Assignment 1 Solutions

Jonathan Thong

2023-03-28

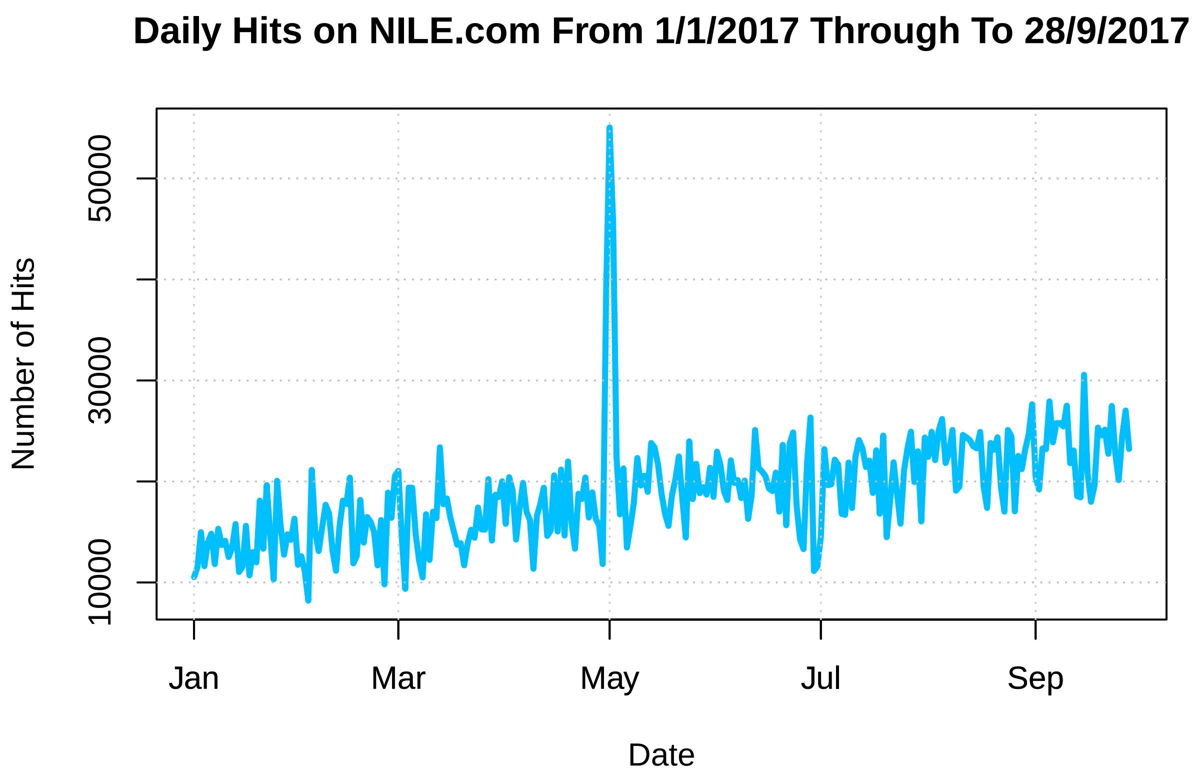
Question 1

You are an analyst working for NILE.COM, a successful online bookseller. You have been tasked with monitoring and forecasting the number of hits (i.e., visits) per day to its website. The file nile.csv on the LMS contains daily hits data for the period 1/1/2017 through to 28/9/2017. You are required to compute all your estimations and plots in R.

a.) Generate an appropriate plot of data and provide a brief description of the observed time series. Make sure to point out any notable visual features/characteristics of the data. (1 Mark)

Let's import and plot the data. The plot should be appropriately scaled with all crucial elements (i.e., axes, title, etc.) correctly labeled (0.5 Marks):

```r
rm(list=ls())

data <- read.csv("nile.csv")

date <- seq(as.Date("2017/1/1"), as.Date("2017/9/28"), by = "day")

data$date <- date

T <- length(date)

time <- seq(1,T)

plot(date,data$Hits,
     main = "Daily Hits on NILE.com From 1/1/2017 Through To 28/9/2017",
     xlab = "Date",
     ylab = "Number of Hits",
     col = "deepskyblue",
     type ="l",
     lwd = 3)
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
```



The data appears to have an upward trend with a large spike in hits occuring around the start of May. (0.5 Marks)

b.) Fit and assess the linear, quadratic and exponential trend models to the data. In your analysis, make sure to include the estimation results as well as plots of the fitted trends that you have generated. Given your results. which trend model would you choose as your forecasting model? (1 Marks)

Let's first estimate the linear trend model, report the estimation results and plot the fitted trend. The plot should be appropriately scaled with all crucial elements (i.e., axes, title, legend etc.) correctly labeled (0.25 Marks).

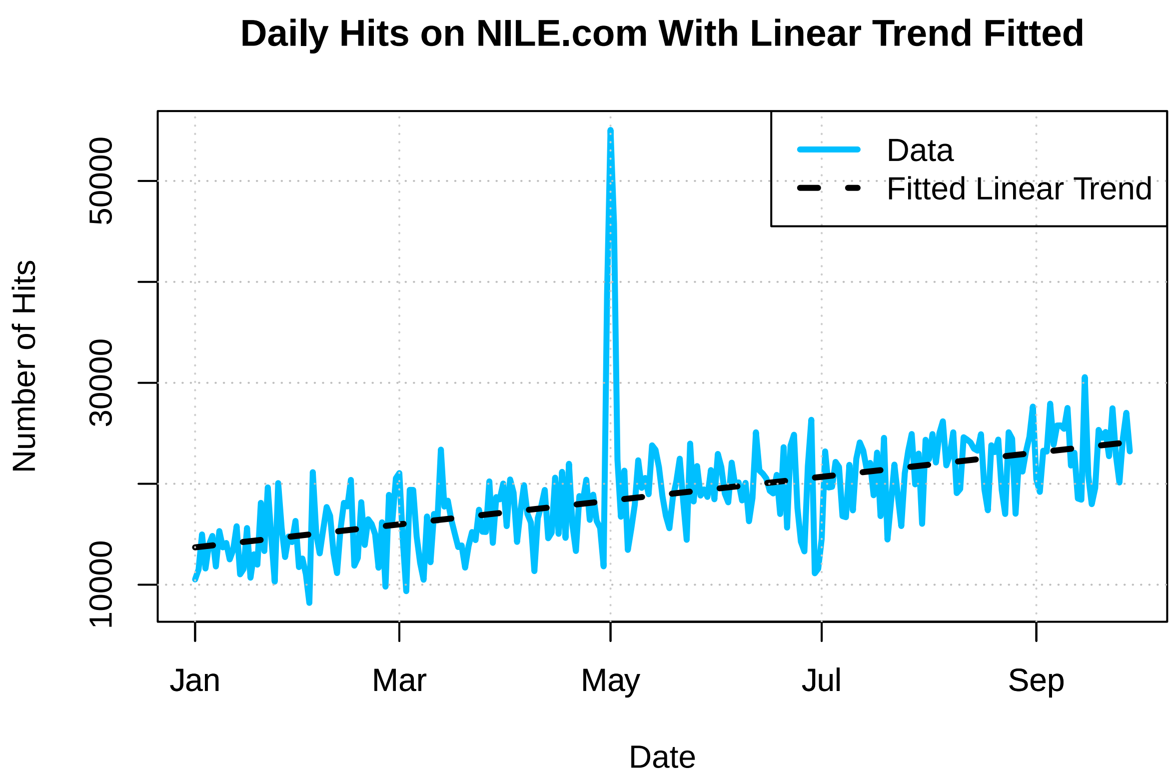
Note: All the relevant objects in our chart need to be clearly labeled. If there is no legend provided in the plot, the fitted trend must be mentioned in the title. Otherwise deduct (0.05 Marks).

```
trendmod1 <- lm(formula = data$Hits ~ time) # This estimates the linear trend model

summary(trendmod1)
```

```
##
## Call:
## lm(formula = data$Hits ~ time)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -9480  -2465    105   1853  36714
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13663.46     527.20   25.92   <2e-16 ***
## time           38.62       3.36   11.49   <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4327 on 269 degrees of freedom
## Multiple R-squared:  0.3293, Adjusted R-squared:  0.3268
## F-statistic: 132.1 on 1 and 269 DF,  p-value: < 2.2e-16
```

```
plot(date,data$Hits,
     main = "Daily Hits on NILE.com With Linear Trend Fitted",
     xlab = "Date",
     ylab = "Number of Hits",
     col = "deepskyblue",
     type ="l",
     lwd = 3)
lines(date, trendmod1$fitted.values, lwd = 3, lty = "dashed")
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
legend(x = "topright",
       legend = c("Data", "Fitted Linear Trend"),
       lty = c("solid", "dashed"),
       lwd = 3.0,
       col = c("deepskyblue","black"))
```



