this task will be repeated every month

b) Fortified wine has the largest market share of the six types of wine. You are asked to focus on fortified wine sales alone and produce as accurate a forecast as possible for the next two months

```
##                     ME     RMSE      MAE         MPE     MAPE      MASE
## Training set 15.49844 285.2321 222.3198 -0.04620971 7.098462 0.7989745
##                     ACF1
## Training set 0.06838359
```

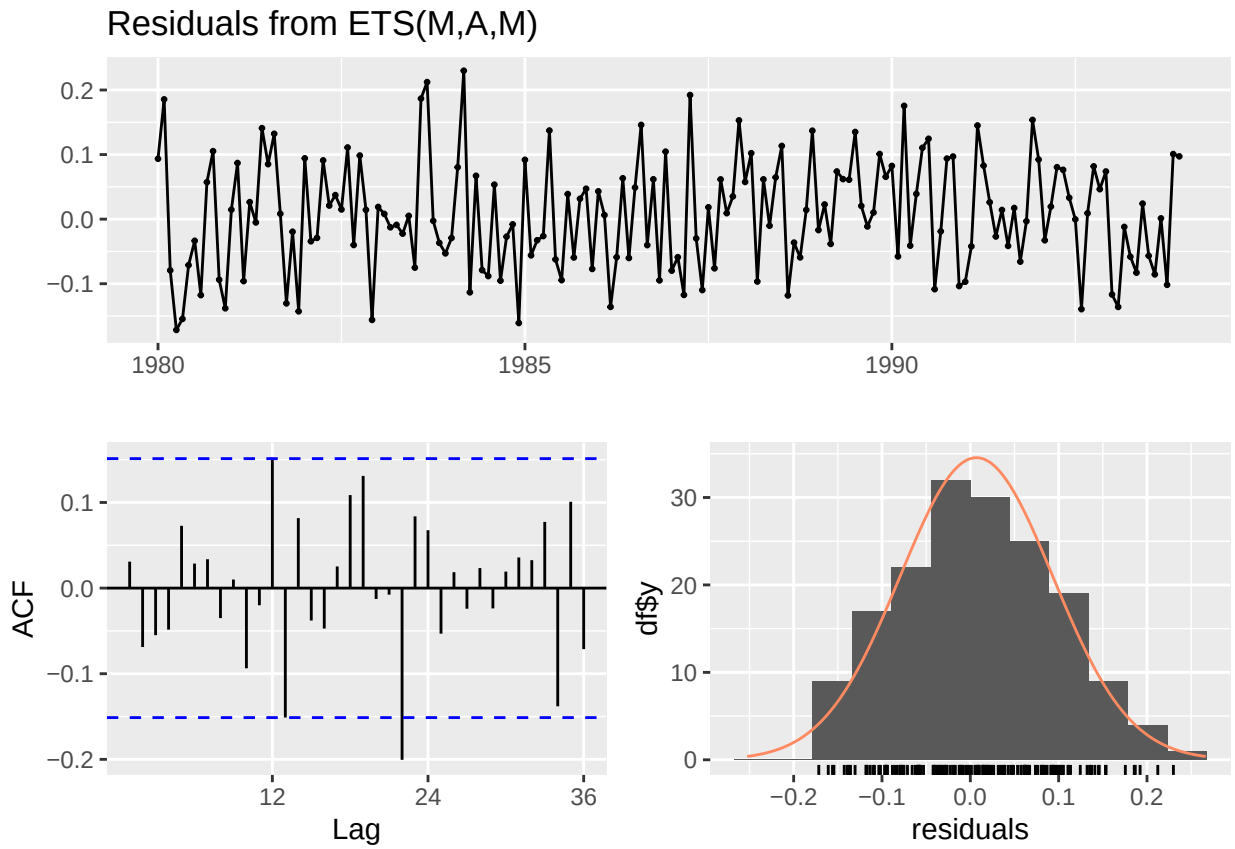
```
##                     ME     RMSE      MAE         MPE     MAPE      MASE
## Training set 15.49844 285.2321 222.3198 -0.04620971 7.098462 0.7989745
##                     ACF1
## Training set 0.06838359
```

```
# Plot the model forecast.
autoplot(window(Fort.train, start = c(1993, 1)), series = 'Training') +
  autolayer(window(Fort.test, end = c(1994, 3)), series = 'Actual') +
  autolayer(Fort.HWES.pred, series = 'Prediction', alpha = .4) +
  theme_classic()
```



