

ADS-506 Mod 2

Module 2 Assignment Exercises

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

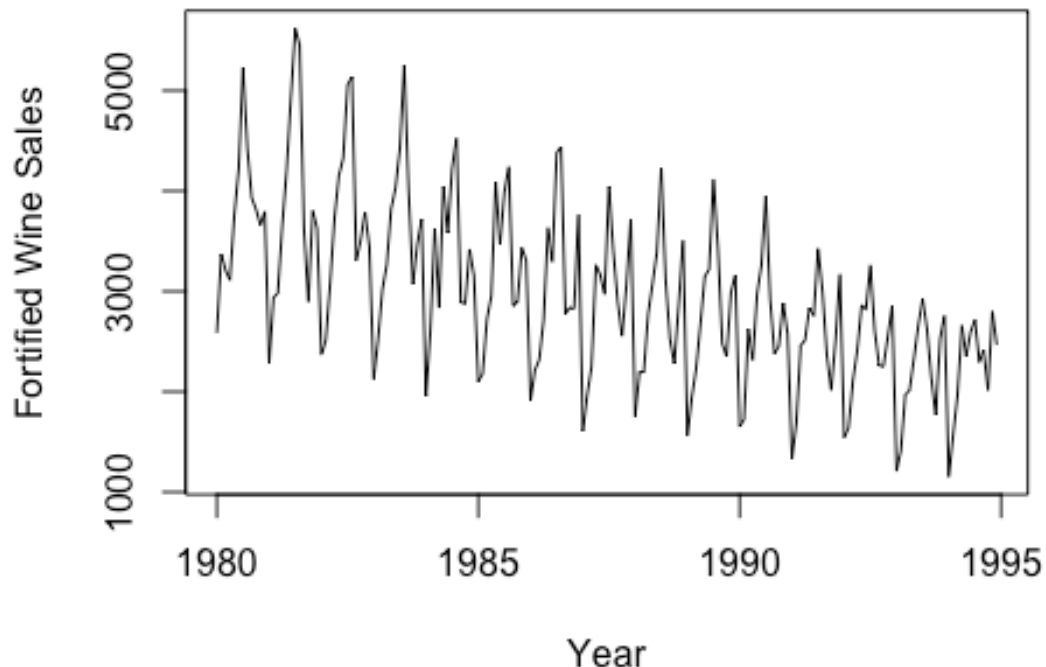
b) Fortified wine has the largest market share of the six types of wine. You are asked to focus on fortified wine sales alone and produce as accurate a forecast as possible for the next two months

• Start by partitioning the data using the period until December 1993 as the training period.

• Apply Holt-Winter's exponential smoothing (with multiplicative seasonality) to sales.

