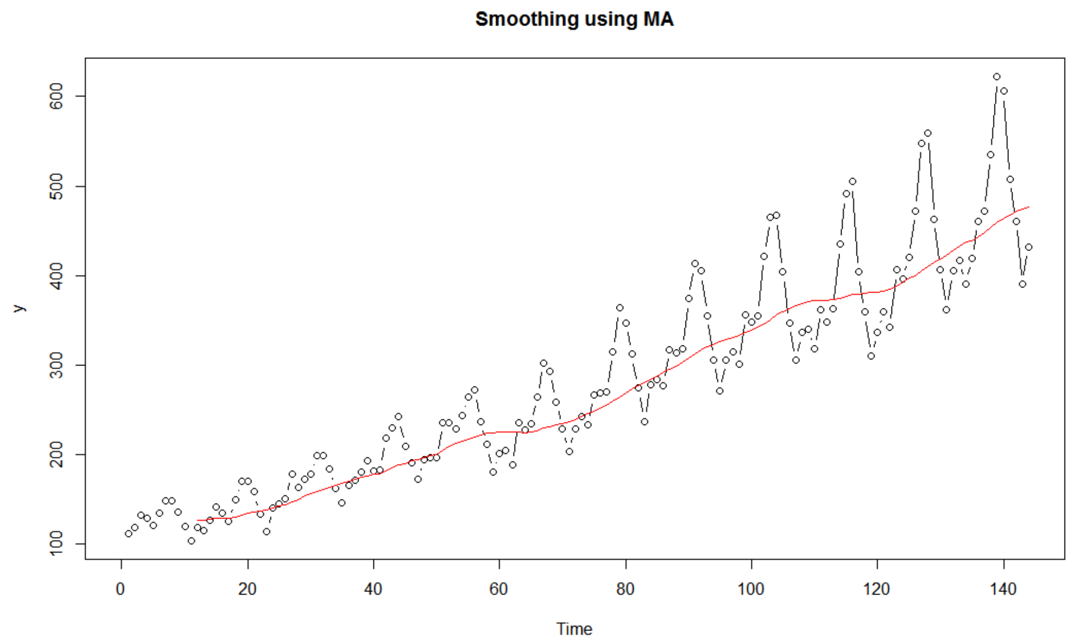
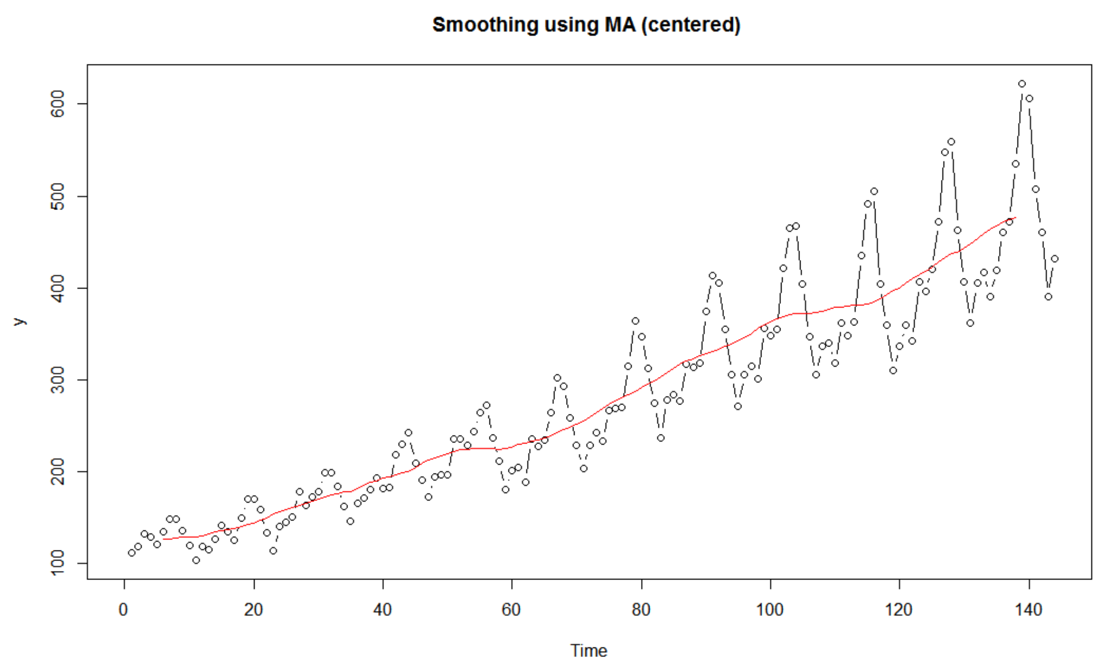
Homework 4

