STAT 443: Lab 1

Alex Grinius(20712972)

xx January, 2022

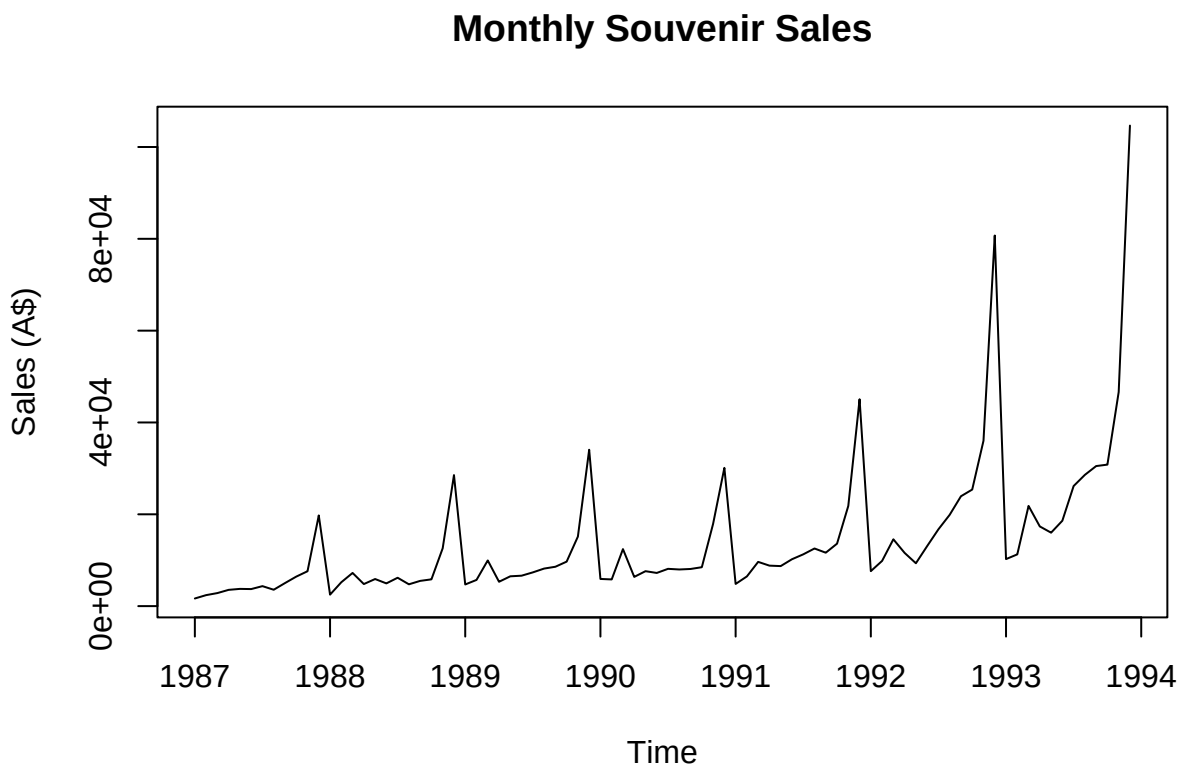
Question 1

(a)

```
souvenir_sales_data <- read.table(file = "souvenir.txt", header = FALSE, sep = "", colClasses = "numeri
```

(b)