create a time series object for fortified wine sales (tsibble is recommended)

```
aus_wine$Date = as.Date(aus_wine$Month, format = "%m/%d/%Y")
wine_ts <- ts(aus_wine$Fortified, start = c(1980, 1), end = c(1994, 12),
frequency = 12)
plot(wine_ts, ylab="Fortified Wine Sales", xlab="Year")
```



```
# create training/validation sets for the model (include at least one year in
the validation set)
fixed.nValid <- 12
fixed.nTrain <- length(wine_ts) - fixed.nValid
train.ts <- window(wine_ts, start = c(1980, 1), end = c(1980, fixed.nTrain))
valid.ts <- window(wine_ts, start = c(1980, fixed.nTrain + 1),
                  end = c(1980, fixed.nTrain + fixed.nValid))

# fit an ETS model to the training set (try different models to find the
best)
fort.train.model <- ets(train.ts, model = "ANN")

# forecast the validation period
fort.train.pred <- forecast(fort.train.model, h=fixed.nValid)

# review accuracy to select the best model
summary(fort.train.pred)

##
## Forecast method: ETS(A,N,N)
##
## Model Information:
## ETS(A,N,N)
```

```
##
## Call:
## ets(y = train.ts, model = "ANN")
##
## Smoothing parameters:
##   alpha = 0.7546
##
## Initial states:
##   l = 2766.7566
##
##   sigma: 730.6318
##
##      AIC      AICc      BIC
## 3080.368 3080.514 3089.739
##
## Error measures:
##
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.8368409 726.2698 576.4105 -4.497327 21.35642 2.071508
##
##           ACF1
## Training set 0.06622061
##
## Forecasts:
##           Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jan 1994      2660.673 1724.3311 3597.016 1228.66139 4092.685
## Feb 1994      2660.673 1487.6780 3833.669  866.73168 4454.615
## Mar 1994      2660.673 1291.3306 4030.016  566.44433 4754.903
## Apr 1994      2660.673 1119.8032 4201.544  304.11577 5017.231
## May 1994      2660.673  965.5444 4355.802   68.19729 5253.150
## Jun 1994      2660.673  824.1975 4497.149 -147.97412 5469.321
## Jul 1994      2660.673  692.9780 4628.369 -348.65699 5670.004
## Aug 1994      2660.673  569.9782 4751.369 -536.76900 5858.116
## Sep 1994      2660.673  453.8232 4867.524 -714.41273 6035.760
## Oct 1994      2660.673  343.4835 4977.863 -883.16278 6204.510
## Nov 1994      2660.673  238.1643 5083.183 -1044.23461 6365.581
## Dec 1994      2660.673  137.2369 5184.110 -1198.58973 6519.937

accuracy(fort.train.pred, valid.ts)

##
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.8368409 726.2698 576.4105 -4.497327 21.35642 2.071508
## Test set     -410.8400834 626.9061 442.7823 -25.868574 27.01972 1.591274
##
##           ACF1 Theil's U
## Training set 0.06622061      NA
## Test set     0.38717574 1.472012

# Use the full data set to fit the best model
fort.model <- ets(wine_ts, model = "ANN")

# forecast the next two months
fort.pred <- forecast(fort.model, h=2)
```

print the last two months of the forecast (this provides the answer to the question)

```
head(fort.pred)
```

```
## $model
```

```
## ETS(A,N,N)
```

```
##
```

```
## Call:
```

```
## ets(y = wine_ts, model = "ANN")
```

```
##
```

```
## Smoothing parameters:
```

```
## alpha = 0.7303
```

```
##
```

```
## Initial states:
```

```
## l = 2785.8638
```

```
##
```

```
## sigma: 721.0309
```

```
##
```

```
## AIC AICc BIC
```

```
## 3307.767 3307.903 3317.345
```

```
##
```

```
## $mean
```

```
## Jan Feb
```

```
## 1995 2507.318 2507.318
```

```
##
```

```
## $level
```

```
## [1] 80 95
```

```
##
```

```
## $x
```

```
## Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
```

```
## 1980 2585 3368 3210 3111 3756 4216 5225 4426 3932 3816 3661 3795
```

```
## 1981 2285 2934 2985 3646 4198 4935 5618 5454 3624 2898 3802 3629
```

```
## 1982 2369 2511 3079 3728 4151 4326 5054 5138 3310 3508 3790 3446
```

```
## 1983 2127 2523 3017 3265 3822 4027 4420 5255 4009 3074 3465 3718
```

```
## 1984 1954 2604 3626 2836 4042 3584 4225 4523 2892 2876 3420 3159
```

```
## 1985 2101 2181 2724 2954 4092 3470 3990 4239 2855 2897 3433 3307
```

```
## 1986 1914 2214 2320 2714 3633 3295 4377 4442 2774 2840 2828 3758
```

```
## 1987 1610 1968 2248 3262 3164 2972 4041 3402 2898 2555 3056 3717
```

```
## 1988 1755 2193 2198 2777 3076 3389 4231 3118 2524 2280 2862 3502
```

```
## 1989 1558 1940 2226 2676 3145 3224 4117 3446 2482 2349 2986 3163
```

```
## 1990 1651 1725 2622 2316 2976 3263 3951 2917 2380 2458 2883 2579
```

```
## 1991 1330 1686 2457 2514 2834 2757 3425 3006 2369 2017 2507 3168
```

```
## 1992 1545 1643 2112 2415 2862 2822 3260 2606 2264 2250 2545 2856
```

```
## 1993 1208 1412 1964 2018 2329 2660 2923 2626 2132 1772 2526 2755
```

```
## 1994 1154 1568 1965 2659 2354 2592 2714 2294 2416 2016 2799 2467
```

```
##
```

```
## $upper
```