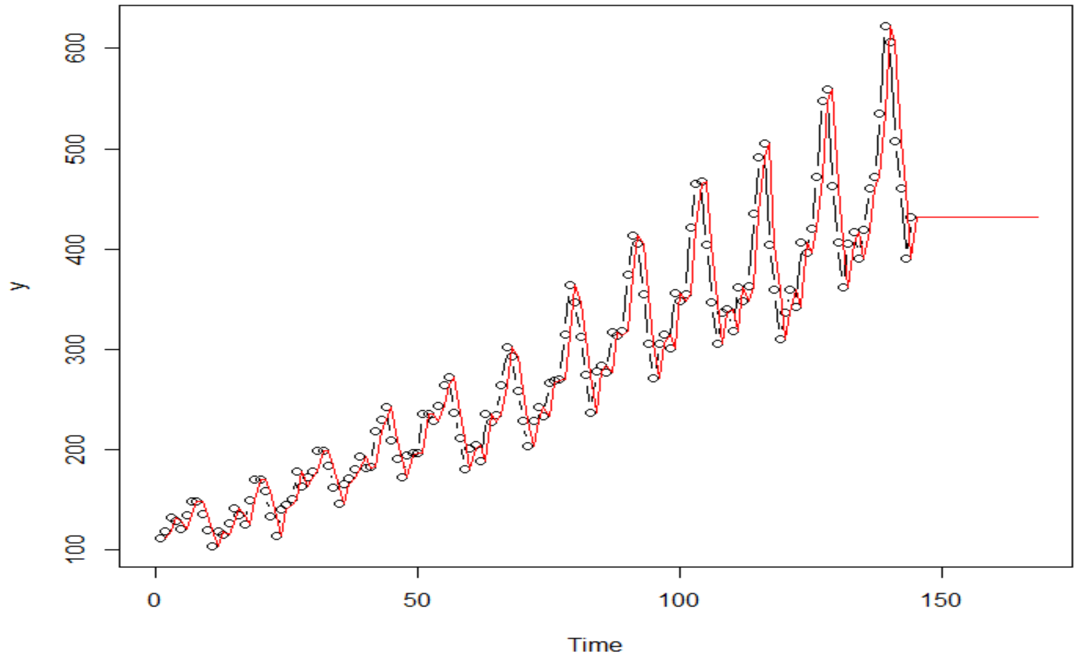
1 a) The data was centered using moving average. Since the data is year-wise, the window length m was chosen as 12 to smooth out the seasonality.



The smoothed MA lags the actual data by a half

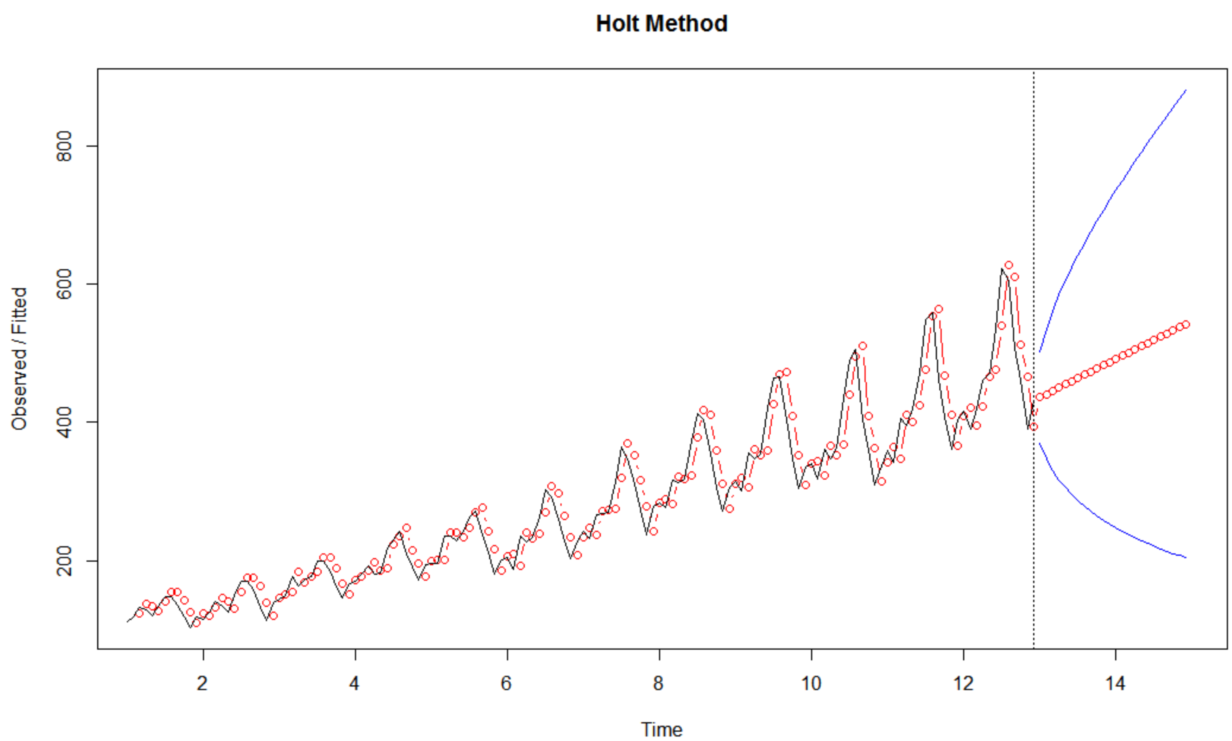


b) The Holt model without the trend and seasonality component was used to compute the optimal alpha for EWMA smoothing. The alpha obtained is 0.9999339. The forecast for the next 24 months is 431.9972.
With EWMA, the forecast for t + k time period, y_{t+k} is the same as the level at time t, L_t. Thus, all the forecasts are identical to the step at time L_{144} since only the level is used in the model for EWMA.



> alpha
[1] 0.9999339

c) The Holt method forecasts are shown below. Since this model accounts for the trend (compared to EWMA which does not), the k-step forecasts are not identical to each other. This model only accounts for the trend component and not the seasonality, so the forecasts obtained increase linearly. The forecast at time y_{t+k} are given by $L_t + k \times T_t$ where T_t is the trend at time t. The optimal parameters are alpha = 1 and beta = 0.003218516.