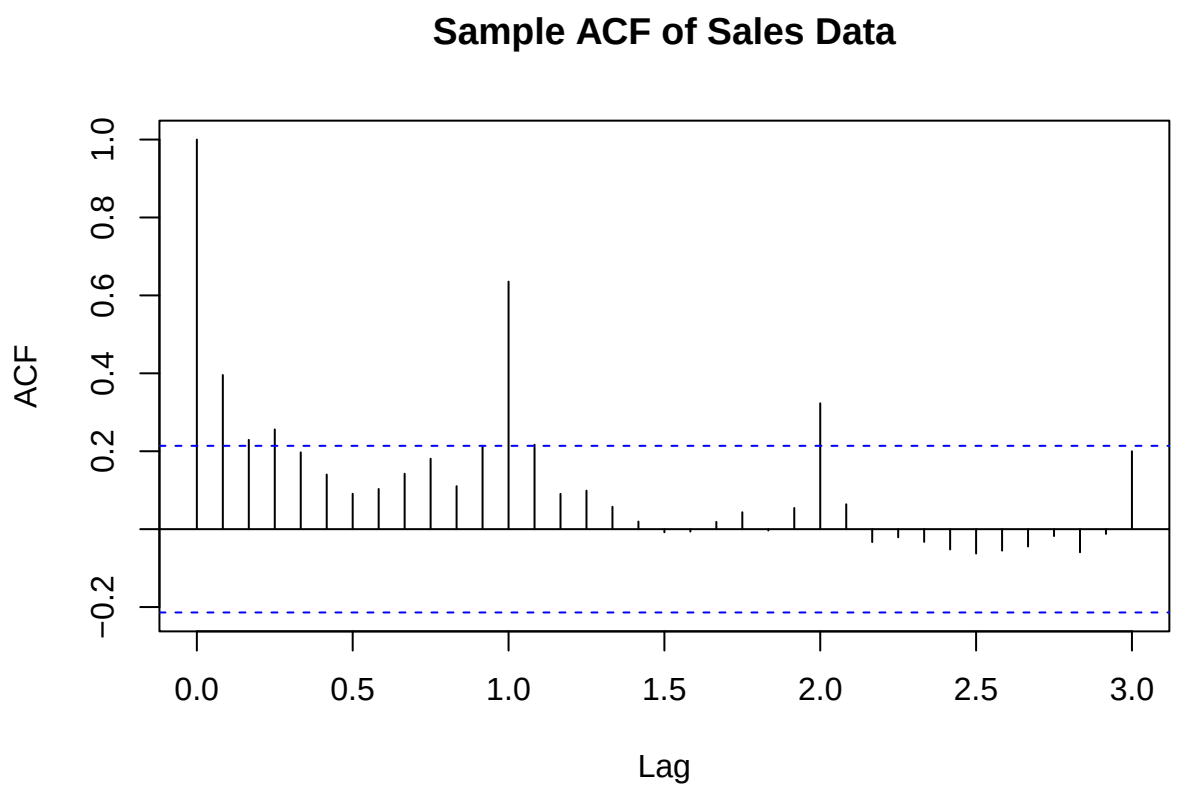
```
sales_time_series <- ts(souvenir_sales_data, start = c(1987, 1), frequency = 12)
plot.ts(sales_time_series, main = "Monthly Souvenir Sales", ylab = "Sales (A$)")
```



This graph shows clear seasonality with it's repeated peaks. It appears to be multiplicative since the peaks increase overtime, but the baseline remains about the same

(c)

```
acf(sales_time_series, main = "Sample ACF of Sales Data", lag.max = 36)
```



Seasonality is shown again with sinusoidal correlations, decaying with time.