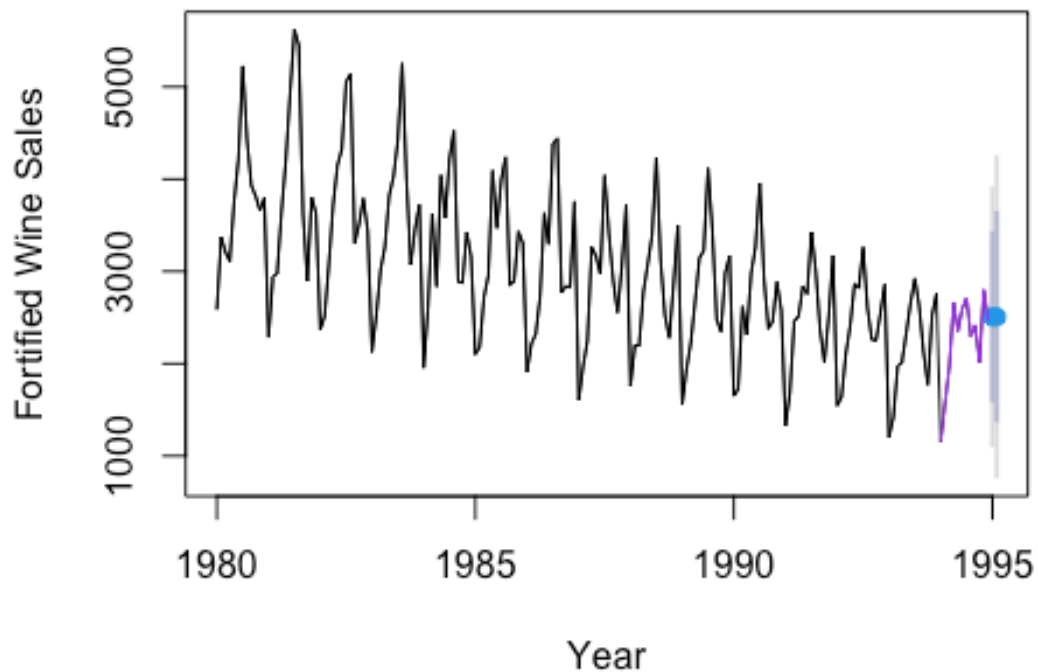
```
## 80% 95%
```

```
## Jan 1995 3431.356 3920.512
```

```
## Feb 1995 3651.513 4257.213
##
## $lower
##           80%           95%
## Jan 1995 1583.280 1094.1235
## Feb 1995 1363.123  757.4226

plot(fort.pred, ylab = "Fortified Wine Sales", xlab = "Year")
  lines(train.ts, col = "black")
  lines(valid.ts, col = "purple")
```