> HWData
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = y, gamma = FALSE, seasonal = "additive")

Smoothing parameters:
 alpha: 1
 beta : 0.003218516
 gamma: FALSE

Coefficients:

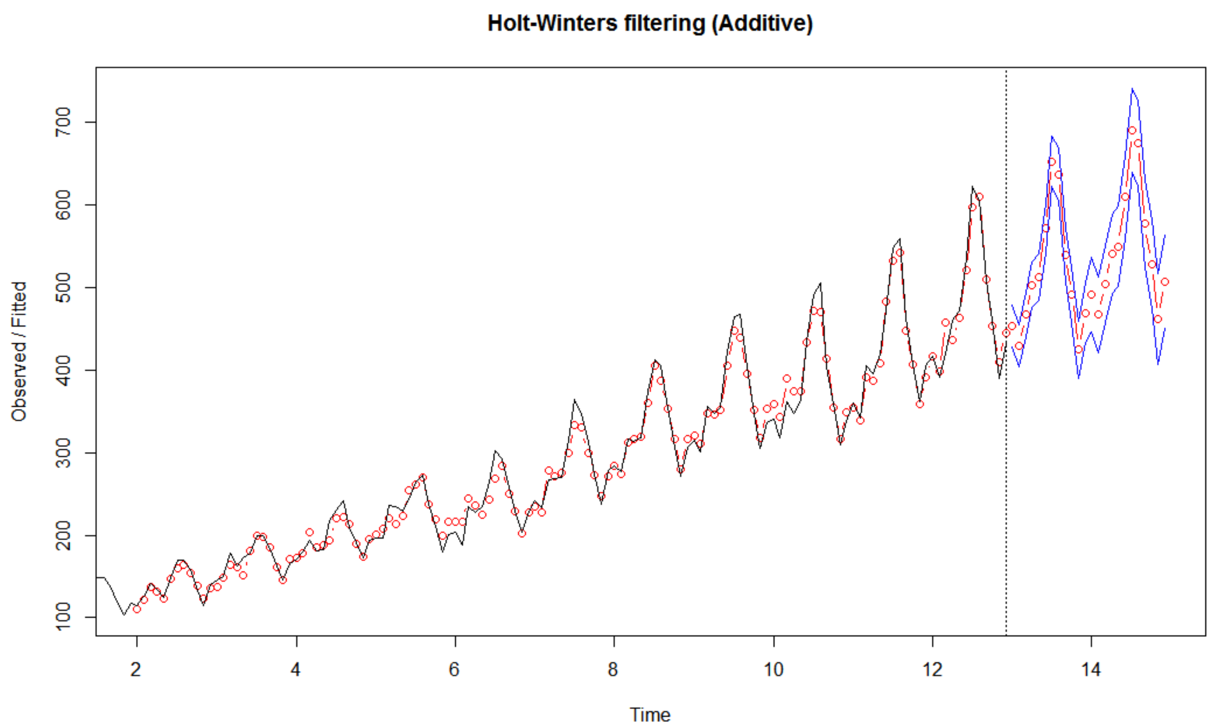
```
        [,1]
a 432.000000
b   4.597605
```

> HWDataPred

	fit	upr	lwr
Jan 13	436.5976	503.0933	370.1019
Feb 13	441.1952	535.3858	347.0046
Mar 13	445.7928	561.3379	330.2477
Apr 13	450.3904	584.0248	316.7560
May 13	454.9880	604.6357	305.3404

Jun 13 459.5856 623.7793 295.3920
Jul 13 464.1832 641.8168 286.5497
Aug 13 468.7808 658.9829 278.5788
Sep 13 473.3784 675.4407 271.3162
Oct 13 477.9761 691.3084 264.6437
Nov 13 482.5737 706.6754 258.4719
Dec 13 487.1713 721.6108 252.7317
Jan 14 491.7689 736.1696 247.3681
Feb 14 496.3665 750.3963 242.3366
Mar 14 500.9641 764.3275 237.6006
Apr 14 505.5617 777.9938 233.1296
May 14 510.1593 791.4208 228.8978
Jun 14 514.7569 804.6307 224.8831
Jul 14 519.3545 817.6424 221.0666
Aug 14 523.9521 830.4724 217.4318
Sep 14 528.5497 843.1351 213.9644
Oct 14 533.1473 855.6431 210.6515
Nov 14 537.7449 868.0078 207.4821
Dec 14 542.3425 880.2391 204.4459

d) An additive Holt-Winters model was fit to the data, the results of which are shown below. The optimal parameters, as shown below, are : alpha = 0.24795, beta = 0.03453, gamma = 1. A small value for alpha means that previous estimates of the levels are given more weight to compute the current level at time t. A small beta indicates that the trend at previous time periods is given more weight to compute the current trend. gamma = 1 indicates that the seasonality at time t is just the difference between the current value and the current level (so the seasonality coefficients indicate this differences). The Seasonality coefficient at “t” is just the weighted sum of the seasonality coefficient at “t-s” (s= 12 months), and the difference between current value and current level. The seasonality indices are negative, then positive and then change to negative again. The largest seasonal effect occurs in the months of July and September (largest positive difference of the current value and current level), which makes sense since travelling activity peaks during the summer months.