c) Create a time plot of the residuals from the Holt-Winter's exponential smoothing.

```
checkresiduals(Fort.HWES.model)
```



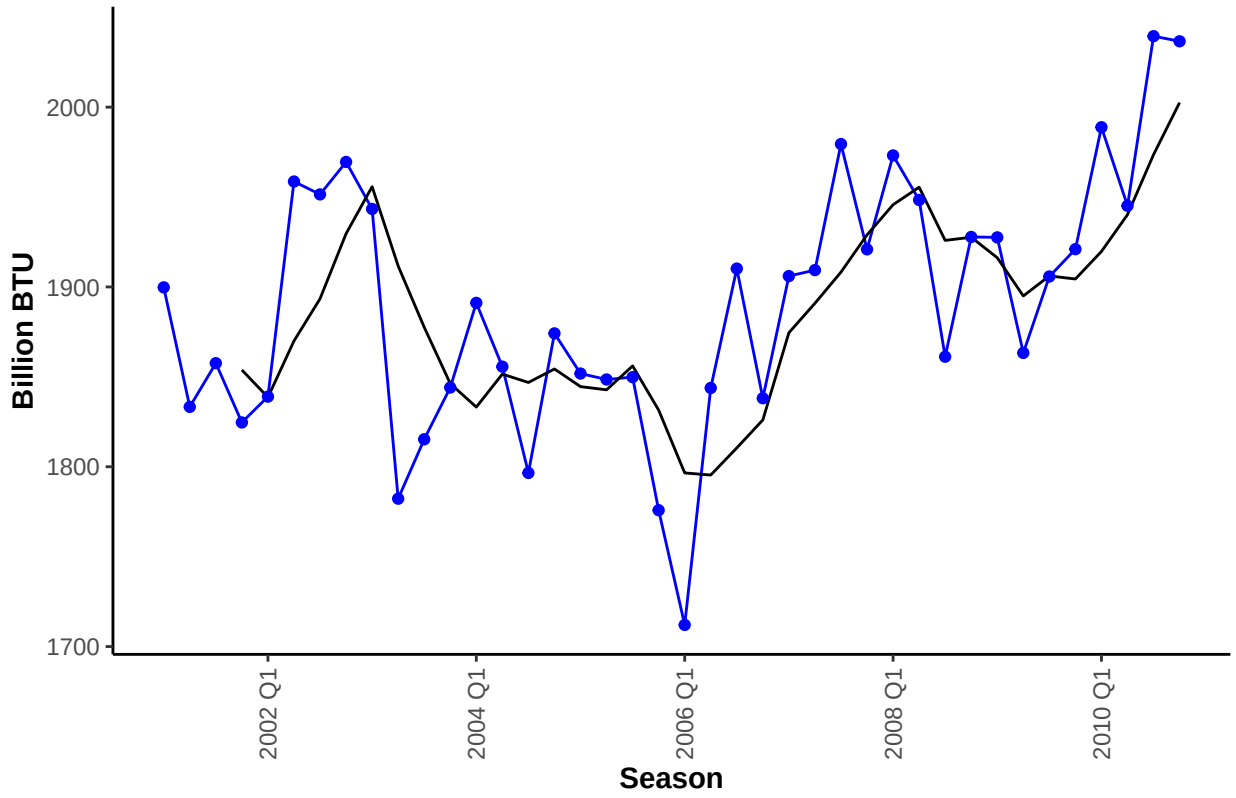

```
##
##  Ljung-Box test
##
## data:  Residuals from ETS(M,A,M)
## Q* = 31.318, df = 8, p-value = 0.0001233
##
## Model df: 16.   Total lags used: 24
```