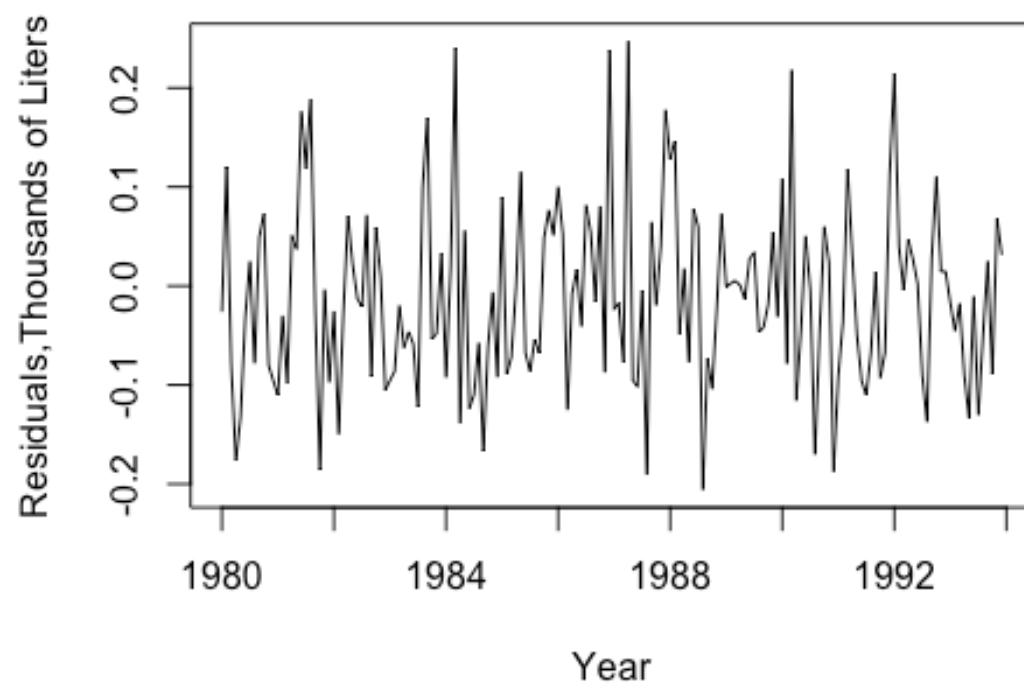
Forecasts from ETS(A,N,N)



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

```
fort.train.holt <- ets(train.ts, model = "MAA") # MAA model to fit Holt-
Winter's exponential smoothing
```

```
plot(fort.train.holt$residuals, xlab = "Year",
      ylab = "Residuals,Thousands of Liters")
```



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.

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

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

head(NaturalGas)

## # A tibble: 6 × 2
##   Quarter    NaturalGas
##   <date>         <dbl>
## 1 2001-01-01     1900.
## 2 2001-04-01     1833.
## 3 2001-07-01     1858.
## 4 2001-10-01     1825.
## 5 2002-01-01     1839
## 6 2002-04-01     1959.