Question 2

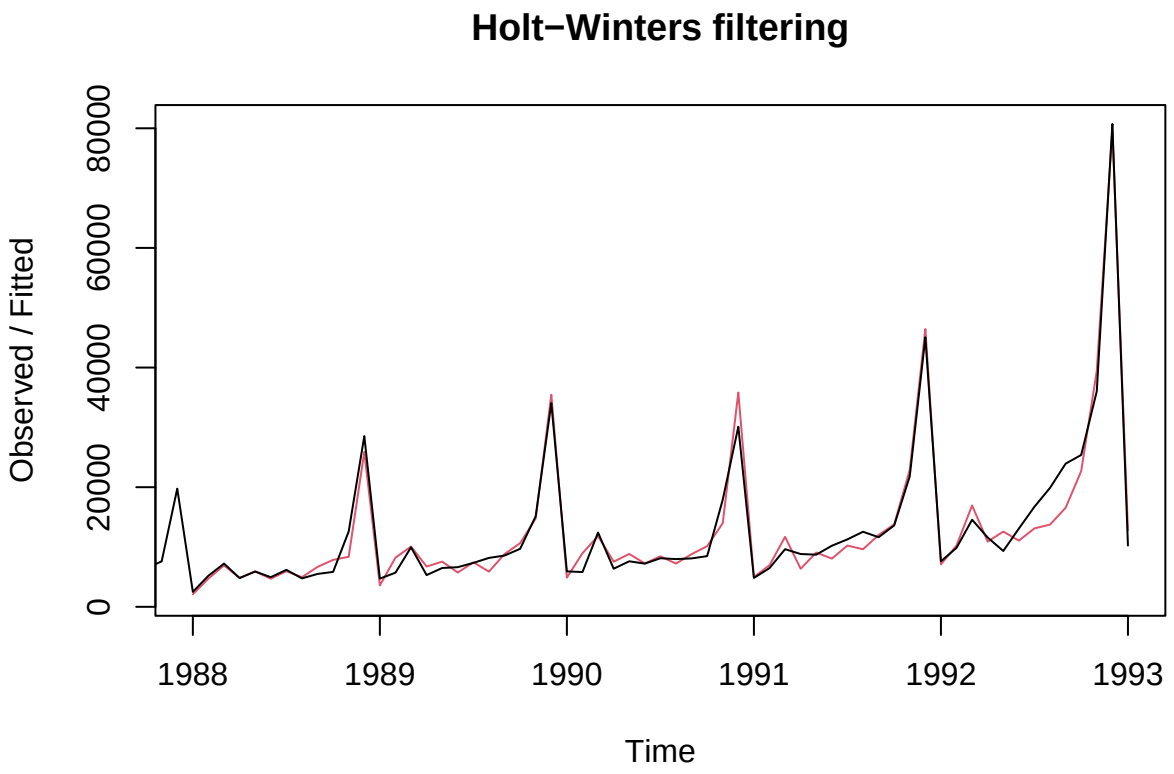
(a)

```
training_data <- window(sales_time_series, start = c(1987, 1), end = c(1993, 1))
holt_winters_model <- HoltWinters(training_data, seasonal = "multiplicative")
holt_winters_model$coefficients
```

```
##            a            b           s1           s2           s3           s4
## 2.579865e+04 4.953699e+02 6.159633e-01 9.869614e-01 7.026077e-01 7.061663e-01
##           s5           s6           s7           s8           s9          s10
## 7.761045e-01 8.690603e-01 8.417587e-01 8.517196e-01 9.110375e-01 1.398782e+00
##          s11          s12
## 2.943041e+00 4.312388e-01
```

(b)