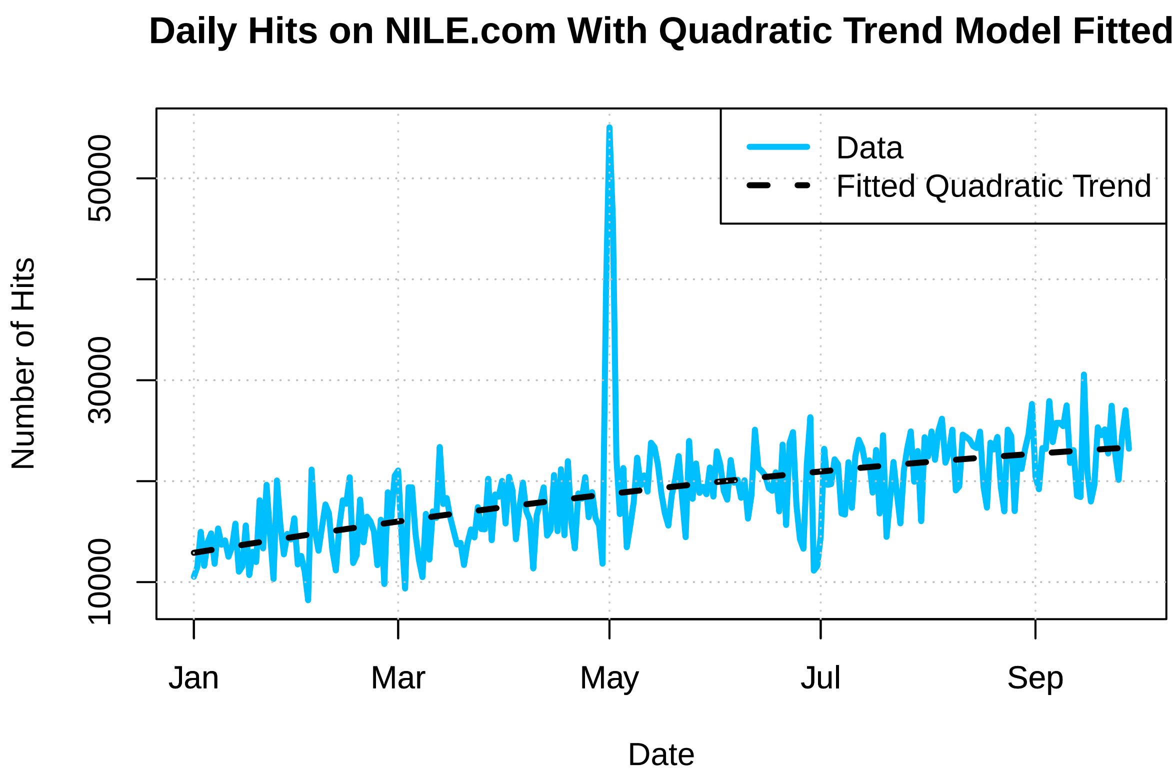
Then, we estimate the quadratic trend model, report the estimation results and plot the fitted trend. The plot should be appropriately scaled with all crucial elements (i.e., axes, title, legend etc.) correctly labeled (0.25 Marks).

Note: All the relevant objects in our chart need to be clearly labeled. If there is no legend provided in the plot, the fitted trend must be mentioned in the title. Otherwise deduct (0.05 Marks).

```r
timesq <- time^2

trendmod2 <- lm(formula = data$Hits ~ time + timesq) # This estimates the quadratic trend model

summary(trendmod2)
```

```
##
## Call:
## lm(formula = data$Hits ~ time + timesq)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -9756  -2359     61   1963  36325
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.285e+04  7.932e+02  16.198  < 2e-16 ***
## time         5.655e+01  1.347e+01   4.199 3.65e-05 ***
## timesq      -6.592e-02  4.794e-02  -1.375     0.17
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4320 on 268 degrees of freedom
## Multiple R-squared:  0.334,  Adjusted R-squared:  0.329
## F-statistic: 67.21 on 2 and 268 DF,  p-value: < 2.2e-16
```

```r
plot(date,data$Hits,
     main = "Daily Hits on NILE.com With Quadratic Trend Model Fitted",
     xlab = "Date",
     ylab = "Number of Hits",
     col = "deepskyblue",
     type ="l",
     lwd = 3)
lines(date, trendmod2$fitted.values, lwd = 3, lty = "dashed")
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
legend(x = "topright",
       legend = c("Data", "Fitted Quadratic Trend"),
       lty = c("solid", "dashed"),
       lwd = 3.0,
       col = c("deepskyblue","black"))
```



Finally, we estimate exponential trend mode, report the estimation results and plot the fitted trend. The plot should be appropriately scaled with all crucial elements (i.e., axes, title, legend etc.) correctly labeled (0.25 Marks):

Note: All the relevant objects in our chart need to be clearly labeled. If there is no legend provided in the plot, the fitted trend must be mentioned in the title. Otherwise deduct (0.1 Marks).