# a) Reproduce the time plot with the overlying MA(4) Line.

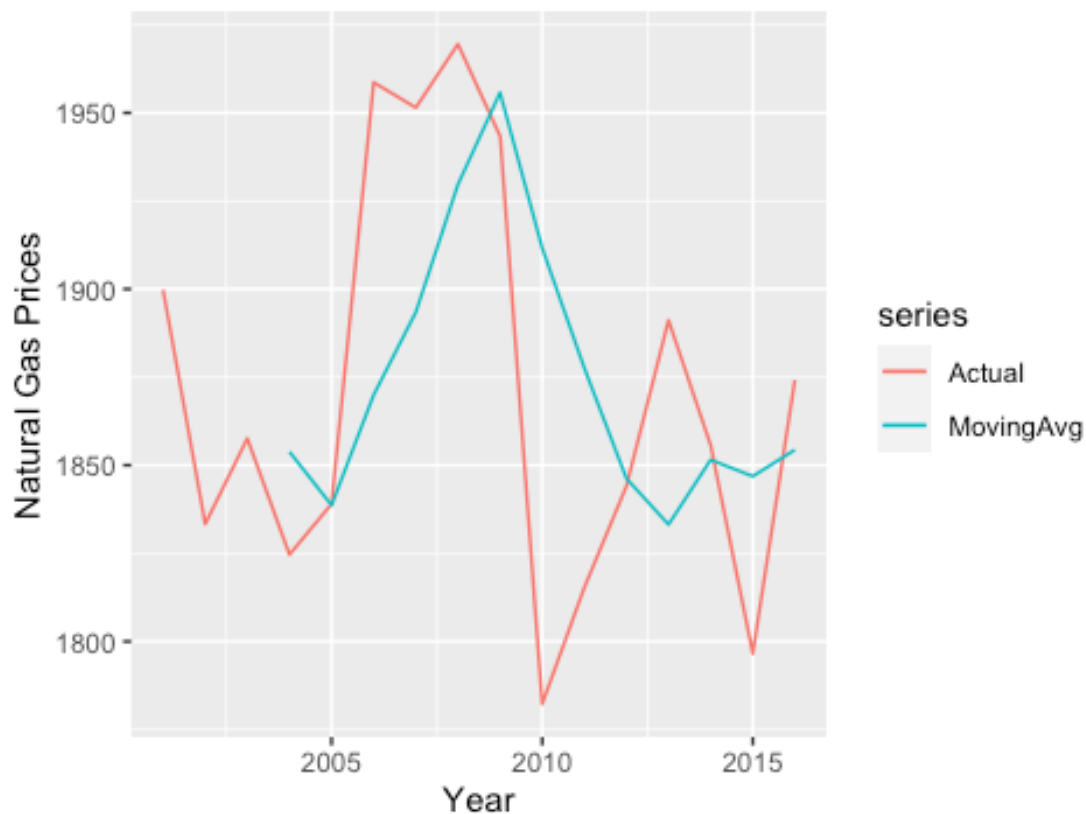
start_date = as.Date("2001-01-01")
end_date = as.Date("2004-10-01")

ng_df <- NaturalGas %>%
  filter(Quarter >= start_date, Quarter <= end_date)
head(ng_df)

## # A tibble: 6 × 2
##   Quarter    NaturalGas
##   <date>         <dbl>
## 1 2001-01-01     1900.
## 2 2001-04-01     1833.
## 3 2001-07-01     1858.
## 4 2001-10-01     1825.
## 5 2002-01-01     1839
## 6 2002-04-01     1959.

natural_gas <- ts(ng_df$NaturalGas, start = c(2001, 1), frequency = 1)
ng.ma <- rollmean(natural_gas, k = 4, align = "right")