5.9
Natural Gas Sales : Figure 5.14 shows a time plot of quarterly natural gas sales (in billions of BTU's) of a certain company, over a period of 4 years. The company's analyst is asked to use a moving average model to forecast sales in Winter 2005.

a) Reproduce the time plot with the overlaying MA(4) line.

```
#dataset comes from
# https://fred.stlouisfed.org/series/NATURALGASD11


# Load Data
NaturalGas <- read_csv("Data/NaturalGas.csv",
                       col_types = cols(Quarter = col_date(format = "%m/%d/%Y"),
                                        NaturalGas = col_double()))


# Create Time Series Object
gas <- ts(NaturalGas$NaturalGas, start = c(2001, 1), frequency = 4)


# Create rolling moving average.
gas.ma <- rollmean(gas, k = 4, align = "right")


autoplot(gas, series = "Actual", color = "blue") +
  autolayer(gas.ma , series = "MovingAvg",color = "black") +
  geom_point(color = "blue") +
  theme_classic() +
  labs(x = "Season",
     y = "Billion BTU") +
  theme(axis.title.y = element_text(face="bold"),
        axis.title.x = element_text(face="bold"),
        axis.text.x = element_text(angle = 90,
                                   vjust = 0.5,
                                   hjust=1),
        legend.position = "top",
        legend.direction = "horizontal") +
  zoo::scale_x_yearqtr(format = '%Y Q%q')
```

```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
```



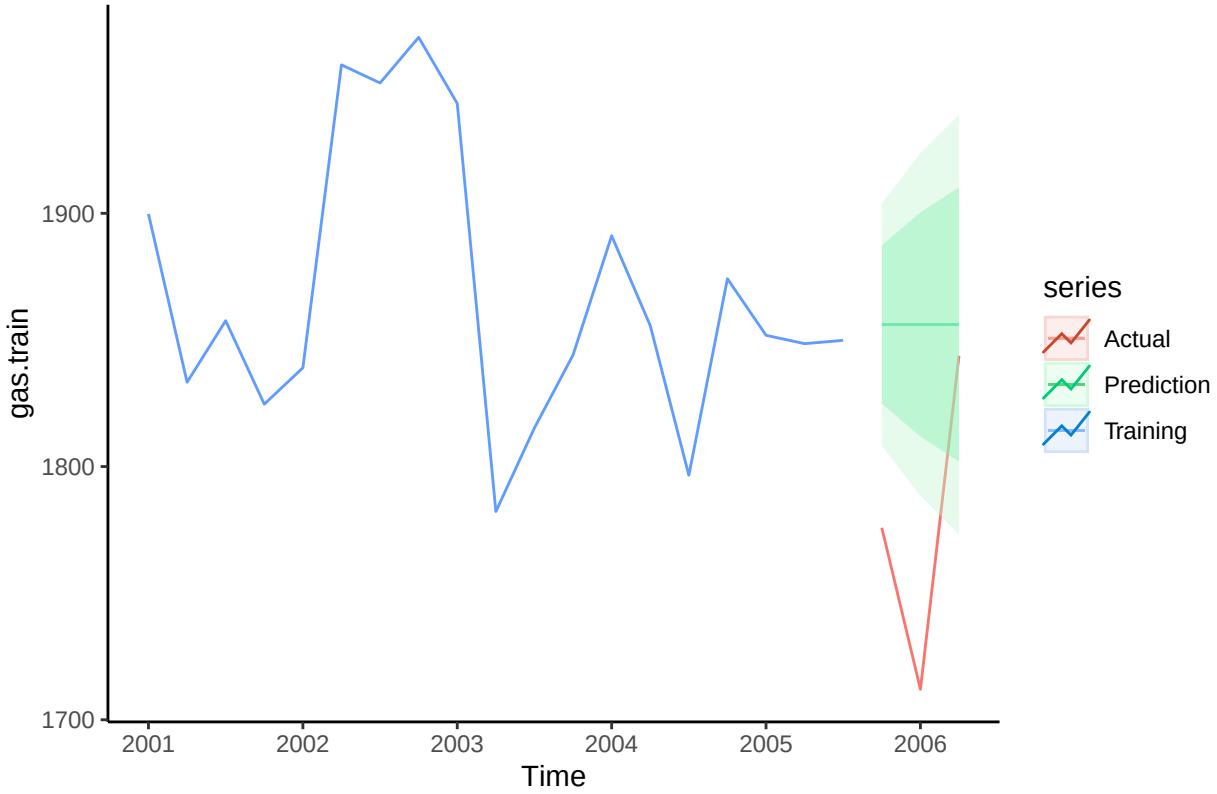
b) What can we learn about the series from the MA line?