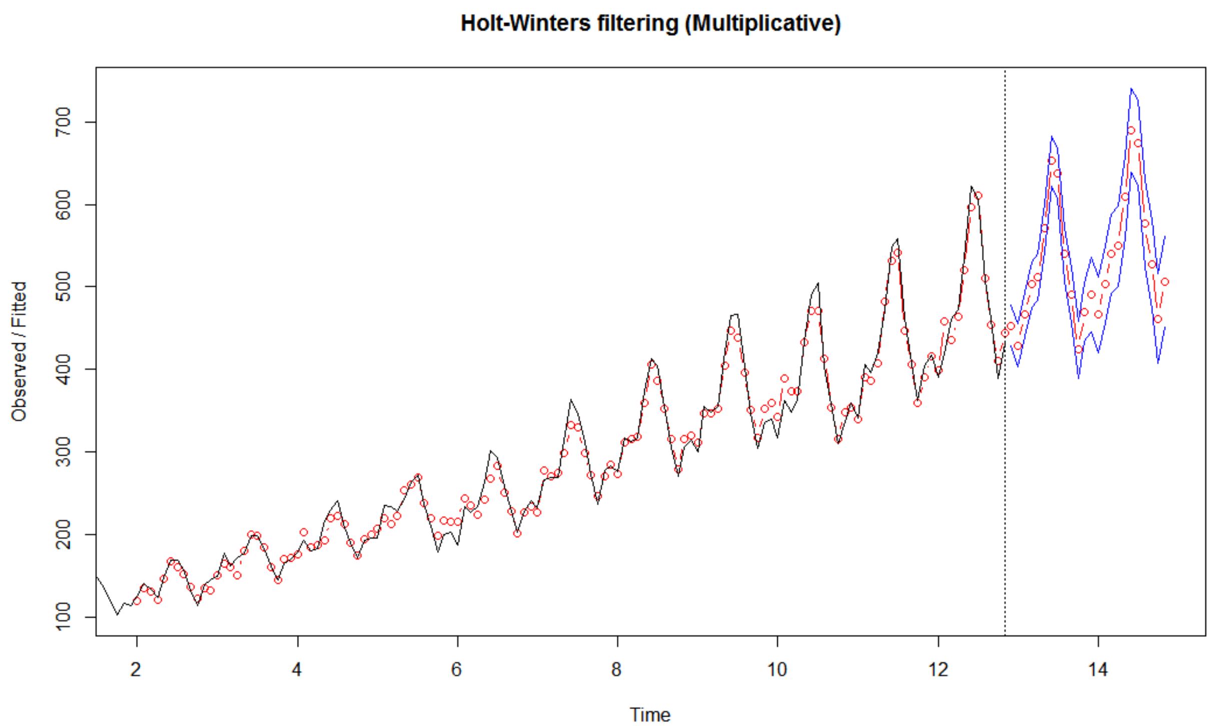

```
> HWdataA
Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:
HoltWinters(x = y, seasonal = "additive")

Smoothing parameters:
 alpha: 0.2479595
 beta : 0.03453373
 gamma: 1

Coefficients:
          [,1]
a   477.827781
b     3.127627
s1  -27.457685
s2  -54.692464
s3  -20.174608
s4   12.919120
s5   18.873607
s6   75.294426
s7  152.888368
```

```
s8  134.613464
s9   33.778349
s10 -18.379060
s11 -87.772408
s12 -45.827781
```

The predicted values for the two-year period (1961-1962) are shown below.

```
> HWdataPredA
            fit      upr      lwr
Jan 13 453.4977 478.5802 428.4153
Feb 13 429.3906 455.2851 403.4960
Mar 13 467.0361 493.7706 440.3015
Apr 13 503.2574 530.8590 475.6558
May 13 512.3395 540.8343 483.8447
Jun 13 571.8880 601.3014 542.4746
Jul 13 652.6095 682.9661 622.2530
Aug 13 637.4623 668.7858 606.1387
Sep 13 539.7548 572.0684 507.4412
Oct 13 490.7250 524.0511 457.3988
Nov 13 424.4593 458.8197 390.0988
Dec 13 469.5315 504.9475 434.1155
Jan 14 491.0292 535.9665 446.0920
Feb 14 466.9221 512.7540 421.0901
Mar 14 504.5676 551.3190 457.8162
Apr 14 540.7889 588.4841 493.0938
May 14 549.8710 598.5338 501.2083
Jun 14 609.4195 659.0732 559.7657
Jul 14 690.1411 740.8087 639.4734
Aug 14 674.9938 726.6978 623.2897
Sep 14 577.2863 630.0487 524.5239
Oct 14 528.2565 582.0989 474.4141
Nov 14 461.9908 516.9343 407.0473
Dec 14 507.0630 563.1283 450.9977
```

e) A multiplicative Holt-Winters model was fit to the data, the results of which are shown below. The optimal parameters, as shown below, are : alpha = 0.2756, beta = 0.0327, gamma = 0.8707. The Seasonality coefficient at “t” is just the weighted sum of the seasonality coefficient at “t-s” (s= 12 months), and the ratio of current value to current level.



> HWdataM
Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = y, seasonal = "multiplicative")

Smoothing parameters:
alpha: 0.2755925
beta : 0.03269295
gamma: 0.8707292

Coefficients:

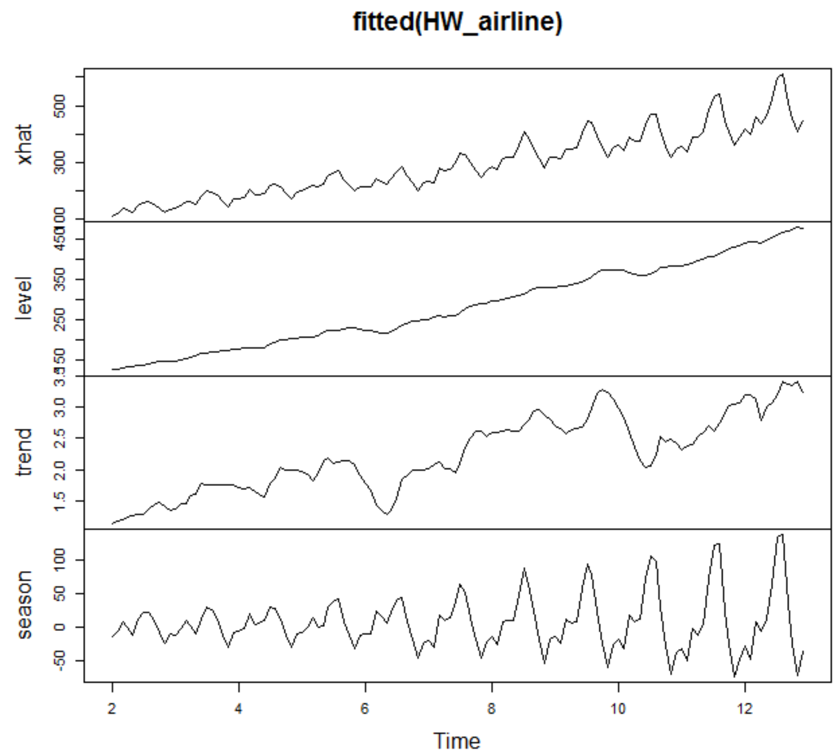
	[,1]
a	469.3232206
b	3.0215391
s1	0.9464611
s2	0.8829239
s3	0.9717369
s4	1.0304825
s5	1.0476884
s6	1.1805272
s7	1.3590778
s8	1.3331706

s9	1.1083381
s10	0.9868813
s11	0.8361333
s12	0.9209877

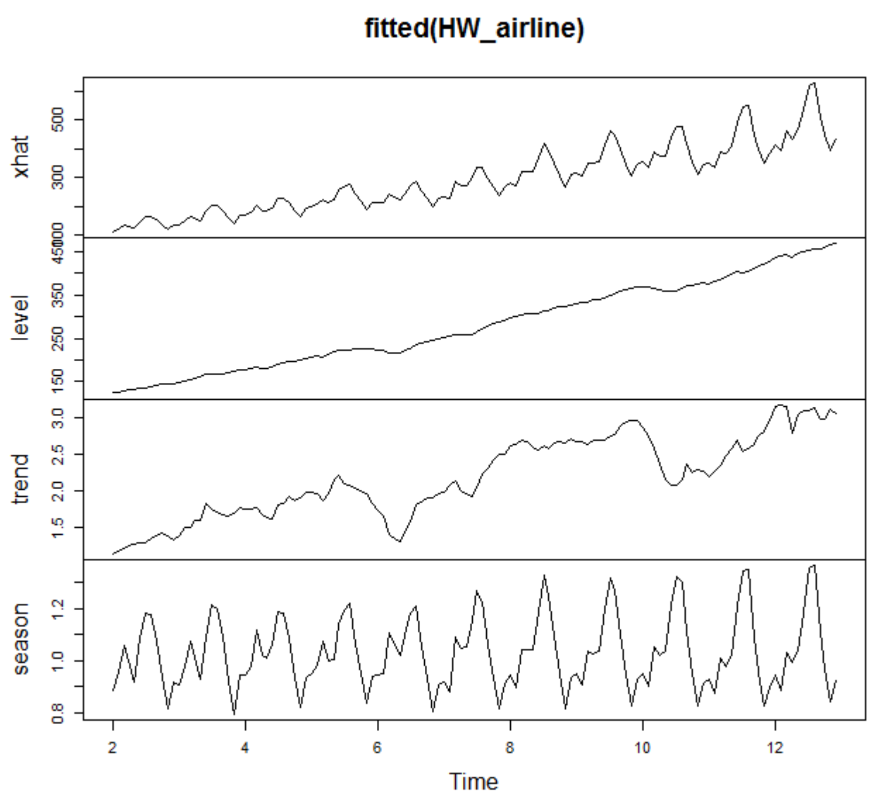
The seasonality indices increase from the second to the seventh month, and then decrease till the eleventh. The largest coefficients occur again in months of July and August, , which makes sense since travelling activity peaks during the summer months. The predicted values for the 1961-1962 period are shown below.

```
> HWdataPredM
         fit      upr      lwr
Jan 13 447.0559 466.8057 427.3061
Feb 13 419.7123 440.2920 399.1326
Mar 13 464.8671 486.7712 442.9630
Apr 13 496.0839 519.3350 472.8329
May 13 507.5326 531.9278 483.1375
Jun 13 575.4509 602.1935 548.7083
Jul 13 666.5923 696.5558 636.6288
Aug 13 657.9137 688.6454 627.1821
Sep 13 550.3088 578.9777 521.6398
Oct 13 492.9853 520.9553 465.0153
Nov 13 420.2073 446.9458 393.4688
Dec 13 465.6345 487.9686 443.3004
Jan 14 481.3732 517.8126 444.9337
Feb 14 451.7258 488.0308 415.4207
Mar 14 500.1008 538.8928 461.3088
Apr 14 533.4477 574.3831 492.5122
May 14 545.5202 587.8399 503.2005
Jun 14 618.2550 664.8185 571.6915
Jul 14 715.8704 768.3289 663.4118
Aug 14 706.2524 759.2423 653.2626
Sep 14 590.4954 638.2882 542.7027
Oct 14 528.7681 574.2084 483.3279
Nov 14 450.5242 492.7194 408.3290
Dec 14 499.0281 535.8450 462.2112
```

f) The fitted values plot of the additive model in (d) is given below:



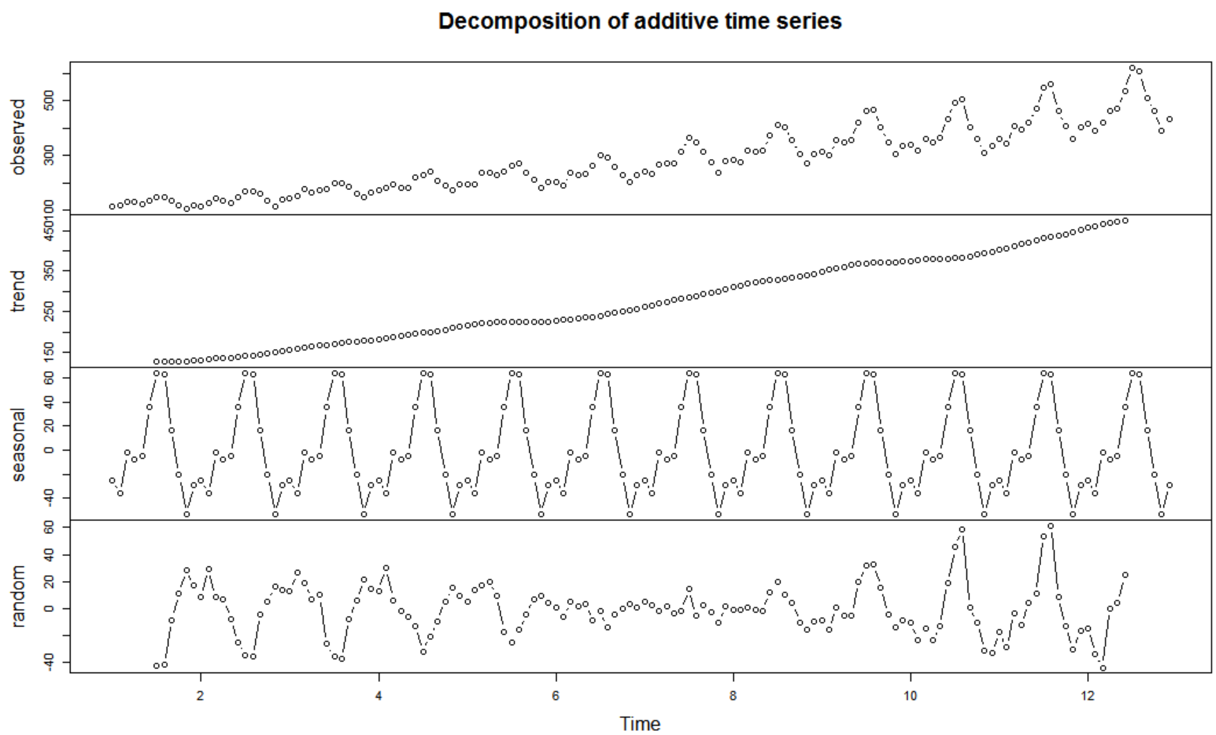
The fitted values plot of the multiplicative model in (e) is given below:

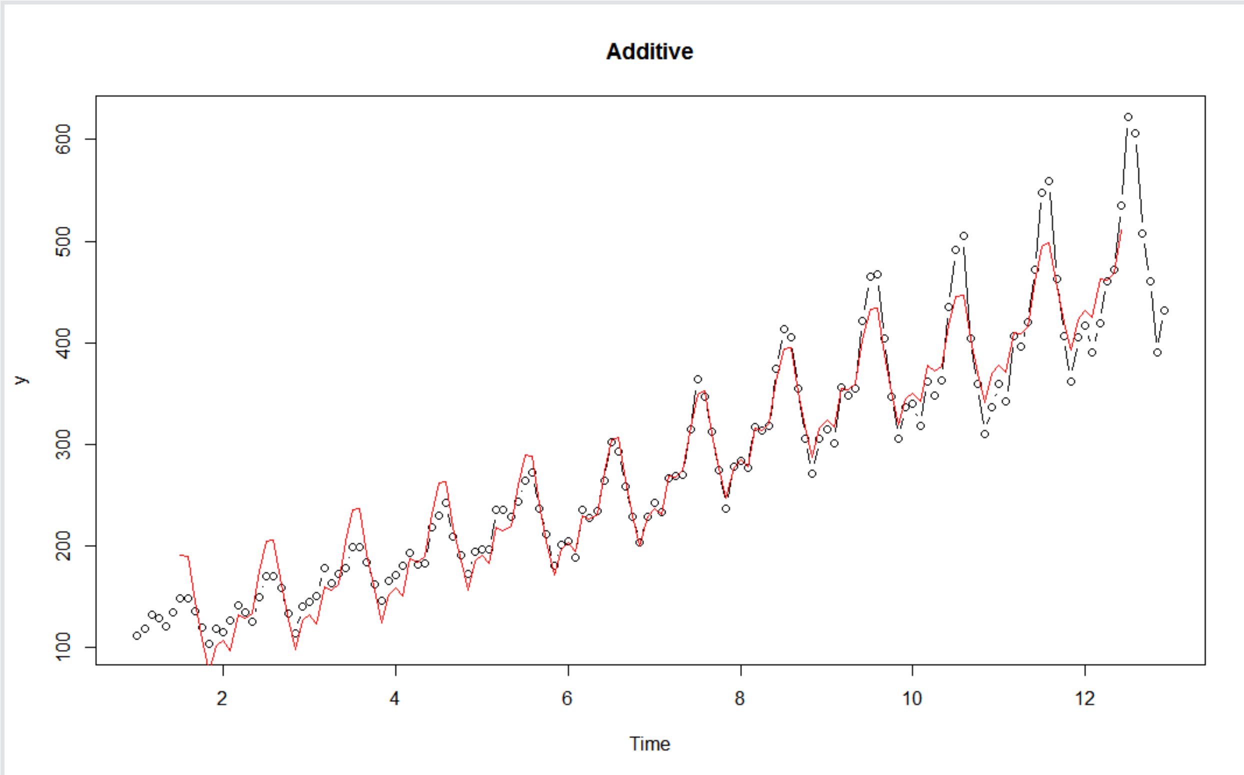


The multiplicative model captures the seasonality much better than the additive model. We can see above that the amplitude of the seasonality is proportional to the trend. This is better captured by the multiplicative model than the additive model. From the fitted model, we can see that the seasonality is fairly constant for the multiplicative model, as compared to the additive one.