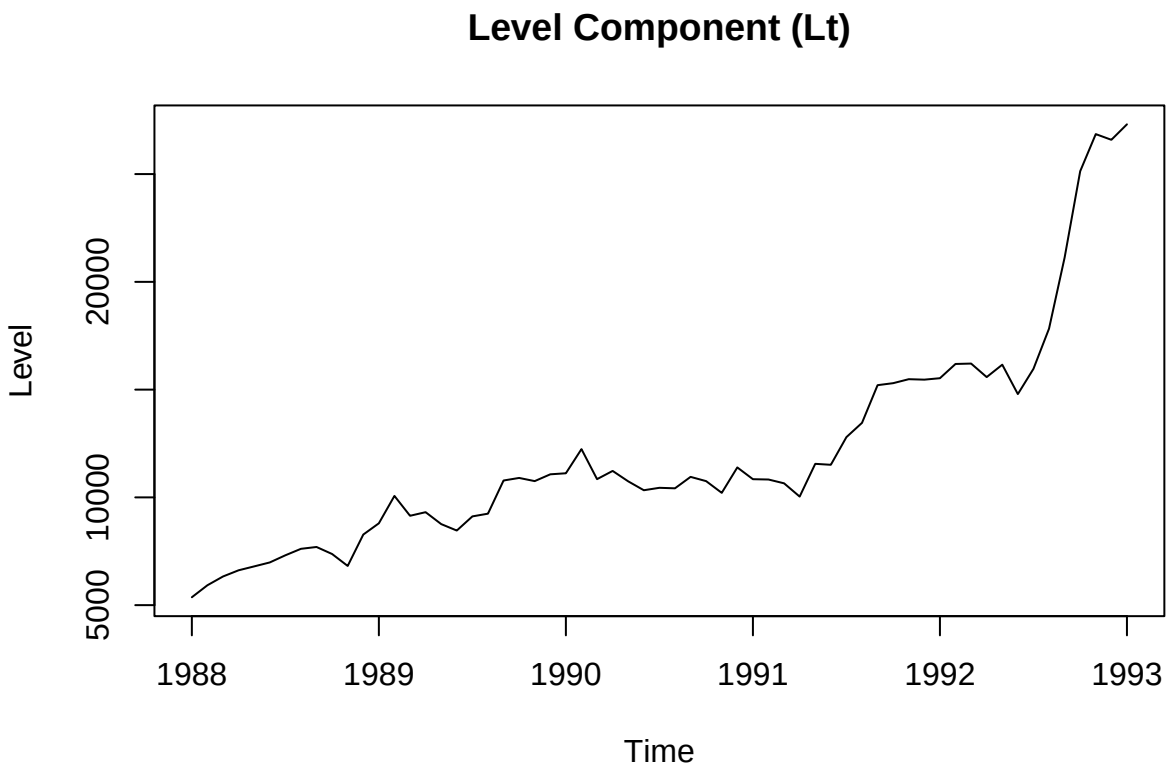
```r
plot(holt_winters_model)
```



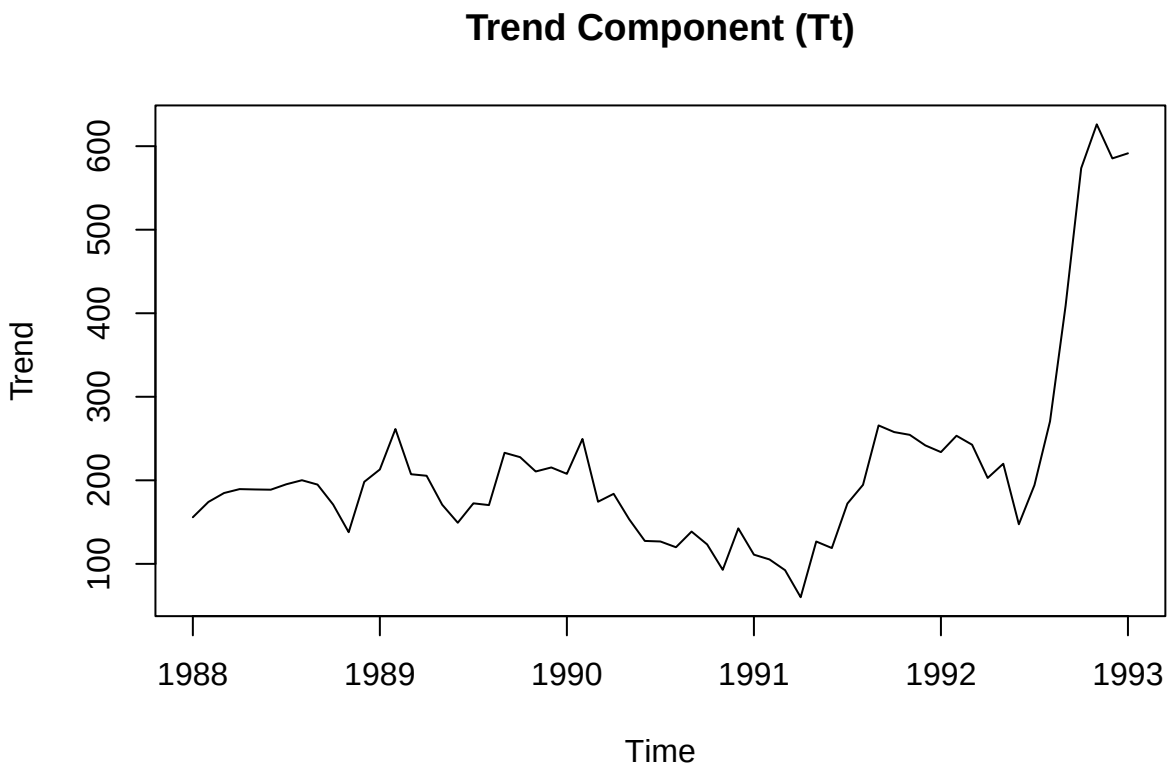
Question 3

(a)