Answer There is a declining trend in gas sales during this period

c) Run a moving average forecaster with adequate season length. Are forecasts generated by this method expected to over-forecast, under-forecast, or accurately forecast actual sales? Why?

Answer We end up over-forecasting actual sales. This seems like a shock event happens in the winter of 2005 and early 2006 that is unforeseeable in the data. This could be considered an outlier event.

```
# Test Train Partitioning
gas.train <- window(gas, end = c(2005, 3))
gas.test <- window(gas, start = c(2005, 4), end = c(2006, 2))

# Create rolling moving average.
gas.ma.model <- rollmean(gas.train, k = 4, align = "right")
gas.ma.pred <- forecast(gas.ma.model, h = 3)


# Plot the model forecast.
autoplot(gas.train, series = 'Training') +
  autolayer(gas.test, series = 'Actual') +
  autolayer(gas.ma.pred, series = 'Prediction', alpha = .4) +
  theme_classic()
```


gas.train
1900
1800
1700
2001
2002
2003
2004
2005
2006
Time
series
Actual
Prediction
Training

6.6 Forecasting Australian Wine Sales : Figure 6.26 shows time plots of monthly sales of six types of Australian wines (red, rose, sweet white, dry white, sparkling, and fortified) for 1980-1994. The data is available in AustralianWines.xls. The units are thousands of liters. You are hired to obtain short-term forecasts (2-3 months ahead) for each of the six series, and this task will be repeated monthly

b) Fortified wine has the largest market share of the six types of wine considered. You are asked to focus on fortified wine sales alone and produce as accurate as possible forecasts for the next 2 months. **Answer** See Below

- Start by partitioning the data using the period until December 1993 as the training period. **Answer** See Below
- Fit a regression model to sales with a linear trend and seasonality. **Answer** See Below