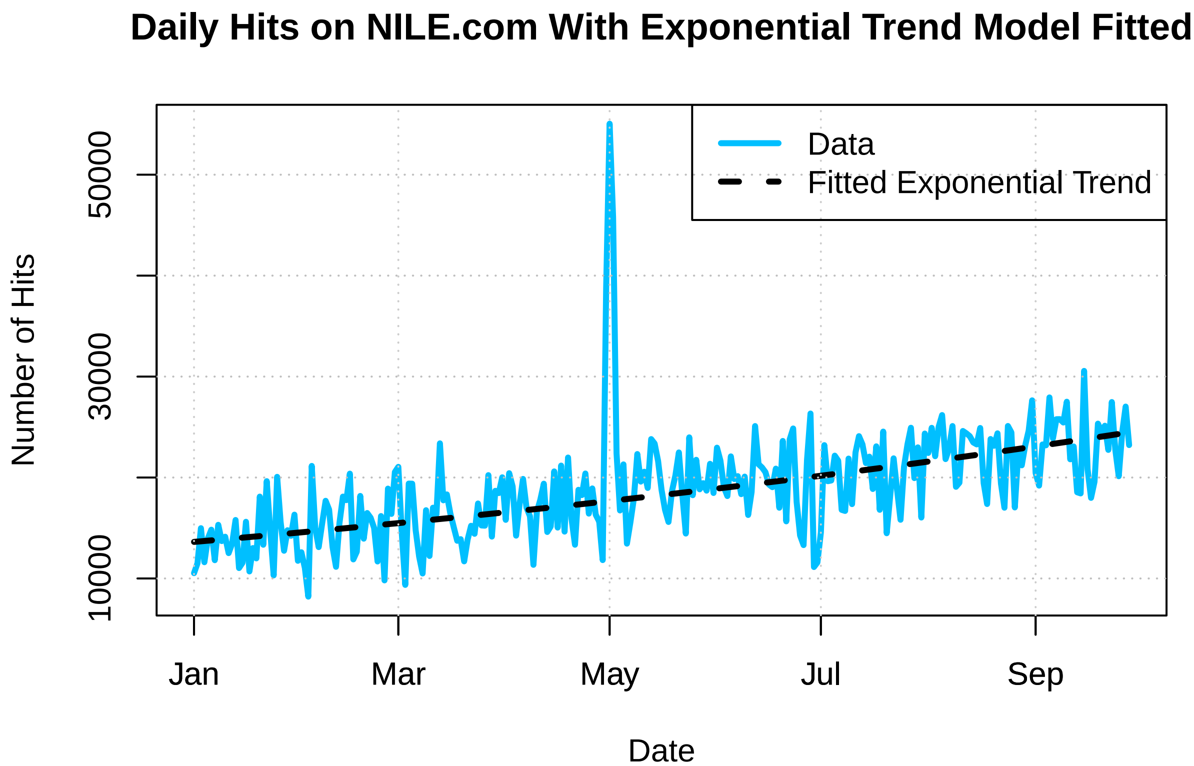
Note: The student should plot the fitted trend in the scale of the original data. We want to be able to visually compare the three trend specifications. If only the log scale plot is provided, deduct (0.1 Marks).

```
loghits <- log(data$Hits)

trendmod3 <- lm(formula = loghits ~ time) # This estimates the exponential trend model

summary(trendmod3)
```

```
##
## Call:
## lm(formula = loghits ~ time)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.59004 -0.12382  0.02281  0.11182  1.13613
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.5173291  0.0248469   383.0   <2e-16 ***
## time        0.0021698  0.0001584    13.7   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.204 on 269 degrees of freedom
## Multiple R-squared:  0.411,  Adjusted R-squared:  0.4088
## F-statistic: 187.7 on 1 and 269 DF,  p-value: < 2.2e-16
```

```
plot(date,data$Hits,
     main = "Daily Hits on NILE.com With Exponential Trend Model Fitted",
     xlab = "Date",
     ylab = "Number of Hits",
     col = "deepskyblue",
     type ="l",
     lwd = 3)
lines(date, exp(trendmod3$fitted.values), lwd = 3, lty = "dashed")
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
legend(x = "topright",
       legend = c("Data", "Fitted Exponential Trend"),
       lty = c("solid", "dashed"),
       lwd = 3.0,
       col = c("deepskyblue","black"))
```



Having estimated and plotted the three trend specifications, we can evaluate their fit using information criteria. To compute the correct AIC and BIC for the exponential trend model, we have to estimate it using the **nls** function:

```
abar <- exp(coef(trendmod3)[1])

bbar <- coef(trendmod3)[2]

hits <- data$Hits
trendmod3.a <- nls(formula = hits~a*exp(b*time), start=list(a=abar,b=bbar))
```

```r
infocrit <- data.frame(model = c("Linear", "Quadratic","Exponential"), AIC=c(AIC(trendmod1), AIC(trendmo

infocrit
```

```
##         model      AIC      BIC
## 1      Linear 5311.077 5321.884
## 2   Quadratic 5311.173 5325.581
## 3 Exponential 5313.861 5324.667
```

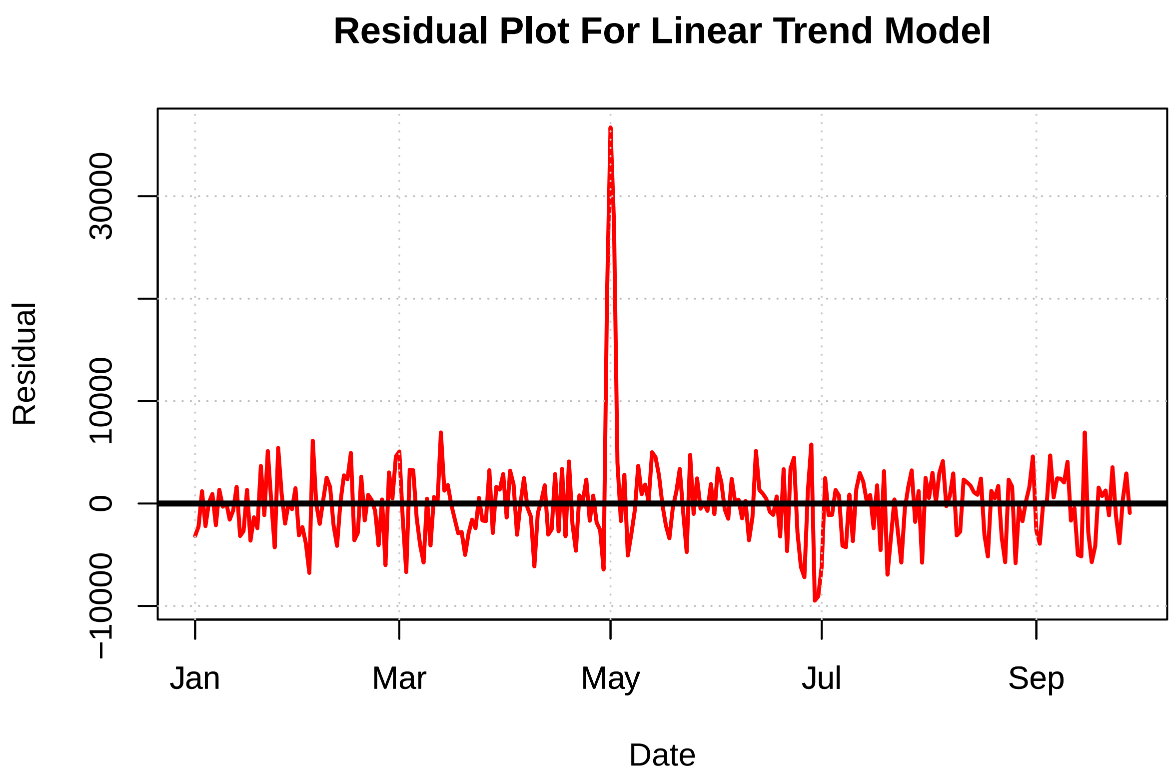
According to the AIC and BIC, the linear trend model is preferred.(0.25 Marks)

Note: The quadratic trend model is clearly not a reasonable choice since the coefficient on the quadratic term is not statistically significant. The exponential trend model should not be chosen as it yields the highest AIC. If the exponential or quadratic trend models are chosen over the linear trend model, deduct (0.3 Marks)

c.) For a few days in late April and early May, website hits were much higher than usual during a big sale. Do you find evidence of a corresponding group of outliers in the residuals from your trend models? Do they influence your trend estimates much? How should you treat them? (1 Mark)

The spike in hits during the sale should be clearly apparent in the residuals of each trend model.

```r
plot(date, trendmod1$residuals,
     main = "Residual Plot For Linear Trend Model",
     xlab = "Date",
     ylab = "Residual",
     col = "red",
     type = "l",
     lwd = 2.0)
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
abline(0,0, lwd = 3) # This is just a horizontal line at zero
```



In principle, we could exclude the outlying observations from the estimation sample, however, given that these data points are not erroneous – they simply reflect increased demand during a sale – this may not be a good option, particularly as it would lead to a gap in the sample. Deduct (0.6 Marks) if this is the suggested approach.