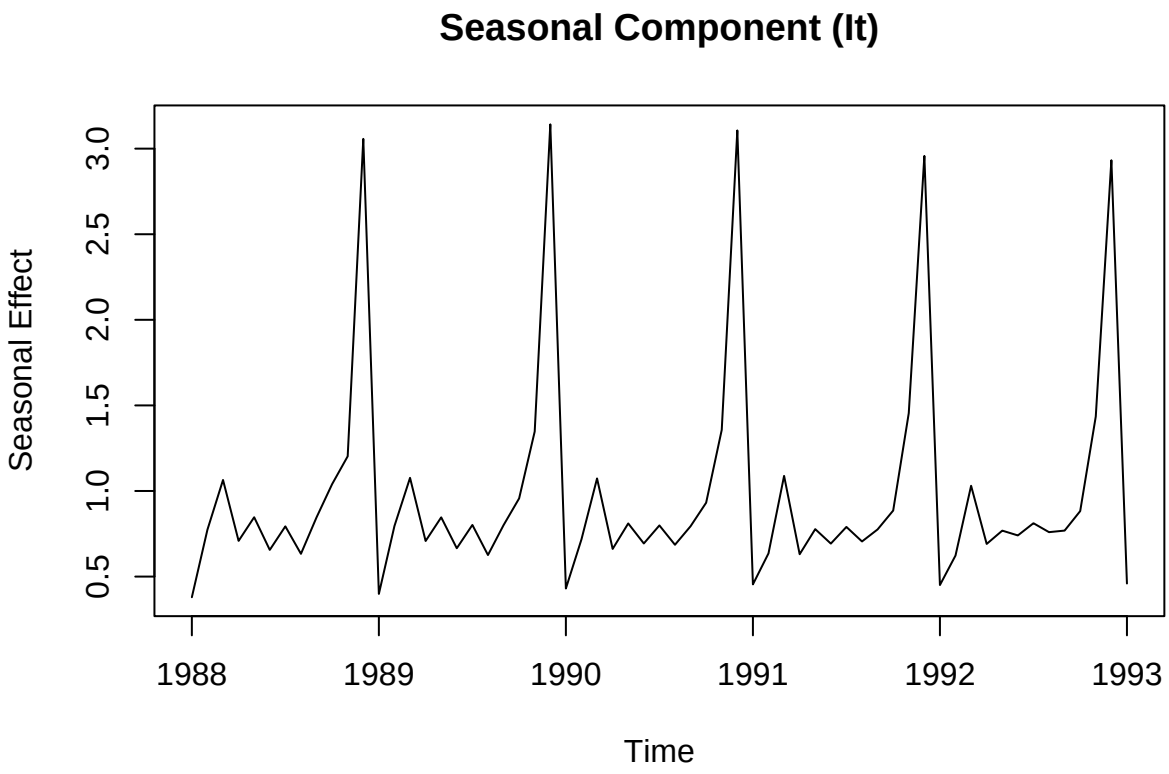
```r
fitted_values <- fitted(holt_winters_model)
plot.ts(fitted_values[, "level"], main = "Level Component (Lt)", ylab = "Level")
```




```
plot.ts(fitted_values[, "trend"], main = "Trend Component (Tt)", ylab = "Trend")
```