autoplot(natural_gas, series = "Actual", ylab = "Natural Gas Prices", xlab =
"Year") +
  autolayer(ng.ma, series = "MovingAvg")
```



- b) What can we learn about the series from the MA line? There is a declining trend in gas sales during this period

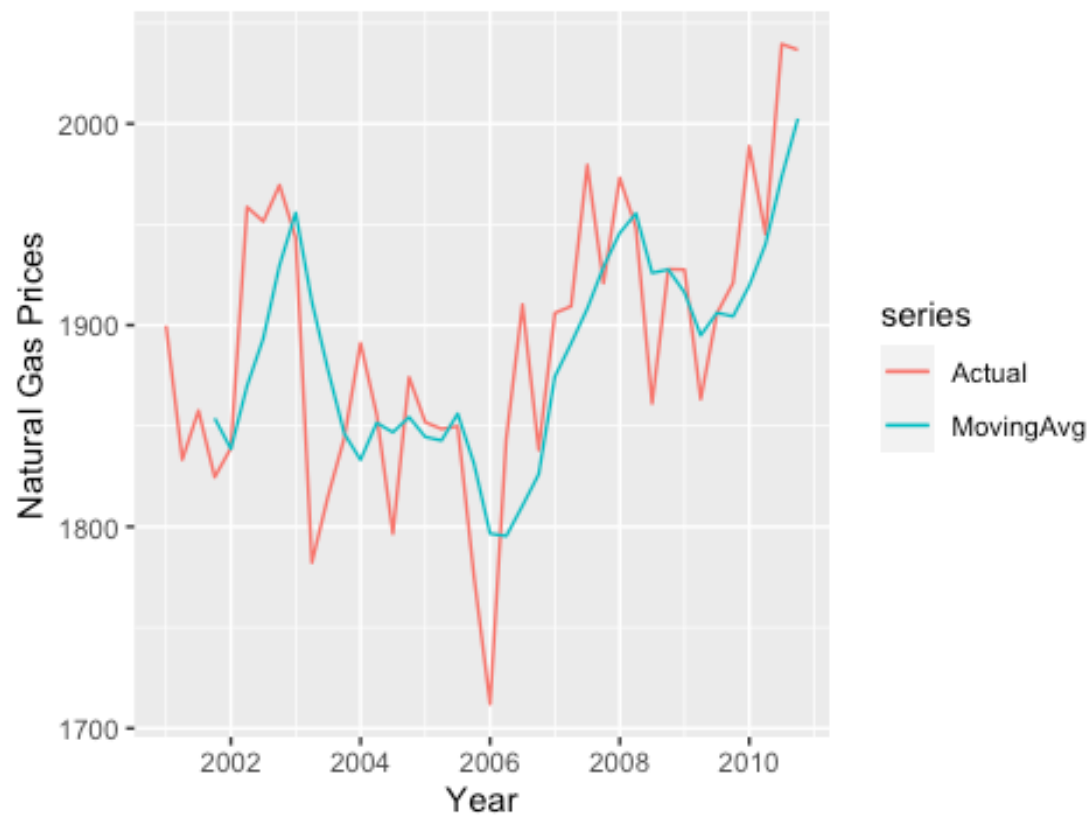
The moving average line reveals that there is a downward trend on fig 5.14. This suggest declining gas prices over 2001 to 2004.

- 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?

With the data points set at 4 and the season length ran throughout the dataset, the forecast seems to be under-forecasting peaks and over forecasting dips compared to the actual sales so the seasonal pattern is smooth out. It does seem to forecast the general cyclical trend relatively accurate.

```
nat_gas <- ts(NaturalGas$NaturalGas, start = c(2001, 1), frequency = 4)
natg.ma <- rollmean(nat_gas, k = 4, align = "right")

autoplot(nat_gas, series = "Actual", ylab = "Natural Gas Prices", xlab =
"Year") +
  autolayer(natg.ma, series = "MovingAvg")
```

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.
 - Start by partitioning the data using the period until December 1993 as the training period.
 - Fit a regression model to sales with a linear trend and seasonality.
- i. Create the “actual vs. forecast” plot. What can you say about model fit?
- ii. Use the regression model to forecast sales in January and February 1994

Use the code from 5.8 as a template for this question