Alternatively, we could use a dummy variable to identify the sale period. The use of a dummy variable in the regression would eliminate the spike in the residuals. Moreover, the model with the dummy included yields a much lower AIC and BIC!

```
sale <- integer(T)

sale[120] = 1
sale[121] = 1
sale[122] = 1

trendmod4 <- lm(formula = data$Hits ~ time + sale)

summary(trendmod4)
```

```
##
## Call:
## lm(formula = data$Hits ~ time + sale)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9198.2 -2218.1   368.4  2146.9  8528.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13242.604    382.334   34.64   <2e-16 ***
## time           39.392      2.431   16.20   <2e-16 ***
```

```
## sale        28512.968   1817.892   15.69   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3131 on 268 degrees of freedom
## Multiple R-squared:  0.6503, Adjusted R-squared:  0.6477
## F-statistic: 249.2 on 2 and 268 DF,  p-value: < 2.2e-16
```

```
AIC(trendmod4)
```

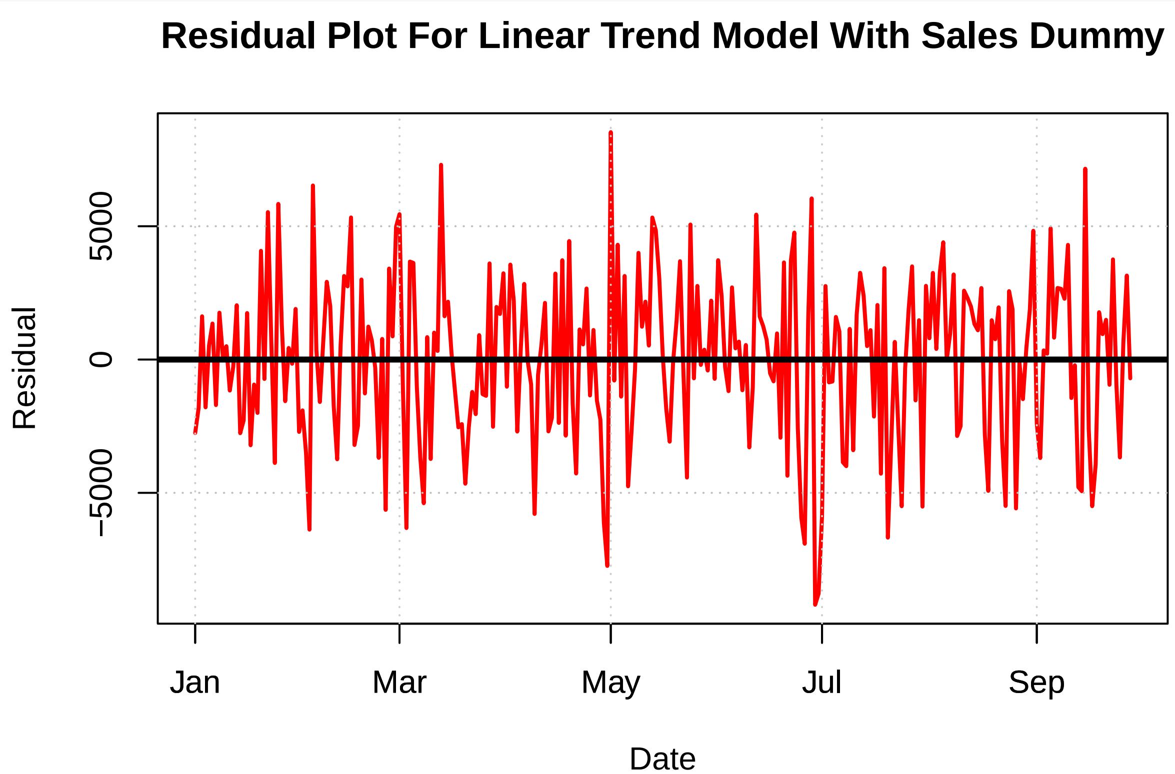
```
## [1] 5136.588
```

```
BIC(trendmod4)
```

```
## [1] 5150.997
```

Looking at the residuals from the dummy variable linear trend model, we can see that the spike has been eliminated:

```
plot(date, trendmod4$residuals,
     main = "Residual Plot For Linear Trend Model With Sales Dummy",
     xlab = "Date",
     ylab = "Residual",
     col = "red",
     type = "l",
     lwd = 2.0)
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
abline(0,0, lwd = 3) # This is just a horizontal line at zero
```



Note: If the dummy variable is used, award full marks. If no dummy is used and the outlier data is kept in the sample (i.e., the basic trend model is used), deduct (0.2 Marks).

d.) Using your preferred model, assess the significance of day-of-week effects in the hits to the NILE.COM website. (1 Marks)

Estimating the day of the week effects requires us to define a set of dummy variables that correspond to each of the seven days of the week. Let's start by generating our day of the week dummy variables:

```
D1 <- rep(c(1,0,0,0,0,0,0), ceiling(T/7))
D2 <- rep(c(0,1,0,0,0,0,0), ceiling(T/7))
D3 <- rep(c(0,0,1,0,0,0,0), ceiling(T/7))
D4 <- rep(c(0,0,0,1,0,0,0), ceiling(T/7))
D5 <- rep(c(0,0,0,0,1,0,0), ceiling(T/7))
D6 <- rep(c(0,0,0,0,0,1,0), ceiling(T/7))
D7 <- rep(c(0,0,0,0,0,0,1), ceiling(T/7))

D1 <- D1[1:T]
D2 <- D2[1:T]
D3 <- D3[1:T]
D4 <- D4[1:T]
D5 <- D5[1:T]
D6 <- D6[1:T]
D7 <- D7[1:T]
```

If we estimate the linear trend model with the full set of dummies (i.e. omitting the constant), we observe that the coefficients on all of the dummies are pretty close to one another.

Note: Need to include chosen trend in this specification. If no trend is specified and the model only includes the day of the week dummies, deduct (0.2 Marks)

```
trendmod5 <- lm(formula = data$Hits ~ 0 + time + D1 + D2 + D3 + D4 + D5 + D6 + D7)
summary(trendmod5)
```

```
##
## Call:
## lm(formula = data$Hits ~ 0 + time + D1 + D2 + D3 + D4 + D5 +
##     D6 + D7)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -8753  -2299   -124   1901  35865
##
## Coefficients:
##       Estimate Std. Error t value Pr(>|t|)
## time    38.661      3.359   11.51   <2e-16 ***
## D1   13741.387    826.001   16.64   <2e-16 ***
## D2   14507.240    827.836   17.52   <2e-16 ***
## D3   14130.707    829.681   17.03   <2e-16 ***
## D4   14378.124    831.535   17.29   <2e-16 ***
## D5   12929.361    833.399   15.51   <2e-16 ***
## D6   12982.078    836.339   15.52   <2e-16 ***
## D7   12898.365    838.171   15.39   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4325 on 263 degrees of freedom
## Multiple R-squared:  0.9529, Adjusted R-squared:  0.9515
## F-statistic: 665.2 on 8 and 263 DF,  p-value: < 2.2e-16
```

Note: If the significance day of week effects are inferred simply from the fact that the estimated coefficients in the above regression are significantly different from zero, deduct (0.3 Marks). We want to know whether the coefficients are different from one another!