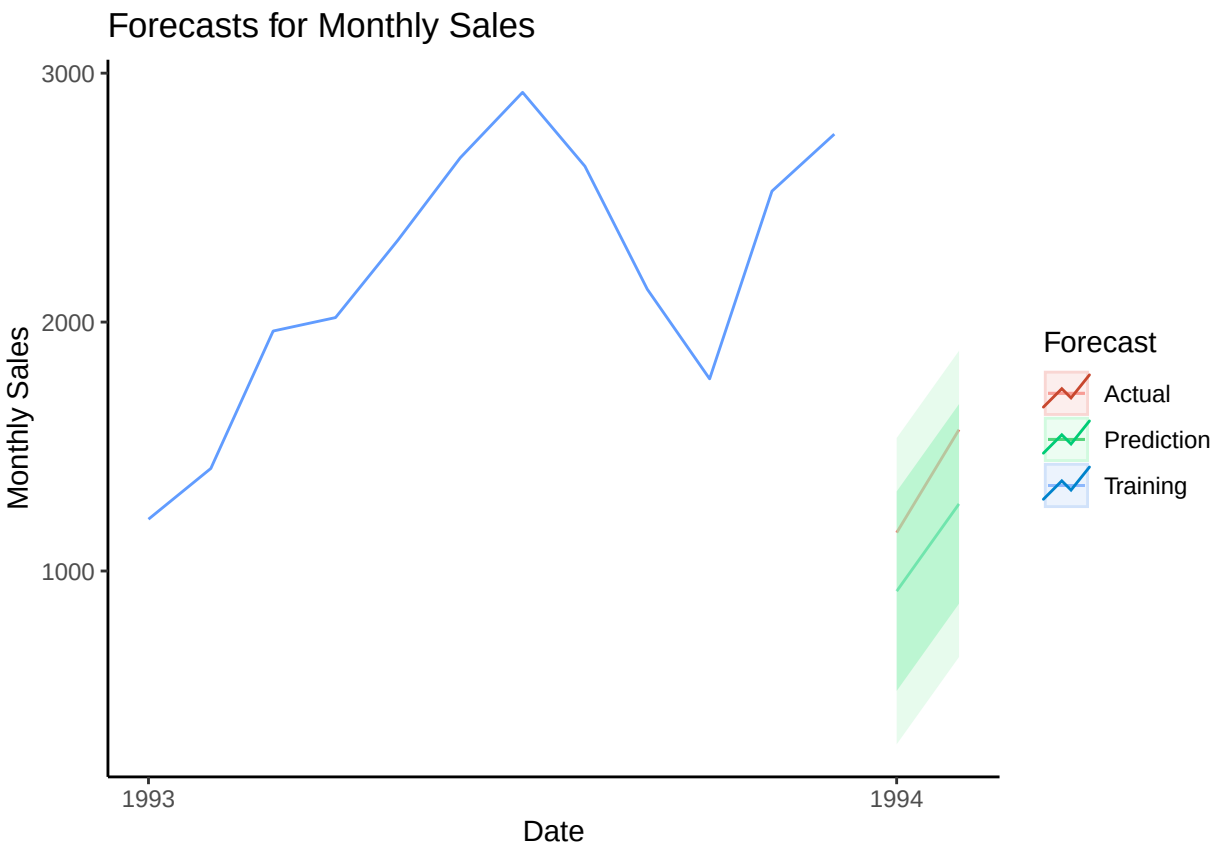
i. Create the "actual vs. forecast" plot. What can you say about model fit? **Answer** The model predict pretty well looking at the "actual vs. forecast" plot we see the actual sales is within the prediction intervals of the model.

ii. Use the regression model to forecast sales in January and February 1994 **Answer** Forecasted Sales for January are 919. Forecasted sales for February are 1270.

```
# Create Time Series Object
Fortified <- ts(AustralianWines$Fortified, start = c(1980, 1), frequency = 12)

# Training and test partitions
Fort.train <- window(Fortified, end = c(1993, 12))
Fort.test <- window(Fortified, start = c(1994, 1), end = c(1994,2))


# Fit a regression model with trend and seasonality.
fort.lm.model <- tslm(Fort.train ~ trend + season, Fort.train)
fort.lm.pred <- forecast(fort.lm.model, h = 2)

fort.lm.pred
```

```
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Jan 1994       918.8819 518.1185 1319.645 303.7771 1533.987
## Feb 1994      1269.5962 868.8328 1670.360 654.4914 1884.701
```

```
# Plot Results
autoplot(window(Fort.train, start = c(1993, 1)), series = "Training") +
  autolayer(Fort.test, series = 'Actual') +
  autolayer(fort.lm.pred, series="Prediction", alpha = .4) +
  xlab("Date") +
  ylab("Monthly Sales") +
  ggtitle("Forecasts for Monthly Sales") +
  theme_classic() +
  guides(colour=guide_legend(title="Forecast"))
```


Forecasts for Monthly Sales
3000
2000
1000
Monthly Sales
1993
1994
Date
Forecast
Actual
Prediction
Training