```
# Data Frame:
#wine_ts

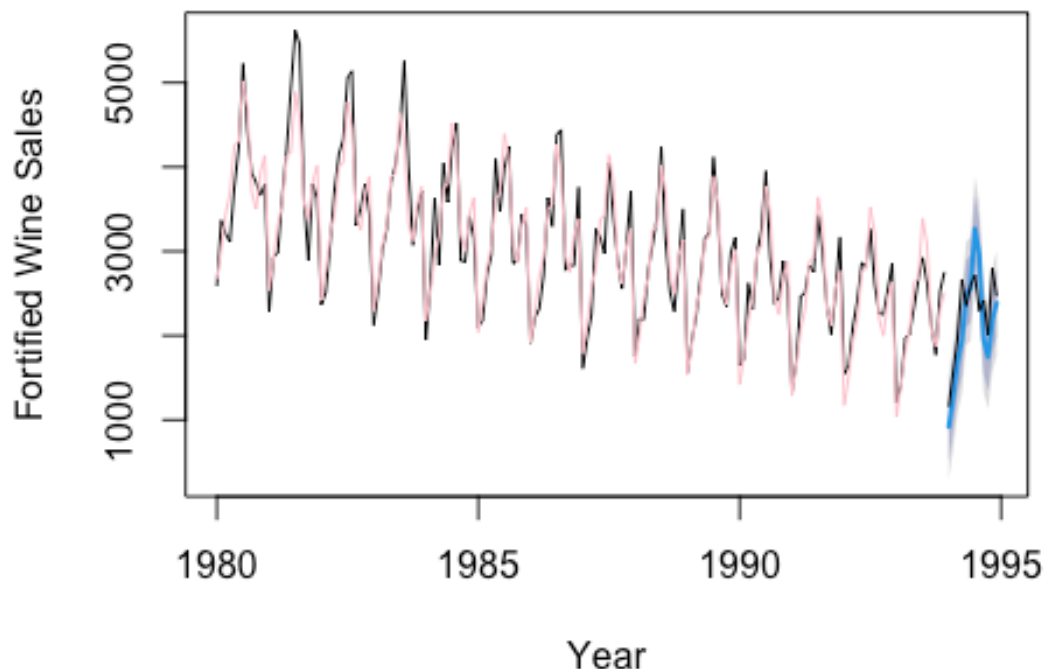
# Partitioned data set:
#train.ts
#valid.ts

# Fit regression model:
fort.lmt.model <- tslm(train.ts ~ trend + season)

# Prediction:
fort.lmt.pred <- forecast(fort.lmt.model, h = 12)

# Plot fitted trend and seasonal predicted line:
plot(fort.lmt.pred, ylab = "Fortified Wine Sales", xlab = "Year")
  lines(fort.lmt.pred$fitted, col = "pink")
  lines(valid.ts, col = "black")
```

Forecasts from Linear regression model



The model fit in pink curve follows the same downward trend as the training data with the seasonal spikes slightly under projected at the beginning and becomes slightly over projected as the dips as well as peaks progress in the later years leading to 1995.

Forecast from Jan to Dec 1994:

```
head(fort.lmt.pred)
```

```
## $model
```

```
##
```

```
## Call:
```

```
## tslm(formula = train.ts ~ trend + season)
```

```
##
```

```
## Coefficients:
```

```
## (Intercept)      trend    season2    season3    season4
```

```
season5
```

```
##      2679.65      -10.42      361.13      791.98      1048.40
```

```
1619.39
```

```
##      season6      season7      season8      season9      season10
```

```
season11
```

```
##      1691.24      2410.01      2116.22      1115.71      922.20
```

```
1373.62
```

```

##      season12
##      1583.18
##
##
## $mean
##      Jan      Feb      Mar      Apr      May      Jun      Jul
## 1994  918.8819 1269.5962 1690.0247 1936.0247 2496.5962 2558.0247 3266.3819
##      Aug      Sep      Oct      Nov      Dec
## 1994 2962.1676 1951.2390 1747.3104 2188.3104 2387.4533
##
## $lower
##      [,1]      [,2]
## Jan 1994  518.1185  303.7771
## Feb 1994  868.8328  654.4914
## Mar 1994 1289.2614 1074.9199
## Apr 1994 1535.2614 1320.9199
## May 1994 2095.8328 1881.4914
## Jun 1994 2157.2614 1942.9199
## Jul 1994 2865.6185 2651.2771
## Aug 1994 2561.4042 2347.0628
## Sep 1994 1550.4757 1336.1342
## Oct 1994 1346.5471 1132.2056
## Nov 1994 1787.5471 1573.2056
## Dec 1994 1986.6899 1772.3485
##
## $upper
##      [,1]      [,2]
## Jan 1994 1319.645 1533.987
## Feb 1994 1670.360 1884.701
## Mar 1994 2090.788 2305.130
## Apr 1994 2336.788 2551.130
## May 1994 2897.360 3111.701
## Jun 1994 2958.788 3173.130
## Jul 1994 3667.145 3881.487
## Aug 1994 3362.931 3577.272
## Sep 1994 2352.002 2566.344
## Oct 1994 2148.074 2362.415
## Nov 1994 2589.074 2803.415
## Dec 1994 2788.217 3002.558
##
## $level
## [1] 80 95
##
## $x
##      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
## 1980 2585 3368 3210 3111 3756 4216 5225 4426 3932 3816 3661 3795
## 1981 2285 2934 2985 3646 4198 4935 5618 5454 3624 2898 3802 3629
## 1982 2369 2511 3079 3728 4151 4326 5054 5138 3310 3508 3790 3446
## 1983 2127 2523 3017 3265 3822 4027 4420 5255 4009 3074 3465 3718
## 1984 1954 2604 3626 2836 4042 3584 4225 4523 2892 2876 3420 3159

```