In order to discern whether there exist significant day of week effects, we can omit one of the day dummies and re-estimate the model with a constant. This has the effect of treating the day whose dummy which we have omitted as the baseline. The coefficients on the remaining dummies now have the interpretation of deviations from the baseline. Any statistically significant day of week effects would then present themselves as significant coefficients. Looking at our results we can clearly see that there none.

```
trendmod6 <- lm(formula = data$Hits ~ time + D1 + D2 + D3 + D4 + D5 + D6)
summary(trendmod6)
```

```
##
## Call:
## lm(formula = data$Hits ~ time + D1 + D2 + D3 + D4 + D5 + D6)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -8753  -2299   -124   1901  35865
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12898.365    838.171  15.389   <2e-16 ***
## time           38.661      3.359  11.510   <2e-16 ***
## D1            843.022    985.936   0.855    0.393
## D2           1608.875    985.913   1.632    0.104
## D3           1232.342    985.901   1.250    0.212
## D4           1479.759    985.901   1.501    0.135
## D5             30.996    985.913   0.031    0.975
## D6             83.713    992.287   0.084    0.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4325 on 263 degrees of freedom
## Multiple R-squared:  0.3449, Adjusted R-squared:  0.3275
## F-statistic: 19.78 on 7 and 263 DF,  p-value: < 2.2e-16
```

e.) Your manager would like a forecast of the number of visits on the website for the remaining year. Using your results from parts (a) to (d), select a final model to use for forecasting. Use your model to produce appropriate interval forecasts through the end of 2017. Make sure to provide an interpretation of your forecast. (1 Mark)

Our analysis tells us that we should be using the linear trend model without day of the week effects to forecast. Since we want to forecast until the end of the year, we need to set our forecast horizon appropriately:

```
h <- 365 - T

horizon <- data.frame(time = seq(from = T+1, to = T+h))
```

Since our manager has asked us to forecast the number of visits on the website, we want to make sure that we compute a prediction interval. The prediction interval represents a range of values that is likely to contain the future number of visits. Since we are using the default, this probability will be 95%. (0.5 Marks)

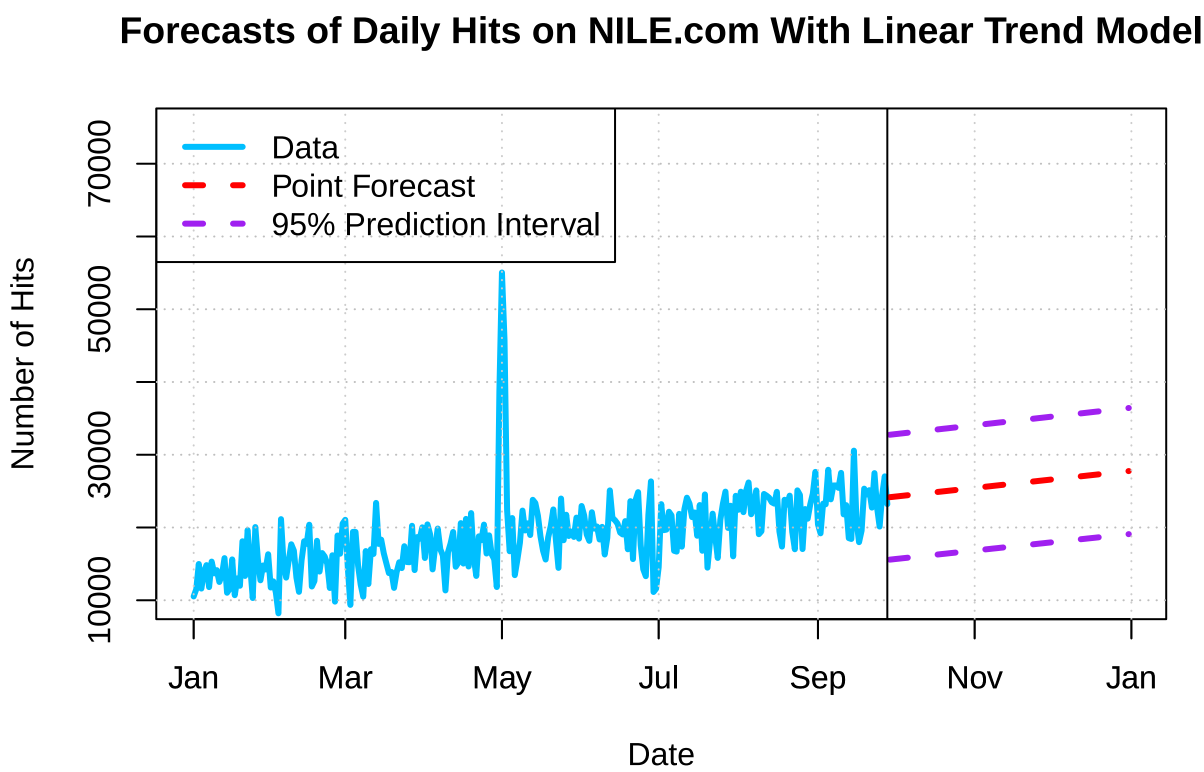
Note: Students need to explicitly choose a prediction interval. Simply replicating my code which produces both the confidence and prediction intervals without making a choice should incur a deduction of (0.3 Marks).

The plot should be appropriately scaled with all crucial elements (i.e., axes, title, etc.) correctly labeled (0.5

Marks):

Note: All the relevant objects in our chart need to be clearly labeled. If there is no legend provided in the plot, the fitted trend must be mentioned in the title. Otherwise deduct (0.1 Marks).

```r
forecast <- predict(trendmod1, newdata = horizon, interval='prediction')

forecast <- data.frame(forecast)

date.for <- seq(date[T], by = "day", length.out = h+1)

date.for <- date.for[1:h+1]

datenew <- c(date,date.for)

hitsnew <- c(data$Hits,rep(NA,h))

forecast.fit <- c(rep(NA,T),forecast$fit)
forecast.upr <- c(rep(NA,T),forecast$upr)
forecast.lwr <- c(rep(NA,T),forecast$lwr)

plot(datenew,hitsnew,
     main = "Forecasts of Daily Hits on NILE.com With Linear Trend Model",
     xlab = "Date",
     ylab = "Number of Hits",
     ylim = c(10000,75000),
     col = "deepskyblue",
     type ="l",
     lwd = 3)
lines(datenew, forecast.fit, lwd = 3, lty = "dashed", col = "red")
lines(datenew, forecast.lwr, lwd = 3, lty = "dashed", col = "purple")
lines(datenew, forecast.upr, lwd = 3, lty = "dashed", col = "purple")
grid(nx = NA, ny = NULL,
     col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
abline (v = datenew[T])
legend(x = "topleft",
       legend = c("Data", "Point Forecast", "95% Prediction Interval"),
       lty = c("solid", "dashed", "dashed"),
       lwd = 3.0,
       col = c("deepskyblue","red","purple"))
```



Question 2

The file applerev.xlsx contains Apple Inc.'s quarterly revenues (measured in billions of USD). The sample period starts at the third quarter of 2010 and ends at the fourth quarter of 2022. You are required to compute all your estimations and plots in R.

a.) Generate an appropriate plot of data and provide a brief description of the observed time series. Make sure to point out any notable visual features/characteristics of the data. (1 Mark)