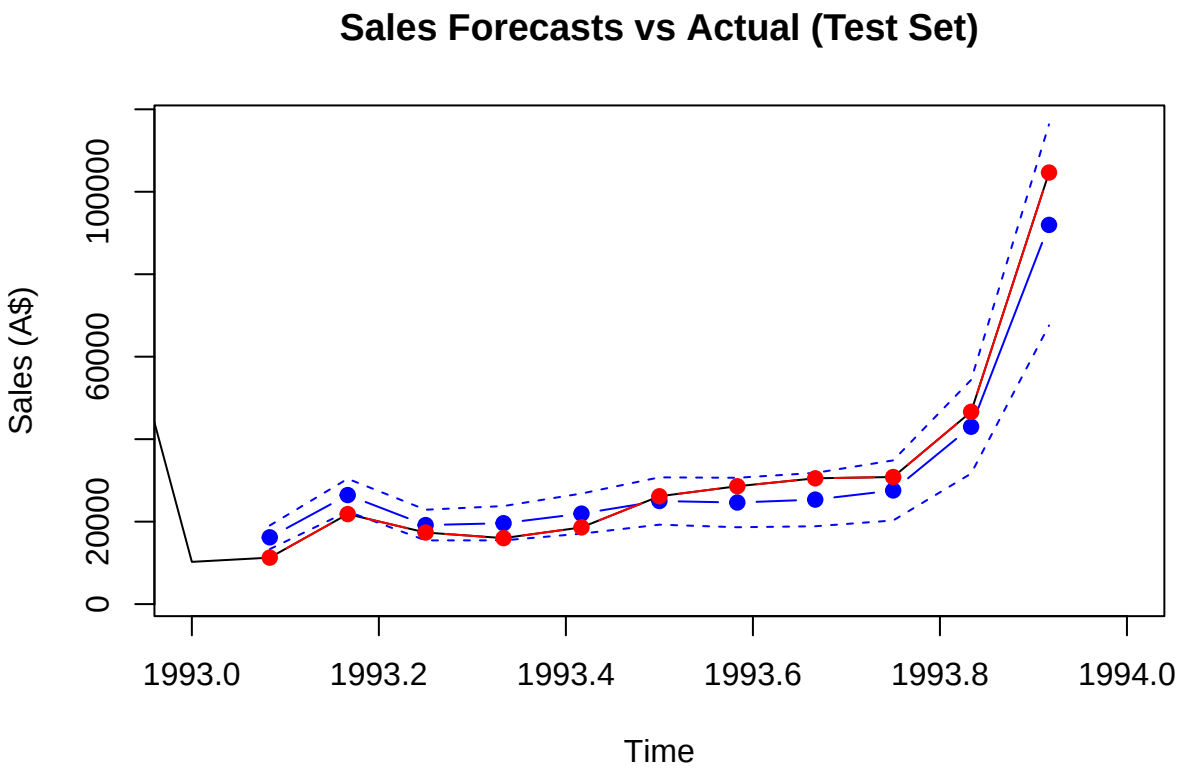

```
plot.ts(fitted_values[, "season"], main = "Seasonal Component (It)", ylab = "Seasonal Effect")
```



Question 4

```
test_data <- window(sales_time_series, start = c(1993, 2), end = c(1993, 12))
predictions <- predict(holt_winters_model, n.ahead = 11, prediction.interval = TRUE, level = 0.95)
```

```
plot.ts(sales_time_series, xlim = c(1993, 1994),
        ylim = c(min(sales_time_series, predictions), max(sales_time_series, predictions)),
        main = "Sales Forecasts vs Actual (Test Set)", ylab = "Sales (A$)")
lines(ts(predictions[, "fit"], start = c(1993, 2), frequency = 12), col = "blue", lty = 1, type = "b", p
lines(ts(predictions[, "upr"], start = c(1993, 2), frequency = 12), col = "blue", lty = 2)
lines(ts(predictions[, "lwr"], start = c(1993, 2), frequency = 12), col = "blue", lty = 2)
lines(test_data, col = "red", lty = 1, type = "b", pch = 19)
```



The forecast is pretty accurate, the predicted upper and lower bounds contain the majority of the actual data. The first 4 months do sit right on the lower bound though.

Question 5

```
forecast_values_Feb_Apr <- predictions[1:3, "fit"]
print("Forecast values for February, March, and April 1993:")
```

```
## [1] "Forecast values for February, March, and April 1993:"
```

```
print(forecast_values_Feb_Apr)
```

```
## [1] 16196.15 26440.10 19170.49
```

Question 6

The seasonality appears to be multiplicative (q1), I would consider applying a log transform to the time series. The log transform would aid in bringing all the data closer together, and reduce the magnitude of variance in the later seasons.