```
## 1985 2101 2181 2724 2954 4092 3470 3990 4239 2855 2897 3433 3307
## 1986 1914 2214 2320 2714 3633 3295 4377 4442 2774 2840 2828 3758
## 1987 1610 1968 2248 3262 3164 2972 4041 3402 2898 2555 3056 3717
## 1988 1755 2193 2198 2777 3076 3389 4231 3118 2524 2280 2862 3502
## 1989 1558 1940 2226 2676 3145 3224 4117 3446 2482 2349 2986 3163
## 1990 1651 1725 2622 2316 2976 3263 3951 2917 2380 2458 2883 2579
## 1991 1330 1686 2457 2514 2834 2757 3425 3006 2369 2017 2507 3168
## 1992 1545 1643 2112 2415 2862 2822 3260 2606 2264 2250 2545 2856
## 1993 1208 1412 1964 2018 2329 2660 2923 2626 2132 1772 2526 2755
```

Accuracy:

```
accuracy(fort.lmt.pred, valid.ts)
```

```
##                               ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -1.353415e-14 286.2136 222.371 -0.5182601 7.20901 0.7991585
## Test set      1.354991e+02 426.5323 362.690 6.9974184 16.25365 1.3034381
##                               ACF1 Theil's U
## Training set 0.1593305      NA
## Test set      0.2161928 0.9317167
```

```
summary(fort.lmt.pred)
```

```
##
## Forecast method: Linear regression model
##
## Model Information:
##
## Call:
## tslm(formula = train.ts ~ trend + season)
##
## Coefficients:
## (Intercept)      trend      season2      season3      season4
##      2679.65      -10.42      361.13      791.98      1048.40
##      1619.39
##      season6      season7      season8      season9      season10
##      1691.24      2410.01      2116.22      1115.71      922.20
##      1373.62
##      season12
##      1583.18
##
##
## Error measures:
##                               ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -1.353415e-14 286.2136 222.371 -0.5182601 7.20901 0.7991585
##                               ACF1
## Training set 0.1593305
##
## Forecasts:
##      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
## Mar 1994	1690.0247	1289.2614	2090.788	1074.9199	2305.130
## Apr 1994	1936.0247	1535.2614	2336.788	1320.9199	2551.130
## May 1994	2496.5962	2095.8328	2897.360	1881.4914	3111.701
## Jun 1994	2558.0247	2157.2614	2958.788	1942.9199	3173.130
## Jul 1994	3266.3819	2865.6185	3667.145	2651.2771	3881.487
## Aug 1994	2962.1676	2561.4042	3362.931	2347.0628	3577.272
## Sep 1994	1951.2390	1550.4757	2352.002	1336.1342	2566.344
## Oct 1994	1747.3104	1346.5471	2148.074	1132.2056	2362.415
## Nov 1994	2188.3104	1787.5471	2589.074	1573.2056	2803.415
## Dec 1994	2387.4533	1986.6899	2788.217	1772.3485	3002.558