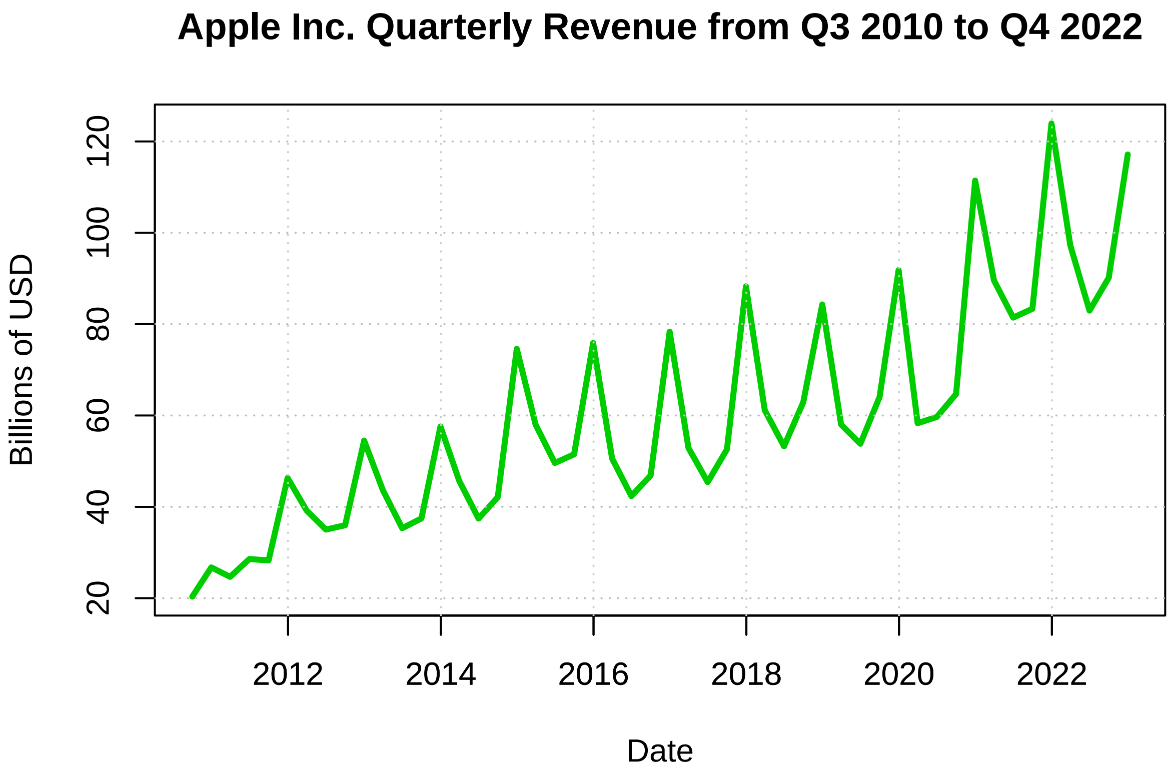
The plot should be appropriately scaled with all crucial elements (i.e., axes, title, etc.) correctly labeled (0.4 Marks).

Looking at the plot of the data, we can see that there is a clear upward trend and quarterly seasonal pattern. We also note that the magnitude of the seasonal fluctuations are increasing with the level of the time series (0.6 Marks).

Note: Students need to mention that the magnitude of the seasonal fluctuations are increasing with the level of the time series. Failure to mention this incurs a deduction of (0.2 Marks)

```
rm(list =ls())

library(readxl)

apple = read_excel("applerev.xlsx", sheet = "Sheet1", col_names = "rev", range = "B2:B51")

date = seq(as.Date("2010/9/30"), as.Date("2022/12/31"), by = "quarter")

plot(date, apple$rev,
     main = "Apple Inc. Quarterly Revenue from Q3 2010 to Q4 2022",
     xlab = "Date",
     ylab = "Billions of USD",
     col = "green3",
```

```
      type = "l",
      lwd = 3.0)
grid(nx = NA, ny = NULL,
      col = "grey", lwd = 1)
abline(v=axis.Date(1, date), col="grey80", lty = "dotted")
```



b.) Using the steps outlined in the lecture slides, compute an appropriate decomposition of the data into its trend-cycle, seasonal and residual/remainder components. Generate appropriate plots of these components and make sure to justify any choices that you have made. (3 Marks)

From our visual inspection of the data in part (a) it should be clear to us that we should be using a multiplicative decomposition. This is the most appropriate approach as the magnitude of the seasonal fluctuations are clearly increasing with the level of the series. (1 Mark)

Note: If students pursue an additive decomposition, deduct (0.6 Marks)

Since our data is observed at a quarterly interval, we compute the trend cycle component using a centered moving average: (0.2 Marks)

```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
trend.cycle <- ma(apple$rev, order = 4, centre = TRUE)
```

Then, we compute the de-trended series $\frac{y_t}{\hat{T}_t\hat{C}_t}$: (0.2 Marks)

```
detrend <- apple$rev/trend.cycle
```

Then, we proceed to compute the seasonal component for each quarter by taking the average of the de-trended values for each quarter. We can do this easily using a dummy variable regression: (0.2 Marks)

```r
T = length(date)

s = 4

nyears = ceiling(T/s)

Q3 <- rep(c(1,0,0,0), nyears)
Q4 <- rep(c(0,1,0,0), nyears)
Q1 <- rep(c(0,0,1,0), nyears)
Q2 <- rep(c(0,0,0,1), nyears)

Q3 <- Q3[1:T]
Q4 <- Q4[1:T]
Q1 <- Q1[1:T]
Q2 <- Q2[1:T]

seasmod <- lm(formula = detrend ~ 0 + Q3 + Q4 + Q1 + Q2)

summary(seasmod)
```

```
##
## Call:
## lm(formula = detrend ~ 0 + Q3 + Q4 + Q1 + Q2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.11348 -0.03538 -0.01084  0.03555  0.13200
##
## Coefficients:
##    Estimate Std. Error t value Pr(>|t|)
## Q3  0.87411    0.01571   55.64   <2e-16 ***
## Q4  1.33345    0.01571   84.88   <2e-16 ***
## Q1  0.96416    0.01504   64.11   <2e-16 ***
## Q2  0.83611    0.01504   55.59   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0521 on 42 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.9976, Adjusted R-squared:  0.9974
## F-statistic:  4375 on 4 and 42 DF,  p-value: < 2.2e-16
```

Then, we have to re-scale the estimated coefficients so that they sum exactly to 4. (0.2 Marks)

To see how to do this, suppose that we have four numbers $\{x_1, x_2, x_3, x_4\}$ and we have that their sum is equal to k

$$\sum_{i=1}^{4} x_i = k$$

Then, if we multiply each number by $\frac{4}{k}$ we have

$$\sum_{i=1}^{4} \frac{4}{k} x_i = \frac{4}{k} \sum_{i=1}^{4} x_i = \frac{4}{k} k = 4$$

```r
seas.coef <- seasmod$coefficients

seas.coef.adj <- seas.coef*4/sum(seas.coef)
```

Note: If the adjustment is not made, deduct (0.2 Marks)

Having computed the seasonal indices, we can then proceed to generate the seasonal component by by stringing together these quarterly values, and then replicating the sequence for each year of data:

```r
seasonal <- rep(seas.coef.adj, nyears)

seasonal <- seasonal[1:T]
```