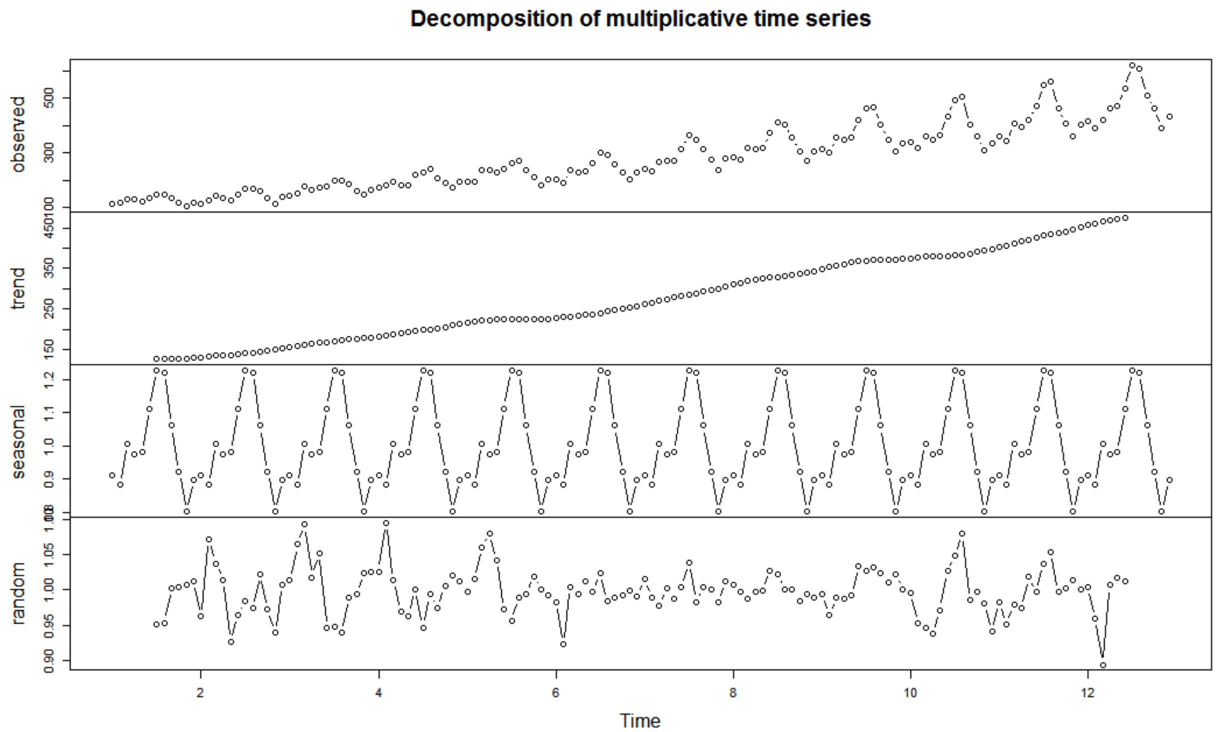
2.

a) The decomposition of the time series into trend, seasonal and random components using additive model is given below. The variability in the data is not completely captured by the trend and seasonal components, as we notice some cyclical patterns in the random component. The second plot overlays the original time series with the sum of trend and seasonal components in red, and we see that the additive model overestimates the amplitudes in the early time periods and underestimates amplitudes in the later ones.