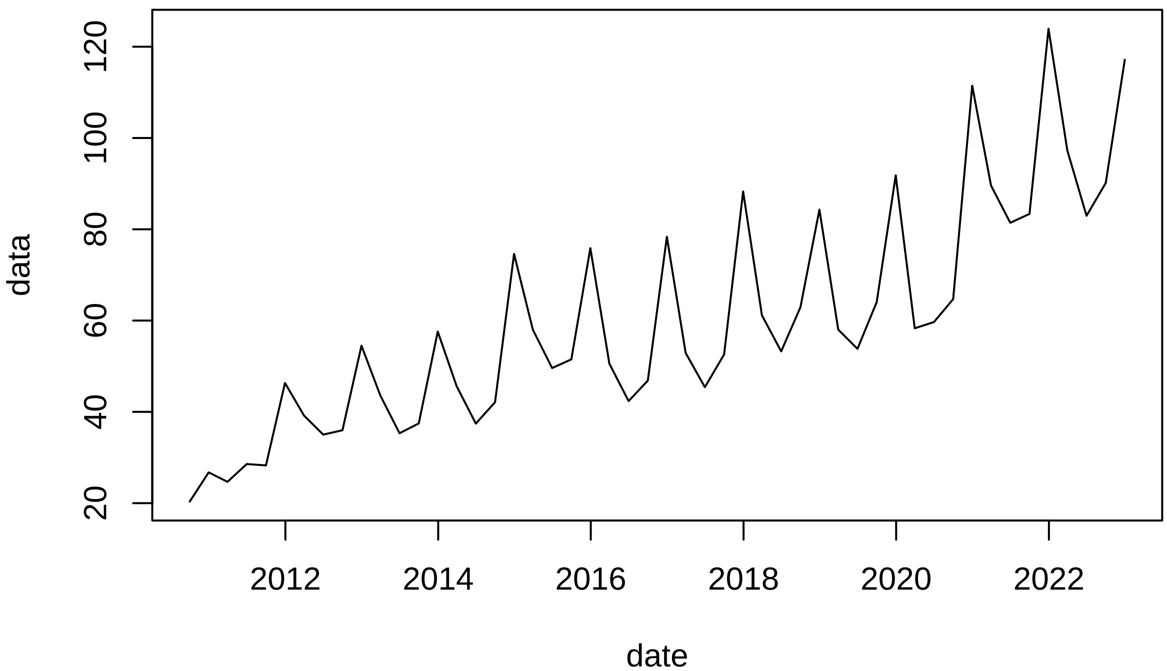
Finally, the remainder component is calculated by dividing out the estimated seasonal and trend-cycle components (0.2 Marks)

$$\hat{R}_t = \frac{y_t}{\hat{T}_t \hat{C}_t \hat{S}_t}$$

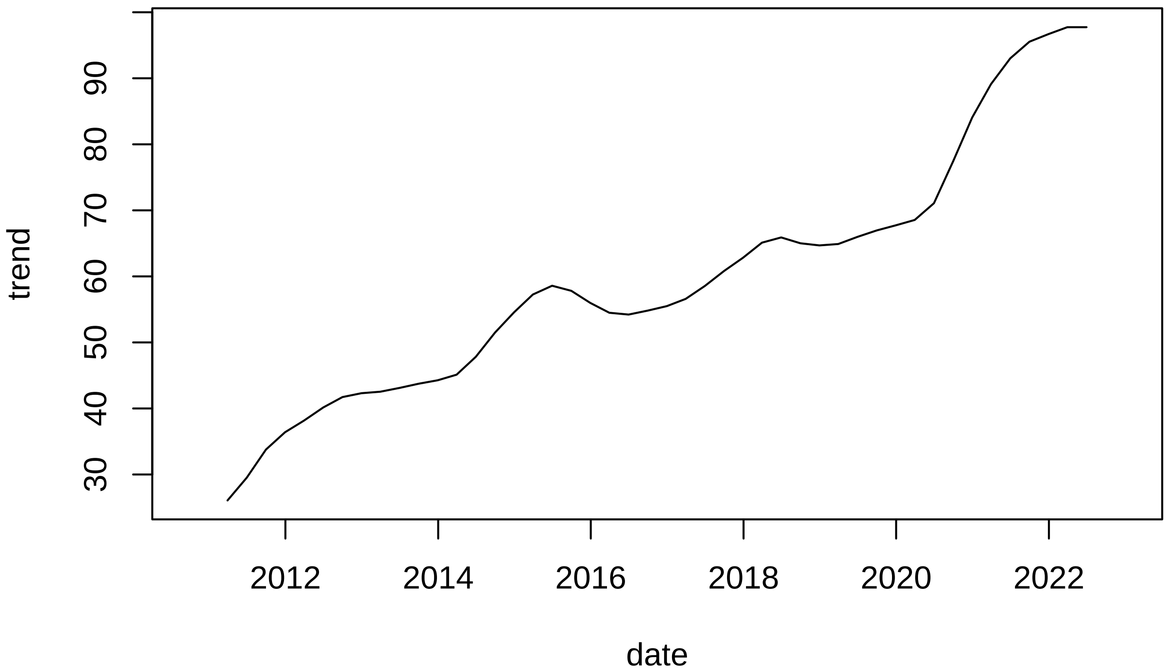
```r
remainder <- apple$rev/(trend.cycle*seasonal)
```

Now that we have all of our elements computed, we can plot them (1 Mark):

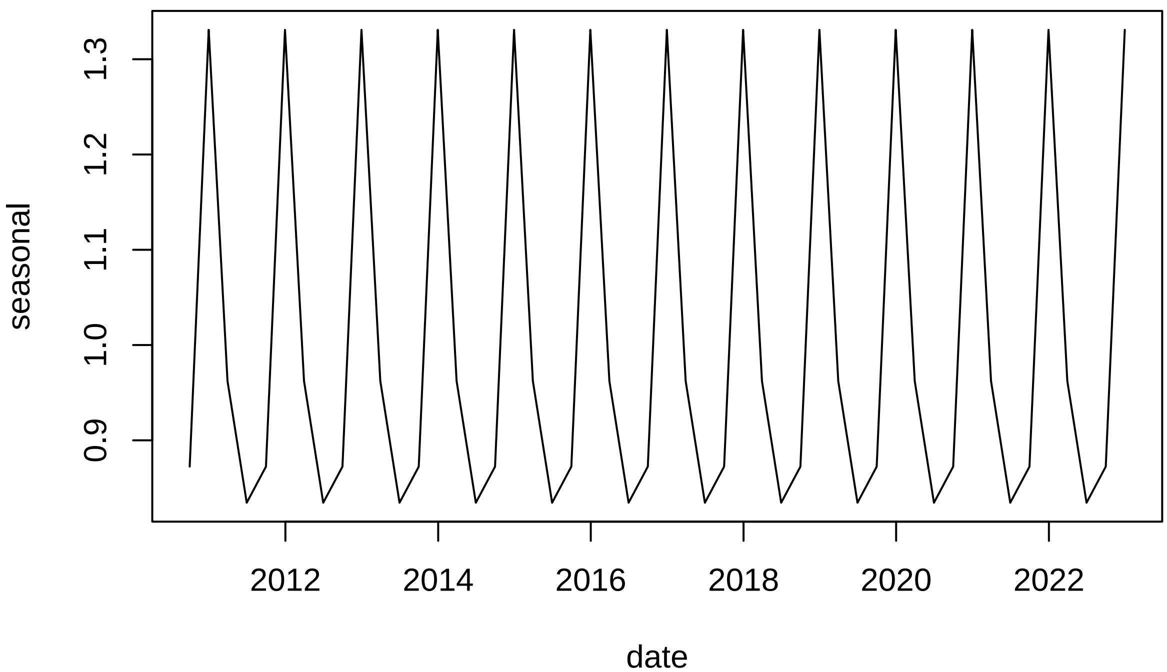
```r
plot(date,apple$rev, type = 'l', ylab = "data")
```



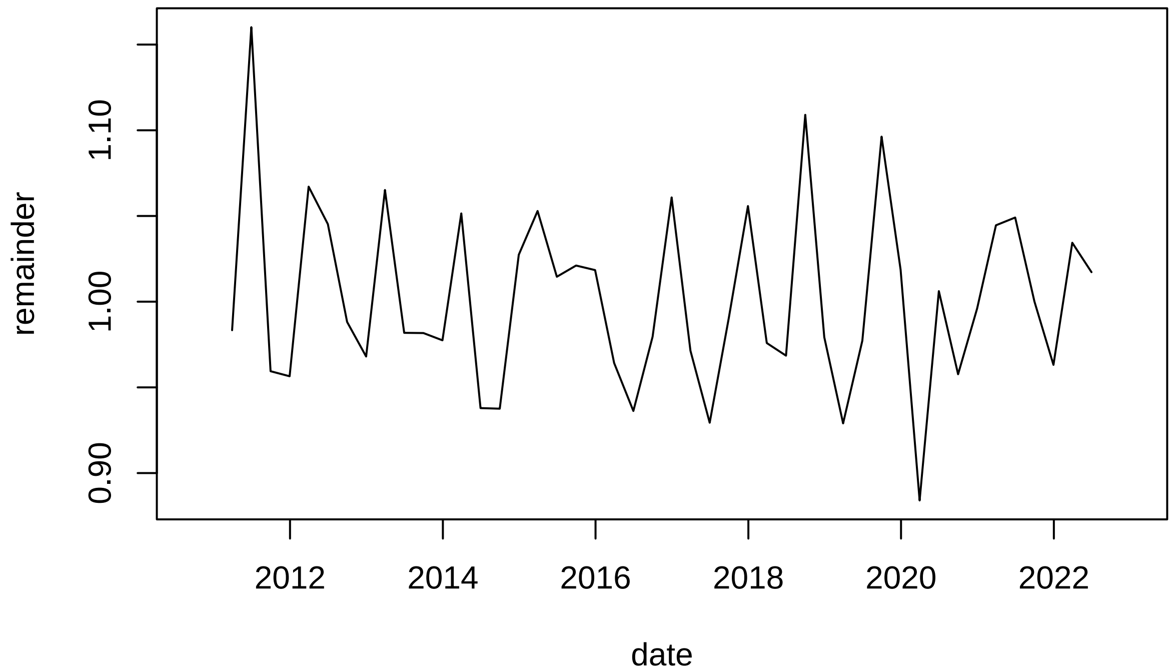

```r
plot(date,trend.cycle, type = 'l', ylab = "trend")
```




```
plot(date,seasonal, type = 'l', ylab = "seasonal")
```



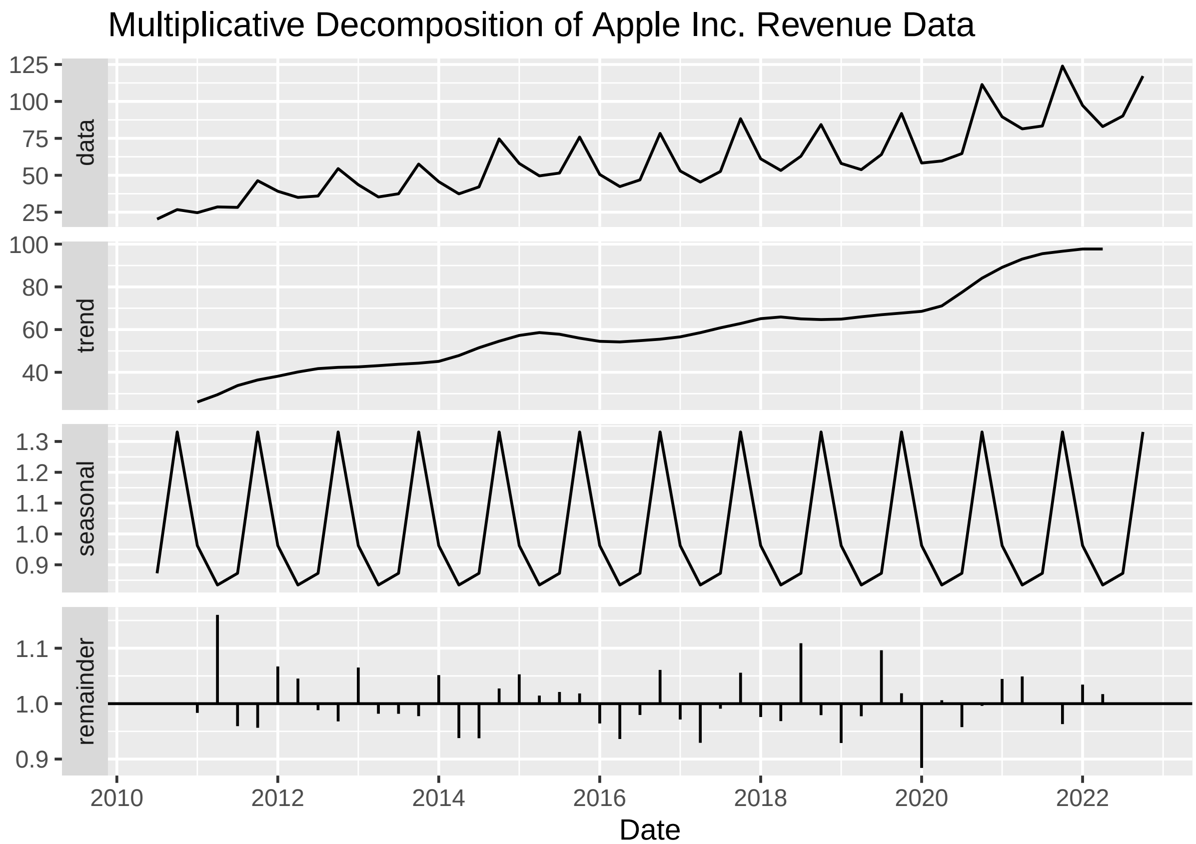

```
plot(date,remainder, type = 'l', ylab = "remainder")
```



c.) Compute the same decomposition using the decompose() and compare it with the decomposition that you have computed in part (b). (1 Mark)

Using the **decompose()** function, we obtain (0.5 Marks):

```
aapl.rev <- ts(apple$rev, start = c(2010,3), end = c(2022,4), frequency = 4)

aapl.decomp <- decompose(aapl.rev, type = "multiplicative")

autoplot(aapl.decomp,
         main = "Multiplicative Decomposition of Apple Inc. Revenue Data",
         xlab = "Date")
```



If we compare the components produced by the decompose function with the ones that we computed in part (b), we should see that they are exactly the same! (0.5 Marks)

Note: Deduct (0.5 Marks) if no comparison is made. If a comparison is made and the numbers are not exactly the same, deduct (0.2 Marks)

```
compare <- data.frame(trendcycle1 = trend.cycle, trendcycle2 = aapl.decomp$trend, seasonal1 = seasonal,

head(compare)
```

```
##   trendcycle1 trendcycle2 seasonal1 seasonal2 remainder1 remainder2
## 1          NA          NA 0.8724027 0.8724027         NA         NA
## 2          NA          NA 1.3308425 1.3308425         NA         NA
## 3    26.07125    26.07125 0.9622770 0.9622770  0.9833479  0.9833479
## 4    29.51125    29.51125 0.8344778 0.8344778  1.1601332  1.1601332
## 5    33.77500    33.77500 0.8724027 0.8724027  0.9594303  0.9594303
## 6    36.39625    36.39625 1.3308425 1.3308425  0.9564868  0.9564868
```