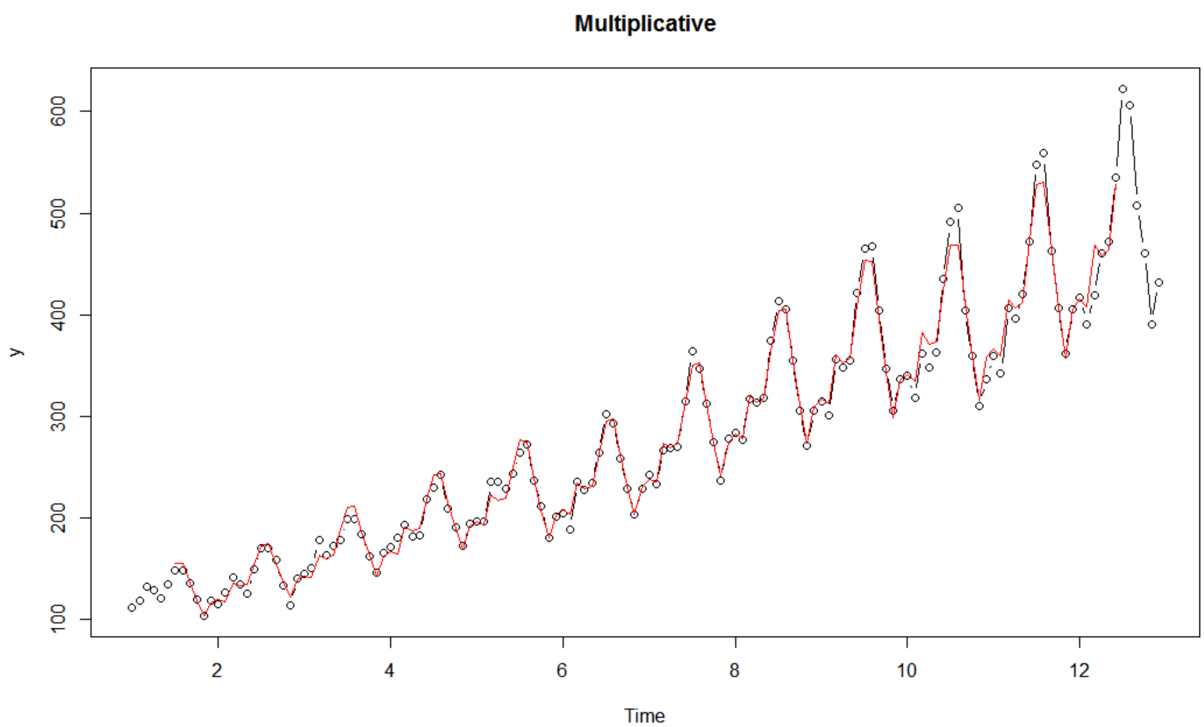




b) The decomposition of the time series into trend, seasonal and random components using multiplicative model is given below. The variability in the data is captured much better in this model compared to the additive one, as we see cyclical patterns in the random component, but on a much smaller scale. The second plot overlays the original time series with the product of the trend and seasonal components. We notice that while the model still slightly overestimates the time components in the early time periods and underestimates in the later ones, it performs much better than the additive model in representing the data. The random component is also pretty small, in the range of 0.9 and 1.1 for the multiplicative model.



Decomposition of multiplicative time series
observed
trend
seasonal
random
Time



Multiplicative
y
Time

Appendix: Source Code

```
#### Load Data ####
setwd("C:\\Users\\Sanjeevni\\code\\msia420")
data <- read.csv("C:\\Users\\Sanjeevni\\Documents\\1 - Northwestern\\2015 Winter\\Predictive
Analytics\\Data\\HW4_data.csv", header=F)
names(data) <- "trend"

### 1 ####
#####################################################################
## A ##
par(mfrow=c(1,1))
y<-ts(data[[1]], frequency=1)
m=12;n=length(y)  #m = MA window length, k = prediction horizon
plot(y,type="b",xlim=c(0,n), main = "Smoothing using MA")
MAdata<-filter(y, filter=rep(1/m,m), method = "convolution", sides = 1)
lines(MAdata,col="red")

plot(y,type="b",xlim=c(0,n), main = "Smoothing using MA (centered) ")
MAdata<-filter(y, filter=rep(1/m,m), method = "convolution", sides = 2)
lines(MAdata,col="red")

#####################################################################
## B ##
par(mfrow = c(1,1))
y<-ts(data[[1]], frequency=1)
k=24;n=length(y)  #k = prediction horizon
HWData<-HoltWinters(y, seasonal = "additive", beta=FALSE, gamma = FALSE)
alpha=HWData$alpha
#for (alpha in c(1,.2,.05)){
  plot(y,type="b",xlim=c(0,n+k), main="EWMA")
  EWMAdata1<-alpha*filter(y, filter=1-alpha, method = "recursive", sides = 1, init=y[1]/alpha)
  yhat=c(NA,EWMAdata1,rep(EWMAdata1[n],k-1))
  lines(yhat,col="red")
#}
#repeat for alpha = 1, 0.2, 0.05
####################################################################
## C ##

par(mfrow = c(1,1))
y<-ts(data[[1]], deltat=1/12)
```

```
k=24;n=length(y)  #k = prediction horizon
HWData<-HoltWinters(y, seasonal = "additive", gamma = FALSE)
HWDataPred<-predict(HWData, n.ahead=k, prediction.interval = T, level = 0.95)
plot(HWData,HWDataPred,type="b", main="Holt Method")
HWData
# alpha: 1
# beta : 0.02095455

#################################################################
## D ##

y<-ts(data[[1]], deltat=1/12)  #sampling interval corresponds to 1/12 the seasonality period.
Could instead specify frequency = 12
k=24;n=length(y)  #k = prediction horizon
HWdataA<-HoltWinters(y, seasonal = "additive")
HWdataPredA<-predict(HWdataA, n.ahead=k, prediction.interval = T, level = 0.95)
plot(HWdataA,HWdataPredA,type="b", main="Holt-Winters filtering (Additive)")
########

#################################################################
## E ##

y<-ts(data[[1]], frequency=12)  #sampling interval corresponds to 1/12 the seasonality period.
Could instead specify frequency = 12
k=24;n=length(y)  #k = prediction horizon
HWdataM<-HoltWinters(y, seasonal = "multiplicative")
HWdataPredM<-predict(HWdataM, n.ahead=k, prediction.interval = T, level = 0.95)
par(mfrow=c(1,1))
plot(HWdataM,HWdataPredM,type="b", main="Holt-Winters filtering (Multiplicative)")

par(mfrow=c(1,2))
plot(HWdataM,HWdataPredM,type="b")
plot(HWdataA,HWdataPredA,type="b")

#################################################################
### 2 ###
## A ##
par(mfrow=c(1,1))
y<-ts(data[[1]], deltat=1/12)
Dectrade<-decompose(y, type = "additive")
plot(Dectrade,type="b")
Dectrade
##
```

```
y_hat<-Dectrade$trend+Dectrade$seasonal
plot(y,type="b", main="Additive")
lines(y_hat,col="red")

## B ##
par(mfrow=c(1,1))
y<-ts(data[[1]], deltat=1/12)
Dectrade<-decompose(y, type = "mult")
plot(Dectrade,type="b")
Dectrade
##
y_hat<-Dectrade$trend*Dectrade$seasonal
plot(y,type="b", main="Multiplicative")
lines(y_hat,col="red")
```